Fuel 345 (2023) 128180

Contents lists available at ScienceDirect

Fuel

journal homepage: www.elsevier.com/locate/fuel

Full Length Article

Deep learning for multiphase segmentation of X-ray images of gas diffusion layers

Mehdi Mahdaviara [a], Mohammad Javad Shojaei [b], Javad Siavashi [a], Mohammad Sharifi [a,*], Martin J. Blunt [c]

[a] *Department of Petroleum Engineering, Amirkabir University of Technology, Tehran Polytechnic, 15875-4413 Tehran, Iran*
[b] *School of Computing, Newcastle University, Newcastle, UK*
[c] *Department of Earth Science and Engineering, Imperial College London, London SW7 2BP, UK*

ARTICLE INFO

Keywords:
Image processing
Deep learning
Semantic segmentation
Gas diffusion layer

ABSTRACT

High-resolution X-ray computed tomography (micro-CT) has been widely used to characterise fluid flow in porous media for different applications, including in gas diffusion layers (GDLs) in fuel cells. In this study, we examine the performance of 2D and 3D U-Net deep learning models for multiphase segmentation of unfiltered X-ray tomograms of GDLs with different percentages of hydrophobic polytetrafluoroethylene (PTFE). The data is obtained by micro-CT imaging of GDLs after brine injection. We train deep learning models on base-case data prepared by the 3D Weka segmentation method and test them on the unfiltered unseen datasets. Our assessments highlight the effectiveness of the 2D and 3D U-Net models with test IoU values of 0.901 and 0.916 and f1-scores of 0.947 and 0.954, respectively. Most importantly, the U-Net models outperform conventional 3D trainable Weka and watershed segmentation based on various visual examinations. Lastly, flow simulation studies reveal segmentation errors associated with trainable Weka and watershed segmentation lead to significant errors in the calculated porous media properties, such as absolute permeability. Our findings show 43, 14, 14, and 3.9% deviations in computed permeabilities for GDLs coated by 5, 20, 40, and 60 w% of PTFE, respectively, compared to images segmented by the 3D Weka segmentation method.

1. Introduction

Three-dimensional imaging has transformed our understanding of otherwise opaque porous materials and the flow processes within them [1–3]. Image segmentation is a crucial part of this process which detects the solid skeleton, the pore space, and the fluid phases within the pore space [4–9], which depends on the quality of the images [10]. This has made image processing an active research area [11] with applications ranging from natural porous media, including biological tissues [12–15] and subsurface geomaterial [1,3,6,16,17] to manufactured porous environments such as gas diffusion layers (GDLs) used in proton exchange membrane fuel cells (PEMFC) [18–20].

Fuel cells are energy conversion devices that generate green electrical energy with minimal to zero greenhouse gas emissions. The ability of fuel cells to produce green energy offers a significant advantage over fossil fuels [21]. GDLs are an essential component of fuel cells that allow the reactants, oxygen and hydrogen, to enter the reaction site in catalyst layers, transport electrons, while allowing the by-products of the reaction (water and heat) to escape during fuel cell operation [22,23]. The pore spaces are the pathways for the fluids, while the carbon fibres are responsible for transferring electrons and heat. GDLs are usually partially coated with hydrophobic PTFE, as they are originally water-wet, to allow the simultaneous flow of water and gas [23,24]. GDLs are also present in electrolysers which act in reverse, using an electrical current to split water into hydrogen and oxygen.

Full cell performance strongly depends on GDL properties such as thickness, pore size distribution (PSD), tortuosity, permeability, and wettability [25–29]. In recent years, high-resolution X-ray imaging has been applied to study these properties in more depth to optimise fuel cell performance [18,30]. Researchers usually follow three main steps: (I) imaging; (II) image processing, in which images are denoised and segmented into discrete phases; and (III) computations, in which different parameters such as pore radii, contact angle, and pore size distribution are calculated from the segmented images. Subsequently, transport phenomena [31–34], evaporation [35,36], thermal and electrical conductivity [37,38], fluid distribution [39–41], and permeability

* Corresponding author.
E-mail address: m_sharifi@aut.ac.ir (M. Sharifi).

https://doi.org/10.1016/j.fuel.2023.128180
Received 26 January 2023; Received in revised form 7 March 2023; Accepted 19 March 2023
Available online 12 April 2023

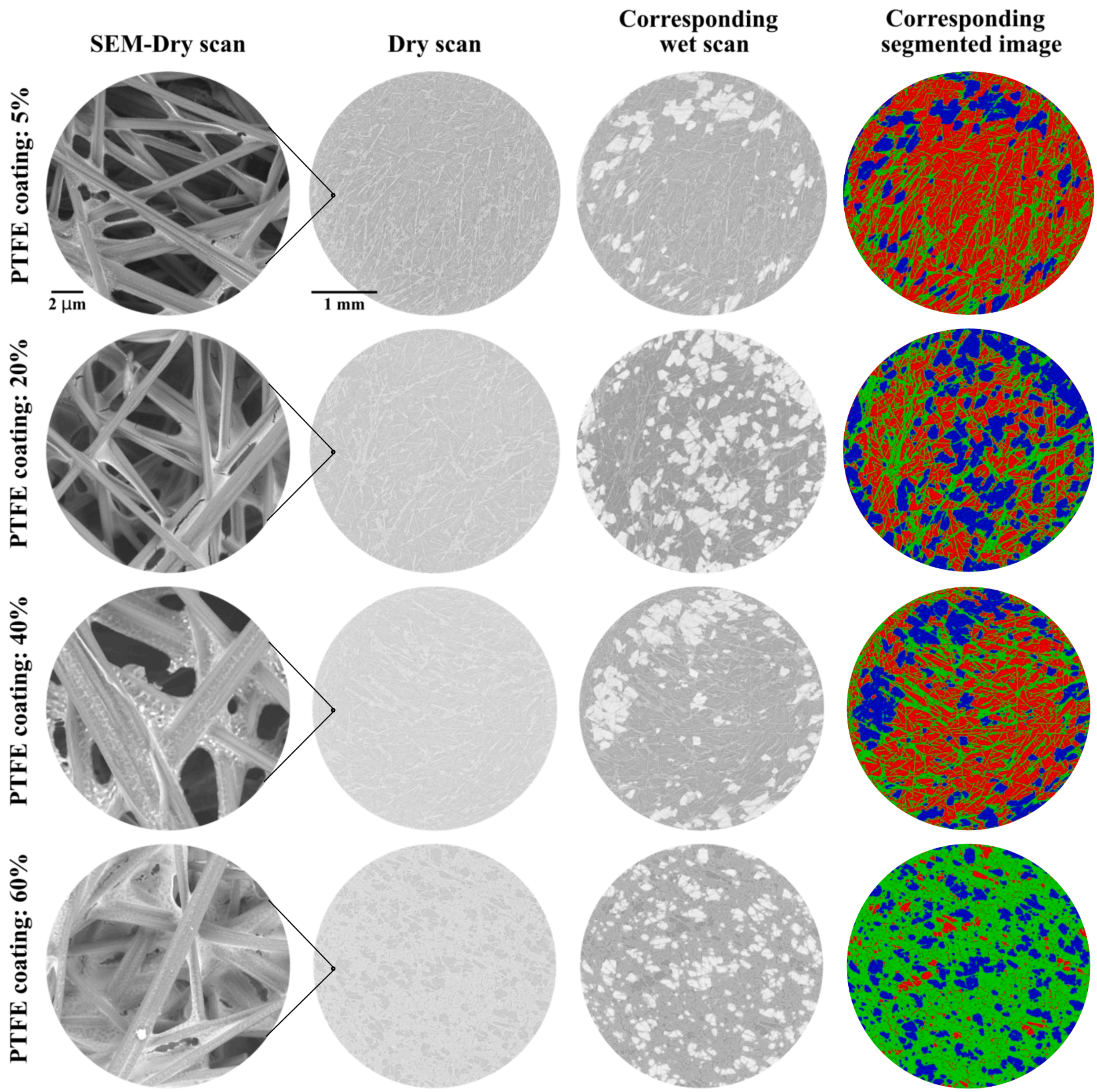

Fig. 1. Examples of SEM and X-ray images of different GDLs as presented in the legend. In the segmented images, the aqueous phase (brine), air, and fibres are shown in blue, red, and green, respectively. The voxel sizes of the SEM and micro-CT images were 0.01 and 2.05 µm, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Information about the scanner device settings and the resolution/dimensions of the output images.

Parameter	Value
Photon source energy	26 keV
The exposure time	900 ms
The number of projections per sample	2880
Voxel size	2.05 µm

[42–44] can be investigated using simulation studies. In image segmentation, the original grayscale images are portioned into discrete phases. In the case of GDL images, for instance, the segmentation algorithm recognizes and labels the fibres (containing both the fibres themselves and the PTFE binder), air and water. The segmentation can be challenging depending on the level of contrast between the intensity of objects, especially at the borders. Predictions of transport properties from simulations on segmented images are significantly influenced by segmentation errors which necessitates employing robust approaches to meticulously detect and delineate the phases in GDL images.

Traditional segmentation approaches can be categorised into manual, automatic, and semi-automatic techniques. Manual segmentation is a simple approach that takes advantage of user knowledge but requires extensive time and effort, which makes it impractical for large data with no reproducibility [45]. An alternative is using automatic tools to segment objects without the user's involvement. However, researchers have found they are less powerful than manual methods for detecting objects [46]. Hence, semi-automatic techniques were introduced to take advantage of both manual and automatic techniques [47–52]. Semi-automatic segmentation can be implemented using different methods, such as thresholding, in which the images are segmented using pre-defined intensity threshold values by the user. For

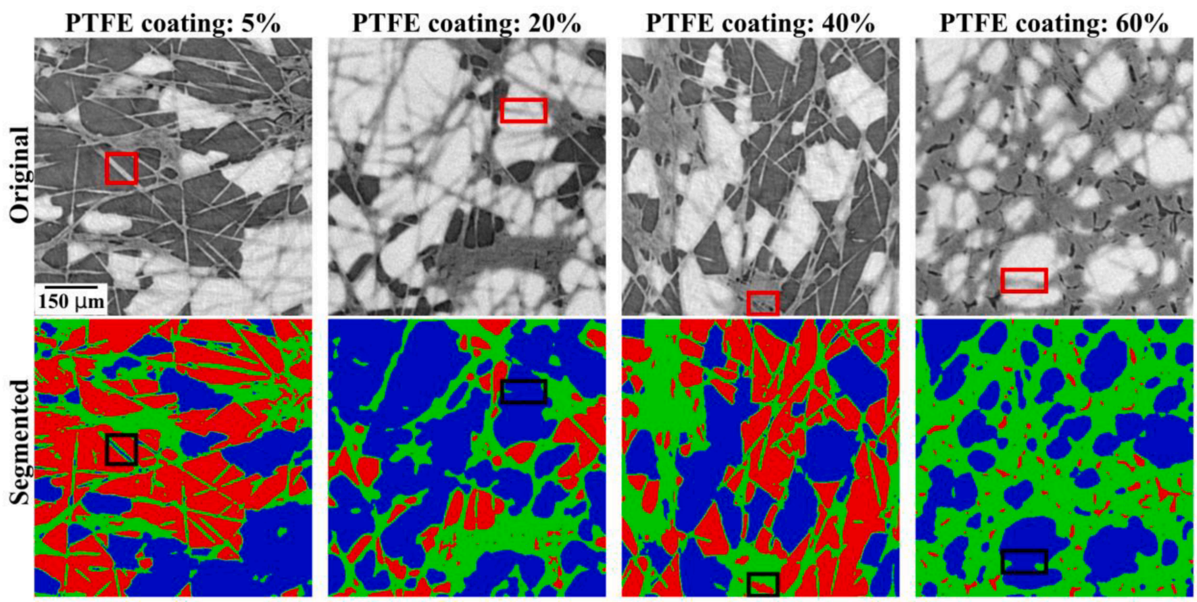

Fig. 2. Samples of the original GDL images (with 384 × 384 pixels) and the corresponding segmented images obtained using trainable 3D Weka segmentation. In the segmented images, the aqueous phase, air, and fibres are shown in blue, red, and green, respectively. Rectangles depict examples of possible errors in the segmentation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

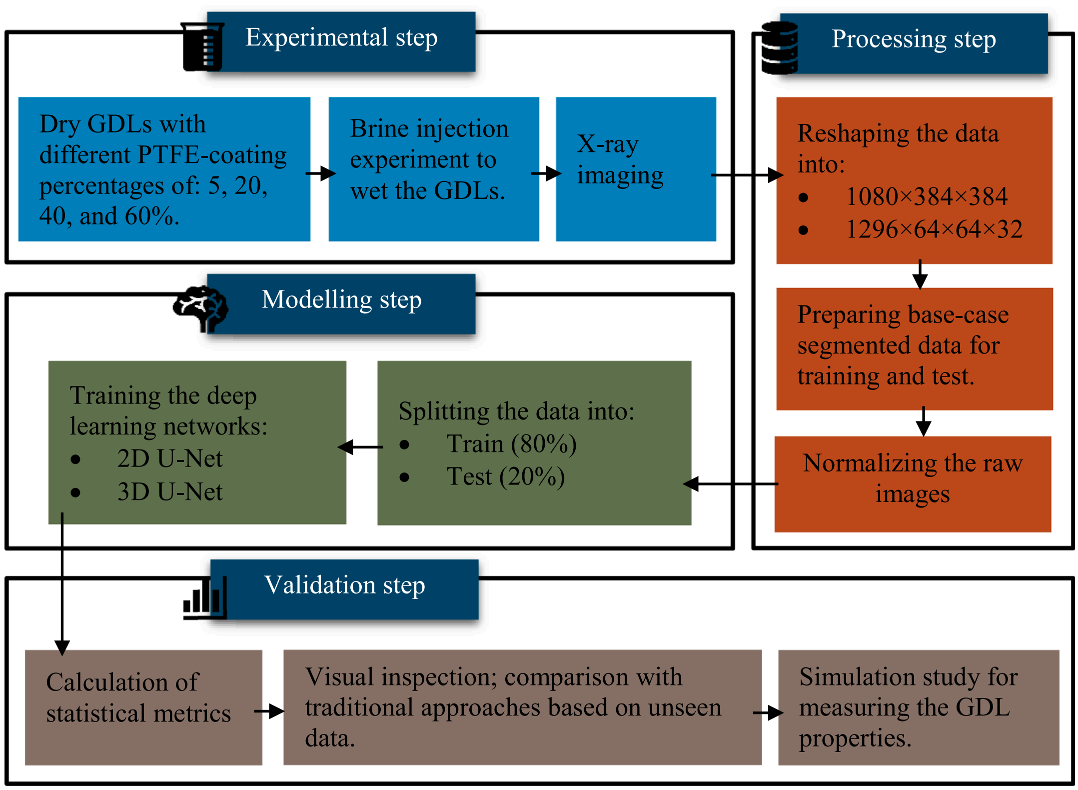

Fig. 3. The workflow for developing deep-learning-based models for simultaneous segmentation of wet images of GDLs.

example, region-based thresholding, such as watershed [53], selects a marker for each object defined by the user and groups analogous adjacent pixels. Edge-based thresholding recognises edge pixels of an object using edge-detection operators, and cluster-based algorithms such as trainable Weka [54] which clusters the pixels to reproduce typical shapes in the images. However, the major issue is the user bias introduced by tunable parameters. Moreover, the performance of the segmentation tools depends on the image quality, the application of a filter as a pre-processing step, and the image segmentation algorithm. In particular, image segmentation is very challenging for wet GDL images, with segmentation into solid, liquid and gas phases, as the contrast between the phases is low, and the intensities of the image are not unimodal [55].

Deep learning (DL) has emerged as a powerful tool to overcome these

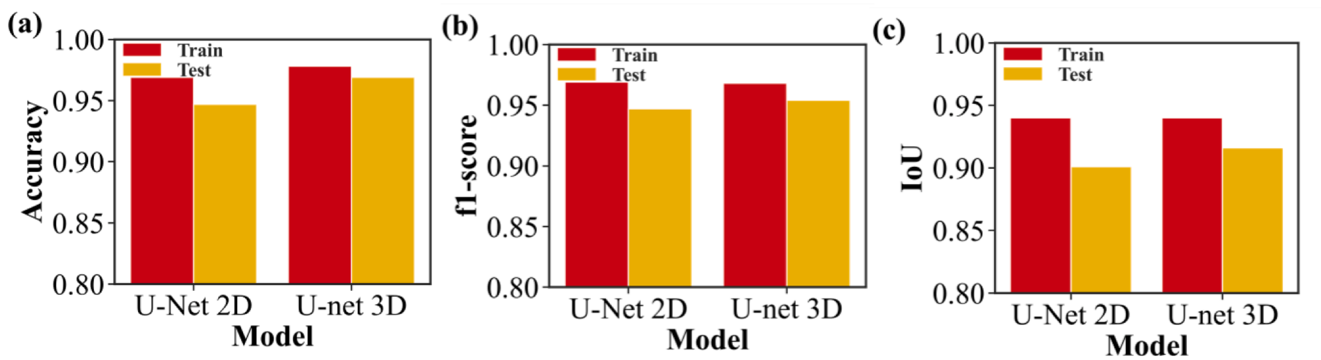

Fig. 4. Accuracy metrics: (a) accuracy, (b) f1-score, and (c) IoU metrics for U-Net 2D and U-Net 3D networks, Eqs. (4)-(6).

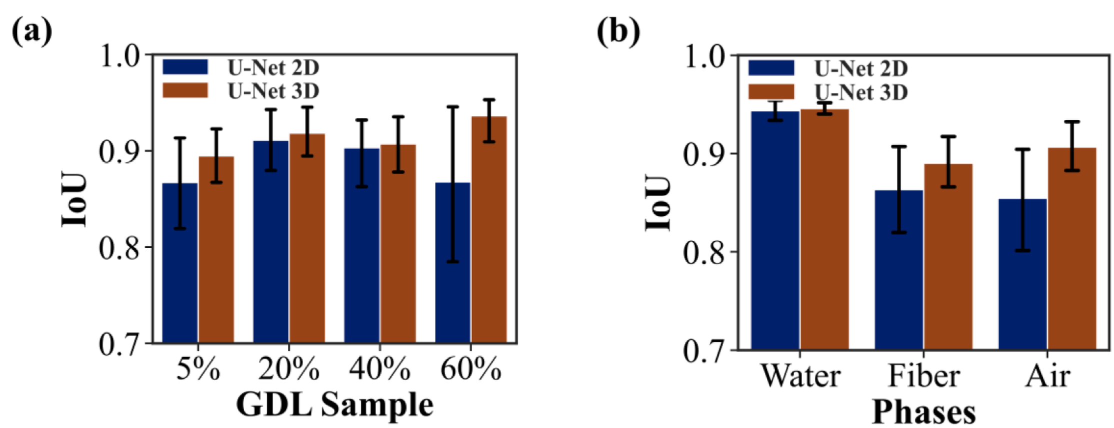

Fig. 5. A comparison between the performance of the 2D and 3D U-Net networks for segmentation of (a) the samples with different coating percentages and (b) various phases of each sample (water, fibre, and air). The error bars of subplot (a) indicate the range of IoUs of different phases, while that of (b) subplot implies the range of IoU values pertinent to GDLs with different PTFE coating percentages.

limitations. The DL algorithms learn high-level features in an incremental manner without explicit intervention of expertise. There are three different approaches to implementing DL models: supervised learning, where all input data are labelled; unsupervised learning, which works based on the unlabelled data; and semi-supervised learning, where the user labels some part of the dataset. Segmentation is usually performed in a supervised manner using original and ground truth segmented images as the input and outputs of the network, respectively. The segmentation's performance depends on the model type, configuration, and architecture. Many studies have assessed these factors for the segmentation of medical scans [45,56–60] and X-ray images of subsurface porous deposits [61–66]. However, there have been few studies on applying DL models for X-ray image segmentation of GDLs due to lack of data and the complex structure of the wet scans [55].

This paper evaluates the performance of deep learning models for simultaneous segmentation of X-ray images of GDLs with different coating percentages into water, air, and PTFE-coated fibre (or fibre for brevity) phases. To this end, we trained individual 2D and 3D U-Net architectures using unfiltered scans of wet GDLs and the corresponding base-case segmented images. Subsequently, the segmentation quality of deep learning models is compared to traditional approaches using statistical metrics, visual inspection, and flow simulations. Our results showed high accuracy and generality of the individual 2D and 3D models for multiphase segmentation of various GDL images without filtration and parameter tuning. Hence, the main contribution of this paper in fuel cell studies is facilitating accurate characterization of GDLs using X-ray images. This brings about a deeper understanding of the pore-scale phenomena inside the GDLs, which aids improving the performance of fuel cells.

The remainder of this study is organised as follows. The second section presents a brief overview of the experimental procedure, materials, X-ray imaging, and image processing. In the third section, deep learning networks and the training procedure are described in more detail, followed by the results and discussion, and conclusions.

2. Experimental procedure and materials

We imaged four GDL sets of AvCarb MGL 370 Carbon Papers with different coating percentages using a Heliscan X-ray scanner. More detail about the experiments can be found in our previous work [18]. The thickness and diameter of the circular GDLs were 0.37 and 6 mm, respectively. The samples were coated with 5, 20, 40, and 60 w% of a PTFE hydrophobic agent (the percentages are the mass ratio of the added PTFE to the untreated GDL). The porosity of these GDLs was calculated at 0.69, 0.67, 0.66, and 0.33 from the segmented images, respectively. Initially, both SEM and X-ray images of the samples were acquired with no water present (dry scan), as shown in Fig. 1. Then brine was injected at a rate of 7 µl/h and stopped after breakthrough. Another scan (wet scan) was taken after 2 h when the system had reached equilibrium.

More information about the imaging configurations for our experiments is provided in Table 1.

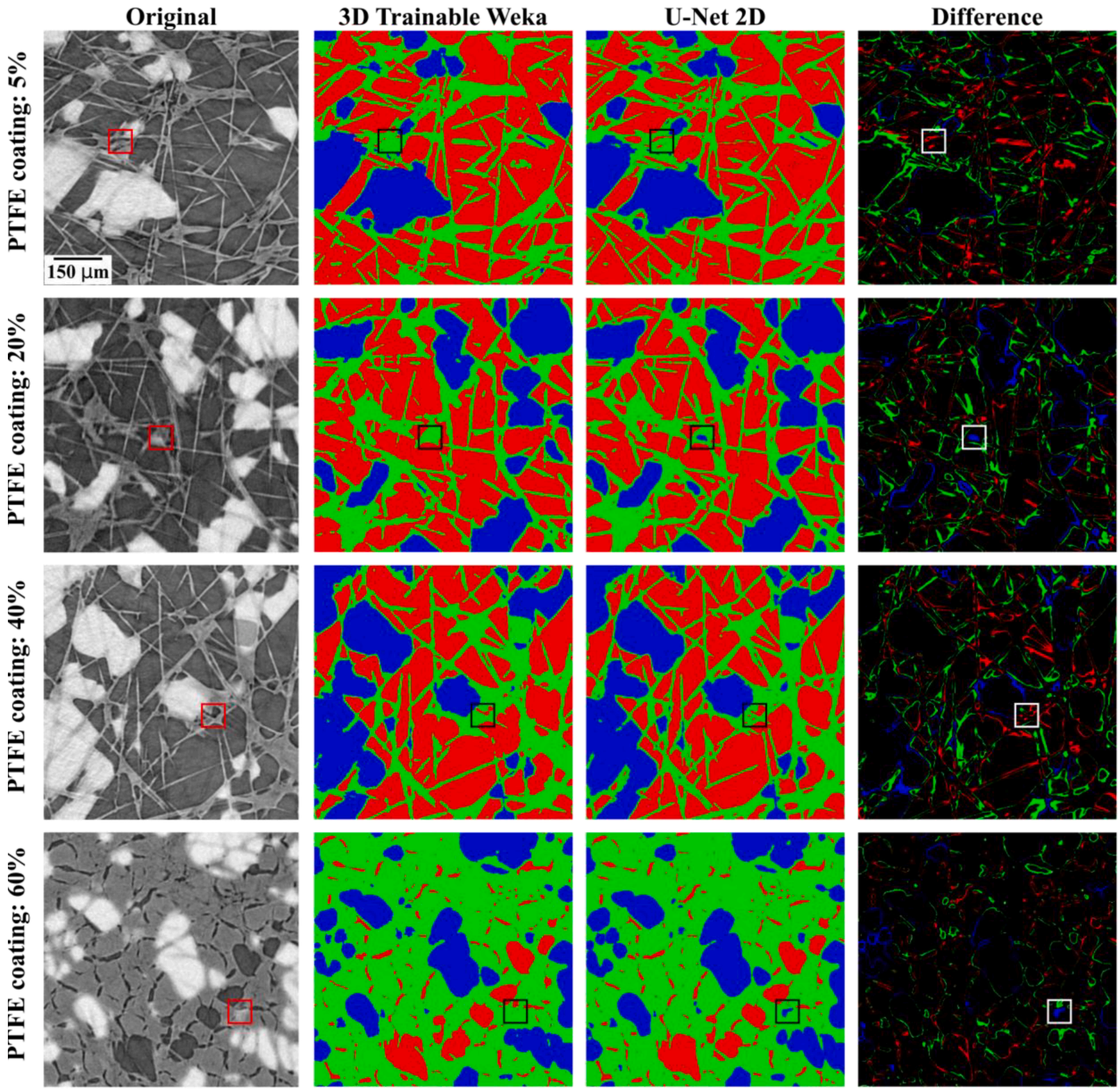

Fig. 6. A visual comparison between the performance of the trainable Weka and the U-Net 2D network. Blue, red, and green represent water, air, and fibre phases, respectively. In the fourth column, the areas with zero differences are depicted in black. Some regions where the U-Net outperforms 3D trainable Weka are highlighted by the squares, as examples. The images with dimensions of 384 × 384 were randomly selected from the test sub-data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Methodology

3.1. Image processing

We cropped the circular wet scans to 1152 × 1152 × 30 voxel volumes. The histogram of the image intensity with voxel frequency, ranging from 0 to 255, is presented in Fig. B.1 for each GDL. The figure shows the histogram is not easily divided into three distinct phases, and the intensity of the phases is not unique for all GDL sets, emphasising the challenge associated with the segmentation of all GDL images with a unique model. In this research, we explore the application of deep learning to segment these three phases accurately and simultaneously.

The semi-automatic 3D trainable Weka segmentation (ImageJ plugin version 1.53) [54] was employed for preparing the base-case dataset. In this approach, the user manually selects regions of interest (ROI) from each phase in a few images. Subsequently, the selected ROIs and the corresponding labels are introduced to a traditional machine learning algorithm such as Random Forest [67] to train to classify in a pixel-by-pixel manner. The rest of the images are segmented automatically thereafter. Our segmentation results showed that 3D trainable Weka segmentation is more accurate for providing base-case data than the watershed segmentation method. We applied this semi-automatic segmentation slice-by-slice instead of the whole dataset to obtain highly precise base-case datasets for training the deep learning model. Moreover, we applied the median filter to all images to remove noise while preserving the edges to improve Weka segmentation accuracy. Examples of the GDL images with the corresponding segmented data are represented in Fig. 2. While the segmentation correctly identifies the phases in many regions, there are still some errors, as highlighted with rectangles in Fig. 2.

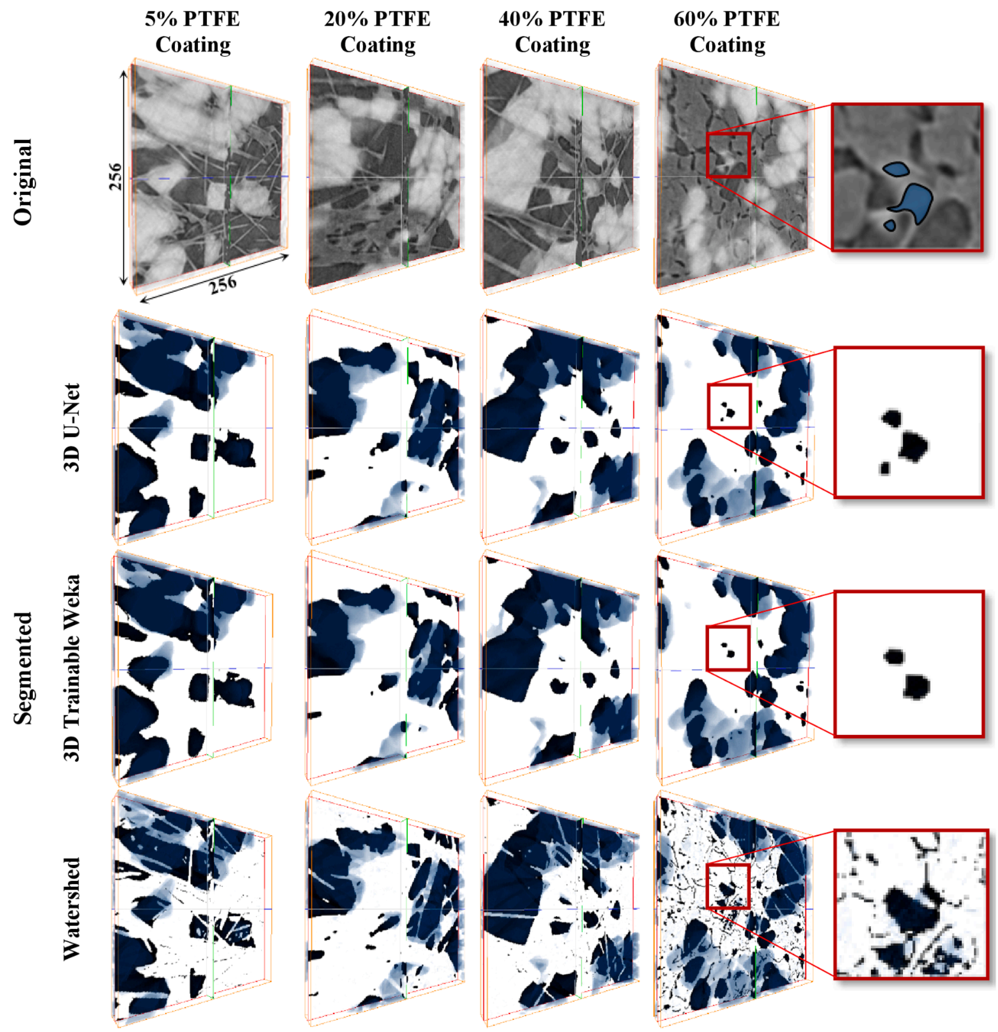

Fig. 7. A comparison between the performance of the U-Net 3D network, trainable Weka, and watershed segmentation for recognising and delineating the water in the GDL images (32 × 256 × 256). For a clear visualisation, the results of the 16th slice (the middle of the sample) are shown.

3.2. Training the 2D and 3D models

In this research, we used grayscale and Weka-segmented data (base-case) to train 2D and 3D autoencoders to segment unfiltered GDL images with different coating percentages accurately and simultaneously. Autoencoders are feed-forward backpropagation artificial neural networks (ANNs) that can learn data encodings without human supervision [68]. By training the network to extract the most salient features of an input image dataset, autoencoders can learn a representation (encoding) through a dimensionality reduction algorithm for higher-dimensional datasets [69,70]. In this case, autoencoders are similar to Principal Component Analysis (PCA), a widely-used data analysis method [68]. However, autoencoders are more flexible due to their ability to execute both linear and non-linear transformations during encoding, whereas PCA is limited to linear behaviour. Autoencoders discover significant features in the input data by minimising the reconstruction error between the input and output data [68,71]. Autoencoders have three major components in their structure: encoder, bottleneck, and decoder. The encoder is responsible for getting the input data and reducing its size to compress the data into a more compact form while keeping major features of the autoencoder structure [72,73]. The bottleneck is the most crucial layer of the network as it stores the main features of the dataset and transforms them from the encoder to the decoder [72–75]. It is worth noting, with a narrower bottleneck, overfitting is less likely to occur. The decoder is the last component that decompresses the reassembled data once it has been encoded, and then the results are checked against the ground truth (base-case) data [72–75]. Each autoencoder has the same number of neurons in its input and output layers to reconstruct the input data from lower dimensional features [68,71]. Unlike a regular data compression technique, autoencoders compress those data that are sufficiently close to the dataset on which they were trained.

Various convolutional encoder-decoder architectures such as SegNet [76] and U-Net [77,78] have been introduced for image segmentation. SegNet uses a combination of convolutional and max-pooling (down-sampling) layers to compress information into a bottleneck and generate an input representation. The decoder then rebuilds input information (using deconvolutional layers) to build a segmented map, highlighting the main features and classifying them appropriately. However, some of the essential features of the input data could be lost by feeding through multiple encoder blocks, but this issue has been addressed in the U-Net network. In the U-Net structure, skip connections have been used to bypass the bottleneck by connecting the encoder and the decoder directly [77,78]. So, the decoder takes feature maps from different representations and integrates them into a single map to reduce data loss.

In this study, 2D and 3D U-Net structures were trained to segment wet GDL X-ray images. Fig. 3 shows a flowchart summarising the steps we followed from the experiments to train and validate deep learning models. As previously mentioned, in the experimental step, four dry

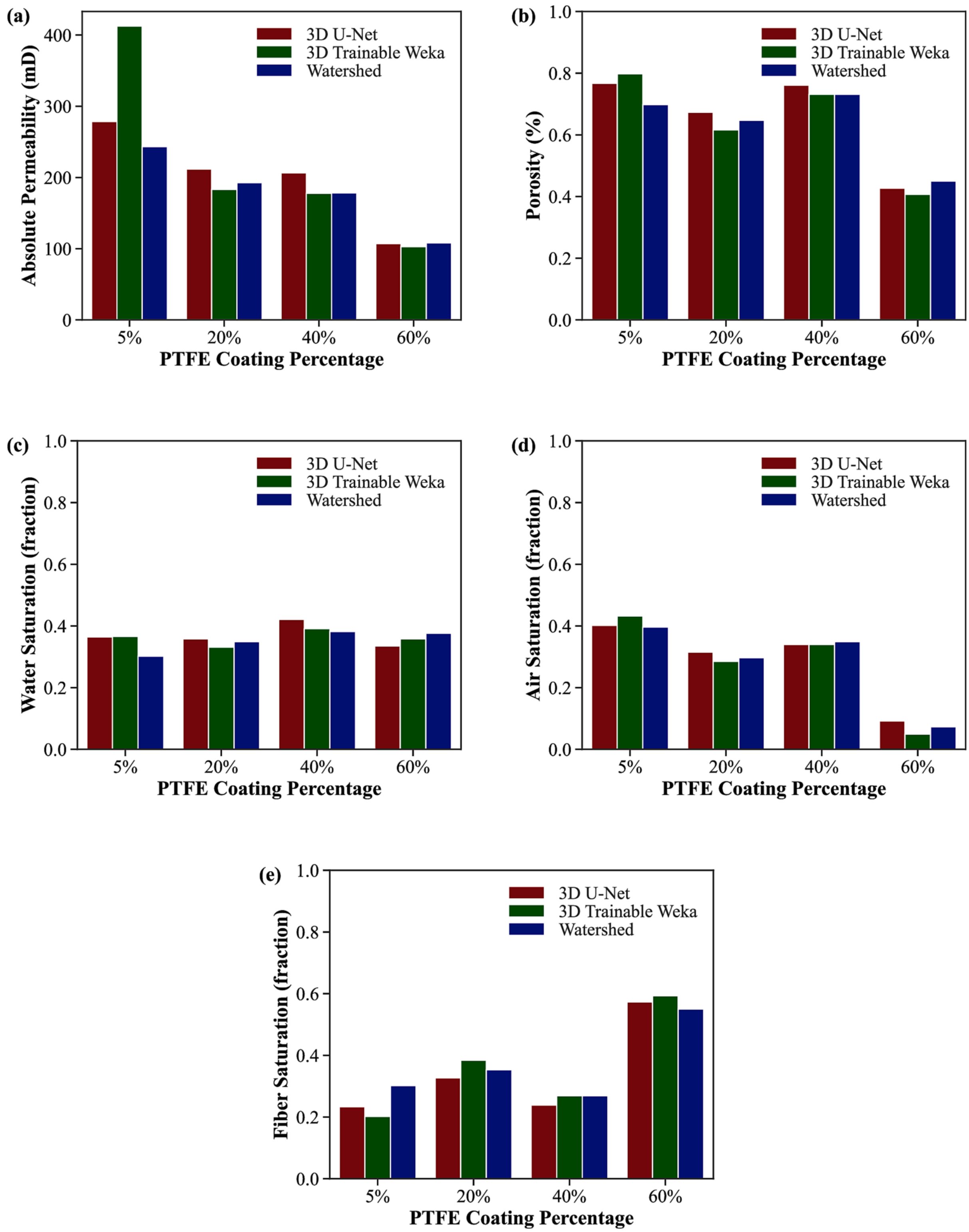

Fig. 8. The values of (a) absolute permeability (mD), (b) porosity (%), (c) water saturation (fraction), (d) air saturation (fraction), and (e) fibre saturation (fraction) for different GDL images (coated by 5, 20, 40 and 60 w% of PTFE) that have been segmented using various approaches of 3D U-Net, 3D trainable Weka, and watershed segmentation.

GDLs with different PTFE-coating percentages from 5 to 60% were wetted by brine injection and imaged using a high-resolution micro-CT device. In the processing step, a total of 1080 grey-scale images of 384 × 384 pixels (comprising 270 scans per GDL set) and 1296 volumes with 64 × 64 × 32 voxels (324 images per GDL set) were prepared for 2D and 3D modelling, respectively. The Weka-segmented images were assumed as base-case (ground truth) masks. After image normalisation, 80% of them were used as the training dataset and the other 20% as the test (validation) dataset. Subsequently, the 2D and 3D U-Nets were trained in a supervised manner for the segmentation task. Both the U-Net structures used the Adam optimiser [79] with a learning rate of 10^{-4} for their optimisation algorithm. A combination of focal loss [80] and dice

Table A1
The training and testing metric values of accuracy, F1-score, and IoU for the 2D and 3D U-Net autoencoders.

Model	Sub-data	Accuracy	F1-score	IoU
U-Net 2D	Train	0.969	0.969	0.940
	Test	0.947	0.947	0.901
U-Net 3D	Train	0.978	0.968	0.940
	Test	0.969	0.954	0.916

Table A2
The IoU values for the 2D and 3D U-Net networks for different PTFE coating percentages (5%, 20%, 40%, and 60%) and various phases of water, fibre, and air.

Model		Water		Fibre		Air	
		Train	Test	Train	Test	Train	Test
U-Net 2D	5%	0.929	0.919	0.797	0.772	0.901	0.884
	20 %	0.966	0.959	0.872	0.854	0.914	0.901
	40 %	0.939	0.938	0.927	0.805	0.912	0.898
	60 %	0.874	0.861	0.956	0.943	0.940	0.940
U-Net 3D	5%	0.936	0.931	0.852	0.856	0.937	0.856
	20 %	0.959	0.957	0.884	0.887	0.936	0.887
	40 %	0.950	0.948	0.872	0.865	0.943	0.865
	60 %	0.943	0.943	0.948	0.957	0.871	0.957

loss [81] was used as the loss function. Focal loss is defined by [80]:

$$FL = -(1-p_t)^{\gamma}\log(p_t) \tag{1}$$

where p_t is the probability of the ground truth class and γ is a tuning parameter to put more focus on misclassified examples. Dice loss is widely used for semantic segmentation tasks, especially in medical applications, and is defined [82]:

$$DL = 1 - \frac{2\sum_{i=1}^{N} p_i g_i}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} g_i^2} \tag{2}$$

where p_i stands for the predicted probability of the i^{th} pixel and g_i refers to the ground truth of the i^{th} pixel. The final loss function is a combination of these two:

$$Total\ loss = DL + (w \times FL) \tag{3}$$

where w is a weighting factor that controls the influence of focal loss. To the trade-off between bias and overfitting, we found the optimum epoch number for training 2D and 3D U-Net models equal to 20. A schematic of the structure implemented for the 2D U-Net is illustrated in Fig. B.2. The same structure is used for the 3D U-Net, with the difference of using 3D convolution, deconvolution, and max-pooling layers instead of 2D layers. At last, in the validation step, the trained models were assessed in different steps, with the procedures and results presented in the following section.

4. Results and discussion

As discussed before, unique 2D and 3D U-Net models were trained using wet GDL datasets to simultaneously segment all the GDL sets with different coating percentages into water, air, and fibre. We employed various techniques such as calculation of the statistical metrics, visual inspection, and simulation to evaluate the accuracy of the deep learning models (see Fig. 3).

4.1. Validation based on the statistical metrics

The overall train and test performance were assessed by the accuracy, f1-score, and IoU metrics. The accuracy metric is simply the ratio of the number of pixels/voxels with the same segmentation label as the base-case to the total number of pixels/voxels:

$$\text{Accuracy} = \frac{\text{The number of accurate predictions}}{\text{The total number of pixels/voxels}} \tag{4}$$

We also evaluated the dice similarity coefficient [81] (or f1-score) and the Jaccard Index [83] (or so-called intersection over union (IoU)). Consider two images, A and B. A is the U-Net segmented image and B is the base-case Weka-segmented image. The two images have the same total number of voxels, N, and are segmented into a phase labelled i (1, 2, and 3 for air, water, and fibre, respectively). We consider one

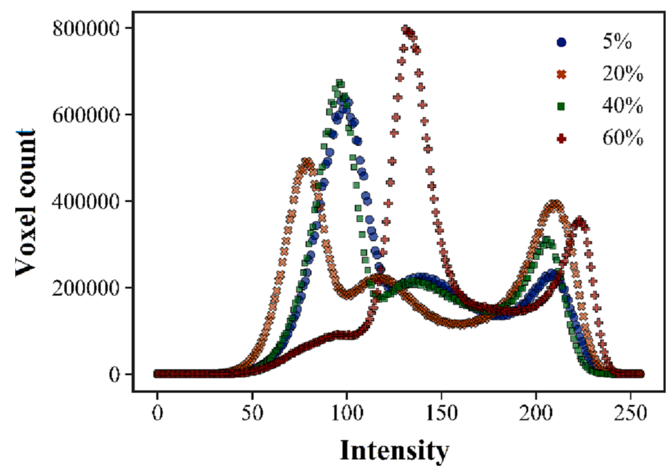

Fig. B.1. The histogram of the four GDLs with different percentages of the PTFE as indicated in the legend. The histogram represents the number of pixels with each grey-scale value.

Table A3
A detailed representation of the values of absolute permeability (mD), porosity (%), water saturation (fraction), air saturation (fraction), and fibre saturation (fraction) pertinent to different GDL images (coated by 5%, 20%, 40% and 60% PTFE) that have been segmented using various methods of 3D U-Net, 3D Trainable Weka, and watershed.

Samples (PTFE%)	Segmentation Method	Water Saturation	Fibre Saturation	Air Saturation	Porosity	Absolute Permeability (mD)
5%	3D U-Net	0.364	0.233	0.402	0.767	278.57
	3D Trainable Weka	0.366	0.202	0.432	0.798	412.48
	Watershed	0.302	0.302	0.396	0.698	243.34
20%	3D U-Net	0.358	0.327	0.315	0.673	211.89
	3D Trainable Weka	0.331	0.384	0.285	0.616	183.24
	Watershed	0.349	0.353	0.297	0.647	192.69
40%	3D U-Net	0.421	0.239	0.340	0.761	206.50
	3D Trainable Weka	0.391	0.269	0.340	0.731	177.78
	Watershed	0.382	0.269	0.349	0.731	178.31
60%	3D U-Net	0.335	0.573	0.092	0.427	107.09
	3D Trainable Weka	0.358	0.593	0.049	0.407	102.84
	Watershed	0.376	0.550	0.073	0.450	108.17

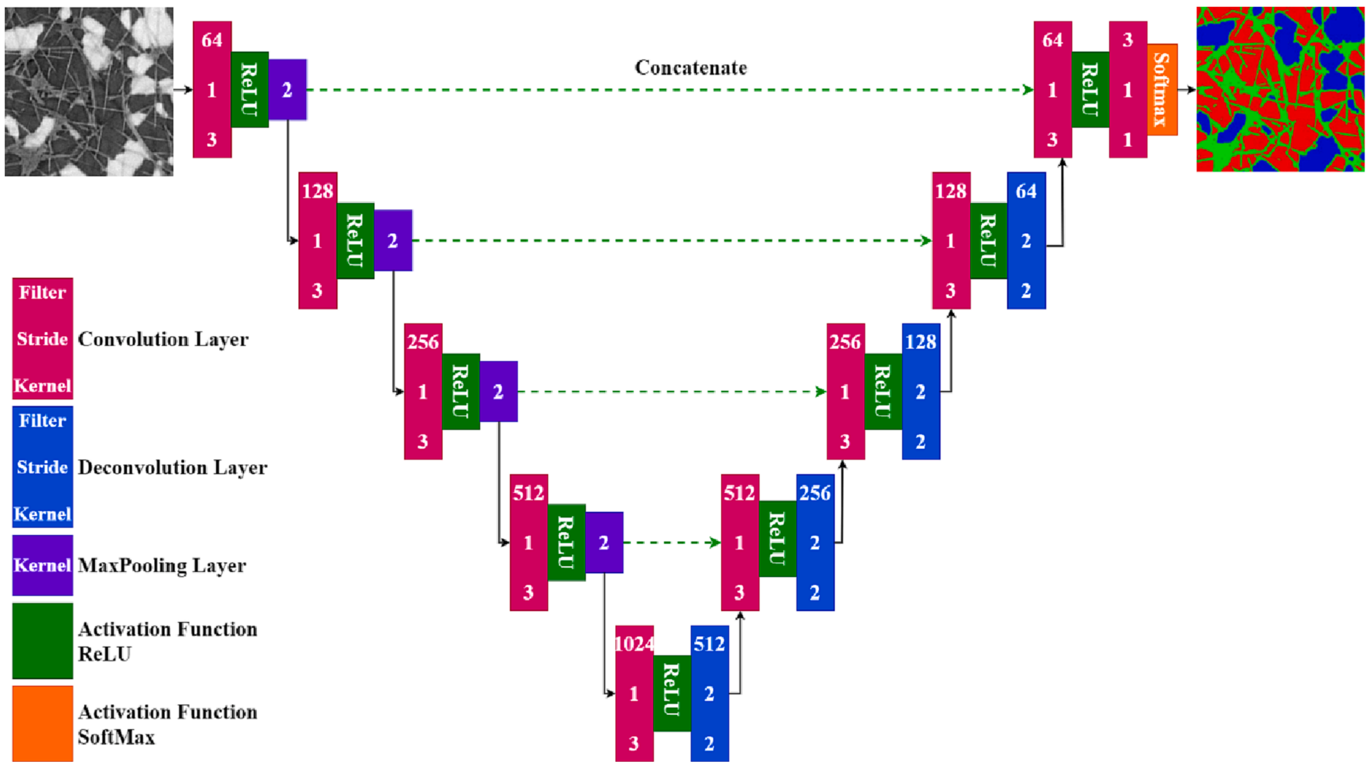

Fig. B.2. The architecture of the 2D U-Net autoencoder used in this study. The 3D U-Net uses a similar architecture, with the differences using of 3D convolution, deconvolution, and max-pooling layers instead of 2D layers.

phase at a time. A_i is the number of voxels of label i in image A; B_i is the number of voxels of label i in image B. I_i is the number of voxels for which the label is i in *both* images. U_i is the number of voxels for which the label is i in *either* image A or image B. $N \geq U_i \geq A_i, B_i \geq I_i \geq 0$.

We then define the f1-score, F_i, as follows:

$$F_i = \frac{2I_i}{A_i + B_i} \tag{5}$$

The IoU, IU_i, is defined as:

$$IU_i = \frac{I_i}{U_i} \tag{6}$$

As explained earlier, the training and test procedures were implemented based on 1080 wet images of 384 × 384 pixels (270 scans per GDL), and 1296 volumes of 64 × 64 × 32 voxels (324 per GDL) for the 2D and 3D U-Nets, respectively. 80% of the data were specified for the training, and the rest were used for testing the models.

First, we calculate the average values of F_i and IU_i for the three phases: $F = (F_1 + F_2 + F_3)/3$; $IU = (IU_1 + IU_2 + IU_3)/3$ and then take the mean over all the images. These are shown in Fig. 4 (see Table A.1 for more details). The mean IoU of 2D and 3D U-Net are 0.940 and 0.940 for training, and 0.901 and 0.916 for testing, respectively. It is worth noting that we used unfiltered images for the training and testing procedure of the models. In contrast, pre-processing was performed to prepare the base-case dataset, highlighting the model's ability to segment low-quality image datasets. The minor differences between the accuracies for training and testing steps revealed our models are robust to over-fitting, which is a common problem in deep learning applications. While the performance of the two networks is similar, the 3D U-Net does slightly better.

To evaluate the 2D and 3D models in more detail, the mean IoU is calculated for each GDL (270 images of 384 × 384 pixels and 324 volumes of 64 × 64 × 32 voxels per GDL set for the 2D and 3D models, respectively) and each phase (water, air, and fibre). The results are presented in Fig. 5 (further details are provided in Table A.2). The minor differences between the IoUs, in Fig. 5(a) reveal that both the 2D and 3D models are well generalised to segment different GDLs and phases. Moreover, Fig. 5(a) shows 3D U-Net model performs better than the 2D U-Net model; this can also be confirmed by Fig. 5(b). Fig. 5(b) also shows that the deep learning-based models achieve better results for segmenting the water phase, due to the higher contrast between the water phase and the other phases.

4.2. Validation based on visual inspection

As shown in Fig. 2, the base-case dataset does not necessarily capture exactly the true phase distribution in the original dataset. Hence, we used visual inspections and qualitative comparisons to evaluate if deep learning models outperformed and improved segmentation compared to the base-case data. Fig. 6 shows an example of the outcome of the 2D U-Net model alongside the image obtained by the 3D trainable Weka method. Visual inspection reveals that the 2D U-Net provides a better segmentation in many areas, as highlighted by the squares in Fig. 6. Accordingly, the differences between these two outcomes (fourth column) do not necessarily imply errors in the U-Net model. It needs to be emphasised that contrary to trainable Weka where segmentation is performed slice by slice, the deep learning network has simultaneously segmented all the unfiltered GDL sets with different structures and textures. These outcomes highlight the advantages of deep learning in terms of time, user effort, reproducibility, and accuracy.

The 3D U-Net model was assessed on the unfiltered unseen data that were excluded from the training/testing procedure. For this purpose, four 3D images (from different PTFE-coated GDLs) with 256 × 256 × 32 voxels were extracted from the original data to validate the 3D U-Net model. A snapshot of the outcome segmented phases is shown in Fig. B.3. Visual examination indicates that one trained model appropriately recognised and delineated the water, air, and fibres for each GDL.

We compared the outcome of the 3D U-Net, 3D trainable Weka, and

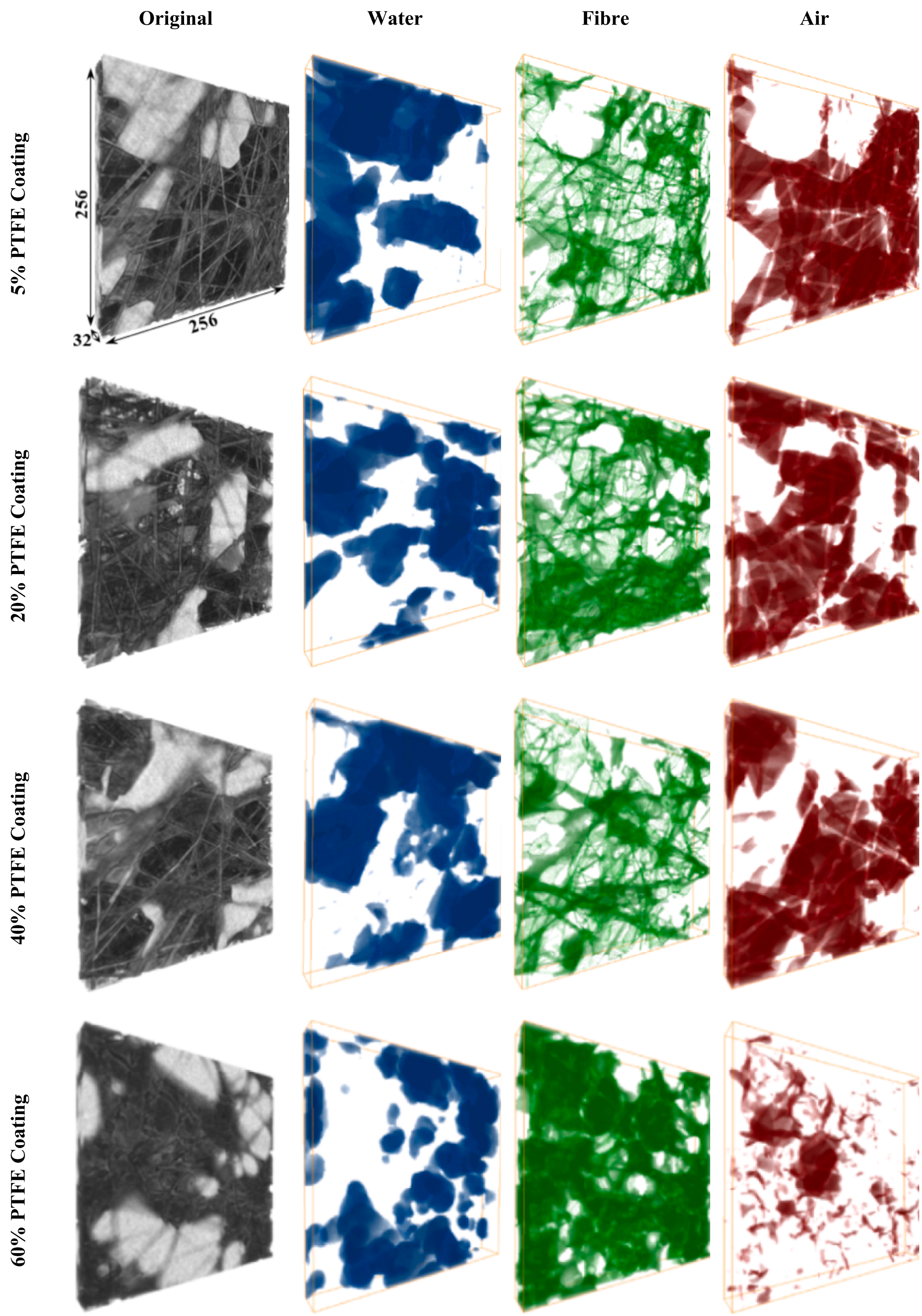

Fig. B.3. A visual representation of the performance of the U-Net 3D model for segmentation of unseen 3D data with 256 × 256 × 32 voxels.

watershed approaches to segment the unseen data. The results for the water phase are depicted in Fig. 7 (see Figs. B.4 and B.5 in Appendix B for the air and fibre phases, respectively). A general view of the images reveals the superiority of the 3D U-Net for segmenting all the phases over the two other methods. A more detailed inspection was performed by focusing on a highlighted square region of the GDL with 60% PTFE coating. As Fig. 7 shows, the water phase is more accurately distinguished by the 3D U-Net compared to the other methods, while under and over-estimations can be observed in the case of the 3D trainable Weka and watershed techniques, respectively. The same results for the fibre (Fig. B.4) and air (Fig. B.5) have been observed. This demonstrates that deep learning can provide better segmentation results than the base-case dataset (generated by 3D trainable Weka), and watershed segmentation. This can be explained by the fact that while the Weka method correctly segmented most of the images, there were mistakes in some small regions (see Fig. 2). However, the deep learning model correctly learns typical structures from the most accurate part of the base-case dataset and then correctly segments all the unseen datasets.

4.3. Simulation of permeability

We performed pore-scale simulations based on the unseen images

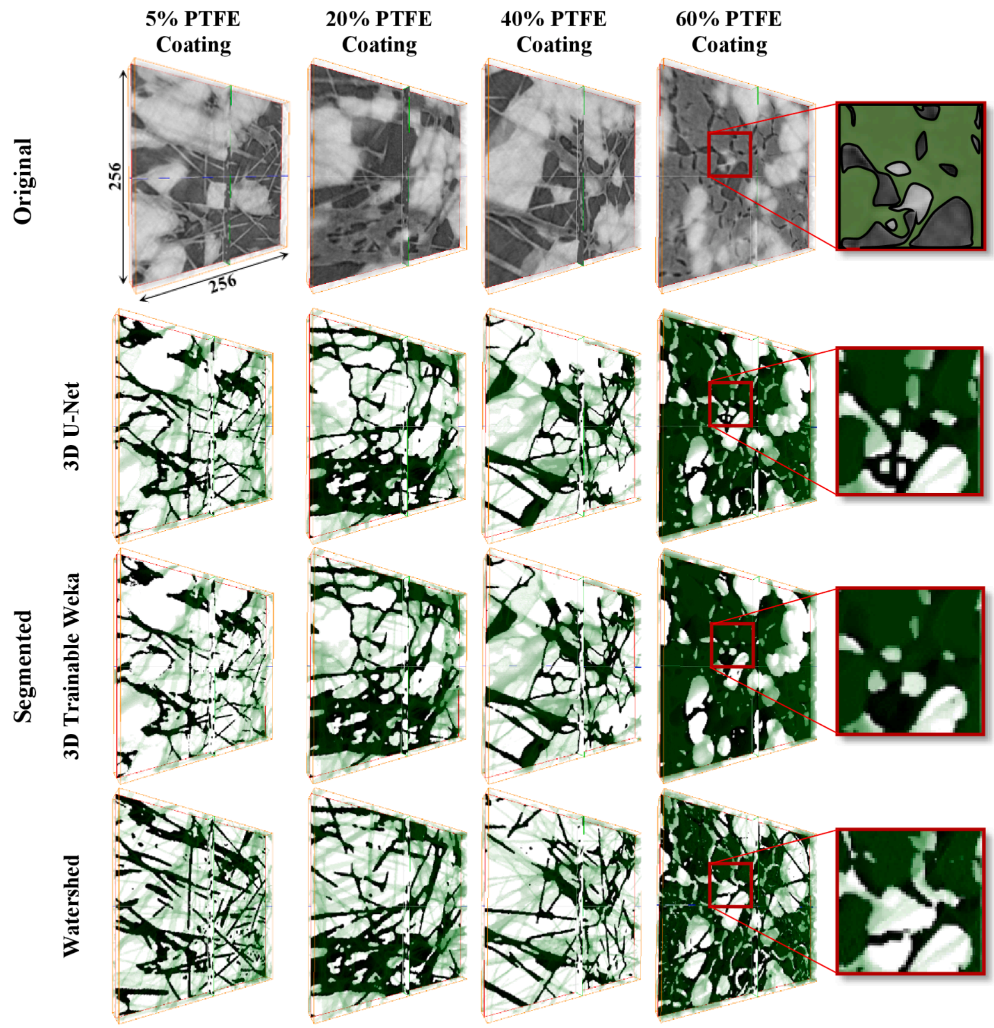

Fig. B.4. A comparison between the performance of the U-Net 3D network, trainable Weka, and watershed for recognising and delineating the fibres in the GDL images (32 × 256 × 256). For a clear visualisation, the 16th slice (the middle of the sample) is shown.

segmented by 3D U-Net to calculate the absolute permeability to highlight the performance of deep learning models. Subsequently, the simulation was repeated based on the Weka and watershed segmented images to analyse how different segmentations affect predictions of flow properties. For this purpose, first, we obtained the saturation of each phase in the unseen images of Fig. B.4 (four images with the dimension of 256 × 256 × 32 voxels) by dividing the voxels related to each phase by the total number of voxels. Accordingly, the porosity of each sample was calculated. Subsequently, single-phase flow of water was simulated on the binarized images (just pore space and fibre) obtained by different segmentation approaches. For executing the flow simulation, the Stokes equation was solved with absolute convergence criteria of 10^{-6} to obtain the pressure and velocity fields with the PerGeos v. 2019.3 software package. The absolute permeability values for each GDL in the z-direction (perpendicular to the GDL layer) were calculated. The results are shown in Fig. 8, and Table A.3. Fig. 8(b)–(e) illustrates how saturation and porosities are affected by segmentation. Although these differences in porosity are small, they have a significant impact on the predicted dynamic parameter, absolute permeability. For instance, the 5% PTFE-coated GDL had estimated porosities of 76.7, 79.8, and 69.8%, based on deep learning, Weka, and watershed methods, and permeabilities of 279, 412, and 243 mD, respectively. Using the 3D U-Net results as the most accurate as discussed above, the Weka method led to a 43, 14, 14, and 3.9% deviation in the absolute permeability calculation for the 5, 20, 40, and 60% PTFE-coated GDLs, respectively. The differences are 13, 7.2, 14, and 1%, respectively, for the watershed segmentation method. Even though the difference in porosity is larger in this case it appears to have less of an effect on the flow.

Overall, our findings highlight the advantage of the 2D and 3D autoencoders for multiphase segmentation of images of GDLs with complex textures. The results indicated that these smart approaches can segment different GDLs with no pre-processing. Furthermore, we showed how segmentation can impact the accuracy of the calculation of porous media properties. The aforementioned results justify the further application of state-of-the-art artificial intelligence techniques to segment complex porous media such as the GDLs considered in this study.

5. Summary and conclusions

The current study was undertaken to improve multiphase segmentation of GDL X-ray images with different coating percentages using two autoencoders: 2D and 3D U-Net. The experimental images were obtained after brine injection through four GDL samples pre-coated by various percentages of a hydrophobic agent, PTFE. 2D and 3D U-Net models were trained in a supervised manner to simultaneously segment the water, air, and PTFE-coated fibres without image processing. The model performance was evaluated using statistical metrics, visual inspection, and calculation of the GDL properties. Our main findings are as follows.

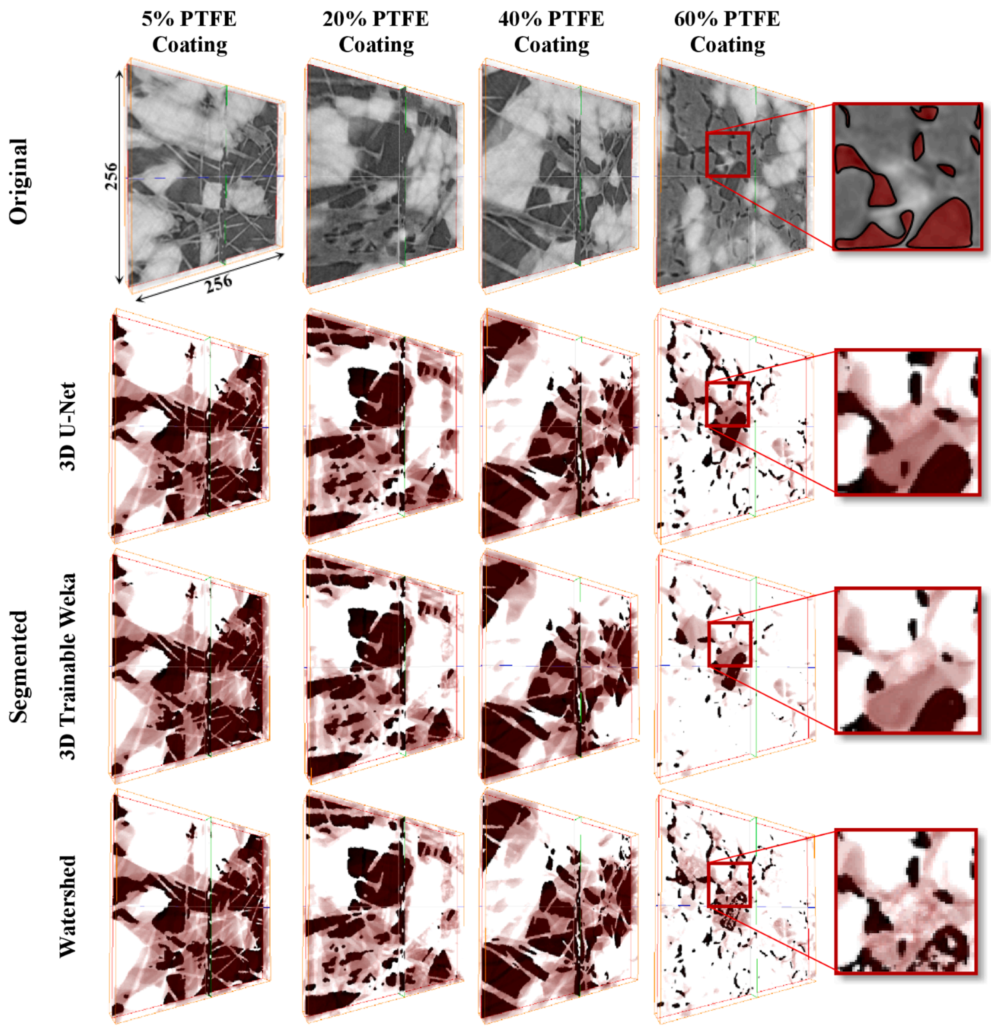

Fig. B.5. A comparison between the performance of the U-Net 3D network, trainable Weka, and watershed for recognising and delineating the air in the GDL images (32 × 256 × 256). For a clear visualisation, the 16th slice (the middle of the sample) is shown.

- Both 2D and 3D U-Net models showed high accuracy with the test IoU values of 0.901 and 0.916, and f1-scores of 0.947 and 0.954, respectively. The 3D models marginally outperform the 2D model.
- A visual inspection revealed that the deep learning model outperforms both 3D trainable Weka and watershed segmentation methods.
- We simulated single-phase flow in the segmented images. Our results indicated a noticeable difference between the absolute permeabilities calculated. Considering the 3D U-Net as the most reliable tool based on the previous analysis, for GDLs coated by 5, 20, 40, and 60% of PTFE, 3D trainable Weka was associated with 43, 14, 14, and 3.9%, and the watershed method with 13, 7.2, 14, and 1% differences, respectively, in the predicted permeability.
- Future work could apply these methods to a wider range of GDL images, and to segment the solid structure and fluid phases for a variety of porous materials.

CRediT authorship contribution statement

Mehdi Mahdaviara: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Visualization. **Mohammad Javad Shojaei:** Conceptualization, Validation, Investigation, Resources, Data curation, Writing – review & editing. **Javad Siavashi:** Conceptualization, Software, Writing – original draft. **Mohammad Sharifi:** Validation, Writing – review & editing, Supervision. **Martin J. Blunt:** Resources, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The codes used in this article can be found online at https://github.com/mma9631/2D-3D-UNet-for-GDL-segmentation.git.

Appendix A

Tables A.1–A.3.

Appendix B

Figs. B.1–B.5.

References

[1] Blunt MJ, Bijeljic B, Dong Hu, Gharbi O, Iglauer S, Mostaghimi P, et al. Pore-scale imaging and modelling. Adv Water Resour 2013;51:197–216.

[2] Cnudde V, Boone MN. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. Earth Sci Rev 2013;123:1–17.
[3] Blunt MJ. Multiphase Flow in Permeable Media: A Pore-Scale Perspective. Cambridge: Cambridge University Press; 2017.
[4] Bakhshian S, Hosseini SA, Shokri N. Pore-scale characteristics of multiphase flow in heterogeneous porous media using the lattice Boltzmann method. Sci Rep 2019;9(1):3377.
[5] Xiao L, Zhu L, Clökler C, Grünzweig A, Wilhelm F, Scholta J, et al. Experimental validation of pore-scale models for gas diffusion layers. J Power Sources 2022;536:231515.
[6] Hashemi L, Blunt M, Hajibeygi H. Pore-scale modelling and sensitivity analyses of hydrogen-brine multiphase flow in geological porous media. Sci Rep 2021;11(1):8348.
[7] Zhang Y, Bijeljic B, Blunt MJ. Nonlinear multiphase flow in hydrophobic porous media. J Fluid Mech 2022;934.
[8] Shams M, Singh K, Bijeljic B, Blunt MJ. Direct Numerical Simulation of Pore-Scale Trapping Events during Capillary-Dominated Two-Phase Flow in Porous Media. Transp Porous Media 2021;138(2):443–58.
[9] Mostaghimi P, Blunt MJ, Bijeljic B. Computations of Absolute Permeability on Micro-CT Images. Math Geosci 2013;45(1):103–25.
[10] Saxena N, Hofmann R, Alpak FO, Dietderich J, Hunter S, Day-Stirrat RJ. Effect of image segmentation & voxel size on micro-CT computed effective transport & elastic properties. Mar Pet Geol 2017;86:972–90.
[11] Andrä H, Combaret N, Dvorkin J, Glatt E, Han J, Kabel M, et al. Digital rock physics benchmarks—Part I: Imaging and segmentation. Comput Geosci 2013;50:25–32.
[12] Missbach-Guentner J, Pinkert-Leetsch D, Dullin C, Ufartes R, Hornung D, Tampe B, et al. 3D virtual histology of murine kidneys –high resolution visualization of pathological alterations by micro computed tomography. Sci Rep 2018;8(1).
[13] Wang LV. Multiscale photoacoustic microscopy and computed tomography. Nat Photonics 2009;3(9):503–9.
[14] Momose A, Takeda T, Itai Y, Hirano K. Phase–contrast X–ray computed tomography for observing biological soft tissues. Nat Med 1996;2(4):473–5.
[15] Rabbani A, Babaei M, Gharib M. Automated segmentation and morphological characterization of placental histology images based on a single labeled image. arXiv preprint arXiv:221003566 2022.
[16] Siavashi J, Najafi A, Ebadi M, Sharifi M. A CNN-based approach for upscaling multiphase flow in digital sandstones. Fuel 2022;308:122047.
[17] Najafi A, Siavashi J, Ebadi M, Sharifi M, Fahimpour J, Koroteev D. Upscaling permeability anisotropy in digital sandstones using convolutional neural networks. J Nat Gas Sci Eng 2021;96:104263.
[18] Shojaei MJ, Bijeljic B, Zhang Y, Blunt MJ. Minimal surfaces in porous materials: X-ray image-based measurement of the contact angle and curvature in gas diffusion layers to design optimal performance of fuel cells. ACS Appl Energy Mater 2022;5(4):4613–21.
[19] Peng X, Kulkarni D, Huang Y, Omasta TJ, Ng B, Zheng Y, et al. Using operando techniques to understand and design high performance and stable alkaline membrane fuel cells. Nat Commun 2020;11(1).
[20] Meyer Q, Ashton S, Boillat P, Cochet M, Engebretsen E, Finegan DP, et al. Effect of gas diffusion layer properties on water distribution across air-cooled, open-cathode polymer electrolyte fuel cells: A combined ex-situ X-ray tomography and in-operando neutron imaging study. Electrochim Acta 2016;211:478–87.
[21] Haile SM. Fuel cell materials and components. Acta Mater 2003;51(19):5981–6000.
[22] Liu CP, Saha P, Huang Y, Shimpalee S, Satjaritanun P, Zenyuk IV. Measurement of Contact Angles at Carbon Fiber–Water–Air Triple-Phase Boundaries Inside Gas Diffusion Layers Using X-ray Computed Tomography. ACS Appl Mater Interfaces 2021;13(17):20002–13.
[23] Pollet BG, Franco AA, Su H, Liang H, Pasupathi S. 1 - Proton exchange membrane fuel cells. In: Barbir F, Basile A, Veziroğlu TN, editors. Compendium of Hydrogen Energy. Oxford: Woodhead Publishing; 2016. p. 3–56.
[24] Park G-G, Sohn Y-J, Yang T-H, Yoon Y-G, Lee W-Y, Kim C-S. Effect of PTFE contents in the gas diffusion media on the performance of PEMFC. J Power Sources 2004;131(1):182–7.
[25] Lee H-K, Park J-H, Kim D-Y, Lee T-H. A study on the characteristics of the diffusion layer thickness and porosity of the PEMFC. J Power Sources 2004;131(1):200–6.
[26] Kleemann J, Finsterwalder F, Tillmetz W. Characterisation of mechanical behaviour and coupled electrical properties of polymer electrolyte membrane fuel cell gas diffusion layers. J Power Sources 2009;190(1):92–102.
[27] Feser JP, Prasad AK, Advani SG. Experimental characterization of in-plane permeability of gas diffusion layers. J Power Sources 2006;162(2):1226–31.
[28] Mularczyk A, Lin Q, Blunt MJ, Lamibrac A, Marone F, Schmidt TJ, et al. Droplet and percolation network interactions in a fuel cell gas diffusion layer. J Electrochem Soc 2020;167(8):084506.
[29] Mularczyk A, Lin Q, Niblett D, Vasile A, Blunt MJ, Niasar V, et al. Operando Liquid Pressure Determination in Polymer Electrolyte Fuel Cells. ACS Appl Mater Interfaces 2021;13(29):34003–11.
[30] Zenyuk IV, Parkinson DY, Connolly LG, Weber AZ. Gas-diffusion-layer structural properties under compression via X-ray tomography. J Power Sources 2016;328:364–76.
[31] Maier M, Dodwell J, Ziesche R, Tan C, Heenan T, Majasan J, et al. Mass transport in polymer electrolyte membrane water electrolyser liquid-gas diffusion layers: A combined neutron imaging and X-ray computed tomography study. J Power Sources 2020;455:227968.
[32] Bosomoiu M, Tsotridis G, Bednarek T. Study of effective transport properties of fresh and aged gas diffusion layers. J Power Sources 2015;285:568–79.
[33] Niu Z, Bao Z, Wu J, Wang Y, Jiao K. Two-phase flow in the mixed-wettability gas diffusion layer of proton exchange membrane fuel cells. Appl Energy 2018;232:443–50.
[34] Wang Y, Xu H, Zhang Z, Li H, Wang X. Lattice Boltzmann simulation of a gas diffusion layer with a gradient polytetrafluoroethylene distribution for a proton exchange membrane fuel cell. Appl Energy 2022;320:119248.
[35] Mularczyk A, Michalski A, Striednig M, Herrendörfer R, Schmidt TJ, Büchi FN, et al. Mass Transport Limitations of Water Evaporation in Polymer Electrolyte Fuel Cell Gas Diffusion Layers. Energies 2021;14(10):2967.
[36] Zenyuk IV, Lamibrac A, Eller J, Parkinson DY, Marone F, Büchi FN, et al. Investigating evaporation in gas diffusion layers for fuel cells with X-ray computed tomography. J Phys Chem C 2016;120(50):28701–11.
[37] Pfrang A, Veyret D, Sieker F, Tsotridis G. X-ray computed tomography of gas diffusion layers of PEM fuel cells: Calculation of thermal conductivity. Int J Hydrogen Energy 2010;35(8):3751–7.
[38] Ye L, Qiu D, Peng L, Lai X. Microstructures and electrical conductivity properties of compressed gas diffusion layers using X-ray tomography. Appl Energy 2022;326:119934.
[39] Khajeh-Hosseini-Dalasm N, Sasabe T, Tokumasu T, Pasaogullari U. Effects of polytetrafluoroethylene treatment and compression on gas diffusion layer microstructure using high-resolution X-ray computed tomography. J Power Sources 2014;266:213–21.
[40] Zenyuk IV, Parkinson DY, Hwang G, Weber AZ. Probing water distribution in compressed fuel-cell gas-diffusion layers using X-ray computed tomography. Electrochem Commun 2015;53:24–8.
[41] Pang Y, Hao L, Wang Y. Convolutional neural network analysis of radiography images for rapid water quantification in PEM fuel cell. Appl Energy 2022;321:119352.
[42] Cawte T, Bazylak A. A 3D convolutional neural network accurately predicts the permeability of gas diffusion layer materials directly from image data. Curr Opin Electrochem 2022;35:101101.
[43] Gao Y, Zhang X, Rama P, Liu Y, Chen R, Ostadi H, et al. Calculating the Anisotropic Permeability of Porous Media Using the Lattice Boltzmann Method and X-ray Computed Tomography. Transp Porous Media 2012;92(2):457–72.
[44] Ismail MS, Hughes KJ, Ingham DB, Ma L, Pourkashanian M. Effects of anisotropic permeability and electrical conductivity of gas diffusion layers on the performance of proton exchange membrane fuel cells. Appl Energy 2012;95:50–63.
[45] Goodarzi Ardakani V, Gambaruto AM, Silva G, Pereira R. A porosity model for medical image segmentation of vessels. Int J Numer Methods Biomed Eng 2022;38(4):e3580.
[46] Falcão AX, Udupa JK, Samarasekera S, Sharma S, Hirsch BE, Lotufo RA. User-Steered Image Segmentation Paradigms: Live Wire and Live Lane. Grap Models Image Process 1998;60(4):233–60.
[47] Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9(1):62–6.
[48] Adams R, Bischof L. Seeded region growing. IEEE Trans Pattern Anal Mach Intell 1994;16(6):641–7.
[49] Marr D, Hildreth E. Theory of edge detection. Proc R Soc Lond B 1980;207(1167):187–217.
[50] Jumb V, Sohani M, Shrivas A. Color image segmentation using K-means clustering and Otsu's adaptive thresholding. Int J Innov Technol Explor Eng (IJITEE) 2014;3(9):72–6.
[51] Oliver A, Munoz X, Batlle J, Pacheco L, Freixenet J. Improving Clustering Algorithms for Image Segmentation using Contour and Region Information. *2006 IEEE International Conference on Automation, Quality and Testing, Robotics.* 2. 2006: 315-20.
[52] Schlüter S, Sheppard A, Brown K, Wildenschild D. Image processing of multiphase images obtained via X-ray microtomography: A review. Water Resour Res 2014;50(4):3615–39.
[53] Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans Pattern Anal Mach Intell 1991;13(06):583–98.
[54] Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics 2017;33(15):2424–6.
[55] Shum AD, Liu CP, Lim WH, Parkinson DY, Zenyuk IV. Using Machine Learning Algorithms for Water Segmentation in Gas Diffusion Layers of Polymer Electrolyte Fuel Cells. Transp Porous Media 2022;144(3):715–37.
[56] Hesamian MH, Jia W, He X, Kennedy P. Deep learning techniques for medical image segmentation: achievements and challenges. J Digit Imaging 2019;32(4):582–96.
[57] Cheng J, Tian S, Yu L, Liu S, Wang C, Ren Y, et al. DDU-Net: A dual dense U-structure network for medical image segmentation. Appl Soft Comput 2022;126:109297.
[58] Wang S, Li C, Wang R, Liu Z, Wang M, Tan H, et al. Annotation-efficient deep learning for automatic medical image segmentation. Nat Commun 2021;12(1).
[59] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. Unet++: A nested u-net architecture for medical image segmentation. Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer 2018:3–11.
[60] Sharma N, Ray AmitK, Shukla KK, Sharma S, Pradhan S, Srivastva A, et al. Automated medical image segmentation techniques. J Med Phys 2010;35(1):3.
[61] Karimpouli S, Tahmasebi P. Segmentation of digital rock images using deep convolutional autoencoder networks. Comput Geosci 2019;126:142–50.
[62] Varfolomeev I, Yakimchuk I, Safonov I. An Application of Deep Neural Networks for Segmentation of Microtomographic Images of Rock Samples. Computers 2019;8(4):72.

[63] Tian X, Daigle H. Machine-learning-based object detection in images for reservoir characterization: A case study of fracture detection in shales. Lead Edge 2018;37 (6):435–42.
[64] Niu Y, Mostaghimi P, Shabaninejad M, Swietojanski P, Armstrong RT. Digital Rock Segmentation for Petrophysical Analysis With Reduced User Bias Using Convolutional Neural Networks. Water Resour Res 2020;56(2).
[65] Wang YD, Shabaninejad M, Armstrong RT, Mostaghimi P. Deep neural networks for improving physical accuracy of 2D and 3D multi-mineral segmentation of rock micro-CT images. Appl Soft Comput 2021;104:107185.
[66] Ar Rushood I, Alqahtani N, Wang YD, Shabaninejad M, Armstrong R, Mostaghimi P. Segmentation of X-Ray Images of Rocks Using Deep Learning. *SPE Annual Technical Conference and Exhibition.* Day 4 Thu, October 29, 2020. 2020.
[67] Breiman L. Random Forests. Mach Learn 2001;45(1):5–32.
[68] Abirami S, Chitra P. Chapter Fourteen – Energy-efficient edge based real-time healthcare support system. In: Raj P, Evangeline P, editors. Advances in Computers. Elsevier; 2020. p. 339–68.
[69] Kunapuli SS, Bhallamudi PC. Chapter 22 – A review of deep learning models for medical diagnosis. In: Kumar P, Kumar Y, Tawhid MA, editors. Machine Learning, Big Data, and IoT for Medical Informatics. Academic Press; 2021. p. 389–404.
[70] Bhuvaneshwari M, Kanaga EGM, Anitha J, Raimond K, George ST. Chapter 7 – A comprehensive review on deep learning techniques for a BCI-based communication system. In: N P, Kautish S, Peng S-L, editors. Demystifying Big Data, Machine Learning, and Deep Learning for Healthcare Analytics. Academic Press; 2021, p. 131-57.
[71] Navamani TM. Chapter 7 – Efficient Deep Learning Approaches for Health Informatics. In: Sangaiah AK, editor. Deep Learning and Parallel Computing Environment for Bioengineering Systems. Academic Press; 2019. p. 123–37.
[72] Sewak M, Sahay SK, Rathore H. An overview of deep learning architecture of deep neural networks and autoencoders. J Comput Theor Nanosci 2020;17(1):182–8.
[73] Pinaya WL, Vieira S, Garcia-Dias R, Mechelli A. Chapter 11-autoencoders. Academic Press; 2020:193-208.
[74] Jiang P, Maghrebi M, Crosky A, Saydam S. Chapter 23 - Unsupervised Deep Learning for Data-Driven Reliability and Risk Analysis of Engineered Systems. In: Samui P, Sekhar S, Balas VE, editors. Handbook of Neural Computation. Academic Press; 2017. p. 417–31.
[75] Altan G, Kutlu Y. Chapter three – Generalization performance of deep autoencoder kernels for identification of abnormalities on electrocardiograms. In: Das H, Pradhan C, Dey N, editors. Deep Learning for Data Analytics. Academic Press; 2020. p. 37–62.
[76] Badrinarayanan V, Handa A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. arXiv preprint arXiv: 150507293 2015.
[77] Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nat Methods 2019;16(1): 67–70.
[78] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer; 2015. p. 234–41.
[79] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980 2014.
[80] Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 2980–8.
[81] Dice LR. Measures of the Amount of Ecologic Association between Species. Ecology 1945;26(3):297–302.
[82] Zhao R, Qian B, Zhang X, Li Y, Wei R, Liu Y, et al. Rethinking Dice Loss for Medical Image Segmentation. IEEE International Conference on Data Mining (ICDM) 2020; 2020:851–60.
[83] Jaccard P. The distribution of the flora in the alpine zone. New Phytol 1912;11(2): 37–50.