



HAL
open science

Artificial neural network approach for multiphase segmentation of battery electrode nano-CT images

Zeliang Su, Etienne Decencière, Tuan-Tu Nguyen, Kaoutar El-Amiry, Vincent de Andrade, Alejandro A. Franco, Arnaud Demortière

► **To cite this version:**

Zeliang Su, Etienne Decencière, Tuan-Tu Nguyen, Kaoutar El-Amiry, Vincent de Andrade, et al.. Artificial neural network approach for multiphase segmentation of battery electrode nano-CT images. npj Computational Materials, 2022, 8 (1), pp.30. 10.1038/s41524-022-00709-7 . hal-03595111

HAL Id: hal-03595111

<https://u-picardie.hal.science/hal-03595111v1>

Submitted on 8 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.




Distributed under a Creative Commons Attribution 4.0 International License

ARTICLE OPEN



Artificial neural network approach for multiphase segmentation of battery electrode nano-CT images

Zeliang Su^{1,2}, Etienne Decencière³, Tuan-Tu Nguyen^{1,2,4}, Kaoutar El-Amiry^{1,2}, Vincent De Andrade⁵, Alejandro A. Franco^{1,2,6,7} and Arnaud Demortière^{1,2,7} 

The segmentation of tomographic images of the battery electrode is a crucial processing step, which will have an additional impact on the results of material characterization and electrochemical simulation. However, manually labeling X-ray CT images (XCT) is time-consuming, and these XCT images are generally difficult to segment with histographical methods. We propose a deep learning approach with an asymmetrical depth encode-decoder convolutional neural network (CNN) for real-world battery material datasets. This network achieves high accuracy while requiring small amounts of labeled data and predicts a volume of billions voxel within few minutes. While applying supervised machine learning for segmenting real-world data, the ground truth is often absent. The results of segmentation are usually qualitatively justified by visual judgement. We try to unravel this fuzzy definition of segmentation quality by identifying the uncertainty due to the human bias diluted in the training data. Further CNN trainings using synthetic data show quantitative impact of such uncertainty on the determination of material's properties. Nano-XCT datasets of various battery materials have been successfully segmented by training this neural network from scratch. We will also show that applying the transfer learning, which consists of reusing a well-trained network, can improve the accuracy of a similar dataset.

npj Computational Materials (2022)8:30; <https://doi.org/10.1038/s41524-022-00709-7>

INTRODUCTION

The X-ray computed tomography (XCT) is a robust characterization tool in which the battery field has shown tremendous interest during the last decade^{1–3}. It provides valuable 3D morphological information on the battery materials and electrode architectures. Its broad range of observation allows us to investigate nanometric particles, up to tens of nm resolution⁴, and the bulk electrode with a large field of view from tens of μm to 1 mm. For instance, nano-XCT techniques have been recently used to study, on the material level, phases spatial distribution⁵, steric changes⁶, 3D oxidation state evolution⁷. The micro-CTs are often employed on the cell level⁸ and operando studies⁹. The use of synchrotron sources, the emergence of fast imaging detectors, and the advanced in situ/operando characterization spawn soaring data quantities and lead to unprecedented challenges in image processing and data management. The raw dataset often contains few gigabytes of projection and is then reconstructed into a stack of tomograms that typically includes a billion voxels for the analysis. The 3D analysis and electrochemical simulation are usually preceded by a step of semantic segmentation, which consists of digitally partitioning each voxel of the raw stack of tomograms (Fig. 1a left part) into different phases (Fig. 1a right part).

The segmented volume of XCT can be used as an input of electrochemical models^{10–14} to simulate electrochemical performance, which helps to understand transport phenomena in the electrode and to design a better electrode architecture. Pietsch et al.¹⁵ has firstly discussed the impact of segmentation on the determination of morphological and transport properties for commercial anode materials. They studied each parameter in the XCT image post-processing and the thresholding in segmentation.

They observed that the variation of porosity and tortuosity due to the difference in segmentation could become considerable. Therefore, the data processing and segmentation step should be done carefully in the battery field. For nano-XCT data (e.g., Fig. 1c-top histogram), where high signal-to-noise-ratio is challenging to obtain and a wide variety of artifacts is present, the straightforward grayscale thresholding approach is not accurate enough, especially for complex composite materials because of the histogram overlapping.

Up to date, in the tomography field, people investigate intensively machine learning approaches to accelerate the image segmentation^{16,17}, such as coupling fixed feature extractors and a machine learning classifier^{18,19}. Over the past decade, thanks to advances in the computing power, large-scale convolutional neural networks (CNN, see Methods) have become easier to train. The CNN has thrived in automated segmentation and other similar computer vision problems in various fields such as satellite or astrological images^{20,21}, facial recognition²², camera-assisted vehicle autopilot^{23,24}, and medical imaging^{24–29}. XCT images of battery materials contrast with the examples above as they typically contain crystals, agglomerates, polymers and porosities with complex morphologies and architectures to maximize electrochemical reaction rate. Liu et al.³⁰ investigated the degradation of a Li-ion NMC material with a Mask R-CNN that provided a quantitative instances-level particle identification despite the particle cracking. Labonte et al.³¹ studied the binarization of a graphite anode micro-XCT dataset with a more sophisticated 3D neural network capable of providing a segmentation uncertainty map with a stochastic neural network.

¹Laboratoire de Réactivité et Chimie des Solides (LRCS), CNRS UMR 7314, Université de Picardie Jules Verne, Hub de l'Energie, Rue Baudelocque, 80039 Amiens Cedex, France.

²Réseau sur le Stockage Electrochimique de l'Energie (RS2E), CNRS FR 3459, Hub de l'Energie, Rue Baudelocque, 80039 Amiens Cedex, France. ³MINES ParisTech – PSL Research University, CMM, Center for Mathematical Morphology, 35 rue Saint Honoré, Fontainebleau, France. ⁴Renault Technocentre, 78084 Guyancourt, France. ⁵Advanced Photon Source, Argonne National Laboratory, Lemont, USA. ⁶Institut Universitaire de France, 101 Boulevard Saint Michel, 75005 Paris, France. ⁷ALISTORE-European Research Institute, CNRS FR 3104, Hub de l'Energie, Rue Baudelocque, 80039 Amiens Cedex, France. ✉email: arnaud.demortiere@cnrs.fr

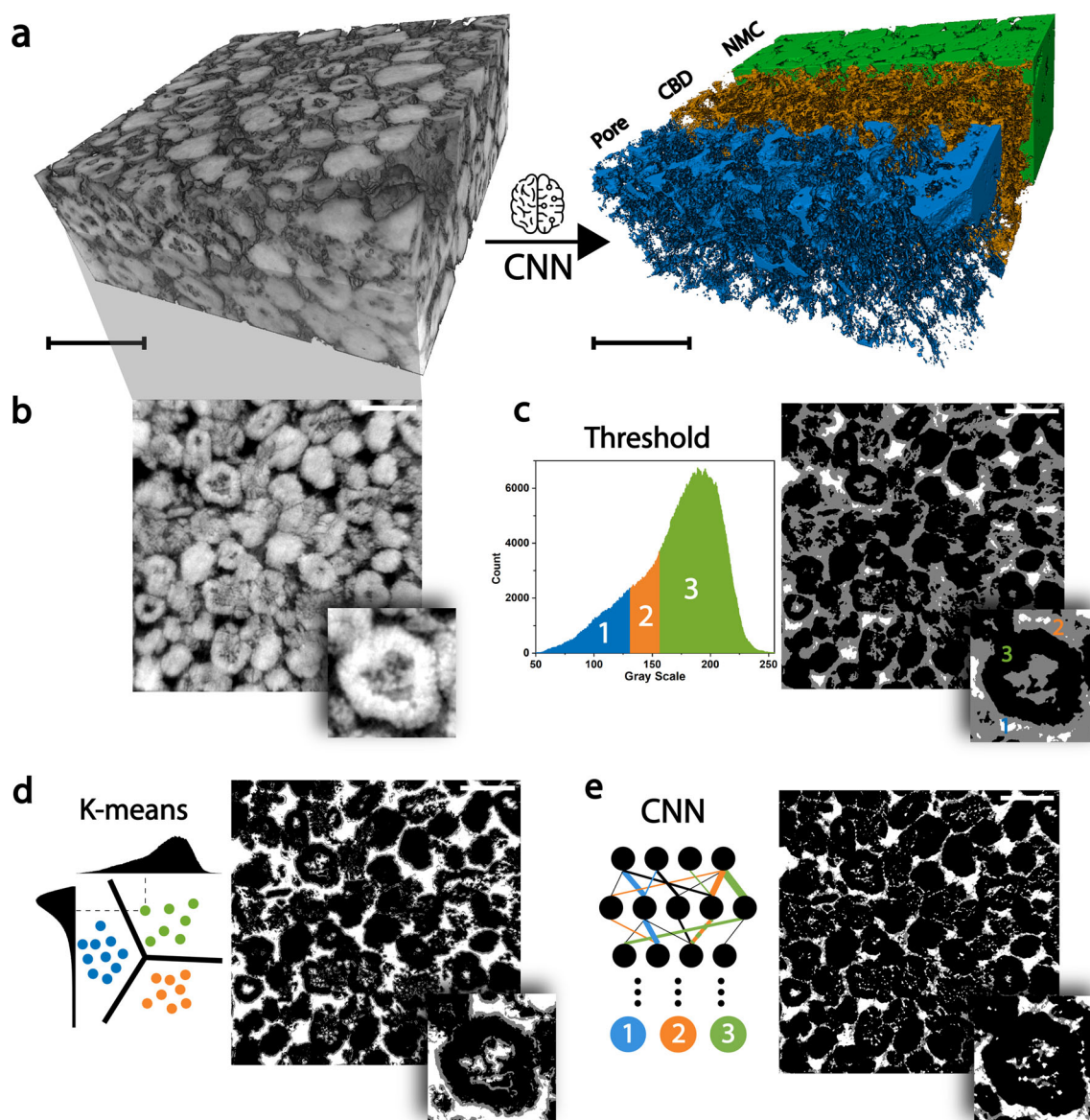


Fig. 1 Comparison of multiphase segmentation by different methods. **a** Volume rendering illustrating the objective of turning raw 3D volume of NMC₁ into a segmented one. NMC/CBD/pore are rendered respectively in green/orange/blue colors. **b** Selected cross-section from the previous raw volume with a zoom on an NMC particle and its surroundings. **c** The histogram of **(b)** on the left, and on the right histogram thresholding result based on the theoretical volume fractions, where the black, gray, and white colors represent respectively NMC, carbon-binder-domain, and pores. **d** An illustration of segmentation applying automatic k-means on a map constructed by the gray level and Sobel filtered values of the raw tomogram. On the right, the result of this method. **e** On the left, a schematic of artificial neural network, and on the right the output of CNN. All scale bars are 8 μm .

For the first $\text{LiNi}_{0.5}\text{Mn}_{0.2}\text{Co}_{0.3}\text{O}_2$ (hereafter namely dataset NMC₁) cathode material, the goal is to distinguish the three phases in the electrode presented in the Fig. 1b: (a) the white NMC active material where the lithium is stored, (b) the carbon binder domain (CBD) of a mixture of polymer and carbon black surrounding the NMC, which maintains the mechanical cohesion of the material, and (c) the porosity impregnated by the liquid electrolyte where the ions circulate during the electrochemical reaction. The use of the thresholding approach (Fig. 1c) or the automatic K-means method (Fig. 1d) applying on a 2D histogram leads to an overestimation of the CBD phase and a coarse separation of the interfaces. These as-segmented volumes with the NMC particles are firmly surrounded by the CBD. For instance, the use of these volumes might induce a poor exchange on the NMC surface and result in a biased electrochemical simulation.

Our current contribution (Fig. 1e) expands the portfolio for accurate multiphase segmentation of battery CT images with a portable neural network architecture. We discuss the impact of hidden segmentation bias which has often been overlooked when applying an automatic algorithm. This article is organized as follows. First, we will present the workflow of training a network from scratch improve the performance of a CNN by tuning the hyperparameters (HPs). Thence, we will identify the cognitive bias diluted in the labeled data and quantify their potential impacts on the material properties characterization. Finally, our approach will be cross-validated with other battery nano-CT data. We will also show that the accuracy can be improved by reusing the kernels of a pre-trained network, namely transfer learning.

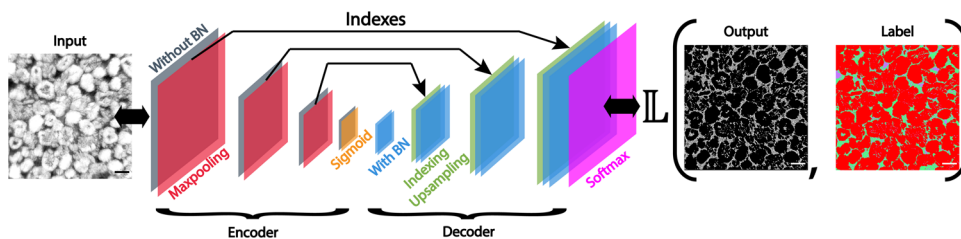


Fig. 2 A schematic representation of LRCS-Net with the input image and the composition of loss function. The image scale bars are of 5 μm .

RESULTS AND DISCUSSION

CNN architecture and hyperparameters tuning

LRCS-Net (Fig. 2) used throughout this work has been optimized to segment efficiently nano-CT images of the battery electrode and is derived from Seg-Net²⁴ and Xlearn³² artificial neural networks (explanation of neural network refers to Methods and the structural optimization is shown in Supplementary Note 2). On the encoder side, the LRCS-Net contains in total five layers of convolution with three indexing max-pooling (MP) and ends with a sigmoid function instead of a leaky function applied in the rest of the network. On the decoder side, eight layers of convolution with three up-sampling receiving indexes at the first convolution layer of each block. The model has less trainable parameters than the frequently used network U-net²⁶ in semantic segmentation in other fields. The throughput of CT images per second can reach twice as much as in the U-net and the prediction speed for a volume of billion voxels is a quarter faster. The Supplementary Note 3 explains intuitively the functioning of this network by visualizing the flow of images within it.

To explore the best performance of the CNN, the HPs should be optimized for each dataset. In contrast to the trainable weights, the HPs are tunable by experimenters. They control the size of the network and determines the convergence of the training process. Comparing with the enormous datasets in the domain of object detection, real-world tomography datasets of battery materials contain fewer classes.

The CNN is prone to overfitting certain class than other if the HPs are badly initiated. As such, the network can be easily trapped by a poor local minimum that predicts only the majority class. We call it a major class pitfall as the accuracy is stuck at the value of volume fraction of the major class. In the case of NMC, this is reflected by a low variation plateau of accuracy around 80%, which corresponds to the volume fraction of NMC₁ dataset. This is due to the unbalanced quantity of different phases in the training data. And at the beginning of training, inferring the majority phase costs less and minimizes the loss faster. For a deeper understanding of the HPs' influence on the CNN's performance and finding the reasonable interval for each HP for the current CNN, an investigation is conducted below (using the platform SegmentPy, see Supplementary Fig. 5 and Methods).

Figure 3b-e plot the average of validation accuracy in solid lines and the standard deviation in the colored area during the training. The learning rate is a parameter that controls the momentum of the trainable variables during the backpropagation. A higher learning rate leads to an instability of the local minimum, while a lower one traps the network in a poor local minimum. This HP is delicate to tune. For example, a constant learning rate drives to poor minima and accuracy. In contrast, a periodically decreasing one with a decay ratio of 0.3 has an optimal convergence (Fig. 3b, decay applied at the end of each epoch. An epoch defines the entire training dataset). Nevertheless, reducing the ratio to 0.1 starts to reduce variation and limit performance. The batch size is another HP handling the parallelization while updating the weights of the training process. We see in Fig. 3c that by controlling the total amount of training images, a small batch size with less parallelization can lead to better convergence. Two other

important HPs are those defining the number of convolution channels and the kernel size. Figure 3d and e show that increasing the CNN size does not necessarily drive to better performance and can result in overfitting.

One should note that the HPs could have interactions^{33,34} among them. To illustrate this, Fig. 3f plots the accuracies sorted in descending order with different combinations of HPs. We see that the value of the initial learning rate should be carefully chosen for obtaining better accuracies. Here, small batches (in purple) are prioritized, which is in accordance with Fig. 3c. Other HPs, on the other hand, do not have a clear trend on the optimization. In contrast to the trainable parameters in the CNN that receive feedbacks from the loss by gradients (Methods), seeking the best combination of HPs is indeed a black box guessing problem that can only be found by trials (ad hoc approach). Random search³⁵ and Bayesian search^{34,36} based on the gaussian process are methods that could help to refine the HPs.

Reveal the influence of biases diluted in ground truth (GT)

Segmenting large battery material volumes always involves automatic (e.g., Otsu, watershed) or semi-automatic (e.g., the current supervised-learning) methods. In most of the papers, the result of segmentation is used directly for quantitative measurements, although it is justified qualitatively or sometimes the justification is even missing. Unless images of higher resolution of the exact same labeled zone by coupling with the FIB-SEM³⁷ is available, deploying the CNN for segmenting the XCT images must deal with the uncertainty. Apart from visually judging and inspecting the metrics such as the accuracy, there is no other efficient way of qualifying the segmentation. In applicative cases as shown with our previous examples, the inconsistency among the training, validation, and testing dataset due to the uncertainty causes impasse such that although the prediction is visually satisfying, the accuracy is stuck at about 90%.

In this section, we will use training CNN as paradigm to discuss the uncertainty and the origin of this roof of performance and try to quantify its impacts on the post material properties determination while dealing with real-world data. For this, we will discuss alongside with the results of two experiments: a survey of the degree of discrepancies between the experimenters and training several neural networks on slightly different labels to evaluate the resulting material properties.

First and foremost, applying a supervised-learning method will dilute human bias in the training process. Using CNN in semantic segmentation problem is to train a neural network to approach an ideal function \mathcal{F}_{ideal} that transforms the input tomographic volume \mathbf{V}_{raw} into a ground truth segmented volume: $\mathbf{CNN}_{ideal}(\mathbf{V}_{raw}) \sim \mathbf{GT}_{ideal} = \mathcal{F}_{ideal}(\mathbf{V}_{raw})$ with \mathbf{W} the trainable parameters in the network. Here, the CNN can also be generalized to other parametrized automatic methods. One should bury in mind that one chooses only a subset of the volume to manually generate labels for the training. It will always accompany with some human bias $\mathbf{GT}_{ideal} + \epsilon_{cog|exp,raw} = \mathbf{GT}_{manual}$, where $\epsilon_{cog|exp,raw}$ is the cognitive bias. From our experiences, this bias $\epsilon_{cog|exp,raw}$ with the subscripts mainly depends on experimenter and the

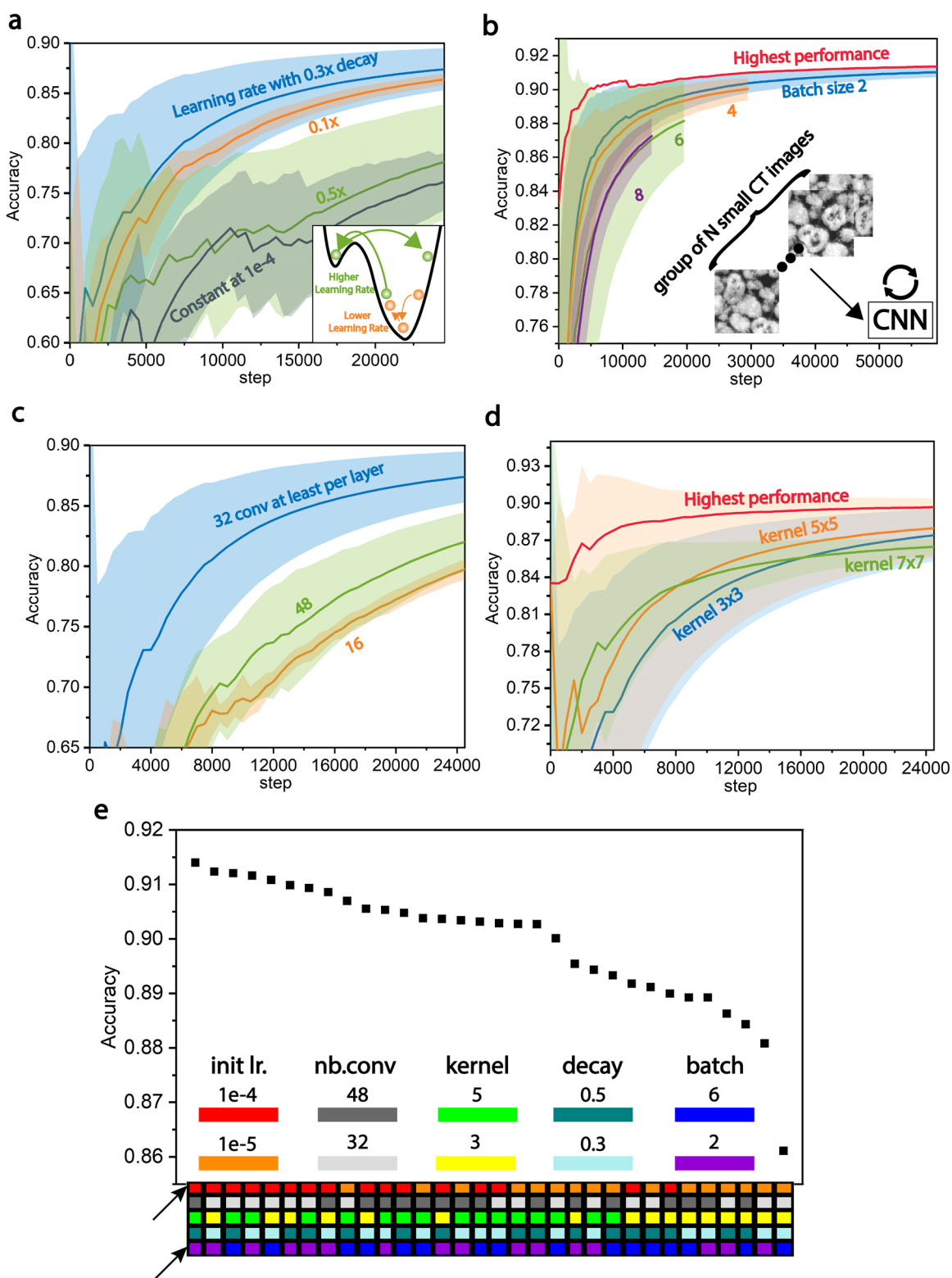


Fig. 3 The statistical score comparison of important hyper-parameters. (HPs) performed on the validation dataset (the unseen data during training). The solid lines and corresponding colored areas throughout this figure represent respectively the mean and standard deviation of validation scores of three trainings. **a** The constant learning rate turns out to overfit, which is often ascribed as inflexibility or bad generalization of the CNN onto the unseen data. An optimal decay ratio of learning rate is found to be 0.3 for this set of hyperparameters. **b** Controlling the total quantities of training data, small batches are preferred for better convergence. **c** A middle value of 32 for the number of convolutional kernels leads to the best accuracy. **d** 5×5 kernels leads to optimum performance. **e** Sorted validation scores in descending order of a broader hyperparameter search. The X axis is replaced by a grid of colors representing used hyperparameters for achieving the corresponding score. The initial learning rate shows particular importance for obtaining higher accuracies. The performance depicted in blue throughout (b–e) are obtained with the Dice loss function, leaky activations, and BN implementation in the decoder and other training details refer to Supplementary Table 1–2.

quality of the raw data. And the validation (Fig. 3)/test (Fig. 5d) datasets is to compare the $CNN_{train|W_{trained}}(V_{raw}^{test})$ with $GT_{manual}^{valid/test}$, where the subsets of train, valid, and test do not intersect one another. We see that the $\epsilon_{cog|exp,raw}$ intervenes three times in the process, once in training the CNN with GT_{manual}^{train} and the others in the $GT_{manual}^{valid/test}$. The origin of the performance roof is because such bias changes in different subsets. Notably, other methods than neural network cannot get rid of such bias as the experimenter needs to at some point verify the output of the method and to make further improvements.

Thereby, the first survey experiment (see Supplementary Note 4) aims at determining and showing the degree of $\epsilon_{cog|exp,raw}$. The results showed that there could be at least $\sim 10\%$ difference in the segmentation collected from different people. We see that by comparing their results to the commonly accepted GT, the main differences lie mostly on the interface between phases for the example (given in Supplementary Fig. 6). Moreover, the magnitude of such difference (Supplementary Table 3) is in accordance with the last few percent of CNN accuracy in Fig. 3. As explained above, this is because the ϵ_{cog} is not fixed and is unavoidably diluted in the whole labelling process. In other words, the ceiling of performance can be interpreted as an indicator of the experimenter's self-consistency of labeling data and the degree of uncertainty in the segmentation. To reduce the segmentation ambiguity, one can couple the XCT with other techniques such like chemical DRX-CT³⁸ and ptychography-XCT³⁹. However, the resolution of DRX-CT or the acquisition time of ptychography-XCT should be improved. The second experiment is to give an estimation of the influence of the cognitive bias on the segmentation with a larger statistic. With the previous experiment, we understood that the segmentation ambiguity locates mainly on the interface. Additionally, Supplementary Fig. 8 shows raw tomograms and a line profile perpendicular to an NMC-CBD interface. One can see that the sharp border in Supplementary Fig. 8 corresponds to a slope of 10 voxels in width. In absence of larger samples of expert GTs for NMC₁ dataset, we established an algorithm to simulate perturbed GTs not exceeding the interval of 10 voxels on the interfaces to train different CNNs (detailed in Methods). The algorithm consists of locating all the interfaces and choosing a part of them to push/pull by random units. New predictions from these LRCS-Nets are evaluated by volume fractions, surface area, and another metric the intersection of Union (IoU). The use of the latest is because the overall accuracy does not reflect the balance between classes in multiphase segmentation. The network has no guarantee of converging toward a minimum of good quality. For instance, it could tilt in a particular class but still achieve decent accuracy (e.g., having all possible NMC particles correctly segmented, but mostly wrong for the others in a majority class trap). IoU for each class is a more common metric in the semantic segmentation to assess whether the network is trained in an imbalanced manner. It is calculated by dividing the common area of the predicted segmentation and the ground truth by their union.

Due to the computational cost, we first validated this algorithm in 2D with a thousand repetitions. The 2D histogram of interface voxels in a thousand simulated GTs roughly underlies the Gaussian shape with a full width at half maximum of 10 voxels. It is shown in Fig. 6a as a green mask on the raw tomogram. The mask has a darker green color when the count is high and transparent when it is zero. Figure 6b depicts the 3D histogram of purple interfaces for a hundred simulated perturbed 3D GTs.

Fifteen CNNs are then trained with labeled images generated from the same training dataset and evaluated by the common ground truth in the test dataset. HPs use the best combination of HPs obtained with the previous NMC₁ datasets. Figure 6c represents the IoU distributions of the 3D predictions of these networks and the variance of the overall accuracy. The NMC phase

has the most stable IoU dispersion of $92.7 \pm 0.2\%$, which contrasts with the CBD $37.6 \pm 1.4\%$ and the pores $65.9 \pm 1.3\%$. Figure 6d shows the ratio of the surface area and volume fraction for the three phases. We see that the higher surface area to volume ratio results in smaller IoUs, confirming our previous finding of the uncertain area. CBD is the most difficult to segment among these three phases in this dataset and tends to have inconsistencies between experimenters. Potential ways to improve IoUs of thin objects could be to use higher resolution and smaller FoV with interlaced scans or other advanced XCT techniques^{4,40} or reconstruction algorithms⁴¹.

The 15 CNNs trained from the perturbed data are used to predict 15 volumes. The volume fractions and interfaces for each of these volumes are plotted in Fig. 6d & e. The 3D interfaces vary in intervals of $2.5 \pm 0.3\%$, $2.1 \pm 0.3\%$, $2.6 \pm 0.2\%$ respectively for NMC-CBD, NMC-pore, and CBD-pore (Fig. 6e). We see that the accuracy deviations (Fig. 6c) evaluated on the test data results in $<1\%$ of the variance for the 3D predictions (Fig. 6e). Note that the interface is deliberately expressed as a percentage of voxels instead of nm^{-1} to avoid ambiguity as there are various extrapolations of tomographic voxels to a surface, such as taking the diagonal triangle or an arbitrary constant value, which will result in different values. Threshold, as described in Fig. 1c on NMC₁, resulted in 0.47% surface area voxels, which is 4–5 times less surface area than by the CNN segmentation.

Validation of LRCS-Net via various datasets

In the previous sections, the battery data segmentation routine and the influence of the human bias diluted in the datasets have been shown. In this section, we try to generalize our approach on different tomographic datasets of battery materials. A similar dataset from NMC with the same composition but higher loading was used for training a second network using transfer learning. Two other datasets of battery materials with different morphologies will also be shown.

For the second dataset of NMC (denoted NMC₂ hereafter), instead of initializing the kernel randomly in the beginning of the training, we recover all the well-trained kernels in the best-trained model (denoted LRCS-Net₁) with the NMC₁ dataset (Fig. 5a). This is called transfer learning (Table 1). The Kernels of the LRCS-Net were saved in four different advancements during the previous training. Different starting learning-rates were applied (in Fig. 4a, a descending order of starting learning-rate from blue: $1e-4$, orange: $1e-4 \times 0.3^{(N-1)}$, to green: $1e-4 \times 0.3^N$. N is epoch number, at the end of which saved the state of LRCS-Net₁^N). A control experiment is carried out with a random initial state of the network and with NMC₂ dataset.

Unlike training from scratch, resumed trainings begin directly above 80% accuracies since the kernels have already been trained. These starting points of transfer learning, from the different depth of resuming point of LRCS-Net₁^N, increase and then stabilize around 83%. A final gain of more than 2% on average was obtained, which is in accordance with the conclusions of Yosinski et al.⁴² that the transfer learning of all kernels leads to a better generalization of the network. The green curves show that lower starting learning rates give higher accuracies. However, this performance gain stabilizes after resuming from steps after 30k of LRCS-Net₁^N, indicating that the benefits of generalization from a trained model is limited. Nevertheless, this finding is still beneficial

Table 1. Summary of the training data for the transfer learning.

	LRCS-Net ₁	Transfer learning	Control
Trained with dataset	NMC ₁	NMC ₂	NMC ₂
Initial state of CNN	Random	LRCS-Net ₁ ^N	Random

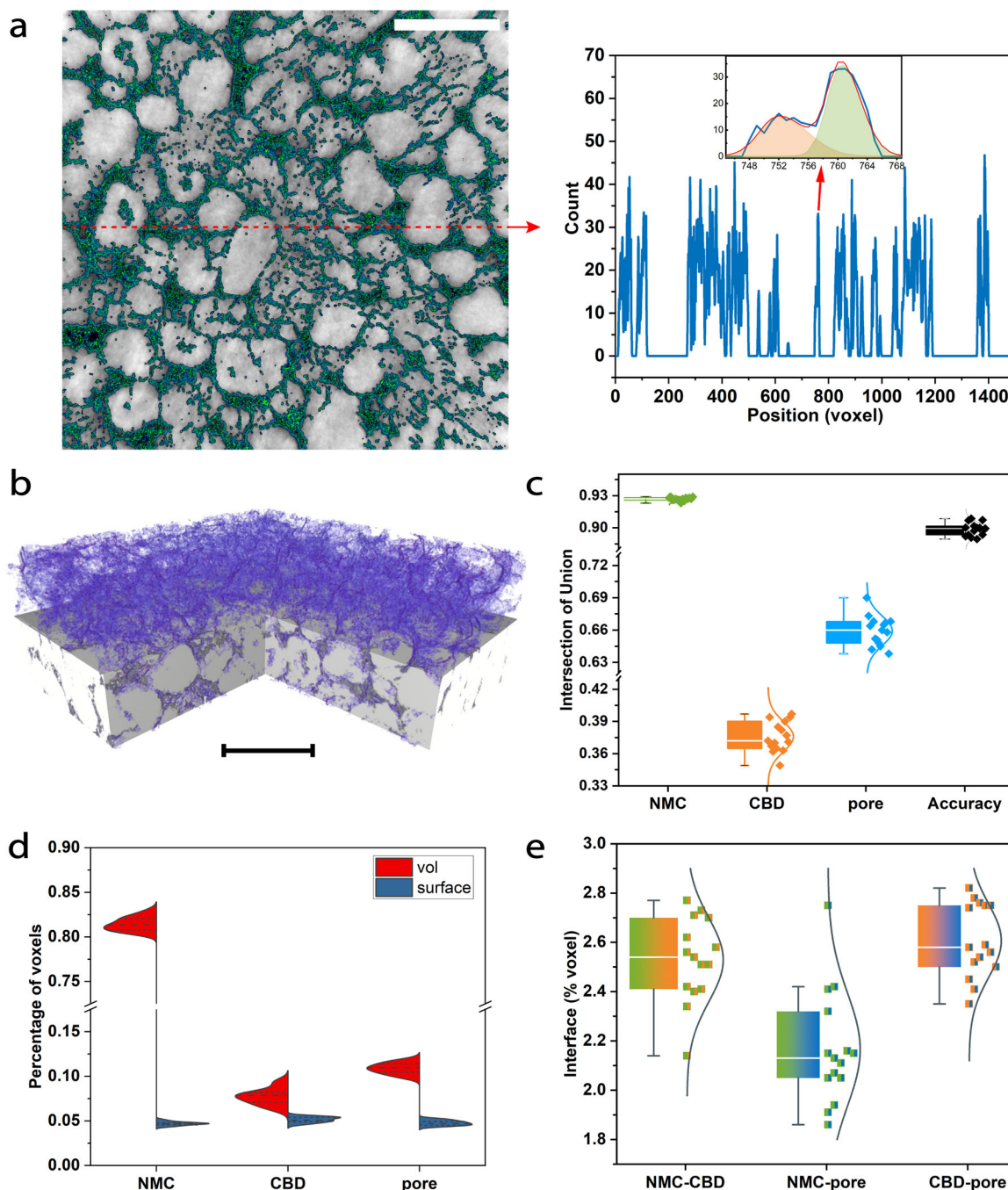


Fig. 4 Evaluation of the uncertainty impact. **a** A raw tomogram masked by the 2D histogram of interphases of a thousand simulated labeled images in the training dataset. The middle line profile of the histogram is plotted in blue and zoom onto two frontiers with a width at middle-height of about ten voxels. The orange and green undercurve area are fitted by gaussian distributions to guide the eye. **b** The histogram of interfaces extended to 3D for 100 synthetic 3D volumes **(c)** the evaluation of IoU and accuracy of 15 CNNs trained with simulated training data **(d)** distribution of the surfaces and volume fractions of the predicted 3D volume from the said CNNs. **(e)** The dispersion plot of three types of interfaces segmented by these 15 CNNs. The scale bar represents 10 μm .

in accelerating the segmentation of tomographic data as the convergence of the learning curves in transfer learning are steeper than the ones of training from scratch. We have successfully demonstrated that LRCS-Net can achieve reasonably high accuracy by receiving only a single segmentation example image and improve accuracy and convergence speed by transferring already trained kernels.

In addition to these two NMC datasets, the IoUs for a pristine binder-free carbon nanotube cathode material for Li-O₂ battery (Fig. 5b) and another dataset of the same cathode material in the

recharged state (Fig. 5c) are shown. These materials made of low Z elements have weak X-ray attenuation coefficient. Therefore, these two additional datasets are obtained using a different imaging technique, i.e., the Zernike Phase Contrast⁴³. The morphology of these materials and the complications of segmentation differed from the previous Li-ion cathode.

Figure 5d summarizes the incremental IoUs obtained by LRCS-Net comparing them to the threshold for all these X-ray nano-CT datasets. We find that our CNN exceeds the 4% threshold in terms of total accuracy. And the IoUs for all classes are above the

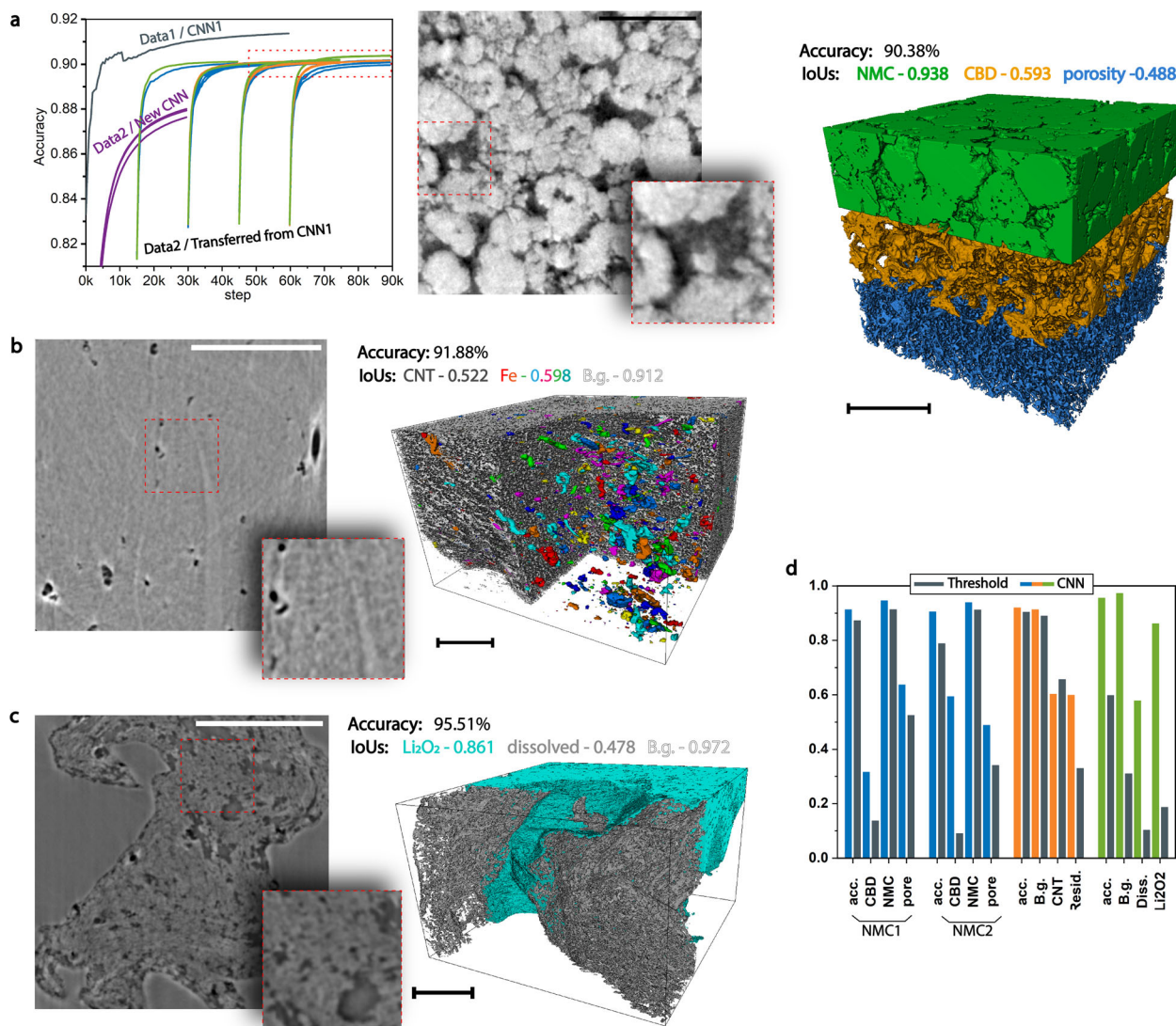


Fig. 5 Performances for other battery datasets. **a** Accuracy comparison between training from scratch and the ones that starts by kernel transfer from a well-trained LRCS-Net. Data1 and 2 are similar NMC materials from two electrodes of different loading (mg/cm²). The purple curve of trainings from scratch with NMC₂ have lower accuracies than the transfer learnings (blue/orange/green ones of descendant initial learning-rates) from the NMC₁ trained model (gray curve). A selected slice of the raw tomogram and an inset zoom is shown. On the top right, the 3D segmentation volume of NMC₂ for three phases is depicted. **b** A cross-section of raw volume and the 3D rendering of the segmentation of this carbon nanotube binder-free pristine material for Li-O₂ battery. The gray color represents the carbon. The colored inclusions in the electrode separated by the 3D watershed algorithm after the segmentation. **c** The same material was treated by acid to remove the iron particles then went through a full discharge-charge round-trip. On the left-hand side, a cross-section of this material. And on the right-hand side, the cyan color domain corresponds to the non-dissolved Li₂O₂, and the gray colors are the dissolved ones. **d** Comparison between the threshold and LRCS-Net on their accuracies and Intersection over Union (IoU) for two datasets of NMC, the binder-free cathode of Li-O₂ battery, and another recharged state cathode of Li-O₂ battery. The scale bars in (a–c) are 10 μm.

threshold, indicating that the improvement in segmentation is well balanced for each class. The IoU of the CBD phase is generally the lowest of these three classes because it includes the smallest objects.

In Fig. 5b, the pristine cathode contains tightly entangled carbon nanotubes and residual iron particles and other inclusions from the fabrication of nanotube. We have segmented three phases: nanotubes, in which gray-level is closed to the background; impurities, which present a strong contrast to X-ray and inverted by the phase contrast technique resulting in the darkest color; and the void, brighter than the other classes in the background. The halo artifact surrounding the inclusions is arbitrarily included in the background. In the 3D volume of Fig. 5c, the recharged electrode is segmented differently:

undissolved Li₂O₂ (blue), dissolved domain (dark gray), and background (transparent). The difficulties in segmenting these datasets are as follows. The carbon nanotube in the pristine dataset is extremely thin and almost anchored in the background. The Li₂O₂ and the void in the recharged dataset have the same gray level but have different textures.

Supplementary Fig. 7 shows the synergies between the HPs in LRCS-Net with descending order of scores. Like the NMC₁ dataset, the trend of obtaining better results with small batches and 1e−4 as initial learning rate is again obtained. Compared to the threshold, LRCS-Net improved the IoUs of these datasets. For the pristine dataset, some background noise is included by the threshold method. The IoUs for iron particles and background are improved by LRCS-Net, while the improvement in CNT

segmentation is modest (<0.02). For the recharged dataset, the threshold failed with the threshold method due to the similar gray level of Li_2O_2 and the background. In contrast, the LRCS-Net can distinguish these phases and has higher IoUs.

The NMCs for the high-capacity applications studied in this work have relatively dense NMC particles. Some cracks can be seen due to the calendaring process. The morphology of the particles is different from the lab-used spherical NMC particles^{1,30,44}. On the other hand, the pristine O_2 -cathode has a fourfold higher porosity (83%) than a traditional SP carbon electrode characterized in our previous study⁴³, which can facilitate the diffusion of oxygen and leave more room for lithium peroxide deposition. The tortuosity of this CNT material calculated by¹⁰ averages 1.15 in three directions, which is low and closed to 1 that favors the oxygen diffusion within the structure. The non-total dissolution of the peroxide, as shown by Fig. 5c, indicates that the electrochemistry should be further improved, for example, by using different electrolytes. Throughout these four datasets, the current CNN can achieve the presented performance with a small training dataset of a single raw/GT pair image to achieve accurate segmentation.

DISCUSSION

Nano-XCT data of battery materials is challenging to segment. The overlapping gray-levels and tomographic artifacts are factors that hamper accurate segmentation with traditional methods. We addressed this problem with a small CNN (LRCS-Net) and presented the workflow of training a CNN from scratch within the framework of the open source SegmentPy software. We demonstrated that portable and computationally inexpensive models (LRCS-Net) can also easily achieve decent accuracy and make fast prediction with small training dataset.

This work has been focusing on deploying CNNs for applicative segmentation of multiphase battery materials. At the current state, the HPs tuning is still an unavoidable task in the segmentation routine. Hence, we gave practical examples of HPs tuning and showed their influences on the convergence. Among the studied HPs, we found that the learning rate and batch size are the most sensitive and therefore need to be carefully adjusted. These findings have been verified on two XCT datasets of Li-ion battery cathode and reproducible in two other Li- O_2 battery datasets using phase contrast technique. Furthermore, we have shown the incremental effect of applying transfer learning for the training in a similar dataset.

With a survey approach and a data simulation approach, we have answered several fundamental questions. We have first identified the nature and the region of uncertainty for a NMC dataset by interrogating a group of scientists to segment the same image. The outcome shows it is difficult for people to reach a unanimous consent on voxels near the interface. These areas are also those with ambiguity in the prediction of the network. We have thus further quantified the impact of such uncertainty by comparing the outputs of CNNs trained with synthetic data. We have given the variances of the surface area and the volume fraction of the NMC₁ dataset.

In summary, the current work has not only demonstrated the capability of the CNN but also addressed to a challenging topic of uncertainty in the segmentation for battery CT material, which has been considered as an unquantifiable and often neglected in the field. Finally, we would like to add that, in practice, fine segmentation adjustments can be made afterward, and more tomography slices can be used for composing each dataset.

In perspectives, a profound comparison of LRCS-net with the family of U-Net and its derived forms will be carried out^{45–47}. Other pseudo-3D CNN model uses adjacent slices as 3D input, but 2D convolution kernels as reported in^{48,49} or 3D CNN model, which uses volume as inputs and 3D convolutions by Labonte

et al.³¹ and an associated uncertainty metric⁵⁰ can be further investigated. There are also some emerging automatic techniques^{51,52} searching optimal CNN architecture that could be potentially deployed in our current cases. Future direction might be to train a versatile network with a larger dataset for a specific collection of material. To this end, the reported transfer learning will be a reliable supporting technique. Emerging weak supervised few-shot segmentation methods^{53,54} with a different training fashion is a potential direction in segmenting the materials of similar characteristics with few labeling interventions. Last but not least, more realistic tomographic artifacts such as motion artifacts or ring artifacts can be artificially added to the augmentation to reinforce the network capacity.

METHODS

CNN approach and the fundamentals

A CNN is a branch of deep learning that mainly contains units of convolution. It is a mathematical model that artificially mimics the function of the neural network. For a segmentation task, it is trained to encode the features of the input image and give the associated segmentation on the output side without explicit feature extractors and instructions called by the experimenter.

The basic units of a CNN include (1) a convolutive kernel with trainable variables (or called hereafter weight) that perform feature filtering on the receiving data (Fig. 6a). (2) Max-pooling (MP)/Up-sampling (UP) which modify the dimensions so that the following operations can act on a different scale of data (Fig. 6b). These operators in this work appear in pairs and communicate with each other with indexes. The MPs on the first half of CNN (encoder) transmit the position information of max values to the UPs of the same level in the second half of CNN (decoder). (3) The activation function (e.g., different examples applied in this work in Fig. 6c) is the switch of a neuron that is triggered upon receiving a value greater than the threshold. This function is added after the convolutive kernels to form a complete layer.

A typical representation of CNN (e.g., the optimized LRCS-Net) is shown in Fig. 2a, where the sheets illustrate the layers of these basic units. Other operations are added for specific purposes. For instance, batch normalization (BN) is usually added in the layers to reduce the effect of scale variance of different input channels of the previous layer. BN and its derivative techniques often lead to a faster convergence^{55,56} (Fig. 6d). The soft-max layer converts the output of CNN into a kind of phase probability map. Detailed definitions of all these basic operations can be found in Supplementary Note 1. Stacking these layers sequentially and connecting the indexes bridges, as shown in Fig. 2a, forms a CNN.

The CNN is uniformly and randomly parameterized at the initial state with the method described by Glorot et al.⁵⁷ and should be trained by supervision with a series of raw tomograms as input and corresponding example of segmentation as output. The effective output of the network is compared to a given segmented sample in a loss function (or simply loss hereafter), denoted by \mathcal{L} in Fig. 2a. The loss can be translated, to some extent, as the distance between the result and the expectation. Thanks to the differentiability of all the operations in the network and the propagation derived from the chain rule loss (also called back-propagation, which contrasts with a forward propagation by giving an input image and obtaining an output segmentation), it is possible to calculate the partial derivative for each weight with respect to the loss. We optimize the weights with a gradient descent technique^{58,59}, which consists of shifting each weight by a certain amount against the sign of partial derivative. With a significant number of iterations of computing the forward/backward propagation and leveling weights, the overall network will converge to a point where the predicted result remains as expected. In simple words, the CNN “self-learns” to uncover hidden logic or representations from input images to output segmentation.

Sampling and composition of datasets

Although a tomography experiment can generate a few gigabytes of raw tomograms, annotating phases on tomography images to “teach” CNN can be tedious and extremely time-consuming for some datasets. In the case of the NMC dataset in Fig. 1a, an average of one hour should be considered to obtain a good quality ground truth. On the flip side, CNN is well known, data hungry, and typically fueled by thousands, if not millions of images.

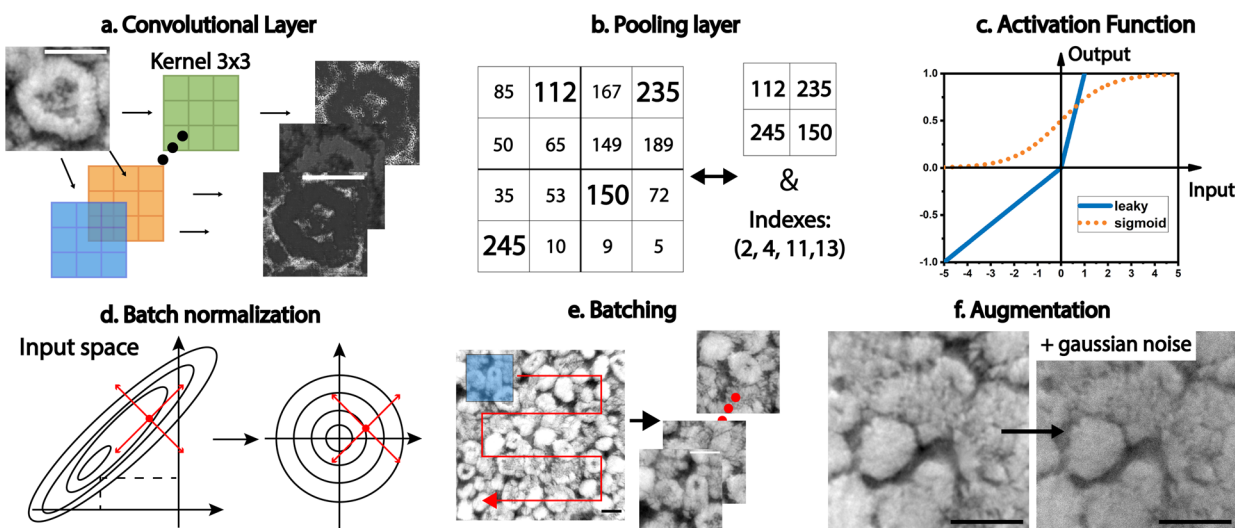


Fig. 6 Illustration of different components of a neural network with the studied architecture. **a–c** Illustrate basic operations such as convolutions, pooling, and activation function in a CNN. **d** The batch normalization layer that adjusts variance between the inputs for faster convergence. **e** The random batching, which generates a substantial quantity of data from a limited number of labeled images. **f** Illustration of the artificial noise augmentation technique applied on the input of the network to increase the data size and improve the insensibility to the noisy data while predicting. The scale bars represent 5 μm .

Limited by the amount of annotated data and given the need to diversify the data for the robustness of the invisible data prediction, two strategies are applied: (a) the small patches are cropped randomly and synchronously in the input image and the labeled image (Fig. 6e). (b) the variation in contrast, noises, and distortions are added at random to the cropped raw tomogram, namely hereafter augmentation (Fig. 6f).

Throughout this work, a single slice of the raw tomographic image is used for the CNN training dataset. Two more slices, perpendicular to the same direction in the studied volume and distant from each other (to avoid similarity), should be chosen and segmented as the validation and test datasets. The training dataset is only used to update the weight, while the validation data is used to assess the predicting accuracy of CNN to invisible images. The training/validation should be repeated if the structure and HPs are adjusted. And the test dataset serves to confirm the performance of the final optimized CNN.

For the transfer learning, a new volume of NMC₂ cathode was used, and in the current study, the already trained NMC₁ dataset was not diluted in the second one. All the data used here is published in TomoBank data repository⁶⁰.

Material preparation

The two studied 3D volumes depicted in Figs. 1a and 5a are Li-ion battery cathode material LiNi_{0.5}Mn_{0.2}Co_{0.3}O supplied by industry. A Zeiss Laser Dissector is used to cut the material into a particular pattern with the central 50 μm of diameter cylinder (Supplementary Fig. 9, the pattern under the optical microscopy). We use a strongly sharpened pencil lead slightly dipped in the epoxy and approach the pattern with a micromanipulator with an angle of 90°. Let the epoxy polymerize for 15 min. We pulled back the pencil lead in the opposite direction, and the cylinder was detached from the bulky electrode. The Li-O₂ battery material is prepared differently. Two binder-free (NanoTech Lab) electrodes from the same patch were made of MWCNTs (purchased from NanoTech Lab) with the filtration method. One of them was cycled in a Swagelok for a complete round-trip between 2 and 4.3 V at a constant current density of 20 mA/g_{carbon}. It is then prepared in a dry room as the cycling products are unstable in the presence of water. The pristine and recharged electrodes are both chopped with a blade, and then a small piece was picked with the same method of epoxy under a microscope. The cycled Li-O₂ cathode is sealed immediately inside a Kapton capillary with Torr Seal after the sampling. As the Kapton is transparent to 8 keV X-ray, the TXM can be directly performed on the capillary, where the samples are protected from the air during the transport and acquisition.

Nano-CT experiment and tomographic reconstruction

The pencil lead with the material was placed on the rotation stage of APS ID-32-C beamline⁶¹. A zone plate condenser at 8 keV energy with a working distance of 3.4 m is used. ~1200 frames of projection with equal angle delta within 180° degree are collected on the fly. The projections are reconstructed by FBP-CUDA in Astra-TomoPy⁶² Python library. To obtain a better contrast, the authors noticed that analytical reconstruction such as FBP is preferred to the alternative algorithm like SIRT, with which it is unable to differentiate the CBD from the background porosity as their grayscale are too closed. A 3D median filter of kernel three and an optional 2D unsharp mask of radius six and weight 0.6 have been applied before all the segmentations in this work.

Synthetic training data algorithm

The algorithm perturbs the training dataset by pushing or pulling random pixels on the interfaces of a segmentation. The operation of this algorithm is to locate all the voxels firstly on three types of interfaces in our multiphase segmentation problem. A 3D median will randomly pick a percentage of these interface voxels to apply a dilatation on either side. It will be corrosion for one phase and dilatation for the other phase of the interface. We found that this algorithm synthesizes more realistic segmentation in 3D than in 2D. This is because there might be interfaces in a neighbor plan (e.g., Fig. 5a) that will not be considered in 2D. Yet in 3D (e.g., Fig. 5b), the consideration of the adjacent plans makes the synthetic results more realistic. At least ten adjacent slices of raw tomogram were well-segmented and used as the input of this algorithm. The two parameters to tune in this algorithm are the surface voxel picking ratio of the interface and the number of iterations. We found that 10% interface for each iteration and five iterations generate the best perturbed data with homogeneous and plausible changes that visibly difficult to distinguish like Supplementary Fig. 6.

Hardware and software

The CNN training is run on a PC with Ubuntu OS equipped with Intel Xeon CPU and Quadro P5000 GPUs. The SegmentPy utilized in this work is an in-house open-source software. Its neural network part is based on TensorFlow and mpi4py. And it is open source and can be downloaded on github.SegmentPy.io.

DATA AVAILABILITY

The battery tomography datasets used in this contribution is available for free download at <https://tomobank.readthedocs.io>.

CODE AVAILABILITY

The SegmentPy python software is available for free download at <https://segmentpy.readthedocs.io>.

Received: 23 November 2020; Accepted: 5 January 2022;

Published online: 09 February 2022

REFERENCES

- Ebner, M., Chung, D.-W., Garcia, R. E. & Wood, V. Tortuosity Anisotropy in Lithium-Ion Battery Electrodes. *Adv. Energy Mater.* **4**, 1301278 (2014).
- Finegan, D. P. et al. In-operando high-speed tomography of lithium-ion batteries during thermal runaway. *Nat. Commun.* **6**, 6924 (2015).
- Pietsch, P. & Wood, V. X-Ray Tomography for Lithium Ion Battery Research: a practical guide. *Annu. Rev. Mater. Res.* **47**, 451–479 (2017).
- Müller, S. et al. Multimodal Nanoscale Tomographic Imaging for Battery Electrodes. *Adv. Energy Mater.* **10**, 1904119 (2020).
- Tan, C. et al. Four-Dimensional Studies of Morphology Evolution in Lithium–Sulfur Batteries. *ACS Appl. Energy Mater.* **1**, 5090–5100 (2018).
- Pietsch, P. et al. Quantifying microstructural dynamics and electrochemical activity of graphite and silicon-graphite lithium ion battery anodes. *Nat. Commun.* **7**, 12909 (2016).
- Yu, Y.-S. et al. Three-dimensional localization of nanoscale battery reactions using soft X-ray tomography. *Nat. Commun.* **9**, 921 (2018).
- Eastwood, D. S. et al. Lithiation-Induced Dilation Mapping in a Lithium-Ion Battery Electrode by 3D X-Ray Microscopy and Digital Volume Correlation. *Adv. Energy Mater.* **4**, 1300506 (2014).
- Vanpeene, V., King, A., Maire, E. & Roué, L. In situ characterization of Si-based anodes by coupling synchrotron X-ray tomography and diffraction. *Nano Energy* **56**, 799–812 (2019).
- Nguyen, T.-T. et al. The electrode tortuosity factor: why the conventional tortuosity factor is not well suited for quantifying transport in porous Li-ion battery electrodes and what to use instead. *Npj Comput. Mater.* **6**, 123 (2020).
- Müller, S. et al. Quantifying Inhomogeneity of Lithium Ion Battery Electrodes and Its Influence on Electrochemical Performance. *J. Electrochem. Soc.* **165**, A339–A344 (2018).
- Chouchane, M., Rucci, A., Lombardo, T., Ngandjong, A. C. & Franco, A. A. Lithium ion battery electrodes predicted from manufacturing simulations: assessing the impact of the carbon-binder spatial location on the electrochemical performance. *J. Power Sources* **444**, 227285 (2019).
- Shodiev, A. et al. 4D-resolved physical model for Electrochemical Impedance Spectroscopy of Li(Ni_{1-x}YMn_xCo_y)O₂-based cathodes in symmetric cells: Consequences in tortuosity calculations. *J. Power Sources* **454**, 227871 (2020).
- Lu, X. et al. 3D microstructure design of lithium-ion battery electrodes assisted by X-ray nano-computed tomography and modelling. *Nat. Commun.* **11**, 2079 (2020).
- Pietsch, P., Ebner, M., Marone, F., Stampanoni, M. & Wood, V. Determining the uncertainty in microstructural parameters extracted from tomographic data. *Sustain. Energy Fuels* **2**, 598–605 (2018).
- Guntoro, P. I., Ghorbani, Y., Koch, P.-H. & Rosenkranz, J. X-ray Microcomputed Tomography (μ CT) for Mineral Characterization: A Review of Data Analysis Methods. *Minerals* **9**, 183 (2019).
- Lombardo, T. et al. Artificial Intelligence Applied to Battery Research: Hype or Reality? *Chem. Rev.* [acs.chemrev.1c00108](https://doi.org/10.1021/acs.chemrev.1c00108) <https://doi.org/10.1021/acs.chemrev.1c00108> (2021).
- Arganda-Carreras, I. et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017).
- Berg, S. et al. ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).
- Demir, I. et al. DeepGlobe 2018: a Challenge to Parse the Earth through Satellite Images. Deepglobe 2018: A challenge to parse the earth through satellite images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 172–181. (2018).
- Tuccillo, D. et al. Deep learning for galaxy surface brightness profile fitting. *Mon. Not. R. Astron. Soc.* **475**, 894–909 (2018).
- Taigman, Y., Yang, M., Ranzato, M. & Wolf, L. DeepFace: closing the Gap to Human-Level Performance in Face Verification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1701–1708 (2014).
- Bochkovskiy, A., Wang, C.-Y. & Liao, H.-Y. M. YOLOv4: optimal speed and accuracy of object detection. Preprint at <https://arxiv.org/abs/2004.10934>.
- Badrinarayanan, V., Kendall, A. & Cipolla, R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 2481–2495 (2017).
- Geurts, P., Irthum, A. & Wehenkel, L. Supervised learning with decision tree-based methods in computational and systems biology. *Mol. Biosyst.* **5**, 1593 (2009).
- Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (eds. Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F.) 234–241 (Springer International Publishing, 2015).
- Shashank Kaira, C. et al. Automated correlative segmentation of large Transmission X-ray Microscopy (TXM) tomograms using deep learning. *Mater. Charact.* **142**, 203–210 (2018).
- Tekawade, A., Sforzo, B. A., Matusik, K. E., Kastengren, A. L. & Powell, C. F. High-fidelity geometry generation from CT data using convolutional neural networks. In *Developments in X-Ray Tomography XII* (eds. Müller, B. & Wang, G.) 67 (SPIE, 2019).
- Simpson, A. L. et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. Preprint at <http://arxiv.org/abs/1902.09063>.
- Jiang, Z. et al. Machine-learning-revealed statistics of the particle-carbon/binder detachment in lithium-ion battery cathodes. *Nat. Commun.* **11**, 2310 (2020).
- LaBonte, T., Martinez, C. & Roberts, S. A. We Know Where We Don't Know: 3D Bayesian CNNs for Uncertainty Quantification of Binary Segmentations for Material Simulations. Preprint at <https://arxiv.org/abs/1910.10793>.
- Yang, X. et al. Low-dose x-ray tomography through a deep convolutional neural network. *Sci. Rep.* **8**, 2575 (2018).
- Schuman, C. D., Plank, J. S., Bruer, G. & Anantharaj, J. Non-Traditional Input Encoding Schemes for Spiking Neuromorphic Systems. in 2019 International Joint Conference on Neural Networks (IJCNN) 1–10 (IEEE, 2019).
- Parsa, M. et al. Bayesian Multi-objective Hyperparameter Optimization for Accurate, Fast, and Efficient Neural Network Accelerator Design. *Front. Neurosci.* **14**, 667 (2020).
- Bergstra, J. & Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **13**, 281–305 (2012).
- Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **104**, 148–175 (2016).
- Moroni, R. et al. Multi-Scale Correlative Tomography of a Li-Ion Battery Composite Cathode. *Sci. Rep.* **6**, 30109 (2016).
- Finegan, D. P. et al. Spatial quantification of dynamic inter and intra particle crystallographic heterogeneities within lithium ion electrodes. *Nat. Commun.* **11**, 631 (2020).
- Müller, S. et al. Deep learning-based segmentation of lithium-ion battery microstructures enhanced by artificially generated electrodes. *Nat. Commun.* **12**, 6205 (2021).
- Nguyen, T. et al. 3D Quantification of Microstructural Properties of LiNi_{0.5}Mn_{0.3}Co_{0.2}O₂ High-Energy Density Electrodes by X-Ray Holographic Nano-Tomography. *Adv. Energy Mater.* **11**, 2003529 (2021).
- Nikitin, V. et al. Distributed Optimization for Nonrigid Nano-Tomography. *IEEE Trans. Comput. Imaging* **7**, 272–287 (2021).
- Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems* (eds. Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. & Weinberger, K. Q.) vol. 27 (Curran Associates, Inc., 2014).
- Su, Z. et al. X-ray Nanocomputed Tomography in Zernike Phase Contrast for Studying 3D Morphology of Li–O₂ Battery Electrode. *ACS Appl. Energy Mater.* **3**, 4093–4102 (2020).
- Ebner, M., Geldmacher, F., Marone, F., Stampanoni, M. & Wood, V. X-Ray Tomography of Porous, Transition Metal Oxide Based Lithium Ion Battery Electrodes. *Adv. Energy Mater.* **3**, 845–850 (2013).
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N. & Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (eds. Stoyanov, D. et al.) 3–11 (Springer International Publishing, 2018).
- Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. & Liang, J. UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Trans. Med. Imaging* **39**, 1856–1867 (2020).
- Pelt, D. M. & Sethian, J. A. A mixed-scale dense convolutional neural network for image analysis. *Proc. Natl Acad. Sci.* **115**, 254–259 (2018).
- Ziabari, A. et al. 2.5D Deep Learning For CT Image Reconstruction Using A Multi-GPU Implementation. In: Proceedings of the 2018 52nd Asilomar Conference on Signals, Systems, and Computers, 2044–2049 (2018).
- Strohmann, T. et al. Semantic segmentation of synchrotron tomography of multiphase Al–Si alloys using a convolutional neural network with a pixel-wise weighted loss function. *Sci. Rep.* **9**, 19611 (2019).
- Krygier, M. C. et al. Quantifying the unknown impact of segmentation uncertainty on image-based simulations. *Nat. Commun.* **12**, 5414 (2021).

51. Shaw, A., Hunter, D., Landola, F. & Sidhu, S. SqueezeNAS: Fast Neural Architecture Search for Faster Semantic Segmentation. in 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) 2014–2024 (IEEE, 2019).
52. Yan, X., Jiang, W., Shi, Y. & Zhuo, C. MS-NAS: Multi-scale Neural Architecture Search for Medical Image Segmentation. in Medical Image Computing and Computer Assisted Intervention-MICCAI 2020 (eds. Martel, A. L. et al.) 388–397 (Springer International Publishing, 2020).
53. Amirreza Shaban, I. E., Shray Bansal, Zhen Liu & Boots, B. One-Shot Learning for Semantic Segmentation. in Proceedings of the British Machine Vision Conference (BMVC) (eds. Tae-Kyun Kim, G. B., Stefanos Zafeiriou & Mikolajczyk, K.) 167.1–167.13 (BMVA Press, 2017).
54. Wang, K., Liew, J. H., Zou, Y., Zhou, D. & Feng, J. PANet: Few-Shot Image Semantic Segmentation With Prototype Alignment. in 2019 IEEE/CVF International Conference on Computer Vision (ICCV) 9196–9205 (IEEE, 2019).
55. Ioffe, S. & Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. in Proceedings of the 32nd International Conference on Machine Learning (eds. Bach, F. & Blei, D.) vol. 37 448–456 (PMLR, 2015).
56. Wu, Y. & He, K. Group Normalization. *Int. J. Comput. Vis.* **128**, 742–755 (2020).
57. Glorot, X. & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, 249–256 (2010).
58. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. Preprint at <https://arxiv.org/abs/1412.6980v9>.
59. Ruder, S. An overview of gradient descent optimization algorithms. Preprint at <https://arxiv.org/abs/1609.04747>.
60. De Carlo, F. et al. TomoBank: a tomographic data repository for computational x-ray science. *Meas. Sci. Technol.* **29**, 034004 (2018).
61. De Andrade, V. et al. Nanoscale 3D imaging at the Advanced Photon Source. *SPIE Newsroom* (2016) <https://doi.org/10.1117/2.1201604.006461>.
62. Pelt, D. M. et al. Integration of TomoPy and the ASTRA toolbox for advanced processing and reconstruction of tomographic synchrotron data. *J. Synchrotron Radiat.* **23**, 842–849 (2016).

ACKNOWLEDGEMENTS

This research is supported by the French Ministry Higher Education, Research and Innovation. The authors are grateful for the participation of the researchers in the workshop of NanOperando (GDR CNRS N°2015) for the ground truth survey (2019/11, Energy Hub, Amiens, France). A.A.F. acknowledges Institut Universitaire de France for the support. The authors are also grateful for the re-lecture by D. Boursier. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

AUTHOR CONTRIBUTIONS

A.D., E.D., and Z.L.S. have conceived the investigation. This work was supervised by A.D, E.D., and A.A.F. The datasets were acquired and reconstructed by Z.L.S., T.-T.N., and V.D. The software was coded by Z.L.S. and partially contributed by K. E.-A., Z.L.S., and A.D. wrote the paper. All authors participated in the discussion and revision of this paper and finally approved this work.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-022-00709-7>.

Correspondence and requests for materials should be addressed to Arnaud Demortière.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022