ROYAL SOCIETY OF CHEMISTRY

PAPER

Check for updates

Convolutional neural networks for high throughput screening of catalyst layer inks for polymer electrolyte fuel cells†

Mohammad J. Eslamibidgoli, [ID] ‡*[a] Fabian P. Tipp,‡[b] Jenia Jitsev, [ID] [c] Jasna Jankovic,[d] Michael H. Eikerling [ID] [ae] and Kourosh Malek*[a]

The performance of polymer electrolyte fuel cells decisively depends on the structure and processes in membrane electrode assemblies and their components, particularly the catalyst layers. The structural building blocks of catalyst layers are formed during the processing and application of catalyst inks. Accelerating the structural characterization at the ink stage is thus crucial to expedite further advances in catalyst layer design and fabrication. In this context, deep learning algorithms based on deep convolutional neural networks (ConvNets) can automate the processing of the complex and multi-scale structural features of ink imaging data. This article presents the first application of ConvNets for the high throughput screening of transmission electron microscopy images at the ink stage. Results indicate the importance of model pre-training and data augmentation that works on multiple scales in training robust and accurate classification pipelines.

1 Introduction

Despite decades of research, the rapid improvement of the effective electrocatalytic activity and durability of catalyst layers (CLs) remains a foremost challenge in driving forth the commercialization of polymer electrolyte fuel cells (PEFCs).[1–5] Accelerating the fabrication process and the evaluation of CLs is crucial in view of two primary goals in PEFC development: drastically reduce the Pt loading and enable operation at high current and power density.[6–11]

Conventional catalyst layers are fabricated from ink-based recipes that include a solid-state catalyst of Pt nanoparticles supported typically on a high-surface-area carbon (Pt/C), an ionomer dispersion (*e.g.*, Nafion or Aquivion), and a water–alcohol mixture as the solvent.[12–14] Ink-based methods allow for the rapid and reproducible manufacturing of membrane electrode assemblies. In ink, Pt/C particles assemble into porous

agglomerates, with small intra-agglomerate pores (primary pores, <10 nm) and larger pores (secondary pores, 30–150 nm) being formed in the inter-agglomerate space.[15,16] Ionomer molecules assemble into a network of skin-type films that partially encapsulate agglomerates, with a thin proton-conducting water film formed between the agglomerate surface and ionomer films. The resulting bi-modal pore size distribution, as well as the heterogeneous wettability of pore surfaces determine the liquid water saturation in the CL as well as other porous fuel cell media, thereby exerting a significant impact on PEFC performance.[17]

Ink-based fabrication affords variations in Pt loading, ionomer content, carbon content and type of carbon, the size distribution of Pt and C particles, wetting properties of pore walls, the thickness of the applied CL, dielectric properties of the ink solvent, ink processing conditions, temperature, and stirring rate. In the past two decades, optimization in this multidimensional parameter space has been extensively pursued.[18–21] Following ink fabrication, the ink is applied to the membrane or deposited onto a substrate by a wet-coating technique to prepare a catalyst coated membrane (CCM) or membrane electrode assembly (MEA). In combination with the CL deposition process, the properties of the starting ink determine the structure, properties and performance of resulting MEAs.[22–26] The correlations among fabrication parameters, ink ingredients, intrinsic materials properties, microstructure, and operating conditions are difficult to unravel. Considerable uncertainty and controversy thus remain regarding the key descriptors that must be monitored and controlled in the

[a] Theory and Computation of Energy Materials (IEK-13), Institute of Energy and Climate Research, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany. E-mail: m.eslamibidgoli@fz-juelich.de; k.malek@fz-juelich.de

[b] Department of Chemistry, University of Cologne, Greinstr. 4-6, 50939 Cologne, Germany

[c] Jülich Supercomputing Center, Forschungszentrum Jülich, 52425 Jülich, Germany

[d] Department of Materials Science and Engineering, University of Connecticut, 97 North Eagleville Road, Unit 3136, Storrs, CT 06269-3136, USA

[e] Chair of Theory and Computation of Energy Materials, Faculty of Georesources and Materials Engineering, RWTH Aachen University, Aachen 52062, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ra05324h

‡ Equal contribution.

manufacturing process of MEAs to deliver cells with optimal performance.

Progress in CL design is achieved, in large part, through modification at the materials level. Ultrahigh activities were obtained by tuning the composition of Pt–Ni alloys and varying the shape and morphology of nanoparticles, *e.g.*, Pt-based catalysts with nanocage,[27] nanoframe,[28] nanowires[29] or octahedral[30] shapes. Other modification strategies were pursued in the engineering of the film thickness, surface morphology, and surface roughness of extended surface catalysts[31] or changing the support material.[32] Local reaction conditions can be improved by embedding catalysts into different electrodes, such as ionomer-impregnated porous gas diffusion electrodes or ionomer-free ultrathin flooded electrodes.[33,34]

Electron microscopy is the most widely employed characterization method in nanoscience. It provides images at (sub-)nanometer resolution, including quantitative information about particle and pore size distributions, crystal structure, materials morphology, and composition.[35–37] Transmission electron microscopy (TEM) has become a standard imaging technique to study the microstructure of MEA components before cell assembly or post-testing.[38–42] High-resolution imaging techniques such as high-resolution TEM (HR-TEM), high-angle annular dark field-scanning transmission electron microscopy (HAADF-STEM), 3D electron tomography, or high-resolution atomic force microscopy (HR-AFM) have yielded 2D or 3D images of catalyst layer microstructures in both inks and deposited layers.[43–49] Many useful features in the TEM images of catalyst layer inks are difficult to distinguish upon visual examination. These features, however, can be automatically determined by machine learning algorithms.

Recent advances in autonomous materials fabrication using lab-scale stationary or mobile robotic systems[50–55] demand a bidirectional and rapid data flow from catalyst fabrication to ink preparation, CCM fabrication, MEA assembly, and a half or complete cell design, including testing and *in situ/ex situ* characterization steps. This article focuses on the subsection of the discovery pipeline from fuel cell ink to device. In this area, data workflows lack algorithms for extracting information from imaging data.[56] An AI-driven algorithm for rapid image processing could assist materials discovery-on-a-chip by improving the data workflow. This approach, complemented by data acquisition units and physical modeling and by high-throughput communication with AI-driven processes during the primary materials design stage, can be utilized for optimization purposes at all levels from ink to CCM and further to complete cell or stack.

2 Methods

Our methodical pipeline for TEM image recognition of catalyst layer inks consists of three steps: (1) data sampling and data augmentation at multiple scales; (2) making use of a network model pre-trained on a large generic natural images dataset, ImageNet, for transfer learning; and (3) training a classifier for ink recognition. In the first step, the model performance is evaluated against the prepared fixed-scale datasets. In the second step, model performance is compared with a custom ConvNet model trained from scratch. Finally, in the last step, we present a fast and accurate approach for ink classification by training a logistic regression model on extracted features obtained from transfer learning.

2.1 Dataset preparation

We used TEM imaging data provided by Automotive Fuel Cell Cooperation (AFCC) for our investigation from two independent experimental studies of catalyst powder agglomeration with Nafion and Aquivion ionomers employed as hinders. Three ink powders were studied with Nafion ionomer, EA50 I33 ink, F50 I33 ink, and V50 I33 ink (EA – graphitized, F – stabilized and V – Vulcan carbon support with Pt from TKK, and 33 wt% Nafion ionomer), and three with Aquivion ionomer, EA50 I33 ink, F50E I33 ink, and V50 I33 ink. Further information about the raw images and initial processing is provided in the ESI.† The ink agglomerates encompass multi-scale structural features, which are, in fact, highly relevant to the physical descriptors of CL performance and lifetime. As exemplified in Fig. 1(a) at different magnification levels, these descriptors include the size, morphology, and distribution of carbon particles and catalyst nanoparticles, ionomer aggregates, or pore networks. As TEM images display structural features at different length scales and resolutions, a ConvNet model's main objective is to extract these features from the input data effectively. Such features should make it possible to accurately identify the optimal *vs.* non-optimal samples or reliably sort large databases in terms of the type of material. Therefore, during the training, the input dataset should be diverse enough to enable the emergence of such multi-scale structural features in the network.

To prepare a diverse input data space, we employed data augmentation techniques[58] *via* region proposal algorithms – primarily designed for the object detection problems.[57] Two different region-specific approaches were employed: (1) a naïve and relatively inefficient Sliding Window (SW) technique and (2) the more efficient Selective Search (SS) algorithm.[57] By comparing the two approaches, we aim to demonstrate the effectiveness of using the latter. Given the high-resolution TEM ink data at large length scales (200–500 nm), these methods also generate many of examples for training and validation purposes required in the deep learning model development. Before applying the region-specific approaches, the images in each class were split into training and hold-out sets; this ensures no leakage between the two sets (see ESI† for details).

As shown in Fig. 1(b), to sample the imaging data using the SW approach, varying window sizes of 100×100, 200×200, 300×300, and 500×500 pixels were slid over the raw ink images (see ESI†). This was performed both on the training set and the hold-out set. At each stop of the window, the corresponding fixed-scale image patch was extracted. This window was initiated at the top left of each image and moved 25-pixel steps to the right before moving back to the leftmost edge. After these steps, the window was moved down by 25 pixels, and this whole procedure was repeated. The main drawback of this technique is that it does not discriminate between parts of the
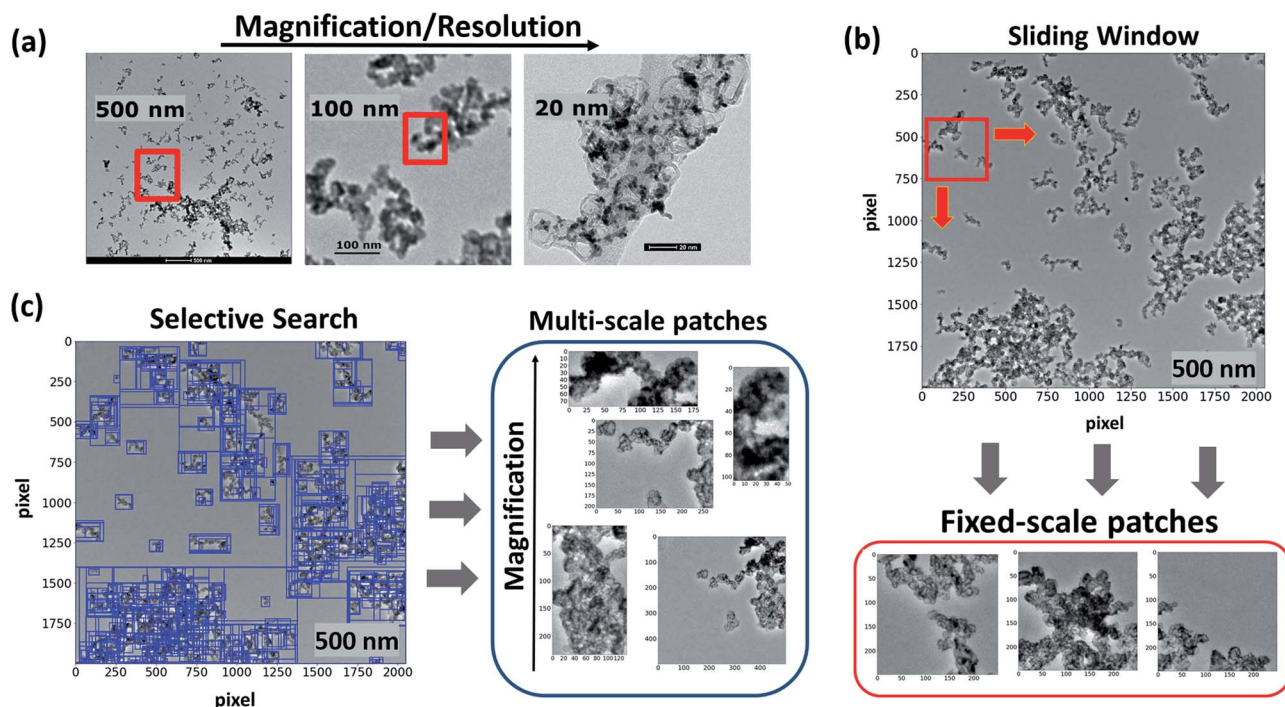
Fig. 1 (a) An example of real-world images of catalyst ink agglomerates at different magnification levels or resolutions. (b) Representation of the sliding window technique to sample fixed-scale patches from high-resolution TEM images of catalyst inks. (c) Representation of the selective search algorithm[57] to sample patches from high-resolution TEM images of inks based on hierarchical grouping of similar regions such as color, size, texture, and shape compatibility.

images that contain the agglomerates from those that are mostly part of the background, which leads to creating a noisy training set. Moreover, the choice of the aspect ratio of sliding window is arbitrary; this can diminish the ability to extract useful features.

To address the first issue and filter out images with a high amount of background, only those images in which the mean grayscale value of the pixels is lower than 115–155 were saved, depending upon their respective classes (the possible values for the pixels range from 0 (black) to 255 (white)). This was performed because the agglomerates are significantly darker than the background, where the lower values of the pixels correspond to a darker tone. Therefore, a lower mean pixel value of an image generally corresponds to a more significant proportion of the image being occupied by the agglomerates. After this initial step, the images were further inspected to remove images containing a large amount of background. As for the second issue, we used various window sizes as described above. Therefore, in our first approach to creating the dataset, we combined several fixed-scale extracted patches using the SW technique, which is inefficient.

As shown in Fig. 1(c), we employed the selective search algorithm as the second approach for data augmentation and extracted patches from the high-resolution ink images.[57] Selective search is a region proposal algorithm designed to localize possible object locations in the input image based on the hierarchical grouping of features and regions subject to the similarity of diverse factors such as the color, texture, lighting conditions, and composition – referred to diversification.

Selective search has outperformed several previous region proposal algorithms in execution time, repeatability, proposal recall, detection, and mean average precision.[59] On our dataset it also resolved both issues with the SW approach: the undesired background patches and the lack of variability in the aspect ratio of the extracted images. This is because SS aims to select the image parts containing agglomerates at different scales and with varying sizes.

Full implementation of the SS algorithm is available in Python, which includes three modes according to various diversification strategies explained in the original paper.[57] For our dataset preparation, the SS in the mode 'quality' was applied to each high-resolution ink image in the initially separated training and hold-out sets. We discarded patches whose length or width was under 100 pixels from this dataset due to their relatively low resolution. Moreover, the ratio between the height and the width or between the width and the height was also calculated; all images with one ratio greater than two were discarded. From these resulting filtered images, about 2000 images for the training set and 2000 for the hold-out set were selected for each class (see Table 1 in the ESI†). It was now ensured that the selected images were sourced in roughly equal proportions.

2.2 ConvNets for image classification

In a typical ConvNet architecture for classification,[60] we provide as input an image and predict as output the probability distribution over possible classes, which sums up to 1. An image is a 3D tensor (row × column × channel). The ConvNet performs

a set of convolutions by sliding over spatial locations of the image and applying learnable kernels (filters or feature detectors) to generate transformed representations known as activation maps that can be interpreted as signaling for particular learned features. Activation maps in hidden layers are as well organized in 3 dimensions. The activation maps have smaller spatial dimensions but larger depth in channel dimension than the original input image. Each activation map channel learns to respond to certain features of the image. As we go deeper into convolutional layers, filters become more complex, reflecting high-level features in the image.[61] For example, in the first convolutional layer channels, may detect low-level features like edges, while second layers may detect various edge compositions and so on. For most ConvNets, the convolution operations are often followed by applying activation functions and pooling operations. Activation functions, such as Rectified Linear Units (ReLU), introduce simple non-linearities into the layer-wise signal transformation that are essential for sufficient model expressiveness while helping to avoid vanishing gradients during model optimization. Pooling constitutes a down-sampling process that operates over each activation map independently, creating a smaller activation map. The max-pooling operation (pooling the maximum pixel value) is common alongside average pooling. Flattening or global average pooling are also applied to convert multiple feature maps to a single vector representation used as an input for a fully connected layer or classifier in the last step.[61]

After a series of the convolutional layers, fully connected or global pooling layers are applied to generate a 1D feature vector. A classifier is then applied to generate the probability distribution over each class. In order to classify the image, the class corresponding to the largest probability value is taken. Once the class probability distribution is generated based on the current ConvNet weights, loss signaling mismatch between the network output and true class assignments is computed. Back-propagation is then applied to compute the gradient of the loss function for each weight in every kernel, and the weights are modified to minimize the loss function.

Due to the high amount of training data, time, and computational resources required to train a ConvNet model from scratch, transfer learning is usually adopted, especially when data is limited, where a pre-trained model is used to extract useful generic features from the image and enable learning. For instance, a pretrained model can be trained on a large generic natural image dataset to solve a classification problem. In our work, we employed the VGG16 model trained on the ImageNet dataset.[62] For our dataset, we found that only the first four pre-trained convolutional layers of VGG16 in conjunction with the corresponding max-pooling layers are sufficient for accurate classification (mini-VGG). The employed architectures are shown in Fig. 2.

3 Results and discussion

The dataset was prepared for the six sample classes EA, F, and V support in Nafion and Aquivion. The overall steps to train and evaluate a ConvNet model after splitting and preprocessing the

data include: (1) choosing the network architecture, (2) initializing the weights, (3) finding a learning rate and regularization strength, (4) minimizing the loss function and monitoring the progress, and (5) evaluating the model on the hold-out set.

To set a benchmark for our study, a relatively shallow custom ConvNet was constructed, shown in Fig. 2(a). Next, we evaluated various scenarios for the training and the hold out sets. Fig. 3 shows the learning curves (validation accuracy *vs.* epoch number) of the different combinations. In the first case, when the network is trained on fixed-scale extracted patches of 100×100 using SW, it performs poorly on the hold-out set generated using SS. In this case, the validation accuracy did not increase above $\approx 25\%$ indicating significant overfitting for the model. Training our custom network on the dataset generated from SW, including 100×100, 200×200, and 300×300 patches, increased the validation accuracy to $\approx 70\%$ on the SS generated hold-out set. This already suggests the importance of applying zoom-in and zoom-out data augmentation in the training phase.[63] Additionally, training on 100×100, 200×200 and 300×300 and 500×500 patches further improved the performance to $\approx 75\%$.

If we train the network on the latter dataset from SW and test it on the SW generated hold-out set, the accuracy increases to 90%. This is expected because, as discussed SS selects patches at a different aspect ratio from SW. Therefore, it is not surprising to see an increased accuracy by $\approx 15\%$ when both the training and hold-out sets have similar aspect ratios. On the other hand, our best validation accuracy was obtained when we trained our custom network on the SS generated dataset. In this case, irrespective of how we prepared the hold out set (using SW or SS), a validation accuracy of $\approx 95\text{–}97\%$ was achieved. This suggests that, given a set of high-resolution ink images at a large length scale, SS can be effectively used to prepare training examples for the training of the robust and accurate ink classification models.

A quicker approach was conducted, in which features were extracted from the first four convolutional layers of VGG16 (Fig. 2(b)), and then the feature vectors were used to train a logistic regression model for classification. Here, the first four
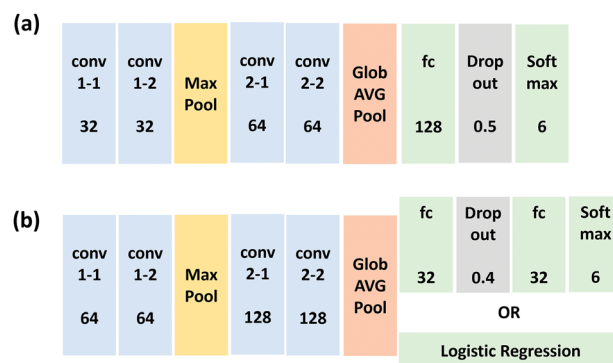

Fig. 2 (a) Architectures of the custom network trained from scratch in this study (b) that for the mini-VGG pretrained model on ImageNet dataset employed for transfer learning for the catalyst ink image classification.
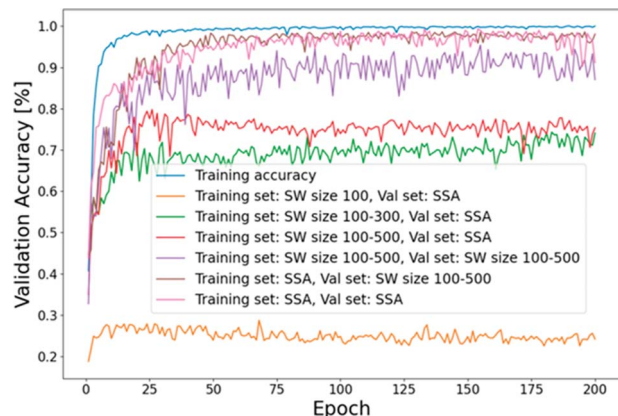
Fig. 3 The learning curves for the validation accuracy of different training and hold-out sets considered in this study (SW: sliding window, SSA: selective search algorithm).
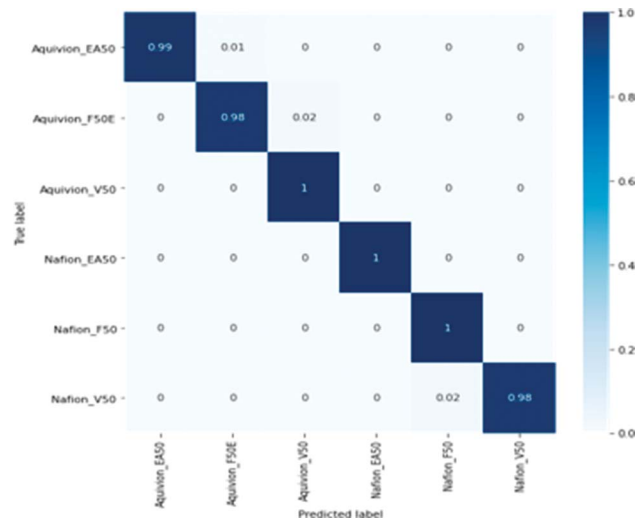

Fig. 4 The confusion matrix obtained for the classification of inks in 6 different sample categories using extracted features from the pre-trained mini-VGG model followed by logistic regression as the classifier.

convolutional layers of the pre-trained VGG16 model were extracted along with two corresponding max-pooling layers. At the end of the last layer, a global average pooling layer was added to obtain a feature vector. The training and validation images were passed through the network, and the resulting feature vectors (128 entries per vector) were saved into NumPy arrays. The features extracted from the training set were then used to fit a logistic regression model much faster than training neural networks. Next, the labels of hold-out set were predicted and compared with the true labels. Fig. 4 shows the normalized confusion matrix for the logistic regression classifier trained on the extracted features from the pre-trained mini-VGG model. The x-axis shows the predicted labels, and the y-axis shows the true labels. The nearly diagonal matrix indicates the high accuracy of the model in the classification of the six considered classes.

Table 1 compares the corresponding classification report in each class obtained from the pre-trained network with that from the custom network. Here, accuracy measures true positive and true negative classifications over the sum of true positive, true negative, false positive, and false negative. *Precision* is defined as the number of true positive (correctly labeled) classifications over the total true positive and false positive classifications. *Recall* is defined as the number of true positive classifications over the total number of true positives and false negatives. F_1-Score is the harmonic mean of precision and recall,

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (1)$$

The transfer learning approach was found to be swift and led to more accurate predictions, as summarized in Table 1. Thus, we conclude that preparing the training set with SS followed by employing a pre-trained model for feature extraction and training the logistic regression model for classification was a fast and accurate process for this problem.

Gradient-weighted Class Activation Mapping (Grad-CAM)[64] was employed to visualize the critical regions based on which

our ConvNet model predicts the class of the ink agglomerates. It employs the class-specific gradient information from the last convolutional layer to produce class-discriminative activation maps. Highlighting these regions makes it easier to interpret how the ConvNet model interacts with ink images for classification. Fig. 5 shows samples from each of the six different classes considered in our study, along with their class-discriminative localization maps. First, it is essential to note that the model does not predict the class based on unwanted noise or the background color. It investigates the relevant regions of the image, such as the edges of agglomerates where ionomer could be localized, the shape of the carbon agglomerates, the primary pores, and the Pt nanoparticles. As these features are highly relevant to the structure-function-property relations of catalyst layers in PEFCs, ConvNets can be considered highly promising for identifying optimal ink microstructures.

To show the efficacy of ConvNets for the classification of inks, a further in-depth analysis was performed to compare the structural characteristics of V50, F50, and EA50 catalyst inks

Table 1 Inner-class evaluation of the logistic regression classifier on the hold out set; features were extracted by pre-trained mini-VGG backbone architecture. Values in parentheses correspond to the custom model trained from scratch

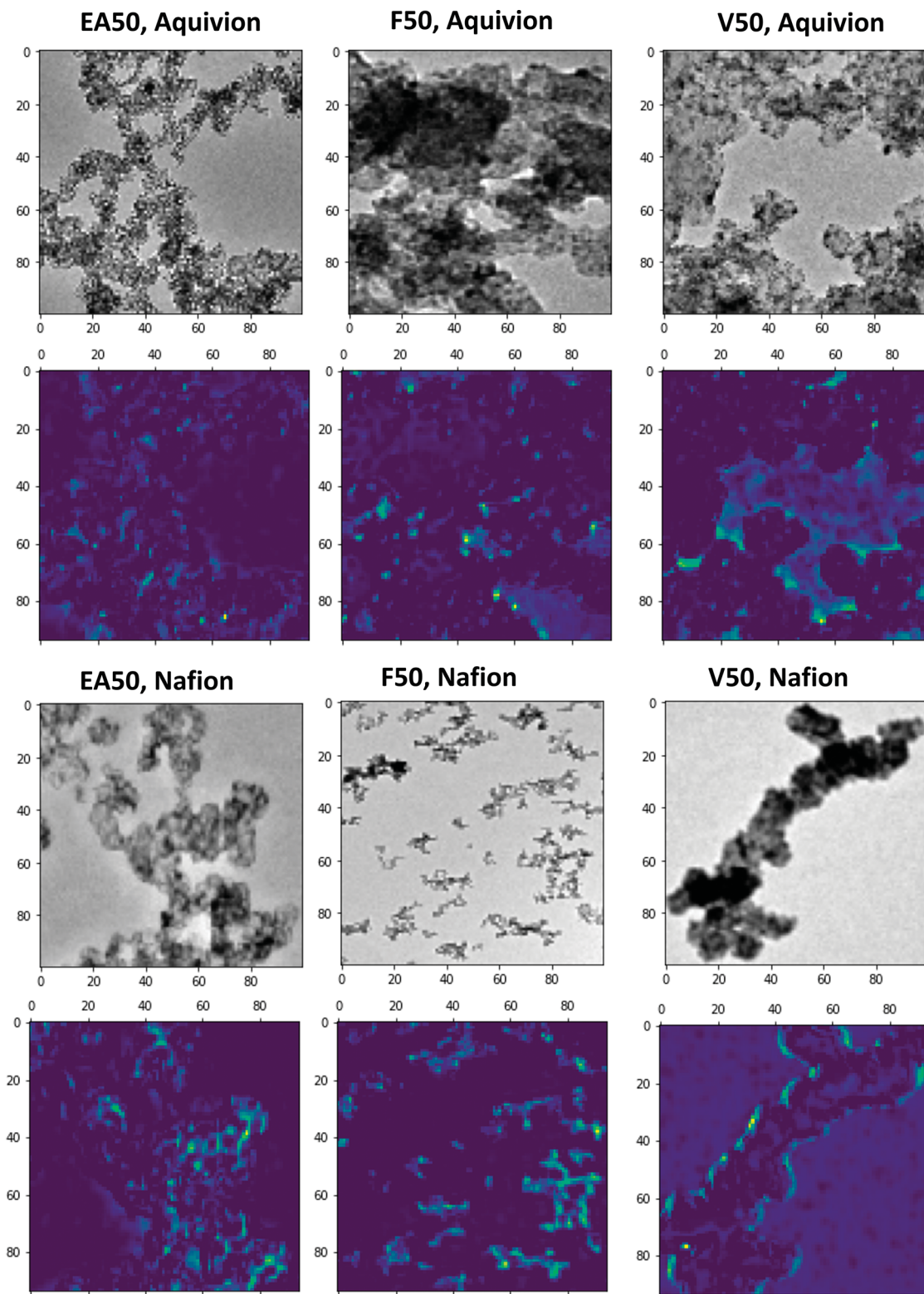| Type	Precision	Recall	F_1-Score	Num images
Aquivion-EA50	1.00 (0.96)	0.98 (0.96)	0.99 (0.96)	1869
Aquivion-F50E	0.98 (0.96)	0.99 (0.95)	0.99 (0.95)	1890
Aquivion-V50	0.99 (0.95)	0.99 (0.97)	0.99 (0.96)	1921
Nafion-EA50	1.00 (0.96)	1.00 (0.96)	0.99 (0.96)	1863
Nafion-F50	1.00 (0.96)	0.99 (0.95)	1.00 (0.95)	1882
Nafion-V50	0.98 (0.95)	1.00 (0.94)	0.99 (0.95)	1607

Fig. 5 Grad-CAM visualization of the trained custom network[64] indicating the regions where our model interacts with ink images to classify them into six different sample categories.

with Nafion. Comparisons were made by conventional image processing using ImageJ software.[65] Images were processed to segment the catalyst agglomerates and determine the surface area of agglomerates' 2D projections. The total number of analyzed agglomerates was 400 for each sample. Features smaller than the smallest carbon particle determined from the

2D TEM image analysis for each catalyst were considered noise and removed from the analysis (*e.g.*, smaller than 28 nm for V50, 16 nm for F50, and 15 nm for EA50). For comparison, the number of catalyst particles per agglomerate was estimated by dividing the total area of the agglomerate by the area of the smallest particle. As shown in Fig. 6(a), V50 comprises the largest agglomerates (with 100–150 particles), compared to F50 and EA50. F50 has mostly the agglomerates of middle size (with 5–80 particles), compared to V50 and EA50. EA50, on the other hand, consists of very large agglomerates (>180 particles) compared to V50 and F50. On the other hand, V50 and EA50 have the smallest agglomerates (<5 particles) compared to F50. A comparison of the agglomerate size distribution of different samples based on the 2D area size of the agglomerates is provided in Fig. 6(b). Table 2 reports the Carbon and Pt average particle size and standard deviation for each catalyst powder. While V50 has larger C particles compared to F50 and EA50, the average Pt particle size is smaller for V50 as compared to F50 and EA50.

Overall, this study presents a first attempt to use ConvNets to classify the 2D TEM images of catalyst layer inks accurately. Understanding how ConvNets learn from these imaging data is a key to develop robust models for swift and accurate characterization workflows. Structural quantification using for example particle or pore size distributions can also be achieved by developing deep learning-based segmentation models, a topic of ongoing research in our group. Systematic experiments should accompany deep learning-based model development CLs' image processing to generate benchmark imaging datasets. Having a statistically representative dataset will further provide essential input for ConvNets to unravel the complex structure-property-function relations. Once trained over various fabrication formulations, operating criteria, and testing conditions, the presented approach based on ConvNet-based models can be used as a practical guideline for rapid screening and optimizing the catalyst ink compositions.

Table 2 Comparison of C support and Pt particle/agglomerate size from 2D TEM images

Type	C particle size in Pt/C (nm)	Pt particle size (nm)
Nafion-V50	38 ± 10	2.5 ± 0.5
Nafion-F50	21 ± 5	2.9 ± 0.5
Nafion-EA50	21 ± 6	3.3 ± 0.9

4 Conclusions

The presented work demonstrated an application of deep learning models based on convolutional neural networks for high throughput screening of transmission electron microscopy images of catalyst layer inks in the polymer electrolyte fuel cells. Results suggest that convolutional neural networks can significantly improve the efficacy of knowledge mining and analyses of ink imaging data by extracting relevant structural features that are difficult to distinguish upon visual examination. It was shown that robust and accurate models should be built upon data augmentation techniques to prepare diverse input data of the ink images at multiple scales. These features comprise crucial structural characteristics such as particle and pore size distributions and the ionomer cluster size distribution, which are related to catalyst layers' performance and lifetime. Indeed, visualizing the activation maps from the last convolutional layer of the trained models indicated that the class-discriminative features are extracted from the relevant physical descriptors. This would suggest that the convolutional neural networks are, for example, capable of rapid discrimination between optimal from non-optimal samples.

While we specifically discussed the results for the catalyst layer ink in polymer electrolyte fuel cells, the algorithms themselves are generic. They can be trained for similar
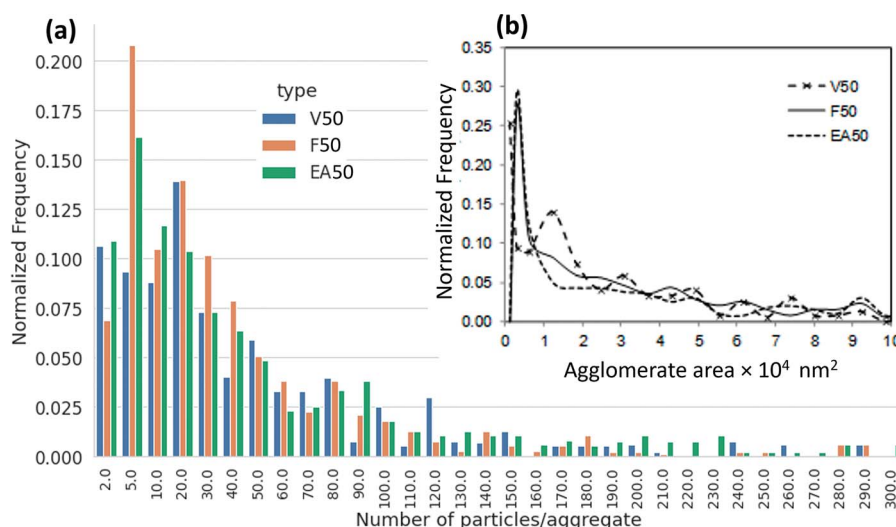

Fig. 6 Agglomerate size distribution for V50, F50, and EA50 with Nafion, (a) represented through the number of particles per agglomerate, (b) based on 2D area size of agglomerates. A total number of analyzed agglomerates for each sample is 400.

fabrication processes, *e.g.,* recognition of the slurry during fabrication of Li-ion electrodes. In the latter case, the algorithms will learn a different set of correlation factors and dependencies that govern the related structure–function relationships.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgements

Notes and references

1 K. Jiao, J. Xuan, Q. Du, Z. Bao, B. Xie, B. Wang, Y. Zhao, L. Fan, H. Wang, Z. Hou, *et al.*, *Nature*, 2021, **595**, 361–369.

2 G. Buvat, M. J. Eslamibidgoli, S. Garbarino, M. Eikerling and D. Guay, *ACS Appl. Energy Mater.*, 2020, **3**, 5229–5237.

3 G. Buvat, M. J. Eslamibidgoli, A. H. Youssef, S. Garbarino, A. Ruediger, M. Eikerling and D. Guay, *ACS Catal.*, 2019, **10**, 806–817.

4 M. J. Eslamibidgoli and M. H. Eikerling, *Electrocatalysis*, 2016, **7**, 345–354.

5 M. J. Eslamibidgoli, P.-É. A. Melchy and M. H. Eikerling, *Phys. Chem. Chem. Phys.*, 2015, **17**, 9802–9811.

6 M. Uchida, Y. Aoyama, N. Eda and A. Ohta, *J. Electrochem. Soc.*, 1995, **142**, 4143.

7 M. Uchida, Y. Fukuoka, Y. Sugawara, N. Eda and A. Ohta, *J. Electrochem. Soc.*, 1996, **143**, 2245.

8 W. Vielstich, A. Lamm and H. Gasteiger, *Handbook of fuel cells. Fundamentals, technology, applications*, 2003.

9 Q. Wang, M. Eikerling, D. Song, Z. Liu, T. Navessin, Z. Xie and S. Holdcroft, *J. Electrochem. Soc.*, 2004, **151**, A950.

10 M. Eikerling, A. A. Kornyshev and A. R. Kucernak, *Phys. Today*, 2006, **59**, 38.

11 M. Eikerling, A. A. Kornyshev and A. A. Kulikovsky, *Physical modeling of fuel cells and their components*, Wiley Online Library, 2007.

12 M. Eikerling and A. Kulikovsky, *Polymer electrolyte fuel cells: physical principles of materials and operation*, CRC Press, 2014.

13 M. J. Eslamibidgoli, J. Huang, T. Kadyk, A. Malek and M. Eikerling, *Nano Energy*, 2016, **29**, 334–361.

14 K. Malek, T. Mashio and M. Eikerling, *Electrocatalysis*, 2011, **2**, 141–157.

15 E. Sadeghi, A. Putz and M. Eikerling, *J. Electrochem. Soc.*, 2013, **160**, F1159.

16 A. Malekian, S. Salari, J. Stumper, N. Djilali and M. Bahrami, *Int. J. Hydrogen Energy*, 2019, **44**, 23396–23405.

17 T. Muzaffar, T. Kadyk and M. Eikerling, *Sustainable Energy Fuels*, 2018, **2**, 1189–1196.

18 M. H. Eikerling, K. Malek and Q. Wang, *PEM fuel cell electrocatalysts and catalyst layers*, Springer, 2008, pp. 381–446.

19 S. Holdcroft, *Chem. Mater.*, 2014, **26**, 381–393.

20 M. Chen, C. Zhao, F. Sun, J. Fan, H. Li and H. Wang, *eTransportation*, 2020, 100075.

21 A. Malek, E. Sadeghi, J. Jankovic, M. Eikerling and K. Malek, *J. Phys. Chem. C*, 2020, **124**, 3429–3438.

22 B. Garlyyev, K. Kratzl, M. Rück, J. Michalička, J. Fichtner, J. M. Macak, T. Kratky, S. Günther, M. Cokoja, A. S. Bandarenka, *et al.*, *Angew. Chem., Int. Ed.*, 2019, **58**, 9596–9600.

23 H. Ishikawa, Y. Sugawara, G. Inoue and M. Kawase, *J. Power Sources*, 2018, **374**, 196–204.

24 S. A. Berlinger, B. D. McCloskey and A. Z. Weber, *J. Phys. Chem. B*, 2018, **122**, 7790–7796.

25 J. Hou, M. Yang, C. Ke, G. Wei, C. Priest, Z. Qiao, G. Wu and J. Zhang, *EnergyChem*, 2020, **2**, 100023.

26 T. Soboleva, K. Malek, Z. Xie, T. Navessin and S. Holdcroft, *ACS Appl. Mater. Interfaces*, 2011, **3**, 1827–1837.

27 X. Tian, X. Zhao, Y.-Q. Su, L. Wang, H. Wang, D. Dang, B. Chi, H. Liu, E. J. Hensen, X. W. D. Lou, *et al.*, *Science*, 2019, **366**, 850–856.

28 C. Chen, Y. Kang, Z. Huo, Z. Zhu, W. Huang, H. L. Xin, J. D. Snyder, D. Li, J. A. Herron, M. Mavrikakis, *et al.*, *Science*, 2014, **343**, 1339–1343.

29 M. Li, Z. Zhao, T. Cheng, A. Fortunelli, C.-Y. Chen, R. Yu, Q. Zhang, L. Gu, B. V. Merinov, Z. Lin, *et al.*, *Science*, 2016, **354**, 1414–1419.

30 X. Huang, Z. Zhao, L. Cao, Y. Chen, E. Zhu, Z. Lin, M. Li, A. Yan, A. Zettl, Y. M. Wang, *et al.*, *Science*, 2015, **348**, 1230–1234.

31 J. Kibsgaard, Z. Chen, B. N. Reinecke and T. F. Jaramillo, *Nat. Mater.*, 2012, **11**, 963–969.

32 S. Ott, A. Orfanidi, H. Schmies, B. Anke, H. N. Nong, J. Hübner, U. Gernert, M. Gliech, M. Lerch and P. Strasser, *Nat. Mater.*, 2020, **19**, 77–85.

33 S. Poojary, M. N. Islam, U. N. Shrivastava, E. P. Roberts and K. Karan, *Molecules*, 2020, **25**, 3387.

34 Y. Zeng, Z. Shao, H. Zhang, Z. Wang, S. Hong, H. Yu and B. Yi, *Nano Energy*, 2017, **34**, 344–355.

35 W. Li, Z. Wang, F. Zhao, M. Li, X. Gao, Y. Zhao, J. Wang, J. Zhou, Y. Hu, Q. Xiao, *et al.*, *Chem. Mater.*, 2020, **32**, 1272–1280.

36 D. Su, *Green Energy Environ.*, 2017, **2**, 70–83.

37 D. H. Anjum, *IOP Conference Series: Materials Science and Engineering*, 2016, p. 012001.

38 J. Jankovic, *ECS Meeting Abstracts*, 2019, p. 1767.

39 S. Pedram, N. Macauley, S. Zhong, H. Xu and J. Jankovic, *ECS Trans.*, 2020, **98**, 197.

40 J. Jankovic, D. Susac, T. Soboleva and J. Stumper, *ECS Trans.*, 2013, **50**, 353.

41 J. Chen, A. Goshtasbi, A. P. Soleymani, M. Ricketts, J. Waldecker, C. Xu, J. Yang, T. Ersal and J. Jankovic, *J. Power Sources*, 2020, **476**, 228576.

42 L. G. Melo, A. P. Hitchcock, J. Jankovic, J. Stumper, D. Susac and V. Berejnov, *ECS Trans.*, 2017, **80**, 275.

43 M. Sabharwal, L. Pant, A. Putz, D. Susac, J. Jankovic and M. Secanell, *Fuel Cells*, 2016, **16**, 734–753.

44 R. L. Borup, A. Kusoglu, K. C. Neyerlin, R. Mukundan, R. K. Ahluwalia, D. A. Cullen, K. L. More, A. Z. Weber and D. J. Myers, *Curr. Opin. Electrochem.*, 2020, **21**, 192–200.

45 Y.-C. Park, H. Tokiwa, K. Kakinuma, M. Watanabe and M. Uchida, *J. Power Sources*, 2016, **315**, 179–191.

46 R. Hiesgen, T. Morawietz, M. Handl, M. Corasaniti and K. A. Friedrich, *Electrochim. Acta*, 2015, **162**, 86–99.

47 S. Helmly, M. Eslamibidgoli, K. A. Friedrich and M. Eikerling, *Electrocatalysis*, 2017, **8**, 501–508.

48 D. Rossouw, L. Chinchilla, N. Kremliakova and G. A. Botton, *Part. Part. Syst. Charact.*, 2017, **34**, 1700051.

49 H. Uchida, J. M. Song, S. Suzuki, E. Nakazawa, N. Baba and M. Watanabe, *J. Phys. Chem. B*, 2006, **110**, 13319–13321.

50 B. P. MacLeod, F. G. Parlane, T. D. Morrissey, F. Häse, L. M. Roch, K. E. Dettelbach, R. Moreira, L. P. Yunker, M. B. Rooney, J. R. Deeth, *et al.*, *Sci. Adv.*, 2020, **6**, eaaz8867.

51 F. Häse, L. M. Roch and A. Aspuru-Guzik, *Trends Chem.*, 2019, **1**, 282–291.

52 B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, R. Clowes, *et al.*, *Nature*, 2020, **583**, 237–241.

53 R. Shimizu, S. Kobayashi, Y. Watanabe, Y. Ando and T. Hitosugi, *APL Mater.*, 2020, **8**, 111110.

54 C. W. Coley, N. S. Eyke and K. F. Jensen, *Angew. Chem., Int. Ed.*, 2020, **59**, 22858–22893.

55 A. Malek, M. J. Eslamibidgoli, M. Mokhtari, Q. Wang, M. H. Eikerling and K. Malek, *ChemPhysChem*, 2019, **20**, 2946–2955.

56 M. Ge, F. Su, Z. Zhao and D. Su, *Mater. Today Nano*, 2020, **11**, 100087.

57 J. R. Uijlings, K. E. Van De Sande, T. Gevers and A. W. Smeulders, *Int. J. Comput. Vis.*, 2013, **104**, 154–171.

58 E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan and Q. V. Le, 2018, arXiv preprint arXiv:1805.09501.

59 J. Hosang, R. Benenson and B. Schiele, 2014, arXiv:1406.6962.

60 A. Krizhevsky, I. Sutskever and G. E. Hinton, *Adv. Neural Inf. Process. Syst.*, 2012, **25**, 1097–1105.

61 Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu and M. S. Lew, *Neurocomputing*, 2016, **187**, 27–48.

62 K. Simonyan and A. Zisserman, 2014, arXiv preprint arXiv:1409.1556.

63 Y. Chen, Y. Li, T. Kong, L. Qi, R. Chu, L. Li and J. Jia, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 9563–9572.

64 R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

65 W. S. Rasband, *ImageJ*, US National Institutes of Health, Bethesda, Maryland, USA, 2011.

66 D. Krause, *Journal of Large-scale Research Facilities, JLSRF*, 2019, **5**, 135.