BIO Web of Conferences **129**, 10007 (2024) https://doi.org/10.1051/bioconf/202412910007
EMC 2024

Nanoparticle size estimation by HR-STEM and generative AI

Henrik Eliasson[1], Mr. Angus Lothian[2], Mr Ivan Surin[3], Dr. Sharon Mitchell[3], Professor Javier Pérez-Ramírez[3], Professor Rolf Erni[1]
[1]Electron Microscopy Center, Empa – Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland, [2]Computer Vision Laboratory, Department of Electrical Engineering, Linköping University, Linköping, Sweden, [3]Department of Chemistry and Applied Biosciences, ETH Zürich, Zürich, Switzerland

Background incl. aims

Scanning transmission electron microscopy (STEM) is commonly used to estimate the particle size distribution in heterogeneous catalysts. A problem with the conventional approach of manually measuring particle diameters in STEM images is that it is labor-intensive and disregards the actual particle structure, making particle size prediction with atomic precision impossible. To overcome these limitations, we propose a machine learning approach based on the CycleGAN architecture. The model learns to map between simulated and experimental images, and a subsequent network is trained to estimate the size of an imaged nanoparticle in terms of number of atoms. This technique is fully automatic and could form the basis for a modern characterization method when combined with automated data acquisition. As catalysts should ideally be studied under reaction conditions, we also explore the possibility of extending this workflow to gas-cell measurements.

Methods

Triggered by our recent experimental results, we have developed a structure generator which can generate realistic but random atomic models of Pt nanoparticles supported on ceria. These models have been used to generate a dataset of 3500 HAADF STEM multislice image simulations. A second dataset of high resolution HAADF STEM images of Pt nanoparticles supported on ceria was recorded on a probe-corrected Titan Themis operated at 300 keV. All experimental images were recorded under identical conditions. Images in both datasets are 128x128 with a pixel size in the range 10-50 pm. The simulated and experimental dataset were used to train a CycleGAN to map between the two sets. We use 2 U-Net generators and 2 PatchGAN discriminators for the CycleGAN. With the trained CycleGAN, we then send the simulated dataset through the "simulation-to-experimental" generator, thus obtaining a new dataset of 3500 fake experimental images with a corresponding particle size ground truth. Another U-Net with a fully connected layer at the end to output a scalar size prediction was then trained on the fake experimental dataset. The trained size-estimator was applied to the experimental dataset and the particle size distribution was compared with that of the manual technique of measuring the particle diameters and assuming a hemispherical particle shape.

Results

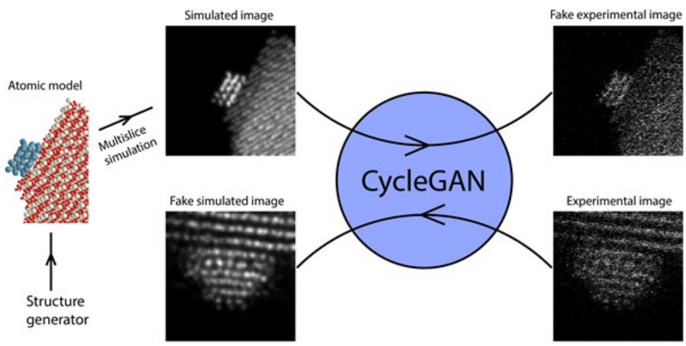
We have successfully trained a cycle-consistent generative adversarial network to map between physical image simulations and experimental HR-STEM images of small Pt nanoparticles supported on cerium dioxide. Passing raw multislice simulated images through the network yields realistic-looking experimental images with noise profile matching that of the real experimental dataset. In the same way, passing an experimental image through the network can be compared to applying a potent denoising algorithm to the image.

We apply our size estimation network to a set of 118 experimentally observed nanoparticles in a Pt/CeO2 catalyst to get the particle size distribution. Most predictions are reasonable and the particle size distribution obtained by the size-estimator network follows that of the manual hemisphere estimation technique. We find that the mean Pt particle size in the catalyst is 117 atoms with our technique, and 154 with the hemisphere estimation technique. Typical fail-cases of our model includes images of particles larger than 500 atoms and particles oriented along a major zone axis. This is likely due to a lack of large particles in the simulated data, as well as a lack of particles in zone axis in the experimental set. This could be addressed by generating more data.

Conclusion

In conclusion, our machine learning approach establishes a connection between physical image simulations and experimental images, making it possible to train supervised machine learning techniques on realistic-looking data with a known atomic model ground truth. We have showcased this by developing a size-estimation network which proves to output reasonable size estimates when applied to experimental images. Combined with automated data acquisition, we envision techniques like this will lead to more comprehensive and accurate characterization of nanoparticle-based heterogeneous catalysts.

Graphic:



Keywords:

Nanoparticles, Generative AI, Catalysis, Simulations

Reference:

Zhu et al., Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, ICCV, 2017.

Ronneberger et al., U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI, 2015.

Khan et al., npj Computational Materials 9, 85 (2023).

Eliasson et al., Nanoscale 15, 19091-19098 (2023).