Supporting Information for

Automated Image Analysis for Single-Atom Detection in Catalytic Materials by Transmission Electron Microscopy

Sharon Mitchell[¶]*, *Ferran Parés*[§], *Dario Faust Akl*[¶], *Sean M. Collins*[±]*, *Demie M. Kepaptsoglou*[¥♯], *Quentin M. Ramasse*[¥#], *Dario Garcia-Gasulla*[§]*, *Javier Pérez-Ramírez*[¶], *and Núria López*[Δ]

[¶] Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, Vladimir-Prelog-Weg 1, 8093 Zurich, Switzerland.

[§] Barcelona Supercomputing Center (BSC), Plaça d'Eusebi Güell 1-3, 08034 Barcelona, Spain.

[±] School of Chemical and Process Engineering and School of Chemistry, University of Leeds, Leeds, LS2 9JT, U.K.

[¥] SuperSTEM Laboratory, SciTech Daresbury Campus, Daresbury, WA4 4AD, U.K.

[♯] Department of Physics, University of York, Heslington, York, YO10 5DD, U.K.

[#] School of Chemical and Process Engineering and School of Physics, University of Leeds, Leeds, LS2 9JT, U.K.

[Δ] Institute of Chemical Research of Catalonia and The Barcelona Institute of Science and Technology, 43007 Tarragona, Spain.

Note S1. Choice of CNN architecture. Convolutional neural network (CNN) architectures have evolved at a rapid pace, and there are several well-established nets for image classification with differing levels of complexity. In our approach to develop a method for characterizing single-atom catalysts (SACs), we compared ResNet18,[1] a deep residual net with 18 layers that is popular benchmark for CNN models for detection problems, with a smaller custom architecture containing fewer layers (SAC-CNN, see **Table S1**). The custom architecture was developed to provide a better fit between network capacity (i.e., the number of parameters) and input dimensionality. In terms of classification performance, ResNet18 architecture obtained a slightly worse accuracy compared to SAC-CNN. However, more significantly, we found that ResNet18 overfits much more than SAC-CNN during training (**Figure S4**), which impacts its proper generalization.

Note S2. Choice of CNN architecture. Computer vision techniques have been widely applied to extract features from grey-level images (e.g., photographs, electron micrographs). The most common approach consists of applying a two-dimensional Gaussian fitting algorithm to the image to detect blob-like image structures.[2,3] For comparison, our proposed deep-learning method was compared to a traditional computer vision (CV) model using a similar protocol. The same procedure was used to generate image crops of equivalent size (512×512 px) from manually-tagged images. The Gaussian kernel is adjusted to fit the approximate atom size and improve detection when filtering through the image, with an optimal value of $\sigma = 0.22$ and no offset. After kernel-filtering, an optimized threshold is applied and a labeling technique is used to obtain the final localization of the predicted atoms. These steps follow the same procedure of the inference pipeline in the deep learning technique after generation of the likelihood landscape, albeit using different parameters (e.g., threshold values).

Table S1. Characteristics of the proposed SAC-CNN for single-atom detection in Pt_1/NC.

Layer	Type	Neurons	Output tensor	Parameters
1	Convolution 2D	32	(*x*,*x*,32)	288
2	Batch normalization 2D	32	(*x*,*x*,32)	64
3	Convolution 2D	64	(*x*,*x*,64)	18432
4	Batch normalization 2D	64	(*x*,*x*,64)	128
5	Convolution 2D	128	(*x*,*x*,128)	73728
6	Batch normalization 2D	128	(*x*,*x*,128)	256
7	Adaptive pooling 2D	-	(3,3,128)	0
8	Fully-connected	128	(128,)	147456
9	Fully-connected	128	(128,)	16384
10	Fully-connected	2	(2,)	256
	Total			**427840**

Table S2. Performance of SAC-CNN for single-atom detection in Pt_1/NC.

Image	Precision	Recall	F1
Test 1	0.615	0.706	0.658
Test 2	0.660	0.700	0.680
Test 3	0.614	0.623	0.619
Test 4	0.684	0.772	0.726
Test 5	0.582	0.639	0.609
Test 6	0.676	0.814	0.739
Test 7	0.574	0.660	0.614
Test 8	0.644	0.644	0.644
Test 9	0.542	0.957	0.692
Test 10	0.695	0.759	0.726
Mean	0.629	0.728	0.671
STD	0.052	0.103	0.049

Table S3. Performance of standard CV approach for single-atom detection in Pt_1/NC.

Image	Precision	Recall	F1
Test 1	0.581	0.735	0.649
Test 2	0.619	0.780	0.690
Test 3	0.506	0.609	0.553
Test 4	0.620	0.842	0.714
Test 5	0.403	0.787	0.533
Test 6	0.524	0.729	0.610
Test 7	0.465	0.623	0.532
Test 8	0.604	0.711	0.653
Test 9	0.458	0.936	0.615
Test 10	0.695	0.722	0.639
Mean	0.536	0.747	0.619
STD	0.076	0.097	0.063

Table S4. Performance of SAC-CNN for single-atom detection in Pt_1/NC after training with half of the training set.

Image	Precision	Recall	F1
Test 1	0.696	0.471	0.561
Test 2	0.758	0.500	0.602
Test 3	0.756	0.449	0.564
Test 4	0.766	0.356	0.487
Test 5	0.773	0.279	0.410
Test 6	0.714	0.509	0.594
Test 7	0.875	0.396	0.546
Test 8	0.813	0.289	0.426
Test 9	0.706	0.766	0.735
Test 10	0.786	0.407	0.537
Mean	0.764	0.442	0.546
STD	0.054	0.139	0.093

Table S5. Performance of AtomSegNet for single-atom detection in Pt_1/NC.

Image	Precision	Recall	F1
Test 1	0.282	0.588	0.381
Test 2	0.277	0.780	0.408
Test 3	0.287	0.565	0.381
Test 4	0.442	0.644	0.524
Test 5	0.278	0.689	0.396
Test 6	0.373	0.695	0.485
Test 7	0.308	0.679	0.424
Test 8	0.291	0.711	0.413
Test 9	0.353	0.872	0.503
Test 10	0.301	0.685	0.418
Mean	0.319	0.691	0.433
STD	0.051	0.084	0.049

Table S6. Performance of SAC-CNN for single-atom detection in Fe_1/ECN (threshold = 0.97).

Image	Precision	Recall	F1
1	0.558	0.795	0.655
2	0.544	0.623	0.581
3	0.875	0.574	0.693
Mean	0.659	0.664	0.643
STD	0.153	0.095	0.047

Table S7. Performance of CV for single-atom detection in Fe_1/ECN (threshold = 0.21).

Image	Precision	Recall	F1
1	0.514	0.781	0.620
2	0.448	0.565	0.500
3	0.814	0.574	0.673
Mean	0.592	0.640	0.598
STD	0.159	0.100	0.072

Table S8. Performance of AtomSegNet for single-atom detection in Fe_1/ECN.

Image	Precision	Recall	F1
1	0.255	0.644	0.366
2	0.193	0.594	0.291
3	0.474	0.607	0.532
Mean	0.307	0.615	0.396
STD	0.121	0.021	0.101

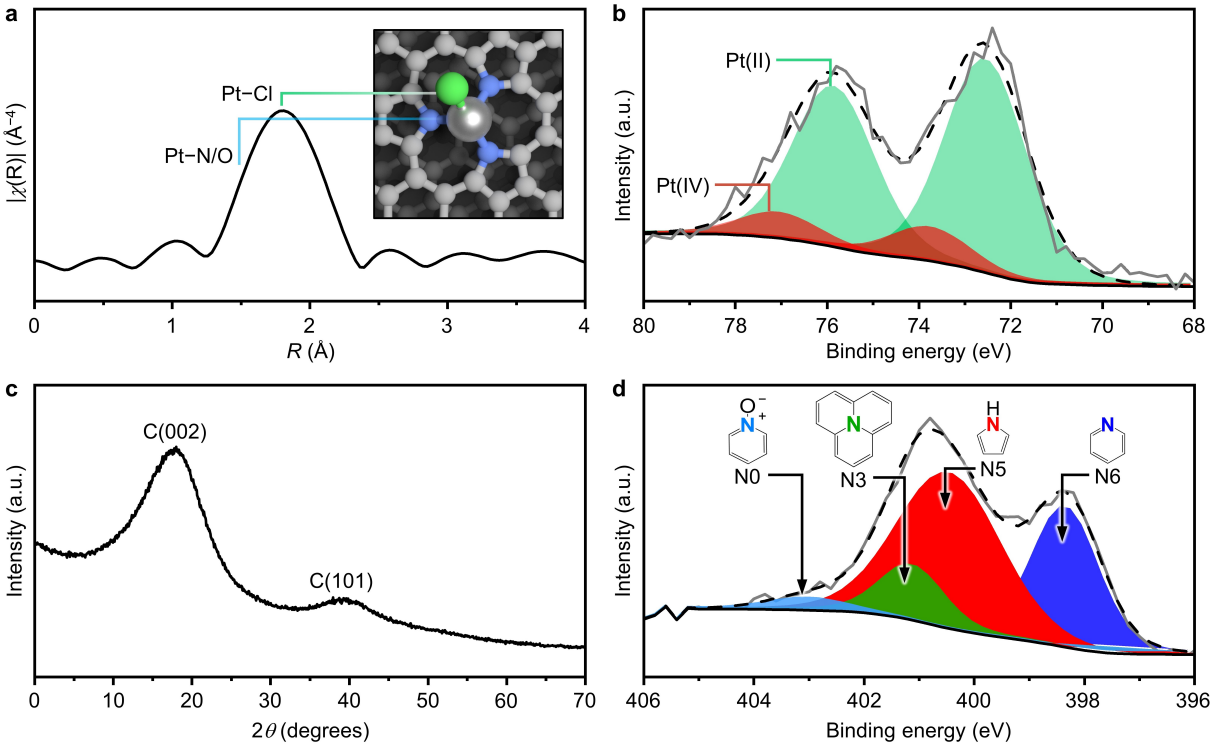

Figure S1. Characterization of the Pt_1/NC single-atom catalyst by bulk spectroscopies. (a) Fourier-transformed extended X-ray absorption spectrum evidencing the predominance of Pt-N/O and Pt-Cl contributions. A tentative metal site structure is shown inset. (b) Pt 4*f* photoelectron spectrum with fitted contributions assigned to Pt(II) and Pt(IV). (c) Powder X-ray diffraction pattern evidencing broad reflections consistent with the long-range graphitic structure of the nitrogen-doped carbon host. (d) The N 1*s* photoelectron spectrum fitted with components assigned to pyridine-N-oxide (N0), graphitic nitrogen (N3), pyrrolic (N5) and pyridinic (N6) groups. The results fully agree with previously reported characterization data for this catalyst.[4]

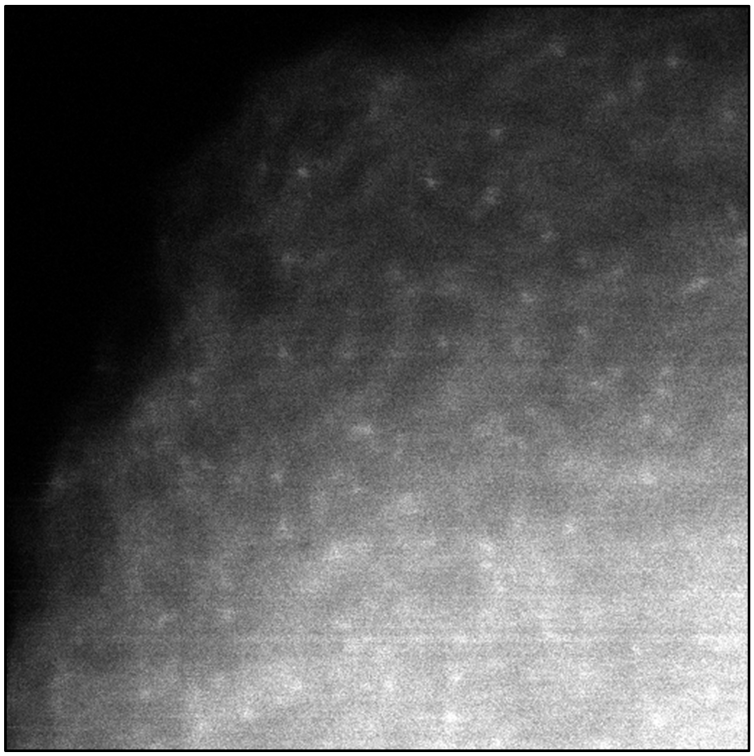

Figure S2. Typical unprocessed AC-STEM image of the Pt_1/NC single-atom catalyst. The image corresponds to 15×15 nm.

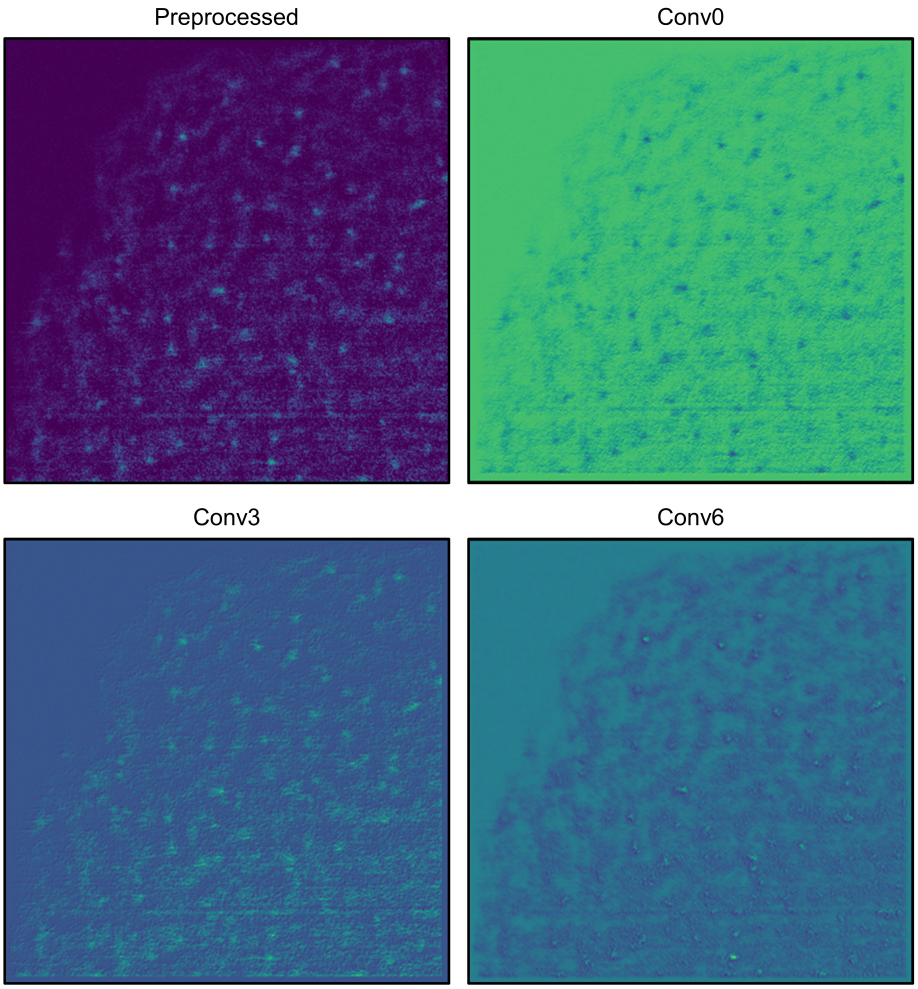

Figure S3. Visualizations of the outputs of the three convolutional layers of the SAC-CNN model applied to a typical preprocessed AC-STEM image (15×15 nm) of Pt_1/NC. The first convolutional layer (Conv0) tends to learn simple visual patterns which activate on pixel contrast (typical of image noise) but not on homogeneous blobs, characteristic of catalyst atomic features. The next convolutional layer (Conv3), which works at a higher visual scale, already focuses on the areas containing potential atoms, significantly reducing the relevance of background noise. Finally, the last convolutional layer (Conv6) produces an activation map with highly probable and well bounded atom locations. This map is used by the following fully-connected layers in the SAC-CNN network to make the final prediction regarding the presence of atoms.

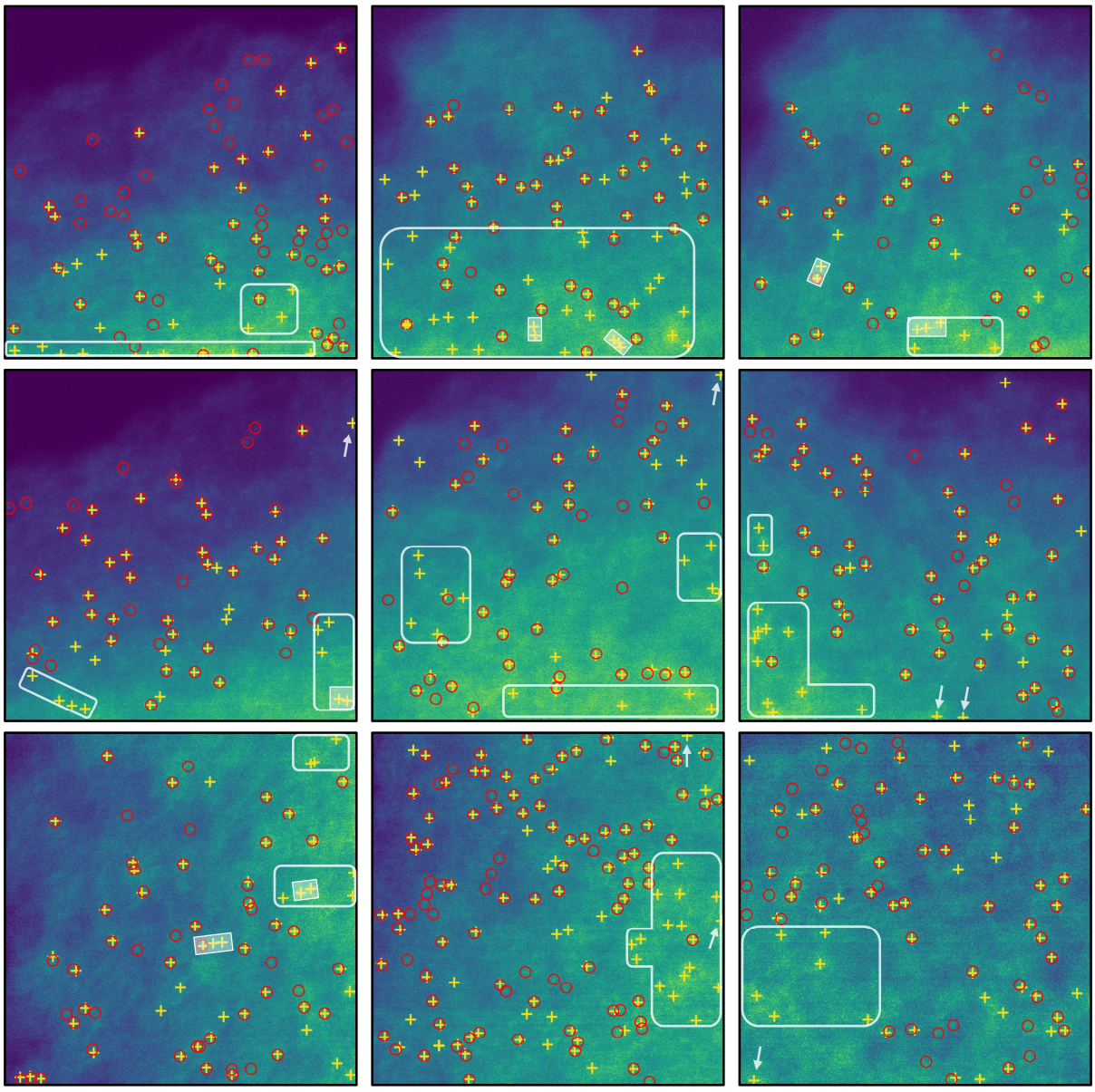

Figure S4. Performance of the SAC-CNN model. Comparison of the manually-labeled (red circles) and SAC-CNN-detected (yellow crosses) single atoms over the validation set. The automated method shows consistently improved performance in the brightest image areas (white boxes). It also detects atoms close to the image edges (white arrows) and to other atoms (shaded boxes). All images correspond to 15×15 nm.

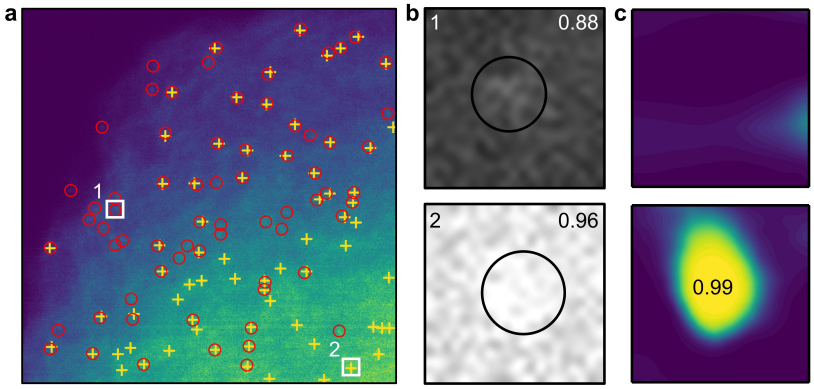

Figure S5. Perception of image contrast. Comparison of the contrast of two regions in the Aberration-corrected scanning transmission electron microscopy (AC-STEM) image (15×15 nm) of the platinum single-atom catalyst 1) where a manually-tagged feature was not detected by the SAC-CNN and 2) vice versa. The zooms (21×21 px, corresponding to the window size of the model) of these regions highlight the challenge of identifying single atoms by simple visual inspection (contrast-to-background ratios of the circled features are indicated). In contrast, the corresponding likelihood maps (indicating the associated probability) generated by the SAC-CNN clearly distinguish differences in the intensity patterns.

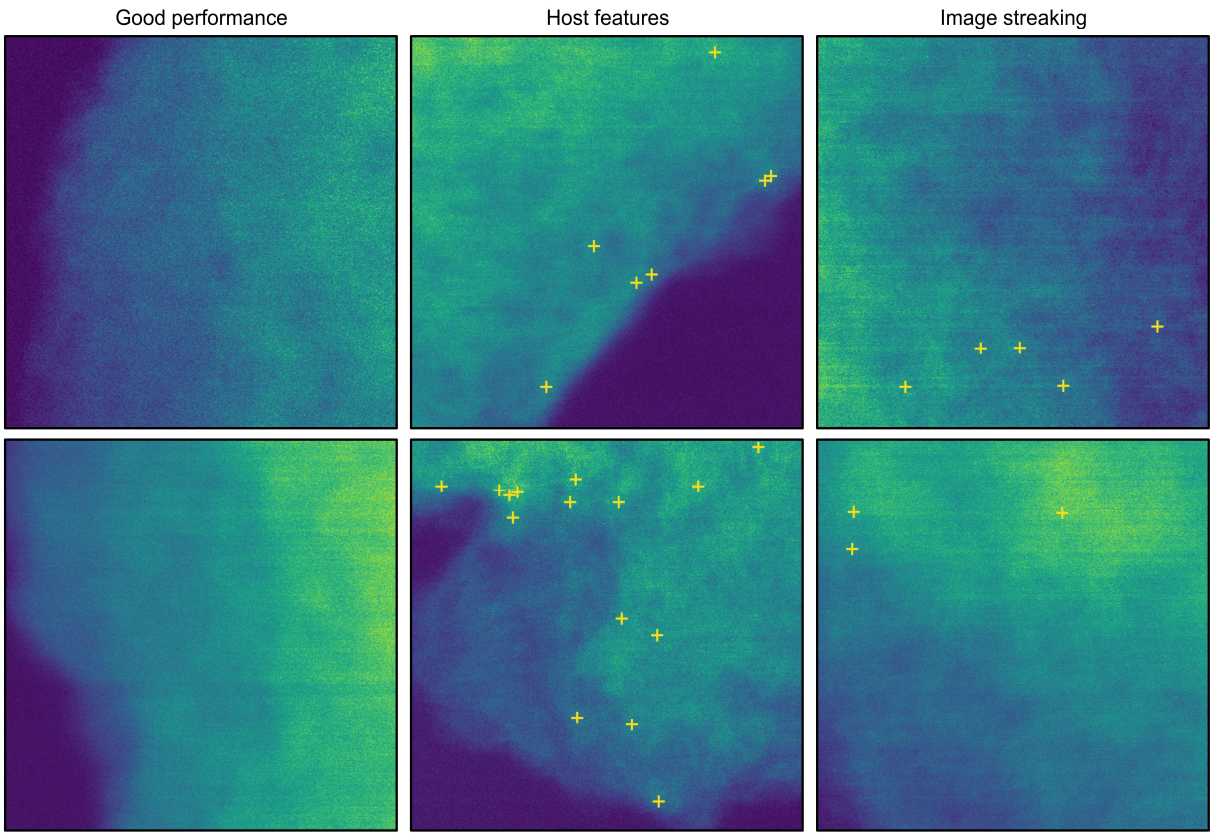

Figure S6. Performance of SAC-CNN model on AC-STEM images of the metal-free nitrogen-doped carbon support. In general, few atoms were detected upon analysis of AC-STEM images of the metal-free carrier with the SAC-CNN (84% of images contained <5 detections), confirming the robustness of the approach. As illustrated for representative examples, when detection occurred it typically resulted from streaking in the image (identified by the occurrence of multiple atom detections on the same horizontal line) or structural variations in the carrier. All images correspond

to 15×15 nm (34.3 px nm^{-1}).

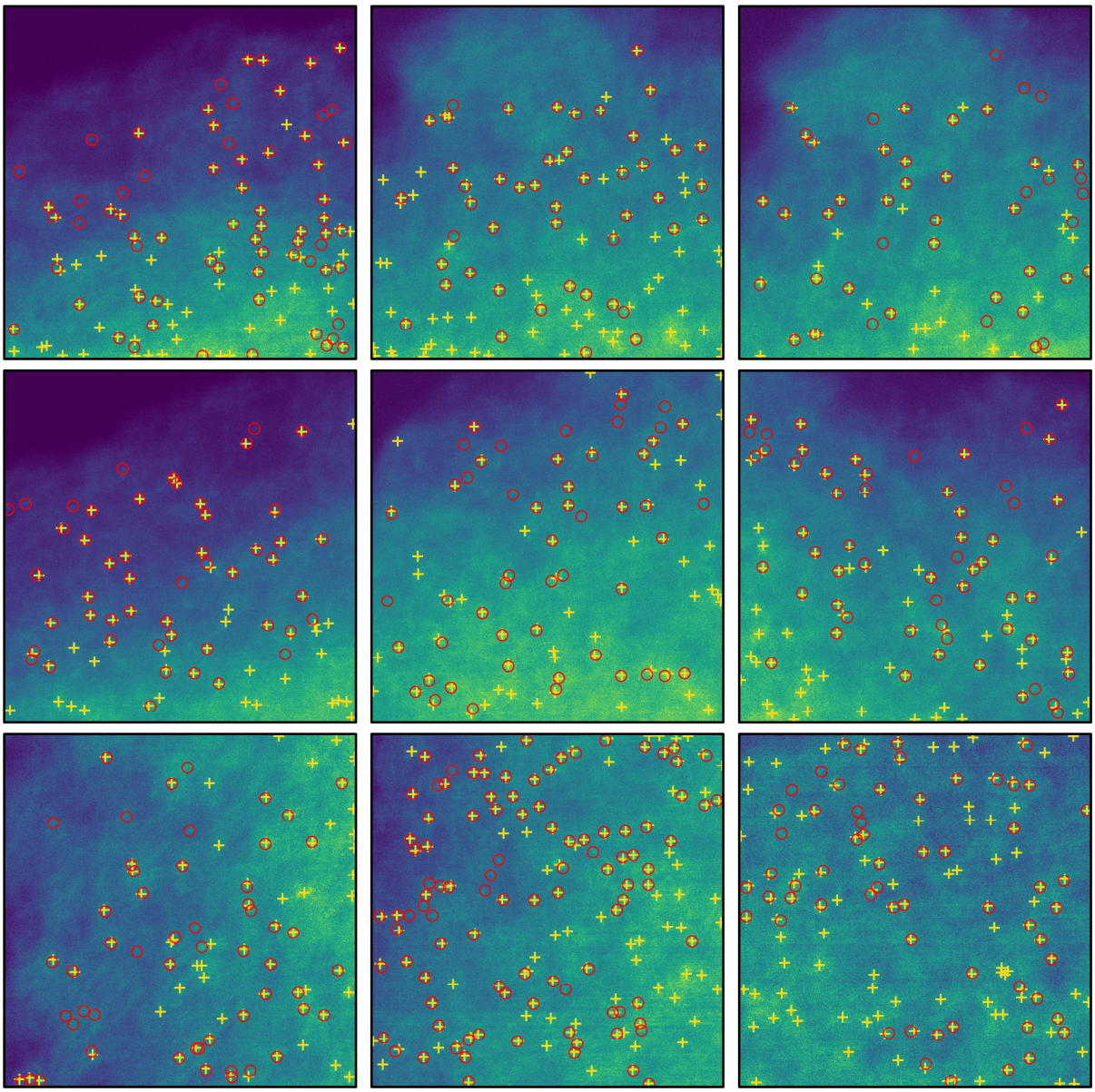

Figure S7. Performance of CV when applied to the validation set of AC-STEM images of Pt_1/NC. The CV method typically results in a slightly higher recall but lower precision than the SAC-CNN model (corresponding images in **Figure S2**) because in general it detects more atoms per image. All images correspond to 15×15 nm.

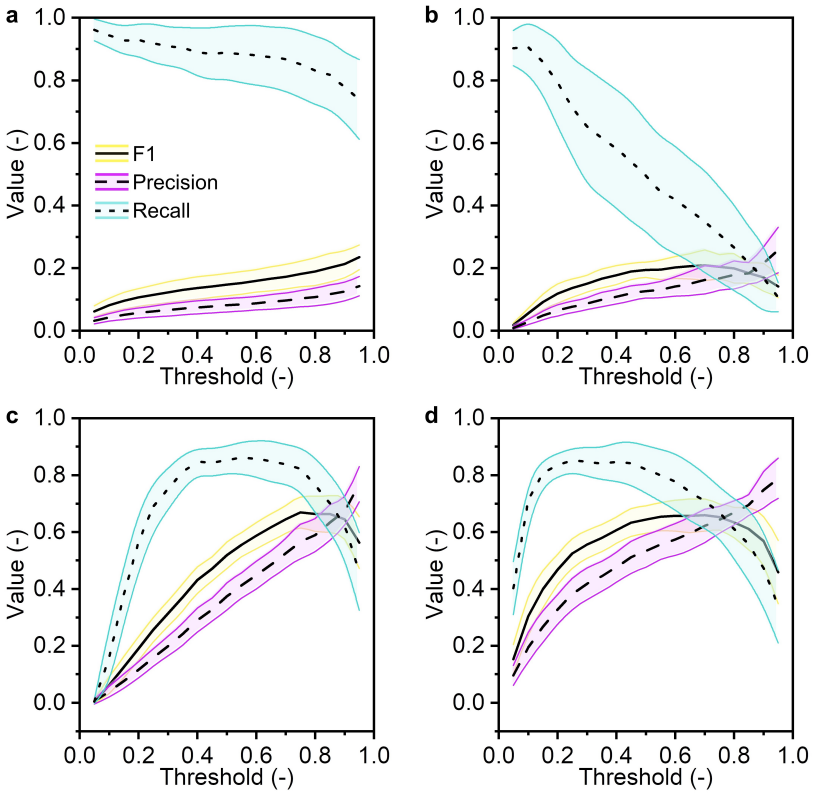

Figure S8. Comparison of the F1, precision, and recall scores (shaded areas indicate the associated standard deviation) as a function of likelihood thresholds for the ResNet18 architecture trained (a) under the same conditions of our SAC-CNN model or (b) by optimizing the number of training epochs. Significantly inferior F1 scores compared to the customized model were observed in both cases, with notably lower precision at high thresholds and rapidly dropping recall in the case of (b). Performance of the SAC-CNN architecture with modified training conditions (c) reducing the number of training epochs to 12 or (d) including negative crops from images of the metal-free sample. Neither change significantly altered the performance.

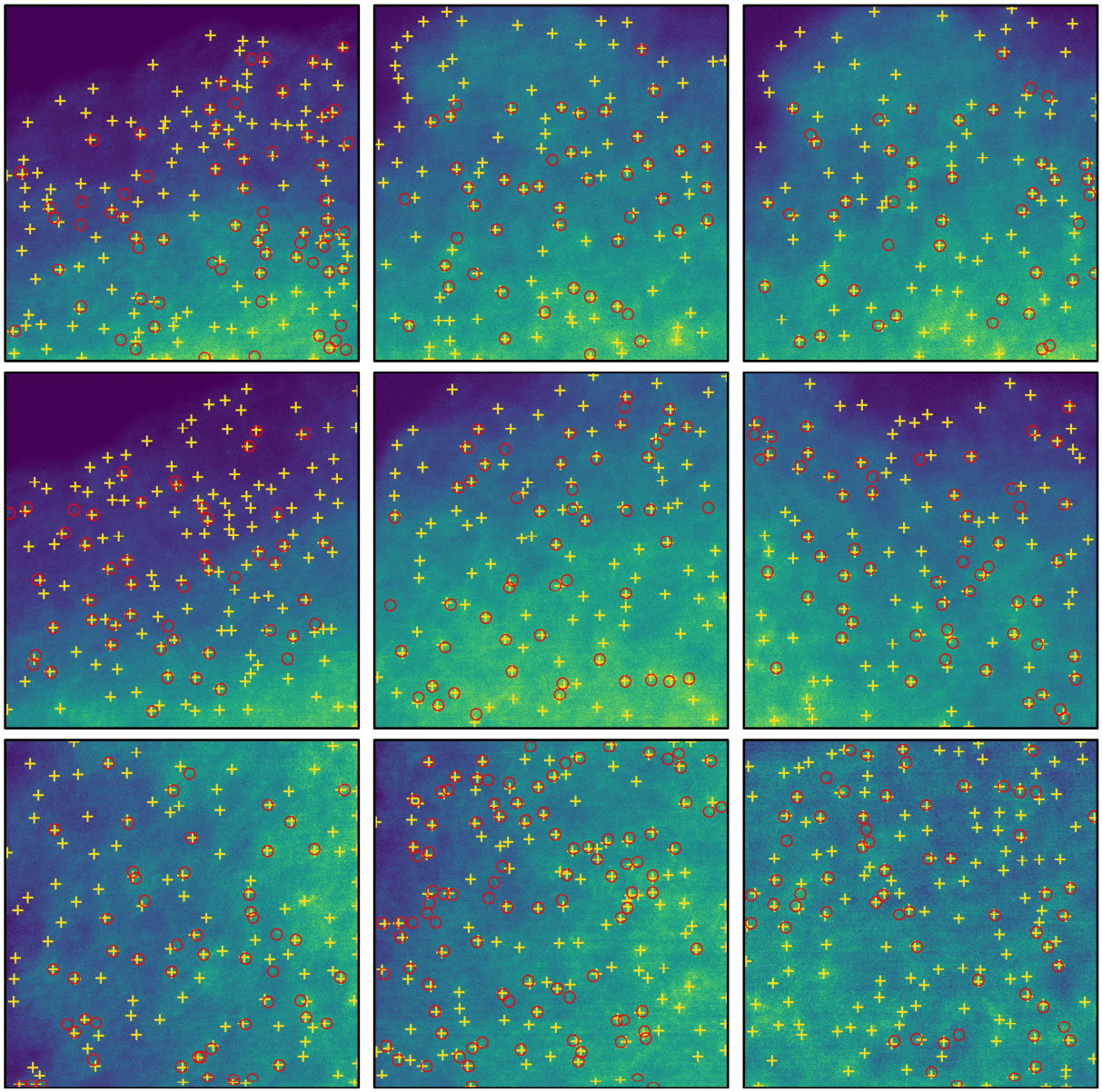

Figure S9. Performance of AtomSegNet when applied to the validation set without additional training or specific optimization. A much higher number of detections (yellow crosses) compared to the SAC-CNN model (corresponding images in **Figure S2**) is evidenced across the sample. While this includes many of the manually-labeled atoms (red circles) there also appear to be a number of false positive detections. All images correspond to 15×15 nm.

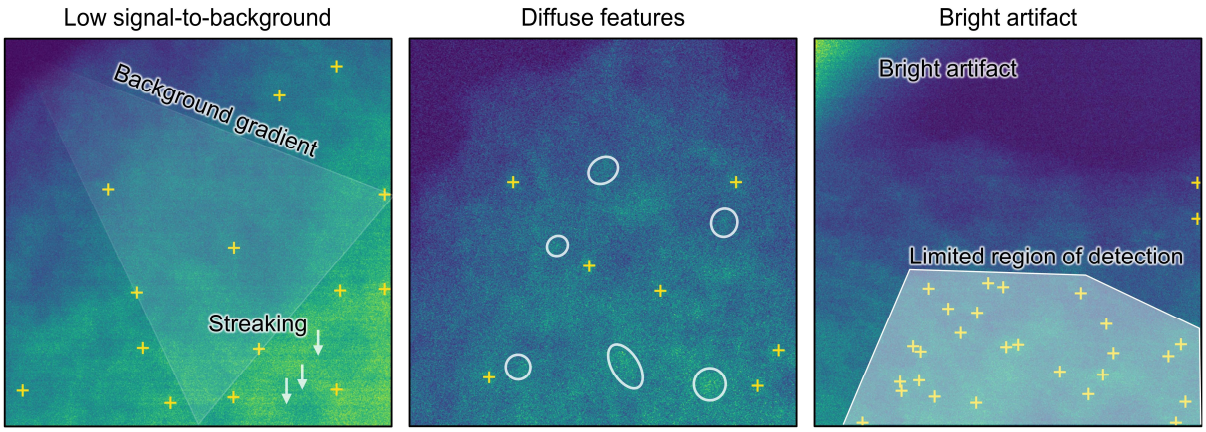

Figure S10. Analysis of outliers in the full data set. The images identified as outliers in **Figure 3** display significant deviations from the expected nearest-neighbor distances. Inspection of the images reveals that many atomic positions were not recognized due to flattening of contrast (e.g., in thick regions of the sample) or increased feature size (e.g., through defocus), resulting in higher values of r_{Measured} than those expected for an analogous random arrangement of atoms. Bright artifacts in the image (such as those arising from nearby thick and out-of-focus sample regions) can also lead to excessive flattening of the contrast during pre-processing, resulting in a bias towards atom detection in only delimited areas of the whole sample. All images correspond to 15×15 nm (34.3 px nm^{-1}).

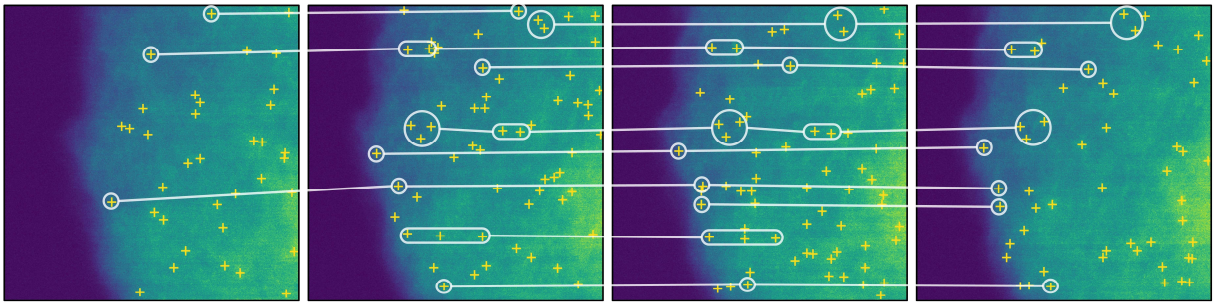

Figure S11. The atomic positions detected by the SAC-CNN model on consecutively-acquired identical location frames of the platinum single-atom catalyst. A subset of the atoms remains constant in each of the images (see connected outlines), while others change location. All images correspond to 15×15 nm (34.3 px nm^{-1}).

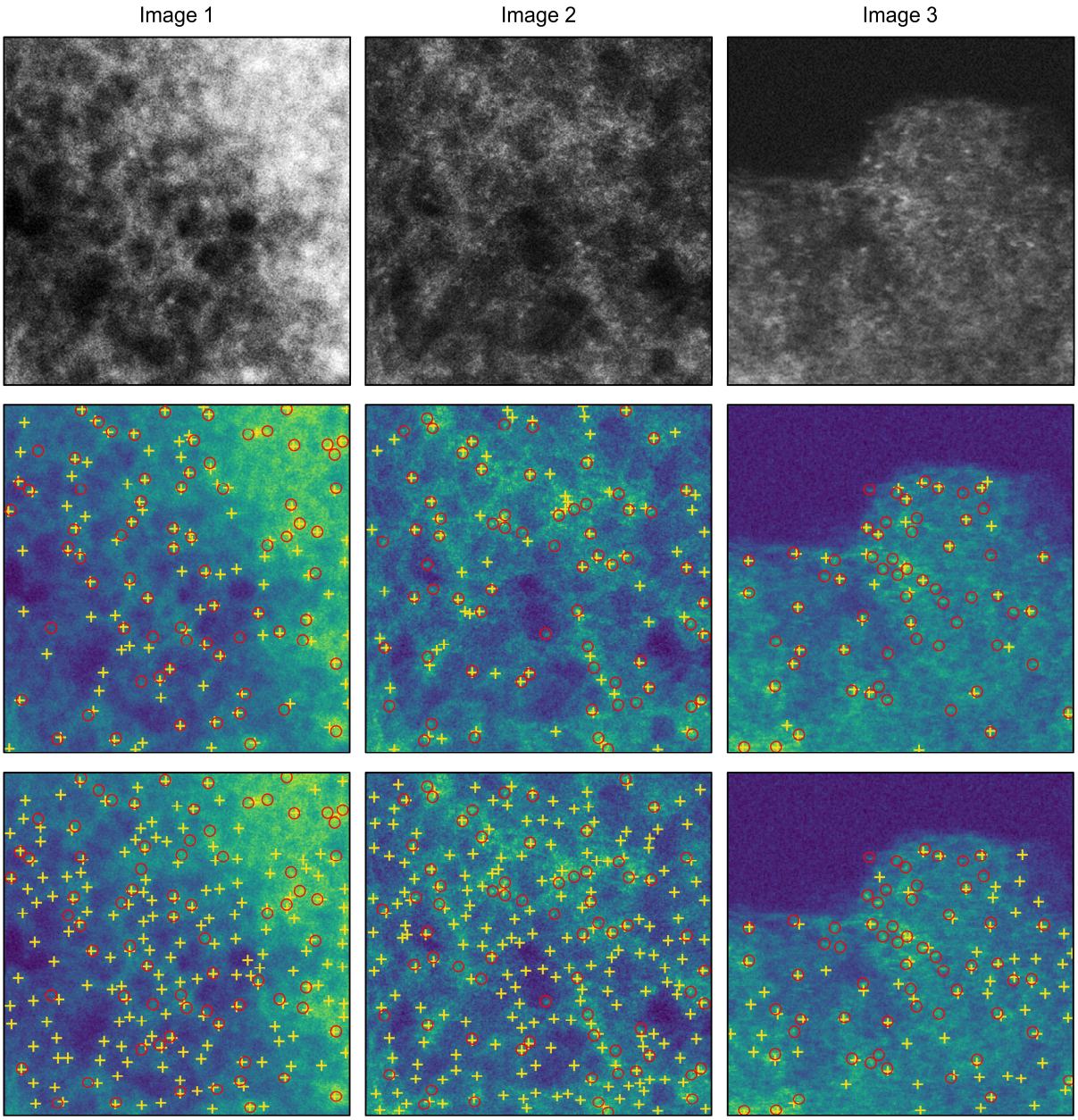

Figure S12. Unprocessed AC-STEM images of the Fe_1/C_3N_4 single-atom catalyst (top row) and output images showing the atom detections (yellow crosses) determined by CV (middle row) and AtomSegNet (bottom row). Red circles denote the manually labeled atom positions. Images 1 and 2 correspond to 10×10 nm (34.1 px nm^{-1}) and Image 3 is 7.44×7.44 nm (34.4 px nm^{-1}).

Supplementary References

(1) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition in Proceedings of the IEEE conference on computer vision and pattern recognition. **2016**, 770–778.

(2) Lindeberg, T. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *Int. J. Comp. Vis.* **1993**, *11*, 283–318.

(3) Lindeberg, T. Image matching using generalized scale-space interest points. *J. Math. Imaging Vis.* **2015**, *52*, 3–36.

(4) Kaiser, S. K.; Fako, E.; Manzocchi, G.; Krumeich, F.; Hauert, R.; Clark, A. H.; Safonova, O. V.; López, N.; Pérez-Ramírez, J. Nanostructuring unlocks high performance of platinum single-atom catalysts for stable vinyl chloride production. *Nat. Catal.* **2020**, *3*, 376–385.