Supporting information

Deep neural network analysis of nanoparticle ordering to identify defects in layered carbon materials

Daniil A. Boiko
Evgeniy O. Pentsak
Vera A. Cherepanova
Evgeniy G. Gordeev
Valentine P. Ananikov*

Contents

Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky Pr. 47, Moscow 119991, Russia; val@ioc.ac.ru

Role of type, amount and distribution of defects

Table S1. References for importance of determination of type, amount and distribution of defects.

Area	**Role of defects**	**Defining characteristic of defects**	**Ref.**
Electronic devices	tune the bandgap, electronic properties, and transport properties	density and distribution of defects	1
	define electronic and mechanical properties	characteristics of grain boundaries	2–4
	define transport ability in the vicinity of the Fermi level	presence localized states introduced by grain boundaries	5
	define mechanical strength	location, number and size of grain boundaries and void defects	6–8
Carbocatalysts	carbocatalysis of hydrogenation	type of defects (Stone-Wales defects)	9
	carbocatalysis of oxidative coupling reaction	type of defects (structural defects)	10
	carbocatalysis of reductive hydrogen atom transfer reactions	type of defects (localized π-edge states)	11
Supported catalysts	reduce leaching	binding strength	12
	prevention of agglomeration	size of the defect, binding strength	12
	provide charge transfer between metal and support	binding strength	12
	define stability of carbon supports	amount, distribution and type of defects	13

Monte-Carlo simulations: non-uniform surface case

Fig. S1. Mask of reactive areas

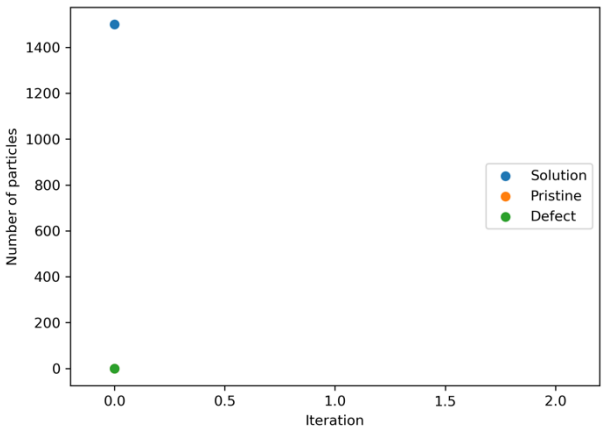
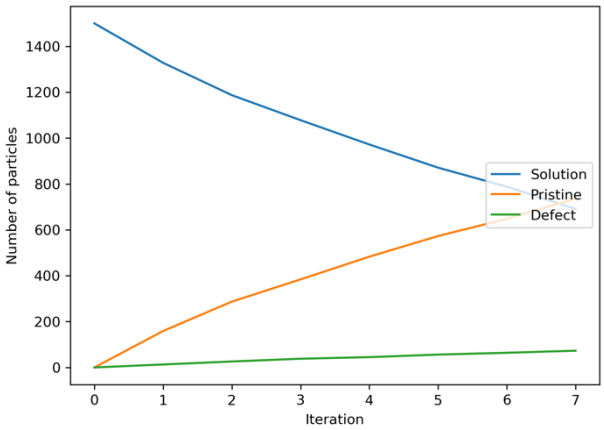


Fig. S2. The initial state of the system



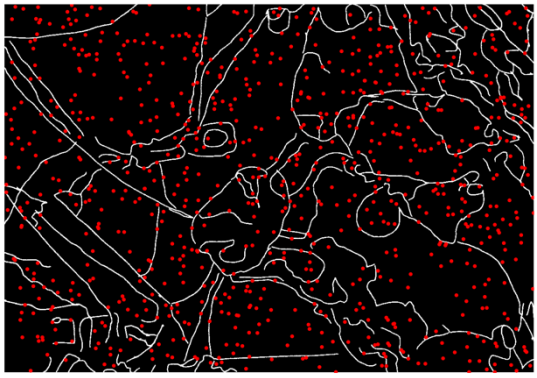
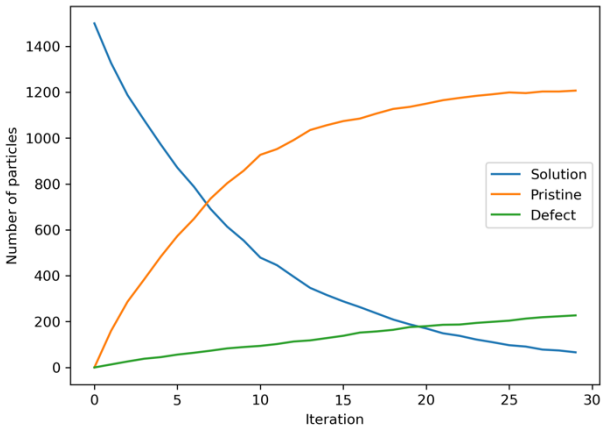


Fig. S3. The simulation after 7 iterations



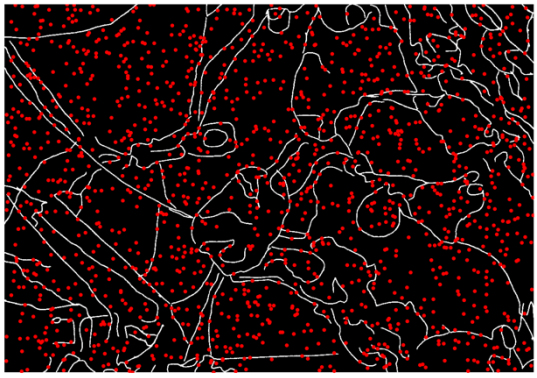
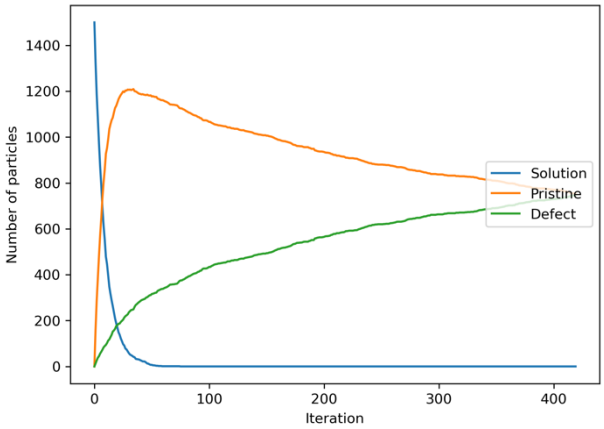


Fig. S4. The simulation after 29 iterations



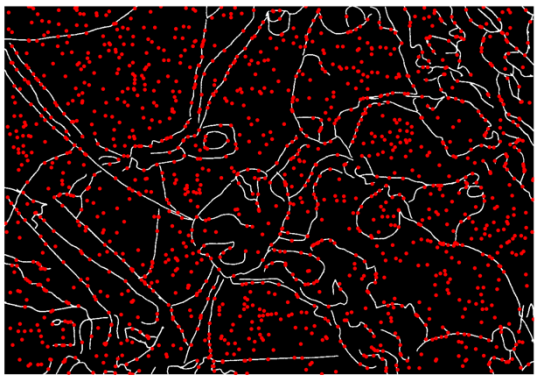
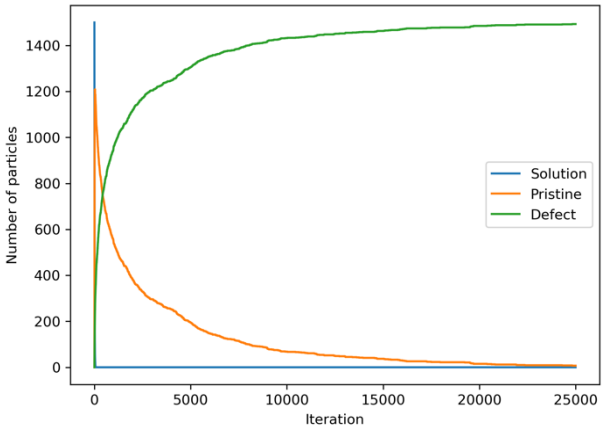


Fig. S5. The simulation after 419 iterations



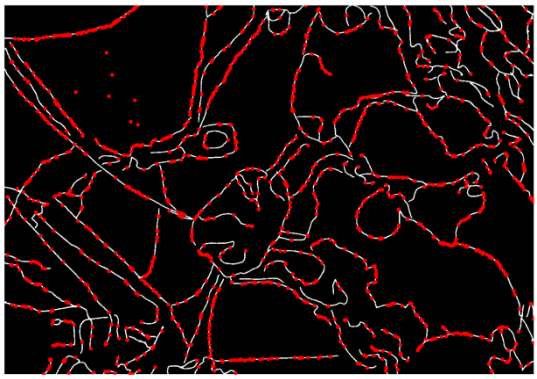


Fig. S6. The simulation after 25 000 iterations

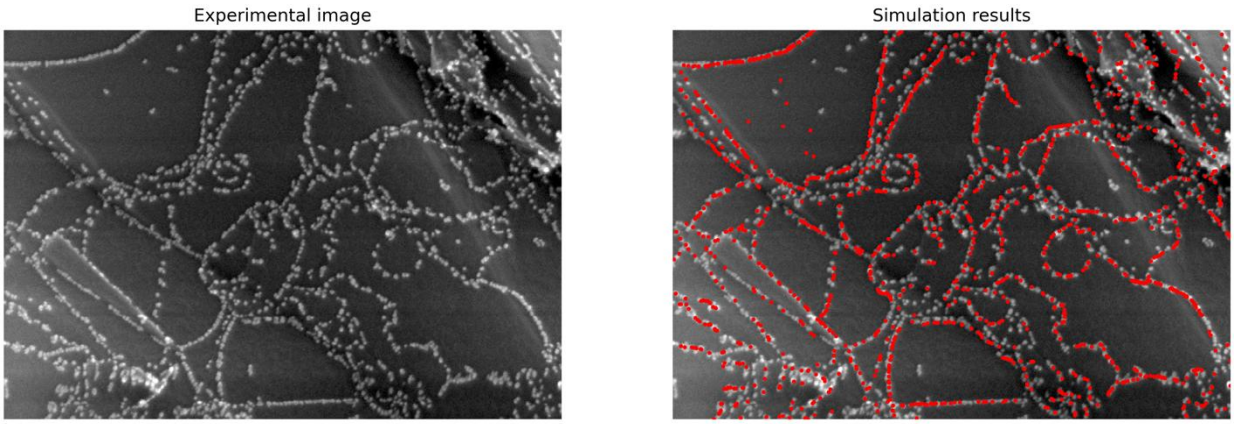


Fig. S7. The comparison between experimental image and results of the simulation

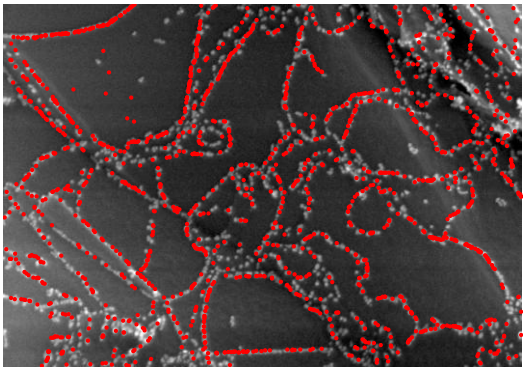


Fig. S8. Experimental image with overlaid simulation results

Video movie reflecting the simulation process is available in the Movie-1.mp4 file.

Monte-Carlo simulations: smooth surface case

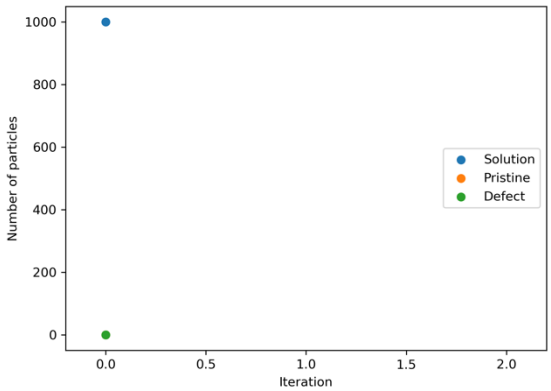
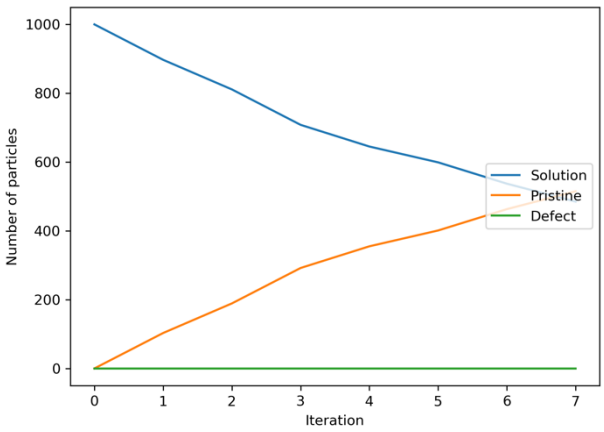


Fig. S9. The initial state of the system



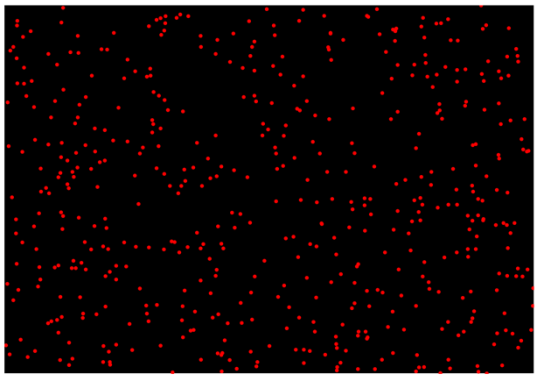
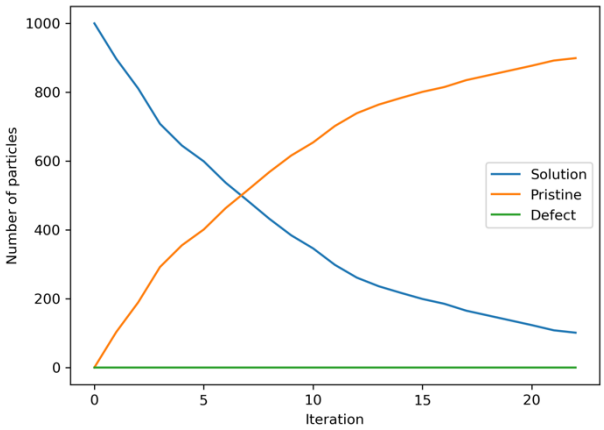


Fig. S10. The simulation after 7 iterations



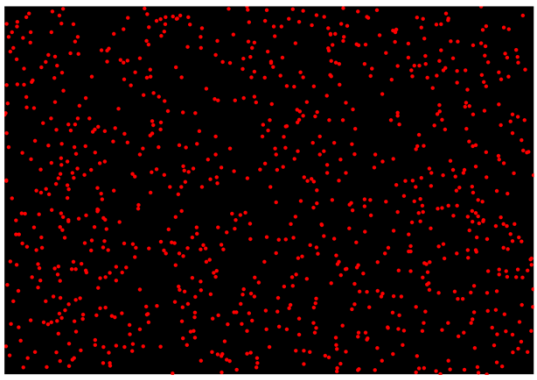
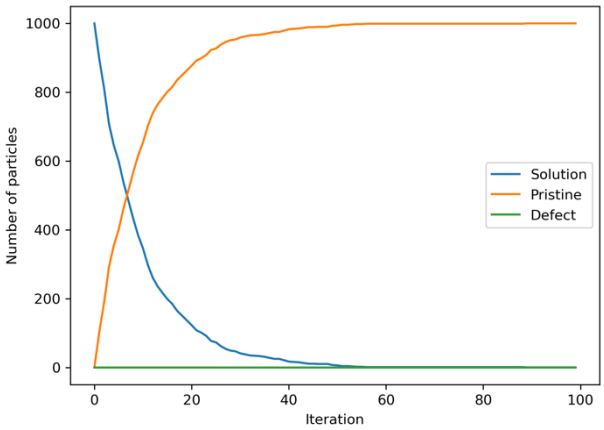


Fig. S11. The simulation after 22 iterations



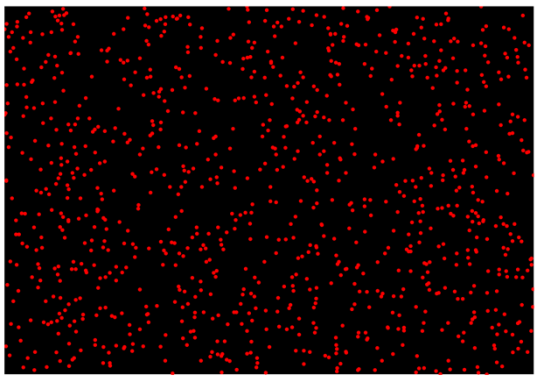


Fig. S12. The simulation after 100 iterations

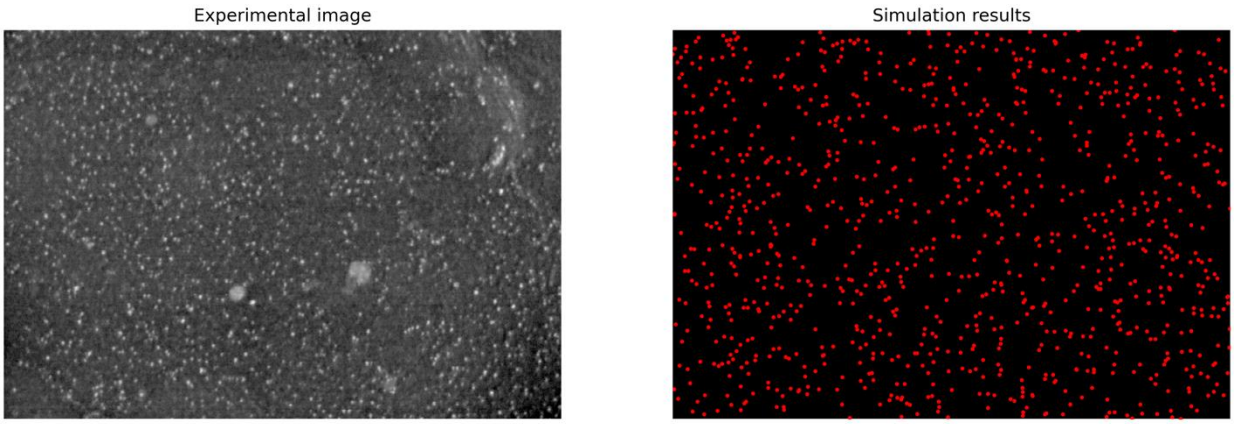


Fig. S13. The comparison between experimental image and results of the simulation

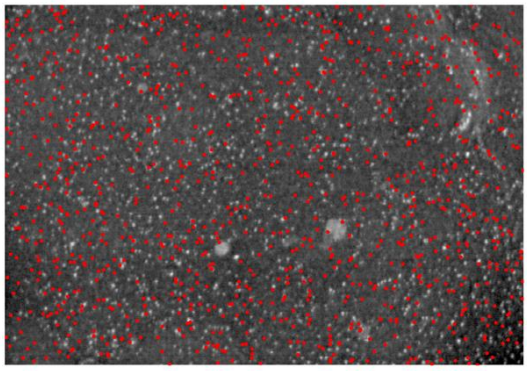


Fig. S14. Experimental image with overlaid simulation results

Video movie reflecting the simulation process is available in the Movie-2.mp4 file.

Degree of order determination

The degree of order can be determined by finding particle positions first and then checking whether these positions were drawn from 2D uniform distribution. This was done using one-sample two-dimensional Kolmogorov-Smirnov test.[14]

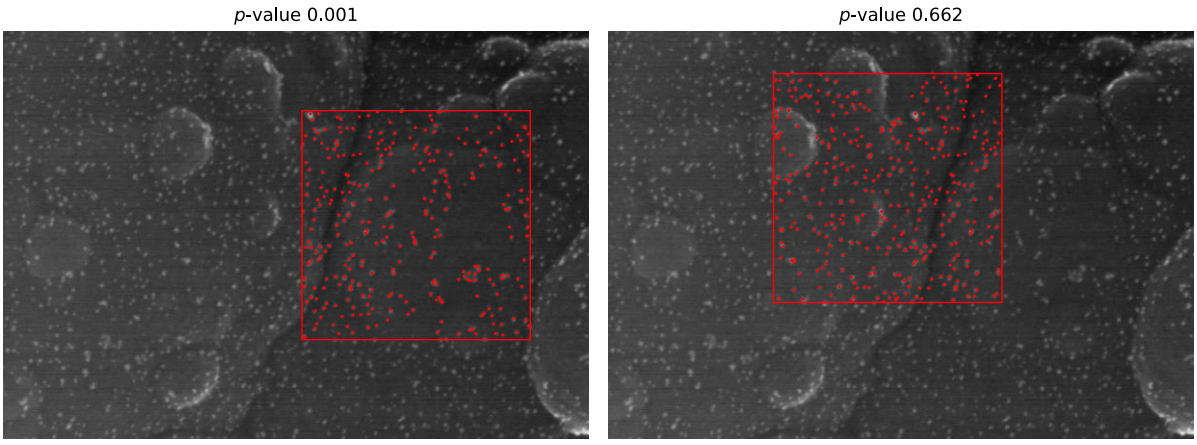


Fig. S15. Two patches of the image of the disordered material showing different degree of order

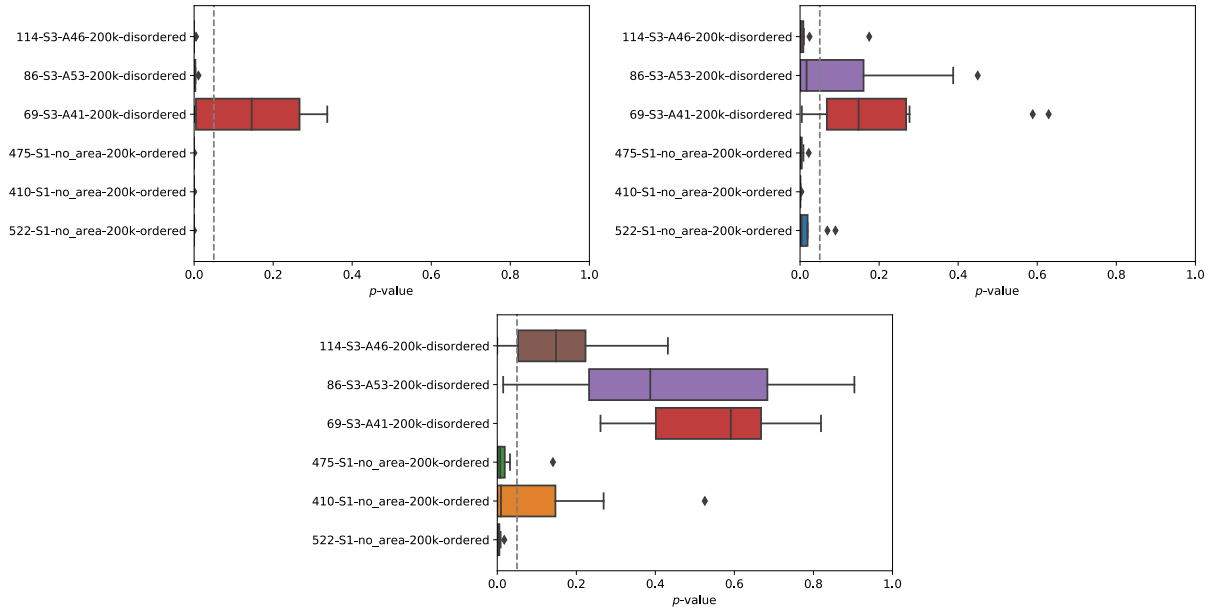


Fig. S16. Distribution of *p*-values for 10 sample patches in multiple ordered and disordered images. Top left 750 px, top right 500 px, bottom 250 px. Image names are the same as in previously published dataset[15]

Large variation of the results confirms need for more robust methods for determination of degree of order.

Classification neural network

Training

Model was trained on 227x227 patches of the images provided in the dataset.[16]

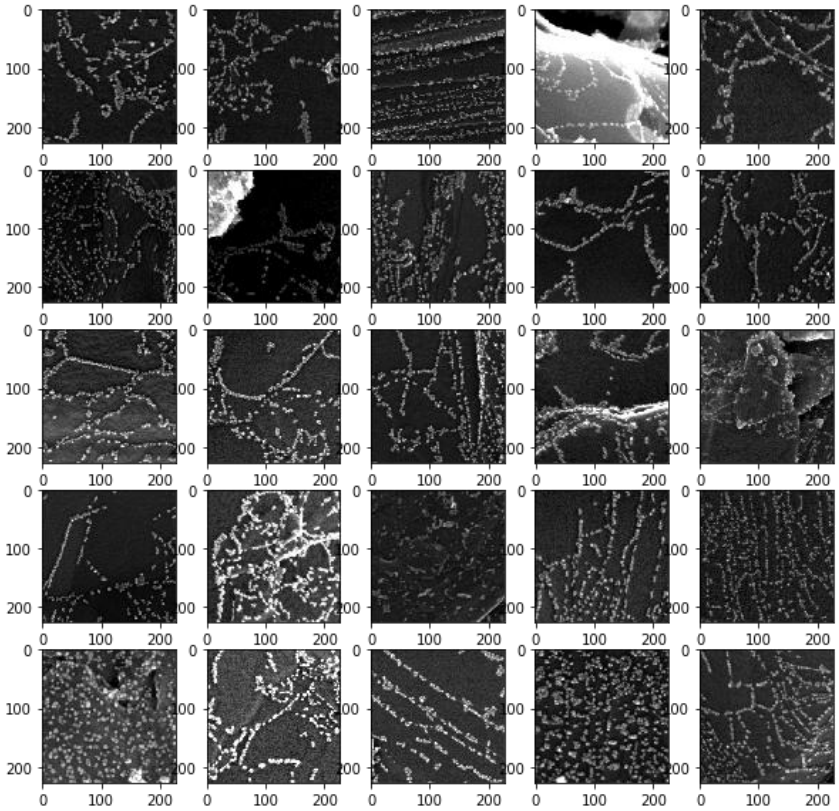


Fig. S17. Example of training images

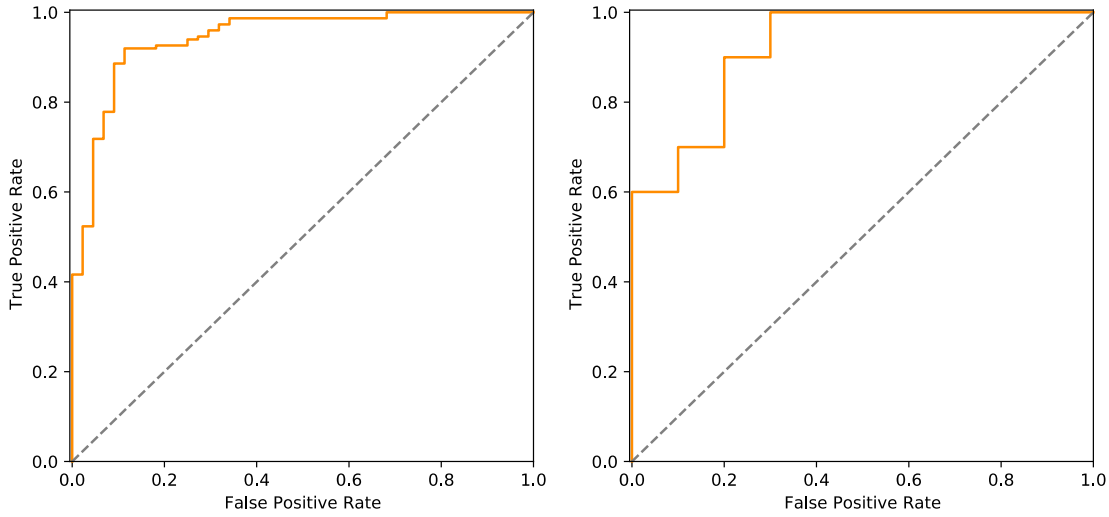


Fig. S18. ROC-curves for AlexNet model (left — whole validation dataset, right — human evaluation dataset)

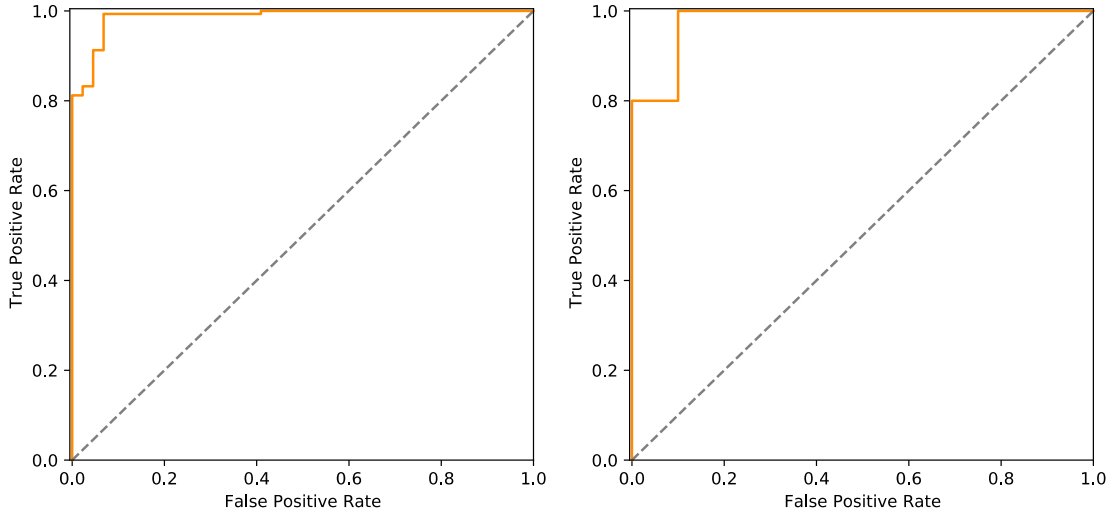


Fig. S19. ROC-curves for ResNet model (left — whole validation dataset, right — human evaluation dataset)

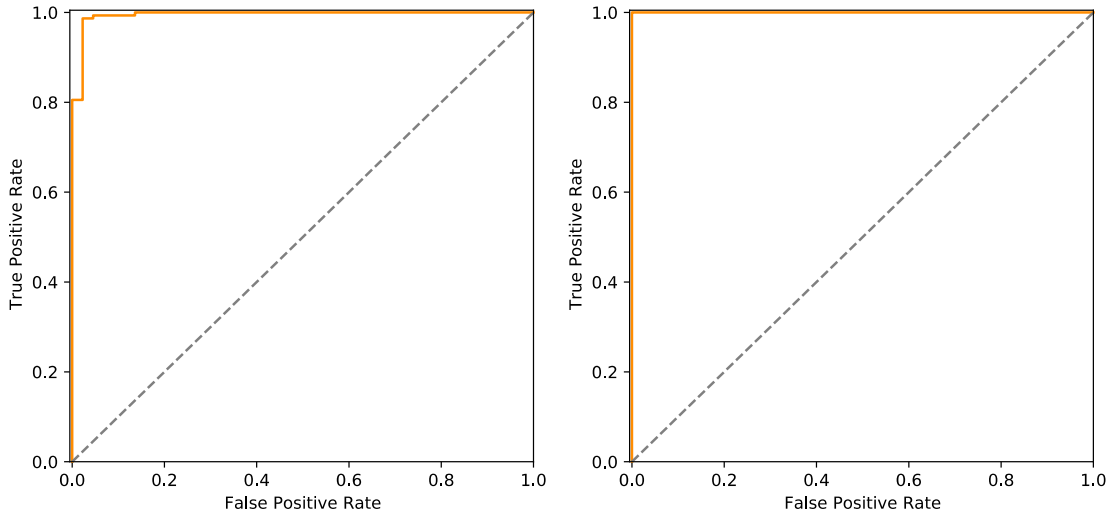


Fig. S20. ROC-curves for VGG model (left — whole validation dataset, right — human evaluation dataset)

Model interpretation

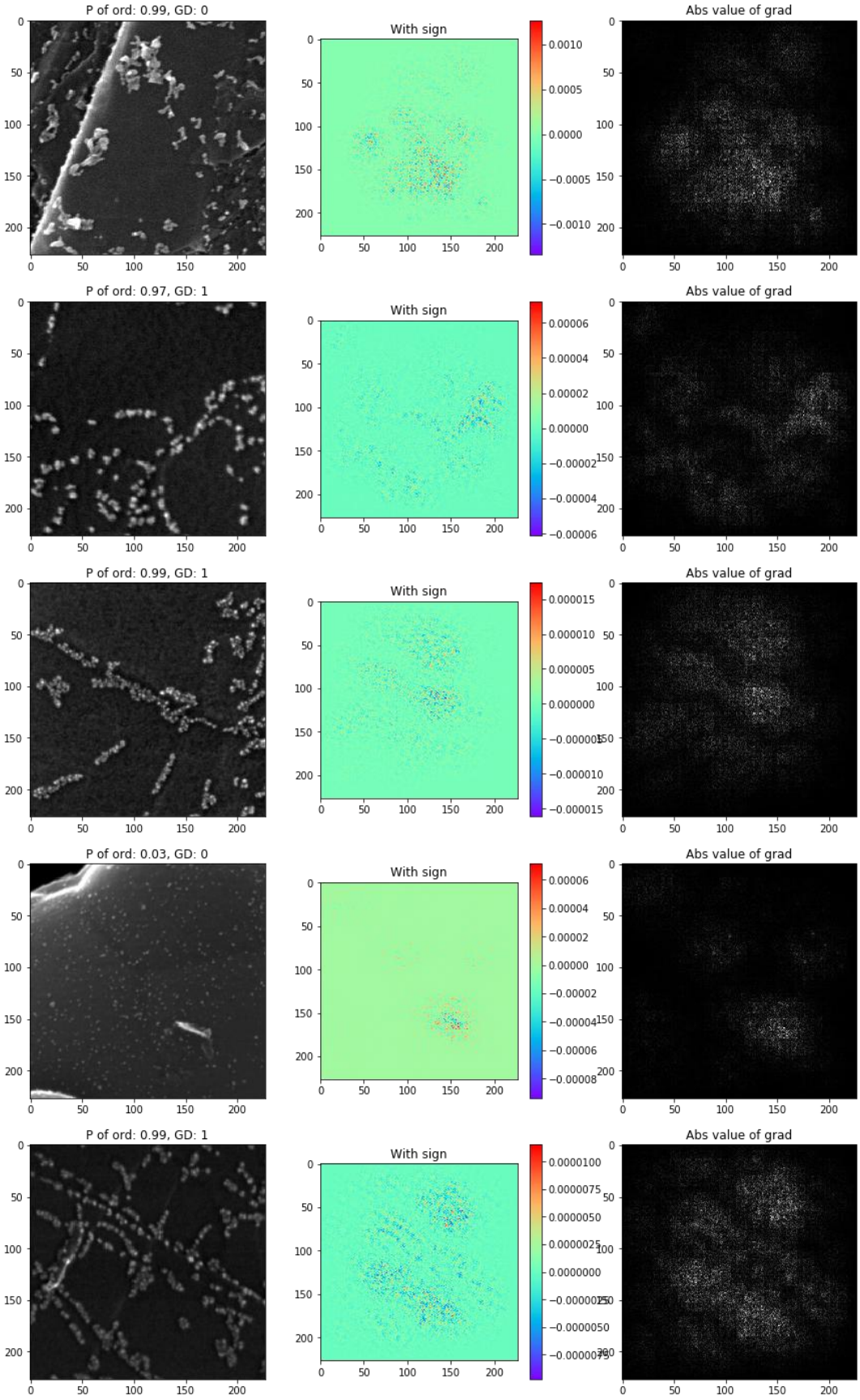


Fig. S21. Validation dataset: gradients of loss function with respect to the image. Left — source images, middle — gradients, right — absolute value of gradients

Human evaluation

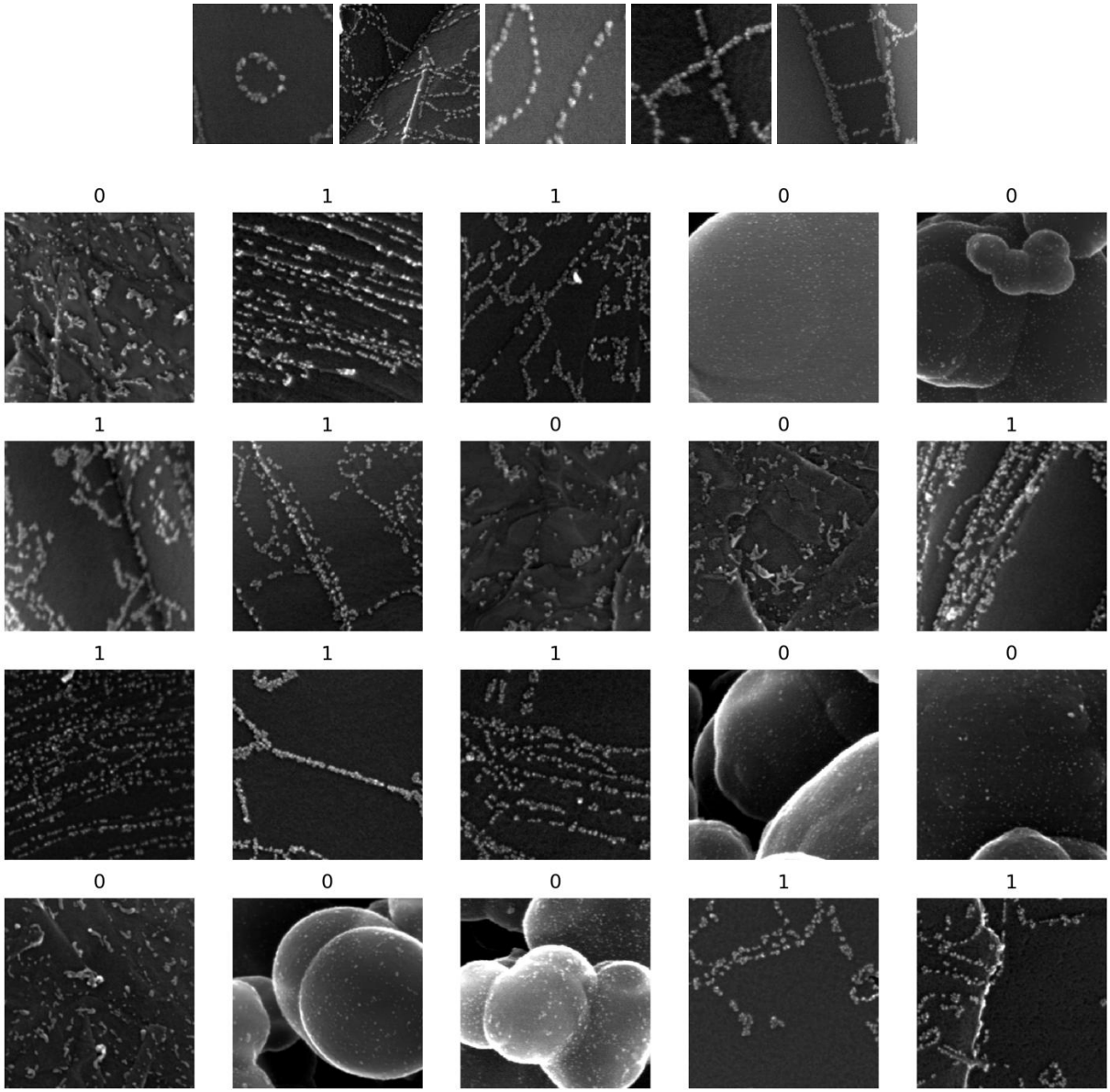


Fig. S22. Images selected for human evaluation (test images are shown on top). “1” — ordered, “0” — disorderd

Human training was performed using following images:

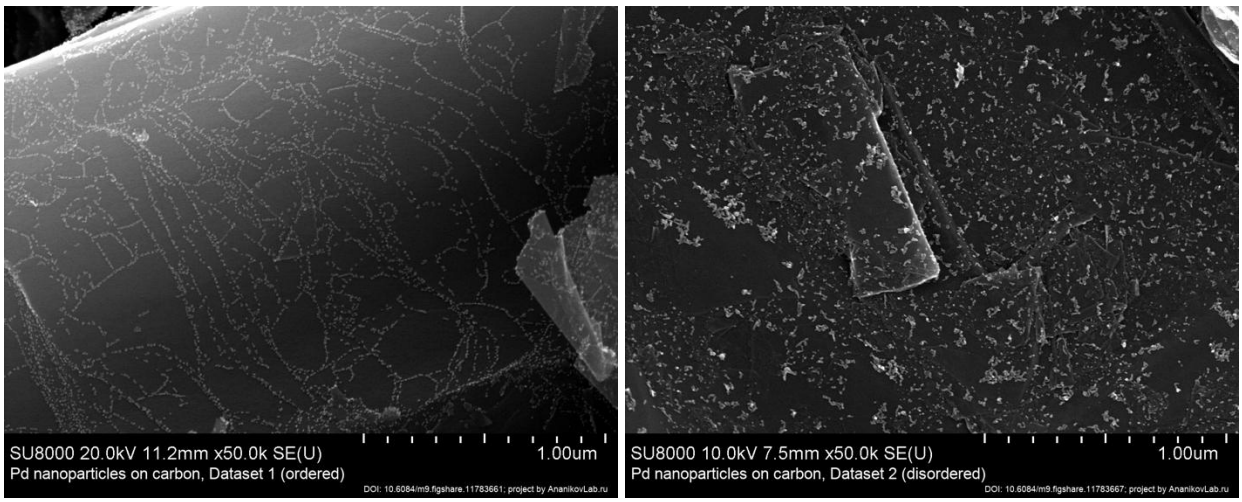


Fig. S23. Images, used to train humans (ordered location of nanoparticles – left; disordered – right)

The images were accompanied with a description of what we mean by order, where the nanoparticles are, and how to classify properly.

It was assumed, that some people would not read/understand the description and the task properly. To identify these answers, we included five very simple cases of images with high degree of order. Respondents' performance in these questions is shown below. In overall, 245 people were involved.

Table S2. Percentage of people who answered correctly to the test questions

Question / metric	**Correct answers**
1	87.8%
2	96.3%
3	94.7%
4	93.1%
5	98.0%
All test questions	79.6%

Then we calculated percentage of correct answers and classification metrics for all the respondents and for the people, who answered all test questions correctly, i.e. passed the test.

Table S3. Poll statistics by question/metric (passing the test means answering correctly each of the first five test questions as summarized in Table S1)

Question / metric	**Correct answers from those responders, who did not pass test questions**	**Correct answers from those responders, who passed test questions**
6	98.0%	98.5%
7	78.0%	90.8%

8	50.0%	73.3%
9	80.0%	93.3%
10	86.0%	94.4%
11	22.0%	48.7%
12	62.0%	93.8%
13	98.0%	97.4%
14	98.0%	97.4%
15	78.0%	80.5%
16	90.0%	95.4%
17	82.0%	99.5%
18	68.0%	89.7%
19	80.0%	94.9%
20	82.0%	95.4%
21	96.0%	97.4%
22	84.0%	93.8%
23	86.0%	92.3%
24	62.0%	88.2%
25	46.0%	76.9%
Precision score	0.866	0.962
Accuracy score	0.763	0.896
Recall score	0.638	0.837

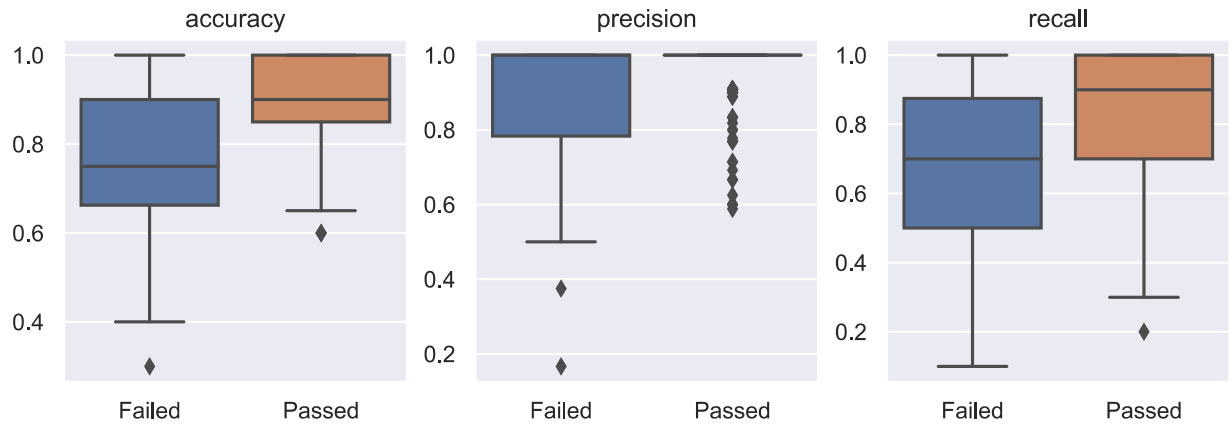


Fig. S24. Distribution of metrics for the poll responders

Segmentation neural network

Training details

Search for the architecture was done, using custom-build wrapper for Segmentation Models PyTorch package.[17] All network were trained, using Adam optimizer[18] with 10^{-4} learning rate. The networks were trained on a single NVIDIA 1080 TI.

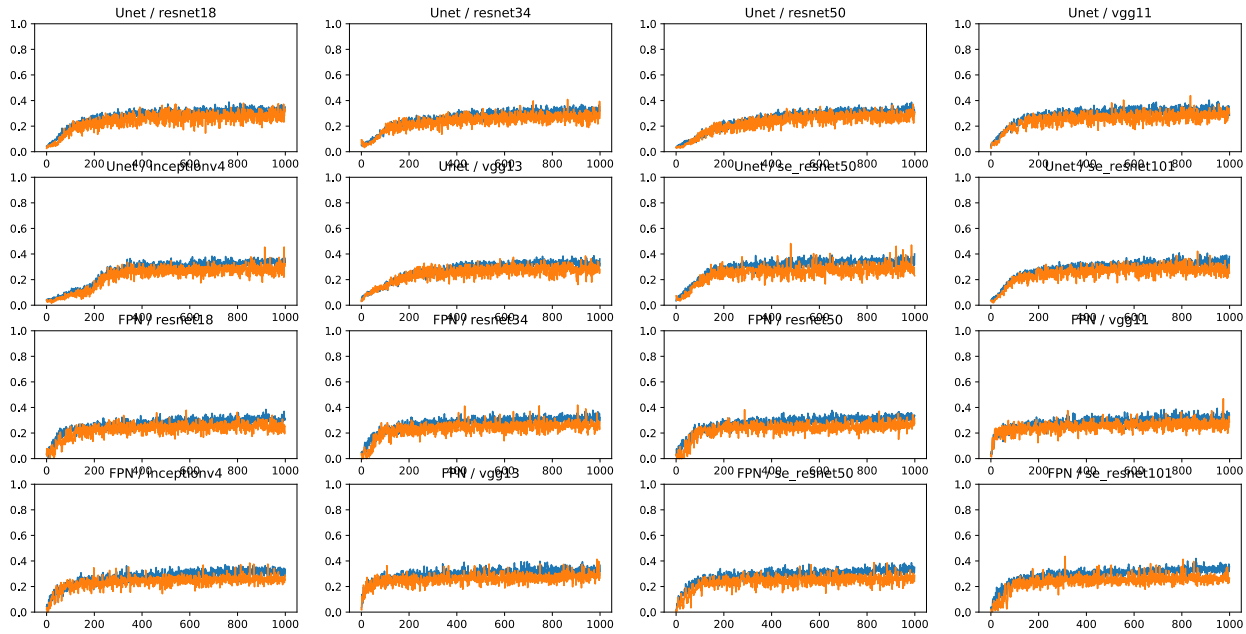


Fig. S25. Learning curves for models, evaluated during grid-search. Metric — ROC AUC score

Then the models were retrained but learning rate with lowered 10-fold every 750 epochs.

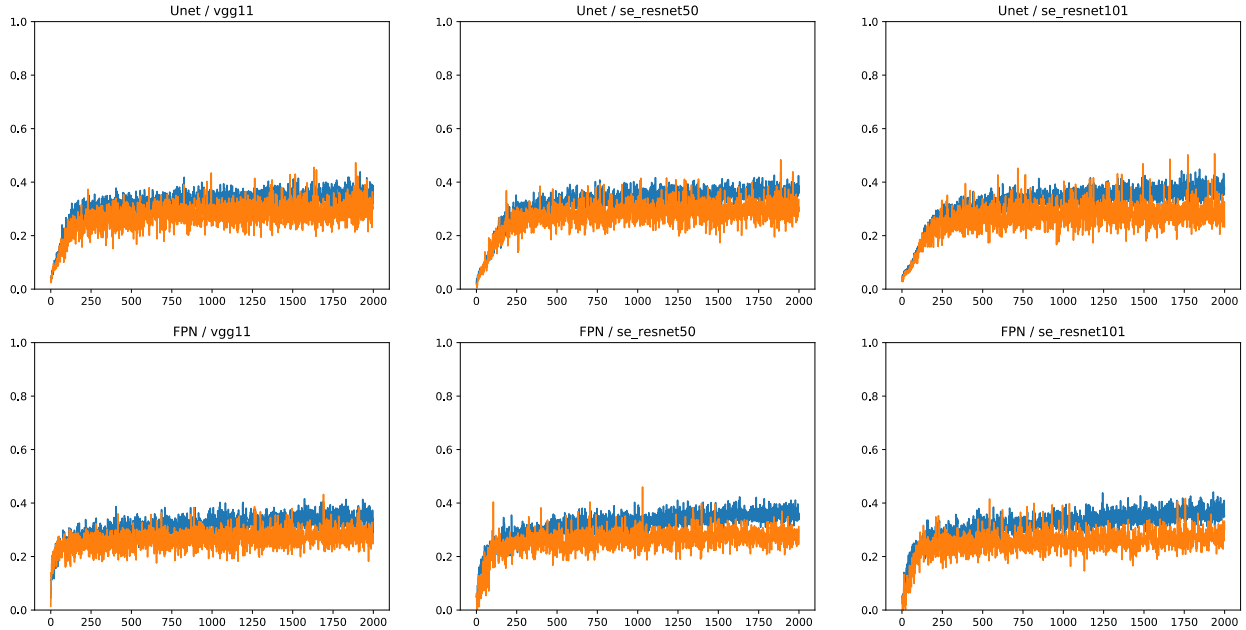


Fig. S26. Learning curves for the best models. Metric — ROC AUC score

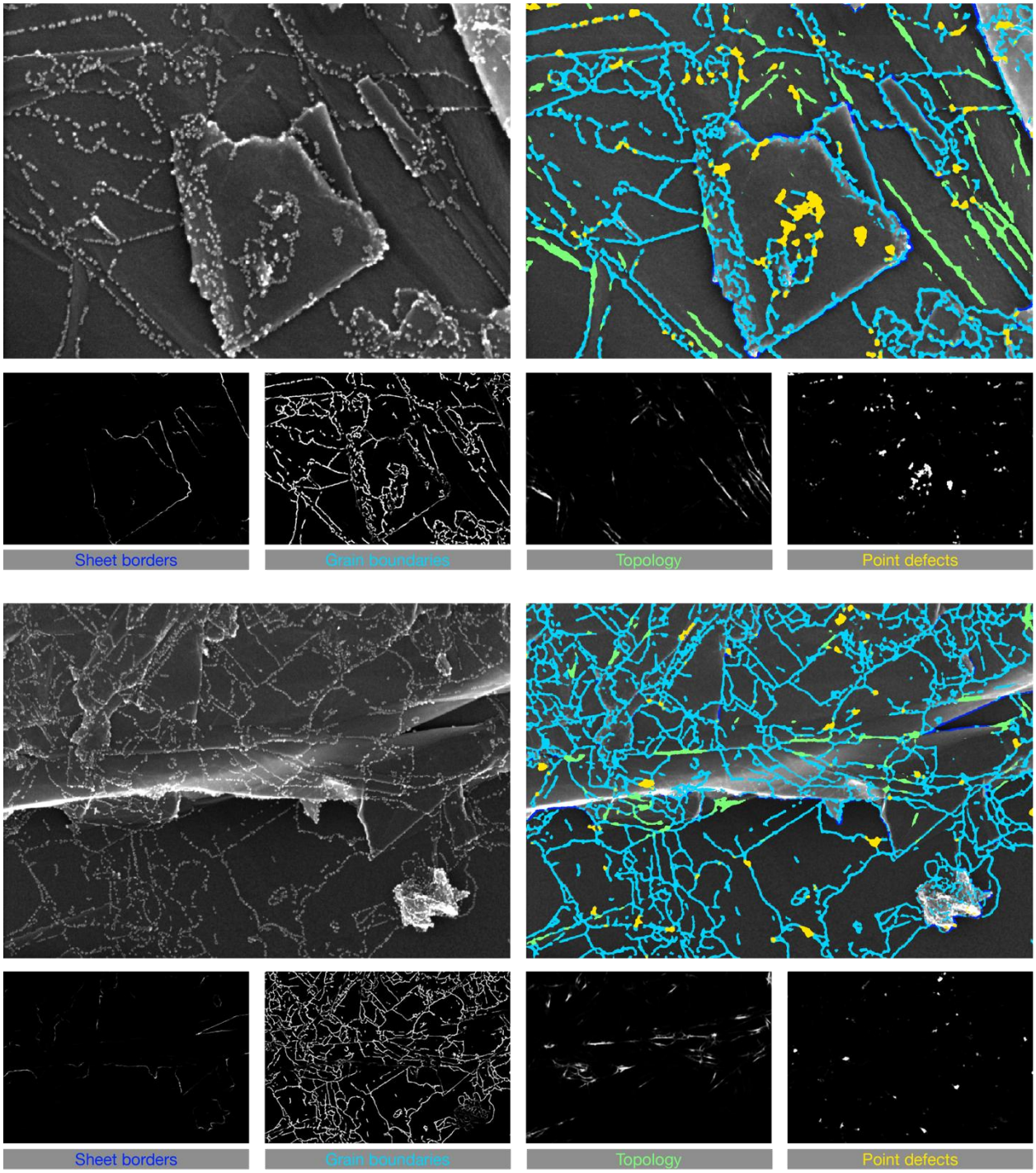


Fig. S27. Other examples of segmentation neural network outputs

References

1. Shakeri, M. Effect of randomly distributed asymmetric stone-wales defect on electronic and transport properties of armchair graphene nanoribbon. *Superlattices Microstruct.* **128**, 116–126 (2019).
2. Tapasztó, L. *et al.* Mapping the electronic properties of individual graphene grain boundaries. *Appl. Phys. Lett.* **100**, 053114 (2012).
3. Zhang, J. & Zhao, J. Structures and electronic properties of symmetric and nonsymmetric graphene grain boundaries. *Carbon N. Y.* **55**, 151–159 (2013).

4. Ayuela, A., Jaskólski, W., Santos, H. & Chico, L. Electronic properties of graphene grain boundaries. *New J. Phys.* **16**, 083018 (2014).
5. Sun, J. *et al.* Electronic and transport properties of graphene with grain boundaries. *RSC Adv.* **6**, 1090–1097 (2016).
6. Jing, N. *et al.* Effect of defects on Young's modulus of graphene sheets: a molecular dynamics simulation. *RSC Adv.* **2**, 9124 (2012).
7. Zhang, J. & Zhao, J. Mechanical properties of bilayer graphene with twist and grain boundaries. *J. Appl. Phys.* **113**, 043514 (2013).
8. Park, Y. & Hyun, S. Size Effect of Defects on the Mechanical Properties of Graphene. *J. Korean Phys. Soc.* **72**, 681–686 (2018).
9. Yermakov, A. Y. *et al.* Hydrogen Dissociation Catalyzed by Carbon-Coated Nickel Nanoparticles: Experiment and Theory. *ChemPhysChem* **14**, 381–385 (2013).
10. Xie, J. *et al.* Reduced graphene oxide-catalyzed oxidative coupling reaction of 4-methoxyphenol in aerobic aqueous solution. *Carbon N. Y.* **121**, 418–425 (2017).
11. Savaram, K. *et al.* Dry microwave heating enables scalable fabrication of pristine holey graphene nanoplatelets and their catalysis in reductive hydrogen atom transfer reactions. *Carbon N. Y.* **139**, 861–871 (2018).
12. Luo, Z. & Khanna, S. N. Carbon-Carbon Cross-Coupling Reactions. in *Metal Clusters and Their Reactivity* 143–162 (Springer Singapore, 2020). doi:10.1007/978-981-15-9704-6_9.
13. Huang, H. & Wang, X. Pd nanoparticles supported on low-defect graphene sheets: for use as high-performance electrocatalysts for formic acid and methanol oxidation. *J. Mater. Chem.* **22**, 22533 (2012).
14. Taillon, G. KS2D: 2 Dimensional Kolmogorov-Smirnov test for goodness-of-fit. *GitHub repository* (2021).
15. Boiko, D. A., Pentsak, E. O., Cherepanova, V. A. & Ananikov, V. P. Electron microscopy dataset for the recognition of nanoscale ordering effects and location of nanoparticles — Dataset 1 (ordered). doi:10.6084/m9.figshare.11704215.
16. Boiko, D. A., Pentsak, E. O., Cherepanova, V. A. & Ananikov, V. P. Electron microscopy dataset for the recognition of nanoscale ordering effects and location of nanoparticles. *Sci. Data* **7**, 101 (2020).
17. Yakubovskiy, P. Segmentation Models Pytorch. *GitHub repository* (2020).
18. Kingma, D. P. & Ba, J. L. Adam: A method for stochastic optimization. in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* (2015).