Artificial intelligence for imaging data analysis in materials science: microscopy and behind

A G Okunev1,2, M Yu Mashukov1, N N Sankova1,2 A V Nartova1,2 and A V Matveev4
1 Novosibirsk State University, Pirogova str., 1, Novosibirsk, 630090, Russia
2 Boreskov Institute of Catalysis SB RAS, pr. Ac. Lavrentieva, 5, Novosibirsk, 630090, Russia
E-mail: matveev.nsu@mail.ru

Abstract. In various fields of microscopy, a fairly common task is to count and determine the size of the studied objects. Sometimes this is a one-time tedious task that is done by lab personnel without the programming skills. In this paper, we present an approach that allows researchers to perform automated processing of research imaging data using our services and even to train deep neural networks on their own in no code mode.

1. Introduction
Advances in imaging technologies have had a significant impact on all areas of science and technology, providing the opportunity to study the objects with high resolution in space and time. Advances in the physical foundations of microscopy have made it possible to create devices that can be used to produce images at the level of individual atoms. Often, when studying physical, chemical, and biological objects, it is important to determine their statistical parameters (quantity, size, and area). At the same time, the operator needs to process thousands of objects to obtain statistically reliable results. To solve these problems, various methods and technologies based on computer vision, machine learning, neural networks, etc. are being actively developed, primarily for biology purposes [1-3].

We have developed two services based on the use of deep machine learning:

- an online service ParticlesNN for recognizing nanoparticles in images of scanning tunnelling microscopy (STM) and transmission electron microscopy (TEM).
- cloud service DLgram, using Telegram messenger as an interface, which can be used for training a neural network and recognizing any type of object.

2. The online service ParticlesNN
We have developed the online service ParticlesNN [4], http://particlesnn.nsu.ru, which is based on the Cascade Mask-RCNN neural network.

2.1. Scanning tunnelling microscopy
Training was performed on a dataset containing 23 STM images with 5157 nanoparticles. As a result, the trained neural network recognized nanoparticles in the verification set with 0.93 precision. The
accuracies for mean particle size calculated from predicted contours compared with ground truth were in the range of 0.87–0.99, figure 1 [5, 6].

Figure 1. Example of nanoparticles recognition on a STM image (a). Boundary particles are highlighted in pink, which can be excluded when calculating the average particle size to get a more accurate result. A histogram of the size distribution of recognized nanoparticles is shown on (b). STM image of Pt nanoparticles deposited on highly-oriented pyrographite was obtained using an SPM 100 VT (RHK Technology, USA).

2.2. Transmission electron microscopy

Training was performed on a dataset containing 21 TEM images with 1114 nanoparticles. As a result, the trained neural network recognized nanoparticles in the verification set with 0.71 precision. The accuracies for mean particle size calculated from predicted contours compared with ground truth were in the range of 0.73–0.84, figure 2.

Figure 2. Transmission electron microscopy image of Pt nanoparticles on alumina catalyst: 223 particles were recognized by the network (a). A histogram of the size distribution of recognized nanoparticles is shown on (b). TEM image was obtained using a JEM-2010, Jeol Co., Japan, with the lattice resolution of 0.14 nm with accelerating voltage of 200 kV.

The online service ParticlesNN processes the results obtained by determining the size of the contours and their statistical parameters – the degree of coverage, concentration, average size, and distribution of objects by size.
The Particles NN online service differs from other software products (CellProfiler [1], Ilastik [2], ImageJ/Fiji [3]) in the following features:

- it is possible to process images that contain significant noise and artifacts without pre-processing.
- a user can adjust automatically defined contours in order to refine the recognition results.
- the option of simultaneous statistical processing of several images is implemented.
- the processing results are displayed in the form of a histogram and tables, in which information on all identified objects is available – their coordinates, dimensions.
- the refinement of particle contours is implemented using the fitting procedure with a Gaussian 2D distribution.
- the functions of correct accounting of objects on the image border are implemented.

![Figure 3. Scanning electron microscopy (SEM) images obtained using a Hitachi TM-1000 microscope of polystyrene microspheres: a training image with 2 labeled particles (a); 281 particles recognized by network, 35 are not identified (b).](image)

![Figure 4. Optical microscopy images obtained using an optical microscope Biomed-5 with Levenhuk M1000 Plus camera at a magnification of x100 of polystyrene microspheres: a training image with the 3 labelled particles (a); 807 particles recognized by network, 104 are not identified (b).](image)

3. The cloud service DLgram
We implemented a revolutionary subshot learning approach that allows researchers to train a deep neural network using a handful of the instances of objects of interest. The trained network able to effectively recognize the rest of the object instances on the same image as well as correctly segment similar objects.
on other images, as shown in figures 3-4. Training and inference can be implemented in no code mode and perform automated processing of scientific data using a cloud service in Telegram messenger group referred as t.me/nanoparticles_nsk.

The realized cloud service is especially user-friendly for the following reasons:

- There is no need to install special software.
- it is possible to train a neural network on any type of objects.
- only a few objects should be labelled for training.
- the trained neural network can be used for recognizing identical objects on other similar images.
- a user can adjust automatically defined contours in order to refine the recognition results.
- the statistical processing of the obtained data is implemented.

4. Conclusion

We have introduced two services for automatic image data processing based on the use of deep neural networks:

- online service ParticlesNN with a pre-trained neural network for recognizing nanoparticles in scanning tunneling and transmission electron microscopy images.
- the cloud service DLgram based on the Telegram channel that allows a user to train the network on its own and to use it in future work.

The obtained data can be used to determine their statistical parameters (quantity, size, and area).

Acknowledgments

This work was supported by Ministry of Science and Higher Education of the Russian Federation within the 5-100 Excellence Programme, as well as within the governmental order for Boreskov Institute of Catalysis SB RAS (project AAAA-A21-12101139011-4).

References

[1] McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini B, Karhohs K, Doan M, Ding L, Rafelski S, Thirstrup D, Wiegraebe W, Singh S, Becker B, Caicedo J and Carpenter A 2018 CellProfiler 3.0: Next-generation image processing for biology PLOS Biology 16 1-17
[2] Berg S, Kutra D, Kroeger T, Straehle C, Kausler B, Haubold C, Schiegg M, Ales J, Beier T, Rudy M, Eren K, Cervantes J, Xu B, Beuttenmueller F, Wolny A, Zhang C, Koethe U, Hamprecht F and Kreshuk A 2019 ilastik: interactive machine learning for (bio)image analysis Nat Methods 16 1226-32
[3] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White D J, Hartenstein V, Eliceiri K, Tomancak P and Cardona A 2010 Fiji: An open-source platform for biological-image analysis Nat Methods 9 676-82
[4] Particles N N http://particlesnn.nsu.ru
[5] Okunev A G, Nartova A V and Matveev A V 2019 Recognition of nanoparticles on scanning probe microscopy images using computer vision and deep machine learning International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON) Novosibirsk Russia 0940-3
[6] Okunev A G, Mashukov M Y, Nartova A V and Matveev A V 2020 Nanoparticle recognition on scanning probe microscopy images using computer vision and deep learning Nanomaterials 10 1-16