small methods
www.small-methods.com

Synthetic Image Rendering Solves Annotation Problem in Deep Learning Nanoparticle Segmentation

Leonid Mill, David Wolff, Nele Gerrits, Patrick Philipp, Lasse Kling, Florian Vollnhals, Andrew Ignatenko, Christian Jaremenko, Yixing Huang, Olivier De Castro, Jean-Nicolas Audinot, Inge Nelissen, Tom Wirtz, Andreas Maier, and Silke Christiansen*

Nanoparticles occur in various environments as a consequence of man-made processes, which raises concerns about their impact on the environment and human health. To allow for proper risk assessment, a precise and statistically relevant analysis of particle characteristics (such as size, shape, and composition) is required that would greatly benefit from automated image analysis procedures. While deep learning shows impressive results in object detection tasks, its applicability is limited by the amount of representative, experimentally collected and manually annotated training data. Here, an elegant, flexible, and versatile method to bypass this costly and tedious data acquisition process is presented. It shows that using a rendering software allows to generate realistic, synthetic training data to train a state-of-the art deep neural network. Using this approach, a segmentation accuracy can be derived that is comparable to man-made annotations for toxicologically relevant metal-oxide nanoparticle ensembles which were chosen as examples. The presented study paves the way toward the use of deep learning for automated, high-throughput particle detection in a variety of imaging techniques such as in microscopies and spectroscopies, for a wide range of applications, including the detection of micro- and nanoplastic particles in water and tissue samples.

1. Introduction

Nanoparticles are omnipresent in our daily lifes. They can be found in products ranging from cosmetics, textiles, and foods, and are important in various technological fields such as energy, electronics, medicine, and many more.[1–5] However, as a consequence of industrial processes and man-made pollution, unwanted nanoparticle size distributions and concentrations[6] give rise to concerns with respect to human health and environmental pollution. While the nanoparticles' physicochemical properties (size, shape, surface chemistry, etc.) determine the quality of products,[7,8] such characteristics are also important in order to evaluate the biological impact of nanoparticles at a molecular, cellular, and systemic level for any risk assessment for environmental and human health.[9] Characterizing nanoparticles in a dynamic context and on a case-by-case basis, microscopic imaging techniques including those that use focused electron or ion beams in scanning electron microscopes (SEMs) or helium ion microscopes[10] (HIMs) to generate nanometer scale spatial resolution are frequently applied in the scientific community. Given the substantial information content of digital images, these techniques often benefit from, or require, automated high-throughput data analysis that enables the accurate identification of large numbers of particles in a robust way.

L. Mill, Dr. C. Jaremenko, Prof. A. Maier
Pattern Recognition Lab
Friedrich-Alexander-University Erlangen-Nuremberg
91058 Erlangen, Germany
E-mail: leonid.mill@fau.de

L. Mill, Dr. F. Vollnhals, Prof. S. Christiansen
Institute of Optics, Information and Photonics
Friedrich-Alexander-University Erlangen-Nuremberg
91058 Erlangen, Germany

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/smtd.202100223.

DOI: 10.1002/smtd.202100223

N. Gerrits, Dr. I. Nelissen
Health Unit
Flemish Institute for Technological Research
Mol 2400, Belgium

D. Wolff, L. Kling, Dr. F. Vollnhals, Dr. C. Jaremenko
Institut für Nanotechnologie und korrelative Mikroskopie
91301 Forchheim, Germany

Dr. P. Philipp, Dr. A. Ignatenko, Dr. O. De Castro, Dr. J-N. Audinot, Dr. T. Wirtz
Advanced Instrumentation for Ion Nano-Analytics, Materials Research and Technology Department
Luxembourg Institute of Science and Technology
Belvaux L-4422, Luxembourg

Prof. S. Christiansen
Physics Department
Free University
14195 Berlin, Germany

Prof. S. Christiansen
Correlative Microscopy and Material Data Department
Fraunhofer Institute for Ceramic Technologies and Systems
01277 Dresden, Germany

Several approaches[11–21] have been proposed for automated image analysis of SEM and TEM images. However, most of these approaches rely on single thresholds for the feature separation,[20,21] encounter major difficulties caused by irregular object patterns and noise,[22] or they rely on hand-crafted features for the particle shapes,[14,15] which impair the generalization potential of such algorithms for the characterization of arbitrary nanoparticles or heterogeneous nanoparticle ensembles. In order to handle the complexity of nanoparticle images that include various sizes, shapes, distributions, and shadow variations more sophisticated image analysis approaches are required. With the recent advancements in machine learning and mainly deep learning,[23] deep convolutional neural networks (CNNs)[24–29] have been developed, which are able to learn from data sets containing millions of images[30] to resolve object detection tasks. When trained on such big data sets, CNNs are able to achieve task-relevant object detection performances that are comparable or even superior to the capabilities of humans.[31,32]

In recent years, new methods have been proposed for the analysis of nanostructures in SEM and TEM images that rely on advanced machine and deep learning techniques[33–45] which allow for accurate and high-throughput image analysis. However, as most of these methods use a supervised learning approach, significant human effort is needed to prepare the training data. The main difficulty for the training data lies in the acquisition of a representative data set of nanoparticle images which ideally contain various sizes, shapes, and distributions for a variety of nanoparticle types. Additionally, manual annotation of the acquired data is mandatory to obtain the so called "ground truth" or "labels", which is, in general, error-prone, time-consuming, and consequently costly. Although approaches exist such as "precision learning"[46] or "transfer learning"[47] that reliably work with a substantially reduced amount of training data, a certain data set size is still required that incorporates human effort, that is for manual annotation of the acquired data. To partially solve this problem recent approaches have been introduced that use synthetically generated[38,41] SEM and TEM images as training data for a deep learning based nanoparticle analysis. However, accurate and realistic nanoparticle simulations still remain a complex-to-solve task, especially for the segmentation of stand-alone particles in superimposed 3D particle ensembles. To overcome the aforementioned data limitations, we have developed a semi-automated data synthesis scheme using an open-source rendering software. Particularly, we use Blender[48] in this work, in order to generate synthetic data based on very limited amounts of real microscopic data sets. The proposed workflow enables the generation of a virtually unlimited number of synthetic and photo-realistic microscopic nanoparticle images comprising various types of particles with different sizes, shapes, compositions, and 3D distributions. Thereby, the respective synthetic ground truth segmentation mask is automatically derived for each generated image. We show that the synthetic data is sufficient in terms of realism and size to successfully train a deep learning model for segmentation which is also able to operate on real data. As a demonstration example, we generate photo-realistic synthetic data sets of HIM images of metal-oxide and metal nanoparticles (SiO_2, TiO_2, and Ag) that serve as training data for a state-of-the art[29,49] deep CNN so called U-Net.[28] For experimental validation, manual annotations of the HIM images have been carried out by experts. We show that training the CNN on synthetic data yields similar segmentation accuracy as using real HIM images. Furthermore, we demonstrate the applicability of the proposed method for images containing very complex shaped and distributed Ag particles, nanorods, and nanowires.

2. Novel, Semi-Automated Nanoparticle Segmentation Workflow

Our goal is to extract statistical information on nanoparticle morphology with respect to size, shape, and distribution from high resolution microscopy images acquired with electron- or ion microscopes. However, to guide particle segmentation by deep CNNs, a representative amount of training data is required. This is in general a true limitation due to the fact that microscopy image data is usually not available in sufficiently large amounts to be suitable for robust CNN training purposes. Moreover, we want to overcome the manual image annotation process which is very time-consuming and error-prone. Here, we propose a novel workflow that overcomes the data problem for the deep learning based analysis of particle images. The workflow relies on a very limited number of real reference images, for example, out of electron or ion microscopes, which serve as a blue print for the semi-automated photo-realistic synthetic data generation using a rendering software for example Blender. The proposed nanoparticle segmentation workflow is composed of the steps shown in **Figure 1**. Images of TiO_2 nanoparticles on a silicon wafer surface, taken by a HIM serve as examples for the demonstration of the procedure and the quality of the statistical evaluation of the nanoparticle properties.

The starting point of the segmentation workflow is the acquisition of a limited number of high resolution HIM images of TiO_2 nanoparticles, spread on a (100) silicon wafer. Figure 1a shows an example of such an HIM image. Based on a number of images of that type (see Figure S2, Supporting Information,) a human is able to obtain an overall impression of the average size, shape, potential faceting, and orientation of the TiO_2 nanoparticles with respect to one another as well as the degree of homogeneity of the particle distribution within those images (Figure 1b). The knowledge gained from the initial manual assessment of the nanoparticle data is used to reproduce such features virtually using a rendering software, for example, Blender,[48] and to automatically generate 3D scenes filled with synthetic TiO_2 particle ensembles. During this automated rendering process, details of the individual particles as well as the overall 2D or 3D arrangements on the substrate surface, including potential image artefacts (i.e., such as dirt which may arise from sample preparation procedures), are included in virtual Blender scenes to mimic real TiO_2 particle ensembles. An automatically generated TiO_2 scene is demonstrated in Figure S9, Supporting Information. Details on the automatic Blender scene generation process are provided in Section 4. Subsequently, after a user-defined number of scenes is generated, each Blender scene is automatically being rendered twice. The first rendering process computes a photo-realistic synthetic microscope image of the TiO_2 nanoparticles and potential artefacts, while a subsequent render produces the respective error-free ground truth label image (Figure 1b).

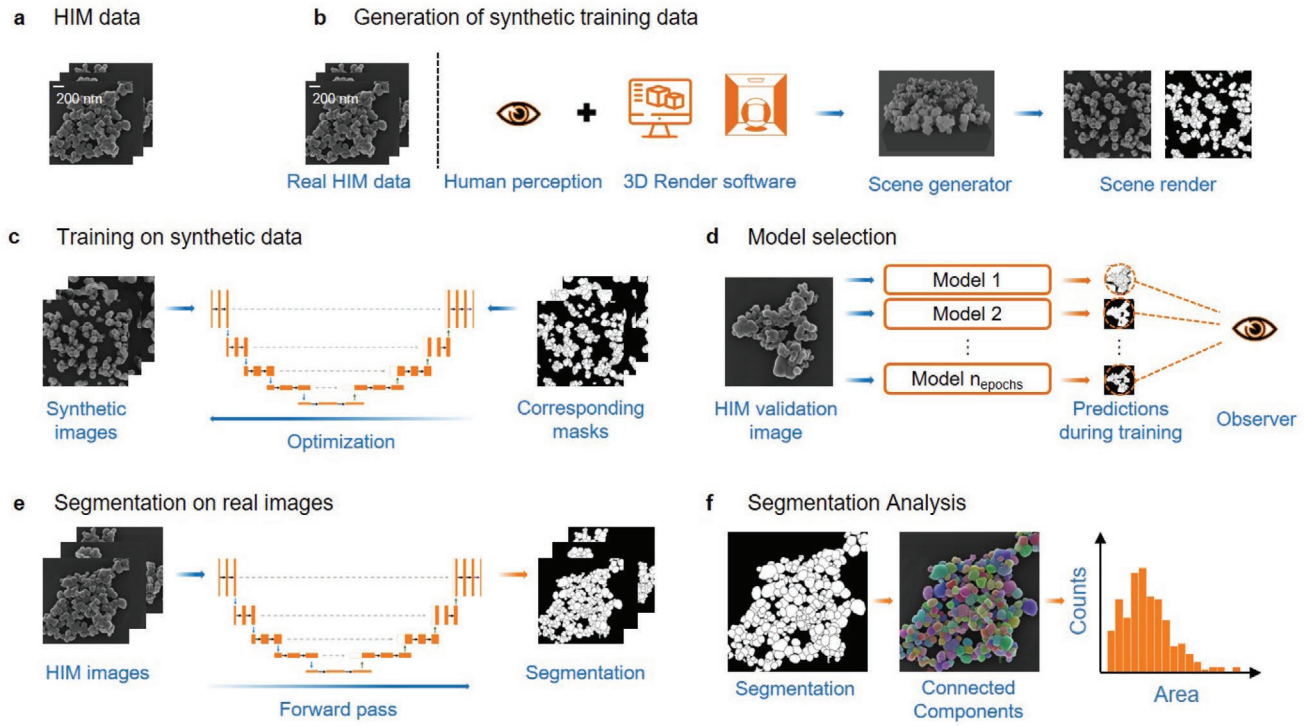

Figure 1. Illustration of the nanoparticle segmentation workflow for real Helium ion microscope (HIM) nanoparticle images using a deep convolutional neural network (CNN) and synthetic training data. a) HIM images of TiO_2 nanoparticle ensembles on a silicon (100) substrate surface. b) A semi-automated procedure to synthesize HIM images of TiO_2 particles based on the knowledge of a limited number of real HIM reference images. Using a render software (here Blender[48]) in combination with the information on particle size, shape, and distribution gained from the reference images it is possible to create synthetic images that mimic real HIM data realistically. During this process, the respective ground truth labels for the synthetic TiO_2 image data are also generated in a fully automated manner. c) The synthetic data set is used to train a deep CNN (here the U-Net[28]) for automated particle segmentation and subsequent quantitative, statistical assessment. d) After training the CNN using synthetic data, the model which predicts the most accurate segmentation on a real validation image is selected for further processing. e) The best CNN model is used to predict the segmentation masks for all experimental microscopy images. f) The accurate prediction of the model permits further statistical image analysis.

With a sufficient amount of realistic synthetic data, a deep CNN can be properly trained on the nanoparticle segmentation task. In our workflow, we use the so called U-Net[28], a the state-of-the art[29,49] deep CNN, which was originally proposed and successfully applied for the segmentation of cells in transmission electron microscope images. Figure 1c shows the U-Net comprising a U-shaped encoder–decoder architecture. The orange boxes represent multi-dimensional feature maps extracted by different convolution layers, while the dashed gray arrows correspond to skip-connections. A detailed description of the architecture is provided by Ronneberger et al.[28]. During the U-Net training process, the network iteratively optimizes its internal parameters to learn the segmentation of nanoparticles based on our synthetic microscopic images by predicting corresponding segmentation masks. To prevent the network from overfitting on the synthetic data, at least one real HIM "validation" image is used during training. In machine learning, "validation data" is used during the training process to evaluate the model performance on hold-out data which is not part of the training data set. Therefore, after each complete training iteration, also referred to as an "epoch" in machine learning, the model predicts segmentation for the validation image (Figure 1d). After the CNN training process, a human observer visually evaluates all predicted segmentation masks for the validation image and selects the model which provides the qualitatively most accurate particle segmentation. Subsequently, the selected model is used to segment all remaining real microscopic (here from a HIM) images (Figure 1e), while subsequently, all sorts of statistical and quantitative information can be deduced for the microscopic HIM images of interest (Figure 1f).

2.1. Synthetic Image Quality

We used the aforementioned semi-automated data generation scheme to generate synthetic microscope images such as the HIM images of SiO_2, TiO_2, and silver nanoparticles of varying overall morphology.

A comprehensive set of real HIM images of such nanoparticles are shown in the in Figures S1–S3, Supporting Information. However, since no human derived manual annotation was performed for silver nanoparticle ensembles in HIM images due to the complexity of the data, in the following we will set our focus on SiO_2 and TiO_2 nanoparticles for the quantitative and qualitative evaluation. **Figure 2** provides a direct comparison of SiO_2 (Figure 2a) and TiO_2 (Figure 2b) particles in HIM images (top row) with the corresponding synthetic photo-realistic images (middle row) and its respective synthetic labels (bottom row).

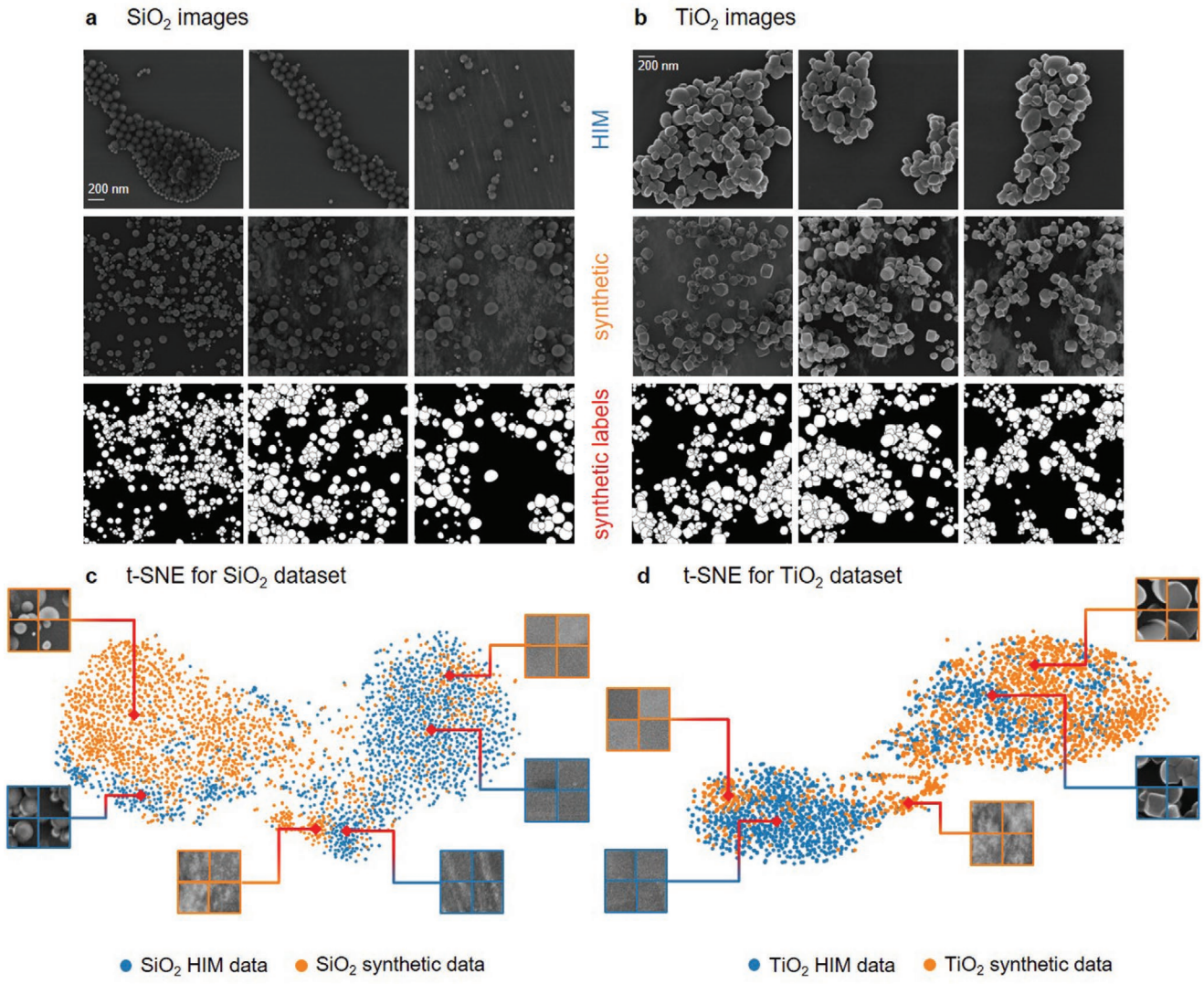

Figure 2. a,b) Comparison of metal-oxide nanoparticles on silicon wafers imaged by a HIM (top row) with synthetically generated images using the proposed semi-automated rendering process (middle row) with its respective synthetic labels (bottom row). a) Spherical SiO_2 nanoparticles with various dimensions on a silicon wafer surface arranged in clusters or a certain degree of alignment in close vicinity to neighboring particles. Also, note the dirt smear on the silicon wafer surface in the rightmost image, which was also covered in the simulation. The synthetic images in the middle row on the contrary show a more statistical distribution of particles compared to the real images, which also include the surface impurity feature. b) Microscopic HIM images of complex shaped and distributed TiO_2 nanoparticles. c,d) t-SNE visualization of the SiO_2 and TiO_2 data sets. Each data point corresponds to an image patch of identical size (144 × 144 px) of real (blue) and synthetic (orange) images. Both t-SNE plots show a distinct separation between background and foreground (particle) data points, while nearby and overlapping points indicate the high similarity between synthetic and real data. c) The t-SNE result for the SiO_2 data set shows two clusters of particle images (left cluster), background images (right cluster) with an additional cluster for image patches, that show surface impurity. The small overlap between real and synthetic SiO_2 particle images indicates that the synthetic images are similar, but do not represent a perfect match of the target HIM data d) t-SNE plot for the TiO_2 particles. The large overlap between synthetic and real particle images (right cluster) indicates that the synthetic images represent the real TiO_2 HIM data quite well.

The real SiO_2 particles essentially show a very regular, spherical shape, which was also considered in the corresponding synthetic SiO_2 simulations (see Figure S5, Supporting Information). The arrangement of the SiO_2 particles in the real HIM images is not random, the particles align in close vicinity, ideally in clusters or rows of particles. In contrast, the particles in the synthetic images are distributed randomly (see Figure S6, Supporting Information) throughout the substrate in order to generate very complex distribution in contrast to the real images and thus increase the variance in the data. The TiO_2 particles on the other hand form agglomerates, are arranged in sheets and have complex shapes. Therefore, in order to cover these characteristics in the synthetic data, we designed a range (we chose four) of virtual 3D so called "template objects" (see Figure S9, Supporting Information) that were randomly combined to complex agglomerates in the TiO_2 render scenes during the automated image creation process (see Figure S10, Supporting Information). For both particle types (SiO_2 and TiO_2), we were able to mimic the characteristics of all relevant particle characteristics in terms of shape and distribution while

comprising a more complex localization of particles. Additionally, due to the presence of substrate surface impurity ("dirt") that stems from imperfect sample preparation, we extended our simulation with a randomized impurity texture creation (see Figure 1b; Figures S4 and S8, Supporting Information) to account for such features. Further details on the impurity texture are provided in Section 4. Note that in contrast to manually annotated images, our synthetic labels do not contain any mislabeling and provide error-free, consistent particle contour lines as displayed in Figure 2a,b (bottom row). Further synthetic images of SiO_2 and TiO_2, as well as the synthetic representation of more complex silver (Ag) nanoparticles that align in from of particle wires, are shown in Figures S11– S13, Supporting Information.

To assess the similarity between the real and synthetic HIM nanoparticle images, we analyzed the SiO_2 and TiO_2 data sets using the t-distributed stochastic neighbor embedding[50] (t-SNE). t-SNE is an unsupervised dimensionality reduction method which is primarily used to visualize high-dimensional data. In simpler terms, t-SNE provides for an idea of how data is arranged in a high-dimensional space. In general, nearby and overlapping data points in a t-SNE plot indicate similar data, while distant data points correspond to significant differences in the data. The resulting t-SNE plots for the SiO_2 and TiO_2 nanoparticles are depicted in Figure 2c,d, respectively. Each scatter point corresponds to an image patch or sample of the size of 144 $\times$ 144 px (with one pixel covering an area of 0.976 $\times$ 0.976 nm^2) of a real (blue) or synthetic (orange) HIM image. More technical details on the sample extraction and data visualization methods are provided in Section 4. In both t-SNE plots, a distinct separation between particle and background images can be observed. The nearby and overlapping blue and orange points indicate the overall high similarity between synthetic and real images. Moreover, for the SiO_2 data, the samples that contain substrate surface impurity in both real and synthetic data points show a high comparability which suggests a very accurate dirt simulation. However, the small overlap between real and synthetic SiO_2 particle images indicates that although the synthetic images are comparable to the real HIM images, the simulation does not perfectly match the target HIM data. While the real data shows a higher variance of background images in terms of noise distribution and pixel intensities, the opposite holds for the synthetic particle data. This can be explained by the different parameters for the simulation process that result in the generation of a high variance of particle images. These images contain for example, complex SiO_2 particle ensembles and a high variety of bright and dark particle instances caused by various shadowing effects in combination with varying light source intensities. This leads to an overall higher variance for the synthetic particle images which is also reflected in the SiO2 t-SNE plot (Figure 2c). More technical details about the data generation are provided in Section 4. Additionally, in both plots an imbalance between synthetic and real particle and background data points can be observed. In this context, synthetic SiO_2 and TiO_2 images contain significantly more particle than background data. This is an intended and expected behavior which results from our assumption that the simulation of random particle distributions with a high number of particles leads to a high variance of complex synthetic particle ensembles. As a consequence, it is expected that a deep CNN trained on complex synthetic data is more robust to variations in real images.

2.2. Model Selection

Since the U-Net was trained on synthetic data only (Figure 1c) real validation data is needed to assess the model performance for real HIM particle data. To select the best model, we compared two approaches for the model selection in a user study with seven participants, consisting of three experts on segmentation and four non-experts. The first approach, referred to as "analytical", requires the time-consuming manual annotation of at least one real HIM validation image, which is subsequently used as ground truth (GT) to select the best model based on the highest F1 score.[51] The F1 score is a metric which is often used to assess the segmentation accuracy of a model compared to a reference segmentation. The second approach relies on a human observer to choose the best model according to the qualitatively (visually) best segmentation performance (Figure 1d). Further details on the user study and model selection are presented in the Methods section. **Figure 3**a illustrates the models, defined by the epoch number, that were selected by each participant for the SiO_2 (orange) and TiO_2 (blue) data sets, respectively. The plot illustrates that the expert group (P5–P7) mostly selected models toward the end of the training process, while no significant trend could be observed for the non-expert group (P1–P4). A detailed qualitative evaluation of the selected segmentation masks showed that in contrast to the non-expert group, the experts tend to choose models that demonstrated a higher capability of separating particles in more detail, producing more accurate contour lines, whereas non-experts relied on the overall segmentation performance only. Figure 3b visualizes the segmentation accuracies of the selected (manually and "analytically") five best models on the validation image based on the mean F1 score.

For both particle types (SiO_2 and TiO_2), the non-expert group (P1–P4) as well as the experts (P5–P7) consistently selected models that provided almost perfect segmentation results on the validation image (SiO_2 - F1 mean (SD): 0.92 (0.01); TiO_2 - F1 mean (SD): 0.93 (0.004)). However, the models selected by the "analytical" approach demonstrate marginally more accurate segmentation performances (SiO_2 - F1 mean (SD): 0.93 (0.001); TiO_2 - F1 mean (SD): 0.93 (0.001)). A qualitative (visual) comparison between the two approaches is displayed in Figure 3c. This figure visualizes the validation images, the segmentation masks predicted by the model with the best F1 score according to the "analytical" approach, the segmentation of a representative expert-selected model as well as the manual GT reference segmentation. Both approaches resulted in a model selection that showed a very high overall segmentation accuracy on the validation image compared to the GT reference. However, the model selected by the "analytical" approach lacks a distinct separation of individual TiO_2 particles, while the expert implicitly selected a model that accounted for this behavior. Moreover, the expert-selected model produces more precise particle contour lines over the "analytical" approach. Since statistical analysis of nanoparticle images requires an accurate separation of particles, we further rely on expert-selected models for the particle analysis in the next

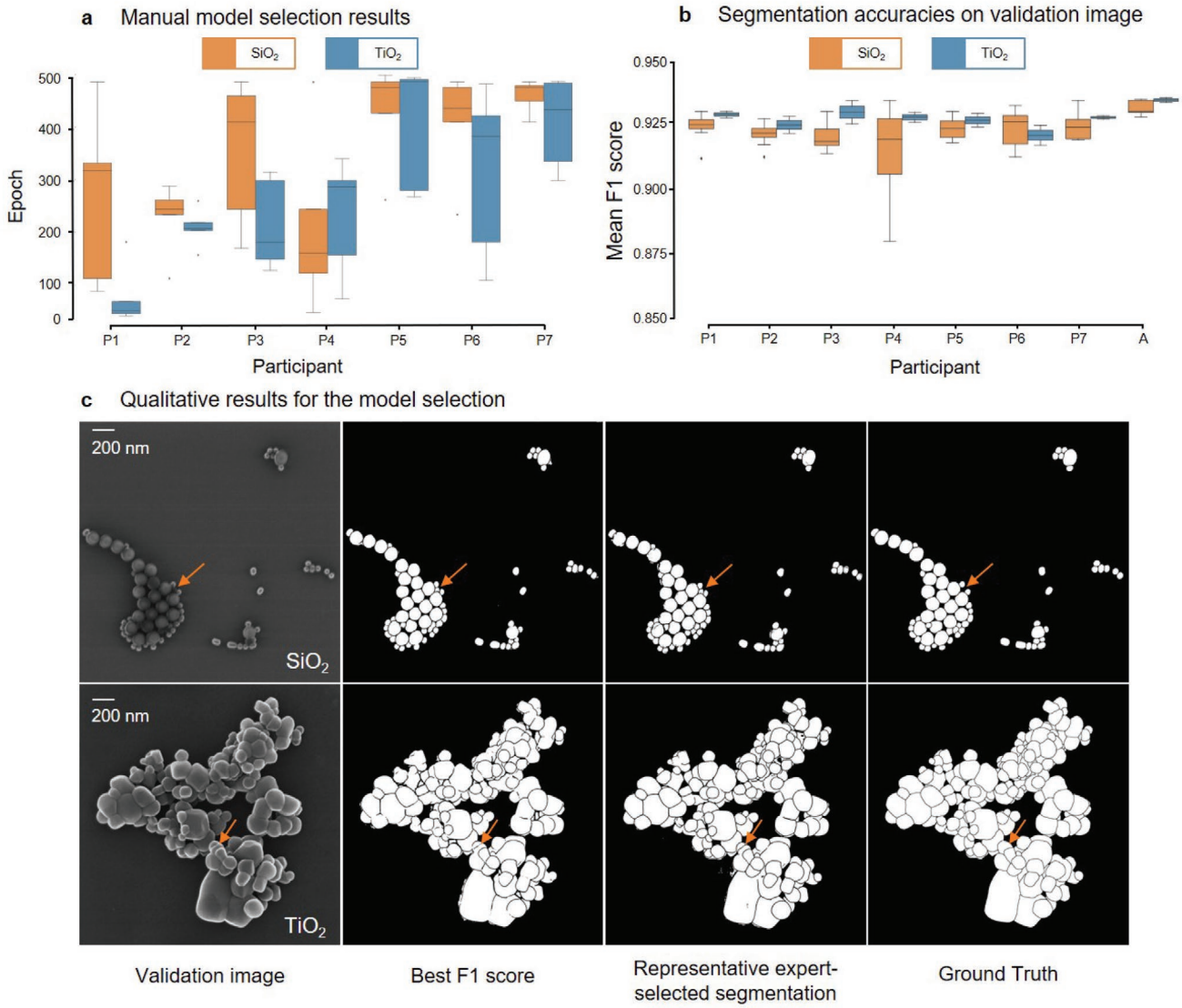

Figure 3. Comparison of two model selection approaches to find the best U-Net model based on real HIM validation images of SiO_2 and TiO_2 nanoparticles. While the "analytical" approach relies on a manual annotation of the validation image in order to find the best model, the second approach is based on the qualitative (visual) evaluation of the segmentation masks predicted by each model. To compare both methods, a user study with seven participants was conducted with non-experts (P1–P4) and experts (P5–6) for segmentation. According to its visual perception, each participant had to choose five best network models based on the model's segmentation performance for a real HIM validation image. a) Visualization of the five models, defined by the epoch number, that were selected by each of the study participants. b) F1 scores (*y*-axis) averaged over the five chosen models by each participant (P1–P7), using the manually annotated GT as reference segmentation for the validation image. Additionally, the diagram shows the mean F1 score for the "analytical" model selection approach, for comparison reasons also averaged over the five best models. c) Visualization of the validation images SiO_2 (top row) and TiO_2 (bottom row), the corresponding segmentation masks predicted by the model with the best F1 score selected by the "analytical" approach, the segmentation of a representative expert-selected model as well as the manual GT reference segmentation. Although the analytically selected model and the representative expert-selected model both provide an accurate overall segmentation, the expert-selected model is more suitable to separate individual particles (TiO_2) which is mandatory for statistical particle analysis. The orange arrows emphasize segmented particles by the U-Net that do not occur in the manually annotated ground truth image.

section. Moreover, we want to emphasize that a human-based model selection is preferable due to the time-consuming and error-prone process of manually annotating a validation image which is required for the "analytical" approach. Additionally, it was observed that training on synthetic images also resulted in a robust particle segmentation for the validation images, even for non-trivial cases as illustrated by orange arrows in Figure 3c, which highlight particles that are segmented as individual particles by the model but are missing in GT segmentation.

2.3. Segmentation Performance and Statistical Particle Analysis

Finally, after the model selection, the best model is used for the segmentation of all remaining real HIM images (Figure 1e). In

a Quantitative segmentation results

	Model	Precision	Recall	Accuracy	F1	AP	Warping Error [0.10^{-6}]	$\bar{N}_{particles}$	manual $\bar{N}_{particles}$
SiO_2	U-Net$_{real}$	**0.955**	**0.947**	**0.992**	**0.950**	**0.694**	**2169**	**81.88**	83
	U-Net$_{sim}$	0.945	0.912	0.990	0.930	0.614	6949	84.63	
TiO_2	U-Net$_{real}$	**0.975**	**0.913**	**0.971**	**0.943**	**0.529**	**23243**	128.40	152
	U-Net$_{sim}$	0.950	0.900	0.961	0.923	0.443	37738	**136.71**	

b Qualitative segmentation results

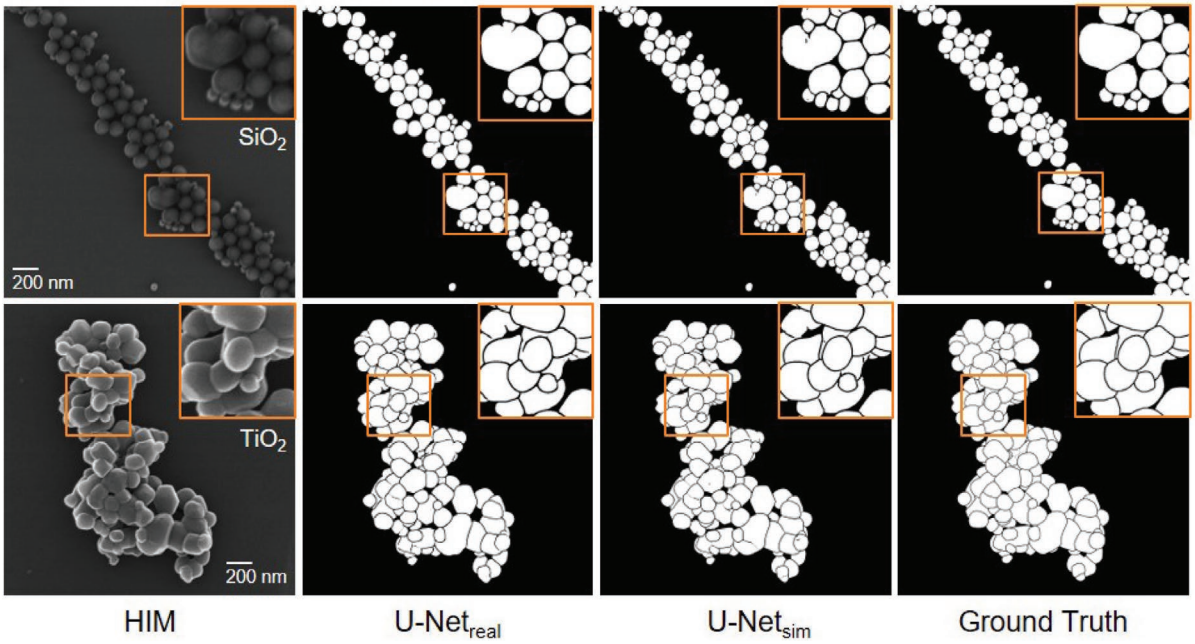

Figure 4. Segmentation results for the U-Net trained on real data (U-Net$_{real}$) compared with the performance of the U-Net which was trained on simulated data (U-Net$_{sim}$). a) Quantitative segmentation results using precision, recall, accuracy, F1 score, average precision (AP), and the warping error as evaluation metrics (see Section 4 for details). The average number of particles detected by the models is denoted as "$\bar{N}_{particles}$", while the number of individual particles that are present in the manual GT annotation is given by "manual $\bar{N}_{particles}$". b) Qualitative segmentation results for an exemplary selected real SiO_2 (top row) and TiO_2 (bottom row) HIM image. Both, the quantitative as well as qualitative results show almost perfect segmentations predicted by U-Net$_{real}$ and U-Net$_{sim}$ for both particle types compared to the manually annotated ground truth segmentation indicating the human-comparable segmentation capabilities of the models.

order to compare the U-Net that was trained on simulated data (U-Net$_{sim}$) with a baseline performance, we manually annotated all real SiO_2 and TiO_2 HIM images and trained a U-Net on real data (U-Net$_{real}$) for each particle type, respectively. Note that the manual annotation of SiO_2 and TiO_2 HIM images was performed by one individual each. Due to the small number of real HIM images (SiO_2: 9, TiO_2: 8) in comparison to the number of synthetically generated images (SiO_2: 180, TiO_2: 180), U-Net$_{real}$ was trained in a leave-one-out cross-validation[52] setup to assess its general segmentation performance. Further details on the training and the post-processing used to enhance the segmentation quality are provided in Section 4.

Figure 4a provides a quantitative comparison of U-Net$_{real}$ and U-Net$_{sim}$ for the SiO_2 and TiO_2 HIM data. According to precision, recall, accuracy, F1 score,[51] average precision[53] (AP), and the warping error[54] U-Net$_{real}$ (F1 score SiO_2: 0.950, F1 score TiO_2: 0.943) marginally outperforms the U-Net$_{sim}$ (F1 score SiO_2: 0.930, F1 score TiO_2: 0.923) for both particle types (see Figure S20, Supporting Information, for additional quantitative results based on the AP metric). However, both models provide nearly perfect segmentations. Additionally, for the SiO_2 data, the average number of individual particles ($\bar{N}_{particles}$) detected by U-Net$_{real}$ (81.88) and U-Net$_{sim}$ (84.63) matches almost ideally the average number of 83 particles (manual $\bar{N}_{particles}$) that are present in the manually annotated SiO_2 GT images. On the other hand, while the number of TiO_2 particles identified by U-Net$_{sim}$ (136.71) is more accurate than the number of particles detected by U-Net$_{real}$ (128.40), a considerable difference to

manual $\bar{N}_{particles}$ (152) can be observed. An explanation for this finding lies in the post-processing used to improve the segmentation output of the U-Nets (U-Net$_{real}$ as well as U-Net$_{sim}$). It removes small objects below a certain (pixel) area size using a morphological operation called "area opening", which results in a smaller amount of particles that are present in the segmentation masks after the post-processing (see Figures S26–S29, Supporting Information). Figure 4b visualizes the qualitative segmentation results for an exemplary selected real SiO_2 (top row) and TiO_2 (bottom row) HIM image. According to the visual impression, U-Net$_{real}$ and U-Net$_{sim}$ provide almost perfect segmentations in comparison to the manually annotated GT segmentation, which also reflect the quantitative results. A direct comparison of the qualitative segmentation results for all real HIM images is illustrated in the Figures S21–S29, Supporting Information, including particle size distributions for the GT reference as well as for the segmentations predicted by U-Net$_{real}$ and U-Net$_{sim}$. Both, the quantitative as well as qualitative results indicate the human-comparable segmentation capability of U-Net$_{real}$ and U-Net$_{sim}$. At this point, we want to highlight the extraordinary results produced by U-Net$_{sim}$. Although trained on synthetic data only, it is able to achieve impressive segmentation accuracies on real HIM data, which is not only comparable to a U-Net trained on real data but also to the segmentation capability of humans.

Due to the human-comparable segmentation accuracy of the U-Net$_{sim}$, it is possible to derive various features of real-world particles including their sizes, shapes, localization, or distributions using a connected component analysis (CCA). Therefore, we analyzed the segmentations of U-Net$_{sim}$ using a CCA to assess the particle size distributions of the real HIM SiO_2 and TiO_2 images as depicted in **Figure 5**a,b. A direct comparison of the particle size distributions for all HIM images based on the segmentations of U-Net$_{real}$, U-Net$_{sim}$, and the manual annotations is visualized in the Figures S21–S29, Supporting Information. While the top row shows raw HIM SiO_2 and TiO_2 images, the second row displays the corresponding post-processed U-Net$_{sim}$ segmentations. The third row visualizes an overlay of the original HIM images with a connected component labeling (CCL) based on the U-Net$_{sim}$ segmentation. In a CCL each color is associated with an individual particle. Note that due to the limited number of colors used in a CCL, neighboring particles may occur in the same color although being separated, individual particles. The CCL results for all HIM images, including

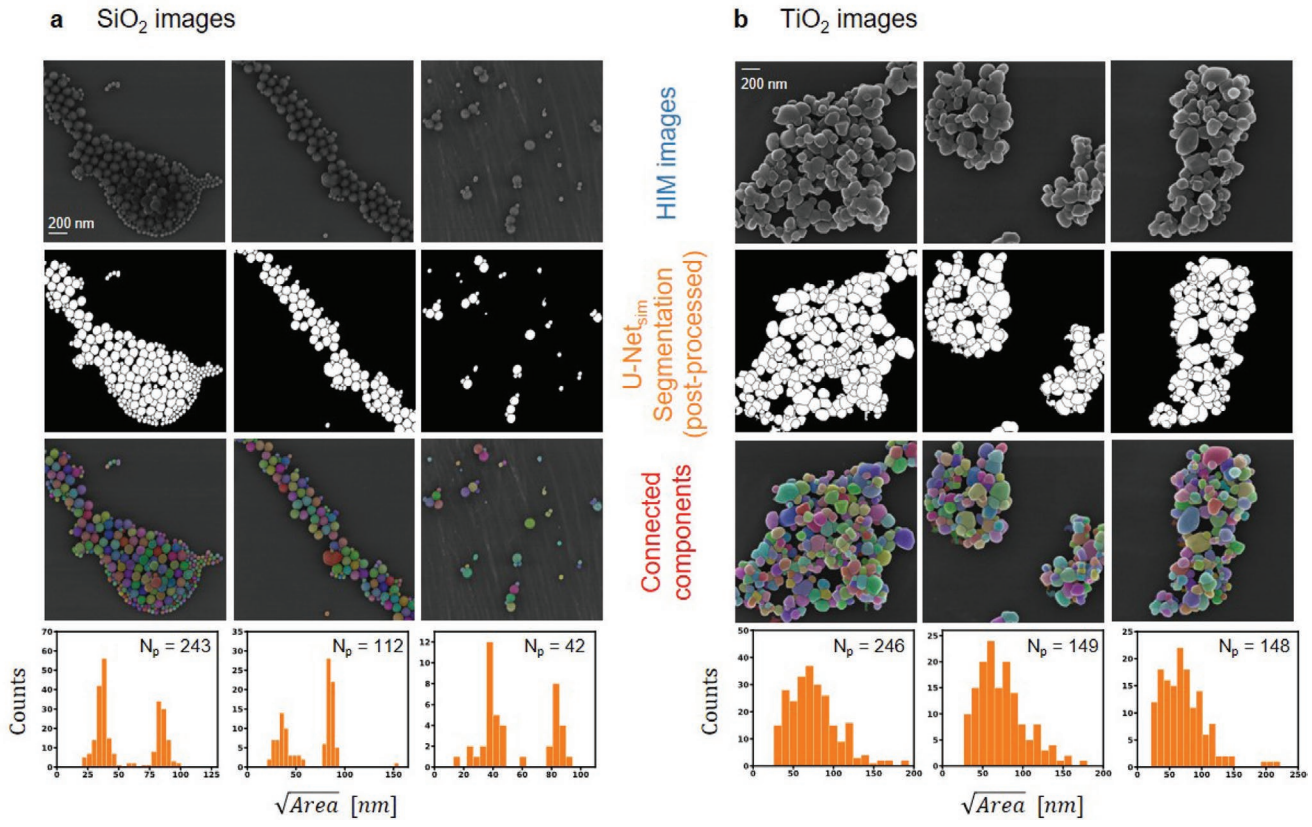

Figure 5. a) Analysis of SiO_2 and b) TiO_2 nanoparticle images. The top row shows raw HIM image data, whereas the second row visualizes the post-processed segmentation prediction of the U-Net$_{sim}$, which was trained on synthetic data only. The third row displays an overlay of the original HIM image with a connected component labeling (CCL) of the post-processed segmentation. In a CCL each color is associated with an individual particle. Note that due to the limited number of colors used in a CCL, neighboring particles may occur in the same color although being separated, individual particles. The bottom row provides a particle size distribution, based on the accurate U-Net$_{sim}$ prediction, for each HIM image which highlights the particle frequency with respect to the square-root of the particle area size denoted nanometer (nm). N_p represents the number of particles. While the SiO_2 histograms reveal a bimodal character, separating the two groups of small and large particles, TiO_2 histograms are characterized by a log-normal distribution.

the complex Ag data, are provided in Figures S30–S38, Supporting Information. Using the CCA, we can derive accurate particle size distributions (Figure 5, bottom row), which highlight the number of particles with respect to the square-root of the particle area size in nanometer (nm). While the SiO_2 histograms (Figure 5a) reveal a bimodal character, separating the two groups of small and large particles, TiO_2 histograms (Figure 5b) are characterized by a log-normal distribution with a modal value of ≈60 nm. Moreover, due to the accurately simulated substrate surface impurity in the synthetic training data, image artefacts as a result of sample preparation related dirt in real SiO_2 HIM images does not affect the segmentation accuracy of the U-Net_{sim} which has properly been trained to deal with this type of image features.

3. Conclusion

In summary, we propose a segmentation workflow for complex nanoparticles in high resolution microscopic images which relies on deep learning in combination with a semi-automated synthetic data generation pipeline based on photo-realistic rendering. With this approach an unlimited number of realistic synthetic images can be created with its respective error-free ground truth labels which subsequently can be used to effectively train a deep CNN. The so trained CNN can subsequently be used to accurately segment nanoparticles in microscope images. We have demonstrated the applicability of the workflow on experimental HIM data sets of SiO_2 and TiO_2 nanoparticles spread on a silicon (100) wafer surface. We showed that the segmentation accuracy of a state-of-the art deep CNN trained only on synthetic data was comparable to segmentations carried out by microscopy experts. Moreover, we have demonstrated quantitative and statistical analysis of microscopy imaged nanoparticles based on the automated deep learning CNN segmentation.

We are confident that the method presented in this work has the potential to solve the training data bottleneck and the annotation problem for automated image analysis approaches and paves the way toward a wider use of deep learning in a variety of microscopy applications. It permits the implementation of automated high-throughput particle segmentation and characterization methods for all sorts of applications based on microscopic images. This can be particularly relevant for studies related to nano-toxicology and other fields such as nano- and bio-medicine, consumer product efficacy testing, and anti-counterfeiting. The time-consuming, costly, and error-prone process of acquiring and manually annotating a representative amount of real data in order to use the power of deep learning is overcome by the proposed approach. Yet, although the presented method provides very accurate particle segmentations, especially for overlapping nanoparticles, it is not possible to extract the exact particle shapes. Therefore, in order to derive the exact shapes of overlapping particles which greatly determine their physiochemical properties, additional data is needed. Additionally, the semi-automated data generation procedure still relies on human input (i.e. for the design of all relevant 3D scene parameters, template particles, shaders, etc.) in order to achieve photo-realistic renders. However, with the recent achievements in differentiable rendering,[55,56] we assume that it will be possible to further automate this process.

4. Experimental Section

Sample Preparation and Data Acquisition: SiO_2 nanoparticles with two different diameters and food grade TiO_2 nanoparticles (E171) with a size distribution of 20 to 240 nm, both deposited on silicon chips (reference AGAR: G3390-10), were obtained from the "Laboratoire National de métrologie et d'Essais". The polyvinylpyrrolidon (PVP)-coated Ag nanowires (PL-AgW50-10mg) with an average diameter of 40–50 nm and a length of up to 50 µm were purchased from PlasmaChem GmbH (Berlin, Germany). A stable suspension was obtained by using a standard protocol.[57] For imaging, the Ag nanowires were sprayed onto a silicon wafer. Secondary electron images of the particles were obtained on a Zeiss ORION NanoFab equipped with a SIMS add-on[58,59] using the helium ion beam at an impact energy of 25 keV and a beam current of 0.5 pA. The images have a size of 2048 × 2048 px with a pixel scale of 1.0309 px nm^{-1}.

Scene and Synthetic Image Generation Process: For the synthetic data generation (Figure 1b), the open-source render software Blender[48] (Version 2.79.7) was used. In a first step, virtual template particles (see Figures S5 and S9, Supporting Information, for SiO_2 and TiO_2, respectively) were manually modeled based on the particle features in the real HIM images. In this context, the authors' want to note that due to lack of additional data that would allow for the reconstruction of a 3D shape of a single real particle, they were not able to perform a quantitative evaluation of how well the modeled template particles mathematically and statistically fit with actual data. In addition to the template particles, a template Blender scene was manually created that contains the substrate as well as the light source. Afterward, the parameters for the light source and the particle shaders were set manually to achieve a photo-realistic appearance of the particles in the rendered images. Once the correct parameters are set, the python application programming interface (API) of Blender was used to automatically generate a user-defined number of scenes filled with randomly distributed, duplicated and scaled template particles (more details are provided in Figures S4 and S8, Supporting Information, for SiO_2 and TiO_2, respectively). To cover the substrate surface impurity, which was present in some HIM images, a random dirt texture was generated for each scene. The dirt texture creation was based on the diamond-square algorithm,[60] also known as the "random midpoint displacement fractal". After all scenes were generated, each scene was rendered twice in an automated manner. While the first rendering process computes a photo-realistic HIM image, the shaders for the substrate and particles were changed for the second rendering process to obtain an error-free ground truth label. Additionally, aspects of domain randomization[61] were applied by randomly varying the brightness of each scene's light source in the photo-realistic renders (see Figures S11 and S12, Supporting Information, for SiO_2 and TiO_2, respectively). This way, brighter as well as darker particle images were generated, which increases the variance for the particle appearance in the synthetic data set. An automatically generated 3D scene with the corresponding label scenery is demonstrated in Figure S6 and S10, Supporting Information, for SiO_2 and TiO_2, respectively. Since the image quality produced by the renderer was adjustable and higher compared to the real HIM images in terms of sharpness, contrast, and noise level, the image quality of the synthetic, photo-realistic images was decreased by introducing aliasing effects in combination with additive Gaussian noise. The latter explains the evenly distributed synthetic background images in the t-SNE plots as shown in Figure 2. Aliasing was achieved by upsampling the rendered images from 507 × 507 px to 2031 × 2031 px using bilinear interpolation. Also, the synthetic labels were post-processed using binarization with a subsequent erosion (morphological operation) to strengthen the border area between particles.

In summary, for each particle type (SiO_2, TiO_2 and Ag) 180 corresponding scenes were generated with their respective synthetic

photo-realistic images and labels. The automated process of generating the particle scenes, rendering the photo-realistic images and the labels, as well as the post-processing amounted to ≈3.5 h for the SiO_2 particles, 5.5 h for the TiO_2 type and 16 h for the Ag images. In this context, a Nvidia GeForce GTX 1070 graphics processing unit (GPU) supported the scene renderings while the scene building process was performed by an Intel Xeon W-2102 central processing unit (CPU). The significant increase in time for the Ag images was caused by sophisticated Blender build-in post-processing edge enhancement algorithms in order to achieve accurate contour lines for the virtual objects, especially for the Ag nanowire images (see Figure S13, Supporting Information). The time spent to find and set the correct render settings including the correct light source parameters and shaders was not included in the calculation as it solely depends on the operators' level of experience for the used render software.

Pre-Processing and Data Augmentation: Prior to the training of the deep CNN, common data augmentation strategies were applied in order to improve the model performance[62] by increasing the variance and diversity of the training data. For this purpose, the pixel intensities were normalized within the range of [0, 1] and contrast limited adaptive histogram equalization[63] (CLAHE) was applied to enhance the contrast of each image in the training data set. Additionally, 90° image rotations, flipping, zoom, intensity changes, and Gaussian-distributed noise was used.

Convolutional Neural Network Architecture and Training: The U-Nets were implemented and trained on the TensorFlow[64] (version 1.12) framework. The original U-Net architecture proposed by Ronneberger et al.[28] as depicted in Figure 1c with the same number of feature maps for the encoder and decoder part was used. Solely the cropping operation was excluded from the skip connections and only one filter kernel was applied in the last 1×1 convolution layer at the end of the decoder part. Additionally, batch normalization[65] was applied after each convolution layer to reduce overfitting and stabilize the learning process. The weights were initialized with the method proposed by He et al.[66] As activation function, the rectified linear unit[67] (ReLU) was used. The training of the network was performed using stochastic gradient descent (SGD) with a patch size of 400×400 px and a mini-batch size of 2. Each image in the training batch was generated by sub-sampling a randomly chosen image from the synthetic data set. Prior to feeding the mini-batch into the network, data augmentation was applied as described in the previous subsection. The number of iterations was set to 150. U-Net_{sim} was trained on 180 synthetic images for 500 epochs. U-Net_{real} was trained in a leave-one-out cross-validation setup on nine SiO_2 and eight TiO_2 images, respectively. Due to the limited amount of real data available the number of epochs for U-Net_{real} was set to 300 and the best network state was chosen according to the minimal validation error. To compute a representative probability map as network output, a pixel-wise sigmoid was applied on its last feature map. The pixel-wise sigmoid is defined as:

$$p(x) = \frac{1}{1+e^{-x}} \tag{1}$$

where x denotes the pixel intensity. The binary segmentation mask was extracted with a threshold of 0.51 on the probability map. In other words, pixels that had a probability of more than 51% were classified as particle pixels. As loss function cross-entropy loss was applied. The networks (U-Net_{real} and U-Net_{sim}) were trained with the Adam[68] optimization algorithm using a constant learning rate of 0.001 and the default optimizer parameters in TensorFlow. For each particle type (SiO_2, TiO_2, and Ag), a separate U-Net model was trained on a Nvidia GeForce GTX 1070 GPU. The U-Net training for 500 epochs took ≈10 h.

Data Set Comparison and Visualization: The dimensionality reduction using t-SNE as illustrated in Figure 2c,d is based on features extracted by a VGG16,[25] a deep CNN, which is implemented in the PyTorch framework[69] and pre-trained on the ImageNet[30] data set. The ImageNet is a large data set for object detection tasks containing over 14 million images with more than 2000 categories. Since the VGG16 model was pre-trained on the ImageNet and achieved very high object detection accuracies,[25] it is assumed that the model had learned to extract relevant features to distinguish between various data sets. For data set comparison, image patches of the size of 144×144 px were processed by the VGG16, such that real and synthetic SiO_2 and TiO_2 images were represented as data points after the dimensionality reduction. The synthetic image patches were extracted by randomly sampling from all of the generated 180 synthetic images. Due to the small amount of HIM data, the real image patches were extracted sequentially from each of the nine HIM SiO_2 and eight TiO_2 images, respectively. This results in 3528 data points for SiO_2 particles (1764 real and synthetic image patches) and 3136 data points for the TiO_2 particles (1568 real and synthetic image patches). Subsequently, a principal component analysis (PCA)[70] was applied to reduce the dimensionality of each feature vector from 1000 to 33 components while covering 90% of its variance. Afterward, t-SNE[50] was used to further reduce the dimensionality and to facilitate visualization. The t-SNE algorithm was applied with a perplexity value of 50, a learning rate of 10, and a maximum number of 5000 iterations for the optimization. The squared Euclidean distance was used as distance metric. Both PCA and the t-SNE algorithm were applied using the scikit-learn[71] machine learning python module.

Time and Quality Assessment for Photo-Realism: The time it took to assess whether a synthetic image correctly matches with real data strongly depends on the user and was a very subjective aspect. However, based on the authors' experience, the qualitative as well as partially quantitative evaluation to check for photo-realism took a couple of minutes. During this process, specific aspects of the synthetic images were compared with real data. These specific aspects were for example, background noise, the characteristic edge-effect, and the background-particle or particle–particle transitions. While the quantitative comparison besides the t-SNE approach consisted of a simple extraction of intensity profiles to assess the similarity for noise and transitions (see Figure S7, Supporting Information, for example intensity profiles extracted from a synthetic and a real HIM SiO_2 nanoparticle image), the qualitative comparison focuses on the overall photo-realism in terms of shader selection and particle shape distributions.

Model Selection: In order to facilitate the understanding of the model selection step in the workflow (Figure 1d), the commonly used approach for selecting an appropriate model in machine learning is explained in the following. When a machine learning algorithm, that is, a deep CNN, is trained in a supervised fashion to solve a specific task, in general three stages must be passed in order to select the best model: training, validation, and testing. While training, the algorithm fits on the training data set in order to perform a specific task (i.e. segmentation). Specifically, during this phase the algorithm iteratively tunes its internal parameters, the so-called "weights" (weighted connections between neurons) in an artificial neural network, after each training step to improve its decision-making. In this context, a machine learning algorithm that is trained on a data set is termed "model". To prevent the model from overfitting on the training set, the validation data is used to assess the model performance after each training iteration using a pre-defined evaluation metric - the so-called "validation loss". The model, which achieves the minimal loss on the validation set is selected as the best model and is evaluated afterward on the previously unseen test data.

For the manual model selection, a user study was conducted with seven participants, consisting of three experts on segmentation and four non-experts. A U-Net was trained for 500 epochs on synthetic SiO_2 and TiO_2 data, respectively. For both particle types, a representative real HIM validation image was selected and manually annotated. The validation images for SiO_2 and TiO_2 as well as the corresponding manual GT annotations are visualized in Figure 3c. After training, each participant was asked to choose five models out of 500 (Model 1 to Model n_{epochs} in Figure 1d) that performed best based on the visual evaluation of the model's segmentation performances on the real validation image.

Post-Processing: The segmentation output of the U-Nets (U-Net_{real} and U-Net_{sim}) was post-processed in order to enhance the segmentation quality. For this purpose, an area opening on the binary segmentation mask with an area size of 400 px for the SiO_2 and 600 px for the TiO_2 images was utilized to remove noise that falls below a certain area size. Subsequently, a distance transform watershed[72,73] algorithm using the MorphoLibJ[73] plugin of the open-source ImageJ/Fiji[74] software

was applied to further separate particles with touching borders. In this context, the default settings were used with a dynamic of 20 for the TiO_2 particles and a normalized output combined with a dynamic of four for the SiO_2 images. The time required to analyze a single HIM image (U-Net prediction with subsequent post-processing) took ≈5–6 s per image. The U-Net_{sim} predictions together with the post-processed segmentation masks for all real HIM SiO_2 and TiO_2 are provided in Figures S14–S19, Supporting Information.

Metrics: For the quantitative evaluation several measures were utilized, including accuracy, precision, recall, and the Dice similarity index,[51] also known as F1 score. The metrics are defined as:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{4}$$

$$F1 = \frac{2 \cdot TP}{2 \cdot TP+FP+FN} \tag{5}$$

True positives (TP) and true negatives (TN) denote pixels, that were classified correctly as particles or background, respectively. False positives (FP) and false negatives (FN) were misclassified pixels that do not appear in the manual ground truth segmentation. The values for the measures vary from 0 to 1, while a value of 0 represents the worst possible accuracy whereas 1 denotes a perfect segmentation result. In literature,[75] a F1 score of 0.7 or higher was already considered as a good segmentation. Besides the F1 score, the warping error[54] and the average precision[53] (AP) was reported. The warping error is a metric that penalizes pixels that lead to topological disagreements but tolerates disagreements over boundary location. To compute the warping error, the Fiji Trainable Weka Segmentation tool was used.[76] The AP is an instance segmentation metric used to assess the quality of segmentation on a per-particle basis. It is the main metric by the Common Objects in Context[77] (COCO) segmentation challenge. This metric represents the area under the precision-recall curve and was computed for a range of pixel-wise intersection over union (IoU) thresholds in the interval [0.5, 0.95] in steps of 0.05. In this context, the IoU accounts for the overlap between a predicted particle instance segmentation and its corresponding GT instance label. The authors' refer to AP using a specific IoU threshold of 0.5 or 0.75 as AP_{50} and AP_{75}, respectively. Additional instance segmentation results based on the AP metric are provided in Figure S20, Supporting Information.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation program (Grant agreement No. 720964) as well as from the European Research Council (ERC) (Grant agreement No. 810316).

Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Author Contribution

L.M. conceived the idea and designed and implemented the workflow together with D.W. L.K., N.G., P.P., F.V., A.I., A.M., and S.C. contributed to the design of the data simulation and the workflow. T.W., J.N.A., and P.P. defined the needs for image segmentation. J.N.A and O.D.C. recorded the SE images on the HIM. I.N. contributed to the writing of the introduction section and commented on the manuscript. Y.H. and C.J. helped with the evaluation of the results. L.M. conducted the simulations, implemented and trained the deep learning models, and evaluated the results. L.M., N.G., L.K., F.V., C.J., I.N., Y.H., A.M., and S.C. discussed the results and suggested improvements. L.M. and S.C. iterated the first version of the manuscript and all authors contributed to the final text.

Data Availability Statement

The data that supports the findings of this study are available in the supplementary material of this article.

Keywords

helium ion microscopy, image analysis, machine learning, nanoparticles, segmentation, toxicology

Received: February 25, 2021
Revised: April 17, 2021
Published online: May 3, 2021

[1] M. E. Vance, T. Kuiken, E. P. Vejerano, S. P. McGinnis, M. F. Hochella Jr, D. Rejeski, M. S. Hull, *Beilstein J. Nanotechnol.* **2015**, *6*, 1769.
[2] S. E. Lohse, C. J. Murphy, *J. Am. Chem. Soc.* **2012**, *134*, 15607.
[3] Q. Sun, Y. A. Wang, L. S. Li, D. Wang, T. Zhu, J. Xu, C. Yang, Y. Li, *Nat. Photonics* **2007**, *1*, 717.
[4] E. C. Dreaden, A. M. Alkilany, X. Huang, C. J. Murphy, M. A. El-Sayed, *Chem. Soc. Rev.* **2012**, *41*, 2740.
[5] Z. Zhang, J. Wang, X. Nie, T. Wen, Y. Ji, X. Wu, Y. Zhao, C. Chen, *J. Am. Chem. Soc.* **2014**, *136*, 7317.
[6] M. Bundschuh, J. Filser, S. Lüderwald, M. S. McKee, G. Metreveli, G. E. Schaumann, R. Schulz, S. Wagner, *Environ. Sci. Eur.* **2018**, *30*, 1.
[7] A. Kongkanand, K. Tvrdy, K. Takechi, M. Kuno, P. V. Kamat, *J. Am. Chem. Soc.* **2008**, *130*, 4007.
[8] M. A. Mackey, M. R. Ali, L. A. Austin, R. D. Near, M. A. El-Sayed, *J. Phys. Chem. B* **2014**, *118*, 1319.
[9] S. Mülhopt, S. Diabaté, M. Dilger, C. Adelhelm, C. Anderlohr, T. Bergfeldt, J. Gómez de la Torre, Y. Jiang, E. Valsami-Jones, D. Langevin, D. Langevin, I. Lynch, E. Mahon, I. Nelissen, J. Piella, V. Puntes, S. Ray, R. Schneider, T. Wilkins, C. W. H.-R. Paur, *Nanomaterials* **2018**, *8*, 311.
[10] G. Hlawacek, V. Veligura, R. van Gastel, B. Poelsema, *J. Vac. Sci. Technol., B: Nanotechnol. Microelectron.: Mater., Process., Meas., Phenom.* **2014**, *32*, 020801.
[11] Z. Z. Wang, *IEEE Trans. Ind. Inf.* **2016**, *12*, 962.
[12] K. Al-Dulaimi, J. Banks, I. Tomeo-Reyes, V. Chandran, in Proc. - Int. Conf. on Pattern Recognition, IEEE, Piscataway, NJ **2017**, 81.
[13] C. Park, J. Z. Huang, J. X. Ji, Y. Ding, *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 1.
[14] M. Mirzaei, H. K. Rafsanjani, *Micron* **2017**, *96*, 86.
[15] Y. Meng, Z. Zhang, H. Yin, T. Ma, *Micron* **2018**, *106*, 34.
[16] C. R. Laramy, K. A. Brown, M. N. O'Brien, C. A. Mirkin, *ACS Nano* **2015**, *9*, 12488.

[17] S. Zafari, T. Eerola, J. Sampo, H. Kälviäinen, H. Haario, *IEEE Trans. Image Process.* **2015**, *24*, 5942.
[18] C. R. Arroyo, A. Debut, A. V. Vaca, C. Stael, K. Guzman, B. K. L. Cumbal, *Biol. .Med.* **2016**, *8*, 1.
[19] M. van Sebille, L. J. van der Maaten, L. Xie, K. Jarolimek, R. Santbergen, R. A. van Swaaij, K. Leifer, M. Zeman, *Nanoscale* **2015**, *7*, 20593.
[20] E. A. Grulke, X. Wu, Y. Ji, E. Buhr, K. Yamamoto, N. W. Song, A. B. Stefaniak, D. Schwegler-Berry, W. W. Burchett, J. Lambert, A. J. Stromberg, *Metrologia* **2018**, *55*, 254.
[21] E. McLeod, T. U. Dincer, M. Veli, Y. N. Ertas, C. Nguyen, W. Luo, A. Greenbaum, A. Feizi, A. Ozcan, *ACS nano* **2015**, *9*, 3265.
[22] D. Riccio, N. Brancati, M. Frucci, D. Gragnaniello, *IEEE J. Biomed. Health Inf.* **2019**, *23*, 437.
[23] Y. Lecun, Y. Bengio, G. Hinton, *Nature* **2015**, *521*, 436.
[24] K. He, G. Gkioxari, P. Dollar, R. Girshick, *Proc. IEEE Int. Conf. Computer Vision*, IEEE, Piscataway, NJ **2017**, 2980.
[25] K. Simonyan, A. Zisserman, *arXiv:1409.1556*, **2014**.
[26] K. He, X. Zhang, S. Ren, J. Sun, in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2016**, 770.
[27] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, *Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2016**, 779.
[28] O. Ronneberger, P. Fischer, T. Brox, in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, Springer, Berlin **2015**, 234–241.
[29] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, O. Ronneberger, *Nat. Methods* **2019**, *16*, 67.
[30] J. Deng, W. Dong, R. Socher, L. Li, K. Li, L. Fei-Fei, *Proc. IEEE Computer Society Conf. Computer Vision Pattern Recognition*, IEEE, Piscataway, NJ **2009**, 248.
[31] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van DenDriessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, *Nature* **2016**, *529*, 484.
[32] A. Krizhevsky, I. Sutskever, G. E. Hinton, *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097.
[33] I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, H. Suchowski, *Light: Sci. Appl.* **2018**, *7*, 1.
[34] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, M. Soljačić, *Sci. Adv.* **2018**, *4*, eaar4206.
[35] J. Madsen, P. Liu, J. Kling, J. B. Wagner, T. W. Hansen, O. Winther, J. Schiøtz, *Adv. Theory Simul.* **2018**, *1*, 1800037.
[36] Y. Zhu, Q. Ouyang, Y. Mao, *BMC Bioinf.* **2017**, *18*, 1.
[37] F. Wang, H. Gong, G. Liu, M. Li, C. Yan, T. Xia, X. Li, J. Zeng, *J. Struct. Biol.* **2016**, *195*, 325.
[38] A. Cid-Mejías, R. Alonso-Calvo, H. Gavilán, J. Crespo, V. Maojo, *Comp. Methods Prog. Biomed.* **2021**, *202*, 105958.
[39] A. G. Okunev, M. Y. Mashukov, A. V. Nartova, A. V. Matveev, *Nanomaterials* **2020**, *10*, 1285.
[40] J. P. Horwath, D. N. Zakharov, R. Megret, E. A. Stach, *npj Comput. Mater.* **2020**, *6*, 1.
[41] L. Yao, Z. Ou, B. Luo, C. Xu, Q. Chen, *ACS Cent. Sci.* **2020**, *6*, 1421.
[42] P. Muneesawang, C. Sirisathitkul, *J . Nanomater.* **2015**, *2015*, 790508.
[43] M. Ilett, J. Wills, P. Rees, S. Sharma, S. Micklethwaite, A. Brown, R. Brydson, N. Hondow, *J. Microscopy* **2020**, *279*, 177.
[44] X. Wang, J. Li, H. D. Ha, J. Dahl, T. Head-Gordon, P. Alivisatos, *JACS Au* **2021**, *1*, 316.
[45] B. Lee, S. Yoon, J. W. Lee, Y. Kim, J. Chang, J. Yun, J. C. Ro, J.-S. Lee, J. H. Lee, *ACS Nano* **2020**, *14*, 17125.
[46] A. K. Maier, C. Syben, B. Stimpel, T. Würfl, M. Hoffmann, F. Schebesch, W. Fu, L. Mill, L. Kling, S. Christiansen, *Nat. Mach. Intell.* **2019**, *1*, 373.
[47] S. J. Pan, Q. Yang, *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345.
[48] Blender Foundation, *Home of the Blender Project - Free and Open 3D Creation Software* **2018**, http://www.blender.org.
[49] R. Hemelings, B. Elen, I. Stalmans, K. Van Keer, P. De Boever, M. B. Blaschko, *Computer. Med. Imaging Graphics* **2019**, *76*, 101636.
[50] L. v. d. Maaten, G. Hinton, *J. Mach. Learning Res.* **2008**, *9*, 2579.
[51] L. R. Dice, *Ecology* **1945**, *26*, 297.
[52] M. W. Browne, *J. Math. Psychol.* **2000**, *44*, 108.
[53] M. Everingham, S. A. Eslami, L. Van Gool, C. K. Williams, J. Winn, A. Zisserman, *Int. J. Comput. Vision* **2015**, *111*, 98.
[54] V. Jain, B. Bollmann, M. Richardson, D. R. Berger, M. N. Helmstaedter, K. L. Briggman, W. Denk, J. B. Bowden, J. M. Mendenhall, W. C. Abraham, K. M. Harris, N. Kasthuri, K. J. Hayworth, R. Schalek, J. C. Tapia, J. W. Lichtman, H. S. Seung, In *2010 IEEE Computer Society Conf .on Computer Vision and Pattern Recognition*. IEEE,, Piscataway, NJ **2010**, 2488.
[55] M. Nimier-David, D. Vicini, T. Zeltner, W. Jakob, *ACM Trans. Graphics* **2019**, *38*, 203.
[56] G. Loubet, N. Holzschuch, W. Jakob, *ACM Trans. Graphics* **2019**, *38*, 1.
[57] G. M. DeLoid, J. M. Cohen, G. Pyrgiotakis, P. Demokritou, *Nat. Protoc.* **2017**, *12*, 355.
[58] T. Wirtz, P. Philipp, J. Audinot, D. Dowsett, S. Eswara, *Nanotechnology* **2015**, *26*, 434001.
[59] T. Wirtz, O. De Castro, J.-N. Audinot, P. Philipp, *Annu. Rev. Anal. Chem.* **2019**, *12*, 523.
[60] A. Fournier, D. Fussell, L. Carpenter, *Commun. ACM* **1982**, *25*, 371.
[61] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, In *2017 IEEE/RSJ Int. Conf.on Intelligent Robots and Systems*, IEEE, Piscataway, NJ **2017**, 23–30.
[62] J. Wang, L. Perez, *arXiv:1712.04621*, **2017**.
[63] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, K. Zuiderveld, *Comput. Vision Graph. Image Process.* **1987**, *39*, 355.
[64] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, et al., *arXiv:1603.04467*, **2016**.
[65] S. Ioffe, C. Szegedy, *arXiv:1502.03167*, **2015**.
[66] K. He, X. Zhang, S. Ren, J. Sun, in *Proc. of the IEEE Int. Conf. on Computer Vision*, IEEE, Piscataway, NJ **2015**1026–1034.
[67] X. Glorot, A. Bordes, Y. Bengio, in *Proc. of the Fourteenth Int. Conf. on Artificial Intelligence and Statistics*, **2011**315–323.
[68] D. P. Kingma, J. Ba, *arXiv:1412.6980*, **2014**.
[69] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, *arXiv preprint arXiv:1912.01703* **2019**.
[70] S. Wold, K. Esbensen, P. Geladi, *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37.
[71] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Mach. Learning Res.* **2011**, *12*, 2825.

[72] P. Soille, L. Vincent, in *Visual Communications and Image Processing'90: Fifth in a Series*, vol. *1360*, International Society for Optics and Photonics, Bellingham, WA **1990**, 240–250.
[73] D. Legland, I. Arganda-Carreras, P. Andrey, *Bioinformatics* **2016**, *32*, 3532.
[74] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676.
[75] M. Ghafoorian, N. Karssemeijer, T. Heskes, I. W. van Uden, C. I. Sanchez, G. Litjens, F.-E. de Leeuw, van B. Ginneken, E. Marchiori, B. Platel, *Sci. Rep.* **2017**, *7*, 5110.
[76] I. Arganda-Carreras, V. Kaynig, C. Rueden, K. W. Eliceiri, J. Schindelin, A. Cardona, H. S. Seung, *Bioinformatics* **2017**, *33*, 2424.
[77] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C. L. Zitnick, In *European Conf. on Computer vision*, Springer, Berlin **2014**, 740–755.