Automated crystal system identification from electron diffraction patterns using multiview opinion fusion machine learning

Jie Chen[1], Hengrui Zhang[1], Carolin B. Wahl[2,3], Wei Liu[4], Chad A. Mirkin[2,3,5], Vinayak P. Dravid[2,3], Daniel W. Apley[4], & Wei Chen[1]

[1]Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA
[2]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA
[3]International Institute for Nanotechnology, Northwestern University, Evanston, IL 60208, USA
[4]Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA
[5]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA

Abstract: A bottleneck in high-throughput nanomaterials discovery is the pace at which new materials can be structurally characterized. Although current machine learning (ML) methods show promise for the automated processing of electron diffraction patterns (DPs), they fail in high-throughput experiments where DPs are collected from crystals with random orientations. Inspired by the human decision-making process, a framework for automated crystal system classification from DPs with arbitrary orientations was developed. A convolutional neural network was trained using evidential deep learning, and the predictive uncertainties were quantified and leveraged to fuse multiview predictions. Using vector map representations of DPs, the framework achieves an unprecedented testing accuracy of 0.94 in the examples considered, is robust to noise, and retains remarkable accuracy using experimental data. This work highlights the ability of ML be used to accelerate experimental high-throughput materials data analytics.

1. Introduction

Progress in the field of automated materials discovery relies on high-throughput experimentation geared towards the rapid identification of key structure-function relationships that define material properties[1] (**Fig. 1**a). Recently, polymer pen lithography (PPL)[2], a cantilever-free approach to synthesizing nanoparticle megalibraries, has enabled the synthesis of billions of unique nanomaterials in a single experiment[3, 4]. Nevertheless, the throughput of this approach, which relies on 90,000 pyramidal tips to independently synthesize distinct materials, far outpaces the rate of characterization (via screening) and analysis. Four-dimensional scanning transmission electron microscopy (4D-STEM)[5, 6] is ideally suited for identifying crystal structures on the nanometer scale at arbitrary locations on a specimen, including specimens that are nanoparticles in a megalibrary on a substrate. 4D-STEM allows spatially resolved electron diffraction patterns (DPs) to be collected. The structural information gleaned from such DPs is critical in identifying nanoparticle structure-function relationships (because symmetry determines the number of independent components of property tensors). Experimental structure data from DPs are also amenable to conversion to computation-ready data as inputs for computational modeling. However, the identification of the type of crystal systems (e.g., triclinic, monoclinic, etc.) from large 4D-STEM datasets is not trivial; it is impossible to

manually extract structural information from every single DP, due to the volume and complexity of the data and data processing.

Autonomous experimentation[7], which integrates machine learning (ML)-driven interpretation of data collected in previous measurements to guide subsequent measurements, requires, in the case of DPs, the rapid, automated extraction of crystallographic information in real-time. Indeed, ML methods are capable of extracting information from complex and unstructured datasets and have proven useful in the analysis of large volumes of data for pattern recognition[8, 9, 10] and crystal structure identification from DPs[11, 12, 13]. Convolutional neural networks (CNNs) are among the most widely used ML methods for evaluating diffraction data in image format, and they have been applied to crystal structure identification from two-dimensional (2D)[8, 11, 14, 15, 16, 17] and one-dimensional (1D) diffraction data[9, 18, 19, 20, 21, 22, 23]. Still, the use of CNN models in the context of DPs is limited; the data must be collected from specific low-order zone axes (high-symmetry orientations in a crystal that contain the most information) in order to maximize the amount of information in the DP and facilitate its analysis by both humans and machines. However, for these models to be accurate, the crystal DPs must be obtained from the same zone axes as the training data, which is not possible in every experimental system. For example, it is not feasible to physically tilt samples to specific zone axes in most high-throughput experimental workflows, including in the context of nanoparticle megalibraries, to collect DPs with known zone axes. Thus, the zone axes of the DPs obtained from 4D-STEM experiments on nanoparticle megalibraries are unknown, and so cannot be matched with the appropriate training data. Moreover, a single DP usually provides only 2D (projected) crystallographic information. To obtain a full description of a material in three dimensions (3D), it is typically necessary to observe the crystal from different orientations via tilting experiments (**Fig. 1**a). In light of these issues, in order to implement autonomous ML-based methods for high-throughput structural screening of vast nanomaterial libraries, it is necessary to develop ML models that can interpret DP information from arbitrary orientations to permit the direct reconstruction of 3D structural information without the need to physically tilt the sample.

When human experts determine the structure of a single crystalline sample using electron diffraction, they first tilt the sample to some low-order zone axis and record a DP, and then tilt the sample to other low-order zone axes to collect additional DPs until they are confident in their conclusion about the crystal structure[24]. Considering that artificial intelligence (AI) aims to mimic human intelligence, it follows that researchers would seek to develop ML algorithms guided by the operational principles of human brains[25, 26, 27] for crystal system identification. Herein, we describe a multiview opinion fusion machine learning (MVOF-ML) framework for crystal system classification with DPs from multiple arbitrary zone axes that is compatible with the nanoparticle megalibrary platform and that does not require sample titling. Given the spotty distribution of diffraction peaks (Bragg disks; particularly for DPs from random, off-axis orientations), vectorized location and intensity information is used as the representation of the DPs instead of the traditional images. We show that this more compact representation of diffraction information leads to greater ML efficiency; a multistream CNN customized from PointNet[28] is utilized to handle this form of point cloud data. The CNN model takes on the role of a human expert making predictions (or forming opinions) from DPs with arbitrary zone axes. Our crystal system classification and its corresponding uncertainty are determined via a sequential decision-making process based on a series of sets of partial information

from multiple views. The classifications ("opinions") and their related uncertainties are fused from different views, and these fused uncertainties reflect the confidence in making the final classification decision and guide whether to continue versus stop collecting additional DP input data. Our MVOF-ML framework achieves a high testing accuracy (percentage of correct predictions) of 0.94 on simulated data and is shown to be robust even in the presence of different sources of noise designed to mimic real experimental conditions. Finally, this framework is applied to experimental data obtained from 4D-STEM datasets of nanoparticle libraries (representative subsets of a megalibrary synthesized by Dip-Pen Nanolithography (DPN)[29], the precursor to PPL, where it performs with similar accuracy. Rather than applying off-the-shelf ML methodology, we built a custom CNN in which the convolution operations are modified in order to arrive at kernels that can be merged to incorporate specific physical principles embedded in the data (for example, radial and angular adjacency information, which encodes symmetry information in a DP). The success of the approach highlights the value of developing physics-informed network ML architectures and incorporating them into frameworks that mirror human scientific reasoning processes to accelerate data analysis and processing across multiple fields.

2. Results

2.1 A MVOF-ML framework for automated crystal system identification from DPs with arbitrary zone axes

A multiview opinion fusion machine learning (MVOF-ML) framework for automated crystal system identification in 4D-STEM is developed (**Fig. 1**b), wherein a series of DPs is analyzed to make a final decision on the crystal system. To realize this framework, a CNN is trained to predict the probability that a crystal belongs to each crystal systems (the "opinion") and then extract the model's confidence for that opinion (i.e., the uncertainty quantification)[30, 31, 32]. Next, the opinions are automatically fused under consideration of uncertainty during the accumulation of DPs. A final decision on the classification of the crystal system is made once the confidence is above (or the uncertainty is below) a desired threshold.

The MVOF-ML framework utilizes a custom CNN architecture that incorporates the physics of the system captured implicitly by the DP (**Fig. 2**). Our CNN uses vector maps (**Fig. 2**a), which contain point cloud data with the locations (r, θ) and intensities (I) of the Bragg disks, as inputs, as opposed to images used in other ML models. There are three branches before a fully connected layer, for which the inputs are the vector maps corresponding to the same DP, sorted in different orders r-θ-I, θ-r-I, and I-r-θ, respectively (**Fig. 2**b) (r-θ-I represents the vector map sorted first by r, then θ, and then intensity I, for example). The sorting of the vector map affects the ability of the following convolutional layer to learn how adjacent Bragg disks, which encode structural symmetry, are correlated. The kernels defining adjacency in the convolutional layers are illustrated by the rectangles inside the DPs (**Fig. 2**b), and they are also shown in black squares containing two rows in the vector maps (**Fig. 2**a). The PointNet[28] framework has been widely used for deep learning on point sets for 3D classifications, thus it is expected that this CNN would work well as a starting point with the 3D point representation (r, θ and I) of Bragg disks in a DP. We modified this standard PointNet CNN architecture according to the physics encapsulated in the point cloud inputs (r, θ, and I) (**Fig. 2**c), using appropriately structured

kernels to incorporate the physical adjacency information directly into the network architecture. This modified PointNet CNN is then applied to each input vector map that is gleaned from subsequent DPs. At the end, an output layer contains seven neurons corresponding to the seven crystal systems (**Fig. 2**d).

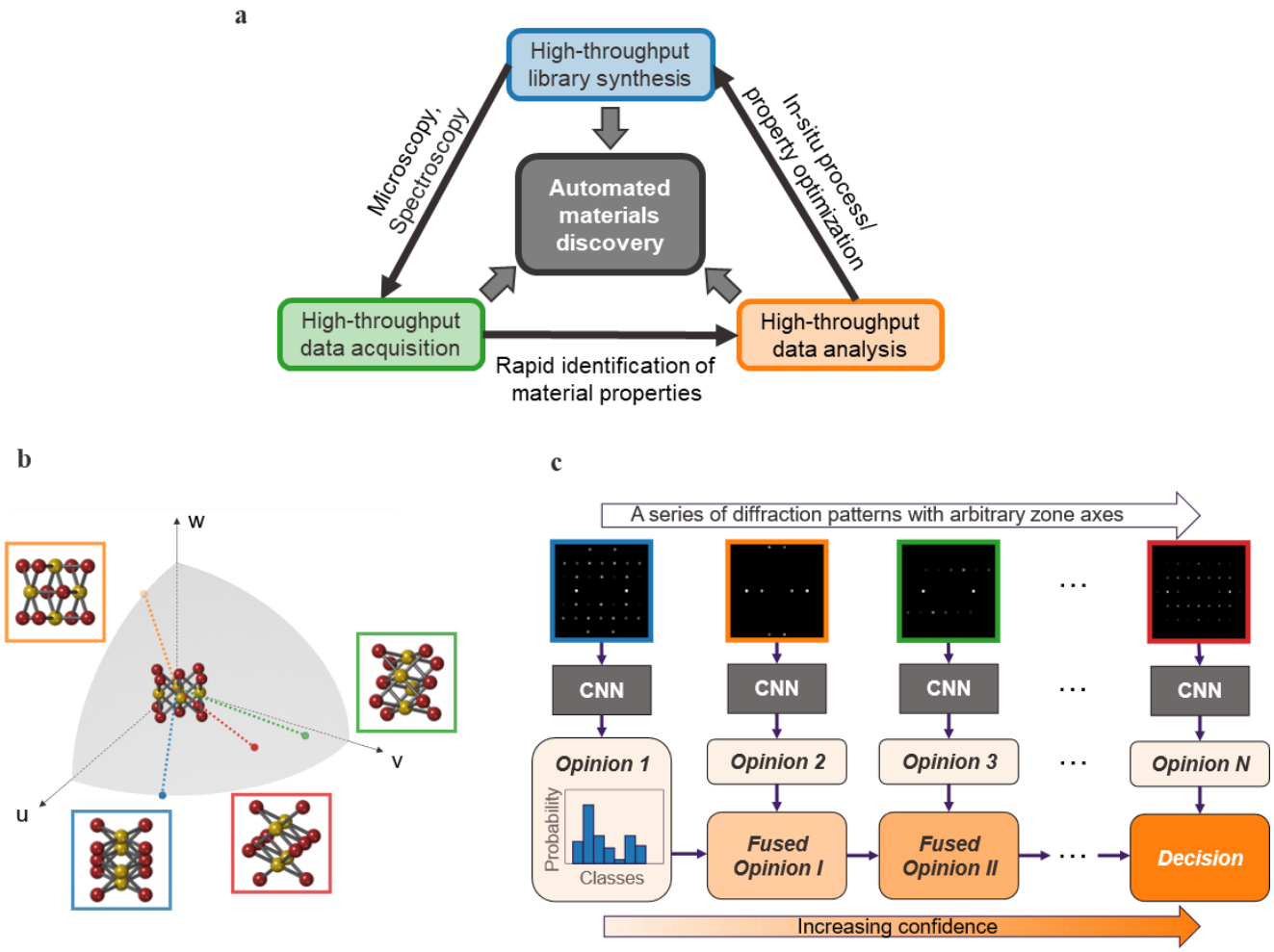


Fig. 1 Automated crystal structure identification from diffraction patterns with arbitrary zone axes. a, While high-throughput synthesis of nanoparticles is achievable, their characterization and analysis remain challenging due to the lack of high-throughput screening methods and data analytics technologies. These tools are necessary to comprehend the structure-function relationships that define material properties, which are vital for autonomously discovering new materials with desired properties. This paper aims to develop machine learning methods for high-throughput analysis. **b**, A crystal can be observed from different orientations. Each orientation corresponds to a zone axis (orientation) of the crystal. From each orientation, a DP can be obtained that is a 2D projection in reciprocal space of the 3D crystal structure. **c**, A framework for automated crystal structure identification from DPs with arbitrary zone axes. A CNN is trained using simulation data from multiple orientations. After training, the DPs from different zone axes of the crystal are consecutively used as inputs to a CNN. For each DP, the CNN model predicts the probabilities of different crystal systems and forms an opinion, and the uncertainty is quantified for each prediction. The individual opinions are iteratively fused to form the final decision. The DPs in **Fig. 1**c correspond to crystal orientations in **Fig. 1**b with the same frame color. The increase of opacity in option/decision blocks in the bottom two rows in **Fig. 1**c indicates the increase of confidence (or the decrease of uncertainty).

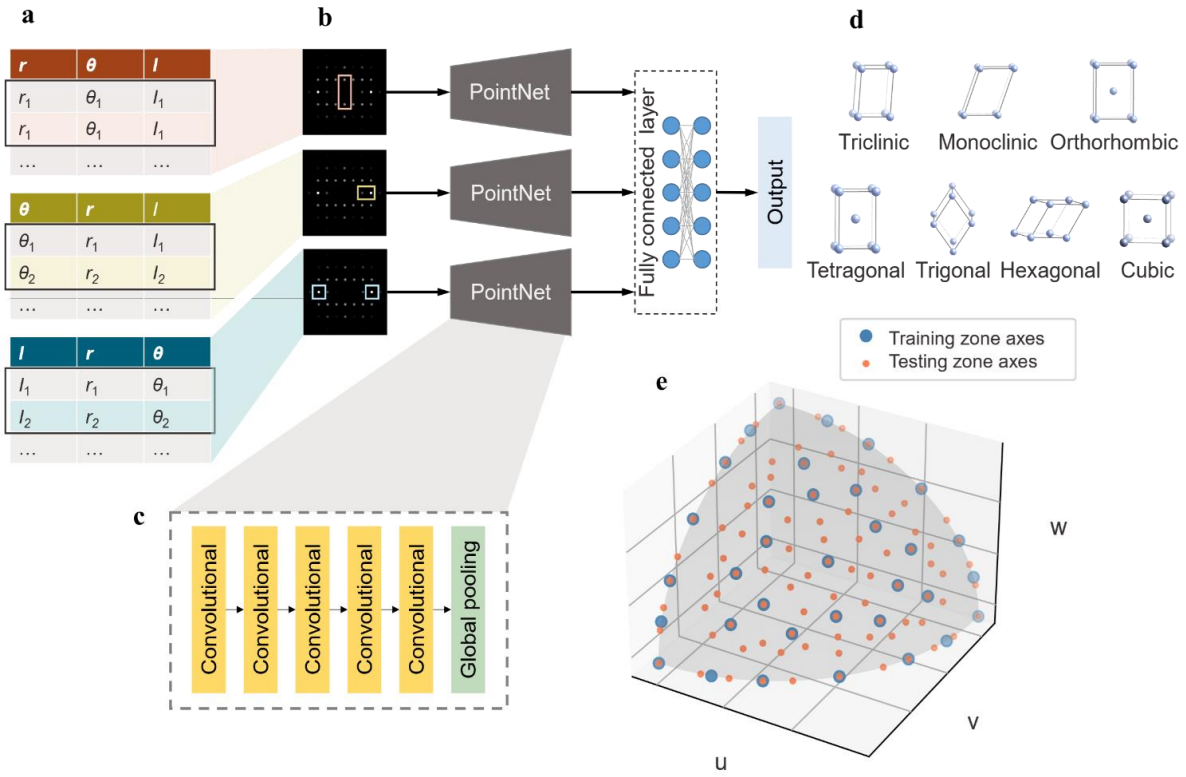


Fig. 2 Convolutional neural network architecture and dataset. a, **b**, **c**, and **d** are CNN architecture used for both training and testing, and **e** is the training and testing datasets. **a**, The DPs are represented in the form of vector maps, which are tabular data in which there are three columns (rows correspond to the identified Bragg disks in the DP): the distance to the center Bragg disk *r*, the angle θ from the horizontal, and the intensity *I*. The vector maps are sorted according to *r*-θ-*I*, θ-*r*-*I*, and *I*-*r*-θ. **b**, The three sorted vector maps, corresponding to different kernel types, are used as inputs to the three PointNets. The outputs of the PointNets are concatenated and passed to the fully connected layers. The output layer contains seven neurons representing seven crystal systems (triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, and cubic). **c**, The PointNet contains several convolutional layers followed by the global pooling layers. **d**, The seven crystal systems as outputs. **e** Distribution of training and testing zone axes in 3D space. Space is represented by a 1/8 sphere. Thirty-four zone axes were sampled for training and 94 were sampled for testing.

In the multistream CNN, the input to each branch consists of the same DP information; however, the vector map containing the point cloud data is sorted in three different ways to achieve three different kernels (SI Fig. S1). We tested this multistream CNN with the combination of the three kernel types and compared its accuracy to the accuracy of single-stream CNNs with only one of three kernel types. The squares in the first three DP images illustrate a kernel with size 2. The results show that the accuracy is the highest when we combine the three kernel types in the multistream CNN (SI Fig. S2), and thus we subsequently adopt this architecture. By utilizing multiple kernel types, the CNN is able to more effectively home in on various physical features of the different adjacent Bragg disks and exploit the actual physical information encoded in the DPs to learn and make accurate predictions.

Traditionally, CNNs perform a classification task by outputting probabilities of particular crystal systems (i.e., opinions), and a classification is usually achieved by applying a SoftMax activation

function at the output layer and minimizing the cross-entropy loss function. Conventionally, it is equivalent to minimizing the cross-entropy loss function and maximizing the likelihood, assuming a multinomial distribution with the parameters as class probabilities. However, such a setup does not account for the confidence of forming an opinion (i.e., the uncertainty). Uncertainty quantification is desired in our MVOF-ML framework since it is used to guide the fusion of the opinions. Therefore, we trained our CNN using the principle of evidential deep learning[33, 34]. The uncertainty is explicitly modeled via

$$u + \sum_{k=1}^{K} b_k = 1, \quad (1)$$

where u is the uncertainty mass, b_k is the belief mass for the k-th class, and K is the total number of classes. In evidential learning, the neural network forms opinions as Dirichlet distributions that can model the probability distribution of each class. The parameters of the Dirichlet distribution are calculated from the neural network outputs. The loss function is the marginal likelihood calculated using the Dirichlet distribution.

The multiview opinions are automatically fused sequentially, two opinions at a time, to capture how live experimental data will be presented to the CNN. Opinions from the first two different views are fused with uncertainty as outlined above. After that, a fused opinion is formed, which is then fused with the third opinion, and so on. During the opinion fusion, uncertainties are also fused. This process stops when the uncertainty is below a desired threshold. The Dempster-Shafer theory[35, 36, 37, 38] implies the following properties[35] that are appropriate for decision making. First, when fused with each new opinion, the uncertainty mass successively decreases, indicating an increasing confidence in decision making, ensuring that the procedure automatically stops analyzing new DPs when enough evidence is accumulated. Second, when two uncertainties are high, the fused uncertainty remain high. For this reason, only the opinions with high confidence (low uncertainty) play a significant role in the decision-making process. Third, the accuracy can be improved during opinion fusion; this is true when the belief mass of the correct class in the new opinion is higher than the maximum belief mass in the previous opinion. Although the accuracy can decrease when fused with a wrong prediction, the amount of degradation is limited when the uncertainty associated with the wrong prediction is large.

2.2 Classification with uncertainty quantification

DPs with different zone axes were simulated for crystal structures from the Materials Project database[39]. Since it is impossible to train the neural network using data from all possible zone axes, a number of diverse and representative zone axes were sampled and tested in 3D space (**Fig. 2**e). For training and testing, one DP from a random orientation is used as input. The classification results, shown in a confusion matrix, have true labels as well as predicted labels (**Fig. 3**a). The summation of each row is the total amount of data per class (i.e., 329,000). The diagonal numbers show correct classifications, and the off-diagonal numbers show misclassifications. For predictions made using single input DPs (one zone axis), the testing accuracy was 0.55. We hypothesized that the accuracy was low for two key reasons. First, any single input DP is only a 2D projection of a 3D crystal structure, and it may not contain sufficient information to uniquely classify the crystal system. Second, the CNN was trained using the data from a fixed and finite number of zone axes, so the zone axes of most of the testing data were different from those of the training data. To explore the former hypothesis, we trained and tested the CNN using two or three DPs (taken along different zone axes) as simultaneous inputs (*vide infra*, SI Fig. S3), and the accuracy of the prediction made with multiple DP inputs is improved

(0.94), which is expected as this provides the CNN with information in 3D. However, since the DPs will have arbitrary orientations for actual samples, this solution cannot be applied to improve accuracy in real high-throughput experiments. So, moving forward, we only input one DP at a time. In particular, we also tested the CNN using test cases comprised of a single DP input at the same orientations as the training data, and a higher accuracy of 0.78 (versus the 0.55 stated above) was attained (SI Fig. S4). This implies that the CNN is more accurate in classifying a DP if the orientation is similar to one that it has already seen. This bodes well because we note that our CNN architecture is not designed to achieve the highest accuracy for a one-shot prediction. Instead, it is designed to achieve a high accuracy as data accumulates, so that it is meaningful in the context of high-throughput experimental data acquisition. We present classification results obtained after fusing information from multiple zone axes in the next section.

The uncertainty of the CNN model was analyzed across the entire testing data set using two metrics: uncertainty mass and information entropy. The uncertainty mass (u in Eq. (1)) describes the overall uncertainty of the classification according to the evidence gathered from data[35]. The information entropy (H) is the average amount of information (or "surprise") of a random variable defined as

$$H = \sum_{k=1}^{K} p_k \log_2 p_k, \tag{2}$$

where p_k is the probability that the CNN model assigns to the k-th class. A higher u or H implies that the model is more uncertain about its prediction. For each test case, u and H are calculated and the distributions of all test cases are presented (**Fig. 3**). Both the mean uncertainty mass and the mean entropy are higher for misclassifications than for correct classifications (**Fig. 3**b,c). This result is reasonable because the model is more uncertain on the samples it predicts wrongly. The testing data is further categorized according to seen and unseen zone axes. The model is more confident when the zone axes were already seen by the model (i.e., they were used for training) (**Fig. 3**d,e). Again, the data show that the CNN trained with evidential deep learning is aware of the uncertainty in its predictions, which are leveraged to guide the decision-making process while the DPs are accumulated.

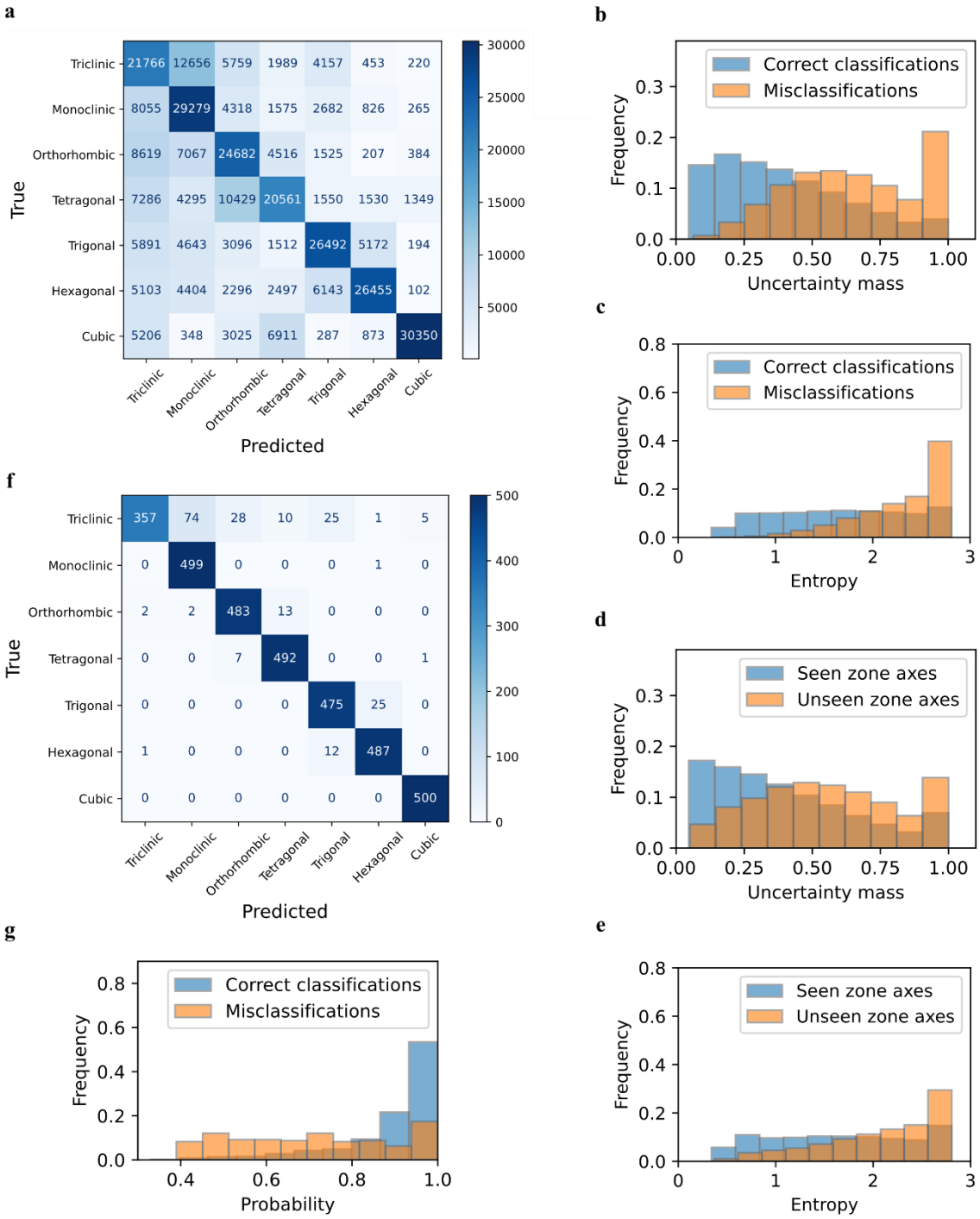


Fig. 3 Results of classfication and uncertainty quantification before (a – e) and after (f and g) multiview opinion fusion. a, The confusion matrix shows the true labels versus predicted labels. The accuracy is 0.55 before the fusion of information from multiple zone axes. **b** – **e**, Uncertainty quantification. **b** and **c** are the distribution of uncertainty mass and information entropy categorized by correct classifications and misclassifications. The model is more uncertain on the samples it predicts wrong. **d** and **e** are the distribution of uncertainty mass and entropy categorized by seen and unseen zone axes, respectively. The model is more uncertain on the samples where the testing zone axes are different from the training zone axes. **f**, Confusion matrix after opinion fusion. The overall accuracy is 0.94. The worst case is triclinic (accuracy 0.71). For other crystal systems, the accuracy is between 0.95 and 1.00. The total amount of data for each class is different from that seen in **Fig. 3**a. This is because, in this case, we treated each crystal as one sample after opinion

fusion, while above we treated the DPs from different zone axes of a crystal as separate datapoints before fusion. **g**, The testing process for each crystal is repeated multiple times. An arbitrary series of DPs with different zone axes are used for opinion fusion at each time. The final decision is the class with the majority vote. The vote percentages are higher for correct classifications than misclassifications. Thus, a correct decision is associated with high confidence.

2.3 Classification fusion from multiple random zone axes

After training, the CNN can be applied to make predictions when given a DP with any arbitrary zone axis. The MVOF-ML framework (**Fig. 1**b) is implemented and its power is illustrated (**Fig. 4**) using a triclinic $LiCo(PO_3)_3$ bulk structure randomly selected from Materials Project database (mp-26015) as an example. First, a DP orientation is randomly selected (zone axis [2 1 4]). The classification result given by the CNN model shows that the most probable classification is trigonal. The uncertainty mass for the first DP is 0.39. The uncertainty mass is higher than the threshold of 0.1, and so the second DP is collected. This representation is analogous to human reasoning, where a decision cannot be made if one is not confident in the conclusion. The CNN model classifies the second DP with another randomly selected zone axis, [2 1 0], as triclinic with an uncertainty mass of 0.22. The fused opinion indicates that both triclinic and trigonal are possible. The fused uncertainty mass is 0.16, which is still higher than the threshold of 0.1, so another DP with a random zone axis is analyzed, and the process continues until the fused uncertainty is lower than the threshold. The final fused opinion yields a decision that the crystal symmetry is triclinic, the correct decision, after five iterations. This example highlights that, although a CNN model may misclassify a single DP, the opinion fusion framework makes the correct overall decision in the end after the uncertainty decreases to a level where the CNN is confident in the opinion.

Next, the overall predictive performance of our CNN framework was examined on the whole testing data set from the Materials Project database (see Methods). The confusion matrix is shown in **Fig. 3**f. The overall accuracy is 0.94, which is higher than other models in the current literature used for crystal system classification. All cubic crystals are correctly identified (i.e., the accuracy is 1) (**Fig. 3**f), and the accuracy was the lowest for the triclinic (accuracy 0.71). For the other crystal systems, their accuracies were between 0.95 and 1.00. If we do not leverage uncertainty decision-making in our framework, fusing the opinions for the DPs of all 94 testing zone axes for a single crystal, the accuracy is 0.54. This accuracy is similar to that seen before the opinion fusion (single DP input, see Section 2.2). These results highlight that the uncertainty quantification is just as important as the number of informational inputs in the ultimate decision-making process. Our CNN operates analogous to a human assigning different degrees of "trust" to an opinion depending on the degree of uncertainty associated with it.

The results in **Fig. 3**f were obtained by majority vote, where the testing process for a crystal is repeated multiple times (100 in this example). Each time a series of DPs with different randomly chosen zone axes were used, the final decision is the classification with the majority vote. The higher the vote percentage, the higher the algorithm's confidence is in making a final decision. The vote percentages for correct classifications and misclassifications (from **Fig. 3**f) are shown in **Fig. 3**g. The median and mean vote percentages for correct classifications versus misclassifications are 0.99 vs. 0.67, and 0.90 vs. 0.69, respectively. Thus, in these examples, when a correct decision is made, it is associated with a high level of confidence.

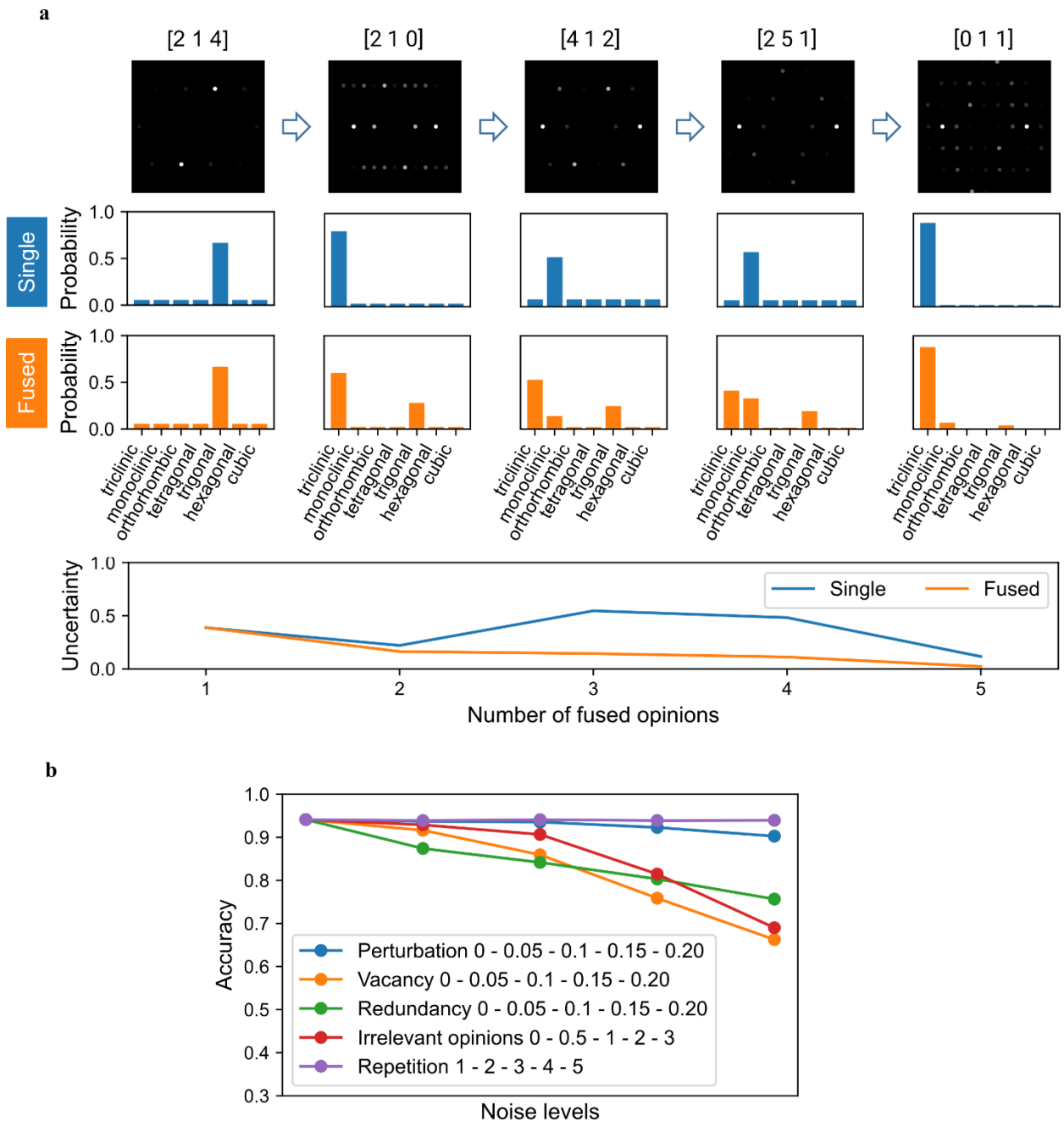


Fig. 4 An illustration for decision-making and tolerance to noise. a, The triclinic mp-26015 from the Materials Project database is tested. Row 1 shows the DPs with corresponding randomly selected zone axes collected sequentially. Probability distributions (opinions) of seven crystal systems predicted from the CNN model are presented in row 2. Row 3 shows the probability distribution formed by fusing the current opinion with the previous fused opinion. During the fusion of opinions, the uncertainty mass decreases (row 4). "Single" and "Fused" correspond to a single or fused opinion or uncertainty, respectively. It should be noted that the zone axes shown next to the DPs are only for visualization purposes and are not used for classification. **b**, Tolerance to noise. Perturbation, vacancy, and redundancy are added as sources of noise to the simulated DPs to alleviate Assertion 1. The five numbers refer to a value of σ for perturbation and proportion of the total amount of Bragg disks for vacancy and redundancy, and they represent five different levels of noise. Irrelevant opinions are DPs from other crystals to investigate the performance of the MVOF-ML framework when Assertion 2 is

relaxed. Repetition refers to fusing opinions from the same DPs by duplicating the candidate pool, corresponding to a relaxation of Assertion 3. The numbers in the legend are the noise levels.

2.4 Tolerance to noise

The results presented thus far are based on the following idealized, but still reasonably representative, conditions on the DPs:

Condition 1. *The simulated DPs are idealized with no noise*,

Condition 2. *The DPs used to make a decision are from the same crystal*, and

Condition 3. *The DPs used to make a decision are from varied zone axes*.

Now, these conditions are relaxed, and the performance of the MVOF-ML framework is investigated in more realistic noisy scenarios. The same trained model using the simulation data is tested. To alleviate Condition 1, we tested the performance of the framework on more realistic noisy DPs. Three types of noise - perturbation, vacancy, and redundancy - were introduced to the ideal simulated DPs. For perturbation, the location and intensity of the Bragg disks were altered. The perturbed Bragg disks were obtained by multiplying the location and intensity values in the vector map with random numbers drawn from an $\mathcal{N}(1,\sigma)$ Gaussian distribution, where σ is the standard deviation. As σ increases, the DPs contain Bragg disks at more strongly perturbed locations and/or intensities than their true values. For vacancy and redundancy, we randomly remove or add Bragg disks, analogous to "missing" points or detecting "false positives" in a DP that are not part of the structure. Any added Bragg disks are sampled from the range of the original vector maps. The data show that the accuracy decreases gradually with increasing noise, but there is no sudden or significant drop in accuracy when noise is introduced (**Fig. 4**b). When the noise level is low (noise level between 0.05 and 0.1 for all the three cases), the accuracies remain higher than 0.8. Thus, the MVOF-ML framework is relatively robust to those three types of noise.

In the above opinion fusion process, we assume that all of the fused opinions are multiple views of the same crystal (Condition 2). This condition is true only when a single crystalline phase is present in the material. However, when a specimen contains multiple crystal structures, the 4D-STEM dataset will include DPs corresponding to different crystals, and as a result, the opinions for different crystals may be fused. When incorporating an opinion irrelevant to the crystal being classified, it will negatively influence the decision process or lead to the wrong decision. In the original candidate pool (all possible orientations that can be collected for testing), there are 94 DPs for each crystal corresponding to 94 testing zone axes. To test the tolerance of the method to irrelevant DPs, when classifying test cases, we introduced the DPs of other crystal structures to the candidate pool, randomize the pool, and then conduct the classification. The number of irrelevant DPs varies from 0.5 to 3 times of the original pool size (i.e., 94). As expected, the accuracy decreases as the number of irrelevant DPs increases (**Fig. 4**b "Irrelevant opinions 0 - 0.5 - 1 - 2 - 3"); however, it remains higher than 0.9 even when the number of DPs corresponding to the true crystal is only half of the size of the candidate pool.

In the above analyses, we assume that every time an opinion for an additional DP with an additional zone axis is fused with another, and the additional zone axis is one that has not been used before (Condition 3), which will not always be true in real experiments. To test the robustness of the method to repeatedly fusing opinions for the same zone axis, we duplicated the original candidate pool. Thus, when multiple DPs are randomly sampled from the pool for opinion fusion, it is possible that

the same DPs are repeatedly fused. The number of duplicates varies from 2 to 5. The data show that the accuracy does not decrease if the candidate pool is duplicated ≤ 5 times **Fig. 4**b "Repetition 0 - 2 - 3 - 4 - 5").

2.5 Demonstration on experimental data

To test the performance of the MVOF-ML framework under realistic experimental scenarios, experimental data obtained from AuCu nanoparticle libraries are used as inputs (SI Fig. S5). Data from different regions of a nanoparticle were treated as DPs from different orientations of the same crystal. In total, 17 nanoparticle datasets (SI Fig. S6) with cubic crystal systems are tested. The experimental data processing for an example nanoparticle (**Fig. 5**a) is depicted (SI Fig. S7). The DPs distributed at different scan positions (white squares in **Fig. 5**a) are shown (**Fig. 5**b) as well as the DPs of all the nanoparticles (SI Figs. S8-S24). At each scan position, the locations and intensities of the Bragg disks are extracted and are used as inputs in the MVOF-ML. The CNN model was retrained on the simulated data (the real experimental data are the test data) with disk intensity higher than 0.2. This intensity thresholding removes all low-intensity disks that would not be observed experimentally to provide a more realistic training dataset. The accuracy of the MVOF-ML framework on the simulated testing data was 0.80. The results for experimental data are as follows.

Out of 17 nanoparticles, 16 were correctly classified (**Fig. 5**c). Nanoparticles 6, 9, and 14 are the same as nanoparticles 5, 8, and 13 at different magnifications, respectively, and they maintain the expected classification trends. The one wrong prediction (trigonal as opposed to cubic, particle 17 in Fig. 5c and SI Fig. 24) is based on the only single crystalline nanoparticle, for which all DPs are from the same zone axis, which makes a unique classification impossible. In this particular case, the particle's orientation is near the cubic <101> zone axes, which exhibit the same symmetry as the trigonal <101>. Accordingly, the correct opinion - "cubic" - is classified as the second probable class for that nanoparticle. Cases like these are inevitable when analyzing individual nanoparticles and highlight the power of our MVOF-ML framework to accumulate data. In a high-throughput experimental data acquisition process, the framework would be allowed to accumulate data across particles to arrive at a "global opinion" for a particle array, eliminating the possibility of incorrect classification due to an insufficient number of crystal orientations being acquired. In addition, cases where there are multiple predictions with comparable probabilities could not only be resolved in a similar fashion, but the number of iterations it takes for the framework to arrive at its final opinion could serve as a metric of difficulty for classification, pinpointing structures of particular interest for manual examination. These results show that the high accuracy observed on simulated data is maintained on noisy experimental data, which underscores the importance of designing robust systems for use in an experimental context and highlights the effectiveness of this MVOF-ML framework.

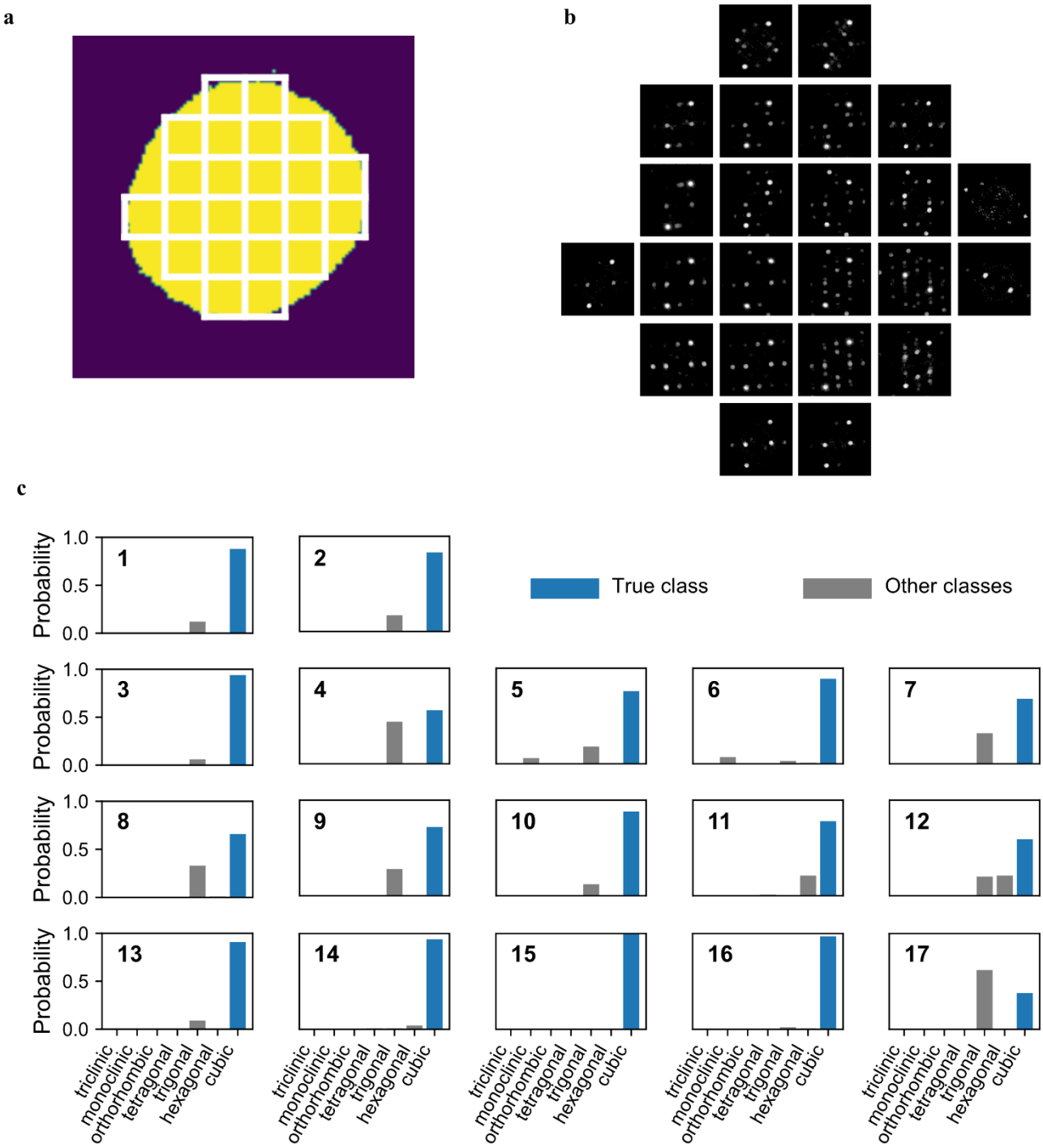


Fig. 5 Experiment data and results. **a**, An example of a nanoparticle and its scan positions in 4D-STEM. **b**, The DPs corresponding to different scan positions. **c**, Classification results for the 17 nanoparticles. Sixteen out of 17 nanoparticles were correctly classified as cubic crystal system.

3. Discussion

In this work, DPs are described as vector maps containing position and intensity information of Bragg disks to encapsulate the parts of the data in a condensed format that contains physical information. We compared the predictive performance of the CNN model using vector maps versus using images as inputs. Due to the difference in types of inputs (i.e., images and tabular data), we used

different CNN architectures for each input type. For grayscale image inputs, we adopted the typical CNN with several combinations of a convolutional and max pooling layer followed by the fully connected layer and the output layer. For vector map inputs, we used a modified PointNet CNN architecture, as described earlier. We tested the predictive performance given 1, 2, and 3 images or vector maps as inputs (corresponding to 1, 2, or 3 multiple DPs for each case), and we found that the accuracy is always higher when vector maps are used as DP descriptors (SI Fig. S25). In addition, we again observed that the accuracy improves significantly when two or three DPs are input to the CNN compared to one DP (*via supra*). The results shown in this section are representative of the situation where testing data are obtained from the same zone axes as the training data, a common practice in the literature when ML is used for crystal structure classification from DPs. However, in real experiments, the zone axes of the collected crystal diffraction data may be unknown and random, especially in high-throughput workflows, including those involving nanoparticle megalibraries. Our MVOF-ML framework (Section 2.1) can avoid that problem, significantly increasing its practical application in such workflows.

In summary, we have developed a framework for automated crystal system identification from DPs with arbitrary zone axes for high-throughput material discovery. Inspired by the human reasoning process from multiple perspectives (i.e., DPs of a crystal taken from different zone axes), the problem is modeled as a sequential decision-making process by multiview opinion fusion. The location and intensity information of the Bragg disks in the DP images are extracted and used as inputs for ML. We designed a custom CNN to capture the physical parameters embedded with the DPs by working off the adjacency relations between Bragg disks. The CNN was trained in the context of evidential deep learning, which provides opinions (classification probabilities) and quantifies uncertainties. The decision is made through the fusing of the opinions from multiple zone axes and is automated through the guidance of the quantified uncertainty. Our framework achieves high testing accuracy and is shown to be robust under noisy scenarios involving Bragg disk perturbation, vacancy and redundancy. Finally, we have demonstrated the framework on real experimental data with high accuracy (0.94). To enhance the applicability of the MVOF-ML method to classify deeper levels of crystal structures (e.g., space groups), future work will implement retrieval of space groups and lattice parameters from the DPs to more completely describe the structures of the materials. In the context of experimental materials discovery, where data can now be generated faster than it can be analyzed, this work sets the stage for integrating ML frameworks into high-throughput experimental workflows, not only delivering the much-needed processing power, but also providing a basis for autonomous decision-making in the materials discovery pipeline.

4. Methods

4.1 Simulation data preparation

The data were generated using an automated workflow modified from Reference[40]. We retrieved the crystallographic information framework (CIF) files from the Materials Project database[39]. In these files, it was necessary to regenerate the symmetry information (i.e., space groups and symmetry operations). Using the FINDSYM software[41], we obtained the crystal system for each structure. We

developed an electron diffraction simulator using the open-source Pymatgen package[42]. Given an incident beam direction, the simulator generates the positions and intensities of Bragg disks, from which we can construct vector maps.

We adopted the following principle for zone axis sampling. The sampling covers zone axes [*u v w*] in the 1/8 sphere (due to symmetry) (**Fig. 2**e), with integer *u*, *v* and *w* values in the range from 0 to 5 (except all zeros). For two zone axes that are close (i.e., the angle between them is below a threshold), only the lower order ones will be sampled. The threshold for training and testing data are 10° and 5°, respectively. Thus, the testing data contains DPs from zone axes unseen in the training data, which matches the scenario of applying the trained model to real experimental data. In total, there are 34 zone axes for training and 94 for testing. Training and testing sets each include 500 randomly sampled crystal structures per crystal system. For each crystal, DPs are simulated with the above zone axes. The resulting training data size is 500 × 7 × 34 = 119,000 and testing data size is 500 × 7 × 94 = 329,000.

We transform the location of a Bragg disk from Cartesian coordinates (the horizontal location *x* and the vertical location *y*) to polar coordinates (the distance to the center Bragg disk *r* and the angle from the horizontal θ). Thus, the vector map is presented by a matrix whose three columns are *r*, θ and intensity *I*. Each row represents one Bragg disk in the DP.

4.2 Experimental data preparation

Chloroauric acid hydrate ($HAuCl_4$ xH_2O, ≥ 99.9% trace metals basis), copper(II) nitrate hydrate ($Cu(NO_3)_2$ xH_2O, 99.999%), nitric acid (HNO_3, 70%, redistilled, ≥ 99.999% trace metals basis), hexamethyldisilazane (HMDS, 99.9%), and hexane (anhydrous, 95%) were purchased from Sigma-Aldrich. Poly(ethylene oxide)-block-poly(2-vinylpyridine) (PEO-b-P2VP, Mn = 2.8-b-1.5 kg/mol, polydispersity index = 1.11) was purchased from Polymer Source, Inc. Dip-Pen Nanolithography (DPN) 1D pen arrays (type M, without a gold coating) were purchased from Advanced Creative Solutions Technology, LLC. Silicon nitride membranes for TEM (amorphous, thickness = 15 nm) were purchased from Ted Pella, Inc.

The synthesis of nanoparticle arrays has been described previously[43]. In brief, silicon nitride membranes were functionalized with HMDS and precursor inks were prepared by adding PEO-b-P2VP, metal precursors, and HNO_3, at desired ratios in deionized water. Plasma-treated DPN pen arrays were mounted onto an AFM (Park Systems XE-150) at a relative humidity of 80% and coated with the precursor ink. Nanoreactors were deposited onto the functionalized silicon nitride membranes by repeatedly bringing the cantilever tips in contact with the surface. Nanoparticles were formed from the nanoreactors through thermal annealing in a H_2 atmosphere.

All 4D-STEM datasets (128×128×128×128 pixels) were acquired on a JEOL JEM-ARM300F (S)TEM equipped with a cold field emission gun operated at 80 kV on a Gatan STELA hybrid-pixelated camera operated at 2000fps.

The identification plots of the 17 nanoparticles are presented in SI Fig. S6. Each nanoparticle varies in shape and composition of grains. The locations and intensities of the Bragg disks are extracted from the DPs. The locations are represented as the centers of the Bragg disks. Since the true Bragg disk intensities are unknown for the experimental data, we approximate intensities using the cross-correlation intensities that measure the extent that a point is a true Bragg disk. The process to extract the data for a typical nanoparticle contains local averaging, symmetrizing, background subtraction,

Bragg disk identification, calibration and recentering (SI Fig. S7). Since the raw DPs of each scan position from the 4D-STEM datasets are noisy, we averaged DPs of each scan position for analysis to reduce the noise. In the end, the coordinates are centered and transformed from pixel distances to inverse nanometers.

4.3 Framework implementation

Algorithm 1: Algorithm for automated crystal system identification

```
/*Training*/
Train a CNN using simulated diffraction data with multiple zone axes

/*Testing*/
Repeat n times:
    Collect a DP
    Prediction with uncertainty
    While uncertainty > threshold do
        Collect next DP
        Prediction with uncertainty
        Fuse predictions and fuse uncertainties
    End
End
```

Inspired by the human reasoning process, the framework for automated crystal system identification for 4D-STEM is developed and presented (**Fig. 1**b), and the algorithm is shown in Algorithm 1. First, a machine learning model (a modified version of PointNet CNN in this work) is trained to make predictions in place of a human expert. Next, a DP simulated from an arbitrary orientation is used as input in the CNN. The CNN predicts the probability that a crystal belongs to each crystal system, regarded as one opinion. In addition to an opinion about what probability it assigns to each class, we also determine how confident it is in that opinion (i.e., the uncertainty quantification)[30, 31, 32]. Next, another DP is collected from a different arbitrary orientation, and the same process is conducted to form the second opinion as well as its uncertainty. Following that, the first and second opinions are fused, and their uncertainties are also fused. If the fused uncertainty is still high, another new opinion will be formed from an additional DP. The previously fused opinion and uncertainty will be fused with the new opinion and uncertainty. Fused Opinion II is more confident (less uncertain) about what structure the crystal belongs to as information is accumulated (**Fig. 1**b). That process is repeated until the fused uncertainty is below a sufficiently small, specified threshold. In other words, when enough DPs are analyzed, decisions regarding the crystal system can be finally made with a high confidence. The decision is the final fused opinion. The uncertainty threshold in Algorithm 1 is a user-defined value. A lower threshold corresponds to a more conservative decision.

4.4 CNN architecture

The details of the CNN architecture (**Fig. 2**b) are shown in Table 1. The three PointNet braches have the same architecture. The inputs are matrixes. Each row represents a Bragg disk. Three columns represet the distance to the center Bragg disk r, the angle from the horizontal line θ, and intensity I, respectively. In the convolutional layers, the kernel moves in a 1D direction from top to bottom of the

input matrixes since each row is treated as a unit in PointNet. The filters of the convolutional layers extract specific features. A significant difference between PointNet and regular CNN is that there are no pooling layers right after the convolutional layers. Instead, there is a global pooling layer after stacking all convolutional layers to condense a filter as one value. In the original PointNet, the kernel is 1D with size of 1 for the invariance of point sequence in 3D space, which results in the ignorance of information among adjacent points. In this work, the kernel size is set to be greater than 1 to extract and learn information among adjacent Bragg disks, which increases the accuracy. The condensed features from the three PointNets are concateated before being processed by two fully connected layers. At the end, there is an output layer with 7 neurons representing 7 crytal systems. The Rectified Linear Unit (ReLU) activation function is used for all convolutional layers, the fully connected layer, and the output layer. The optimizer is Adam with learning rate of 0.001. The CNN was trained with a batch size of 32 and 10 epochs under the framwork of Tensorflow.

Table 1 CNN layers and specifications.

Layers		**Specifications**
PointNet 1	Input Layer	Shape: (800, 3)
	Convolutional layer	Kernel size: 5 × 3, Filter size: 32, Activation: ReLU
	Convolutional layer	Kernel size: 3 × 1, Filter size: 32, Activation: ReLU
	Convolutional layer	Kernel size: 3 × 1, Filter size: 32, Activation: ReLU
	Convolutional layer	Kernel size: 3 × 1, Filter size: 64, Activation: ReLU
	Convolutional layer	Kernel size: 3 × 1, Filter size: 256, Activation: ReLU
	Pooling layer	Global max pooling
PointNet 2		Same architecture as PointNet 1
PointNet 3		Same architecture as PointNet 1
Concatenate layer		Concatenate outputs from PointNets 1, 2, and 3
Fully connected layer 1		Size: 256, Activation: ReLU
Fully connected layer 2		Size: 128, Activation: ReLU
Output layer		Size: 7, Activation: ReLU

4.5 Deep learning for classification with uncertainty quantification

Classification with uncertainty quantification is achieved through evidential deep learning[33]. The main idea is that the uncertainty is explicitly modeled, together with class probabilities, through Eq. (1), where the belief mass and uncertainty mass are obtained as $b_k = e_k/S$, $k = 1, 2, \dots, K$, and $u = K/S$, respectively, where e_k is the evidence for the k-th class, and $S = \sum_{k=1}^{K}(e_k + 1)$. The CNN outputs are e_k. The activation function at the last layer is the Rectified Linear Unit (ReLU) activation to ensure the non-negativity of e_k. The expected probability of the k-th class is $\hat{p}_k = \alpha_k/S$, where $\alpha_k = e_k + 1$.

The neural network is trained by minimizing the loss function expressed using the Dirichlet distribution. The Dirichlet distribution is a probability density function that models the probability distribution of K class probabilities $\mathbf{p} = [p_1, \dots, p_K]$. It is expressed as

$$D(\mathbf{p}|\boldsymbol{\alpha}) = \begin{cases} \dfrac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} p_k^{\alpha_k - 1} & \text{for } \mathbf{p} \in \mathcal{S}_K \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

where $\mathcal{S}_K$ is a set of K probabilities

$$S_K = \{\mathbf{p} | \sum_{k=1}^{K} p_k = 1 \text{ and } 0 \le p_k \le 1\}$$

and $B(\boldsymbol{\alpha})$ is the multinomial beta function. $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]$ is a set of parameters characterizing the Dirichlet distribution function. The K parameters of Dirichlet distribution represent the evidence for K classes.

Let Θ be the weights and biases of a neural network and $\mathbf{y} = [y_1, \ldots, y_K]$ be a one-hot encoded vector representing the true label. The following loss function is adopted

$$L(\Theta) = \int \|\mathbf{y} - \mathbf{p}\|_2^2 \, D(\mathbf{p}|\boldsymbol{\alpha}) d\mathbf{p} + \lambda_t KL[D(\mathbf{p}|\tilde{\boldsymbol{\alpha}}) \| D(\mathbf{p}|\mathbf{1})], \tag{4}$$

where $\|\mathbf{y} - \mathbf{p}\|_2^2$ is the sum of squares loss.

The first term in the loss function (4) can be rewritten as

$$\int \|\mathbf{y} - \mathbf{p}\|_2^2 \, D(\mathbf{p}|\boldsymbol{\alpha}) d\mathbf{p} = \sum_{k=1}^{K} \left(y_k - \frac{\alpha_k}{S}\right)^2 + \frac{\alpha_k(S-\alpha_k)}{S^2(S+1)}, \tag{5}$$

where the first and second terms represent prediction error and variance, respectively.

The second term in the loss function (4) is a Kullback-Leibler (KL) divergence term, which serves as regularization, where $\mathbf{1}$ is a vector with all the K elements being ones, and $\tilde{\boldsymbol{\alpha}} = \mathbf{y} + (\mathbf{1} - \mathbf{y}) \cdot \boldsymbol{\alpha}$ is a vector after removing the evidence of the true class. $\lambda_t = \min(1, t/50) \in [0,1]$ gradually increases with the increase of epoch t. The KL divergence can be expressed as

$$KL[D(\mathbf{p}|\tilde{\boldsymbol{\alpha}}) \| D(\mathbf{p}|\mathbf{1})] = \log\left(\frac{\Gamma(\sum_{k=1}^{K} \tilde{\alpha}_k)}{\Gamma(K) \sum_{k=1}^{K} \Gamma(\tilde{\alpha}_k)}\right) + \sum_{k=1}^{K} (\tilde{\alpha}_k - 1) \left[\psi(\tilde{\alpha}_k) - \psi\left(\sum_{k=1}^{K} \tilde{\alpha}_k\right)\right], \tag{6}$$

where $\Gamma(\cdot)$ and $\psi(\cdot)$ are the gamma function, and the digamma function, respectively. This term is to ensure that the evidence for a sample approaches zero if the model cannot classify it correctly.

4.6 Multiview classification fusion

Dempster's combination rule is used to combine two opinions under uncertainty[35]. Consider that the joint masses $\{\{b_k^1\}_{k=1}^K, u^1\}$ and $\{\{b_k^2\}_{k=1}^K, u^2\}$ are obtained from two different views. The fused belief mass b_k and the uncertainty mass u can be expressed as

$$b_k = \frac{1}{1 - C}(b_k^1 b_k^2 + b_k^1 u^2 + b_k^2 u^1) \tag{7}$$

and

$$u = \frac{1}{1-C} u^1 u^2, \tag{8}$$

respectively, where $C = \sum_{i \neq j} b_i^1 b_j^2$ measures the quantity of conflicts between two different views.

Acknowledgements

Funding. The project described was supported by the Sherman Fairchild Foundation, Inc, the Air Force Office of Scientific Research award FA9550-22-1-0300, and Center for Hierarchical Materials Design (Chimad) under National Institute of Standards and Technology (NIST) 70NANB19H005. This work made use of the EPIC facility of Northwestern University's NUANCE Center, which has received support from the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-1542205); the MRSEC program (NSF DMR-1720139) at the Materials Research Center; the International Institute for Nanotechnology (IIN); the Keck Foundation; and the State of Illinois,

through the IIN. The authors thank Dr. Jenny K. Hedlund Orbeck and Dr. Sarah H. Petrosko for helpful discussions and manuscript preparation.

Author contributions. **Jie Chen**: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Hengrui Zhang**: Conceptualization, Software, Writing - review & editing. **Carolin B. Wahl**: Conceptualization, Methodology, Software, all materials synthesis, characterization, and experimental data collection, Validation, Writing - review & editing. **Wei Liu**: Conceptualization, Software, Validation, Writing - review & editing. **Chad A. Mirkin**: Conceptualization, Formal Analysis, Writing - review & editing, Supervision. **Vinayak P. Dravid**: Conceptualization, Writing - review & editing, Supervision. **Daniel W. Apley**: Conceptualization, Writing - review & editing, Supervision. **Wei Chen**: Conceptualization, Writing - review & editing, Supervision.

Competing interest. C.A.M and C.B.W have financial interests in Mattiq Inc. which could potentially benefit from the outcomes of this research. All other authors declare no competing financial interest.

Data and code availability. The crystallographic information framework (CIF) files, and the codes for simulation of DPs with arbitrary zone axes, CNN model training, and opinion fusion for decision making are available at https://github.com/jcj7292/Multiview-Opinion-Fusion-Machine-Learning.

References

1. Pham T-H, Qiu Y, Zeng J, Xie L, Zhang P. A deep learning framework for high-throughput mechanism-driven phenotype compound screening and its application to COVID-19 drug repurposing. *Nature Machine Intelligence* **3**, 247-257 (2021).
2. Huo F, Zheng Z, Zheng G, Giam LR, Zhang H, Mirkin CA. Polymer pen lithography. *Science* **321**, 1658-1660 (2008).
3. Kluender EJ, *et al.* Catalyst discovery through megalibraries of nanomaterials. *Proceedings of the National Academy of Sciences* **116**, 40-45 (2019).
4. Chen P-C, *et al.* Polyelemental nanoparticle libraries. *Science* **352**, 1565-1569 (2016).
5. Ophus C. Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microscopy and Microanalysis* **25**, 563-582 (2019).
6. Ziatdinov M, Ghosh A, Wong CY, Kalinin SV. AtomAI framework for deep learning analysis of image and spectroscopy data in electron and scanning probe microscopy. *Nature Machine Intelligence* **4**, 1101-1112 (2022).
7. Stach E, *et al.* Autonomous experimentation systems for materials development: A community perspective. *Matter* **4**, 2702-2726 (2021).
8. Aguiar J, Gong ML, Unocic R, Tasdizen T, Miller B. Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning. *Science advances* **5**, eaaw1949 (2019).

9. Lee J-W, Park WB, Lee JH, Singh SP, Sohn K-S. A deep-learning technique for phase identification in multiphase inorganic compounds using synthetic XRD powder patterns. *Nature communications* **11**, 1-11 (2020).
10. Chen J, Liu Y. Fatigue modeling using neural networks: A comprehensive review. *Fatigue & Fracture of Engineering Materials & Structures* **45**, 945-979 (2022).
11. Tiong LCO, Kim J, Han SS, Kim D. Identification of crystal symmetry from noisy diffraction patterns by a shape analysis and deep learning. *npj Computational Materials* **6**, 196 (2020).
12. Li S, Liu Y, Chen D, Jiang Y, Nie Z, Pan F. Encoding the atomic structure for machine learning in materials science. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **12**, e1558 (2022).
13. Chen D, *et al.* Automating crystal-structure phase mapping by combining deep learning with constraint reasoning. *Nature Machine Intelligence* **3**, 812-822 (2021).
14. Kaufmann K, *et al.* Crystal symmetry determination in electron diffraction using machine learning. *Science* **367**, 564-568 (2020).
15. Aguiar JA, Gong ML, Tasdizen T. Crystallographic prediction from diffraction and chemistry data for higher throughput classification using machine learning. *Computational Materials Science* **173**, 109409 (2020).
16. Kim J, Tiong LCO, Kim D, Han SS. Deep learning-based prediction of material properties using chemical compositions and diffraction patterns as experimentally accessible inputs. *The Journal of Physical Chemistry Letters* **12**, 8376-8383 (2021).
17. Ra M, Boo Y, Jeong JM, Batts-Etseg J, Jeong J, Lee W. Classification of crystal structures using electron diffraction patterns with a deep convolutional neural network. *RSC advances* **11**, 38307-38315 (2021).
18. Ziletti A, Kumar D, Scheffler M, Ghiringhelli LM. Insightful classification of crystal structures using deep learning. *Nature communications* **9**, 1-10 (2018).
19. Li Y, Dong R, Yang W, Hu J. Composition based crystal materials symmetry prediction using machine learning with enhanced descriptors. *Computational Materials Science* **198**, 110686 (2021).
20. Vecsei PM, Choo K, Chang J, Neupert T. Neural network based classification of crystal symmetries from x-ray diffraction patterns. *Physical Review B* **99**, 245120 (2019).
21. Lolla S, Liang H, Kusne AG, Takeuchi I, Ratcliff W. A semi-supervised deep-learning approach for automatic crystal structure classification. *Journal of Applied Crystallography* **55**, (2022).
22. Oviedo F, *et al.* Fast and interpretable classification of small X-ray diffraction datasets using data augmentation and deep neural networks. *npj Computational Materials* **5**, 1-9 (2019).
23. Park WB, *et al.* Classification of crystal structure using a convolutional neural network. *IUCrJ* **4**, 486-494 (2017).
24. Lin K-L. Phase identification using series of selected area diffraction patterns and energy dispersive spectrometry within TEM. *Microscopy Research* **2**, 57 (2014).
25. Poo M-m. Towards brain-inspired artificial intelligence. *National Science Review* **5**, 785-785 (2018).
26. Ziatdinov M, *et al.* Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations. *ACS nano* **11**, 12742-12752 (2017).
27. Wu Y, *et al.* Brain-inspired global-local learning incorporated with neuromorphic computing. *Nature Communications* **13**, 65 (2022).

28. Qi CR, Su H, Mo K, Guibas LJ. Pointnet: Deep learning on point sets for 3d classification and segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*) (2017).
29. Chai J, Huo F, Zheng Z, Giam LR, Shim W, Mirkin CA. Scanning probe block copolymer lithography. *Proceedings of the National Academy of Sciences* **107**, 20202-20206 (2010).
30. Chen J, Yu Y, Liu Y. Physics-guided mixture density networks for uncertainty quantification. *Reliability Engineering & System Safety* **228**, 108823 (2022).
31. Chen J, Liu Y. Probabilistic physics-guided machine learning for fatigue data analysis. *Expert Systems with Applications* **168**, 114316 (2021).
32. Lu J, Zhan Z, Apley DW, Chen W. Uncertainty propagation of frequency response functions using a multi-output Gaussian Process model. *Computers & Structures* **217**, 1-17 (2019).
33. Sensoy M, Kaplan L, Kandemir M. Evidential deep learning to quantify classification uncertainty. *Advances in neural information processing systems* **31**, (2018).
34. Hu J, *et al.* MaterialsAtlas.org: a materials informatics web app platform for materials discovery and survey of state-of-the-art. *npj Computational Materials* **8**, 65 (2022).
35. Han Z, Zhang C, Fu H, Zhou JT. Trusted Multi-View Classification with Dynamic Evidential Fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2022).
36. Wu B, Qiu W, Huang W, Meng G, Huang J, Xu S. A multi-source information fusion approach in tunnel collapse risk analysis based on improved Dempster–Shafer evidence theory. *Scientific Reports* **12**, 3626 (2022).
37. Ha M-Q, *et al.* Evidence-based recommender system for high-entropy alloys. *Nature Computational Science* **1**, 470-478 (2021).
38. Liu S, Li X, Hu C, Yao J, Han X, Wang J. Spammer detection using multi-classifier information fusion based on evidential reasoning rule. *Scientific Reports* **12**, 12458 (2022).
39. Jain A, *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials* **1**, 011002 (2013).
40. Wang Y, *et al.* Learning the crystal structure genome for property classification. *Physical Review Research* **4**, 023029 (2022).
41. Stokes HT, Hatch DM. FINDSYM: program for identifying the space-group symmetry of a crystal. *Journal of Applied Crystallography* **38**, 237-238 (2005).
42. Ong SP, *et al.* Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science* **68**, 314-319 (2013).
43. Wahl CB, Aykol M, Swisher JH, Montoya JH, Suram SK, Mirkin CA. Machine learning–accelerated design and synthesis of polyelemental heterostructures. *Science Advances* **7**, eabj5505 (2021).

Supplementary Information for

Automated crystal system identification from diffraction patterns using Multiview Opinion Fusion Machine Learning

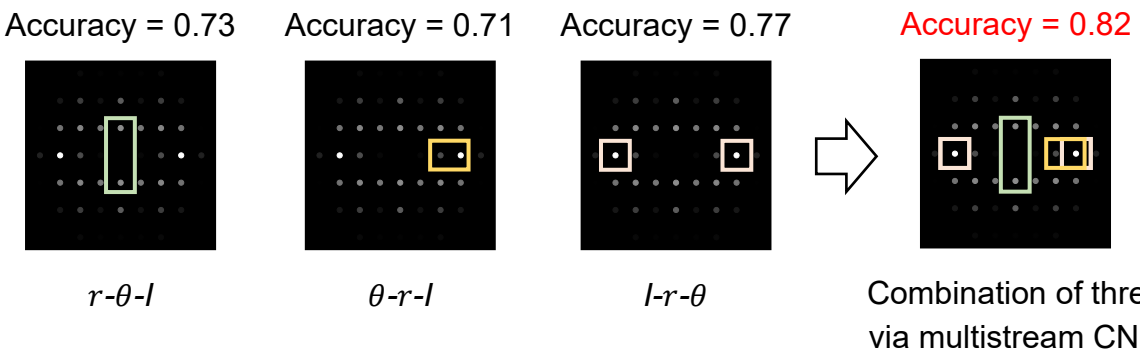


Fig. S1 Predictive accuracy comparison between single- and multistream CNN with different kernel types (discussed in Section 2.1)**.** When combing the three kernel types via multistream CNN, the accuracy is the highest.

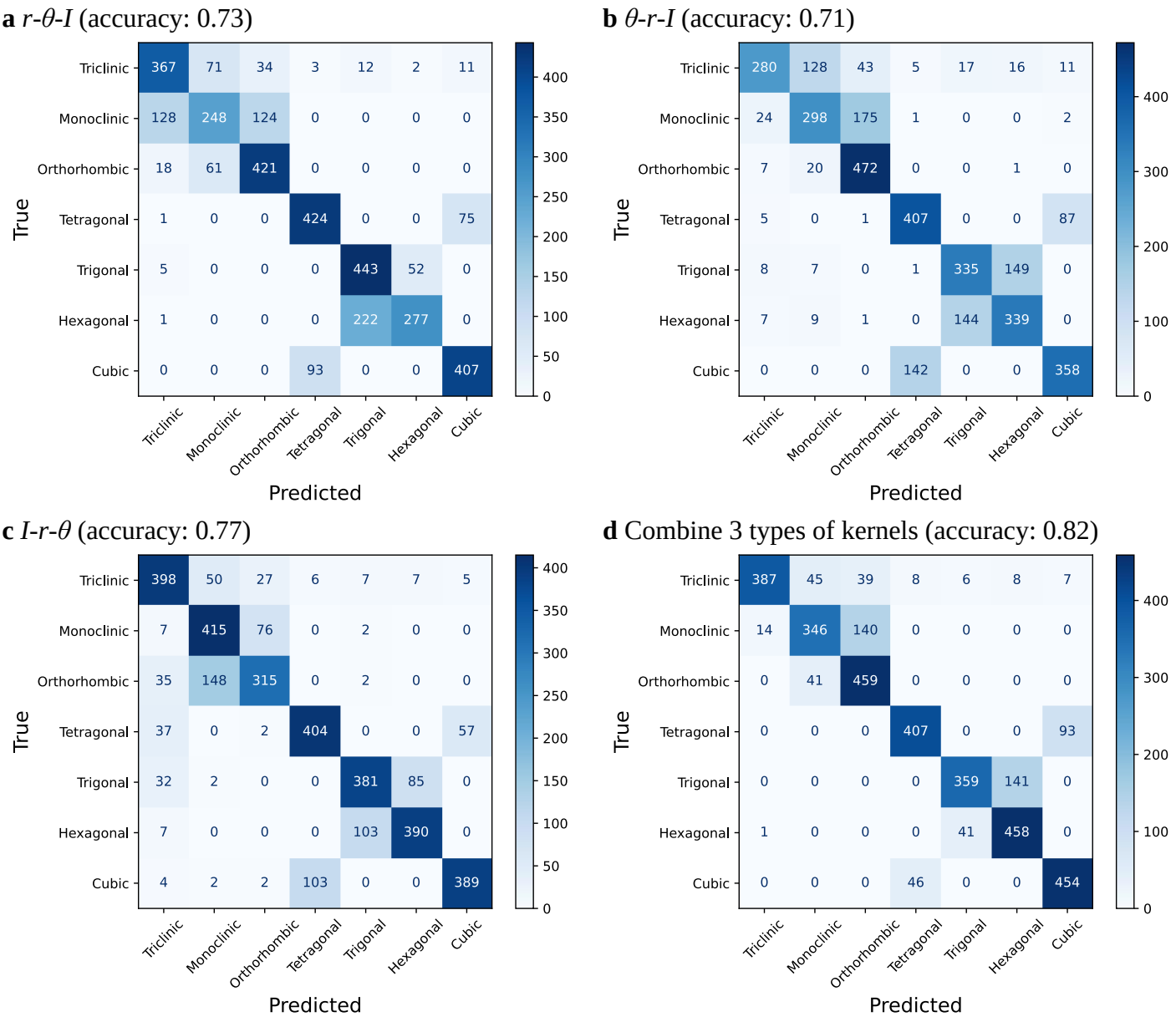
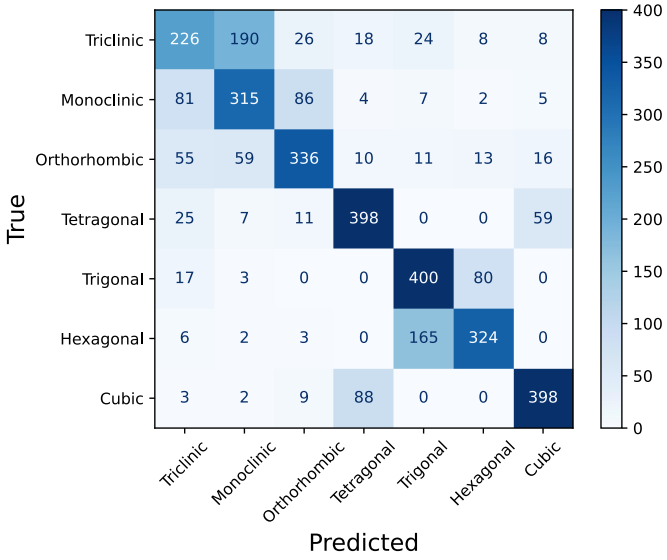
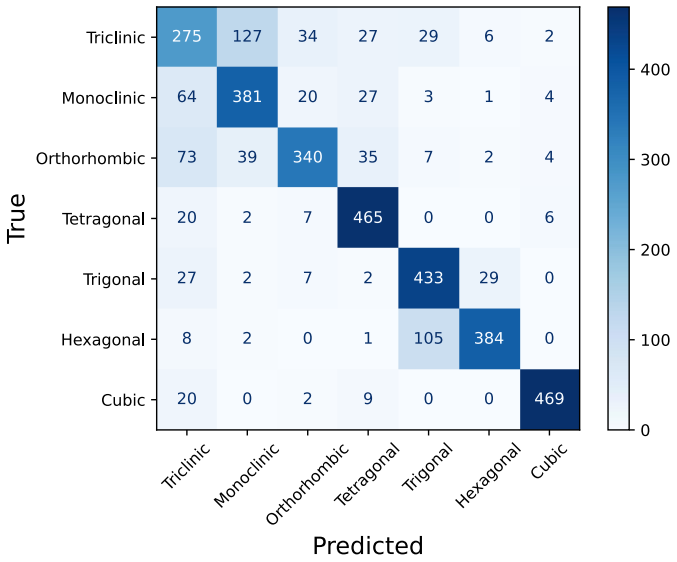


Fig. S2 Predictive accuracy comparison among three different kernel types and their combination (discussed in Section 2.1)**.** Different kernel types correspond to vector maps are sorted in three ways: *r-θ-I*, *θ-r- I*, and *I-r-θ*, respectively. The data has zone axis [001].

a 1 image (accuracy: 0.68, zone axis: [001])

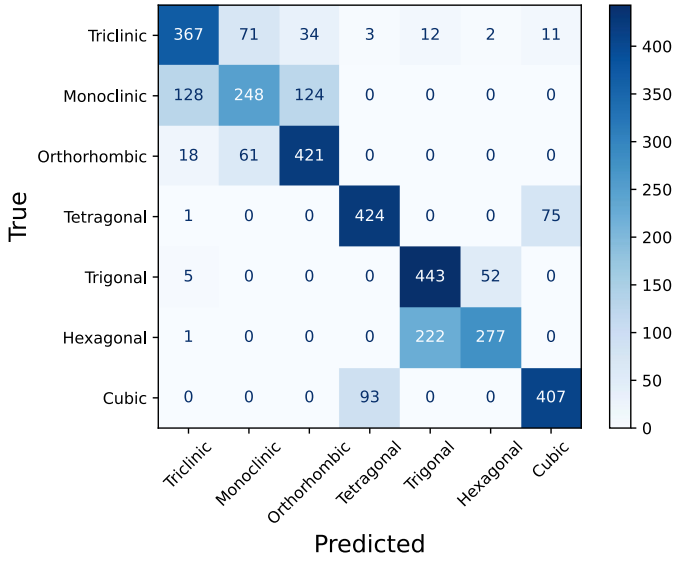


b 2 images (accuracy: 0.78, zone axes: [001], and [100])

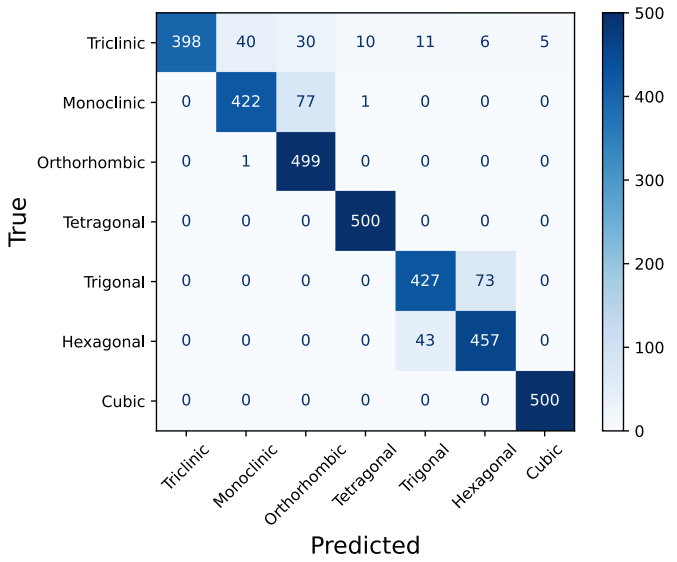


c 3 images (accuracy: 0.84, zone axes: [001], [100], and [010])

d 1 vector map (accuracy: 0.73, zone axis: [001])



e 2 vector maps (accuracy: 0.92, zone axes: [001], and [100])



f 3 vector maps (accuracy: 0.94, zone axes: [001], [100], and [010])

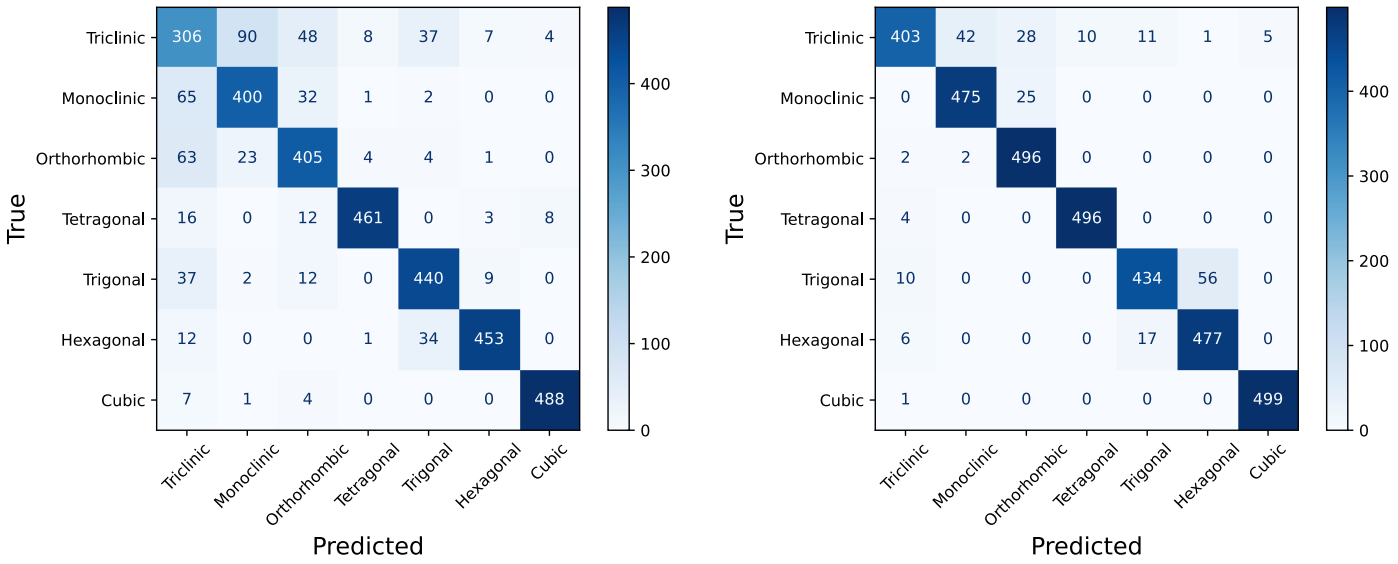


Fig. S3 Predictive results comparison between using images (a, b and c) or vector maps (d, e and f) as inputs (discussed in Section 2.2)**.** The vector maps are sorted by *r*-*θ*-*I*. The accuracy is always higher when vector maps are used as diffraction pattern descriptors. Also, the accuracy can be increased by inputting multiple diffraction patterns. The accuracy improvement is more significant from one to two diffraction patterns than from two to three diffraction patterns.

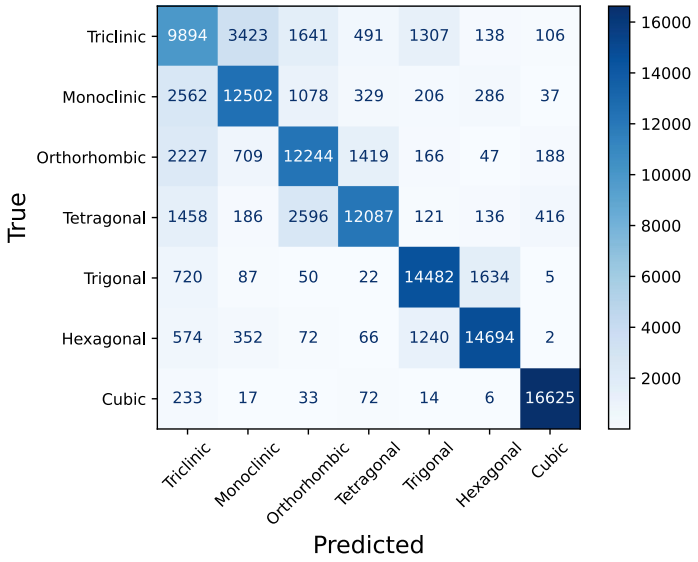


Fig. S4 Predictive results when the CNN is tested using the seen zone axes (discussed in Section 2.2)**.** The CNN is tested using the data with 34 zone axes same as the training data. Results shows an accuracy with a value of 0.78. That is higher than using testing data containing zone axes outside the range of training data.

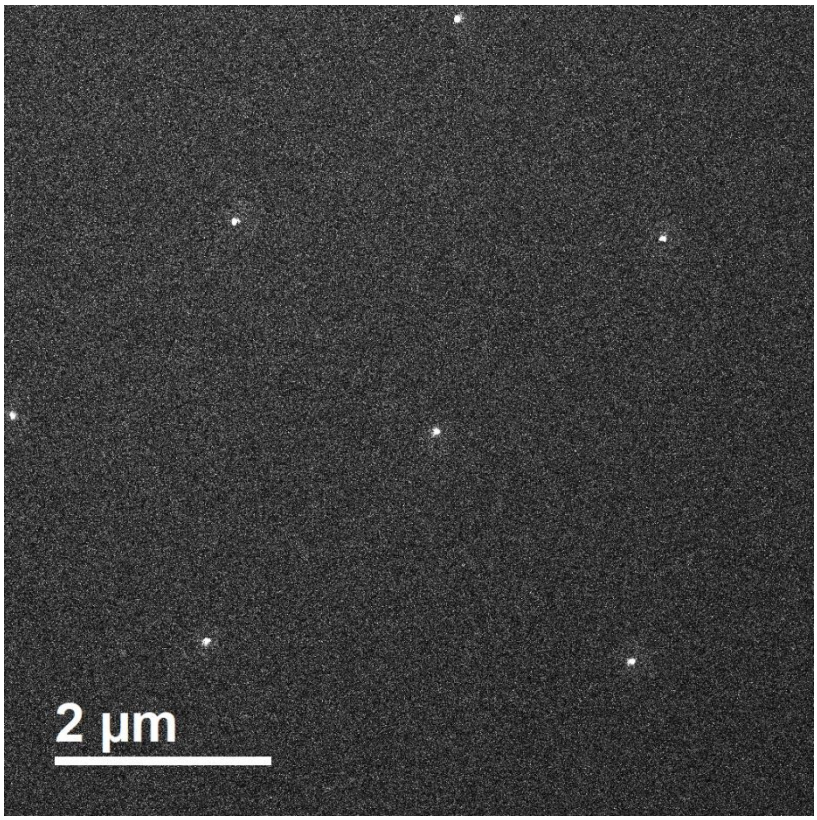


Fig. S5: Dark field STEM image of a nanoparticle array (discussed in Section 2.5)**.**

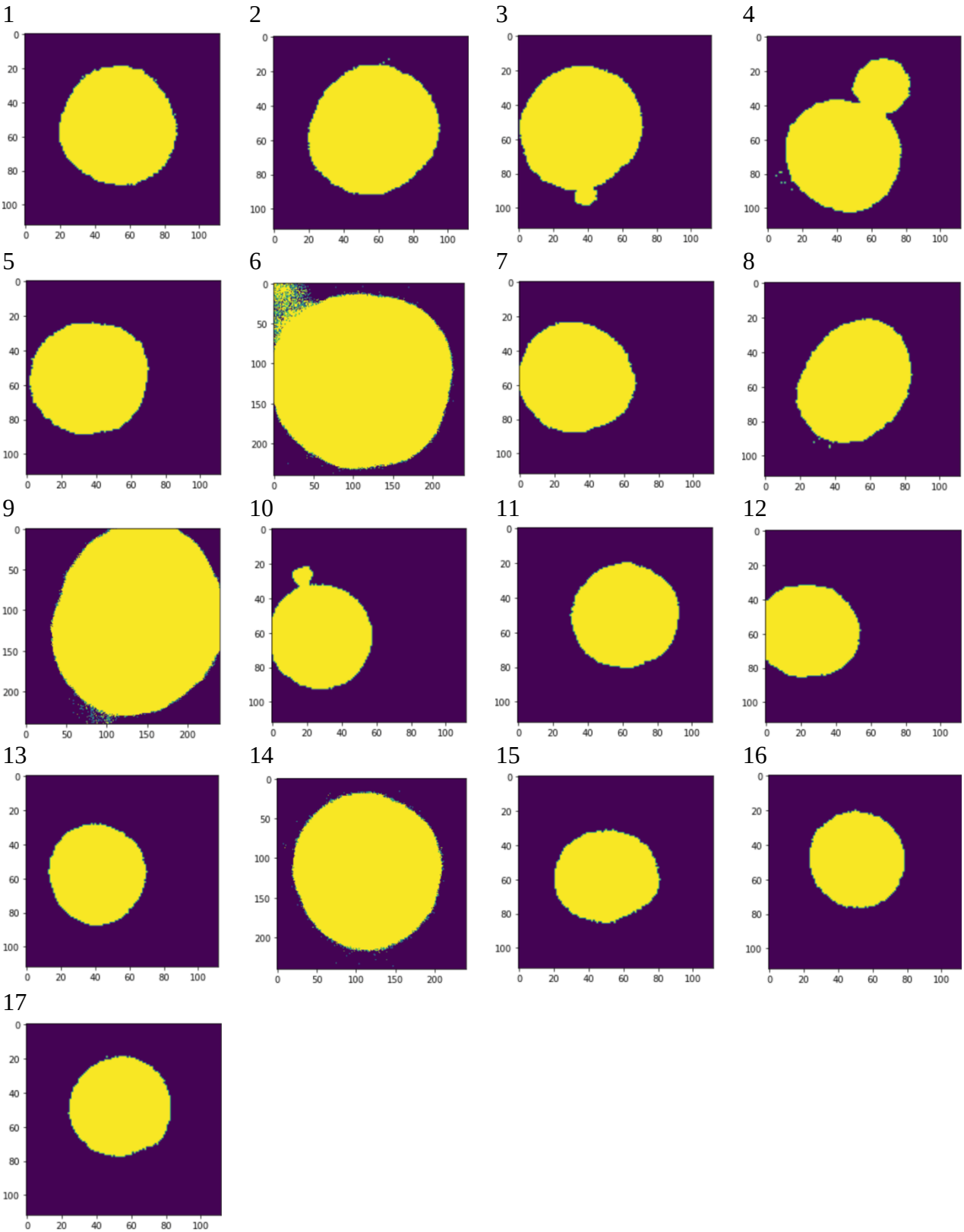


Fig. S6 The nanoparticle identification plots of the tested 17 nanoparticles (Unit: pixel) (discussed in Section 2.5 and 3)**.**

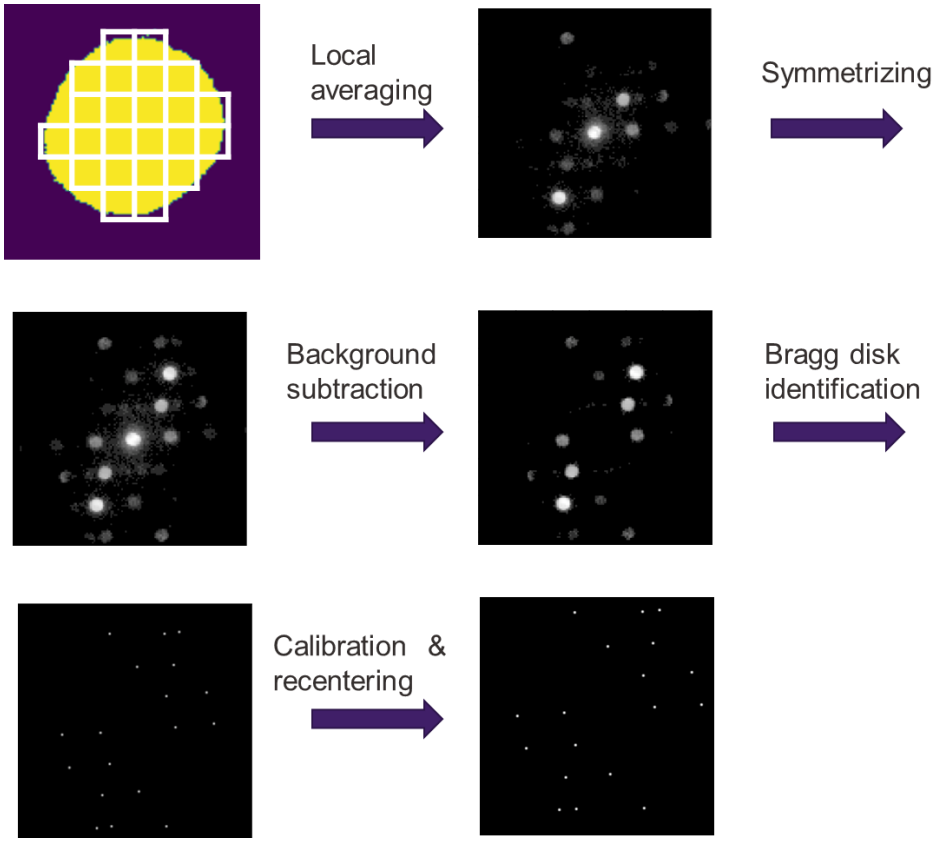


Fig. S7 The flowchart to process diffraction patterns from the nanoparticle (discussed in Section 2.5 and 3)**.** First, the nanoparticle is discretized into different regions (white squares overlayed on the nanoparticle). Next, the diffraction patterns (DPs) are averaged inside each local region (local averaging). Next, the Bragg disks in experimental DPs are symmetrized to compensate missing Bragg disks. Following that, the noisy background is cleared, and the Bragg disks are identified. After recentering around the center disk, the location and intensity of Bragg disks are extracted.

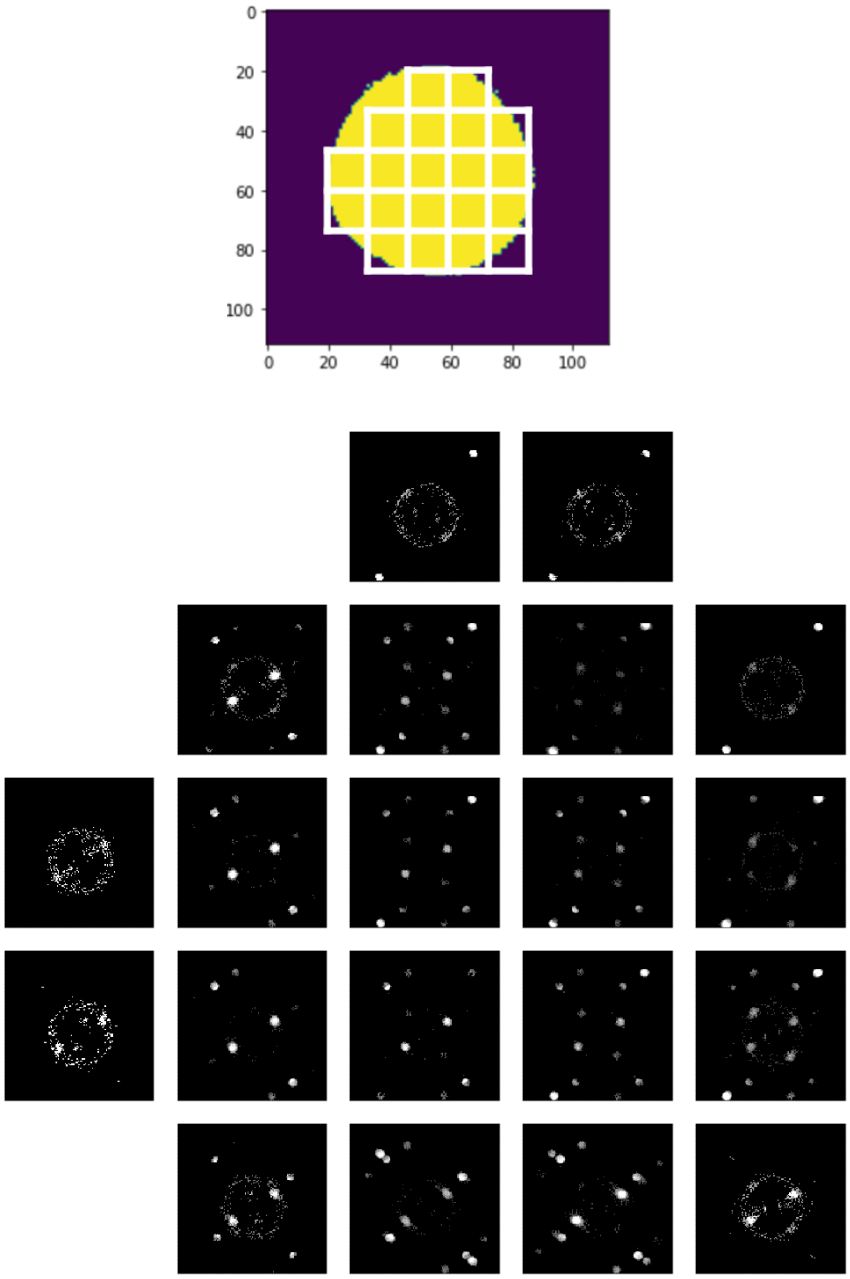


Fig. S8 DPs for Nanoparticle 1 (discussed in Section 2.5)**.**

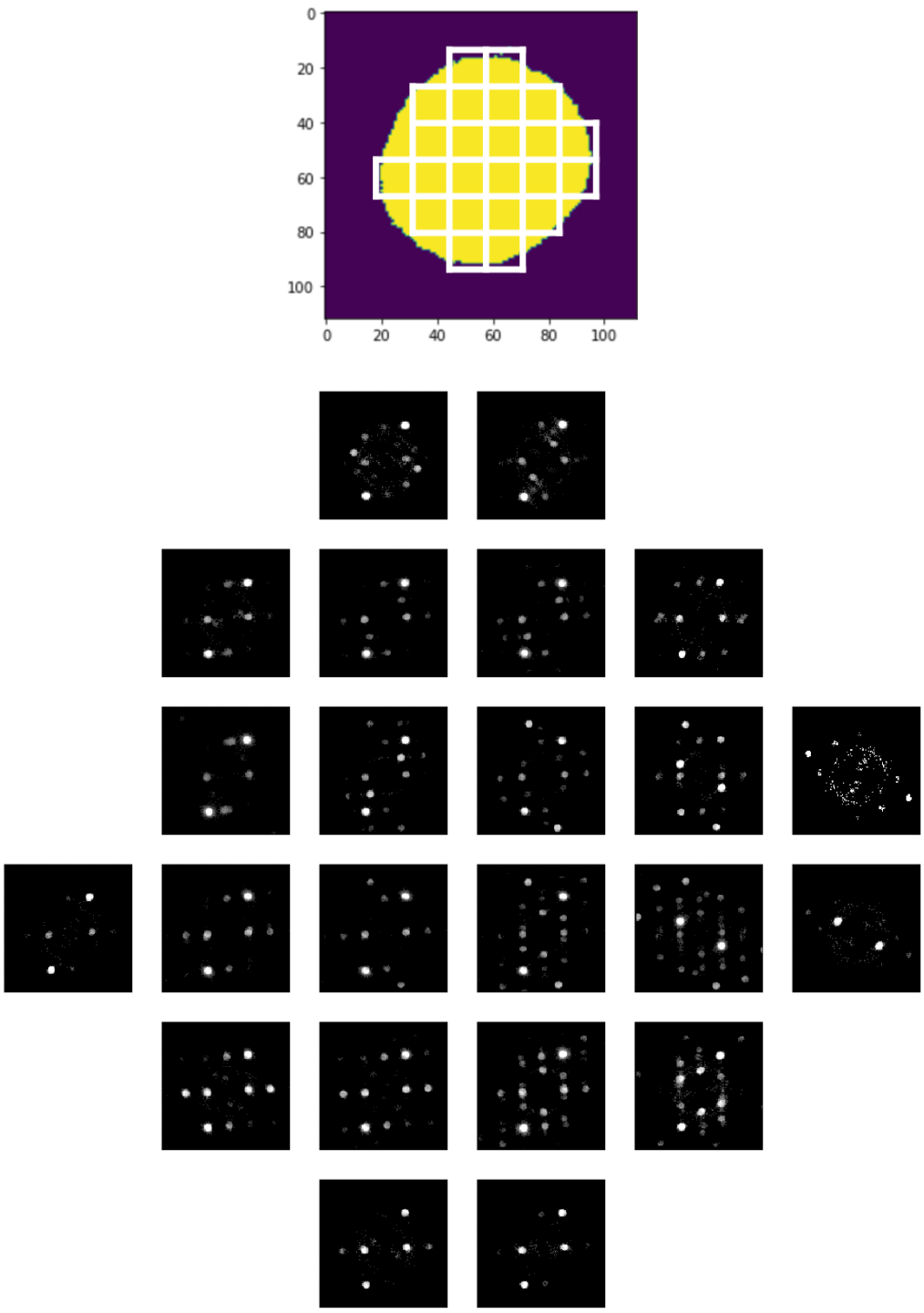


Fig. S9 DPs for Nanoparticle 2 (discussed in Section 2.5)**.**

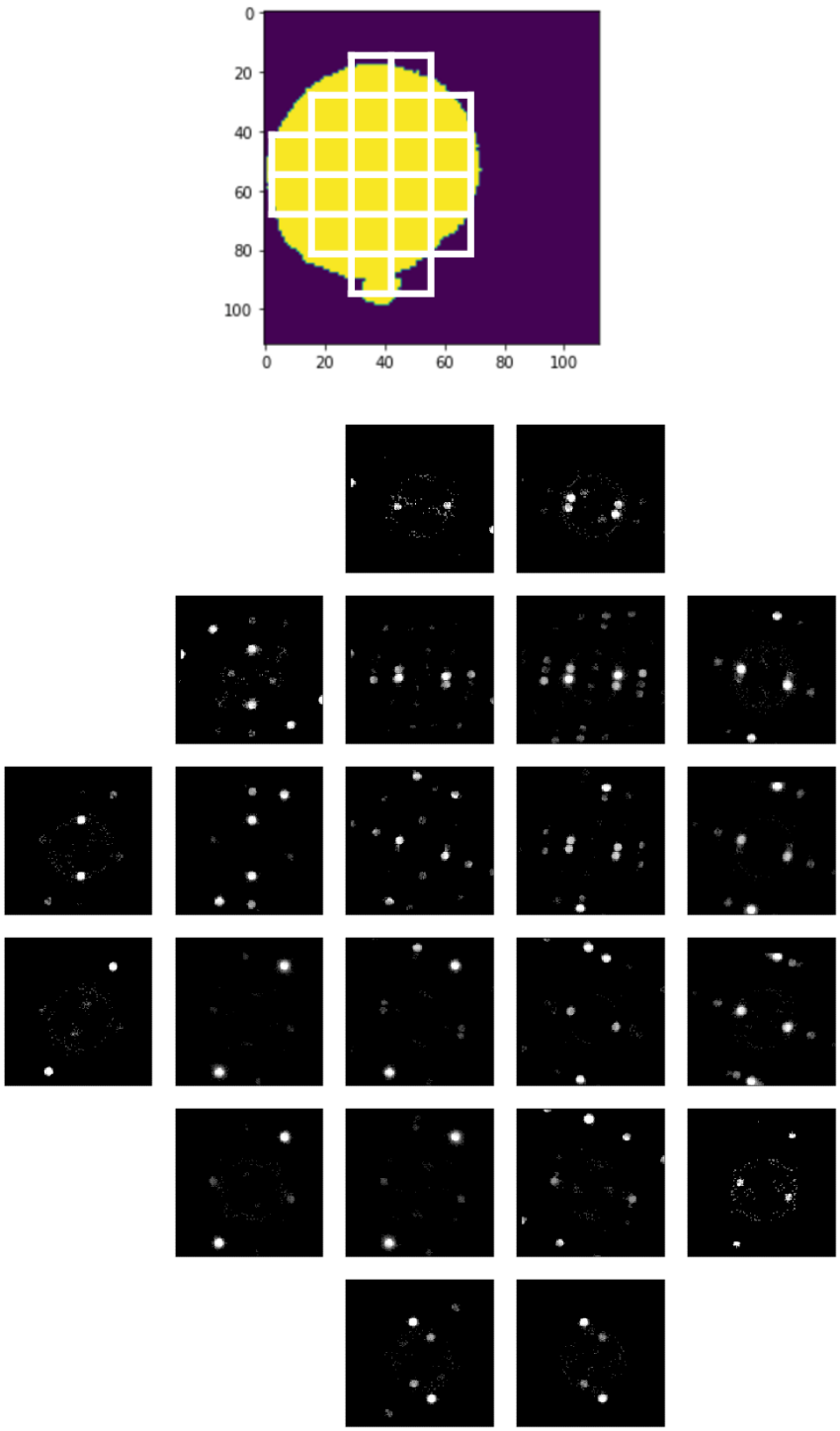


Fig. S10 DPs for Nanoparticle 3 (discussed in Section 2.5)**.**

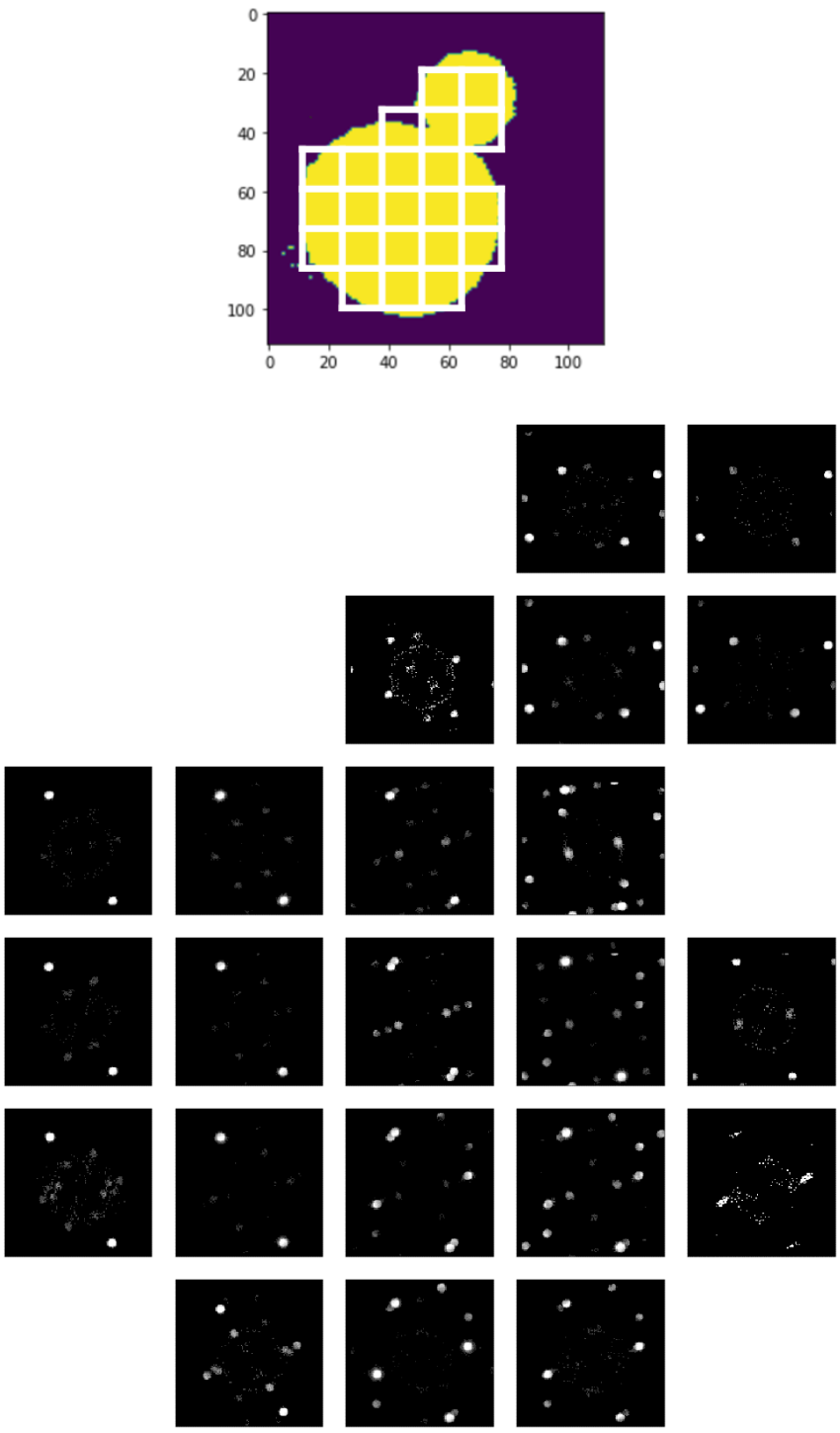


Fig. S11 DPs for Nanoparticle 4 (discussed in Section 2.5)**.**

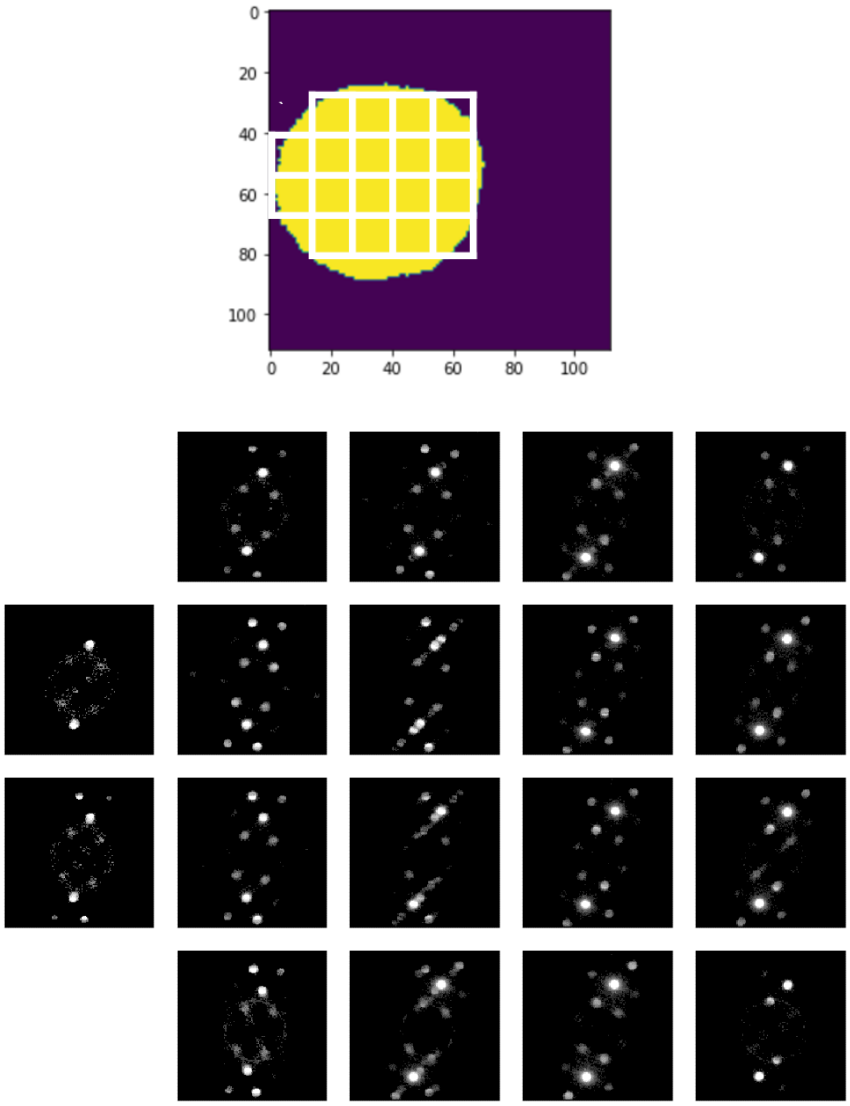


Fig. S12 DPs for Nanoparticle 5 (discussed in Section 2.5)**.**

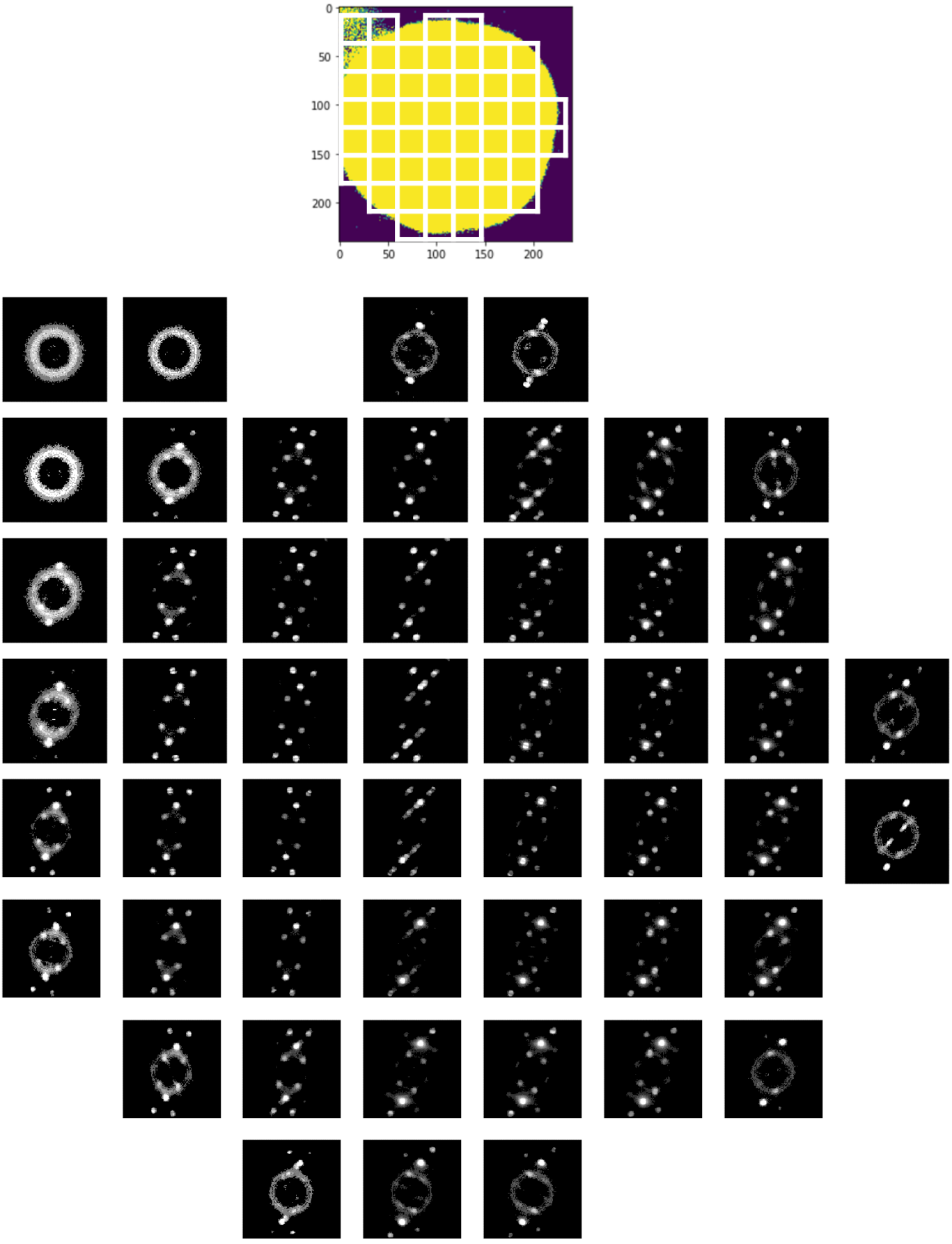


Fig. S13 DPs for Nanoparticle 6 (discussed in Section 2.5)**.**

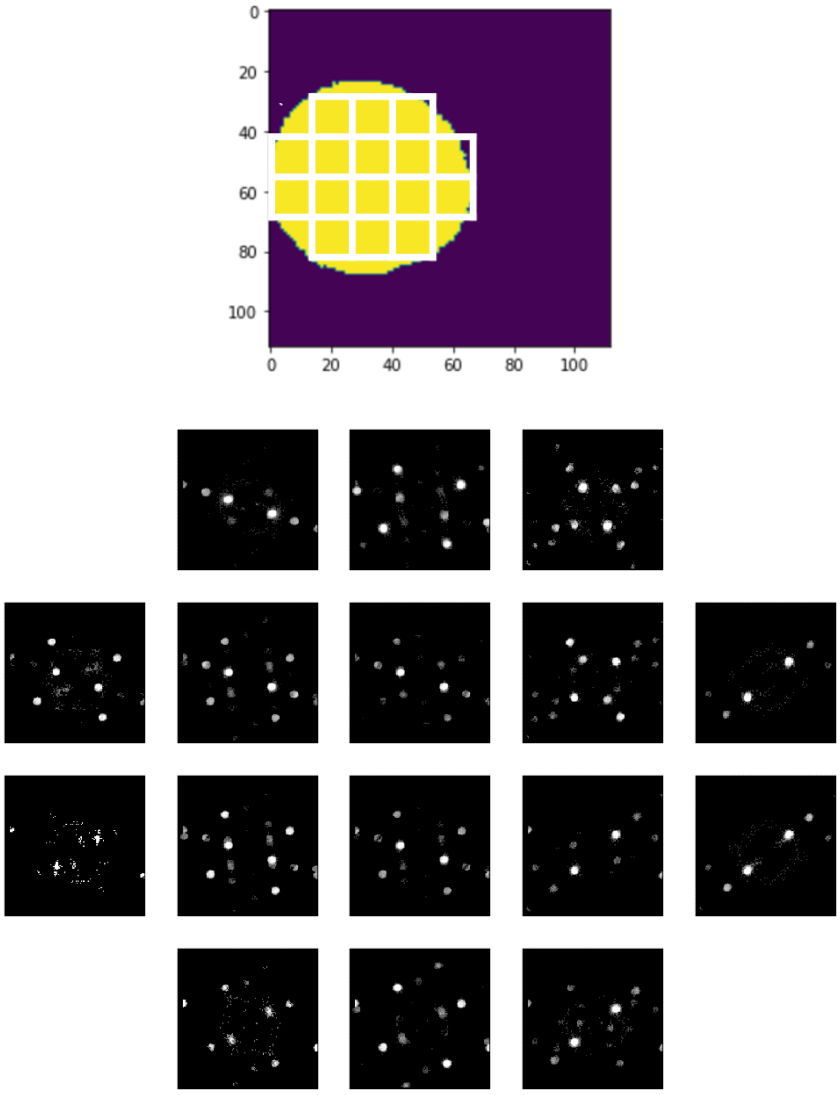


Fig. S14 DPs for Nanoparticle 7 (discussed in Section 2.5)**.**

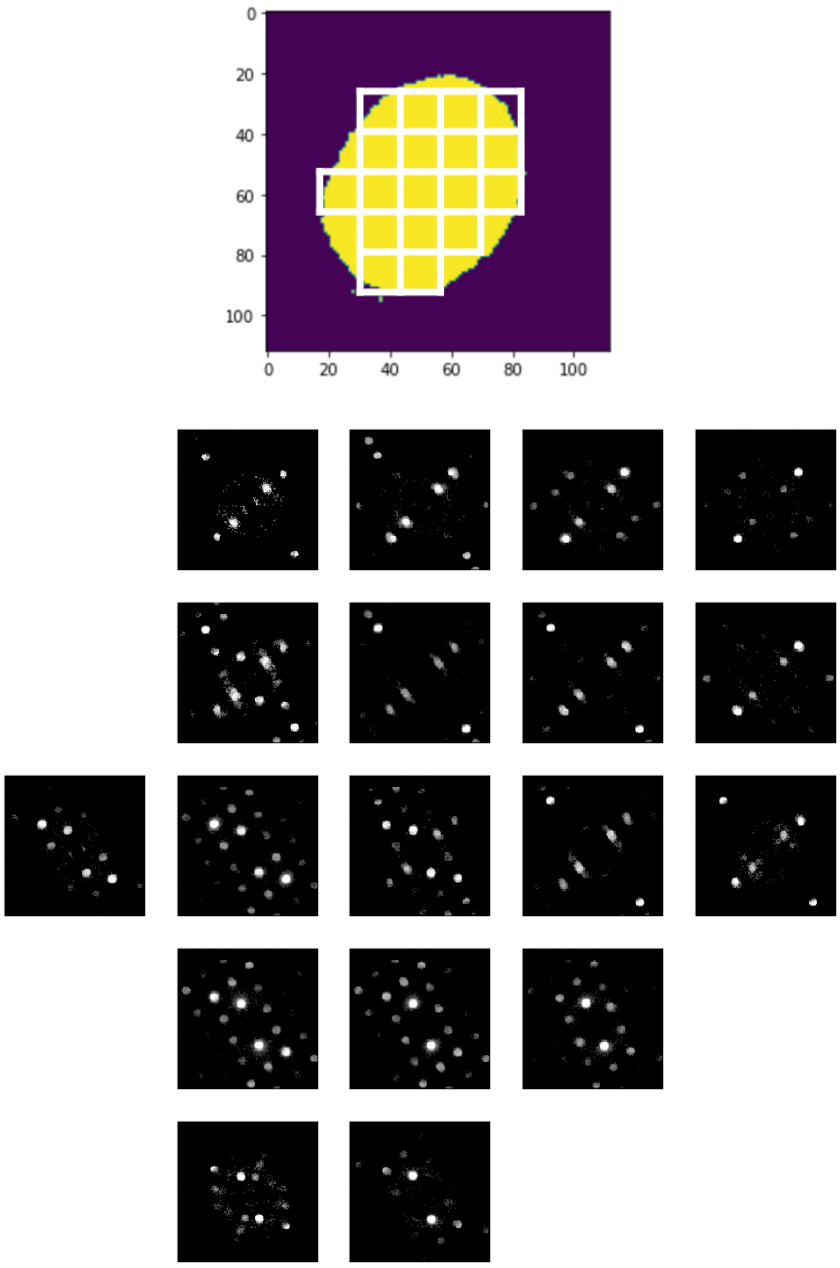


Fig. S15 DPs for Nanoparticle 8 (discussed in Section 2.5)**.**

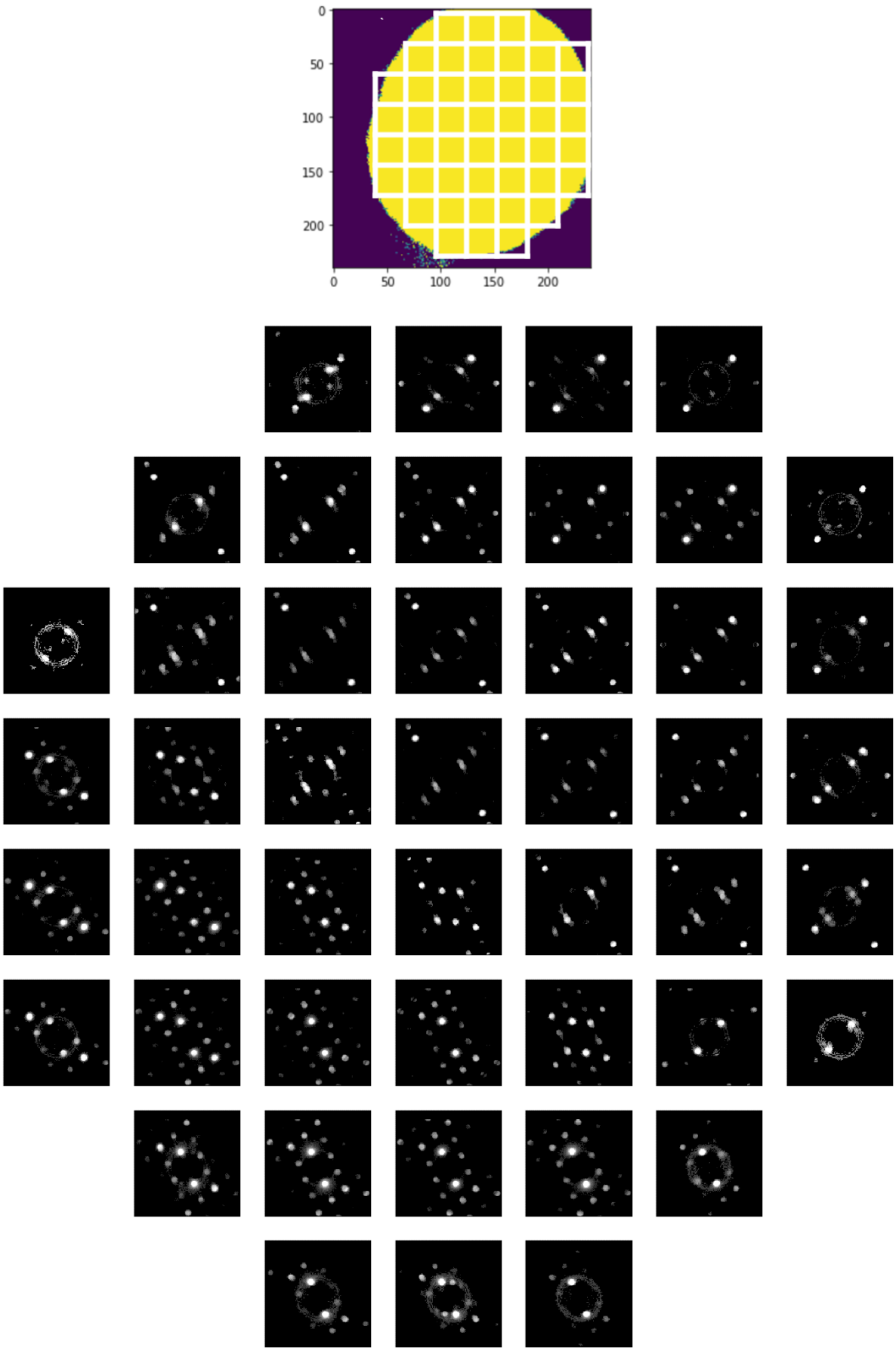


Fig. S16 DPs for Nanoparticle 9 (discussed in Section 2.5)**.**

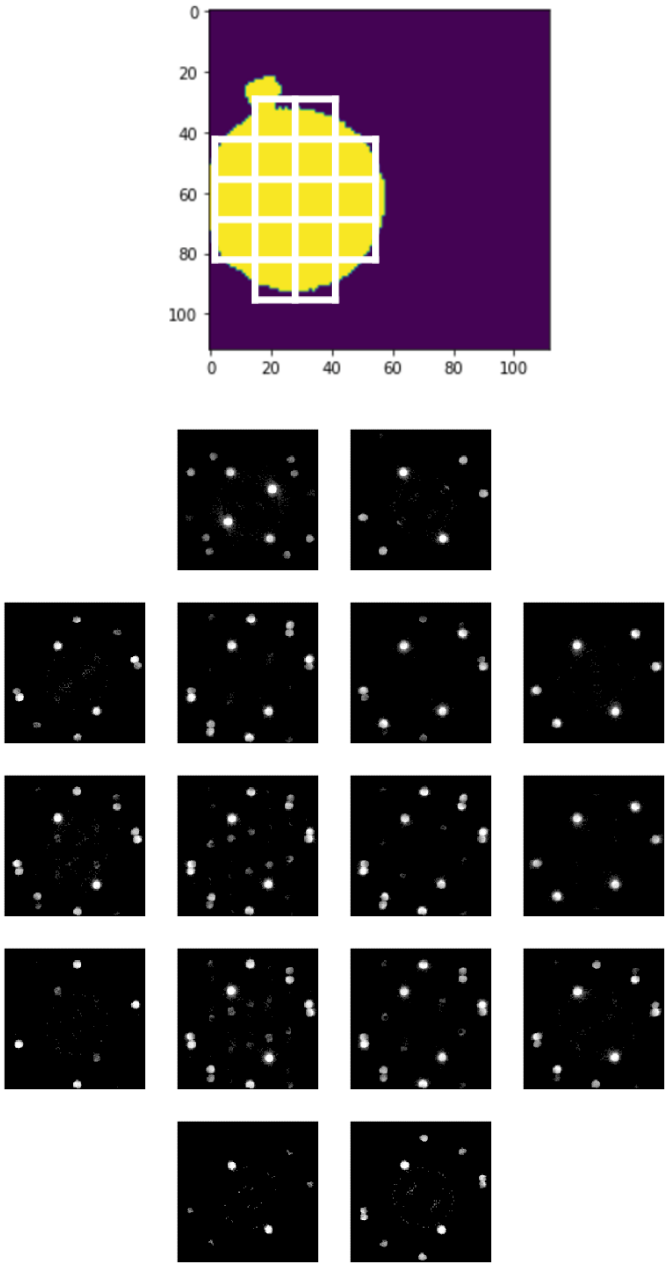


Fig. S17 DPs for Nanoparticle 10 (discussed in Section 2.5)**.**

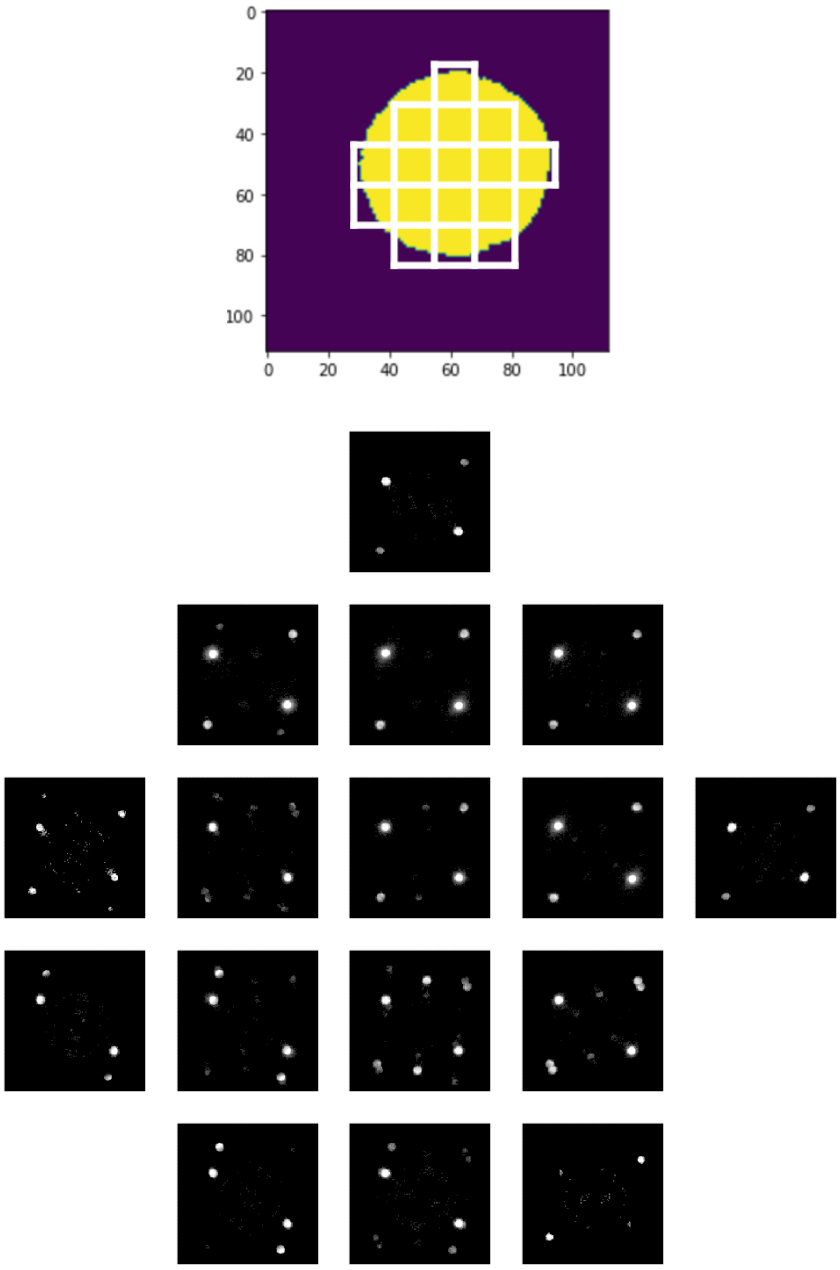


Fig. S18 DPs for Nanoparticle 11 (discussed in Section 2.5)**.**

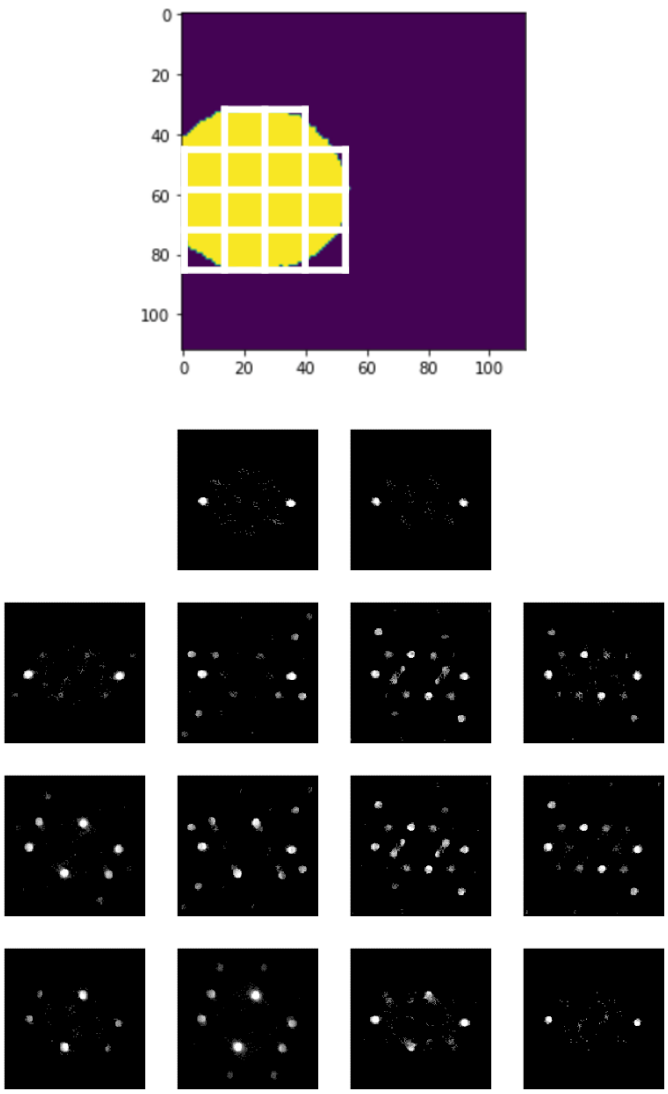


Fig. S19 DPs for Nanoparticle 12 (discussed in Section 2.5)**.**

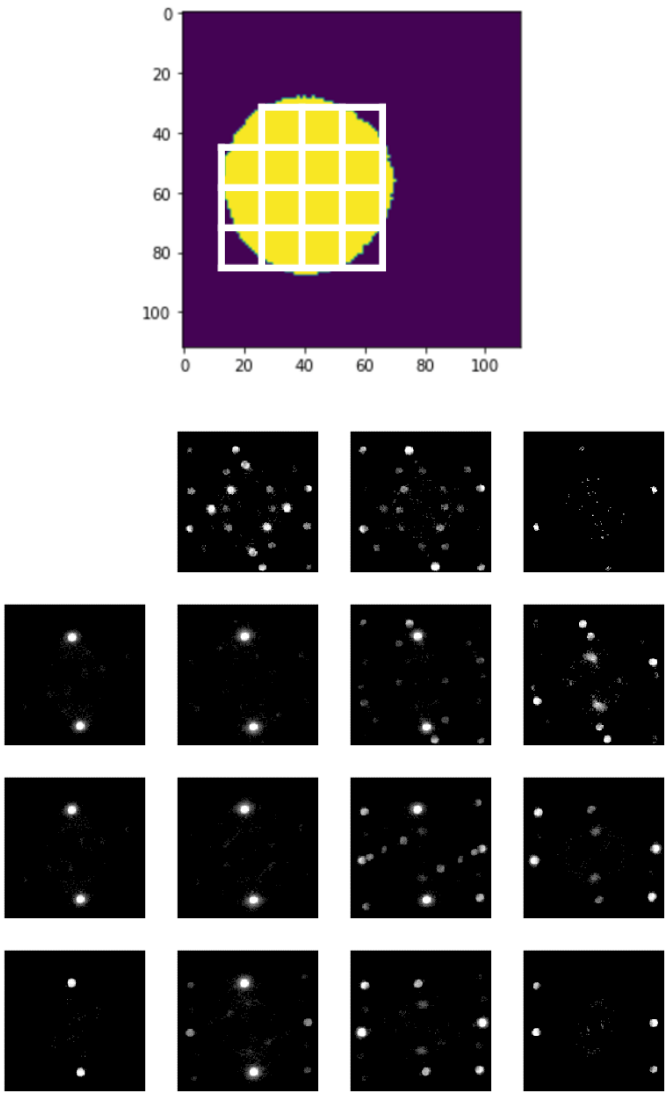


Fig. S20 DPs for Nanoparticle 13 (discussed in Section 2.5)**.**

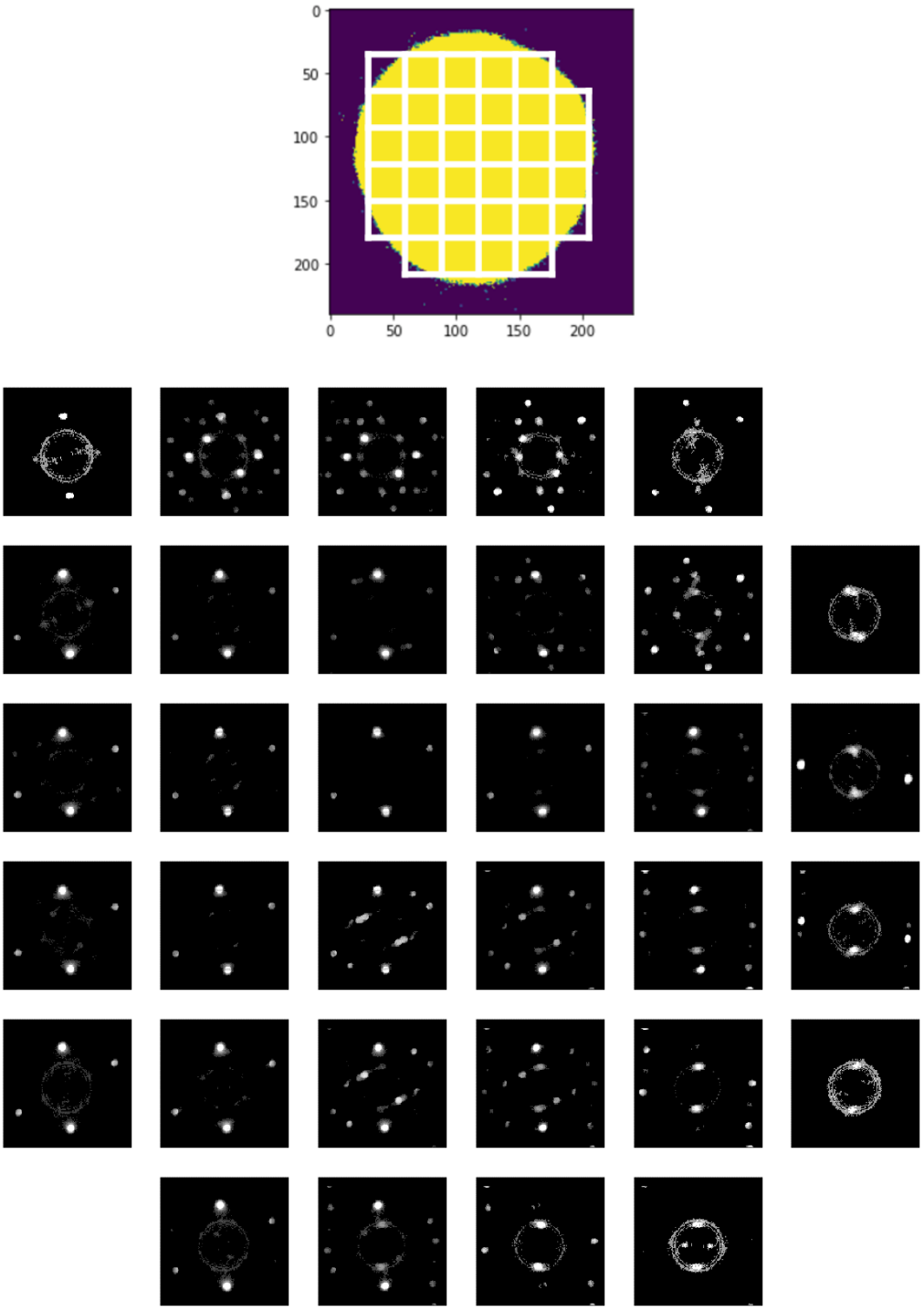


Fig. S21 DPs for Nanoparticle 14 (discussed in Section 2.5)**.**

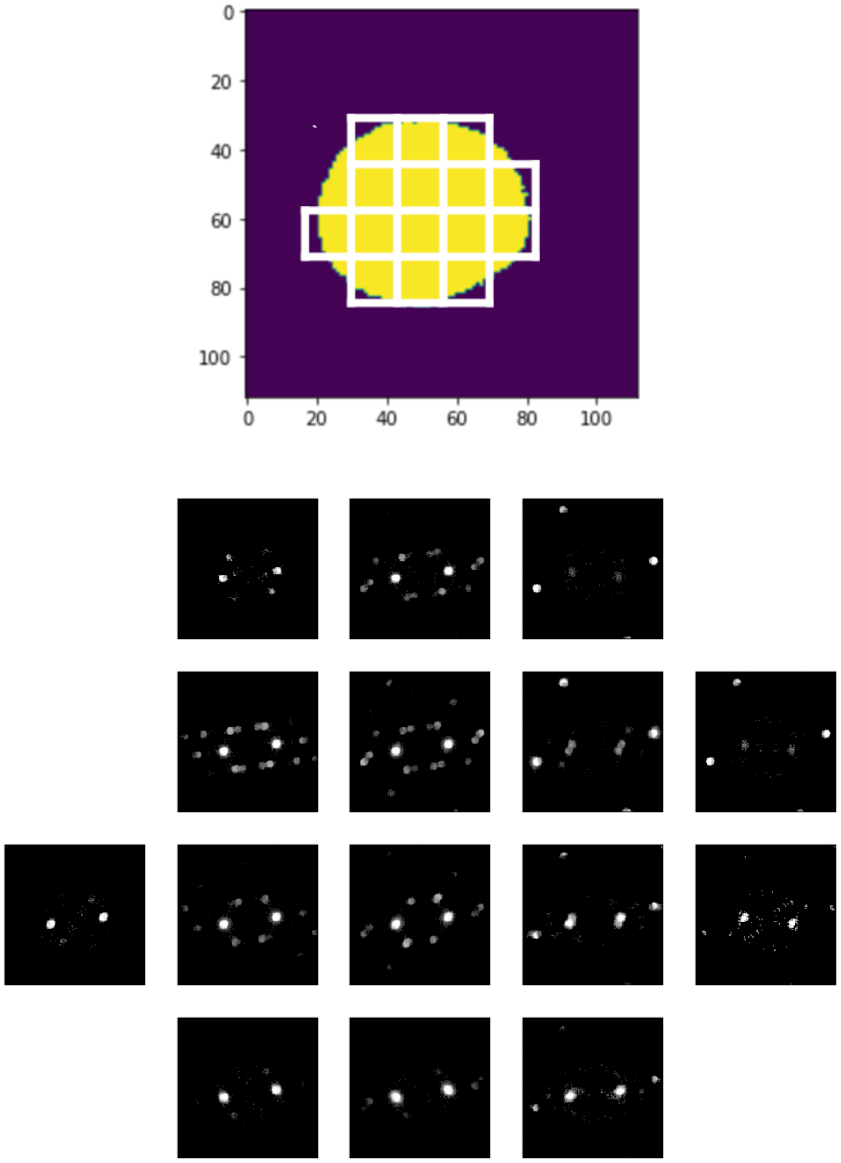


Fig. S22 DPs for Nanoparticle 15 (discussed in Section 2.5)**.**

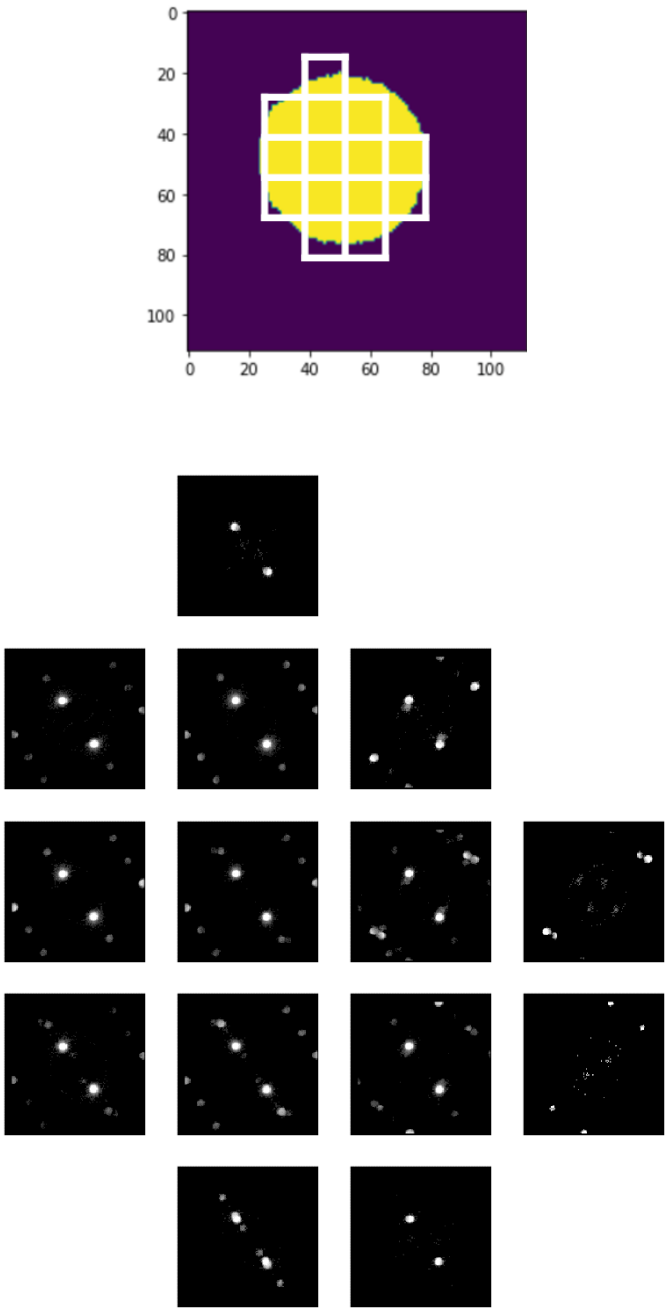


Fig. S23 DPs for Nanoparticle 16 (discussed in Section 2.5)**.**

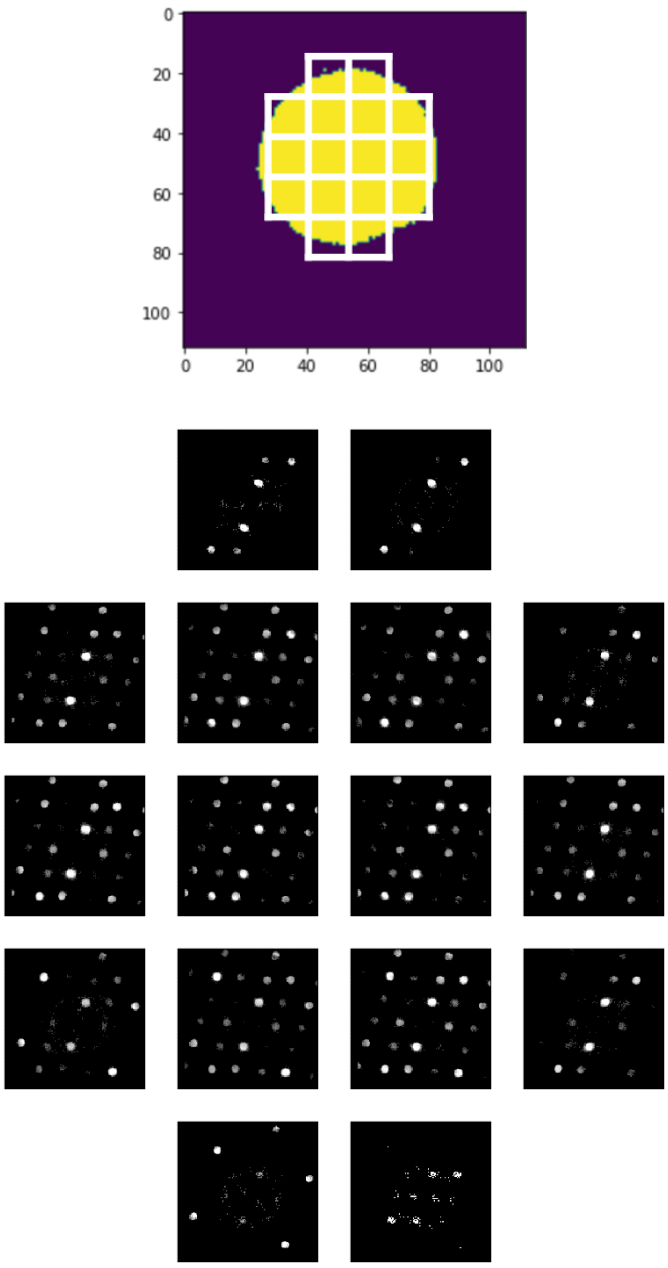


Fig. S24 DPs for Nanoparticle 17 (discussed in Section 2.5)**.**

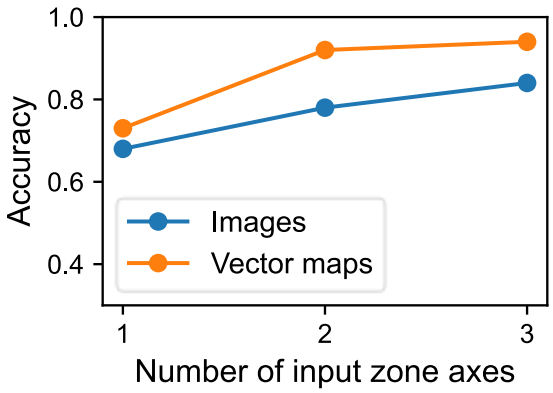


Fig. S25 Influence of image or vector map inputs and number of input zone axes (discussed in Section 3)**.** Predictive accuracy comparison between image and vector maps as inputs. We tested the predictive performance given 1 (zone axis [1 0 0]), 2 (zone axes [1 0 0] and [0 0 1]), and 3 (zone axes [1 0 0], [0 0 1] and [0 1 0]) images or vector maps as inputs. The accuracy is always higher when vector maps are used as DP descriptors.