Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning

J. A. Aguiar1*, M. L. Gong1, R. R. Unocic3, T. Tasdizen2, B. D. Miller1

While machine learning has been making enormous strides in many technical areas, it is still massively underused in transmission electron microscopy. To address this, a convolutional neural network model was developed for reliable classification of crystal structures from small numbers of electron images and diffraction patterns with no preferred orientation. Diffraction data containing 571,340 individual crystals divided among seven families, 32 genera, and 230 space groups were used to train the network. Despite the highly imbalanced dataset, the network narrows down the space groups to the top two with over 70% confidence in the worst case and up to 95% in the common cases. As examples, we benchmarked against alloys to two-dimensional materials to cross-validate our deep-learning model against high-resolution transmission electron images and diffraction patterns. We present this result both as a research tool and deep-learning application for diffraction analysis.

INTRODUCTION
Augmented analysis—algorithm-powered, human-guided data exploration—creates a powerful symbiosis between researchers and their analytical tools. The breadth of data collected simultaneously in the latest generation of scanning transmission electron microscopes (STEMs) presents opportunities for notable advancements in microscopy, multimodal data analytics, and materials research (1). Inside a STEM, materials are imaged for their structure, chemistry, and response to environmental stimuli. Arguably, some of the most useful information gained from this materials characterization approach is the technology to identify materials and their specific atomic arrangements as structure based, with atomic-scale imagery derived from high-resolution imaging or diffraction-based techniques. With the advent of advanced methods of ptychography, precession-based electron diffraction, compressive imaging, and multidimensional imaging—including four-dimensional microscopy—researchers have developed an urgency to identify crystallographic structures and deviations efficiently in more massive datasets (2–6).

The breadth of simultaneously collected data in the latest generation of STEM also presents these same opportunities and more, including accelerated multidimensional data collection and real-time processing. With the native resolution of the microscopes, which varies with each microscope from approximately 0.50 to over 2 Å (limited primarily by spherical aberration), these tools are naturally high-information machines. As experimental techniques, as well as microscopy hardware, evolve—including the recent additions of pixeled detectors—researchers need to address the proliferation of data (7–9). Recent advancements in deep learning have made it possible to analyze human-intractable datasets and perform complex imaging tasks, including segmentation and classification (10–13). However, deep learning and augmented analysis have not yet disrupted the microanalysis and materials community as they have in the computer vision community, despite open source implementations of machine learning algorithms in packages such as TensorFlow and Caffe (14, 15).

In practical terms, learning about material structure is not limited by resolution, but by the ability to process information in the context of atomic arrangements regardless of resolution or technique. In the past, researchers have addressed the identification of structure from x-ray and neutron diffraction data using curve-fit models, Rietveld analysis, peak matching, exhaustive search, and first-hand material knowledge (16–19). Past studies from electron microscopy further reveal the complications of identifying phases and structures. Off-axis and obscure axis orientations complicate diffraction data complexity, potentially causing researchers to overlook identifying diffracting peaks. Texturing issues are addressed, in part, by precession electron diffraction, electron ptychography, and multidimensional STEM imaging. Now, though, processing the abundance of information generated by these techniques has further limited our ability to use all forms of information (20, 21). In the past, George and Wang (22) reported a hybrid system to identify diffraction patterns, but this system was not extended to include all material structures and patterns. The opportunity lies in limiting the number of patterns and extending predictive models to all material and crystallographic types. If we were to use modeling and predictive capabilities to determine pending structure without a priori—or ab initio—based knowledge, we could significantly advance experimental research, especially for those materials composed of multiple orientations and phases. Examples include engineered composites and multilayer systems. Without a predictive approach for translating raw structural data from either high-resolution atomic-scale images or diffraction profiles, the ability to identify unknown data is restricted to computationally limited curve-fitting models and expertise.

Advances in computer vision and artificial intelligence present the opportunity for automation and augmentation of crystallographic and high-resolution imaging data (23). Deep-learning models have been applied to many classification, segmentation, and compression challenges in the computer vision community (24–26). A limited number of machine or deep-learning models have reportedly proposed and demonstrated subimage sampling in image segmentation and inpainting (27, 28). Current event tracking and augmentation of microscope feeds, however, are limited primarily by postanalysis implementations, with the main application being in situ microscopy. The barrier to the independent crystallographic structural determination is translating the body of materials knowledge from current databases into accessible data workflows not limited to high-resolution electron microscopes and diffractometers (29).

In this report, we present an approach to expanding the use of deep learning for crystal structure determination based on diffraction or...
atomic-resolution imaging without a priori knowledge or ab initio–based modeling. Various machine learning models, such as Naïve Bayes, decision forest, and support vector machines, were tested before concluding that convolutional neural networks (CNNs) produce the model with the highest accuracy. To determine crystallography from these data types, we have trained a CNN to perform diffraction-based classification without the use of any stored metadata. The CNN model has been trained on a dataset consisting of diffracted peak positions simulated from over 538,000 materials with representatives from each space group. In summary, we assess the growing potential for crystallographic structure predictions using deep learning for high-throughput experiments, augmenting our ability to readily identify materials and their atomic structures from as few as four Bragg peaks.

RESULTS
Deep-learning model for evaluating crystallographic information
We validated the neural network architecture and workflow based on high-resolution STEM imaging and electron diffraction from crystalline strontium titanate (SrTiO$_3$ or STO) islands on a face-centered cubic structured magnesium oxide (MgO) substrate. Figure 1A is an atomic mass contrast STEM image of the overall sample, with the crystalline portion outlined. Using the high-resolution capabilities of an aberration-corrected STEM with sub-Angstrom resolution, Fig. 1B is an atomically resolved high-angle annular dark-field image of the individual Sr, Ti, and O atoms, preferentially oriented along the [100] zone axes configuration illustrated in Fig. 1C by atomic species. Figure 1D is a fast Fourier transformation (FFT) of the atomically resolved image taken along the same orientation, along with this preferred crystallographic direction. On the basis of the FFT, Fig. 1E is the two-dimensional azimuthal integration of the pattern transformed into a one-dimensional profile. The pattern and profile provide the structural classification details for classifying and predicting the structure using our deep-learning model approach.

Exploiting deep learning–based classification for crystallographic information builds and trains on public and established materials databases, including the Open Crystallography Database, Materials Project Database, American Mineralogist Crystallographic Databases, and Inorganic Crystal Structure Database (30–35). Figure 1F is the

Fig. 1. Neural network data architecture and workflow for crystal space group determination from experimental high-resolution atomic images and diffraction profiles. (A) Any material in an STEM, in this case crystalline STO islands distributed on a rock salt MgO substrate, can be simultaneously imaged with (B) high-resolution atomic mass contrast STEM imaging and (C) decoupled with a selective area. (D) FFT to reveal the material’s structural details. (E) On the basis of either electron diffraction or FFT of an atomic image, a two-dimensional azimuthal integration translates this information into a relevant one-dimensional diffraction intensity profile from which the relative peak positions in reciprocal space can be indexed. arb. units, arbitrary units. (F) Seeding the prediction of crystallography is a hierarchical classification using a one-dimensional convolution neural network model replicated at (G), each layer from family to space group forming a nested architecture. On the basis of the derived peak positions in the azimuthal integration profile, the prediction on STO is reported in Table 2.
one-dimensional CNN architecture used to train and evaluate a hierarchical training dataset composed of 571,340 individual crystals divided among seven families, 32 genera, and 230 individual crystallographic space groups. At each level of the hierarchy, we trained a neural network to form a nested hierarchy for classification, as shown in Fig. 1G. Each CNN consists of six convolutional blocks before three dense layers and a Softmax layer for classification. Convolutional blocks are formed from a convolutional layer, a max pooling layer, and an activation layer. The convolutional layers have a kernel size of two and start with 180 channels, narrowing to 45 channels over the six blocks. Starting in the fourth block, dropout is applied after the pooling layer. Dropout starts at 0.1 and scales up to 0.2 and 0.3 in the fifth and sixth blocks, respectively.

On the basis of the input from the atomic-resolution STO STEM image, the ranked order of predictions made from the FFT image-based profiles is as follows: 225 (\(Fm\overline{3}m\)), 221 (\(Pm\overline{3}m\)), 205 (\(Pa\overline{3}\)). Upon validation with the known crystal structure, we determined that the crystal is structured as space group number 221, \(Pm\overline{3}m\), subsequently oriented along the [100] zone axes containing the [200] and [110] family of crystallographic reflections. Using the data workflow and pipeline provides the generalized framework for classifying all known materials and the flexibility to improve on the model in the future.

## Training on diffraction data for hierarchical structural classification

Initially, there were approximately 650,000 individual structures screened for duplicates or other potential errors. A weighting schema was applied for each class and assigned to address the occurrence of overly represented crystal types, noting substantial imbalances among populated crystal families, groups, and space groups ranging from 136,534 to less than 1000 (36–38). The aggregate accuracies and population statistics for each level of the hierarchy are reported in Table 1. We labeled crystal files by family, genera, and species and used the corresponding label for different levels of the hierarchy to train and evaluate models. The table summarizes the population and accuracy overall in seven crystal classes and 32-point symmetry groups. Each file subsequently generated a distinct diffraction profile as a function of the corresponding Bragg angle, using 571,340 crystal structures. We subsequently generated interplanar \(d\)-spacings at each level of the hierarchy to train the model. The resolution in the individual binary signal, consisting of a normalized vector of intensity against the Bragg angle, was set to 0.5°.

Figure 2A is the confusion matrices evaluated over 136,534 randomly chosen structures of 571,340 total structures at the family level, which includes examples of the best and worst classification schemes as shown in Fig. 2B. When compared at the genera level, cubic and orthorhombic confusion matrices illustrate the hierarchy and classification scheme issues of imbalance in the data despite the weighting during training. The example highlights the classification hierarchy analogous to nested network architecture capable of predicting structure at the family, genera, and space group levels. Predictions along the diagonal are true-positive predictions; a concentration of high percentages along the diagonal indicates an accurate model.

## Optimization and modeling pipelining for crystallographic analysis

Because of a lack of grand canonical examples or a human baseline, the deep-learning model was benchmarked by comparing it to other machine-learning algorithms (39). The model architecture was tested at varying depths, numbers of parameters, layer combinations, and

| Table 1. Population statistics and training accuracy evaluated. Reported population for individual structures over family and genera constituting a database of 571,340 total structures. Note that N/A appears for those genera with a single class. |
|---------------------------------|--------|--------|
| **Accuracy (%)**               | **Population** |
| Triclinic                       | 91.04% | 105,200 |
| Pedial                          | N/A    | 16,740  |
| Pinacoidal                      | N/A    | 88,460  |
| Monoclinic                      | 86.73% | 217,156 |
| Sphenoidal                      | 86.50% | 21,705  |
| Domatic                         | 74.95% | 14,997  |
| Orthorhombic                    | 79.75% | 104,526 |
| Rhombic-diphenoidal             | 77.18% | 22,182  |
| Rhombic-pyramidal               | 82.32% | 19,722  |
| Rhombic-dipyrimal               | 67.42% | 62,622  |
| Tetragonal                      | 84.81% | 40,770  |
| Tetragonal-pyramidal            | 65.80% | 1,081   |
| Tetragonal-disphenoidal         | 76.46% | 1,437   |
| Tetragonal-diagonal-Pyramidal   | 96.23% | 5,112   |
| Tetragonal-trapezohedral        | 82.99% | 2,373   |
| Ditetragonal-pyramidal          | 88.16% | 1,450   |
| Tetragonal-scalenohedral        | 84.22% | 4,309   |
| Ditetragonal-diagonal-Pyramidal | 81.14% | 25,008  |
| Trigonal                        | 82.78% | 31,252  |
| Trigonal-pyramidal              | 81.48% | 3,499   |
| Rhombohedral                    | 90.46% | 6,017   |
| Trigonal-trapezohedral          | 94.13% | 2,373   |
| Ditetragonal-pyramidal          | 92.01% | 5,831   |
| Ditetragonal-scalenohedral      | 89.43% | 13,584  |
| Hexagonal                       | 86.07% | 24,147  |
| Hexagonal-pyramidal             | 88.17% | 1,828   |
| Trigonal-dipymidal              | 90.14% | 453     |
| Hexagonal-dipymidal             | N/A    | 2,106   |
| Hexagonal-trapezohedral         | 99.24% | 930     |
| Ditetragonal-pyramidal          | 92.01% | 2,909   |
| Ditetragonal-scalenohedral      | 95.36% | 3,230   |
| Cubic                           | 95.47% | 48,289  |
| Tetartoidal                     | 96.75% | 1,842   |
| Diploida                        | 90.56% | 1,389   |
| Gyroidal                        | 93.68% | 737     |
| Hexagonal-dipymidal             | 97.14% | 6,475   |
| Hexoctahedral                   | 87.29% | 37,846  |

dropout rates. The final model architecture selected consists of six blocks of convolution. Max pooling and dropout were selected on the basis of performance, the number of trainable parameters, and preservation of spatial information and then by tuning the last block of dense layers. We further determined three dense layers before classification was optimal for producing the highest accuracy.

During hyperparameter optimization, a batch size of 1000 was chosen in conjunction with weighting by the class occurrence to increase the prevalence of rarer classes in the gradient of each batch. The number of peaks used to classify structure had a significant effect on the prediction accuracy. Figure 3 compares the number of peaks present with the measured accuracies for each family based on the accompanying confusion matrices. The confusion matrices are organized across seven crystal families at the genera level, which constitutes the class hierarchy, followed by the number of peaks used. An identical effect is seen at the space group level as well. Comparisons to other machine-learning methods, including decision forests, support vector machines, and the Naïve Bayes model, are reported in table S1.

We further evaluated the deep-learning model for real-time analysis on a single graphical processing unit (GPU) desktop machine to evaluate the efficiency, sensitivity, and computing necessity for performing augmented crystallographic determination in an accessible manner. The single GPU machine used at runtime had a GTX 1050ti 3-GB graphics card and a 3.2-GHz i7 quad-core processor. Running on this single GPU desktop, the model can classify batches of 1000 profiles loaded in sequence at a rate of 2600 to 3500 predictions/s. Conversely, when the same profiles are loaded individually, the model classifies significantly more slowly at a rate of 29 predictions/s. Classification speeds are the same for predictions across families, genera, or space groups. Predicting space groups for 48,289 observations consisting of a single family took 13.2 s, which is equivalent to 3525 predictions/s. Subsequently, the ability to classify structure from diffraction in subsecond times allows for significant acceleration in the acquisition and prediction of at least two orders of magnitude based on the current ability of experts to predict and augment the analysis without previous knowledge.

**DISCUSSION**

**Deep learning for crystallographic prediction from high-resolution imaging and diffraction data**

A generalized workflow and accompanying network for crystallographic materials such as STO at the nanometer scale in Fig. 1A...
provide the necessary capability to derive crystallographic structure from high-resolution images. Figure 1B is an atomic-resolution image that translates images into an accompanying crystallographic pattern in Fig. 1C, where an azimuthal integration in Fig. 1D resolves the individual interatomic $d$-spacings and accompanying Bragg angles for subsequent crystal prediction and refinement. The input to the model seeds a deep-learning model for nested prediction, using over 571,340 crystals to provide a capability for deriving crystal structure. Top-two accuracy is used to discriminate between classes.

The deep-learning model was validated on the basis of experimental data acquired in various modalities using several different microscopes. On the basis of the input from atomic-resolution STO, the model predicts space group 225 ($Fm\overline{3}m$) and 221 ($Pm\overline{3}m$) as the two most likely space groups. Through further analysis, we determined that the crystal is structured as space group number 221, $Pm\overline{3}m$, oriented along the [100] zone axes containing the [200] and [110] family of crystallographic reflections.

**Training on diffraction data for hierarchical structural classification**

Initially, there were approximately 650,000 structure-based files that needed to be cleaned and simulated. The files were screened for formatting errors, missing essential information, and simulation errors. The files were used to simulate diffraction profiles while they were checked against the crystal information file (CIF) file to ensure proper simulation. The profiles were further refined to a binary signal of peak positions. The simulated signals contained several prominent peaks and dozens of lesser peaks that may not be present in all experimental settings and differed between electron, x-ray, and neutron experiments. In the case of electron microscopy, the intensity does not necessarily scale against known structure factors and is strongly affected by material texturing. For full-scan x-ray and neutron data in which the intensities scale against known structure factors, an applied threshold to remove peaks below the signal-to-noise ratio to seed the prediction with fewer peaks leads to predicting crystal with a high degree of accuracy, well above 80%. The prominent diffraction peaks are the most reliable indicator of the structure. The previous crystallography analysis tools further corroborate the model. On the basis of a binary representation of the data as a function of peak position, we trained the hierarchical model on signals for each family and genera, removing peak intensity as a variable, which simplifies the representation.

Moving to the binary representation of peak positions eliminated the intensity of the peaks and allowed the models to be applied to several diffraction-based modalities. A simulated diffraction profile consisted of lattice spacings as a function of either Angstroms or integrated Bragg. Because of constraints on the number of learnable parameters, the Bragg angle resolution was 0.5°. On the basis of a survey of peaks at 0.5°, it is a reasonable resolution for classification in the cases of...
60- to 300-kV electron beams, typically operating voltages of modern electron microscopes.

If space group classification can be learned from peak locations alone, then aggregate signals for each family, genera, and species were summed across all members and quantitatively compared. The aggregate signals over crystal family are shared in fig. S7. Peak distribution over genera or families at each level uniquely differentiate notable features. At the family level of the hierarchy, significant overlaps in distributions and similarities among genera caused predictions among families to be unreliable. Once the family of the crystal is determined, prediction accuracies rise into the 80 to 98% range.

Preprocessing the data yielded 571,340 simulated signals. The signals are not uniformly distributed among all classes at any level of the hierarchy. There is a strong preference for higher-symmetry structures within the seven crystal families at the genera level, but there does not appear to be a preference at the space group level. The model was trained on the materials contained within each database. The disparity in membership between classes introduces challenges into training deep-learning models.

Datasets that are highly imbalanced are susceptible to mode collapse. In such a case, predicting the most common class yields a high accuracy without discriminating between materials. To counteract the imbalance present in the crystallographic data, we imposed a weighting schema in training. During training, the relative presence of classification was used to compute a weight that would be applied during the loss calculation. The weight was defined as the ratio of the total number of structures per space group over all space groups. The same weighting schema was performed at the point group level.

In cases where the dataset is relatively balanced, the number of examples in each class is within the same order of magnitude; the weights are similar enough that they do not bias the model. When a dataset is highly imbalanced, the number of examples in each class differs by more than an order of magnitude. This schema penalizes the model for incorrectly predicting a rarer space group more harshly than it rewards the model for correctly predicting a common space group. Weighting the rare classes to be more important to the model during training had an ameliorating effect on the data imbalance but did not eliminate it. Models trained without this weighting schema suffered mode collapse and would not predict rare classes. To account for and further mitigate mode collapse from data imbalance, models were trained on top-one accuracy but evaluated on top-two accuracy. Misclassifications are predominantly to the common class, and top-two accuracy of the ranked predictions allows the model, in many cases, to correct for misclassifications due to imbalances in the data. The relative score from the output of the Softmax layer determines the rank order.

Optimizing a deep-learning framework for crystallographic prediction

Initial test models showed that attempting to classify directly to species produced models with poor accuracy and compounding effects from mode collapse. Using a hierarchical model that assumes family, then predicts genera, and then species leads to higher accuracy at each step. Even with cascading error, the method is substantially more accurate. Comparing the confusion matrices from each family highlights a high level of accuracy. In Fig. 2A, there is a high degree of diagonalization in the confusion matrices, indicating correct predictions.

Without a canonical model or baseline accuracy, model benchmarking was performed by comparing results with other machine-learning algorithms. CNNs outperformed other machine-learning methods, including decision forests, support vector machines, and the Naïve Bayes model. In certain situations, random forests appear to outperform CNNs. Upon delving into the random forest models, it is revealed that random forest models are subject to mode collapse. Despite the fact that models have a high accuracy above 80%, the model has not learned to...
distinguish classes. It predicts the class that comprises 80% of the data every time. CNNs are susceptible to mode collapse as well, which occurs most prominently when classifying a crystal into a family.

Optimizing the deep-learning model required tuning varying architectural and training hyperparameters. Architectural parameters included depth, number of parameters, layer combinations, and dropout rates. The final architecture consists of six blocks of convolution. Max pooling and dropout were selected on the basis of performance, many trainable parameters, and preservation of spatial information. The last block of dense layers was tuned, and three dense layers before classification were optimally tuned, producing the highest accuracy.

During hyperparameter optimization, a batch size of 1000 was chosen in conjunction with weighting by the class occurrence to increase the prevalence of rarer classes in the gradient of each batch. The number of peaks used to classify the structure (see Fig. 3) was the hyperparameter that most significantly affected accuracy. The number of peaks included is determined by a threshold of peak intensity applied to the simulated patterns. Stricter thresholds, 80 to 90% intensity of the most prominent peak, produced signals with fewer peaks. Relaxed thresholds below 50% produced signals with increasingly many peaks. Using a threshold stricter than 90% of max peak intensity almost universally eliminates all but the maximum peak.

When optimizing the structural parameters, we found that introducing dropout in early layers of the network prevented it from learning from the sparse signals. Typically, dropout is implemented to prevent overfitting of data by ignoring portions of noisy signals. A heavily processed binary peak signal that only contains peak locations for three to six peaks is used in training. The binary peak signal is generated from the azimuthal integration of an FFT or selected area diffraction pattern (SADP) as a rotationally invariant profile where individual peak locations can be identified. A number of peak-finding techniques can be used; however, experience has led us to implement and tool a moving window type max voting peak finder that populates a binary signal sampled at less than 0.03 Å in real d-spacing. Such sparse vector representations of the data, including dropout early in the model, could eliminate the signal before propagating to learnable features. This resulted in worse models. Dropout was introduced gradually starting in the third convolution block, starting at 0.10 and increasing to 0.30, in the last convolutional layers to prevent overfitting. The six blocks of convolution, max pooling, and dropout condense the signal but maintain the spatial relevance of the original data.

To supplement the data and provide a more robust training regimen for the limited data, we opted to use cross-validation instead of splitting the data into single training, testing, and validation sets. For cross-validation, the data were split into 10 folds; for each fold, we trained a model on the other nine folds. The resulting models were compared with each other to test for overfitting and generalization. This process was repeated for each level of the hierarchy. Validation was performed using experimental data that the models had not seen in training or evaluation. The training and tuning of models were performed on a high-performance NVIDIA DGX-1 system. However, to ensure that the model could be usable in a setting where high-performance computing resources were not available, we performed speed benchmarking on an entry-level single GPU desktop.

### Optimizing the deep-learning model for augmented diffraction

Alongside predictive accuracy, we considered performance on single GPU machines during model tuning and design. Although trained on an high-performance computing (HPC) machine, the model was designed to be deployable on any readily available single GPU machine or an entry-level cloud-based end point. Although the model is capable of classifying at a rate exceeding 3500 predictions/s for a large batch of preprocessed diffraction signals, this does not consider the time necessary to process diffraction patterns into the appropriate input form. The model is capable of handling large backlogs of data with this predictive speed, but a more realistic workflow for newly generated data would be running small batches or sequentially predicting for each observation. When analyzing each diffraction image through the full pipeline, including the azimuthal integration and peak-finding algorithms, there is a significant slowdown in predictive speed, as is expected. This is the case in a current workflow from raw to the processed peak position. At 22 predictions/s, it is possible to achieve real-time analysis of a live camera feed.

### Evaluating the predictive nature of deep learning for materials

After training and tuning the models on a synthetic dataset, it was necessary to validate the models on experimental data. We selected several well-known materials with known crystallographic structures representative of ongoing materials research programs. The sparse sampling of materials allows us to validate the processing pipeline and hierarchical
model. These materials range from cubic to orthorhombic, wherein each case is shown in Fig. 4. The materials are representatives of a crystal family. We started with higher-symmetry cubic crystals in Fig. 4A from low to high resolution for nanocrystalline CeO₂. On the basis of each accompanying FFT, in Table 2, we report the top three predictions with a relative ranking score of each prediction included in parenthesis. For low-angle CeO₂, space group 225, we predicted Fm3m within the top three for each FFT in Table 2.

Beyond cubic structures, graphene constitutes a hexagonal lattice of individual carbon atoms. Figure 4B is a SAED pattern taken along the primary c axis, where the deep-learning model predicts the structure as space group 194, P6₃/mmc in the top three in Table 2. Predictive capabilities were explored in Fig. 4C for a quantum topological insulating material, BSTS, which is arranged in the trigonal structure with space group 53 (R₃m). Using the same model subsequently for lowest symmetry, the deep-learning model predicts the structure as space group 166, R₃m, in the top three in Table 2 for all four zone axes reported in Fig. 4D for α-phase uranium. Studied with SAED along the primary c axis, the deep-learning model predicts the structure as space group 63, Cmcm, in the top three reported in Table 2.

For all of these predictions, the model displays a level of sensitivity to the number of peaks used for classification. In several cases, more than four peaks were detected. Using more than four peaks can present some ambiguity as to which peaks should be used for classification without a priori knowledge. In these cases, we can use simple heuristics to narrow down the selection of peaks. In the case of electron diffraction, we have ignored peak below the resolution limit 0.5 Å and above 8 Å are diffractive limited. Despite this range, this is a robust model to perform generalized crystallographic classification. For example, in the case of BSTS, two different, valid sets of peaks were fed into the model, generating two different sets of predictions. The correct space group of 166, R₃m, was in the top two results for each set, but the third prediction changed. Future explorations of these different possible detected peak combinations may lead to further improvements of the model and removal of ambiguities in the predictions.

CONCLUSION
Our results and model report on how to reliably extract crystallographic structural information in diffraction using a deep-learning nested framework. We have demonstrated how CNNs can adequately account for imaging and diffraction data to effectively extend the limits of human-centric analysis. As a result, even peaks in diffracted low-signal to noisy images can potentially be extracted. This would confirm previous expectations that a combination of machine- and image-processing routines increases the applicability of electron microscopes and diffraction-based techniques for materials research. On the basis of the results, we report on the classification of crystal structures from as low as four distinct Bragg reflections contained in either an atomic-resolution image or diffraction pattern. The simplicity and efficiency of the method, capable of predicting crystallographic structure, reduce artifacts and robustly address the need for efficient cross-validation, surpassing limitations in crystallography of unknown materials.

MATERIALS AND METHODS
High-resolution TEM and diffraction
Microscopy was collected and assembled for training using Angstrom-to-nanometer-sized probes to collect high-resolution images and diffraction patterns from an FEI Titan, JEOL ARM, and JEOL 2800, each operated at 60 or 200 kV and capable 0.8-Å resolution. This included low-, medium-, and high-angle annular dark-field STEM ranging from 1024 by 1024 to 2048 by 2048 pixel images with sub-100-μs dwell times to examine samples.

For electron-transparent TEM, thin samples of nanocrystalline CeO₂ and α-phase uranium were prepared using an FEI Helios Dual Beam focus ion beam (FIB) instrument. These samples were coated with a layer of C to improve sample conductivity and minimize sample drift inside the FIB. Inside the FIB, a gradient of fine- to coarse-grained ion beam platinum layers were deposited over of an area 15 μm by 3 μm with a thickness of 3 μm. Thin foil lift-out proceeded over this rectangular area with final lamellae measured as 13 μm by 5 μm by 100 nm in total thickness. The final lamellae were lifted and mounted to a molybdenum Omnimode TEM grid for examination. A final cleaning was performed using a 5-kV gallium beam and a beam current of 12 pA to remove material deposits during FIB preparation to reduce milling damage. Nanoparticle samples of Ce₉Gd₂O₁₉ were crushed, filtered, and dispersed onto a TEM holey carbon grid for examination. Researchers minimized experimental imaging and diffraction artifacts. The cameras were set up with quadrant-based gain normalization, and the noise threshold was characterized for each camera. The cameras were kept at their operating temperatures for long periods of time before carrying out the experiments to achieve thermal equilibrium, and long wait periods were enforced between acquisitions to confirm that the dark counts were as constant as possible over the time required for a full experimental acquisition. In this fashion, only unstructured white noise was observed at short (~20 ms) and longer (>1 s) time durations after dark reference removal. After this initial setup, subsequent images and diffraction patterns were collected without adjusting any beam or camera parameters.

Collected two-dimensional diffraction patterns and fast Fourier-transformed diffraction patterns were individually azimuthally integrated and converted into a one-dimensional profile. A small count offset between 0 and 1000 counts was applied randomly to the camera to minimize systematic noise. Individual members of the diffraction series were then aligned, and the background profile was subtracted. Provided the acquisition time for individual images is sufficiently small, systematic or correlated noise (e.g., high voltage instabilities, camera noise, and vibrations) is effectively suppressed.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/10/eaaw1949/DC1

Fig. S1. Confusion matrix for monoclinic genera and space group.
Fig. S2. Confusion matrix for orthorhombic genera and space group.
Fig. S3. Confusion matrix for tetragonal family and genus 15.
Fig. S4. Confusion matrix for trigonal family and genera 20.
Fig. S5. Confusion matrix for hexagonal family and genus 25.
Fig. S6. Confusion matrix for cubic family and genera 31.
Fig. S7. Aggregate signals for triclinic, monoclinic, orthorhombic, and cubic families plotted against two theta values.

Table S1. Comparing models for crystallographic determination.

REFERENCES AND NOTES
29. M. Patel and H. Yoon are acknowledged for providing samples and data for analysis. Additional data related to this paper may be requested from the authors. The software is open source. It can be downloaded at https://github.com/SCIInstitute/DiffractionClassification.git.
30. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The authors declare that they have no competing interests.
32. M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The authors declare that they have no competing interests.
33. T. T. and M. L. G. were involved in developing the deep-learning model for diffraction analysis. J. A. A., R. R. U., and B. D. M. performed the microscopy and measurements.
34. M. Patel and H. Yoon are acknowledged for providing samples and data for analysis. Additional data related to this paper may be requested from the authors. The software is open source. It can be downloaded at https://github.com/SCIInstitute/DiffractionClassification.git.
35. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The authors declare that they have no competing interests.
36. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. The software is open source. It can be downloaded at https://github.com/SCIInstitute/DiffractionClassification.git.

Downloaded from http://advances.sciencemag.org/ on May 16, 2021
Decoding crystallography from high-resolution electron imaging and diffraction datasets with deep learning

J. A. Aguiar, M. L. Gong, R. R. Unocic, T. Tasdizen and B. D. Miller

DOI: 10.1126/sciadv.aaw1949