



## Original Article

# Deep Neural Network Enabled Space Group Identification in EBSD

Kevin Kaufmann<sup>1</sup> , Chaoyi Zhu<sup>2</sup>, Alexander S. Rosengarten<sup>1</sup> and Kenneth S. Vecchio<sup>1,2\*</sup> 

<sup>1</sup>Department of NanoEngineering, UC San Diego, La Jolla, CA 92093, USA and <sup>2</sup>Materials Science and Engineering Program, UC San Diego, La Jolla, CA 92093, USA

### Abstract

Electron backscatter diffraction (EBSD) is one of the primary tools in materials development and analysis. The technique can perform simultaneous analyses at multiple length scales, providing local sub-micron information mapped globally to centimeter scale. Recently, a series of technological revolutions simultaneously increased diffraction pattern quality and collection rate. After collection, current EBSD pattern indexing techniques (whether Hough-based or dictionary pattern matching based) are capable of reliably differentiating between a “user selected” set of phases, if those phases contain sufficiently different crystal structures. EBSD is currently less well suited for the problem of phase identification where the phases in the sample are unknown. A pattern analysis technique capable of phase identification, utilizing the information-rich diffraction patterns potentially coupled with other data, such as EDS-derived chemistry, would enable EBSD to become a high-throughput technique replacing many slower (X-ray diffraction) or more expensive (neutron diffraction) methods. We utilize a machine learning technique to develop a general methodology for the space group classification of diffraction patterns; this is demonstrated within the  $(4/m, \bar{3}, 2/m)$  point group. We evaluate the machine learning algorithm’s performance in real-world situations using materials outside the training set, simultaneously elucidating the role of atomic scattering factors, orientation, and pattern quality on classification accuracy.

**Key words:** convolutional neural network, crystal structure, EBSD, machine learning, space groups

(Received 14 January 2020; revised 24 March 2020; accepted 21 April 2020)

### Introduction

Conventional electron backscatter diffraction (EBSD) is a standard scanning electron microscope (SEM)-based technique used to determine the three-dimensional orientation of individual grains in crystalline materials. Phase differentiation is a necessary component of this technique for the analysis of multi-phase samples and has been of particular interest in the community (Britton et al., 2010; Foden et al., 2019, 2019; Hielscher et al., 2019). However, determining the underlying structure of unknown materials (phase identification) has remained a challenge in EBSD. Currently, Hough- or dictionary-based pattern matching approaches require a “user-defined” set of phases at the onset of analysis (Chen et al., 2015; Nolze et al., 2017; Singh & De Graef, 2017; Singh et al., 2018; Tong et al., 2019). The Hough transform-based method is the most common approach to pattern matching used in commercial systems. Hough-based indexing looks for the diffraction maxima and creates a sparse representation of the diffraction pattern (Lassen et al., 1994). The sparse representation is used with a look-up table of interplanar angles constructed from the set of selected reflectors for phases specified by the user. Beyond requiring the user to have sufficient knowledge of the sample before beginning analysis, the process remains susceptible to structural misclassification

(McLaren & Reddy, 2008; Chen & Thomson, 2010; Karthikeyan et al., 2013). Potential phase identification solutions leveraging energy-dispersive X-ray spectroscopy (EDS) or wavelength-dispersive X-ray spectroscopy (WDS) have been previously demonstrated and adopted commercially (Goehner & Michael, 1996; Nowell & Wright, 2004; Dingley & Wright, 2009). These strategies are effective for single-point identification of crystal structure subject to an expert user’s ability to select the correct phase from the potential matches. Methods utilizing hand-drawn lines overlaid on individual Kikuchi diffraction patterns have been developed for determining the Bravais lattice or point group (Baba-Kishi & Dingley, 1989; Goehner & Michael, 1996; Michael & Eades, 2000; Li & Han, 2015). These represent important milestones for phase identification from EBSD patterns; however, they remain limited by at least one of the following: analysis time per pattern, the need for an expert crystallographer, or necessitating multiples of the same diffraction pattern with different SEM settings (Li & Han, 2015).

Going beyond Bravais lattice and point group identification to determine the space group of the crystal phases is, in general, a challenging task. X-ray diffraction (XRD) and transmission electron microscopy (TEM)-based convergent beam electron diffraction (CBED) are the most common solutions (Post & Veblen, 1990; Ollivier et al., 1997). X-ray diffraction becomes challenging in multi-phase samples owing to overlapping peaks, texture effects, large numbers of peaks for low-symmetry phases, and pattern refinement. Moreover, the information in XRD patterns does not provide detailed microstructure information, such as

\*Author for correspondence: Kenneth S. Vecchio, E-mail: [kvecchio@eng.ucsd.edu](mailto:kvecchio@eng.ucsd.edu)

Cite this article: Kaufmann K, Zhu C, Rosengarten AS, Vecchio KS (2020) Deep Neural Network Enabled Space Group Identification in EBSD. *Microsc Microanal* 26, 447–457. doi:10.1017/S1431927620001506

morphology, location, or grain statistics that can be observed using EBSD's better spatial resolution. However, with careful analysis, it is possible to extract crystal symmetry, phase fractions, average grain size, and dimensionality from XRD (Garnier, 2009). On the other hand, CBED is limited by intense sample preparation, small areas of analysis, and substantial operator experience (Vecchio & Williams, 1987, 1988; Williams et al., 1991). In comparison, EBSD can be performed on large samples (Bernard et al., 2019; Hufford et al., 2019; Wang et al., 2019; Zhu et al., 2020), including three-dimensional EBSD (Calcagnotto et al., 2010), with high precision ( $\sim 2^\circ$ ), high misorientation resolution ( $0.2^\circ$ ) and high spatial resolution ( $\sim 40$  nm) (Chen et al., 2011). Furthermore, the diffraction patterns collected in EBSD contain many of the same features observed in CBED, including excess and deficiency lines and higher-order Laue zone (HOLZ) rings (Michael & Eades, 2000; Winkelmann, 2008). Identification of a space group-dependent property (chirality) was recently demonstrated in quartz using single experimental EBSD patterns (Winkelmann & Nolze, 2015). However, state of the art EBSD software cannot currently classify the collected diffraction patterns to their space group and can misidentify the Bravais lattice. Common examples encountered in EBSD include the difficulty distinguishing  $L1_2$  (space group 221) from FCC (space group 225) or B2 (space group 221) from BCC (space group 229) (Gao et al., 2016; Li et al., 2018; Wang et al., 2018). Inspired by the similarities between CBED and EBSD patterns (Vecchio & Williams, 1987; Cowley, 1990; Michael & Eades, 2000), we propose applying an image recognition technique from the machine learning field to provide an opportunity for real-time space group recognition in EBSD. Given that CBED patterns contain sufficient 3-D structural diffraction detail to allow structure symmetry determination to the space group level (Vecchio and Williams, 1987), and given the considerable similarity between CBED and EBSD patterns, along with the much larger angular view captured in EBSD patterns, and the demonstration of chirality determination in experimental patterns (Winkelmann & Nolze, 2015), it is reasonable to consider space group differentiation in EBSD patterns.

The recent advent of deep neural networks, such as the convolutional neural network (CNN) designed for image data, offer an opportunity to address many of the challenges to autonomously extracting information from diffraction data (Ziletti et al., 2018; Oviedo et al., 2019). CNNs are of particular interest owing to multiple advantages over classical computer vision techniques, which require a multitude of heuristics (Wang et al., 2005; Alegre et al., 2006; Lombaert et al., 2014; DeCost & Holm, 2015; Zhu et al., 2020)—such as detecting Kikuchi bands, accounting for orientation changes, determining band width, etc.—and carrying the burden of developing the logic that defines these abstract qualities. Instead, this deep learning technique (e.g., CNNs) determines its own internal representation of the data, via backpropagation (Rumelhart et al., 1986), such that it maximizes performance at the discrimination task. This is the underlying principle behind deep representation learning (i.e., deep neural networks) (LeCun et al., 2015). CNNs operate by convolving learnable filters across the image, and the scalar product between the filter and the input at every position, or “patch”, is computed to form a feature map. The units in a convolutional layer are organized as feature maps, and each feature map is connected to local patches in the previous layer through a set of weights called a filter bank. All units in a feature map share the same filter banks, while different feature maps in a convolutional layer use different filter banks. Pooling layers are placed after convolutional layers to down

sample the feature maps and produce coarse grain representations and spatial information about the features in the data. The key aspect of deep learning is that these layers of feature detection nodes are not programmed into lengthy scripts or hand-designed feature extractors, but instead are “learned” from the training data. CNNs are further advantageous over other machine learning models since they can operate on the unprocessed image data and the same architectures are applicable to diverse problems. For example, a similar methodology was recently demonstrated to identify the Bravais lattice of an experimental EBSD pattern (Kaufmann et al., 2020). Another recent example has applied a CNN to simulated EBSD patterns from eight materials that are typically confused in conventional analyses (Foden et al., 2019) with exceptional success.

In the present work, it is demonstrated that convolutional neural networks can be constructed to rapidly classify the space group of singular EBSD patterns. This process is capable of being utilized in a real-time analysis and high-throughput manner in line with recent advancements in EBSD technology (Goulden et al., 2018). The method is described in detail and demonstrated on a dataset of samples within the  $(4/m, \bar{3}, 2/m)$  point group. The dataset utilized further allows for studying the impact that scattering intensity factors (Hanson et al., 1964; Wright & Nowell, 2006) have on classification accuracy. Heavier materials tend to have higher atomic scattering factors in electron diffraction, resulting in more visible reflectors for a diffraction pattern from the same space group and three-dimensional orientation. Training on only low or high atomic number materials is found to reduce future classification accuracy. Increasing the range of atomic scattering factors utilized in the training set, even if the minimum and maximum  $Z$  materials within the space group are not both included, alleviates the effects of this physical phenomenon on the neural network's classification abilities. Furthermore, this work provides a brief analysis regarding the effect of two common pattern quality metrics and train/test orientation differences on classification accuracy. The inner workings of this deep neural network-based method are studied using visual feature importance analysis. By allowing a machine learning algorithm to perform EBSD pattern classification to the space group level, a significant advancement in the utilization and accuracy of phase identification by EBSD can be achieved. When combined with chemical information on phases, for example, from energy-dispersive X-ray spectroscopy, this approach can lead to automated phase identification.

## Materials and Methods

### Materials

Eighteen different single-phase materials, comprising 6 of the 10 space groups within the  $(4/m, \bar{3}, 2/m)$  point group, were selected for demonstrating the proposed space group classification methodology. Suitable samples for the remaining four space groups could not be obtained. This point group was chosen as it contains space groups that are very similar structurally and represent a significant classification challenge for conventional EBSD. The method of fabrication and homogenization (if applicable) for each sample is listed in Supplementary Table A1. The homogenization heat treatments were performed for three weeks in an inert atmosphere at temperatures guided by each phase diagram. Samples were mounted, polished to  $0.05 \mu\text{m}$  colloidal silica, and then vibratory polished with  $0.05 \mu\text{m}$  alumina for several hours.

### Electron Backscatter Diffraction Pattern Collection

EBSD patterns (EBSPs) were collected in a Thermo-Fisher (formerly FEI) Apreo scanning electron microscope (SEM) equipped with an Oxford Symmetry EBSD detector. The Oxford Symmetry EBSD detector was utilized in high resolution ( $1244 \times 1024$ ) mode. The geometry of the set-up was held constant for each experiment. The working distance was  $18.1 \text{ mm} \pm 0.1 \text{ mm}$ . Aztec was used to set the detector insertion distance to 160.2 and the detector tilt to  $-3.1$ . The imaging parameters were 20 kV accelerating voltage, 51 nA beam current,  $0.8 \text{ ms} \pm 0.1 \text{ ms}$  dwell time, and 30 pattern averaging.

After collecting high-resolution EBSPs from each material, all patterns collected were exported as tiff images. Supplementary Figure A1 in the Appendix shows example images of the high-resolution diffraction patterns collected from the materials utilized in this study, organized by space group. All collected data for each material were individually assessed by the neural network, and the collection of images for each sample may contain partial or low-quality diffraction patterns, which will decrease the accuracy of their identification. See Supplementary Figure A2 for the inverse pole figures (IPFs) for each material. The IPFs were constructed using the MTEX software package (Bachmann et al., 2010). The data in Supplementary Figure A2 were first plotted using the scale bars set automatically by MTEX to show the fine distribution of the data, and then with the scale bar fixed from 0 to 5 times random for the purpose of demonstrating the data does not approach medium texture levels. Analysis shows the experimental datasets have very low texture, typically in the range of two to three multiples of uniform distribution (M.U.D.) also referred to as times random. Typically, 5–10 is considered medium texture and greater than 10 is considered strong texture.

See Supplementary Figure A3 for histograms of mean angular deviation (MAD) and band contrast (BC) to compare pattern quality for each material. Each plot is also annotated with the mean ( $\mu$ ) and standard deviation ( $\sigma$ ). The purpose of not filtering the test data was to assess the model as it would be applied in practice. Only the training sets were visually inspected to confirm high-quality diffraction patterns (no partial patterns) were utilized in fitting the neural network.

### Neural Network Architecture

The well-studied convolutional neural network architecture Xception (Chollet, 2017) was selected as the basis architecture for fitting a model that determines which space group a diffraction pattern originated from. Supplementary Figure A4 in the appendix details a schematic of the convolutional neural network operating on an EBSP. For a complete description of the Xception architecture, please refer to Figure 5 in (Chollet, 2017).

### Neural Network Training

Training was performed using 400 diffraction patterns per space group, evenly divided between the number of materials the model had access to during training. For example, if the model was given two materials during training, 200 diffraction patterns per material were made available. The validation set contained 100 diffraction patterns per space group, equivalent to the standard 80:20 train/validation split. The test set contains the rest of the patterns that were not used for training or validation. Model hyperparameters were selected or tuned as follows. Adam

optimization with a learning rate of 0.001 (Kingma & Ba, 2014), and a minimum delta of 0.001 as the validation loss were employed for stopping criteria. The weight decay was set to  $1e^{-5}$  following previous optimization work (Chollet, 2017). The CNNs were implemented with TensorFlow (Abadi et al., 2016) and Keras (Chollet, 2015). The code for implementing these models can be found at <https://github.com/krkaufma/Electron-Diffraction-CNN> or Zenodo (DOI: 10.5281/zenodo.3564937).

### Diffraction Pattern Classification

Each diffraction pattern collected, but not used in training ( $>140,000$  images), was evaluated in a random order by the corresponding trained CNN model without further information. The output classification of each diffraction pattern was recorded and saved in a (.csv) file and are tabulated in the Appendix. All corresponding bar plots of these data were generated with MATLAB. Precision and recall were calculated for each material and each space group using Scikit-learn (Pedregosa et al., 2011). Precision (equation 1) for each class (e.g., 225) is defined as the number of correctly predicted images out of all patterns predicted to belong to that class (e.g., 225). Recall is the number of correctly predicted patterns for each class divided by the actual number of patterns for the class (equation 2).

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \quad (1)$$

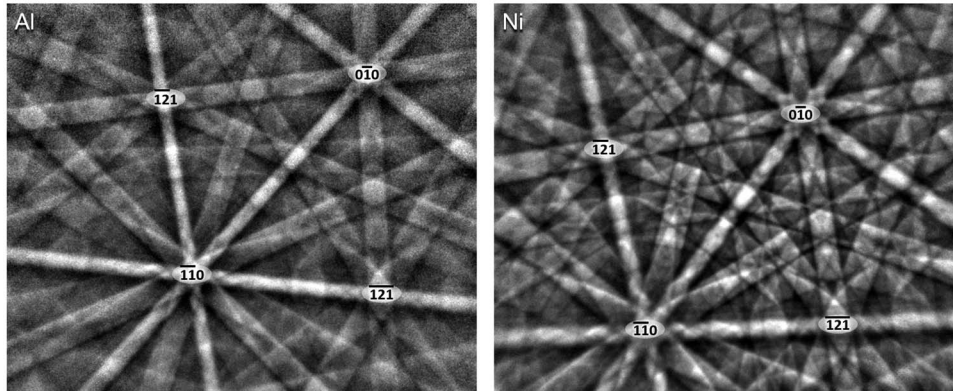
$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \quad (2)$$

### Neural Network Interpretability

Gradient-weighted class activation mapping (Grad-CAM) was employed to provide insight into the deep neural network (Selvaraju et al., 2017). This method computes the importance of local regions in the diffraction image, normalizes them from 0 to 1, and is overlaid as a localization heatmap highlighting the important regions in the image. The “guided” backpropagation modifier was used to achieve pixel-space gradient visualizations and filter out information that suppresses the neurons. This information can be safely filtered out since we are only interested in what image features the neuron detects with respect to the target space group. The gradients flowing into the final convolution layer were targeted for two reasons: (i) convolutional layers naturally retain spatial information unlike the fully-connected layers and (ii) previous works have asserted that with increasing depth of a CNN, higher-level visual constructs are captured (Bengio et al., 2013; Mahendran & Vedaldi, 2016).

## Results and Discussion

The model is first trained on one material from each of six space groups in the  $(4/m, \bar{3}, 2/m)$  point group; there are 10 space groups within the  $(4/m, \bar{3}, 2/m)$  point group, but suitable samples for 4 of these space groups could not be obtained. The first iteration uses the material with the largest formula weighted atomic number ( $\bar{Z}$ ) (i.e., atomic scattering factor) in order to establish a baseline performance. Ta was used as the training material for space group 229 instead of W since training would

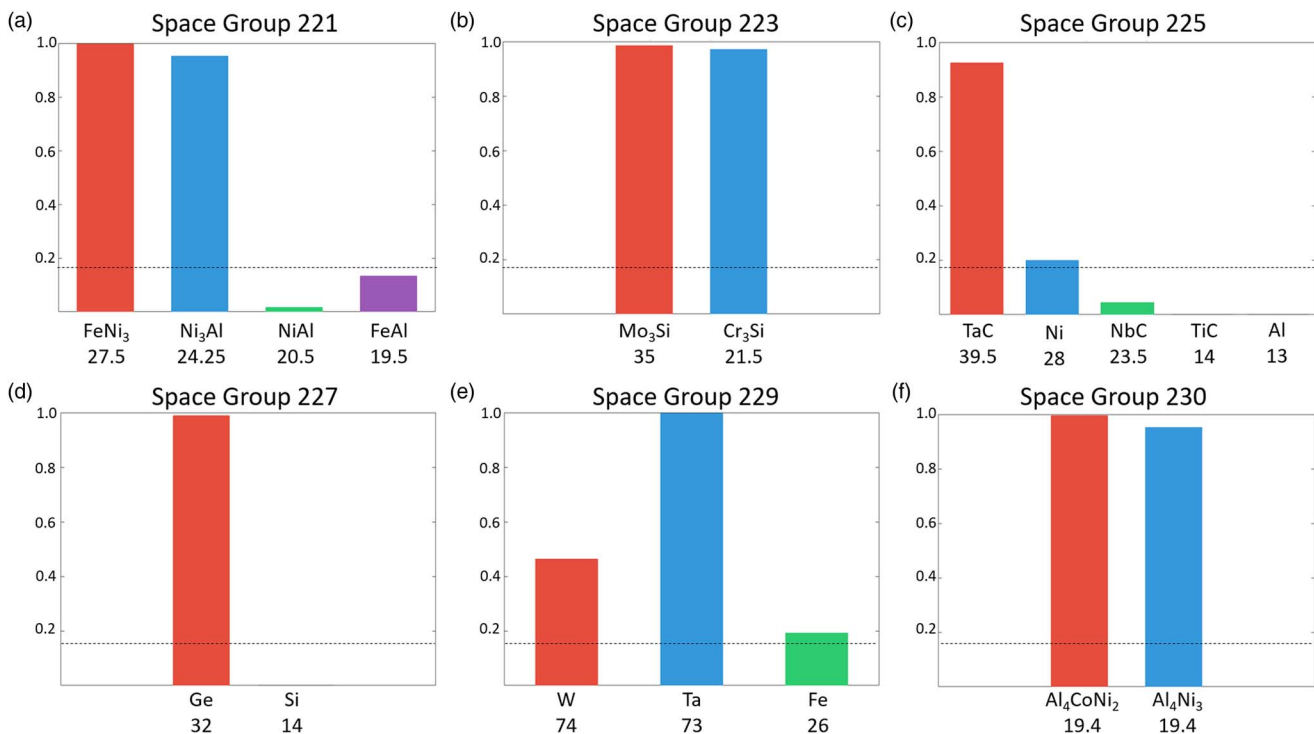


**Fig. 1.** Effect of atomic scattering factors on observable Kikuchi bands. Al ( $Z = 13$ ) and Ni ( $Z = 28$ ). Despite belonging to the same space group, the symmetry information visible for the lower atomic number Al is noticeably reduced compared with Ni.

require nearly 50% of the available W diffraction patterns. Since W and Ta only differ by one atomic number, Ta will serve as a good baseline measurement and provide information about small steps in  $\bar{Z}$ . This baseline will be compared to models trained using materials with a reduced number of visible reflectors resulting from lower atomic scattering factors. Figure 1 demonstrates this effect using a similarly oriented experimental diffraction pattern from Al and Ni with the same four zone axes labeled. The diffraction pattern from Ni has significantly more observable information (Kikuchi bands) since the atomic scattering factor is modulating their intensity. Using a limited number of materials (e.g., one) to represent the entire population of the space group when the neural network is learning the filters that maximize classification accuracy is likely to yield representations that are not as

effective when Kikuchi bands are more or less visible based on diffraction intensity.

Figure 2 shows the normalized accuracy of the first iteration of the model at classifying each of the 18 materials to the correct space group after learning from one material per class. Using only one material during training creates a significant opportunity for the neural network to “invent” representations that are not based on space group symmetry. The number of patterns from a given material that were classified to each space group is given in Appendix Supplementary Table A1. As shown in Figure 2, the model performs significantly better than random guessing (the dashed line at 16.7%) on multiple materials, including 6 of the 12 materials that were not in the training set. Of those 6, it achieves better than 45% accuracy on 4 of them: ( $\text{Ni}_3\text{Al}$ ,



**Fig. 2.** Plots of normalized classification accuracy after fitting the model with the high atomic number materials. (a) Space group 221; trained on  $\text{FeNi}_3$ . (b) Space group 223; trained on  $\text{Mo}_3\text{Si}$ . (c) Space group 225; trained on TaC. (d) Space group 227; trained on Ge. (e) Space group 229; trained on Ta. (f) Space group 230; trained on  $\text{Al}_4\text{CoNi}_2$ . The dashed line represents the chance of randomly guessing the correct space group. The formula weighted atomic number is located below each material.



Cr<sub>3</sub>Si, W, and Al<sub>4</sub>Ni<sub>3</sub>). This is a good indicator that the model is learning useful features for differentiating the space groups, instead of trivial ones that only work during the training process. Furthermore, analysis of the IPFs in Supplementary Figure A2 shows that these materials can have distinctly different orientation distributions. As an example, the Al<sub>4</sub>Ni<sub>3</sub> sample has a greater distribution of data with orientations near [001] in X, Y, and Z than the material used in training (Al<sub>4</sub>CoNi<sub>2</sub>). Yet, the model achieves 95% accuracy on Al<sub>4</sub>Ni<sub>3</sub>. From Supplementary Table A1, the overall accuracy of the model is 51.2%, well above the 16.7% chance of guessing correctly. It is observed that new materials with similar atomic scattering factors to the training material tend to have higher classification accuracy. Mo<sub>3</sub>Si and Cr<sub>3</sub>Si demonstrate that large differences in atomic scattering factor is not the penultimate factor and does not prevent accurate classification when evaluating new materials. There are also several materials with similar average  $Z$  that are rarely misclassified as one another. These include NiAl ( $\bar{Z} = 20.5$ ), FeAl ( $\bar{Z} = 19.5$ ), and Cr<sub>3</sub>Si ( $\bar{Z} = 21.5$ ) as well as Ni<sub>3</sub>Al ( $\bar{Z} = 24.25$ ) and NbC ( $\bar{Z} = 23.5$ ).

Studying the misclassification events in Supplementary Table A1 provides further valuable insight. For example, the FeAl and NiAl samples are both B2, an ordered derivative of the BCC lattice, and the highest number of misclassifications for these two materials belong to space group 229 and 230. Figure 3 shows the difficulty of distinguishing these two space groups, a common problem in the literature (Gao et al., 2016; Li et al., 2018).

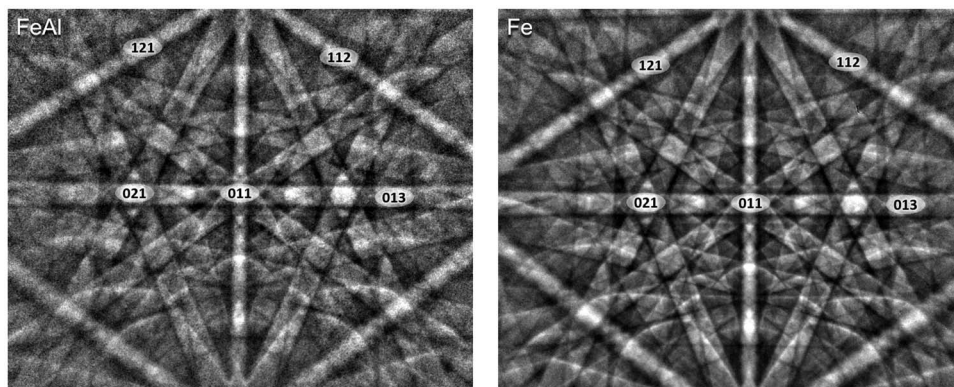
Using the previous model as a baseline, the next three iterations study the effect of swapping the material used to represent a space group with the lowest average atomic number material in the dataset. First, TaC was replaced with Al in space group 225. The results of the exchange is shown in Figure 4 and Supplementary Table A2 in the Appendix. The plots in Figure 4 look largely unchanged except for space group 225 (Fig. 4c), for which materials with the lowest atomic scattering factors are now the most accurately classified. Furthermore, Supplementary Table A2 shows that the incorrect classifications for all other space groups are now primarily space group 225.

In order to confirm the effects observed by exchanging Al for TaC in space group 225 previously, Si is exchanged with Ge in space group 227 (Fig. 5 and Supplementary Table A3). The training set for space group 225 is returned to TaC. The two space groups most affected by this change are 225 and 227, which both have FCC symmetry elements. For both space groups, as well as Fe, the use of Si as the training material causes diffraction

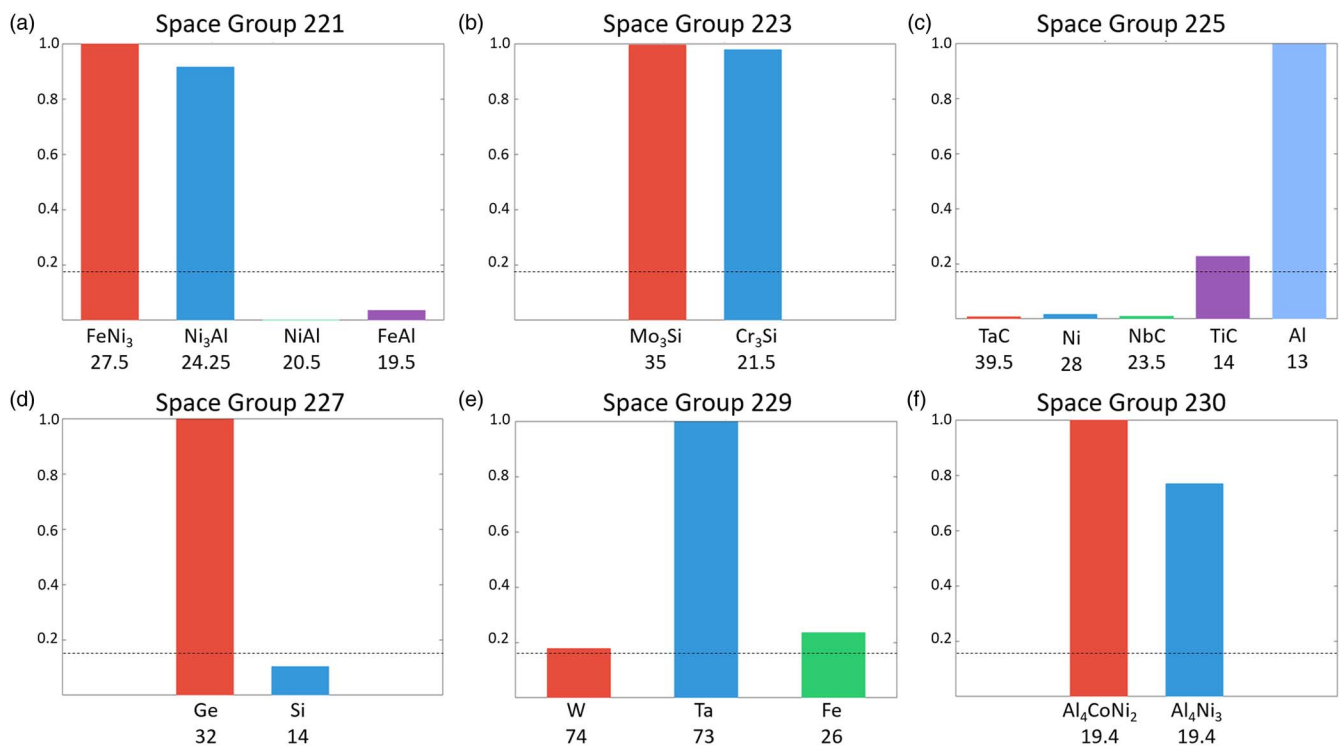
patterns to be primarily misclassified as space group 221. The substitution of Si does result in the beneficial effect of increasing the classification accuracy of materials belonging to space group 221.

The last swap studied was the lower atomic number Fe in place of Ta (Fig. 6). Similar to what was observed when Al was used in space group 225, the increased correct classifications have shifted toward the low atomic number materials within space group 229. Furthermore, the B2 materials are now primarily misclassified as belonging to the BCC symmetry space group 229 (Supplementary Table A4). In comparison to Supplementary Table A1, almost half of the NiAl diffraction patterns were classified as 229 instead of 230. These three examples confirm that providing the neural network with only the lightest or heaviest materials will result in avoidable misclassification events.

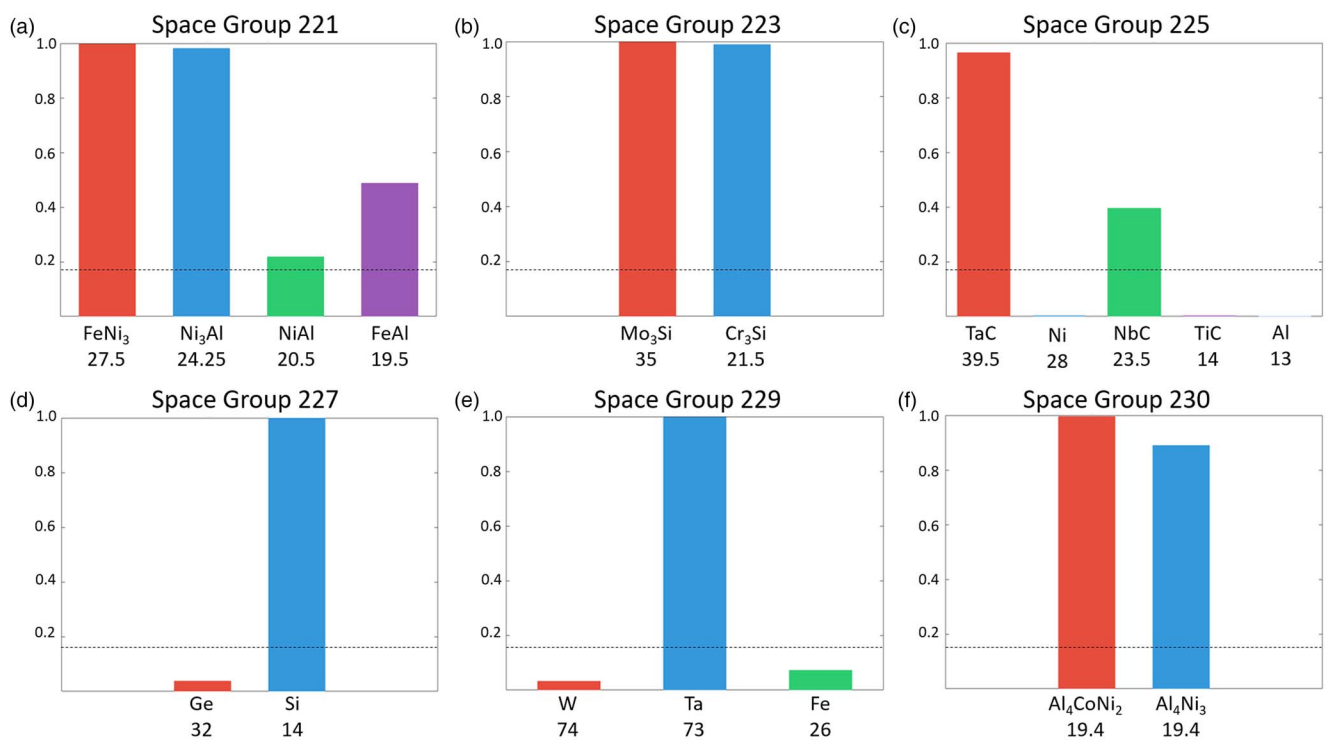
Previously, the model has only been supplied with information from one material to learn from. It should not be surprising that the limited data representing the population of all materials in each class results in misclassification events when extrapolating further away from the training data. To demonstrate the increased accuracy of the technique when the data begins to better represent the population, we add a second material to the training set for space groups 221, 225, 227, and 229. Materials were not added to space group 223 and 230 since performance on those two groups was already well above the probability of the neural network guessing correctly by chance (16.7%) for the second material. The total number of diffraction patterns available to the neural network during training remained fixed as described in the Methods section. Supplementary Figures A5, A6 display the IPFs and histograms, respectively, for the training data used in this model. As evidenced by the low M.U.D. and wide range of pattern quality, the data provided to the model during training is quite diverse. The classification results with new diffraction patterns are shown in Figure 7 and Supplementary Table A5 in the Appendix. Compared to the neural networks trained with only one material, this model has 29% higher accuracy (now 80% correct) and demonstrates significantly improved accuracy on materials not utilized in the training set. For example, diffraction patterns from Ni<sub>3</sub>Al, FeAl, Cr<sub>3</sub>Si, NbC, TiC, and Al<sub>4</sub>CoNi<sub>2</sub> are correctly classified significantly more than random guessing even though the neural network was not provided diffraction patterns from these materials from which to learn. Moreover, significantly increasing the range of atomic scattering factors, such as in Figures 7a, 7c, improves the neural network's performance on materials with scattering factors further outside the range.



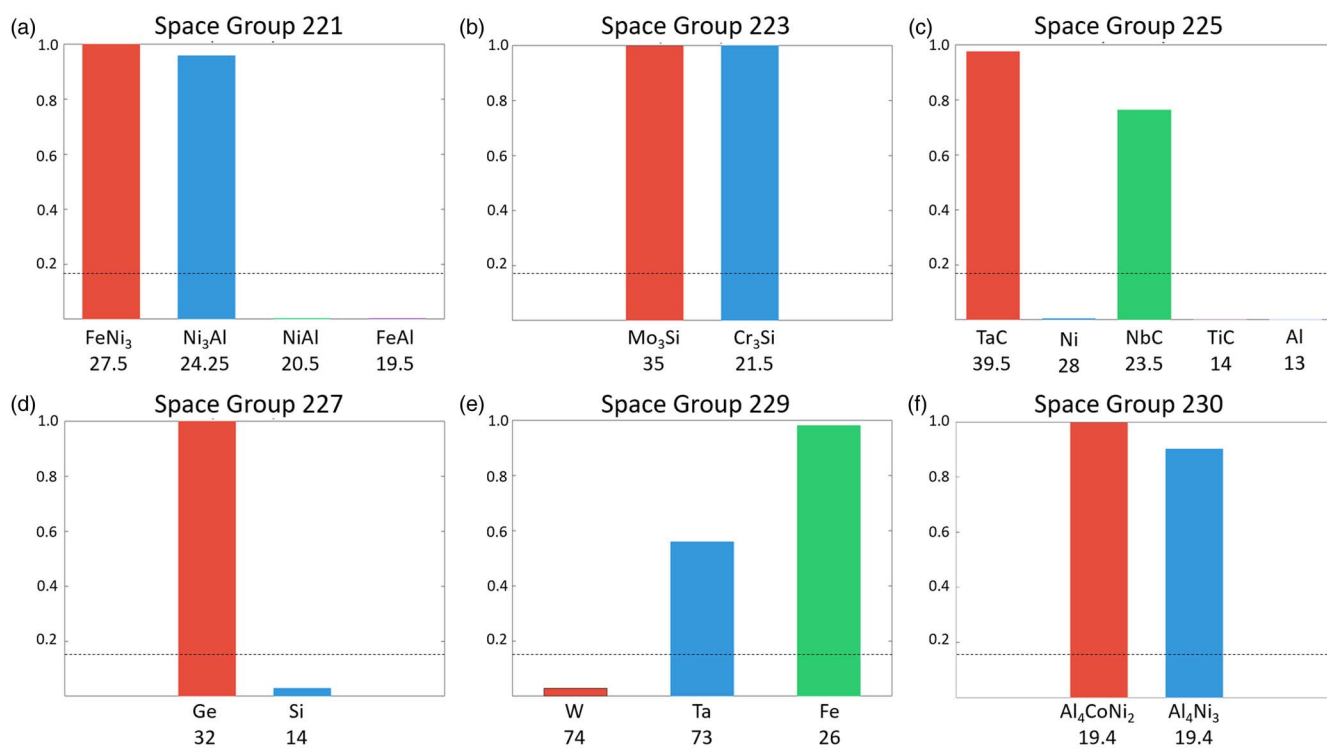
**Fig. 3.** Comparison of diffraction patterns from B2 FeAl (space group 221) and BCC Fe (space group 229). B2 FeAl and BCC Fe can produce nearly identical diffraction patterns despite belonging to two different space groups.



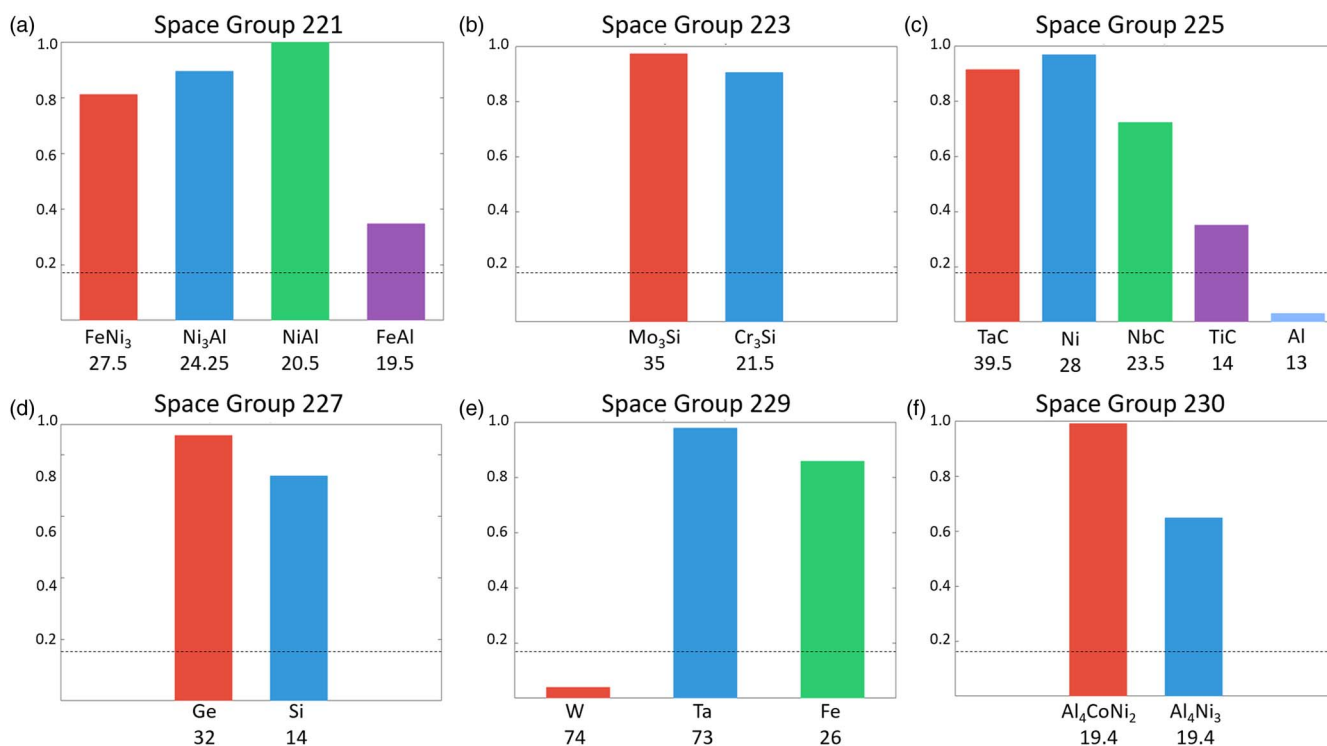
**Fig. 4.** Plots of normalized classification accuracy after fitting the model with a single low atomic number material. (a) Space group 221; trained on FeNi<sub>3</sub>. (b) Space group 223; trained on Mo<sub>3</sub>Si. (c) Space group 225; trained on Al. (d) Space group 227; trained on Ge. (e) Space group 229; trained on Ta. (f) Space group 230; trained on Al<sub>4</sub>CoNi<sub>2</sub>. The dashed line represents the chance of randomly guessing the correct space group. The formula weighted atomic number is located below each material.



**Fig. 5.** Plots of normalized classification accuracy after fitting the model with a single low atomic number material. (a) Space group 221; trained on FeNi<sub>3</sub>. (b) Space group 223; trained on Mo<sub>3</sub>Si. (c) Space group 225; trained on TaC. (d) Space group 227; trained on Si. (e) Space group 229; trained on Ta. (f) Space group 230; trained on Al<sub>4</sub>CoNi<sub>2</sub>. The dashed line represents the chance of randomly guessing the correct space group. The formula weighted atomic number is located below each material.



**Fig. 6.** Plots of normalized classification accuracy after fitting the model with a single low atomic number material. (a) Space group 221; trained on FeNi<sub>3</sub>. (b) Space group 223; trained on Mo<sub>3</sub>Si. (c) Space group 225; trained on TaC. (d) Space group 227; trained on Ge. (e) Space group 229; trained on Fe. (f) Space group 230; trained on Al<sub>4</sub>CoNi<sub>2</sub>. The dashed line represents the chance of randomly guessing the correct space group. The formula weighted atomic number is located below each material.

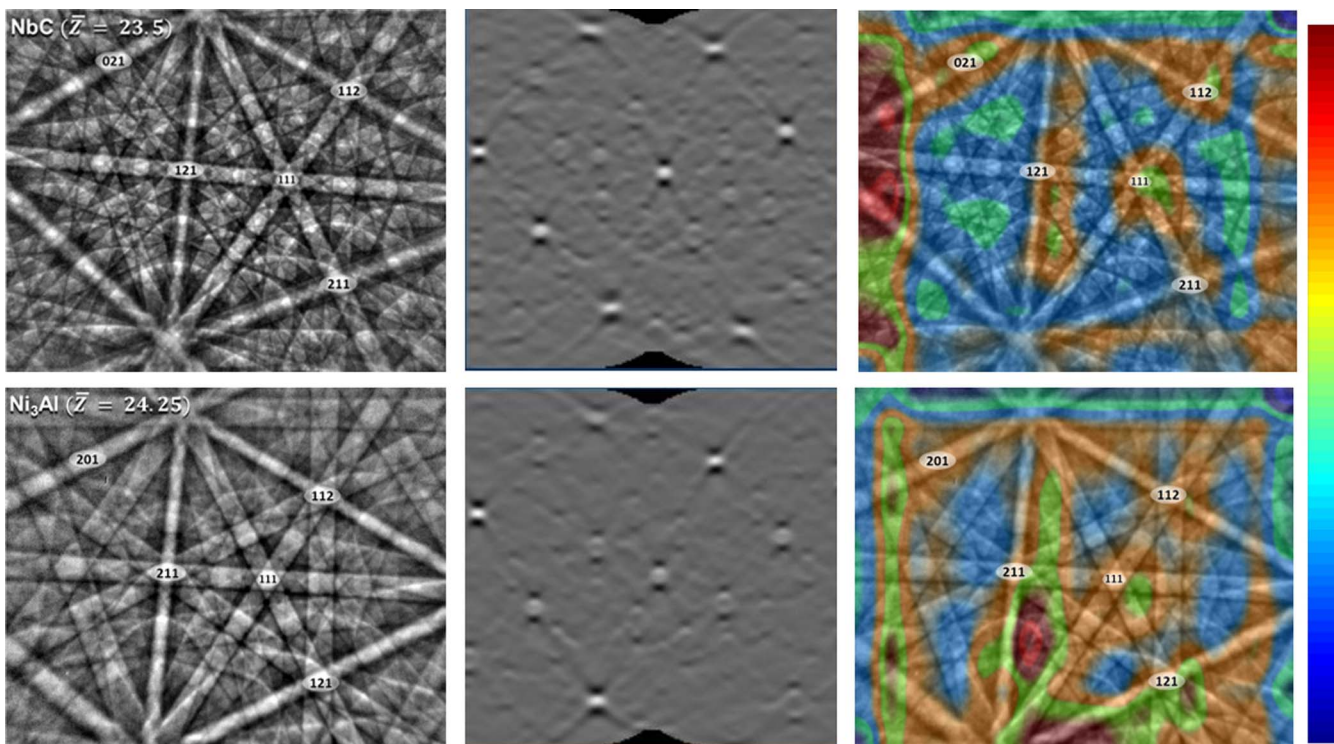


**Fig. 7.** Plots of normalized classification accuracy after fitting the model. (a) Space group 221; trained on FeNi<sub>3</sub> and NiAl. (b) Space group 223; trained on Mo<sub>3</sub>Si. (c) Space group 225; trained on TaC and Ni. (d) Space group 227; trained on Ge and Si. (e) Space group 229; trained on Ta and Fe. (f) Space group 230; trained on Al<sub>4</sub>CoNi<sub>2</sub>. The dashed line represents the chance of randomly guessing the correct space group. The formula weighted atomic number is located below each material.

Aluminum and tungsten are the only two materials where the neural network's classification accuracy is below the probability of randomly guessing the correct answer. An analysis of the orientations and pattern quality for patterns that were correctly identified and misclassified was performed on several materials outside the training set. NbC, Al, and W were selected in order to make a number of comparisons including materials with varying classification accuracy or within the same space group. Supplementary Figure A7 shows that for each material, the IPFs for the correctly classified patterns resembles the IPFs for patterns that were misclassified. In other words, specific orientations do not seem to be more likely to be classified correctly. Supplementary Figure A8 shows the corresponding MAD and BC histograms. While the plots and associated statistics for NbC show patterns with lower band contrast or higher MAD have a slight tendency to be misclassified, there does not appear to be a strong relationship between these pattern quality metrics and classification accuracy.

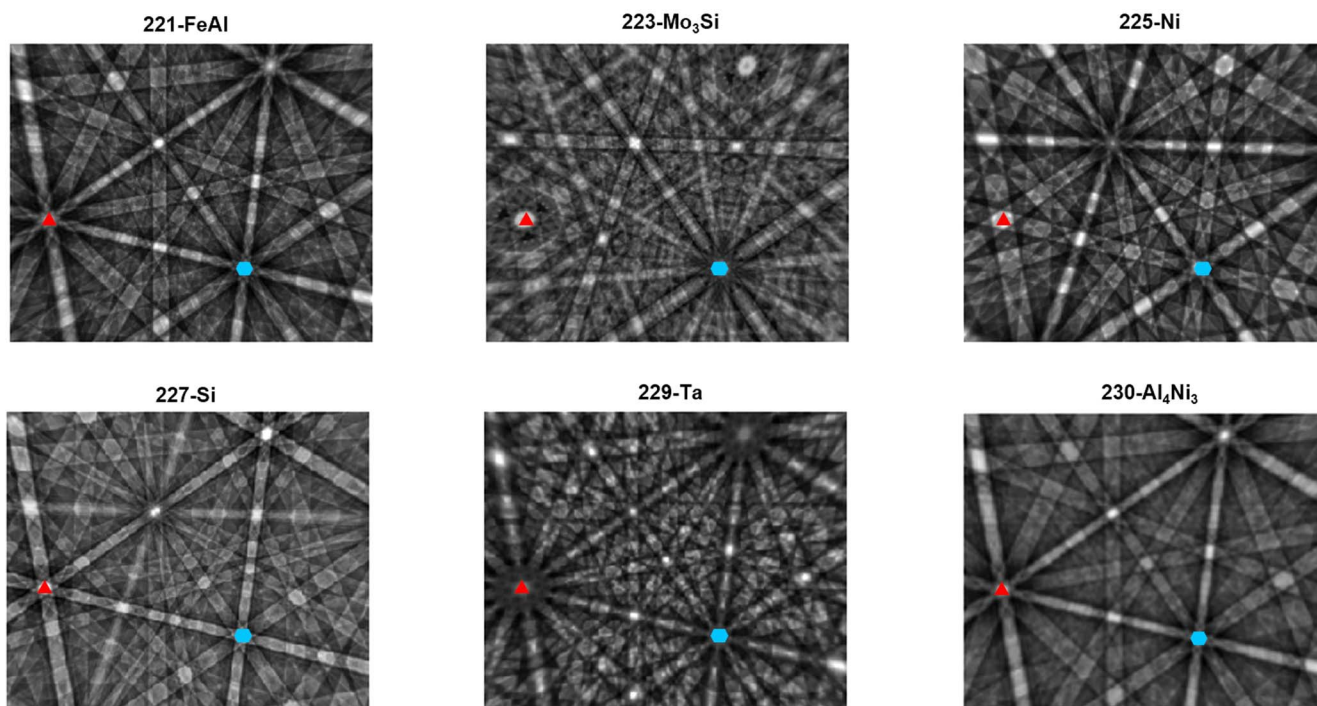
The excellent overall performance of the neural network in correctly identifying the space group necessitates an understandable interpretation of what information is important to the model, since orientation and pattern quality are not seemingly biasing the results of this study. Figure 8 is a set of “visual explanations” for the decisions made by the Hough-based method and the trained CNN from Figure 7. The selected diffraction patterns (Fig. 8 left-side) from NbC and Ni<sub>3</sub>Al are of nearly the same orientation and average atomic number but are from space group 225 and 221, respectively. Importantly, the model had not trained on any EBSPs from these two materials at this point. The Hough-based method produces the “butterfly peak” (Christian et al., 1994) representation (Fig. 8 middle) for matching the interplanar angles to

user selected libraries. The resulting Hough-transforms are nearly identical for both NbC and Ni<sub>3</sub>Al, save for minor differences owing to the small orientation shift. Furthermore, the Hough-based method creates a sparse representation that does not capture the finer detail found within the EBSPs. For example, the [111] and [112] zone axes are each encircled by a discernable HOLZ ring (Michael & Eades, 2000) as well as excess and deficiency features (Winkelmann, 2008). In comparison to the Hough method, the convolutional neural network can learn these features, or lack of them, and associate them with the correct space group. Using gradient-weighted class activation mapping (Grad-CAM) (Selvaraju et al., 2017), we produce a localization heatmap highlighting the important regions for predicting the target class (Fig. 8 right-side). When visually inspecting the importance of local regions to the neural network, it is immediately observed that the neural network finds each of the labeled zone axes to be of high importance (orange in color) in both EBSPs. In fact, the heatmaps look remarkably similar and are concentrated about the same features a crystallographer would use. These features clearly include the zone axes and HOLZ rings. To investigate this further, a diffraction pattern is simulated for each material using EMSOFT (Callahan & De Graef, 2013). In Figure 9, it is first observed that many of the features of the patterns are similar, such as Kikuchi bands and diffraction maxima. These similarities explain why the Hough transform-based method cannot distinguish the two. However, it is also immediately noticeable that equivalent zone axes and surrounding regions have very different appearances. As examples, two sets of equivalent zone axes have been marked with either a red triangle or blue hexagon. The red triangle is the equivalent to the [111]



**Fig. 8.** Comparison of feature detection with Hough-based EBSD and the trained CNN. (Top row from left to right) Experimental EBSP from NbC (space group 225; FCC structure), Hough-based feature detection, and gradient-weighted class activated map. (Bottom row from left to right) Experimental EBSP from Ni<sub>3</sub>Al (space group 221; L1<sub>2</sub> structure), Hough-based feature detection, and gradient-weighted class activated map. The importance scale for the heatmaps goes from dark blue (low) to dark red (high).

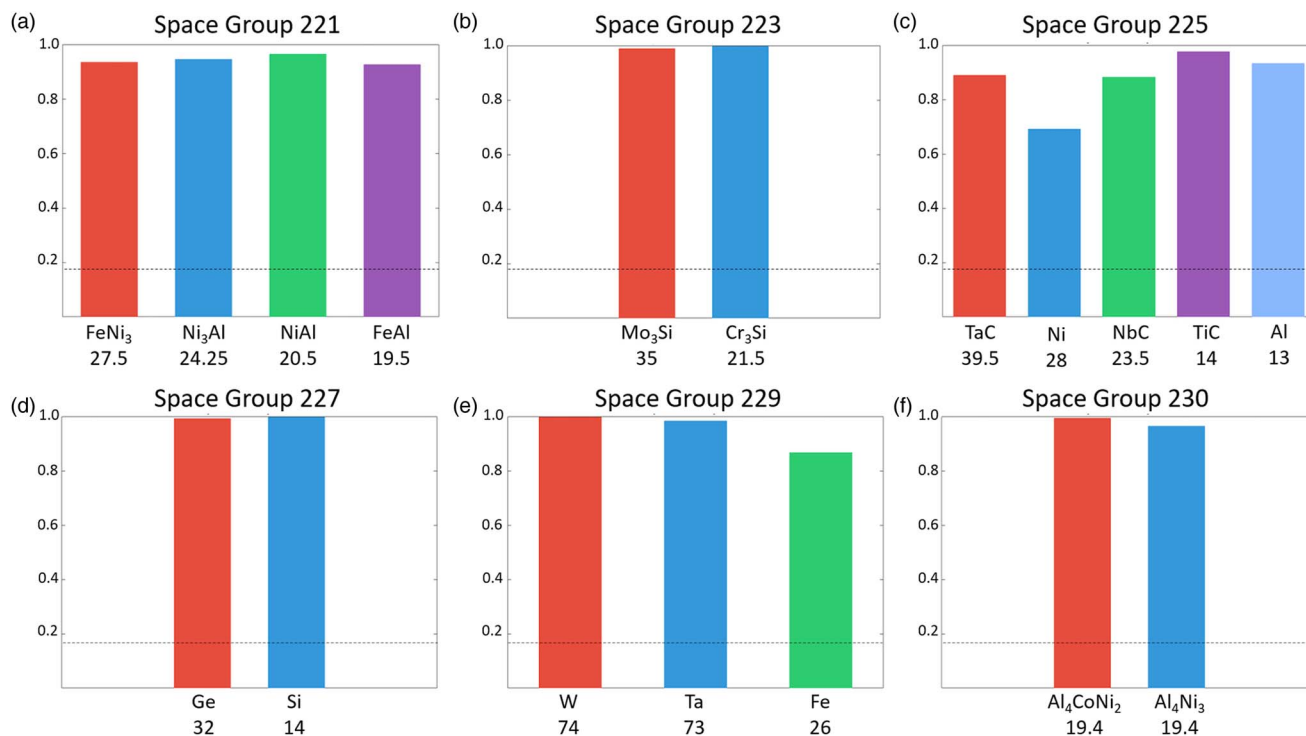




**Fig. 9.** Dynamically simulated EBSPs. One diffraction pattern per space group was simulated to study the expected differences and assess the feature importance observed in experimental patterns. The observed reflectors are similar for each space group; however, attributes nearby the zone axes can vary significantly. Two sets of equivalent zone axes have been indicated with either a red triangle or blue hexagon.

zone axis studied in Figure 8. While the fidelity of the simulated patterns with experimental patterns is not perfect, the existence of the discussed features can be confirmed by comparing them with the experimental patterns previously discussed. For example,

compare the red and blue labeled zone axes in Figure 9 for Ni and Mo<sub>3</sub>Si with the experimental NbC (Fig. 8) and Cr<sub>3</sub>Si (Supplementary Fig. A1) diffraction patterns. The structure and features visible are clearly well correlated.



**Fig. 10.** Plots of normalized classification accuracy after fitting the model with data from each material. (a) Space group 221. (b) Space group 223. (c) Space group 225. (d) Space group 227. (e) Space group 229. (f) Space group 230. The dashed line represents the chance of randomly guessing the correct space group. The formula weighted atomic number is located below each material.

Supplementary Figure A9 helps to elucidate the failure mechanism of this model by studying the activations of the current model, which incorrectly identified the pattern, compared to the activations in the first model, which had correctly identified the same EBSP. In each case, the heatmaps for space group 229 are overlaid onto the diffraction pattern. When correctly identified as belonging to space group 229 (Supplementary Fig. A9, center), significantly more zone axes are given higher importance scores than when misclassified to space group 225 (Supplementary Fig. A9, right). The same is observed with features within the HOLZ rings, such as near the [001] and [101] axis. It is important to note, the similarity between the activations at zone axes, such as [001], [012], and [102], should not be surprising since the activations for class 229 are being studied. Failures such as this can potentially be alleviated as the number of samples, experimental or simulated, used in fitting these models continues to grow. Previous works using experimental (Kaufmann et al., 2020) and simulated (Foden et al., 2019) patterns have further demonstrated the attentiveness of the last layers of the model on the zone axes and surrounding features. The study using simulated patterns also elucidated each convolutional layer's attention to specific aspects of EBSPs including edges and major Kikuchi bands. These studies into the "visual" perception of the model suggest that the network has learned relevant and intuitive features for identifying space groups.

By providing the neural network with diffraction patterns from many materials, the neural network can improve its resiliency to small changes within a space group, simultaneously developing a better understanding of what elements in the image are most useful and learning filters that better capture the information. Figure 10 demonstrates this by supplying the same number of diffraction patterns to learn from as used previously, but evenly divided between all available materials in each space group. Supplementary Table A6 shows the number of images classified to each space group from individual materials. The classification accuracy has increased to 93%, compared to 40–65% for the models that were only provided with one training material per space group (Figs. 2, 4–7) and 80% when using two materials for some of the classes (Fig. 8).

## Conclusion

In this paper, a high-throughput CNN-based approach to classifying electron backscatter diffraction patterns at the space group level is developed and demonstrated. In each study, the CNN is shown to be able to classify at least several materials outside the training set with much better probability than random guessing. Several investigations are conducted to explore the potential for biases owing to crystallographic orientation, pattern quality, or physical phenomena. The number of visible reflectors, directly correlated with atomic scattering factors, is found to have an impact classification accuracy when only low or high atomic number materials are used to fit the model. Increasing the range of atomic scattering factors used in training each class is found to reduce the number of misclassification events caused by large differences in the number of visible reflectors. The dataset for each material and space group were of very low texture and good distributions of pattern quality metrics. Continued inclusion of data, particularly from more materials, will likely increase the robustness of the model when presented with new data. Investigation of the convolutional neural network's inner workings, through visualization of feature importance, strongly indicates the network is using the same features a crystallographer

would use to manually identify the structure, particularly the information within the HOLZ rings. This analysis, combined with IPFs, suggests there is minimal orientation bias present. We believe this method can be expanded to the remaining space groups and implemented as part of a multi-tiered model for determining the complete crystal structure. There are no algorithmic challenges to extending this framework to all 230 space groups, it is only currently limited by the lack of data, which simulated EBSD patterns may help to resolve. This technique should benefit from continued advancements in detectors, such as direct electron detectors, and the framework is expected to be immediately applicable to similar techniques such as electron channeling patterns and CBED. A wide range of other research areas including pharmacology, structural biology, and geology are expected to benefit by using similar automated algorithms to reduce the amount of time required for structural identification.

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/S1431927620001506>

**Acknowledgments.** The authors would like to thank Wenyou Jiang, William Mellor, and Xiao Liu for their assistance. K. Kaufmann was supported by the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program. K. Kaufmann would also like to acknowledge the support of the ARCS Foundation, San Diego Chapter. KV would like to acknowledge the financial generosity of the Oerlikon Group in support of his research group.

## References

- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, Kudlur M, Levenberg J, Monga R, Moore S, Murray DG, Steiner B, Tucker P, Vasudevan V, Warden P, Wicke M, Yu Y, Zheng X & Brain G (2016). TensorFlow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*. Savannah, Georgia, USA: USENIX, pp. 265–283.
- Alegre E, Barreiro J, Cáceres H, Hernández LK, Fernández RA & Castejón M (2006). Design of a computer vision system to estimate tool wearing. *Mater Sci Forum* **526**, 61–66.
- Baba-Kishi KZ & Dingley DJ (1989). Backscatter kikuchi diffraction in the SEM for identification of crystallographic point groups. *Scanning* **11**, 305–312.
- Bachmann F, Hielscher R & Schaeben H (2010). Texture analysis with MTEX- Free and open source software toolbox. *Solid State Phenomena* **160**, 63–68.
- Bengio Y, Courville A & Vincent P (2013). Representation learning: A review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* **35**, 1798–1828.
- Bernard R, Day J & Chin E (2019). Strong olivine lattice preferred orientation in brachinite-like achondrites. In *Lunar and Planetary Science Conference*.
- Britton TB, Maurice C, Fortunier R, Driver JH, Day AP, Meaden G, Dingley DJ, Mingard K & Wilkinson AJ (2010). Factors affecting the accuracy of high resolution electron backscatter diffraction when using simulated patterns. *Ultramicroscopy* **110**, 1443–1453.
- Calcagnotto M, Ponge D, Demir E & Raabe D (2010). Orientation gradients and geometrically necessary dislocations in ultrafine grained dual-phase steels studied by 2D and 3D EBSD. *Mater Sci Eng A* **527**, 2738–2746.
- Callahan PG & De Graef M (2013). Dynamical electron backscatter diffraction patterns. Part I: Pattern simulations. *Microsc Microanal* **19**, 1255–1265.
- Chen CL & Thomson RC (2010). The combined use of EBSD and EDX analyses for the identification of complex intermetallic phases in multicomponent Al-Si piston alloys. *J Alloys Compd* **490**, 293–300.
- Chen D, Kuo JC & Wu WT (2011). Effect of microscopic parameters on EBSD spatial resolution. *Ultramicroscopy*.
- Chen YH, Park SU, Wei D, Newstadt G, Jackson MA, Simmons JP, De Graef M & Hero AO (2015). A dictionary approach to electron backscatter diffraction indexing. *Microsc Microanal*.
- Chollet F (2015). Keras. <https://github.com/keras-team/keras>.

- Chollet F** (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1251–1258.
- Cowley JM** (1990). *Diffraction Physics*, 2nd ed. Amsterdam: North-Holland.
- DeCost BL & Holm EA** (2015). A computer vision approach for automated analysis and classification of microstructural image data. *Comput Mater Sci* **110**, 126–133.
- Dingley DJ & Wright SI** (2009). Phase identification through symmetry determination in EBSD patterns. In *Electron Backscatter Diffraction in Materials Science*, Schwartz A, Kumar M, Adams B & Field D (Eds.), pp. 97–107. Boston, MA: Springer.
- Foden A, Collins D, Wilkinson A & Britton TB** (2019). Indexing electron backscatter diffraction patterns with a refined template matching approach. *Ultramicroscopy* **207**, 112845.
- Foden A, Previero A & Britton TB** (2019). Advances in electron backscatter diffraction. arXiv Prepr arXiv. 1908.04860.
- Gao MC, Yeh J-W, Liaw PK & Zhang Y** (eds) (2016). *High-Entropy Alloys*. Cham: Springer International Publishing.
- Garnier E** (2009). *Powder Diffraction. Theory and Practice*, Dinnebier RE & Billinge SJL (Eds.), pp. xxi + 582. Cambridge: RSC Publishing, 2008. Price (hardcover): GBP 59.00. ISBN (online): 978-1-84755-823-7; ISBN (print): 978-0-85404-231-9. *Acta Crystallogr. Sect. A Found. Crystallogr.* **65**, 51–51.
- Goehner RP & Michael JR** (1996). Phase identification in a scanning electron microscope using backscattered electron kikuchi patterns. *J Res Natl Inst Stand Technol* **101**, 301–308.
- Goulden J, Trimby P & Bewick A** (2018). The benefits and applications of a CMOS-based EBSD detector. *Microsc Microanal* **24**, 1128–1129.
- Hanson HP, Herman F, Lea JE & Skillman S** (1964). HFS atomic scattering factors. *Acta Crystallogr* **17**, 1040–1044.
- Hielscher R, Bartel F & Britton TB** (2019). Gazing at crystal balls-electron backscatter diffraction indexing and cross correlation on a sphere. *Ultramicroscopy* **207**, 112836.
- Hufford LJ, Chin EJ, Perry SE, Miranda EA & Hanchar JM** (2019). Xenolith constraints on rheology of heterogeneous deep crust beneath the Eastern Mojave Desert, California. In *GSA Annual Meeting in Phoenix, Arizona, USA-2019*.
- Karthikeyan T, Dash MK, Saroja S & Vijayalakshmi M** (2013). Evaluation of misindexing of EBSD patterns in a ferritic steel. *J Microsc* **249**, 26–35.
- Kaufmann K, Zhu C, Rosengarten AS, Maryanovsky D, Harrington TJ, Marin E & Vecchio KS** (2020). Crystal symmetry determination in electron diffraction using machine learning. *Science* **367**, 564–568.
- Kingma DP & Ba J** (2014). Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, pp. 1–15.
- Lassen NCK** (1994). Automated Determination of Crystal Orientations from Electron Backscattering Patterns. PhD Thesis.
- LeCun Y, Bengio Y & Hinton G** (2015). Deep learning. *Nature* **521**, 436–444.
- Li L & Han M** (2015). Determining the Bravais lattice using a single electron backscatter diffraction pattern. *J Appl Crystallogr.*
- Li M, Gazquez J, Borisevich A, Mishra R & Flores KM** (2018). Evaluation of microstructure and mechanical property variations in Al<sub>x</sub>CoCrFeNi high entropy alloys produced by a high-throughput laser deposition method. *Intermetallics* **95**, 110–118.
- Lombaert H, Grady L, Pennec X, Ayache N & Cheriet F** (2014). Spectral log-demons: Diffeomorphic image registration with very large deformations. *Int J Comput Vis* **107**, 254–271.
- Mahendran A & Vedaldi A** (2016). Visualizing deep convolutional neural networks using natural pre-images. *Int J Comput Vis* **120**, 233–255.
- McLaren S & Reddy SM** (2008). Automated mapping of K-feldspar by electron backscatter diffraction and application to 40Ar/39Ar dating. *J Struct Geol* **30**, 1229–1241.
- Michael JR & Eades JA** (2000). Use of reciprocal lattice layer spacing in electron backscatter diffraction pattern analysis. *Ultramicroscopy* **81**, 67–81.
- Nolze G, Hielscher R & Winkelmann A** (2017). Electron backscatter diffraction beyond the mainstream. *Cryst Res Technol* **52**, 1600252.
- Nowell MM & Wright SI** (2004). Phase differentiation via combined EBSD and XEDS. *Journal of Microscopy* **213**, 296–305.
- Ollivier B, Retoux R, Lacorre P, Massiot D & Férey G** (1997). Crystal structure of  $\kappa$ -alumina: an X-ray powder diffraction, TEM and NMR study. *J Mater Chem* **7**, 1049–1056.
- Oviedo F, Ren Z, Sun S, Settens C, Liu Z, Hartono NTP, Ramasamy S, DeCost BL, Tian SIP, Romano G, Gilad Kusne A & Buonassisi T** (2019). Fast and interpretable classification of small X-ray diffraction datasets using data augmentation and deep neural networks. *npj Comput Mater* **5**, 60.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M & Duchesnay É** (2011). Scikit-learn: Machine learning in python. *J Mach Learn Res* **12**, 2825–2830.
- Post JE & Veblen DR** (1990). Crystal structure determinations of synthetic sodium, magnesium, and potassium birnessite using TEM and the Rietveld method.
- Rumelhart DE, Hinton GE & Williams RJ** (1986). Learning representations by back-propagating errors. *Nature* **323**, 533–536.
- Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D & Batra D** (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision* October, pp. 618–626.
- Singh S & De Graef M** (2017). Dictionary indexing of electron channeling patterns. *Microsc Microanal* **23**, 1–10.
- Singh S, Guo Y, Winiarski B, Burnett TL, Withers PJ & De Graef M** (2018). High resolution low kV EBSD of heavily deformed and nanocrystalline Aluminium by dictionary-based indexing. *Sci Rep.*
- Tong VS, Knowles AJ, Dye D & Britton TB** (2019). Rapid electron backscatter diffraction mapping: Painting by numbers. *Mater Charact* **147**, 271–279.
- Vecchio KS & Williams DB** (1987). Convergent beam electron diffraction study of Al<sub>3</sub>Zr in Al-Zr AND Al-Li-Zr alloys. *Acta Metall* **35**, 2959–2970.
- Vecchio KS & Williams DB** (1988). Convergent beam electron diffraction analysis of the T1 (Al<sub>2</sub>CuLi) phase in Al-Li-Cu alloys. *Metall Trans A* **19**, 2885–2891.
- Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, Kian Ang K, Kuban DA, Bonnen M, Chang JY & Cheung R** (2005). Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. *Phys Med Biol* **50**, 2887–2905.
- Wang H, Harrington T, Zhu C & Vecchio KS** (2019). Design, fabrication and characterization of FeAl-based metallic-intermetallic laminate (MIL) composites. *Acta Mater* **175**, 445–456.
- Wang T, Komarasamy M, Shukla S & Mishra RS** (2018). Simultaneous enhancement of strength and ductility in an AlCoCrFeNi<sub>2.1</sub> eutectic high-entropy alloy via friction stir processing. *J Alloys Compd* **766**, 312–317.
- Williams DB, Pelton AR & Gronsky R** (eds.) (1991). *Images of Materials*. New York: Oxford University Press.
- Winkelmann A** (2008). Dynamical effects of anisotropic inelastic scattering in electron backscatter diffraction. *Ultramicroscopy* **108**, 1546–1550.
- Winkelmann A & Nolze G** (2015). Chirality determination of quartz crystals using electron backscatter diffraction. *Ultramicroscopy* **149**, 58–63.
- Wright SI & Nowell MM** (2006). EBSD image quality mapping. *Microsc Microanal* **12**, 72–84.
- Zhu C, Kaufmann K & Vecchio KS** (2020). Novel remapping approach for HR-EBSD based on demons registration. *Ultramicroscopy* **208**.
- Zhu C, Wang H, Kaufmann K & Vecchio K** (2020). A computer vision approach to study surface deformation of materials. *Meas. Sci. Technol.* **31**, 055602.
- Ziletti A, Kumar D, Scheffler M & Ghiringhelli LM** (2018). Insightful classification of crystal structures using deep learning. *Nat Commun* **9**, 2775.