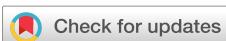


RESEARCH ARTICLE | JUNE 04 2024

Simulation-trained machine learning models for Lorentz transmission electron microscopy

Arthur R. C. McCray  ; Alec Bender  ; Amanda Petford-Long  ; Charudatta Phatak  



APL Mach. Learn. 2, 026120 (2024)
<https://doi.org/10.1063/5.0197138>



Articles You May Be Interested In

Thermal effects on current-related skyrmion formation in a nanobelt

Appl. Phys. Lett. (May 2018)

Stabilization of coupled Dzyaloshinskii domain walls in fully compensated synthetic anti-ferromagnets

AIP Advances (January 2020)

Room-temperature skyrmions in the van der Waals ferromagnet Fe_3GaTe_2

Appl. Phys. Lett. (April 2024)

Simulation-trained machine learning models for Lorentz transmission electron microscopy

Cite as: APL Mach. Learn. 2, 026120 (2024); doi: 10.1063/5.0197138

Submitted: 11 January 2024 • Accepted: 5 May 2024 •

Published Online: 4 June 2024



View Online



Export Citation



CrossMark

Arthur R. C. McCray,^{1,2} Alec Bender,¹ Amanda Petford-Long,^{1,2,3} and Charudatta Phatak^{1,3,a)}

AFFILIATIONS

¹ Material Science Division, Argonne National Laboratory, Lemont, Illinois 60439, USA

² Applied Physics Program, Northwestern University, Evanston, Illinois 60208, USA

³ Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, USA

^{a)} Author to whom correspondence should be addressed: cd@anl.gov

ABSTRACT

Understanding the collective behavior of complex spin textures, such as lattices of magnetic skyrmions, is of fundamental importance for exploring and controlling the emergent ordering of these spin textures and inducing phase transitions. It is also critical to understand the skyrmion–skyrmion interactions for applications such as magnetic skyrmion-enabled reservoir or neuromorphic computing. Magnetic skyrmion lattices can be studied using *in situ* Lorentz transmission electron microscopy (LTEM), but quantitative and statistically robust analysis of the skyrmion lattices from LTEM images can be difficult. In this work, we show that a convolutional neural network, trained on simulated data, can be applied to perform segmentation of spin textures and to extract quantitative data, such as spin texture size and location, from experimental LTEM images, which cannot be obtained manually. This includes quantitative information about skyrmion size, position, and shape, which can, in turn, be used to calculate skyrmion–skyrmion interactions and lattice ordering. We apply this approach to segmenting images of Néel skyrmion lattices so that we can accurately identify skyrmion size and deformation in both dense and sparse lattices. The model is trained using a large set of micromagnetic simulations as well as simulated LTEM images. This entirely open-source training pipeline can be applied to a wide variety of magnetic features and materials, enabling large-scale statistical studies of spin textures using LTEM.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>). <https://doi.org/10.1063/5.0197138>

I. INTRODUCTION

Nanoscale magnetic spin textures are of great interest both for their fundamental physical properties and for potential applications in novel computing paradigms.^{1–3} Of particular significance are topologically non-trivial spin textures, such as skyrmions, which are stable at small sizes and can be efficiently moved with electric currents.^{4–6} Using skyrmions as information carriers has been demonstrated for both neuromorphic and reservoir computing,^{7,8} and this motivates a thorough understanding of the properties of both individual skyrmions and their collective behavior in a skyrmion lattice.^{9,10} Of particular importance is the dynamical response of skyrmions to varying external conditions, which includes applied currents, applied magnetic fields, or changing temperature, and this necessitates *in situ* studies. Quantitative descriptors of the skyrmion size and behavior can then be extracted to gain insights into fundamental skyrmion properties, including

inter-skyrmion interaction potentials and the large-scale orientational or translational ordering behavior.^{11–15}

Lorentz transmission electron microscopy (LTEM) is an increasingly used technique for studying nanoscale magnetic spin textures such as skyrmions. LTEM enables real-space imaging with high spatial resolution of both the magnetic domain structure and sample microstructure.¹⁶ Furthermore, LTEM can be used for a wide range of *in situ* experiments; specialized sample holders enable electrical biasing and controlling the temperature of the sample, and the objective lens can be used to apply an out-of-plane magnetic field while imaging.^{17,18} There has additionally been much progress to increase the time resolution of LTEM imaging, both through the development of cameras with higher frame rates and through stroboscopic techniques.^{19,20}

When imaging with LTEM, the sample is held in a nominally field-free region of the microscope. Imaging is most frequently performed in the Fresnel mode, in which defocused images are

recorded, and the domain walls appear as regions of bright or dark contrast relative to the domains.²¹ The resulting magnetic contrast can be difficult to interpret, and this is frequently the case when imaging complex spin textures, such as skyrmions.^{22,23} Furthermore, a large defocus of several millimeters or more is often required to image magnetic skyrmions due to the weaker magnetic induction component perpendicular to the electron beam direction, and this leads to blurring and overlapping contrast from neighboring features. In the past work, we have shown that by combining experimental LTEM with micromagnetic simulations and LTEM image simulations, it is possible to better interpret LTEM images of complex spin textures,²⁴ but this does not allow for quantitative analysis of spin texture location or ordering in experimental images.

In this work, we present a machine learning (ML) model that is trained on simulated data and applied to extract quantitative information about skyrmion locations and sizes from experimental LTEM images. A pipeline of micromagnetic simulations and LTEM image simulations allows us to generate a set of labeled data on which we can train our model. We demonstrate this by applying the technique to Néel skyrmions, a spin texture for which it is difficult to quantitatively identify the size and position using LTEM. We show that a convolutional neural network (CNN) can be trained to identify skyrmion position, size, and shape from a single defocused LTEM image, information that cannot be extracted either by hand or through other analysis methods. We call a simulation-trained model for solving this task, a SkyrmNet, and we demonstrate its efficacy by showing how the skyrmion lattice evolves under external stimuli in two different materials systems: Fe₃GeTe₂ (FGT), a van der Waals ferromagnet, and in a metallic Pt/Co/Ir multilayer film. While the work in this paper is restricted to looking at the size and distribution

of Néel skyrmions, the open-source training pipeline can be generally applied to perform large-scale statistical studies of other spin textures that are suitable for imaging with LTEM.

II. METHODS

Figure 1 shows the workflow we have developed for generating training data, training a model, and applying the model to experimental data. Each step of the workflow is built on open-source code repositories, and it can be adapted and applied to nearly all materials and magnetic features that are suitable for imaging with LTEM. The workflow begins on the left side of the diagram, with some knowledge about the material parameters for the system of interest. These are used to simulate a wide range of possible sample magnetic configurations using the MuMax micromagnetics package.²⁵

In this work, we are interested in developing a SkyrmNet to perform instance segmentation of Néel skyrmions, i.e., identifying each skyrmion distinct from the background and from its neighbors. We simulated a wide range of Néel domain structures to train our model, including skyrmion lattices with skyrmions of different sizes and densities, Néel stripe domains, and mixed phases of skyrmions and stripe domains. These magnetic configurations were used, along with the relevant material parameters and imaging conditions, to generate simulated LTEM images. Ground truth labels of skyrmion sizes and positions were obtained directly from the magnetization maps, and the images and labels were then used to train our CNN. Once trained, a SkyrmNet can be directly applied to experimental data.

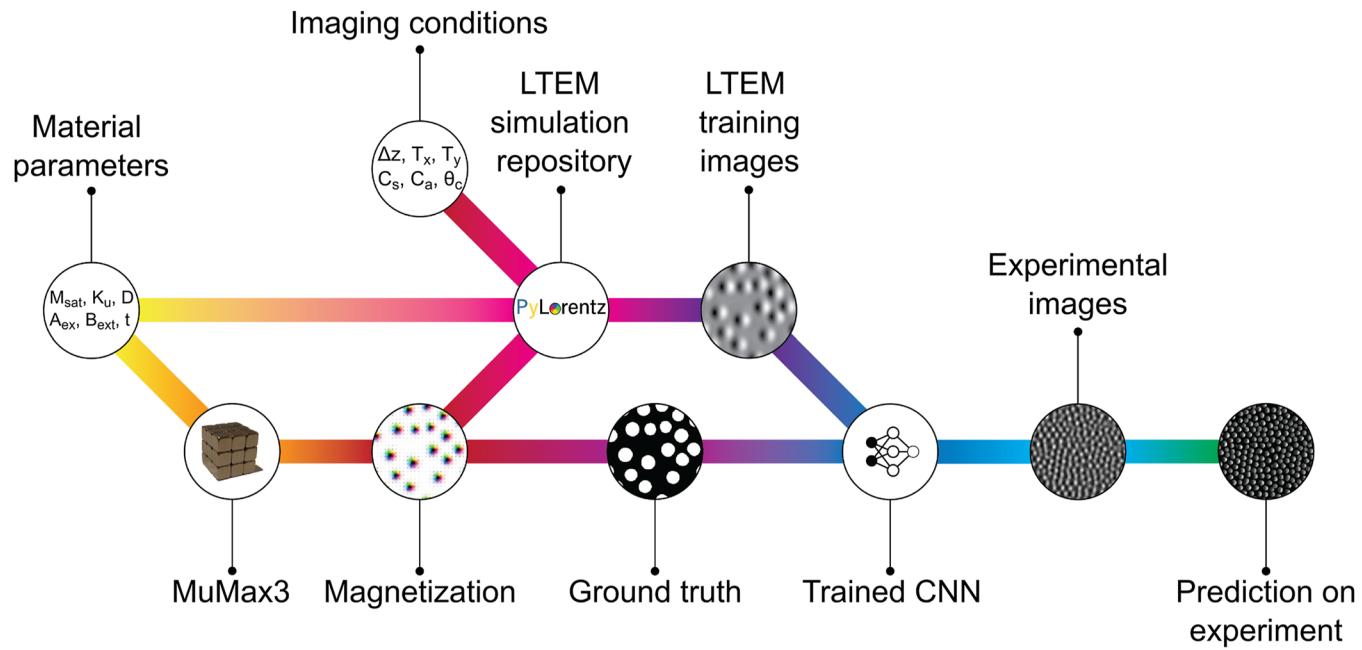


FIG. 1. Flow chart for training and applying a CNN model to LTEM data. The diagram reads from left to right, beginning with material parameters for the sample that will be imaged. These are used to perform micromagnetic simulations of potential magnetic configurations, which are used to create ground truth labels and simulate LTEM images. A model is trained on the simulated data and can be directly applied to experimental images.

A. Simulating Néel skyrmion lattice magnetizations

Generating an appropriate set of magnetization configurations is the most important step of the training process. It is critical that the magnetic configurations in the training set, simulated using a micromagnetics formalism, encompass a broad range of features that could be observed during an *in situ* LTEM experiment. Any features present in an experimental image that are not well represented in the training data will be omitted or incorrectly identified. It is, therefore, important when creating training data to consider the many effects that lead to non-idealized magnetic domain configurations in experimental samples, including pinning effects and local variations in the magnetic parameters. Failing to use a diverse enough training set can lead the model to overfit and incorrectly identify non-existent features in experimental data.

In the work presented here, the micromagnetic parameters that we included were the saturation magnetization, M_{sat} , the exchange stiffness, A_{ex} , the uniaxial anisotropy, K_u , and the interfacial Dzyaloshinskii-Moriya interaction (DMI) strength, D . These magnetic parameters vary within any given sample around the nominally measured, bulk values. Flakes of exfoliated vdW materials, for example, will often have a non-uniform topography, with bends and wrinkles due to strain, and this can alter the magnetic parameters

in those regions.^{26–28} Furthermore, the magnetic parameters of a material are always temperature dependent. We want to apply our model to *in situ* series of LTEM images obtained during heating and cooling, and we must, therefore, be sure that the full range of magnetic domain structures that could be observed experimentally are reflected in the micromagnetic simulations.

We compiled a training set of 10 000 micromagnetic simulations. The simulations were of dimension $(256 \times 256 \times 1)$ cells, with a cell size of $(3 \times 3 \times 27.5)$ nm. In order to accurately reflect experimental domain structures, micromagnetic simulations were run with some of the magnetic parameters varied across the simulated sample. The simulations were given mean and standard deviation values of M_{sat} , A , and D and were divided into 256 in-plane grains. The M_{sat} , A , and D values were then chosen randomly for each grain using the mean and standard deviation for that simulation. The total range of these values was $M_{\text{sat}} \in 1.5 \times 10^5$ – 3.0×10^5 A/m, $A \in 1.1 \times 10^{-11}$ – 1.9×10^{-11} A/m, and $D \in 1 \times 10^{-3}$ – 1.3×10^{-3} A/m². The uniaxial anisotropy was kept constant with a values of $K_u = 1.5 \times 10^6$ J/m³ across all simulations, as was the external applied field value of $B_{\text{ext}} = 0.4$ T. This value of applied magnetic field is much larger than typical experimental values and is often necessary to stabilize skyrmions in micromagnetic simulations.²⁹

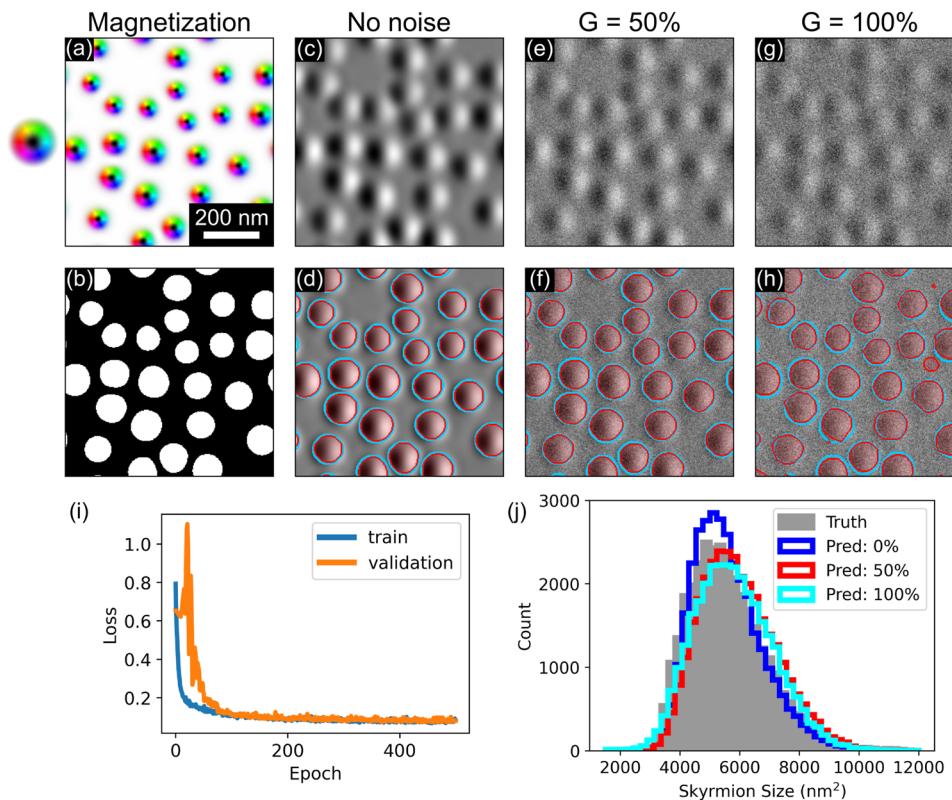


FIG. 2. Application of a SkymNet to simulated data. (a) True magnetization used to simulate input images. (b) Ground-truth label of skyrmion size and position, created from (a). (c) LTEM image simulated from (a) with defocus $\Delta z = -2$ mm and sample tilt angle $T_x = 20^\circ$. (d) SkymNet output of (c), with the identified skyrmions shown in red and overlaid on both the ground truth skyrmion edges (blue) and the input image. (e) and (g) Simulated LTEM images like (c), but with 50% and 100% noise added, respectively. (f) and (h), SkymNet segmentation output overlaid on the ground truth for (e) and (g). (i) Loss plot when training the CNN. (j) Histogram of skyrmion sizes for ground-truth labels and SkymNet predictions with 0%, 50%, and 100% added noise, across the 1000-image test set.

The values used for each energy term were manually determined in order to ensure that a wide range of domain structures were produced while maintaining few simulations with domain structures other than Néel skyrmions and stripes. We performed 90% of the simulations with a simulated grain structure and varying magnetic parameters, while 10% were performed with uniform magnetic parameters across the entire simulated region. This resulted in a diverse training set of magnetic configurations spanning a wide range of Néel skyrmion sizes and densities.

B. Defining the ground truth

The magnetization outputs from MuMax were each used to define a ground truth image. Our ground truth image, for instance segmentation with two classes, is a binary image defining the size and location of each skyrmion. There are multiple ways of defining the skyrmion radius, because the 2π domain wall across a skyrmion has a gradual falloff toward the “bulk” magnetization orientation.³⁰ We define the ground truth skyrmion location from the z -component of the magnetization. The background magnetization points in the $+z$ direction, and we define skyrmions as regions where $M_z <= 0.98$. Figure 2(a) shows the magnetization of a simulated Néel skyrmion lattice, and the ground truth label is shown in Fig. 2(b).

C. Simulating LTEM training images

The LTEM training images were simulated using the PyLorentz package, which is an open-source software suite developed for quantitative image analysis of LTEM images.^{24,31} The simulated magnetization maps and material parameters are input into PyLorentz to simulate the electron phase shift through the sample. Along with microscope parameters, including defocus and aberrations, the phase shift is used to simulate the resulting LTEM image. Figure 2(c) shows a simulated LTEM image from the magnetization in Fig. 2(a), with a sample tilt of 20° around the horizontal x axis and at a defocus of $\Delta z = -2$ mm. Note that despite the separation between adjacent skyrmions visible in the magnetization map, contrast from adjacent skyrmions blends together in the LTEM image. This makes it very difficult to distinguish individual skyrmion sizes and positions.

Unless otherwise noted, the parameters used for the simulated LTEM images in this work are $\theta_x = 20^\circ$, $\theta_y = 0^\circ$, $\Delta z = -2$ mm, accelerating voltage $E = 200$ kV, spherical aberration constant $C_s = 200$ μm , beam coherence $\theta_c = 0.01 \times 10^{-3}$ rad, and defocus spread $\Delta = 80$ nm. These represent reasonable microscope aberrations and imaging conditions for performing LTEM, although it should be noted that the simulation methods used here do not include the effects of structural defects that may be present in the sample.

D. Machine learning with a convolutional neural network

To perform instance segmentation, we implemented a fully convolutional neural network based on a U-net architecture that has proven robust at image segmentation in applications ranging from biological imaging to atomic resolution TEM.³² The 10 000 simulated images were split into training and validation datasets by a ratio of 80:20, with 1000 separately simulated magnetization distributions and images used for test data. Random noise, including Poisson, Gaussian, salt and pepper, and contrast variation, was

added to the training images, so the model could be directly applied to experimental data.

The model was trained for 500 epochs using a cross-entropy loss, and the training loss and validation loss are plotted in Fig. 2(i), which show the rapid convergence of the model. Training the model can be performed quickly; training the model for 500 epochs took less than two hours using an NVIDIA Tesla M10 GPU, and this would be much faster still if performed with a current-generation GPU.

III. RESULTS

The LTEM image shown in Fig. 2(c) was given as an input to our model, and the output is shown in Fig. 2(d), overlaid on the input image. The SkyrmNet-identified skyrmions are shown in red, and the edges of the skyrmions as defined by the ground-truth label are outlined in blue. Due to the large overlap between the ground truth and the prediction, it is evident that the CNN does a good job of identifying the skyrmion locations and sizes, although in the zero-noise case (as shown here), it slightly underestimates the true skyrmion size.

In order to demonstrate the application of a SkyrmNet to a more realistic input image, we added noise to the simulated LTEM input image. Figures 2(e) and 2(g) show the same image as in (c), but with the addition of 50% and 100% Gaussian noise, $G = 50\%$ and $G = 100\%$, respectively, where the level of Gaussian noise refers to the standard deviation of the distribution divided by the mean intensity of the image. When carrying out LTEM imaging, there are many noise sources, including all the types that we added to the training data. We have previously found that our models perform similarly well when given test data that have equivalent levels of pure Gaussian noise or a mix of noise types,³³ and we, therefore, use Gaussian noise here as it is easier to quantify.

The SkyrmNet outputs, given the noisy simulated LTEM images as inputs, are shown in Figs. 2(f) and 2(h) for the 50% and 100% noise cases, respectively. Qualitatively, it seems that the model performed very well for the image with 50% added noise, and the output segmentation appears very similar to the no-noise case shown in Fig. 2(d). In the case of the image with 100% added noise, there are more significant deviations of the prediction from the ground truth, and there are also some small, non-existent skyrmions that are incorrectly identified.

Figure 2(j) shows a histogram that we used to carry out an initial test of the accuracy of our model. It plots all the true skyrmion sizes from the 1000-image test set, as well as the histograms of SkyrmNet-predicted sizes for 0%, 50%, and 100% added noise. We see that in the zero-noise case, the use of SkyrmNet leads to a narrower skyrmion size distribution, thus underestimating the number of skyrmions that are both smaller, and larger, than the average. The SkyrmNet-created histograms for noisy inputs, however, shift to the right, indicating the SkyrmNet starts to overestimate the size of the skyrmions as noise increases. This is interesting because, in Figs. 2(f) and 2(h), it appeared that SkyrmNet was underestimating the skyrmion sizes even with added noise. This demonstrates why it is important to use a large and varied test set that represents a wide range of spin texture sizes and densities. It also shows that cumulative accuracy measurements, such as through a histogram, can give misleading results because incorrect predictions can average out.

A. Accuracy quantification

In order to more thoroughly assess the accuracy of our model, SkyrmNet was used to segment all 1000 images in our test set, and the SkyrmNet output for each image was quantitatively compared to the ground-truth label. Quantifying the quality of a segmentation is non-trivial, with different metrics giving more or less weight to false positives and negatives, and with some metrics better suited to balanced or imbalanced datasets. We, therefore, used three separate metrics: the subset accuracy,³⁴ the F1 score,³⁵ and the Jaccard similarity coefficient.³⁶ Each of these metrics quantifies the similarity between a binary prediction and a binary ground truth, with a value of 1 for identical images, but each metric gives different weights to the recall of the prediction (whether or not the prediction finds all parts of all skyrmions) and the precision of the prediction (whether or not the prediction correctly identifies regions that are not skyrmions). All three metrics were calculated for each image and averaged across the test set, thus giving three final values for the quality of the SkyrmNet segmentation.

One of the magnetization distributions from the test set is shown in Fig. 3(a), with the simulated phase shift and resulting LTEM image shown in Figs. 3(b) and 3(c). Figures 3(d) and 3(e) show the same image as in (c), but with the addition of 10% and 50%

Gaussian noise, respectively. Figure 3(f) shows a plot of the three accuracy metrics for our model as a function of noise added to the input images. All three metrics show that the model performs worse with added noise, but the decrease in performance was not drastic. The model performed particularly well at higher noise levels according to the subset accuracy, while the Jaccard score showed the largest decrease in performance. The Jaccard score is commonly used to evaluate image segmentation tasks, and we find that it gives the lowest score of our three metrics in all cases. This is in large part due to its sensitivity to imbalanced classes, i.e., images for which much less than 50% of the total pixels belong to the skyrmion class. These types of images comprised the majority of our test set, which we believe results in the lower Jaccard score relative to the other metrics.

It can be difficult to take ML models that are trained exclusively on simulated data and then apply them to experimental data. One reason for this is that the experimental data might not be represented in the training dataset. This is especially prevalent for LTEM imaging where, unless the microscope is frequently calibrated, the imaging parameters cannot be known precisely. In particular, the defocus and sample tilt are two parameters that make a large difference in the resulting LTEM image, but are usually not known accurately. In order to assess the resilience of our model to imaging conditions

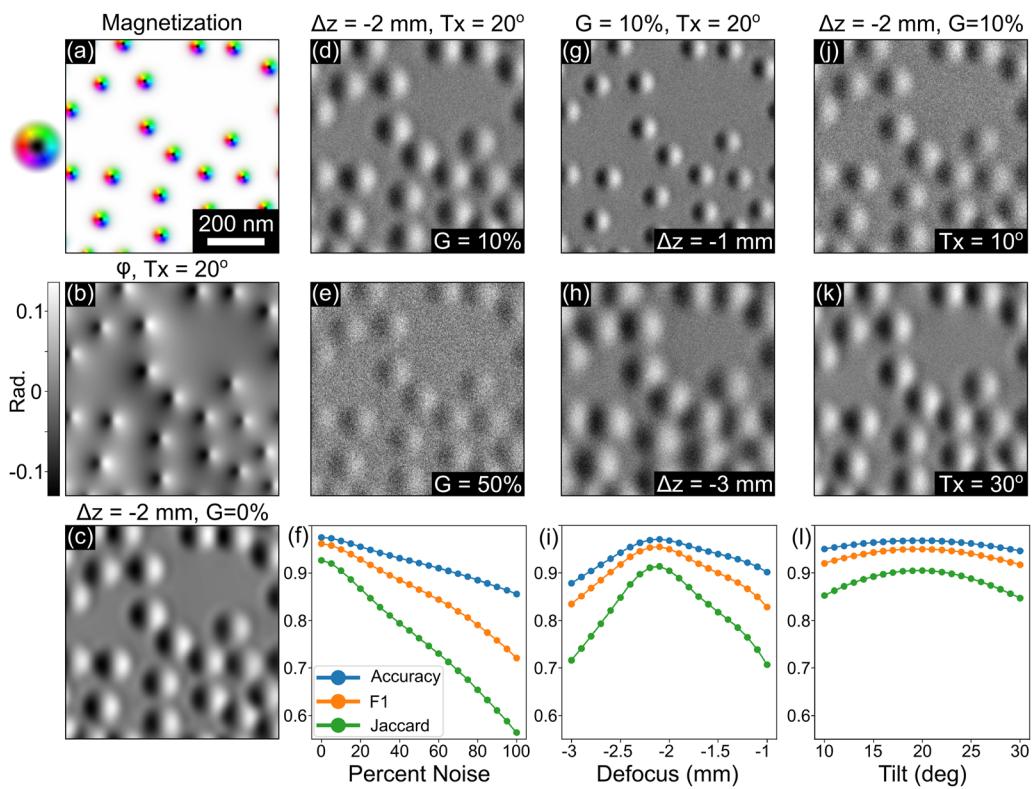


FIG. 3. SkymNet accuracy quantification. (a) Example ground truth magnetization distribution from the test data, (b) corresponding simulated magnetic phase shift for tilt, $T_x = 20^\circ$, and (c) LTEM image simulated from (b) with $\Delta z = -2$ mm. (d) LTEM image from (c) with 10% and (e) 50% added noise. (f) Plot of the SkymNet accuracy vs percent noise using three different quantification metrics. (g) LTEM image simulated from (b) with $\Delta z = -1$ mm and (h) $\Delta z = -1$ mm. (i) Accuracy plots for a SkymNet trained on LTEM images with $\Delta z = -2$ mm, for input images with 10% noise and varying defocus values. (j) and (k) LTEM images simulated from the same magnetization shown in (a) with $\Delta z = -2$ mm but with (j) $T_x = 10^\circ$ and (k) $T_x = 30^\circ$. (l) Accuracy plot for a SkymNet trained on images with $T_x = 20^\circ$, for input images with 10% noise and varying tilt values.

that are different from those used for the training data, we simulated LTEM images for test datasets under a range of imaging parameters and then applied the same SkyrmNet model to all of them. This was the same model as is demonstrated in Fig. 2, and it was trained only on simulated data with $\Delta z = -2$ mm and $T_x = 20^\circ$.

Figures 3(g) and 3(h) show images simulated from the phase shift shown in Fig. 3(b), but with defocus values of $\Delta z = -1$ mm and $\Delta z = -3$ mm, respectively. Both images also have 10% Gaussian noise added. From these images, it is clear how a larger defocus value leads to significant blurring and an apparent increase in skyrmion size. Figure 3(i) plots the accuracy metrics for a SkyrmNet trained only on images with a defocus value of $\Delta z = -2$ mm, which is then applied to our test set simulated with defocus values ranging from -1 to -3 mm, and in all cases with the addition of 10% Gaussian noise. As expected, the prediction quality decreases for input images that were simulated with defocus values that are further from -2 ± 0.2 mm. However, the model is still very accurate within the range of -2 ± 0.2 mm, which represents a moderate defocus offset that is realistic for experimental conditions.³⁷

We next examine how an incorrect sample tilt can affect the segmentation accuracy. Figures 3(j) and 3(k) show two simulated LTEM images with defocus $\Delta z = -2$ mm and 10% added Gaussian

noise, which were simulated at tilt angles of $T_x = 10^\circ$ and 30° , respectively. The image simulated at a 30° tilt shows increased contrast due to the larger component of the skyrmion magnetization that is perpendicular to the beam direction. Figure 3(l) shows a plot of the accuracy metrics of our model as a function of the input tilt angle, with the model trained only on images with a tilt angle of 20° . We see an expected reduction in accuracy as tilt angles move further from 20° , but the reduction is small compared to the reduction in accuracy caused by incorrect defocus values. There is a negligible loss of accuracy when considering the goniometer error margin of $\pm 2^\circ$,³⁸ and the model performs well within a $\pm 5^\circ$ range around the angle used for the training dataset. This is particularly useful when imaging single crystalline samples, as it can be necessary to adjust the tilt angle when imaging to account for the variations in diffraction contrast that can occur during *in situ* experiments. Figure 3(l) shows that a single model can be used for analysis even if the tilt angle must be changed slightly during imaging.

B. Application to *in situ* experimental data

Here, we demonstrate that SkyrmNet can be effectively applied to experimental *in situ* images of skyrmions. We choose two scientifically relevant examples shown in Fig. 4: the temperature

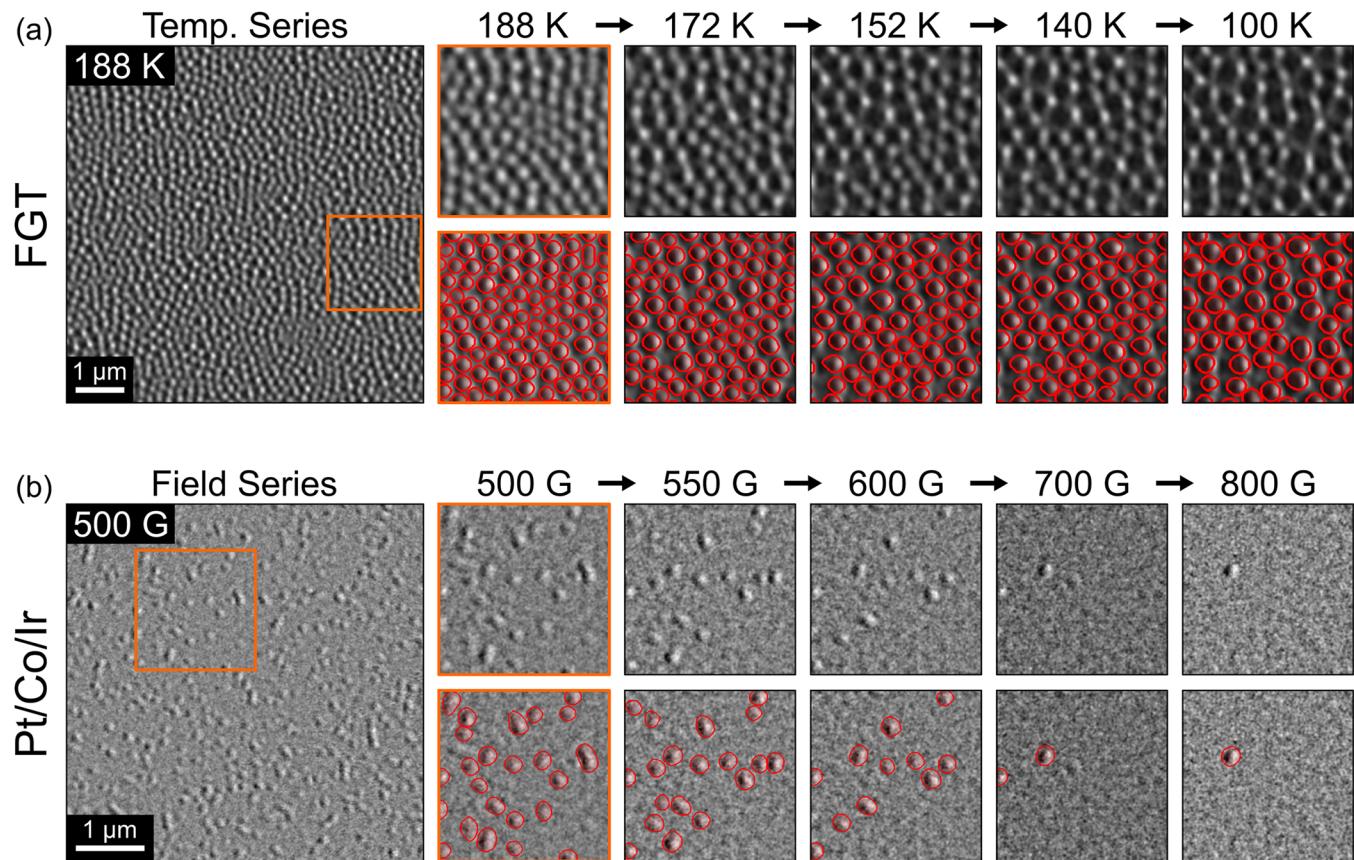


FIG. 4. Application of a SkyrmNet to experimental data. (a) Tracking how a dense skyrmion lattice in FGT evolves as a function of temperature. A full field-of-view LTEM image is shown on the left; the top row shows how skyrmions in the orange box evolve as temperature is reduced from 188 to 100 K; the bottom row shows the SkyrmNet-identified skyrmions (outlined in red) for each image. (b) Tracking how isolated skyrmions in a Pt/Co/Ir metallic multilayer evolve under an increasing applied field from 500 to 800 G.

dependence of a dense Néel skyrmion lattice in a vdW material¹² and the effect of increasing applied field on diffuse skyrmions in a metallic multilayer.³⁹ The experimental data were recorded on a JEOL JEM-2100F TEM instrument in “low-mag” mode. Two SkyrmNet models were trained, each using the same 10 000 micromagnetic simulations and microscope aberration parameters stated previously. The models are different in that they were trained with different defocus and sample tilt values, which were determined by the experimental conditions of each dataset. It is important that these experimental parameters are not dictated by the CNN, as the optimal choice for both defocus and sample tilt is highly dependent on the properties of each individual sample and the imaging conditions used. They must, therefore, be optimized by the microscope user during the experiment. An advantage of the SkyrmNet method is that training a new model for a different microscope or new imaging conditions is fast, and the method is, therefore, easily adaptable to different experimental datasets.

Figure 4(a) shows the application of a SkyrmNet to images of a dense skyrmion lattice in FGT, obtained with a defocus of $\Delta z = -8$ mm, a tilt angle of 22° , and an applied field of 500 G. The image has been bandpass-filtered to reduce noise before segmentation. FGT is an interesting material to study the behavior of skyrmions, because its magnetic parameters change greatly between 220 and 100 K, which is the range between its ordering temperature, T_c , and the minimum temperature that can be easily achieved in the TEM instrument using a liquid nitrogen holder. The most stable skyrmion size is dependent on the magnetic parameters and, therefore, changes as a function of temperature. This leads to the formation of small skyrmions, which nucleate when field-cooling through T_c and are gradually and irreversibly replaced by larger skyrmions as temperature decreases. Studying this process provides insights into the energy landscape of FGT and the nature of the topological stability of the skyrmions.^{40,41}

It is necessary to use a very large defocus when imaging the evolution of Néel skyrmion lattices in FGT. The behavior of the skyrmions is most interesting near T_c , as this is where the topological spin textures stabilize and are most mobile, but at these temperatures, FGT has a small saturation magnetization. This necessitates using a very large defocus value to increase phase sensitivity and to image the spin textures. The large defocus required leads to strong blurring of the spin textures and makes it difficult to accurately determine their size and location, thus highlighting the need for ML-enabled analysis of these images. As shown in Figs. 3(g) and 3(h), higher defocus leads to significantly increased blurring and an increase in the apparent skyrmion size. Without the ML-based method presented here, it would be very difficult to determine true skyrmion sizes in these highly defocused images.

The top row of Fig. 4(a) shows LTEM images as a function of temperature for the same region that is highlighted by the orange box in the left-hand image. The bottom row shows the application of a SkyrmNet to these images, with each skyrmion shown in red and overlaid on the input image. Qualitatively, SkyrmNet has identified all the skyrmions in the image, whether they are densely packed or located near gaps in the lattice that occurred during cooling. Information about the gaps between skyrmions could provide information about the relative strengths of pinning forces and inter-skyrmion repulsion.^{10,30} In addition, from these segmentation data, it would be possible to quantitatively extract the size and

deformation of each skyrmion, which could then be correlated with local lattice order.

Figure 4(b) shows the application of a SkyrmNet to LTEM images of a diffuse skyrmion lattice in a metallic multilayer. These images were recorded while increasing the applied out-of-plane magnetic field, in order to observe how the skyrmions were driven out of the sample. The full field-of-view image is shown on the left and was recorded at room temperature, with a defocus of $\Delta z = -4$ mm, a tilt angle of 30° , and with a 500 G applied field. The skyrmions shown here are similar to reported images of skyrmions in other metallic multilayers.³⁹ We observed many skyrmions that were extended or deformed, and the skyrmion density was low compared to what is seen in FGT. The top row of images in Fig. 4(b) shows how skyrmions shrank and were destroyed as the applied out-of-plane field was increased. The bottom row shows the application of a SkyrmNet to each of these images. We are able to accurately track both the skyrmions as they were driven out of the sample, and also observe how the skyrmion shapes change with the applied field.

We have, therefore, shown that a SkyrmNet, despite being trained only on simulated data, can be applied to both an exfoliated flake of a vdW ferromagnet and a metallic multilayer film. This method is especially useful when applied to *in situ* experiments, as it allows for tracking the changes in skyrmion size and location between subsequent frames. Inference of a CNN is quite fast and can be performed in a second, depending on image size. This method can, therefore, enable nearly real-time analysis when performing *in situ* LTEM experiments.

IV. DISCUSSION

By combining micromagnetic simulations and LTEM image simulations, we can create a training set that enables a CNN to identify the size and shape of Néel skyrmions in experimental LTEM images. This directly enables statistical studies of Néel skyrmion lattices and how they can be collectively manipulated by externally applied stimuli. We hope, however, that the most important feature of this work is not the direct application to Néel skyrmions, but, instead, the broader impact of the workflow we have developed. We have shown that ML models can be trained on simulated LTEM images to accurately identify the extent of features and make quantitative measurements that cannot otherwise be determined. As shown here and in other studies,^{22–24} LTEM image contrast can be very difficult to interpret, especially when imaging complex spin textures such as skyrmions. Micromagnetic simulations allow for direct modeling of these spin textures in a wide range of materials, and these magnetization distributions can be used to create accurate training labels for any magnetic feature of interest. For example, using this pipeline, a CNN could be trained to classify magnetic bubbles by chirality and to distinguish type I and type II bubbles, or to identify antiskyrmions and other spin textures for which the projected magnetic induction is very different from the magnetization.²⁴ Using this method, therefore, a SkyrmNet could be trained to identify more complex spin textures that are otherwise more difficult to identify.

While we hope that this method can be broadly applied to spin textures beyond Néel skyrmions, it does have limitations. Most importantly, a CNN can only be trained to identify features that

are contained within its training set and applying it to experimental data that are not well represented in training can lead to erroneous predictions. Furthermore, we do not think it will be possible to reconstruct the actual magnetization map with any accuracy. Features such as domain wall width can be modeled with micromagnetic simulations, but variations in domain wall width are not always clearly seen in the resulting LTEM images, due to defocus. This is especially noticeable when the in-plane components of the magnetization produce negligible LTEM contrast, such as is the case for Néel domains as demonstrated in this work. The sample must be tilted to induce magnetic contrast, and in this case, the contrast arising from the in-plane component of the magnetization is very subtle. It would, therefore, be difficult to train a model to identify details about the in-plane magnetization while maintaining a wide training dataset. Another constraint is that we have implicitly assumed that the magnetization is nearly uniform through the out-of-plane direction of the sample. LTEM is by nature a transmission technique and is, therefore, sensitive only to the integrated induction and cannot distinguish variations along the beam direction.

Despite these limitations, we believe that our method can be broadly applied to a wide range of materials and spin textures. A major advantage of this method is the relative ease with which a model can be trained and adapted for each experimental dataset. Running the micromagnetic simulations to assemble a training set is the only computationally intensive step of the process, and it takes on the order of days to complete. This is also the most important step of the process because, as previously discussed, these simulations must encompass the full range of spin textures that might be observed experimentally. Once the magnetization distributions have been simulated, creating the ground truths, then simulating the LTEM images, and training a CNN can be completed in a matter of hours. This means that a training set pertaining to a particular material or magnetic domain type can be used to train models for any experimental imaging conditions. This enables the model to conform to imaging parameters as dictated by the experiment, rather than requiring that the experiment be performed with specific, and potentially non-optimal, imaging conditions.

V. CONCLUSIONS

In this work, we have demonstrated that a CNN, trained on simulated data, can be used to identify features in experimental LTEM images that would not otherwise be identified. We have applied this method to performing instance segmentation of *in situ* images of Néel skyrmions in FGT and Pt/Co/Ir, thus obtaining the size and shape of each skyrmion. All the steps necessary for training this model are open-source, and this method can be broadly adapted to identifying magnetic features in any system for which LTEM imaging is appropriate. We believe that it will enable a new type of large-scale, statistical study of magnetic spin textures not previously possible using LTEM.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division. Use of the Center for Nanoscale Materials, an Office of Science user facility, is supported by the U.S.

Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Arthur R. C. McCray: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Software (lead); Writing – original draft (equal). **Alec Bender:** Data curation (equal); Software (equal); Writing – original draft (equal). **Amanda Petford-Long:** Project administration (equal); Supervision (equal); Writing – review & editing (equal). **Charudatta Phatak:** Conceptualization (equal); Project administration (lead); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in the Materials Data Facility, and the code that supports the findings of this study is openly available in GitHub at https://github.com/Art-MC/SkyrmNet_demo.

REFERENCES

- ¹S. Ghosh, P. Rüßmann, Y. Mokrousov, F. Freimuth, and A. Kosma, *J. Appl. Phys.* **133**, 230901 (2023).
- ²F. Hellman, A. Hoffmann, Y. Tserkovnyak, G. S. Beach, E. E. Fullerton, C. Leighton, A. H. MacDonald, D. C. Ralph, D. A. Arena, H. A. Dürr, P. Fischer, J. Grollier, J. P. Heremans, T. Jungwirth, A. V. Kimel, B. Koopmans, I. N. Krivorotov, S. J. May, A. K. Petford-Long, J. M. Rondinelli, N. Samarth, I. K. Schuller, A. N. Slavin, M. D. Stiles, O. Tchernyshyov, A. Thiaville, and B. L. Zink, *Rev. Mod. Phys.* **89**, 025006 (2017).
- ³B. Göbel, I. Mertig, and O. A. Tretiakov, *Phys. Rep.* **895**, 1 (2021).
- ⁴K. Wang, V. Bheemarasetty, J. Duan, S. Zhou, and G. Xiao, *J. Magn. Magn. Mater.* **563**, 169905 (2022).
- ⁵K. Everschor-Sitte, J. Masell, R. M. Reeve, and M. Kläui, *J. Appl. Phys.* **124**, 240901 (2018).
- ⁶A. Fert, N. Reyren, and V. Cros, “Magnetic skyrmions: Advances in physics and potential applications,” *Nat. Rev. Mat.* **2**, 17031 (2017).
- ⁷R. Msiska, J. Love, J. Mulkers, J. Leliaert, and K. Everschor-Sitte, *Adv. Intell. Syst.* **5**, 2200388 (2023).
- ⁸K. Raab, M. A. Brems, G. Beneke, T. Dohi, J. Rothörl, F. Kammerbauer, J. H. Mentink, and M. Kläui, *Nat. Commun.* **13**, 6982 (2022).
- ⁹M. E. Henderson, B. Heacock, M. Bleuel, D. G. Cory, C. Heikes, M. G. Huber, J. Krzywon, O. Nahman-Levesqué, G. M. Luke, M. Pula, D. Sarenac, K. Zhernenkov, and D. A. Pushin, *Nat. Phys.* **19**, 1617 (2023).
- ¹⁰Y. Ge, J. Rothörl, M. A. Brems, N. Kerber, R. Gruber, T. Dohi, M. Kläui, and P. Virnau, *Commun. Phys.* **6**, 30 (2023).
- ¹¹P. Meisenheimer, H. Zhang, D. Raftrey, X. Chen, Y.-T. Shao, Y.-T. Chan, R. Yalisoive, R. Chen, J. Yao, M. C. Scott, W. Wu, D. A. Muller, P. Fischer, R. J. Birgeneau, and R. Ramesh, *Nat. Commun.* **14**, 3744 (2023).
- ¹²A. R. C. McCray, Y. Li, R. Basnet, K. Pandey, J. Hu, D. P. Phelan, X. Ma, A. K. Petford-Long, and C. Phatak, *Nano Lett.* **22**, 7804 (2022).
- ¹³J. Závorka, F. Dittrich, Y. Ge, N. Kerber, K. Raab, T. Winkler, K. Litzius, M. Veis, P. Virnau, and M. Kläui, *Adv. Funct. Mater.* **30**, 2004037 (2020).
- ¹⁴T. Matsumoto, Y.-G. So, Y. Kohno, H. Sawada, Y. Ikuhara, and N. Shibata, *Sci. Adv.* **2**, e1501280 (2016).

- ¹⁵P. Huang, T. Schönenberger, M. Cantoni, L. Heinen, A. Magrez, A. Rosch, F. Carbone, and H. M. Rønnow, *Nat. Nanotechnol.* **15**, 761 (2020).
- ¹⁶C. Phatak, O. Heinonen, M. De Graef, and A. Petford-Long, *Nano Lett.* **16**, 4141 (2016).
- ¹⁷D.-T. Ngo and L. T. Kuhn, *Adv. Nat. Sci.: Nanosci. Nanotechnol.* **7**, 045001 (2016).
- ¹⁸Y. Li, R. Basnet, K. Pandey, J. Hu, W. Wang, X. Ma, A. R. C. McCray, A. K. Petford-Long, and C. Phatak, *JOM* **74**, 2310 (2022).
- ¹⁹T. Shimojima, A. Nakamura, and K. Ishizaka, *Microscopy* **72**, 287 (2023).
- ²⁰T. R. Harvey, N. Rubiano da Silva, J. H. Gaida, M. Möller, A. Feist, S. Schäfer, and C. Ropers, *MRS Bull.* **46**, 711 (2021).
- ²¹M. De Graef, *Introduction to Conventional Transmission Electron Microscopy* (Cambridge University Press, Cambridge, 2003), Vol. 38.
- ²²J. C. Loudon, A. C. Twitchett-Harrison, D. Cortés-Ortuño, M. T. Birch, L. A. Turnbull, A. Štefančič, F. Y. Ogrin, E. O. Burgos-Parra, N. Bukin, A. Laurenson, H. Popescu, M. Beg, O. Hovorka, H. Fangohr, P. A. Midgley, G. Balakrishnan, and P. D. Hatton, *Adv. Mater.* **31**, 1806598 (2019).
- ²³Y. Yao, B. Ding, J. Cui, X. Shen, Y. Wang, W. Wang, and R. Yu, *Appl. Phys. Lett.* **114**, 102404 (2019).
- ²⁴A. R. McCray, T. Cote, Y. Li, A. K. Petford-Long, and C. Phatak, PyLorentz, 2021.
- ²⁵A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sánchez, and B. Van Waeyenberge, *AIP Adv.* **4**, 107133 (2014).
- ²⁶Y. Ga, Q. Cui, J. Liang, D. Yu, Y. Zhu, L. Wang, and H. Yang, *Phys. Rev. B* **106**, 054426 (2022).
- ²⁷A. R. C. McCray, Y. Li, E. Qian, Y. Li, W. Wang, Z. Huang, X. Ma, Y. Liu, D. Y. Chung, M. G. Kanatzidis, A. K. Petford-Long, and C. Phatak, *Adv. Funct. Mater.* **33**, 2214203 (2023).
- ²⁸A. O'Neill, S. Rahman, Z. Zhang, P. Schoenherr, T. Yildirim, B. Gu, G. Su, Y. Lu, and J. Seidel, *ACS Nano* **17**, 735 (2023).
- ²⁹C. Liu, J. Jiang, C. Zhang, Q. Wang, H. Zhang, D. Zheng, Y. Li, Y. Ma, H. Algaidi, X. Gao, Z. Hou, W. Mi, J. Liu, Z. Qiu, and X. Zhang, *Adv. Sci.* **10**, 2303443 (2023).
- ³⁰D. Capic, D. A. Garanin, and E. M. Chudnovsky, *J. Phys.: Condens. Matter* **32**, 415803 (2020).
- ³¹A. R. McCray, T. Cote, Y. Li, A. K. Petford-Long, and C. Phatak, *Phys. Rev. Appl.* **15**, 044025 (2021).
- ³²O. Ronneberger, P. Fischer, and T. Brox, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Springer, Cham, 2015), pp. 234–241.
- ³³A. R. C. McCray, T. Zhou, S. Kandel, A. Petford-Long, M. Cherukara, and C. Phatak, *Npj Comput. Mater.* **10**(1), 111 (2024).
- ³⁴J. Nam, E. Loza Mencía, H. J. Kim, and J. Fürnkranz, in I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.) *Advances in Neural Information Processing Systems*, 30 (Curran Associates, 2017).
- ³⁵A. Tharwat, *Appl. Comput. Inform.* **17**, 168 (2021).
- ³⁶P. Jaccard, *New Phytol.* **11**, 37 (1912).
- ³⁷D. B. Williams and C. B. Carter, *Transmission Electron Microscopy* (Springer, 2009).
- ³⁸M. Hayashida, S. Terauchi, and T. Fujimoto, *Rev. Sci. Instrum.* **82**, 103706 (2011).
- ³⁹A. T. Clark, X. Wang, A. R. Stuart, Q. Wang, W. Jiang, J. E. Pearson, S. G. E. te Velthuis, A. Hoffmann, X. M. Cheng, and K. S. Buchanan, *J. Magn. Magn. Mater.* **563**, 169951 (2022).
- ⁴⁰M. T. Birch, D. Cortés-Ortuño, N. D. Khanh, S. Seki, A. Štefančič, G. Balakrishnan, Y. Tokura, and P. D. Hatton, *Commun. Phys.* **4**, 175 (2021).
- ⁴¹I. S. Lobanov, H. Jónsson, and V. M. Uzdin, *Phys. Rev. B* **94**, 174418 (2016).