Title

Magnetic Hamiltonian parameter estimation using deep learning techniques

Journal

Science Advances, 6(39)

ISSN

2375-2548

Authors

Kwon, HY

Yoon, HG

Lee, C

et al.

Publication Date

2020-09-25

DOI

10.1126/sciadv.abb0872

Peer reviewed

CONDENSED MATTER PHYSICS

Magnetic Hamiltonian parameter estimation using deep learning techniques

H. Y. Kwon[1]*, H. G. Yoon[2], C. Lee[2], G. Chen[3], K. Liu[3,4], A. K. Schmid[5], Y. Z. Wu[6], J. W. Choi[1], C. Won[2]*

Understanding spin textures in magnetic systems is extremely important to the spintronics and it is vital to extrapolate the magnetic Hamiltonian parameters through the experimentally determined spin. It can provide a better complementary link between theories and experimental results. We demonstrate deep learning can quantify the magnetic Hamiltonian from magnetic domain images. To train the deep neural network, we generated domain configurations with Monte Carlo method. The errors from the estimations was analyzed with statistical methods and confirmed the network was successfully trained to relate the Hamiltonian parameters with magnetic structure characteristics. The network was applied to estimate experimentally observed domain images. The results are consistent with the reported results, which verifies the effectiveness of our methods. On the basis of our study, we anticipate that the deep learning techniques make a bridge to connect the experimental and theoretical approaches not only in magnetism but also throughout any scientific research.

INTRODUCTION

Low-dimensional magnetic systems are an interesting topic of research due to their scientific significance and potential applications in next-generation electronic devices. In particular, unique spin textures in two-dimensional magnetic systems, such as magnetic stripe domains (*1*–*5*) and magnetic skyrmions (*6*–*10*), and their connection to various magnetic Hamiltonian parameters have been extensively studied. To quantitatively understand the properties of these magnetic structures, not only experimental methods (*11*–*14*) but also various theoretical methods (*15*–*21*) including numerical calculations have been used. In general, the results from numerical calculations or analytical approaches cannot be directly compared with experimental results, since not all experimental factors can be considered in theoretical calculations. Therefore, the development of a method of directly converting the experimental data to various essential parameters for the theoretical studies is an important issue not only in magnetism but also throughout scientific research in general.

Meanwhile, machine learning techniques, which first emerged decades ago, have been used in diverse fields, showing much enhanced performance and capabilities over conventional techniques. In particular, the deep learning that a machine learning technique using an artificial neural network with multiple layers between the input and output, which is called deep neural network (*22*), have been used to address many scientific challenges, including searching for the ab initio solution of many-electron systems (*23*), predicting protein structures (*24*), solving quantum many-body problems (*25*, *26*), and finding ground states of various magnetic systems (*27*). In particular, a previous study (*28*) showed the potential of deep learning techniques for predicting various properties of simulated low-dimensional magnetic systems.

In this study, we have used deep learning techniques to develop a method to estimate the magnetic Hamiltonian parameters from the experimentally observed images of magnetic domains formed in two-dimensional magnetic systems. The schematic diagram of this study is shown in Fig. 1.

First, energy-minimized spin configurations were generated using Monte Carlo–simulated annealing methods in which the temperature was decreased from above the Curie temperature to zero temperature. During the annealing process, the spontaneous symmetry breaking involves the formation of spin textures, so the domains and domain walls in the generated spin configurations have various morphological characteristics (Fig. 1A). Using these spin configurations as the input, several neural network structures were trained to estimate the magnetic Hamiltonian parameters, which were used to generate the input spin configurations (Fig. 1B). Error analysis was performed to test the validity of the trained network (Fig. 1C). Last, to further prove the effectiveness of the method, we used the trained network to estimate the magnetic Hamiltonian parameters from experimentally observed magnetic domain images using spin-polarized low-energy electron microscopy (SPLEEM) (Fig. 1D) (*29*).

RESULTS

Dataset generation, noise injection, and network capacity

We generated a total spin configuration dataset appearing in a two-dimensional magnetic system under the Hamiltonian in Eq. 1

$$\mathcal{H} = -J\Sigma_{<i,j>}\vec{S}_i\cdot\vec{S}_j - \Sigma_{<i,j>}\vec{\beta}_{ij}\cdot(\vec{S}_i\times\vec{S}_j) - D\ \Sigma_{i,j\neq i}\frac{3(\vec{S}_i\cdot\vec{r}_{ij})(\vec{S}_j\cdot\vec{r}_{ij})-\vec{S}_i\cdot\vec{S}_j\,|\vec{r}_{ij}|^2}{|\vec{r}_{ij}|^5} - K\,\Sigma_i S_{i,z}^2 \quad (1)$$

where J is the exchange interaction strength, β ($\vec{\beta}_{ij} = \beta\hat{\beta}_{ij}$) is the Dzyaloshinskii-Moriya interaction (DMI) strength (*30*, *31*), D is the magnetic dipolar interaction strength, $\vec{r}_{ij}$ is a dimensionless displacement vector between i and j grid sites, and K is the perpendicular

[1]Center for Spintronics, Korea Institute of Science and Technology, Seoul 02792, South Korea. [2]Department of Physics, Kyung Hee University, Seoul 02447, South Korea. [3]Department of Physics, University of California, Davis, Davis, CA 95616, USA. [4]Physics Department, Georgetown University, Washington, DC 20057, USA. [5]National Center for Electron Microscopy, Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [6]Department of Physics, State Key Laboratory of Surface Physics and Advanced Materials Laboratory, Fudan University, Shanghai 200433, China.
*Corresponding author. Email: cywon@khu.ac.kr (C.W.); soky572@gmail.com (H.Y.K.)

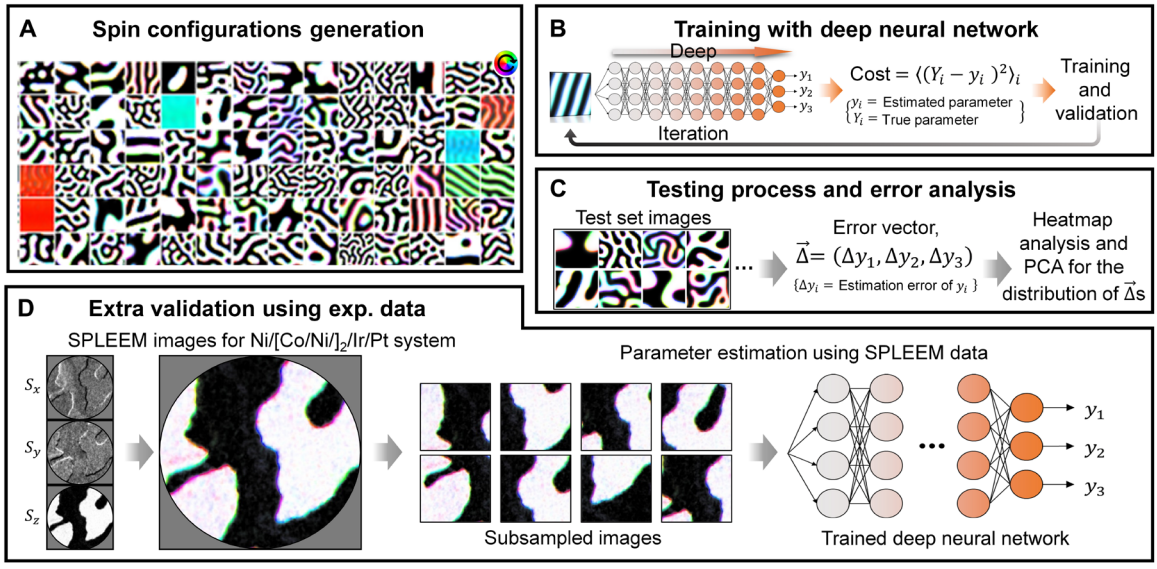

Fig. 1. Schematic diagram of our study. (**A**) Data generation process showing the sampled spin configurations generated through the simulated annealing process. The color wheel indicates the in-plane magnetization directions, and the grayscale indicates the out-of-plane magnetization directions. (**B** and **C**) The training and testing processes used in this study. (**D**) The additional validation process with experimentally observed magnetic domain images.

magnetic anisotropy (PMA) strength. All *ij* pairs are counted once in the summations. Since the range of each parameter used to generate the spin configurations is different, we defined the normalized parameters β_N, K_N, and D_N using the relation, $Y_N = \frac{Y - Y_{\min}}{Y_{\max} - Y_{\min}}$ where Y stands for all β, K, or D and $Y_{\min}$ and $Y_{\max}$ indicate the respective minimum and maximum values of each parameter range. These normalized parameters are used to train the networks, and the estimated parameters from the trained network are transformed into prenormalized values using a reverse process.

The goal of our study is to estimate Hamiltonian parameters from experimentally observed magnetic domain images. Whereas the experimental data have noises for various inevitable reasons, the spin configurations generated by simulation are noiseless in the final state of the simulated annealing process. The total generated spin configurations were divided into three "noiseless" datasets with a ratio of 70, 20, and 10% for training, validating, and testing the neural network, respectively, as shown in fig. S1. To investigate the effect of noises on magnetic domain images, we also constructed "noisy" training and validation datasets by injecting artificial noises with random amplitude into each spin configuration in the noiseless datasets (fig. S2A). In the case of the test datasets, the amplitude of the injected noises, ε_{Ts}, was adjusted step by step for a detailed analysis on the noise resistance of neural networks (fig. S2B). The details about the datasets and training process are described in Methods.

We built three network structures, ResNet18, ResNet50, and customized convolutional neural network (CNN) structures, as described briefly in fig. S3. ResNet (*32*) structures are well-known deep neural networks exhibiting a great performance in extracting features from images. A customized CNN was used to represent a non–deep neural network with relatively less network capacity than the ResNet structures, and the comparison with the latter highlights the importance of a deep neural network. Each neural network was trained separately using the noisy and noiseless datasets, and the mean absolute error (MAE) values were obtained using each test dataset after the training process as shown in Fig. 2.

As the injected noise amplitude ε_{Ts} increases, the accuracy of parameter estimation tends to decrease for all cases (Fig. 2), since the injected noises make it difficult to extract detailed characteristics of magnetic domain structures in the input images. Note that the networks trained using noisy images (left column in Fig. 2) are much more robust on various noise amplitudes in the test dataset than the networks trained only with noiseless images (right column in Fig. 2); there is almost one order difference in scale between the *y* axes ranges in the left and right columns in Fig. 2. The network, overfitted with the noiseless feature of the datasets, loses the accuracy rapidly as the strength of ε_{Ts} increases.

The network capacity also affects noise resistance. The ResNet18 structure, which has a medium size network capacity among the three network structures in our study, shows the most stable performances on various values of ε_{Ts}. Meanwhile, both the ResNet50 and the customized CNN, which have more deep and less deep structures than the ResNet18, are unstable and inaccurate compared to the ResNet18, respectively. The customized CNN is inaccurate for all parameter estimation when $\varepsilon_{Ts} > 0.5$ (left column of Fig. 2), and the ResNet50 trained only with noiseless dataset behaves in an unstable way when it is tested by the dataset with injected noises. This shows that small network capacity can make it difficult to extract sophisticated features from an image and that large network capacity can make overfitting problems serious (*33*). It means that there is a proper network capacity around the ResNet18. The capacity may be further optimized; however, the goal of this study is not to find the most optimal network. Therefore, given that ResNet18 already shows stable performance on noise resistance, the ResNet18 model trained with the noisy dataset was chosen for our further discussion.

Error estimation and error vector distribution

The characteristics of the magnetic domains shown in the input images are important for the trained network to estimate Hamiltonian parameters. The estimation errors are also affected by the domain

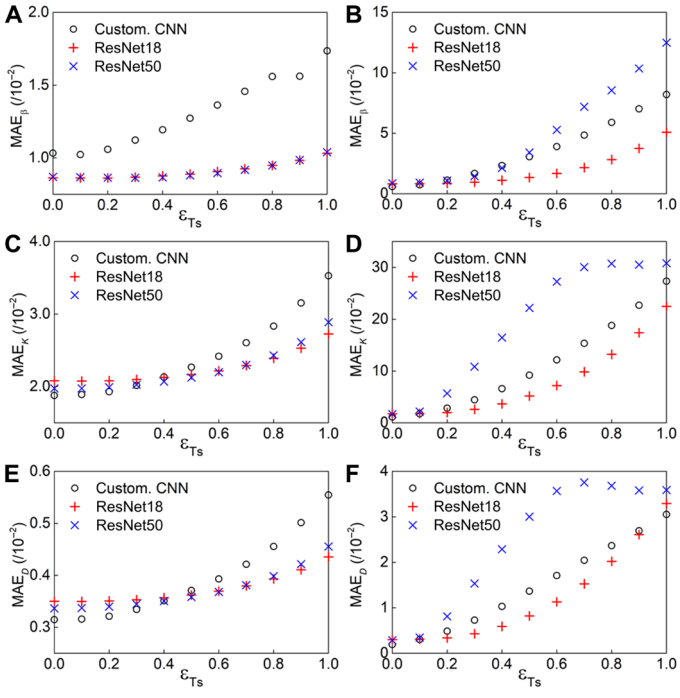

Fig. 2. Noise resistances and network capacities. Estimation errors for each of β (**A** and **B**), *K* (**C** and **D**), and *D* (**E** and **F**). The networks were trained using noisy (left column) and noiseless (right column) datasets. MAE_Y is $\langle | Y_{True} - Y_{Est.} | \rangle_{\varepsilon Ts}$ where *Y* stands for β, *K*, and *D*. The subscripts True and Est. indicate true and estimated parameters, respectively. Data points are calculated using each noisy test dataset with different ε_{Ts} values.

characteristics, and there can be specific relations between them. Figure 3 maps the input domain images and absolute error values from the test dataset on the same K_N and D_N parameter space.

It is found that the accuracies of the parameter estimation for β, *K*, and *D* are not uniform in the parameter space. In the case of β = 0 (first row in Fig. 3), the parameter estimation around $K_N \approx 0$ and $D_N \approx 0$ becomes inaccurate, which is reasonable, since the generated spin configurations under this condition are in-plane single domains without any domain wall structures (uniform color images in the left bottom corner of Fig. 3A). Without domain wall structures, β cannot be estimated correctly, since the DMI directly affects the domain wall characteristics in two-dimensional perpendicular magnetic systems (*18*, *20*). In addition to DMI strength, the anisotropy and dipolar interaction strength are also inaccurately estimated from in-plane single domains, since they have a wide range in $K < 2\pi D$ (*18*) and are not determined to a specific set of values. In the case of β = 3*D* (second row in Fig. 3), there is also a tendency that the accuracy of the parameter estimation decreases when the anisotropy becomes strong, indicated by the gray shades in left side of the heatmaps in Fig. 3 (F to H). One of the possible reasons for this is that other parameter values can generate comparable structural features shown in the domain images. In general, it is difficult to pinpoint what structural features of domain images were extracted by the trained networks and how the trained network estimates Hamiltonian parameters, since the network is too complex to be fully analyzed by a mathematical approach. Additional studies on what characteristics are extracted by the trained network from the input magnetic domain image are necessary to fully develop this approach.

One can notice that the estimation error values of anisotropy and dipolar interaction strength are strongly correlated. To investigate whether the inaccuracies of parameters are related, we defined the error vector as $\vec{\Delta} = (\Delta\beta, \Delta K, 2\pi\Delta D)$, which is composed of three error values, and investigated how $\vec{\Delta}$ are distributed. Figure 3I shows the tips of the error vectors as gray dots in a three-dimensional error space. We note that the estimation errors are not spherically dispersed but rather exhibit a two-dimensional flat distribution in a three-dimensional space. We performed a principal components analysis (PCA) (*34*) on error vectors to see how the uncertainties of estimations are related. The PCA, a statistical procedure using an orthogonal transformation, searches for the principal components that can best explain the variance in the data and reveals the internal structure of the data. The PCA results on the error vectors indicate that a specific axis, $\Delta K \cong 2\pi\Delta D$, is the most dominant principal component; this axis corresponds to a linear relationship between the error values of PMA and dipolar interaction strength. This does not merely mean that the estimation difficulties on *K* are related to those on *D* but rather implies that Δ*K* was compensated by 2πΔ*D* in the parameter estimation; *D* is overestimated/underestimated when *K* is overestimated/underestimated.

This is a notable result: We did not build the algorithm explicitly to include a specific relation between PMA and shape anisotropy from short-range dipolar interaction; yet, the neural network realized spontaneously through the training process that it is the effective anisotropy $K_{eff} (= K - 2\pi D)$ that is the key criterion determining the characteristics of out-of-plane magnetic domain structure, which is a well-known parameter in low-dimensional magnetism. The trained network has also realized that the minimization of $|\Delta K_{eff}| (= |\Delta K - 2\pi\Delta D|)$ is the most important factor for estimating the magnetic Hamiltonian parameters accurately from the two-dimensional spin configurations. One of the main advantages of deep learning techniques is that an efficient method can be obtained through the training process even without intrinsic programming. The spontaneous realization of the importance of K_{eff} by our neural network clearly exemplifies the advantage of deep learning techniques.

We also investigated how much the errors shown in Fig. 3 affect the structural characteristics of magnetic domains. Most of $|\vec{\Delta}|$ values are about 0.01, and these small errors do not significantly affect the characteristics of magnetic domain structures. More discussion is given in section S4.

Validation with SPLEEM data

To verify the validity and effectiveness of the trained network, we applied our network to the experimentally observed magnetic domain images of the Ni/$[Co/Ni]_2$/Ir/Pt(111) system shown in Fig. 4A. In the previous work (*35*), the effective magnitude and sign of the DMI were tuned by adjusting the Ir layer thickness, as the effective anisotropy strength was kept constant (fig. S5).

Figure 4B shows the SPLEEM images of magnetic domains separated by the chiral domain walls. We transformed the raw experimental images to colored images using the same color scheme as in the simulated results (Fig. 4C). Several subpanels with 100 by 100 pixels sizes were cropped from the colored images with 370 by 400 pixels to be used as input data of our trained network. The estimated values of *K* and *D* remain almost constant, while the estimated DMI parameter gradually changes from positive to negative values, as the Ir thickness increases (Fig. 4D). The estimations are consistent with the results reported in the experimental study (*35*).

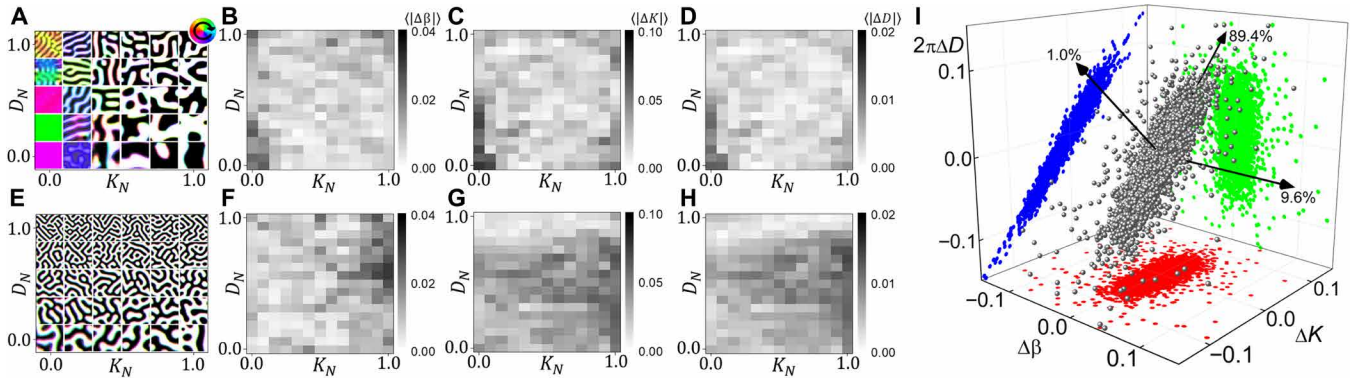

Fig. 3. Input spin configurations and heatmaps of estimation errors. (**A**) Examples of simulated spin configurations and heatmaps for (**B**) $\langle|\Delta\beta|\rangle$, (**C**) $\langle|\Delta K|\rangle$, and (**D**) $\langle|\Delta D|\rangle$ represented in K_N and D_N parameter space when $\beta = 0$. (**E**) Examples of simulated spin configurations and (**F** to **H**) heatmaps representations for each $\langle|\Delta\beta|\rangle$, $\langle|\Delta K|\rangle$, and $\langle|\Delta D|\rangle$ when $\beta = 3D$. The noiseless images are used in (A) and (E). The values located at the same (K_N, D_N) point of each heatmap were averaged. (**I**) Three-dimensional representation of the estimation errors. Nearly 8000 data points are sampled from the test dataset to be displayed in (I), where blue, red, and green dots are projections onto each plane. Black arrows show the principal components, and the numbers near the tips of arrows show the variance ratios from PCA [the most dominant component is along (0.005, 0.689, and 0.725) with an 89.4% variance ratio].

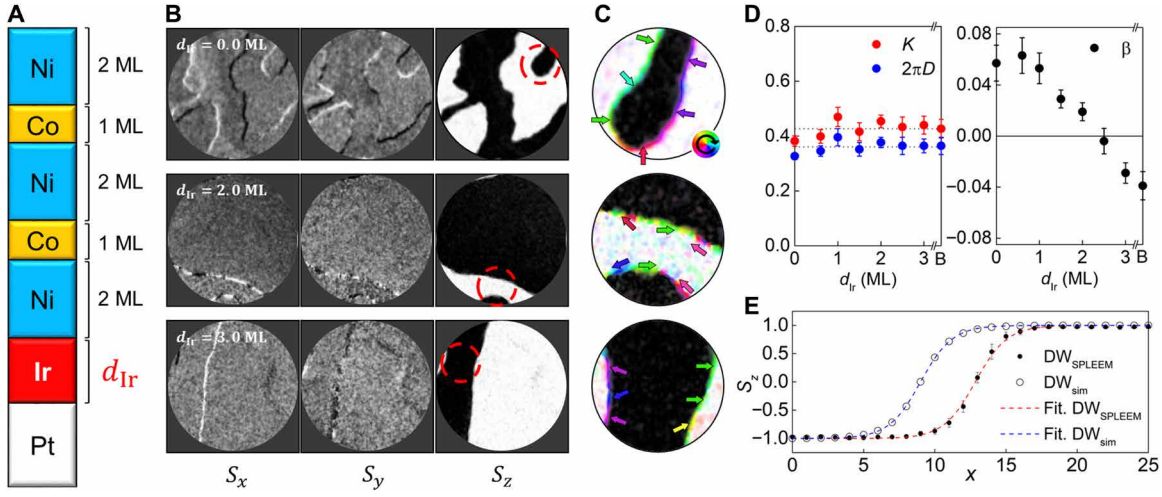

Fig. 4. Estimation results based on SPLEEM images. (**A**) Schematic diagram of the Ni/[Co/Ni]$_2$/Ir/Pt(111) system, where d_{Ir} is the thickness of iridium in the unit of monolayer (ML). (**B** and **C**) Raw (B) and postprocessed (C) images of observed magnetic domain and domain wall structures by SPLEEM. The field of view for (B) and (C) is 10 and 3 μm, respectively. White and black in the three columns in (B) correspond to magnetizations along $+S_x/-S_x$, $+S_y/-S_y$, and $-S_z/+S_z$, respectively. The color wheel represents the direction of in-plane magnetization in (C). (**D**) The K, D, and β parameter values estimated by the trained network using SPLEEM images with various Ir thicknesses. The dotted lines in (D) show the mean values for the estimated K and $2\pi D$. The letter B on the Ir thickness axis indicates the case of a bulk Ir crystal. (**E**) Domain wall (DW) profiles from a SPLEEM image and simulated spin configurations using the estimated parameters when d_{Ir} = 2.5 ML. The DW profiles are shifted so that they do not overlap. The dashed lines in (E) are fitting curves using the function Tanh($(x - c)/w$), where c and w are fitting parameters and x is for a lateral position.

To quantitatively evaluate whether the predicted parameters are proper values, which can be used to simulate the system shown in SPLEEM images, we compared the domain wall profiles from a SPLEEM image and a simulation result using the estimated parameters (Fig. 4E). From the fitting curves, we obtained the fitting parameter for domain wall length scale, w, to be 2.238 and 2.241 from the SPLEEM image and the simulation result, respectively. Since the size of single pixel in a SPLEEM image is around 25 nm (~10-μm field of view/400 pixels), each domain wall width can be approximated as 175.8 and 176.0 nm. This subpixel order difference of domain wall width indicates that the estimated parameters from the SPLEEM images properly simulates the system and suggests that our algorithm can be used to extract magnetic parameters from experimentally observed magnetic domain images.

The long-range dipolar energy constant of the system shown in Fig. 4A is calculated to be 0.171 meV from $\frac{(d_{Co}\mu_{Co} + d_{Ni}\mu_{Ni})^2}{a_{\parallel}{}^3}$, which is used to estimate ferromagnetic multilayered systems in previous studies (*3*, *35*–*37*), where $\mu_B = 9.274 \times 10^{-21}$ erg G^{-1}, $\mu_{Co} = 1.7\ \mu_B$, $\mu_{Ni} = 0.6\ \mu_B$, and $a_{\parallel} = 2.49$ Å. It is known that the DMI considered in this study directly competes with the dipolar interaction to form Néel domain walls in two-dimensional magnetic systems (*20*). Hence, we can convert the range of the estimated DMI parameters to units of energy using a simple relation, $D_L \times \frac{\beta}{D}$, where $D_L = 0.171$ meV. The β written in energy unit when $d_{Ir} = 0$ is 0.187 meV without Ir, and it changes to−0.115 meV, as Ir thickness increases. This value has

similar magnitude with those shown in the previous study (*35*), except when Ir is thin. It is expected that the estimation accuracy will be enhanced if we use a larger dataset with a variety of domain characteristics.

DISCUSSION

We applied a deep learning algorithm to estimate magnetic Hamiltonian parameters from magnetic domain images. By varying magnetic Hamiltonian parameter values, ~80,000 spin configurations were generated through a simulated annealing process by a Monte Carlo method. The effects of noise, which is unavoidable in experimental data, were considered in the generated dataset. The noise-injected dataset was used to train, validate, and test several neural network structures. It was found that deep neural networks are necessary to have sufficient resistance to noise. The error values of the magnetic parameters estimated through the trained network are analyzed using heatmap analysis and PCA to evaluate the performance of the trained network and to investigate whether there are specific relations between the error values. From the analysis, we found that the network learned that the accurate estimation of the effective anisotropy is a key factor to determine the magnetic Hamiltonian parameters from the out-of-plane magnetic domain structures formed in a two-dimensional magnetic system.

We applied the trained network to estimate magnetic Hamiltonian parameters from magnetic domain images observed by SPLEEM and showed that the behaviors of estimated parameters are consistent with the published results in the previous studies. This test demonstrates the effectiveness of applying the deep learning method to determine the strengths of magnetic interactions from experimentally observed images.

METHODS

Details of dataset generation

A large dataset of various spin configurations was generated showing the magnetic domain structures and the corresponding magnetic Hamiltonian parameters. A simulated annealing process using a Monte Carlo method with the magnetic Hamiltonian in Eq. 1 was performed to obtain the spin configurations in a 100 by 100 square grid lattice. We intentionally modulated the annealing rate of the simulation so that the resulting spin configurations do not necessarily represent the ground-state domain configurations. The ground-state spin configurations may be helpful to train the network to extract the characteristics of domain structures, but it can cause an overfitting problem on some limited structural features of magnetic domains in the ground state; i.e., when only using ground states, the network cannot learn the structural diversity of the magnetic domains in local energy minimum states (*38*).

We fixed $J = 1$, and the other parameters were scaled to their ratio with J. Other Hamiltonian parameters were varied as $\beta = -3D \sim 3D$ and $K = 0.1K_{shape} \sim 1.2K_{shape}$, where $K_{shape} = 2\pi D$ and $D = 0.05 \sim 0.13$ to generate a diverse set of magnetic domains on the 100 by 100 grid system. These parameter variation ranges are based on previous studies (*17*, *18*, *20*). The step sizes used to vary each of the β, K, and D parameters were $0.3D$, $0.02K_{shape}$, and 0.005, respectively. As shown in Fig. 3A, the out-of-plane domains are formed when the $\beta = 0$ and the K is stronger than the dipole shape anisotropy K_{shape}, i.e., $K_N > 0$. This result is consistent with the previous studies (*17*, *20*), showing the formation of magnetic stripe domains under the $K > K_{shape}$ conditions. The tendency that the stripe domain width is proportional to J/D was also found in theoretical and numerical calculations of previous studies (*17*). In the presence of DMI, as shown in Fig. 3E, domain width is reduced compared with Fig. 3A, which agrees with previous reports (*18*, *20*). Therefore, within the parameter space used to generate the spin configurations in this study, it is confirmed that the proper magnetic domains were created by the simulated annealing process.

We generated 20 spin configurations for each magnetic parameter set and divided the spin configurations into 14, 4, and 2 spin configurations each for training, validation, and testing datasets, respectively. Since the magnetic domain shape is determined by spontaneous symmetry breaking during the simulated annealing process, each of 20 spin configurations that are generated using the same magnetic parameter set shows different magnetic domain shapes as shown in fig. S1. This means that the total dataset contains essentially no duplicated data. The total number of generated spin configurations was about 80,000.

Noise injection on each dataset

We injected arbitrary noises in spin configurations in our training and validation dataset using $\hat{S}^*(x,y)_i = L2_{\mathrm{Norm.}}(\hat{S}(x,y)_i + \varepsilon_i \hat{R}(x,y)_i)$, where i is for an index of a spin configuration in each dataset, $\hat{S}^*(x,y)$ and $\hat{S}(x,y)$ are noisy and noiseless spin configurations, respectively, ε_i is a randomly chosen noise amplitude between 0 and 1, $\hat{R}(x,y)$ is a unit vector map composed of arbitrary directions, and $L2_{\mathrm{Norm.}}$ indicates the normalization process to make unit vectors. The sampled spin configurations from a noisy dataset are shown in fig. S2A.

For the test dataset, we injected stepped increased noises on them to investigate the robustness of trained networks with various noise amplitudes. The same relation mentioned above is used except that the ε_i is replaced by $\varepsilon_{\mathrm{Ts}}$ increased with 0.1 step size. An example of a noisy test dataset is shown in fig. S2B.

Training process

The cost that should be minimized to train our networks is the mean squared error value between the estimated magnetic Hamiltonian parameters from our networks and true parameters, which were used to generate the input spin configurations. In this process, to avoid the results from the biased training caused by the scale differences among the parameters, the normalized true parameters, β_N, K_N, and D_N, are used to calculate the cost value.

One epoch is defined as one complete run of the training process using the entire training dataset. A total of 1000 epochs were run. A validation dataset was used to monitor the training process at the end of each epoch. We chose the trained neural network model with the lowest validation cost value for the testing process after the whole training process. We shuffled the entire training dataset randomly, and the mini-batches of the dataset were used as the input data with 200 batch size. Adam optimizer (*39*) is used to train the network with a 10^{-4} learning rate.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/39/eabb0872/DC1

REFERENCES AND NOTES

1. M. Kronseder, T. N. G. Meier, M. Zimmermann, M. Buchner, M. Vogel, C. H. Back, Real-time observation of domain fluctuations in a two-dimensional magnetic model system. *Nat. Commun.* **6**, 6832 (2015).

2. M. Yamanouchi, S. Ikeda, F. Matsukura, H. Ohno, A. Jander, P. Dhagat, A. Jander, P. Dhagat, S. Ikeda, F. Matsukura, H. Ohno, Domain structure in CoFeB thin films with perpendicular magnetic anisotropy. *IEEE Magn. Lett.* **2**, 3000304 (2011).
3. C. Won, Y. Z. Wu, J. Choi, W. Kim, A. Scholl, A. Doran, T. Owens, J. Wu, X. F. Jin, H. W. Zhao, Z. Q. Qiu, Magnetic stripe melting at the spin reorientation transition in Fe/Ni/Cu(001). *Phys. Rev. B* **71**, 224429 (2005).
4. O. Hellwig, T. L. Kirk, J. B. Kortright, A. Berger, E. E. Fullerton, A new phase diagram for layered antiferromagnetic films. *Nat. Mater.* **2**, 112–116 (2003).
5. J. E. Davies, O. Hellwig, E. E. Fullerton, G. Denbeaux, J. B. Kortright, K. Liu, Magnetization reversal of Co/Pt multilayers: Microscopic origin of high-field magnetic irreversibility. *Phys. Rev. B* **70**, 224434 (2004).
6. X. Z. Yu, Y. Onose, N. Kanazawa, J. H. Park, J. H. Han, Y. Matsui, N. Nagaosa, Y. Tokura, Real-space observation of a two-dimensional skyrmion crystal. *Nature* **465**, 901–904 (2010).
7. A. Tonomura, X. Yu, K. Yanagisawa, T. Matsuda, Y. Onose, N. Kanazawa, H. S. Park, Y. Tokura, Real-space observation of skyrmion lattice in helimagnet MnSi thin samples. *Nano Lett.* **12**, 1673–1677 (2012).
8. W. Jiang, G. Chen, K. Liu, J. Zang, S. G. E. te Velthuis, A. Hoffmann, Skyrmions in magnetic multilayers. *Phys. Rep.* **704**, 1–49 (2017).
9. S. Woo, K. Litzius, B. Krüger, M. Y. Im, L. Caretta, K. Richter, M. Mann, A. Krone, R. M. Reeve, M. Weigand, P. Agrawal, I. Lemesh, M. A. Mawass, P. Fischer, M. Kläui, G. S. D. Beach, Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets. *Nat. Mater.* **15**, 501–506 (2016).
10. G. Chen, A. Mascaraque, A. T. N'Diaye, A. K. Schmid, Room temperature skyrmion ground state stabilized through interlayer exchange coupling. *Appl. Phys. Lett.* **106**, 242404 (2015).
11. S. Woo, K. M. Song, H.-S. Han, M.-S. Jung, M.-Y. Im, K.-S. Lee, K. S. Song, P. Fischer, J.-I. Hong, J. W. Choi, B.-C. Min, H. C. Koo, J. Chang, Spin-orbit torque-driven skyrmion dynamics revealed by time-resolved x-ray microscopy. *Nat. Commun.* **8**, 15573 (2017).
12. N. Romming, C. Hanneken, M. Menzel, J. E. Bickel, B. Wolter, K. von Bergmann, A. Kubetzka, R. Wiesendanger, Writing and deleting single magnetic skyrmions. *Science* **341**, 636–639 (2013).
13. K.-W. Moon, J. Yoon, J. W. Choi, C. Kim, D.-O. Kim, D. Kim, B. S. Chun, B.-C. Min, C. Hwang, Measuring the magnetization from the image of the stripe magnetic domain. *Phys. Rev. Appl.* **12**, 034030 (2019).
14. N. Saratz, U. Ramsperger, A. Vindigni, D. Pescia, Irreversibility, reversibility, and thermal equilibrium in domain patterns of Fe films with perpendicular magnetization. *Phys. Rev. B* **82**, 184416 (2010).
15. A. B. Kashuba, V. L. Pokrovsky, Stripe domain structures in a thin ferromagnetic film. *Phys. Rev. B* **48**, 10335–10344 (1993).
16. D. Clarke, O. A. Tretiakov, O. Tchernyshyov, Stripes in thin ferromagnetic films with out-of-plane anisotropy. *Phys. Rev. B* **75**, 174433 (2007).
17. H. Y. Kwon, S. S. Hong, J. H. Seok, K. M. Bu, Y. Z. Wu, Z. Q. Qiu, C. Won, A study of the stripe domain phase at the spin reorientation transition of two-dimensional magnetic system. *J. Magn. Magn. Mater.* **322**, 2742–2748 (2010).
18. H. Y. Kwon, K. M. Bu, Y. Z. Wu, C. Won, Effect of anisotropy and dipole interaction on long-range order magnetic structures generated by Dzyaloshinskii-Moriya interaction. *J. Magn. Magn. Mater.* **324**, 2171–2176 (2012).
19. T. H. Johansen, A. V. Pan, Y. M. Galperin, Exact asymptotic behavior of magnetic stripe domain arrays. *Phys. Rev. B* **87**, 060402 (2013).
20. H. Y. Kwon, C. Won, Effects of Dzyaloshinskii-Moriya interaction on magnetic stripe domains. *J. Magn. Magn. Mater.* **351**, 8–15 (2014).
21. I. Lemesh, F. Büttner, G. S. D. Beach, Accurate model of the stripe domain phase of perpendicularly magnetized multilayers. *Phys. Rev. B* **95**, 174423 (2017).
22. J. Schmidhuber, Deep learning in neural networks: An overview. *Neural Netw.* **61**, 85–117 (2015).
23. D. Pfau, J. S. Spencer, A. G. de G. Matthews, W. M. C. Foulkes, Ab-Initio solution of the many-electron Schrödinger equation with deep neural networks. arXiv:1909.02487 [physics.chem-ph]. 5 September 2019.
24. A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D. T. Jones, D. Silver, K. Kavukcuoglu, D. Hassabis, Protein structure prediction using multiple deep neural networks in CASP13. *Proteins* **87**, 1141–1148 (2019).
25. G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
26. Z. Cai, J. Liu, Approximating quantum many-body wave functions using artificial neural networks. *Phys. Rev. B* **97**, 035116 (2018).
27. H. Y. Kwon, N. J. Kim, C. K. Lee, C. Won, Searching magnetic states using an unsupervised machine learning algorithm with the Heisenberg model. *Phys. Rev. B* **99**, 024423 (2019).
28. V. K. Singh, J. H. Han, Application of machine learning to two-dimensional Dzyaloshinskii-Moriya ferromagnets. *Phys. Rev. B* **99**, 174426 (2019).
29. A. T. N'diaye, A. Quesada, Spin-polarized low-energy electron microscopy (SPLEEM), in *Characterization of Materials* (John Wiley & Sons Inc., ed. 2, 2012), pp. 1827–1840.
30. I. Dzyaloshinskii, A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics. *J. Phys. Chem. Solids* **4**, 241–255 (1958).
31. T. Moriya, New mechanism of anisotropic superexchange interaction. *Phys. Rev. Lett.* **4**, 228–230 (1960).
32. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CVPR, 2016), pp. 770–778.
33. D. M. Hawkins, The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **44**, 1–12 (2004).
34. I. T. Jolliffe, *Principal Component Analysis* (Springer-Verlag, ed. 2, 2002).
35. G. Chen, T. Ma, A. T. N'Diaye, H. Kwon, C. Won, Y. Wu, A. K. Schmid, Tailoring the chirality of magnetic domain walls by interface engineering. *Nat. Commun.* **4**, 2671 (2013).
36. Y. Z. Wu, C. Won, A. Scholl, A. Doran, H. W. Zhao, X. F. Jin, Z. Q. Qiu, Magnetic stripe domains in coupled magnetic sandwiches. *Phys. Rev. Lett.* **93**, 117205 (2004).
37. G. Chen, J. Zhu, A. Quesada, J. Li, A. T. N'Diaye, Y. Huo, T. P. Ma, Y. Chen, H. Y. Kwon, C. Won, Z. Q. Qiu, A. K. Schmid, Y. Z. Wu, Novel chiral magnetic domain wall structure in Fe/Ni/Cu(001) Films. *Phys. Rev. Lett.* **110**, 177204 (2013).
38. X. Guo, Y. Yin, C. Dong, G. Yang, G. Zhou, On the class imbalance problem, in *Fourth International Conference on Natural Computation* (IEEE, 2008), pp. 192–201.
39. D. P. Kingma, J. L. Ba, Adam: A method for stochastic optimization, in *3rd International Conference on Learning Representations* (ICLR, 2015), pp. 1–15.

Acknowledgments

Funding: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1A6A3A01091209); by the NRF funded by the Korean Government (NRF-2018R1D1A1B07047114 and NRF-2020R1A5A1104591); by the U.S. National Science Foundation (DMR-1610060 and DMR-1905468); by the UC Office of the President Multicampus Research Programs and Initiatives (MRP-17-454963); by the Office of Science, Office of Basic Energy Sciences of the U.S. Department of Energy under contract no. DE-AC02-05CH11231; by the SMART Center (2018-NE-2861), one of seven centers of nCORE, a Semiconductor Research Corporation program, sponsored by National Institute of Standards and Technology (NIST); by the National Science Foundation of China (nos. 11734006 and 11974079); by the National Research Council of Science and Technology (no. CAP-16-01-KIST); and by the KIST Institutional Program. **Author contributions:** H.Y.K. created the design, developed the methodology, and performed network training. C.L. and H.G.Y. carried out data analysis and prepared Figs. 1 to 3. G.C., K.L., and A.K.S. provided experimental image data and participated in the discussion of the results. J.W.C., Y.Z.W., and C.W. supervised the work progresses and helped with the discussion of the results. All authors contributed to the writing of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 29 January 2020
Accepted 10 August 2020
Published 25 September 2020
10.1126/sciadv.abb0872

Citation: H. Y. Kwon, H. G. Yoon, C. Lee, G. Chen, K. Liu, A. K. Schmid, Y. Z. Wu, J. W. Choi, C. Won, Magnetic Hamiltonian parameter estimation using deep learning techniques. *Sci. Adv.* **6**, eabb0872 (2020).