

Prediction of material parameters of spin textures from LTEM images by machine learning models CNN and ViT

Bohan Li(李博涵), Xinyuan Zhang(张歆媛), Guanhua Chen(陈冠桦), and Lin Chen(陈琳)
Citation: Chin. Phys. B, 2026, 35 (**7**): 077507. DOI: 10.1088/1674-1056/ae5db1
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Hall anomalies in the centrosymmetric triangular-lattice antiferromagnet $GdGa_2$

Sidi Wang(王思迪), Jiyuan Li(李纪源), Yuhao Wang(王宇豪), Keqi Xia(夏克奇), Jing Meng(孟婧), Bocheng Yu(余博丞), Yiqian Hu(胡艺骞), Zheng Li(李峥), Hui Zhang(张慧), Jingzhong Luo(罗晶中), Dongmei Jiang(蒋冬梅), Qingfeng Zhan(詹清峰), Tian Shang(商恬), and Yang Xu(徐杨)
Chin. Phys. B, 2026, 35 (**2**): 027502. DOI: 10.1088/1674-1056/ae2672

Stability and characteristic modes of skyrmions in magnetic nanotubes

Tijjani Abdulrazak, Qizhi Cai(蔡淇智), and Guangwei Deng(邓光伟)
Chin. Phys. B, 2025, 34 (**10**): 107512. DOI: 10.1088/1674-1056/ade071

Coupling of magnon modes in nanodisk with spin texture

Zijie Zhou(周子杰), Junning Zhao(赵俊宁), Xinhui Ma(马心慧), Rong Wang(王熔), and Fusheng Ma(马付胜)
Chin. Phys. B, 2025, 34 (**10**): 107506. DOI: 10.1088/1674-1056/ade426

Magnonic band-pass and band-stop filters with structurally modulated waveguides

Lai-He Feng(冯来和), Mang-Yuan Ma(马莽原), Zhi-Hua Liu(刘智华), Kai-Le Xie(解凯乐), and Fu-Sheng Ma(马付胜)
Chin. Phys. B, 2023, 32 (**6**): 067503. DOI: 10.1088/1674-1056/ac8afb

Magnetic triangular bubble lattices in bismuth-doped yttrium iron garnet

Tao Lin(蔺涛), Chengxiang Wang(王承祥), Zhiyong Qiu(邱志勇), Chao Chen(陈超), Tao Xing(邢弢), Lu Sun(孙璐), Jianhui Liang(梁建辉), Yizheng Wu(吴义政), Zhong Shi(时钟), and Na Lei(雷娜)
Chin. Phys. B, 2023, 32 (**2**): 027505. DOI: 10.1088/1674-1056/aca604

Prediction of material parameters of spin textures from LTEM images by machine learning models CNN and ViT

Bohan Li(李博涵)[1], Xinyuan Zhang(张歆媛)[2], Guanhua Chen(陈冠桦)[2], and Lin Chen(陈琳)[2,†]

[1] *School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing* 210003, *China*
[2] *College of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications, Nanjing* 210023, *China*

(Received 23 December 2025; revised manuscript received 24 March 2026; accepted manuscript online 10 April 2026)

Machine learning (ML) models were trained to predict the Dzyaloshinskii–Moriya interaction constant D and anisotropy constant K of spin textures from Lorentz transmission electron microscopy (LTEM) images. Two ML models, convolutional neural network (CNN) and vision transformer (ViT), were trained, tested and employed to predict the values of D and K. Firstly, training and testing datasets composed of 9300 topological spin textures were prepared using the micromagnetic simulation method with the values of D and K randomly chosen. Secondly, the performance of the CNN model for predicting D and K values was investigated by varying the number of training data, the pooling process, and the number of fully connected layers, which indicated high prediction accuracies. Thirdly, better performance of the ViT model for predicting D and K values was achieved with the coefficient of determination R^2 reaching as high as 0.9987 and 0.9991 for predicting D and K values, respectively. Particularly, the reliance on synthetic data and noise robustness was discussed. Finally, the prediction abilities of CNN and ViT models were evaluated and compared. The results of this research indicate that the ML models can achieve the material parameters directly and effectively from LTEM images, and this process may contribute to the design of advanced devices based on the topological spin textures.

Keywords: spin texture, LTEM, machine learning, ViT model

PACS: 75.25.–j, 75.78.Cd, 87.19.lv, 07.05.Tp

DOI: 10.1088/1674-1056/ae5db1 **CSTR:** 32038.14.CPB.ae5db1

1. Introduction

Utilizing both the charge and the intrinsic spin of the electron, spintronic devices present superior performances such as low power consumption, non-volatility, high speed, *etc.*[1,2] Recently, spintronic devices based on topological spin textures have been attracting more and more attentions, which rely on creating, moving, and detecting nanoscale spin textures.[3,4] Therefore, the morphologies or the magnetic distributions of the spin textures, which are determined by material parameters such as Dzyaloshinskii–Moriya interaction (DMI) constant D and anisotropy constant K, are significant for functionality designs.[5,6] As an antisymmetric interaction, DMI exists widely in non-centrosymmetric magnetic bulk compounds and thin layers with the breaking of inversion symmetry, which favors non-collinear alignment of neighboring spins.[7] Kinds of topologically spin textures can be realized through competitions between DMI, Heisenberg exchange interactions, anisotropy strength and other contributing energies. Therefore, it is important to manipulate the parameter values to realize designed spin texture configurations through experimental material synthesis methods. At meantime, it is also very necessary to obtain the parameter values of the designed spin textures through experimental characterization methods. For example, the K value can be determined through measuring the M–H loops by SQUID, or measuring the resonance field along different crystallographic directions,[8,9] and the D value can be determined by Brillouin light scattering method or propagating spin-wave spectroscopy,[10] which are more complicated.

Particularly, the magnetic distributions of the spin textures are related to the material parameters, which indicates that the parameters should be obtained by analyzing the morphologies using techniques such as Lorentz transmission electron microscopy (LTEM)[11–13] or magnetic force microscopy (MFM).[14–16] So, it is suggested that the well-trained machine learning (ML) models can be employed to analyze the morphologies and to predict the material parameters.[17–19] In recent years, the rapid development of artificial intelligence, particularly neural networks, has introduced a powerful data-driven approach to solving complex physical problems.[20–23] Neural networks, with their ability to automatically extract intricate mappings between input features and output results through extensive data learning, offer a promising alternative for fast and accurate predictions.[24–27] It is reported that Mehmood *et al.* used convolutional neural network (CNN) models to predict the exchange constant and saturation magnetization and identify the vortex helicity from MFM images.[28] McCray *et al.* proposed an AI-enabled phase-retrieval method for LTEM which allows quantitative phase reconstruction of in-situ data and can also enable near real-time quantitative

†Corresponding author. E-mail: cl@njupt.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

magnetic imaging.[29,30]

Compared with MFM, LTEM offers higher spatial resolution by calculating the phase shift of the transmission electron beam.[11] In this work, we employed CNN models and vision transformer (ViT) models to predict the material parameters D and K by analyzing the LTEM images. Before model optimization, training and testing datasets must be prepared, which are LTEM images with known values of D and K. Considering that it is complicated to get large numbers of LTEM images with varied values of D and K experimentally, the micromagnetic simulation method was employed to generate abundant LTEM images with randomly chosen D and K. Subsequently, the CNN and ViT models were trained and optimized. Finally, the prediction abilities of the two ML models were evaluated and compared.

2. Model and method

Figure 1 presents the framework of this work. In this study, we construct and optimize a machine-learning CNN model and a machine-learning ViT model in order to predict the values of the parameters D and K from simulated LTEM images.

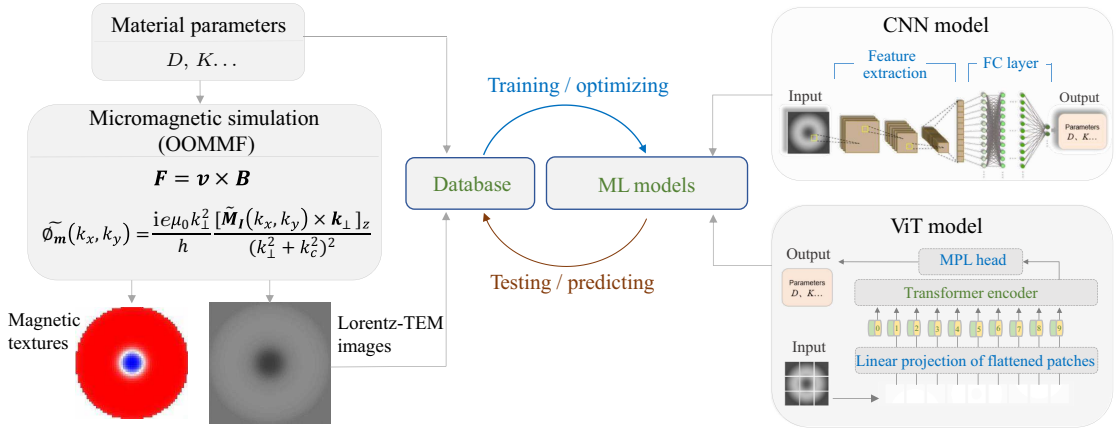


Fig. 1. The framework of this work including generating the datasets by OOMMF, training and optimizing the ML models such as CNN and ViT, and evaluating the prediction abilities of the models.

Firstly, training datasets and testing datasets were generated by micromagnetic simulation software Ubermag with embedded OOMMF codes before training the ML models.[31] Experimentally, LTEM can be performed to image the topologically distributed spin textures. In LTEM, a parallel electron beam with a typical energy of 200 kV–300 kV is incident on the magnetic sample. The charged electrons are deflected by the Lorentz force caused by the magnetic field in the sample. This causes a phase shift $\tilde{\phi}_m(k_x,k_y)$ in the electron beam compared with a beam travelling through vacuum. Here, the Ubermag micromagnetic software is used to construct magnetic samples with different topological spin textures, and based on this, to simulate the phase shift that would be recovered by electron holography. The ground static state of the topological spin textures can be realized by calculating the minimum value of the total energy of the magnetic system $E_{\rm tot}$. $E_{\rm tot}$ can be given by

$$E_{\rm tot} = E_{\rm ex} + E_{\rm K} + E_{\rm zeeman} + E_{\rm DM} + E_{\rm d}, \tag{1}$$

which includes $E_{\rm ex}$ for exchange energy, $E_{\rm K}$ for anisotropy energy, $E_{\rm zeeman}$ for Zeeman energy, $E_{\rm DM}$ for DMI energy, and $E_{\rm d}$ for demagnetization energy. The magnetic phase shift $\tilde{\phi}_m(k_x,k_y)$ of the electron beam can be calculated by[32]

$$\tilde{\phi}_m(k_x,k_y) = \frac{{\rm i}e\mu_0 k_\perp^2}{h}\frac{[\tilde{\boldsymbol{M}}_I(k_x,k_y)\times\boldsymbol{k}_\perp]_z}{(k_\perp^2+k_{\rm c}^2)^2}, \tag{2}$$

where μ_0 is the vacuum permeability, $\boldsymbol{k}$ is the k-vector in Fourier space and $k_\perp$ is defined as

$$k_\perp = (k_x^2+k_y^2)^{1/2}. \tag{3}$$

Specifically, a Tikhonov filter is used to avoid singularities in Fourier space, and $k_{\rm c}$ is the radius of the filter. $\tilde{\boldsymbol{M}}_I(k_x,k_y)$ is the integrated magnetization along the path of the electron beam. Based on the above theory, several software programs are developed to simulate the LTEM images from key parameters such as PyLorentz,[33] Genius,[34] and MALTS.[35] In this work, the mag2exp module in Ubermag was used to generate the LTEM images. The simulation method applied in mag2exp shares the same underlying physical principles with other simulation tools which have been experimentally verified in the field. Due to the fact that mag2exp takes into account a more detailed magnetic moment distribution, the simulated LTEM images would have higher resolution and reliability.[36]

The topological spin textures were realized in an FeGe magnetic disc with a radius of 50 nm and a thickness of 2 nm. In the simulations, the typical parameters in the FeGe are as follows:[37–40] exchange stiffness constant $A_{\rm ex} = 0.4\times10^{-11}$ J/m, saturation magnetization $M_{\rm s} = 3.84\times10^5$ A/m. The external applied static magnetic field H is 0.5×10^5 A/m. The ranges of parameters D (from 0 to 5×10^{-3} J/m^2) and K (from 0 to 10×10^5 J/m^3) used in the simulations are based on

experimentally reported values, but appropriately extended beyond current experimental limits to cover the physically meaningful interval where stable topological spin textures can be generated. The simulated mesh cell is 2 nm×2 nm×2 nm. 9300 different topological spin textures were constructed with the values of D and K randomly chosen. Then, the corresponding LTEM images were collected, which were divided into training datasets with 8500 LTEM images and testing datasets with 800 LTEM images.

In order to ensure that the model can learn the inherent physical correlation between parameters (D, K) and magnetic properties rather than memorizing specific sample features, we constructed the dataset and split the training/testing sets following strict physical and statistical principles. A full set of samples covering the physically reasonable range of D and K was generated. All D and K values were generated randomly within this range, and the corresponding spin texture data (magnetization distribution images and LTEM images) were obtained through micromagnetic simulations without any artificial screening or selection of grid points, ensuring the full dataset has sufficient physical variance. The training data and testing data were chosen randomly from the full set of samples, to ensure that both the training and testing sets uniformly cover the entire distribution range of D and K. In particular, no testing data were accessed or used in any stage of the model training, which completely avoids data leakage in the training process.

Secondly, two ML models, CNN and ViT, were chosen to be trained, optimized, and tested based on the datasets. It is known that the ViT model significantly outperforms traditional CNN model by effectively capturing long-range dependencies through its multi-head self-attention mechanism, thereby overcoming the local receptive field limitations inherent in CNNs.[41–44] This capability is crucial for accurately predicting the multiple parameters from complex LTEM images. As shown in Fig. 1, CNN model employs a hierarchy of convolutional layers with learnable filters to perform automatic feature extraction and translation-invariant classification by progressively aggregating simpler local patterns into more complex and global representations.[41,42] ViT model applies the transformer architecture to compute vision tasks, the core idea of which is to segment the image into patches of fixed size and input them as a sequence into the transformer encoder for processing. The complete ViT model is composed of patch embedding, position embedding, transformer encode, and MLP head.[43,44] Finally, the prediction abilities of the two ML models were compared and evaluated. During evaluating the training effect and the prediction accuracy, the coefficient of determination R^2 and root-mean-square error (RMSE) were introduced to comprehensively assess model performances, which can be calculated by[45,46]

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(D_{\text{Predicted},i} - D_{\text{Actual},i}\right)^2}{\sum_{i=1}^{N} \left(D_{\text{Actual},i} - \overline{D}_{\text{Actual},i}\right)^2}, \tag{4}$$

$$\text{RMSE} = \left(\frac{1}{N}\sum_{i=1}^{N}\left(D_{\text{Predicted},i} - D_{\text{Actual},i}\right)^2\right)^{-2}. \tag{5}$$

3. Results and discussion

3.1. Formation of the topological spin textures

Before training and optimizing the ML models, we studied the stability of the topological spin textures, which composed the training and testing datasets. As shown in Fig. 2(a), the K–D phase diagram presents three different magnetic states, which are ferromagnetic (FM) state, skyrmion state and maze state. Figures 2(b)–2(e) present the spin distributions of the typical topological spin textures from the 9300 simulated images, and figures 2(f)–2(i) present the corresponding simulated LTEM images. From Fig. 2(a) we can see that, larger D is favor to realize maze state, larger K is favor to realize FM state. Especially, skyrmion state can be achieved by manipulating the D and K values precisely.

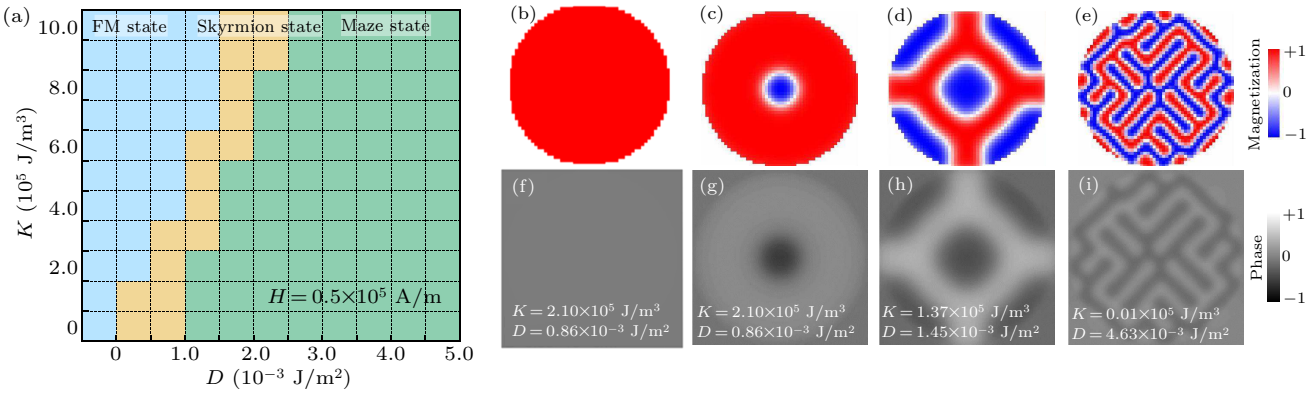


Fig. 2. (a) K–D phase diagram of the topological spin textures with $H = 0.5 \times 10^5$ A/m. (b)–(e) Magnetization distributions of the typical topological spin textures and (f)–(i) corresponding simulated LTEM images.

As shown in Fig. 2(b), FM state presents uniform magnetization distributions, while as shown in Fig. 2(c), a radially symmetrical skyrmion is realized with the spin magnetizations rotating 180° from the core to its boundary. Figures 2(d) and 2(e) show the maze states of the topological spin textures, which present labyrinthine patterns or stripe networks. Meanwhile, the corresponding simulated LTEM images in Figs. 2(g), 2(h), and 2(i) also present the identical phase shift characteristics to the magnetization distributions. Based on the K–D phase diagram, 9300 spin textures are obtained with the value of D randomly chosen in the range from 0 to 5×10^{-3} J/m^2, and K randomly chosen in the range from 0 to 10×10^5 J/m^3. To further demonstrate the phase distributions of the spin textures, we also calculated the K–D phase diagrams when $H = 0$ and $H = 1.0 \times 10^5$ A/m, which can be seen in Fig. S1 in supplementary materials.

3.2. Performance of CNN model for predicting D and K values

Firstly, the CNN model was trained, optimized, and used to predict the values of D and K. Before training the model, the training datasets were standardized to ensure the consistency of the data and the stability of the model. The images were resized to 224×224 pixels to fit the input requirements of the pre-trained model. Average pooling was used in the feature extraction process, and the number of fully connected layers is 4. During training, the MSE was utilized as the loss function, coupled with the Adam optimizer for parameter optimization. The batch size is 8 and the learning rate is 0.001. In each training batch, the model performed forward propagation to generate predictions, subsequently evaluated the performance based on the MSE loss between predicted and actual values, and updated parameters via backpropagation. The entire training process spanned 200 epochs. After training, 200 testing data were used to evaluate the prediction accuracy. Fig. S2 presents the distributions of the training data and testing data, which indicate a uniform and random spatial distribution within the physically reasonable range, with no obvious sample clustering, blank areas, or grid-like aggregation.

Figures 3(a) and 3(b) present the performance of the CNN model on predicting the values of D and K with 2000 training data and testing data. We can see that the model achieves high accuracy in predicting D and K values. The coefficient of determination R^2 can be as high as 0.989 on the training data and 0.988 on the testing data for predicting D values, while 0.987 on the training data and 0.977 on the testing data for predicting K values. In order to evaluate the influence of dataset sizes on the training and predicting effects, three other groups with 500, 1000, and 5000 training data are used to optimize the CNN model, the performances of which can be seen in Fig. S3 in supplementary materials.

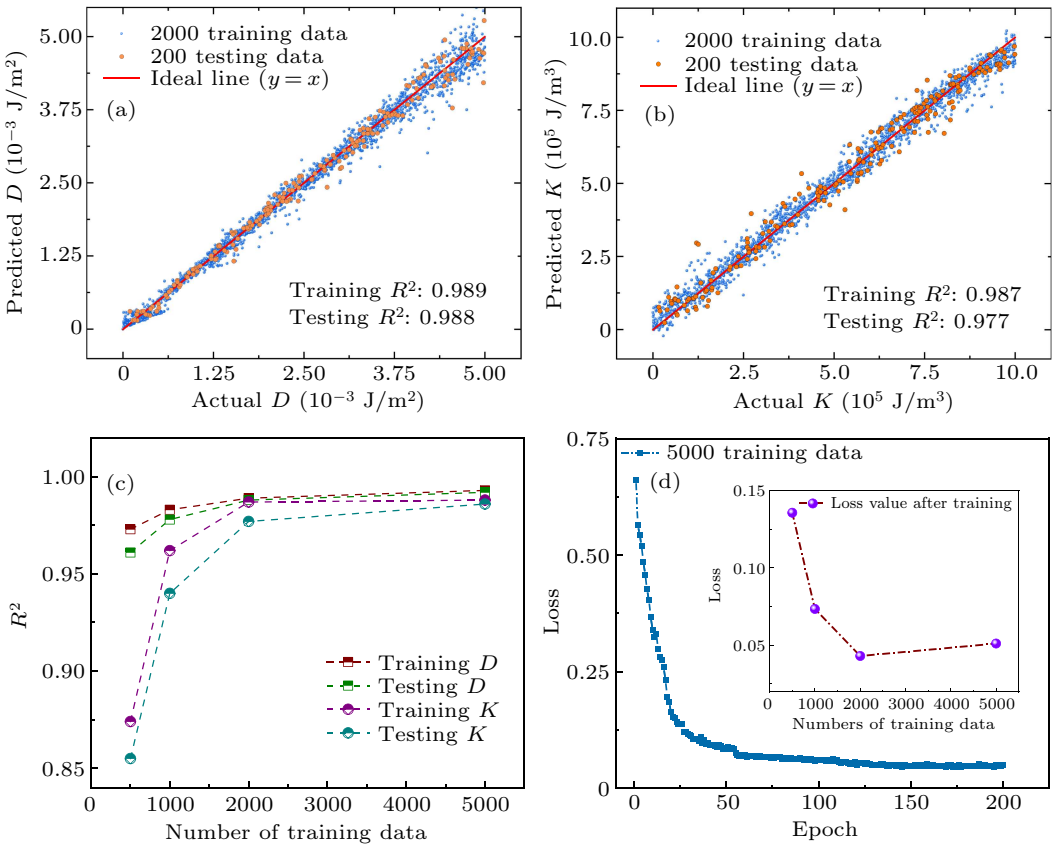


Fig. 3. Performance of CNN model to predict the values of (a) D and (b) K. Different colors represent data in training (blue) and testing (orange) datasets. (c) The value of R^2 with increasing the numbers of training data from 500 to 5000. (d) The loss convergence over training epochs for 2000 training data. The inset shows the final loss value after 200 epochs with increasing the numbers of training data from 500 to 5000.

Figure 3(c) presents the values of R^2 with increasing training data from 500 to 5000. It can be seen that the values of R^2 indicate obviously worse performance on predicting K values compared with predicting D values based on 500 training data. With training data increasing from 500 to 2000, the performance on predicting both D and K values improves significantly, especially for predicting K values. Then, the values of R^2 increase slightly when the training data increases from 2000 to 5000, which indicates slight improvements in predictive ability. Meanwhile, we examine the values of the loss function with different training dataset sizes. Figure 3(d) presents the loss convergence over training epochs for 2000 training data and the inset presents the loss value at the 200th epoch with increasing training data. We can see that the loss value at the 200th epoch keeps dropping with the training data increasing from 500 to 2000, and then increases slightly with the training data increasing to 5000, which also indicates that a dataset of 2,000 is already adequate to achieve a good prediction performance.

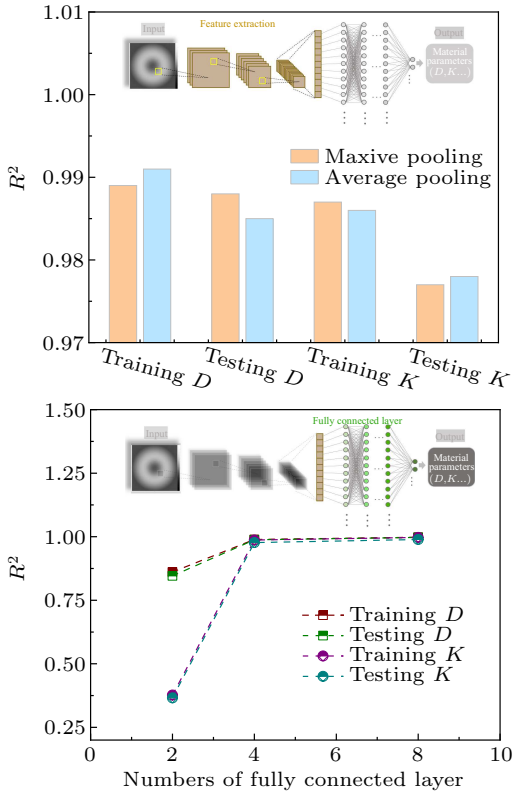


Fig. 4. (a) The values of R^2 for predicting D and K values using max pooling and average pooling. (b) The values of R^2 for predicting D and K values with the number of fully connected layer increasing from 2 to 8.

Further, we tried to optimize the CNN model in order to get better performance. Here, we compare the prediction abilities using max pooling and average pooling which are two commonly chosen methods in the feature extraction process.[46] Figure 4(a) presents the values of R^2 for predicting D and K values with 2000 training data and 200 testing data. Considering the main purpose of pooling is spatial invariance, dimensionality reduction and feature emphasis, it can be seen that the R^2 values using max pooling are smaller than the R^2 values using average pooling on training data for predicting D values and on testing data for predicting K values. Meanwhile, the R^2 values using max pooling are larger than the R^2 values using average pooling on training data for predicting K values and on testing data for predicting D values. The results show that the two pooling strategies yield similar overall prediction accuracy and exhibit no significant difference in model performance for the magnetic texture parameter identification task.

In addition, we investigated the performance of the CNN model with different numbers of fully connected layers. Figure 4(b) presents the R^2 for predicting D and K values when the number of fully connected layers is 2, 4, and 8. The training data and testing data are still 2000 and 200, respectively. We can see that the R^2 values are lower than 0.5 for predicting K values and lower than 0.9 for predicting D values with 2 layers. Then, the model performs better with 4 layers and 8 layers. Particularly, a slight improvement is achieved when the layer number increases from 4 to 8, which indicates a saturation in the number of layers.

3.3. Performance of ViT model for predicting D and K values

Secondly, the ViT model was trained, optimized, and used to predict the values of D and K. Before model training, data preprocessing was performed which was the same as that for the CNN model. After completing the data preprocessing and loading, we employed a custom ML model based on the ViT-Base architecture. The patch size used for the input images was 16×16. During training, the MSE was utilized as the loss function, coupled with the Adam optimizer for parameter optimization. The batch size is 8 and the learning rate is 0.001. The entire training process spanned 200 epochs. The performance of the ViT model on predicting the values of D and K with 500 training data and 2000 training data are presented in Figs. 5(a), 5(b) and 5(c), 5(d), respectively. It can be seen that better performance is achieved in predicting D and K values than when using the CNN model. The R^2 values can reach as high as 0.9987 on the training data and 0.9879 on the testing data for predicting D values, while they are 0.9991 on the training data and 0.9880 on the testing data for predicting K values with 2000 training data points. Even when the number of training data points decreases to 500, the R^2 values are still higher than 0.99 on the training data and the testing data for predicting D values. Meanwhile, the R^2 value is higher than 0.99 on the training data for predicting K values, but achieves a smaller value of 0.9745 on the testing data. Particularly, it can be seen in Fig. 5(c) that the ViT model performs worse in prediction accuracy when the D values range from 0 to about

0.625×10^{-3} J/m^2, while it achieves excellent prediction accuracy when the D values range from 0.625×10^{-3} J/m^2 to 5.0×10^{-3} J/m^2. Meanwhile, the prediction abilities for D values are obviously better than those for K values, which also show good performance. In addition, two other groups with 1000 and 5000 training data points were used to optimize the CNN model, the performance of which can be seen in Fig. S4 in supplementary materials. Meanwhile, we investigated the performance of the ViT model with different batch sizes. Figure S5 presents the values of R^2 with the batch size increasing from 4 to 32 for predicting the values of D and K based on 2000 training data points and 200 testing data points. It is indicated that the variations in the R^2 values for predicting both D and K values based on the testing data are distinct, especially for the R^2 in predicting the value of K, with the values first increasing when the batch number changes from 4 to 8 and then decreasing when the batch number changes from 8 to 32. Particularly, the values of R^2 corresponding to both D and K values based on the training data vary slightly with the changing batch numbers and indicate good performance.

In particular, it can be noticed that the significant outlier phenomena in the prediction of parameter D with small D values by the ViT model can be observed. It is suggested that when the D value is small, the system tends to be in the ferromagnetic state, which can also be seen in the phase diagram, and its corresponding LTEM characteristics are not sufficiently distinct, which makes this parameter difficult to predict accurately. LTEM images of the typical FM states with small D values can be seen in Fig. S6 of supplementary materials.

To evaluate the prediction abilities of the ViT model and the CNN model, we compared the values of R^2 and RMSE as we increase the number of training data from 500 to 5000, as shown in Fig. 6. The variations of R^2 and RMSE values indicate that the ViT model presents distinctly better performance than the CNN model, especially when the number of training data decreases to 500. Particularly, the improvements in prediction ability are more obvious for predicting the values of K than predicting the values of D.

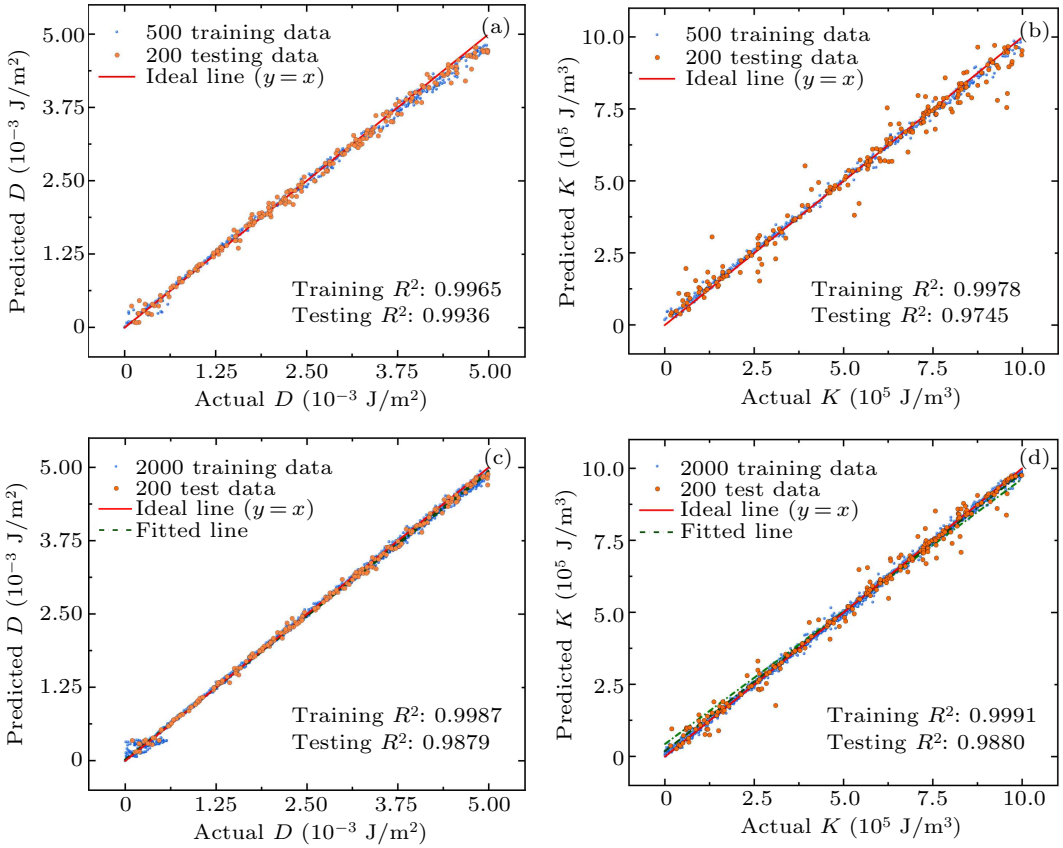


Fig. 5. Performance of ViT model to predict the values of (a) D with 500 training data, (b) K with 500 training data, (c) D with 2000 training data, and (d) K with 2000 training data. The insets list the values of R^2 on training data and testing data.

It is suggested that the ViT model demonstrates superior performance in predicting parameters K and D from LTEM images through its self-attention mechanism and positional encoding. The self-attention mechanism in ViT removes the local receptive field constraint and naturally matches the dual local–global nature of magnetic textures. At the patch level (16×16), self-attention accurately encodes fine local structures such as the skyrmion core, magnetization rotation, and local defect features.[43,44] In particular, through global token interaction, self-attention directly models long-range spa-

tial dependencies across the entire image, including the global symmetry and collective behavior of magnetic textures — key physical properties that cannot be fully represented by local CNN operations. In addition, magnetic systems are governed by long-range interactions (DM interaction, exchange interaction, dipolar interaction). The global receptive field of self-attention is inherently more consistent with the non-local nature of magnetism than the local inductive bias of CNNs. Therefore, the ViT model demonstrates stronger performance than the CNN model in this specific task of analyzing magnetic texture images. Specifically, we have constructed a dedicated dataset of skyrmions with continuously tuned domain wall widths without considering any noise or disturbance effects. Quantitative results show that our model can accurately identify small differences in skyrmion domain wall width, with recognition precision far exceeding the limit of human visual discrimination. The performance of the ViT model in predicting the values from 100 testing data points (all skyrmion states) can be seen in Fig. S7 in supplementary materials.

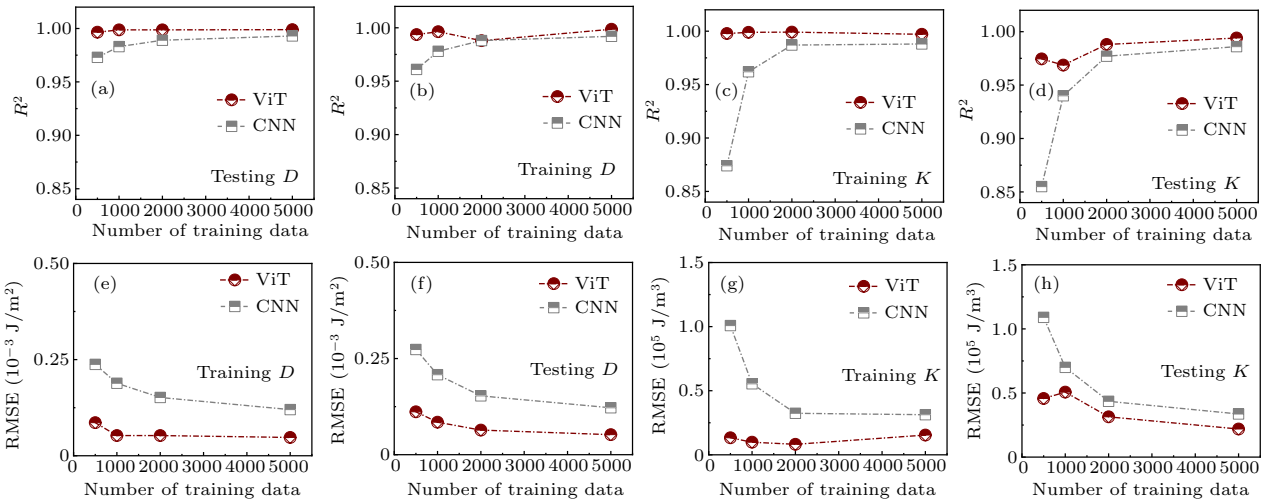


Fig. 6. The values of (a)–(d) R^2 and (e)–(h) RMSE with increasing the numbers of training data from 500 to 5000 on [(a) and (e)] training data and [(b) and (f)] testing data for predicting the D values, [(c) and (g)] training data and [(d) and (h)] testing data for predicting the K values using CNN model (grey symbols) and ViT model (dark red symbols).

In particular, the reliance on synthetic data and noise robustness should be considered. Our training and validation datasets are entirely generated via the Ubermag/OOMMF micromagnetic simulation framework, which produces idealized LTEM images free of experimental artifacts. Experimental LTEM images usually suffer from some non-ideal effects such as Poisson noise, shot noise, thickness variations, and bend contours.[47,48] Here, we have implemented two types of noise injection schemes. Firstly, we added an irregular fluctuating field $\boldsymbol{h}_{\text{fluct}}$ to modify the Hamiltonian to generate the LTEM images with the non-ideal effects.[44] Secondly, we investigated the noise robustness by injecting Poisson noise into the original simulated LTEM images.

As shown in Fig. 7(a), the domain wall between up-magnetization and down-magnetization becomes indistinct with increasing the strength of the fluctuating σ from 0 mOe (1 Oe = 79.5775 A·m^{-1}) to 30 mOe. Figure 7(b) presents the R^2 on training data and testing data with increasing σ. We can see that the R^2 values remain almost unchanged on the training data and testing data for predicting D values with increasing the strength of the fluctuation. Meanwhile, the R^2 values remain almost unchanged on the training data for predicting K and keep decreasing on the testing data with increasing the strength of the fluctuation. We can see that the R^2 value decreases to 0.7943 on the testing data for predicting K with $\sigma = 30$ mOe, which indicates a poor predictive ability. Fig. S8 presents the distributions of the training data and testing data to indicate a uniform and random spatial distribution.

As shown in Fig. 7(c), the LTEM images of a typical maze state of topological spin texture degenerate obviously with dose level (DOSE) from 1.0 to 0.05 by Poisson noise injection. Similar to the results that we added an irregular fluctuating field to simulate the non-ideal effects, the R^2 values keep almost unchanged on the training data and testing data for predicting D values and on the training data for predicting K with increasing the strength of the fluctuating. Meanwhile, the R^2 values keep decreasing on the testing data with the dose level change to DOSE = 0.05, which gives $R^2 = 0.8323$ indicating a poor prediction ability.

It can be seen that the prediction of parameter D remains relatively stable under the added fluctuating field or Poisson noise, which confirms the robustness of our model to a certain extent. Meanwhile, the prediction accuracy of parameter K decreases significantly when unstable factors are introduced. Compared with DMI constant D, magnetic anisotropy K may be more closely related to domain wall curvature, local magnetization gradient, fine texture and sharpness of domain boundaries, which are all small-scale image features and are easily

affected by a fluctuating field or Poisson noise. To improve the robustness of the K prediction, it is suggested that future work focus on introducing perturbation-resilient feature extraction and adding noise-aware training or adversarial training to enhance anti-interference ability. For clarity and better readability, the performance of the ViT model in predicting the values under the above non-ideal effects can be seen in Figs. S9 and S10 in the supplementary materials.

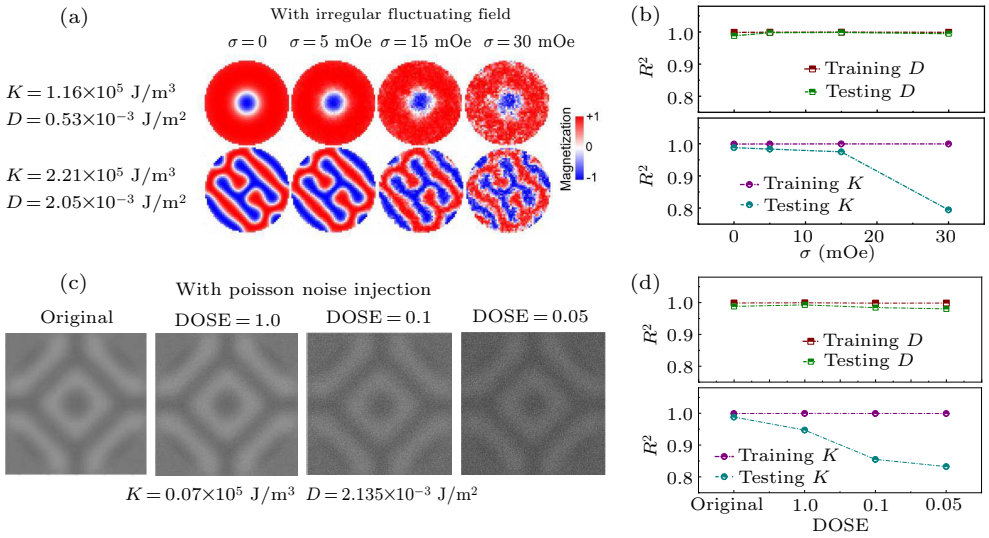


Fig. 7. (a) Magnetization distributions of the typical topological spin textures of the skyrmion state ($K = 1.16 \times 10^5$ J/m^3, $D = 0.53 \times 10^{-3}$ J/m^2) and the maze state ($K = 2.21 \times 10^5$ J/m^3, $D = 2.05 \times 10^{-3}$ J/m^2) with increasing the strength of the fluctuating. (b) R^2 on training data and testing data with increasing the strength of the fluctuating. (c) Simulated LTEM images with different dose level for a typical maze state of topological spin texture ($K = 0.07 \times 10^5$ J/m^3, $D = 2.135 \times 10^{-3}$ J/m^2). (d) R^2 on training data and testing data with different dose level.

4. Conclusion

In conclusion, we used machine learning models, CNN, and ViT, to achieve the material parameters of topological spin textures from simulated LTEM images prepared by the micromagnetic simulation method. Compared with the CNN model, the ViT model presents better performance in achieving D and K values. The coefficient of determination R^2 could reach values as high as 0.9987 and 0.9991 for predicting D and K values, respectively. Finally, the reliance on synthetic data and noise robustness was discussed. Generally, material parameters of topological spin textures such as DMI constant D or anisotropy constant K cannot be obtained through LTEM images directly by experimental methods, but these correlative structure-property relationships can be learned by ML models and the properties can be predicted by the trained ML models. However, we must realize that this ML-assisted parameter prediction work will be more significant and have greater application value if the experimental LTEM images can be used as the training dataset. This process may contribute to the exploration of the topological spin textures and the design of advanced spintronic devices.

References

[1] Fert A, Reyren N and Cros V 2017 *Nat. Rev. Mater.* **2** 17031
[2] Göbel B, Mertig I and Tretiakov O A 2021 *Phys. Rep.* **895** 1
[3] He Q L, Hughes T L, Armitage N P, Tokura Y and Wang K L 2022 *Nat. Mater.* **21** 15
[4] Jiang L X, Li Q C, Zhang X, Li J F, Zhang J, Chen Z X, Zeng M and Wu H 2024 *Acta Phys. Sin.* **73** 017505 (in Chinese)
[5] Yang H X, Liang J H and Cui Q R 2023 *Nat. Rev. Phys.* **5** 43
[6] Zhao X P and Zhao J H 2025 *npj Spintronics* **3** 25
[7] Mühlbauer S, Binz B, Jonietz F, Pfleiderer C, Rosch A, Neubauer A, Georgii R and Boni P 2009 *Science* **323** 915
[8] Kahmann T, Laureen Rösch E, Enpuku K, Yoshida T and Ludwig F 2021 *J. Magn. Magn. Mater.* **519** 167402
[9] Masood K I, Raich R, Jander A and Dhagat P 2024 *J. Magn. Magn. Mater.* **610** 172562
[10] Kuepferling M, Casiraghi A, Soares G, Durin G, Garcia-Sanchez F, Chen L, Back C H, Marrows C H, Tacchi S and Carlotti G 2023 *Rev. Mod. Phys.* **95** 015003
[11] Phatak C, Petford-Long A K and De Graef M 2016 *Curr. Opin. Solid St. M.* **20** 107
[12] Kang S, Töllner M, Wang D, Minnert C, Durst K, Caron A, Dunin-Borkowski R E, McCord J, Kübel C and Mu X 2025 *Nat. Commun.* **16** 1305
[13] Tang J, Kong L Y, Wang W W, Du H F and Tian M L 2019 *Chin. Phys. B* **28** 087503
[14] Marchiori E, Ceccarelli L, Rossi N, Lorenzelli L, Degen C L and Poggio M 2022 *Nat. Rev. Phys.* **4** 49
[15] Habenschaden C, Sievers S, Klasen A, Cerreta A and Schumacher H W 2024 *Rev. Sci. Instrum.* **95** 113704
[16] Kazakova O, Puttock R, Barton C, Corte-León H, Jaafar M, Neu V and Asenjo A 2019 *J. Appl. Phys.* **125** 060901
[17] Allen E A and Tkatchenko A 2022 *Sci. Adv.* **8** eabm7185
[18] Li Q, Fu N, Omee S S and Hu J 2024 *npj Comput. Mater.* **10** 245
[19] Mobarak M H, Mimona M A, Islam M A, Hossain N, Zohura F T, Imtiaz I and Rimon M I H 2023 *Appl. Surf. Sci. Adv.* **18** 100523
[20] Ma W, Liu Z C, Kudyshev Z A, Boltasseva A, Cai W S and Liu Y M 2021 *Nat. Photon.* **15** 77
[21] Liu D Y, Xu L M, Lin X M, Wei X, Yu W J, Wang Y and Wei Z M 2022 *Chip* **1** 100033
[22] He W W, Li J Z, Kong X and Deng L 2024 *Commun. Eng.* **3** 151
[23] Li J, Fang W Q, Jin S J, Suo C Y, Zhang T D, Wu Y L, Xu X D, Liu Y and Yao D X 2025 *AI Sci.* **1** 015001
[24] Karniadakis G E, Kevrekidis I G, Lu L, Perdikaris P, Wang S F and Yang L 2021 *Nat. Rev. Phys.* **3** 422
[25] Choubisa H, Todorović P, Pina J M, Parmar D H, Li Z L, Voznyy O, Tamblyn I and Sargent E H 2023 *npj Comput. Mater.* **9** 117

[26] Shi S W, Chu S B, Xie Y E and Chen Y P 2025 *Phys. Scr.* **100** 015957
[27] Chen G H, Yao J C, Zhu H F, Zhi T, Wang J, Xue J J, Chen L, Tao T and Tao Z K 2025 *Acta Phys. Sin.* **74** 190203 (in Chinese)
[28] Mehmood N, Wang J B and Liu Q F 2022 *J. Appl. Phys.* **132** 043904
[29] McCray A R C, Zhou T, Kandel S, Petford-Long A, Cherukara M J and Phatak C 2024 *npj Comput. Mater.* **10** 111
[30] McCray A R C, Bender A, Petford-Long A and Phatak C 2024 *APL Mach. Learn.* **2** 026120
[31] Beg M, Lang M and Fangohr H 2022 *IEEE Trans. Magn.* **58** 7300205
[32] Beleggia M and Zhu Y 2010 *Philos. Mag.* **83** 1045
[33] Chen X W, Yang L Y, Savchenko A, Yang D H, Shi W, Denneulin T, Kiselev N S, Jin L, Song D S, Dunin-Borkowski R E and Zheng F S 2025 *Phys. Rev. B* **112** 214455
[34] Haug T, Otto S, Schneider M and Zweck J 2003 *Ultramicroscopy* **96** 201
[35] Walton S K, Zeissler K, Branford W R and Felton S 2013 *IEEE Trans. Magn.* **49** 4795
[36] Holt S J R, Lang M, Loudon J C, Hichken T J, Suess D, Cortes-Ortuno D, Pathak S A, Beg M, Zulfiqar K and Fangohr H 2025 *npj Comput. Mater.* **11** 20
[37] Yu X Z, Onose Y, Kanazawa N, Park J H, Han J H, Matsui Y, Nagaosa N and Tokura Y 2010 *Nature* **465** 901
[38] Yu X Z, Kanazawa N, Onose Y, Kimoto K, Zhang W Z, Ishiwata S, Matsui Y and Tokura Y 2011 *Nat. Mater.* **10** 106
[39] Yu X Z, Koshibae W, Tokunaga Y, Shibata K, Taguchi Y, Nagaosa N and Tokura Y 2018 *Nature* **564** 95
[40] Zhao X, Wang L M, Zhang Y F, Han X M, Deveci M and Parmar M 2024 *Artif. Intell. Rev.* **57** 99
[41] Alzubaidi L, Zhang J, Humaidi A J, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel M A, Al-Amidie M and Farhan L 2021 *J. Big Data* **8** 53
[42] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J and Houlsby N 2010 ArXiv abs/2010.11929
[43] Maurício J, Domingues I and Bernardino J 2023 *Appl. Sci.* **13** 5521
[44] Chicco D, Warrens M J and Jurman G 2021 *PeerJ Comput. Sci.* **7** e623
[45] Schmidhuber J 2015 *Neural Networks* **61** 85
[46] Akhtar N and Ragavendran U 2020 *Neural Comput. Appl.* **32** 879
[47] Lyu B Y, Zhao S H, Zhang Y B, Wang W W, Zheng F S, Dunin-Borkowski R E, Zang J D and Du H F 2024 *Sci. China-Phys. Mech. Astron.* **67** 117511
[48] García-Palacios J L and Lázaro F J 1998 *Phys. Rev. B* **58** 14937