Application of convolutional neural networks to spin models studies

A. V. Perzhu [1], E. V. Vasiliev [1,2], A. O. Korol [1,2], D. Yu. Kapitan [1,2]

A. E. Rybin [1,2], V. Yu. Kapitan [1,2] ✉

[1] Far Eastern Federal University, Vladivostok, Russia;

[2] Institute of Applied Mathematics FEB RAS, Vladivostok, Russia

✉ kapitan.vyu@dvfu.ru

Abstract: Nowadays, methods and techniques of Deep Learning are being used in various scientific areas. In this paper, the applying of convolutional neural network was considered in frame of problems from statistical physics and computer simulation of magnetic films. In a frame of the first task, CNN was used to determine critical Curie point for Ising model on 2D square lattice. Obtained results were compared with classical Monte-Carlo method and exact solution. Systems of various lattice sizes and the influence of the size effect on the results' accuracy were considered. Also, authors considered the classical two-dimensional Heisenberg model, a spin system with direct short-range exchange, and studied of its competition with the Dzyaloshinskii-Moriya interaction. A neural network was applied to the recognition of Spiral (Sp), Spiral-skyrmion (SpSk) Skyrmion (Sk), Skyrmion-ferromagnetic (SkF) and Ferromagnetic (FM) phases of the Heisenberg spin system with magnetic skyrmions.

Keywords: Convolutional neural network, Metropolis algorithm, Ising model, Heisenberg model, Magnetic Skyrmion

Применение сверточных нейронных сетей для исследования спиновых моделей

А. В. Пержу [1], Е. В. Васильев [1,2], А. О. Король [1,2]

Д. Ю. Капитан [1,2], А. Е. Рыбин [1,2], В. Ю. Капитан [1,2] ✉

[1] Дальневосточный Федеральный Университет, г. Владивосток, Россия;

[2] Институт прикладной математики ДВО РАН, г. Владивосток, Россия

✉ kapitan.vyu@dvfu.ru

Аннотация. В настоящее время методы глубокого обучения используются в различных научных областях. В данной работе применение сверточной нейронной сети рассмотрено в рамках задач статистической физики и компьютерного моделирования магнитных пленок. В рамках первой задачи CNN использовалась для определения критической точки Кюри для модели Изинга на двумерной квадратной решетке. Полученные

результаты сравнивались с классическим методом Монте-Карло и точным решением. Рассмотрены системы с различными размерами решетки и влияние размерного эффекта на точность результатов. Также авторы рассмотрели классическую двумерную модель Гейзенберга, спиновую систему с прямым короткодействующим обменом, и изучили ее конкуренцию с взаимодействием Дзялошинского-Мория. Нейронная сеть применялась для распознавания Спиральной (Sp), Спирально-скирмионной (SpSk), Скирмионной (Sk), Скирмион-ферромагнитной (SkF) и Ферромагнитной (FM) фаз спиновой системы Гейзенберга с магнитными скирмионами.

Ключевые слова: Сверточная нейронная сеть, алгоритм Метрополиса, модель Изинга, модель Гейзенберга, магнитный скирмион

Introduction

In the fundamental scientific works [1, 2], as well as in modern ones [3-6], much attention is paid to lattice structures. The interactions between spins in the lattice sites can lead to collective behavior and macroscopic effects, for example, as widely known as ferromagnetism or anti-ferromagnetism. Also, recently, structures that have no analogues in natural materials have been actively investigated. This is the reason for the use of supercomputer modelling to study such artificial structures and theoretically predict their properties. Because of supercomputers, it possible to use new classes of algorithms and operate with large and super-large amounts of data to carry out numerical experiments. Numerical methods and computer simulation on a supercomputer are of paramount importance in statistical and mathematical physics, nanophysics, and statistical thermodynamics since supercomputers significantly speed up the solution of various scientific problems [7]. And thanks to the development of machine learning (ML), the software tools for conducting numerical experiments have significantly expanded recently, but scientists are just beginning to reveal the full potential of introducing machine learning methods into their research [8-10].

In our paper, we discussed the applying of CNN in frame of two problems from statistical physics and computer simulation of magnetic films. The first problem is about determination of critical Curie point for Ising model. And the second one is the recognition of different phases of the Heisenberg spin system with magnetic skyrmions.

Research problems and methods

In our work, it was demonstrated that modern machine learning methods can provide new approaches to the study of physical systems within the frame of statistical physics models. For this, the TensorFlow library was used to create a convolutional neural network [11]. In this study, the Metropolis algorithm for Monte Carlo simulation was applied to generate input data for the neural network, and then compared with the results obtained after training the convolutional neural network. We considered two mathematical models of statistical physics: the Ising model with direct exchange and the Heisenberg one with the Dzyaloshinskii-Moriya [12,13] interaction and skyrmions in a system, see more details [7,14]. All values in the work are given in dimensionless values.

A convolutional neural network

We used configurations of spin systems obtained at different simulation parameters for the training and subsequent classification of them in a neural network. To date, the most accurate analysis results are demonstrated by neural networks based on convolutional architecture. We used the TensorFlow library to create a convolutional neural network and to classify our spin systems to different phases.

In our research, we have reduced the problem of determining the phases of spin systems to the problem of image classification - in fact, to the main problem area in which neural networks are used. For recognizing images, CNN accepts them in the RGB format as a three-dimensional matrix. In our case, the convolutional neural network received as input a three-dimensional array representing the components of a spin.

Following this, the convolutional neural network learned, using the training dataset, to highlight the features inherent in one or another spin configuration. Our CNN consists of next layers (main ones), see Figure 1:
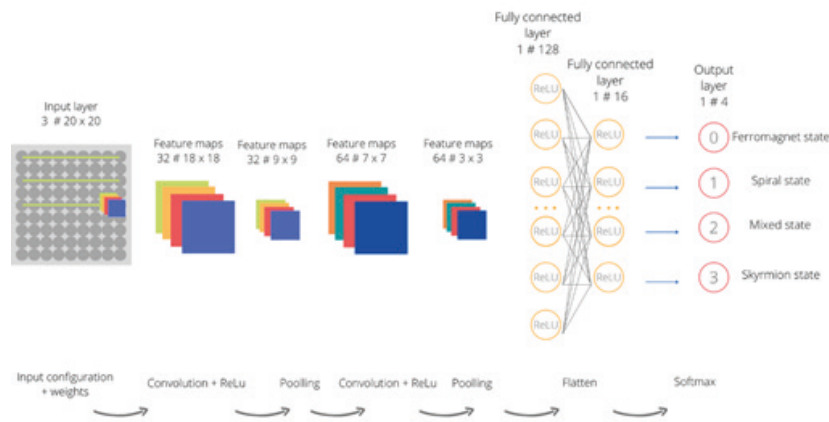

Fig. 1. Architecture of the convolutional neural network

1. Input layer

Input data (configurations of spins), each of the neurons (spins) of which is assigned an initial random weight. The components of a three-dimensional vector were fed to the network input (i.e., the components of Heisenberg spin). The dataset was prepared using Monte Carlo simulation data for training the neural network in state recognition.

2. Convolutional layer with 3×3 filter

When neurons are connected to only a few neurons in the next layer, the layer is said to be convolutional. The convolutional layer acts as a filter that discards the least informative parts of the input data. Each layer has filters (i.e., matrices with weight values). When the filter moves along the matrix of the previous layer, each filter element is multiplied by the value of the neuron, and the values are summed up and written to the feature map.

3. Pooling layer for reducing the dimensions of the data.

4. Fully connected layer

Fully connected layers are used for classification. All layers before the fully connected layer are used to highlight various features that are fed to the input of the classifier. This layer can also be used as the final (output) CNN layer, the result of which is the probability of the input configuration of spins belonging to a certain class.

Results and Discussion
Determination of the second order transition of the Ising model

Different sets of input data of the neural network obtained with different parameters of the Metropolis algorithm for systems of 10×10 and 20×20 Ising spins were used. The obtained data will be used to select the optimal simulation parameters, which will be further used in the study of more complex spin systems. A comparative analysis is carried out with the results of MC modelling and the exact solution of Onsager.

At the first stage, the network was trained on spin configurations obtained during on MC simulation with the following parameters: system size: 10×10, $T = 0.1 ... 5.0$ with a step of 0.01, the number of MC steps for preliminary equilibration of the system: 10000, the number of MC steps for calculating thermodynamic averages in the Metropolis algorithm: 10000, the sample size of configurations for training the network: 50 per one step in temperature, the results are shown in Fig. 2. In Fig. 2, the result of applying a convolutional neural network to the calculation of the critical point T_c was presented in comparison with Onsager's exact solution and the result of MC simulation.
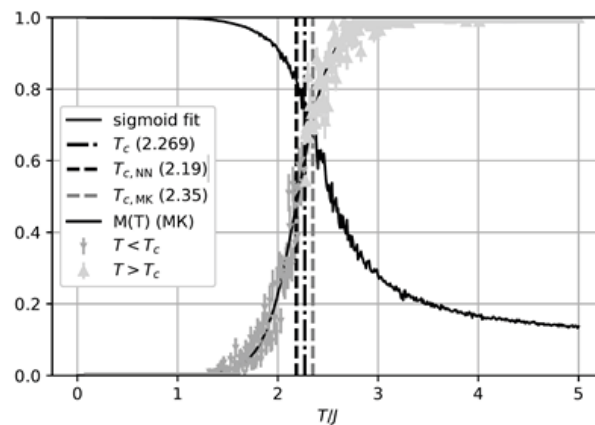

Fig. 2. Results of T_c calculations by various methods

The effect of system size on the accuracy of the obtained results was tested on the system with 20×20 spins. These values were generally similar to the ones given above. It should be noted that an increase of the system size had a positive effect on the results of MC modelling in the calculations of T_c: $T_c = 2.29$, due to a decrease of the influence of the size effect, while the increase in the system size did not significantly affect the results of the neural network operation. The accuracy of the predicted value of the critical temperature, in comparison with the case described above, on average did not change, and in some numerical experiments it even worsened, because network training is based on a probabilistic approach.

The recognition of different phases of the Heisenberg spin system with magnetic skyrmions

Our second way of using the CCN was a data analysis of a study of different phases that appeared depending on the magnitude of the external magnetic field and temperature T at fixed Dzyaloshinskii-Moriya interaction D, see Fig. 3. The diagram in Fig. 3 shows that in the low temperature zone we have ordered phases. Thus, the ground state is the spiral phase, which is observed in the field range $0 - 0.3$; with a further increase in the magnetic field, the spiral phase passes into the skyrmion phase, after which a further transition from one phase to another is observed, up to the temperature range $T > 0.5$, where the system goes into a paramagnetic state. Skyrmions are thermally stable in a fairly wide temperature range, with an external magnetic field H_z from 0.8 to 1.5. The convolutional neural network was used to analyze the data obtained from the Monte Carlo simulations for the recognition of the different phases of the spin system, dependent on the simulation parameters.

One of the conventional methods is to compute the skyrmion number, which is evaluated to keep track of the skyrmion creation process. However, it does not indicate the mixed states of the spin systems very well, depending on the simulation parameters, e.g. a spiral-skyrmion phase, therefore, we use the convolutional neural network in our work.

In a magnetic film, with an increase of the magnetic field strength and temperature, various phases were observed for the flat Heisenberg spin systems: Spiral (S), Labyrinths (L) Spiral-skyrmion (SS) Skyrmion (Sk), Skyrmion-Ferromagnetic (SkF), Ferromagnetic (F), Paramagnetic (P) phases, see Fig. 3. In Skyrmion phase, due to the alignment of the stripes against the magnetic
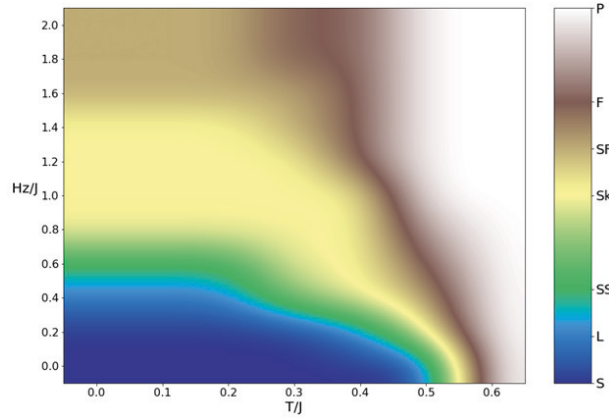
Fig. 3. Phase diagram (T, H_z) at $D = 1.3$

field, stable skyrmions are formed in the system. In these skyrmions, the spins of the nucleus are directed against the magnetic field. In this study, skyrmions of the Bloch type were formed.

Conclusion

The paper considered the application of convolutional neural networks to determine the critical temperature of a second-order phase transition in comparison with performed MC simulations and known solutions. As it was shown above CNN could be successfully used to such problems by reducing them to the problem of classifying spin states at different temperatures. The dependence on the number of Monte Carlo steps and the sample size for the accuracy of training the network and its subsequent application is shown in comparison with the Metropolis algorithm. Systems of various sizes and the influence of the size effect on the accuracy of the results are considered.

The authors also noted the feature of the results obtained using neural networks to determine T_c: if the calculation is performed using the Metropolis algorithm, then always $T_c^{MC} \geq T_c^{exact}$. In turn, in the calculations carried out using convolutional neural networks $T_c^{NN} \geq T_c^{exact}$. The reasons for this behaviour are the subject of future research, during which it is planned to apply neural networks for studying more complex models and lattices.

Also, in the frame of the classical two-dimensional Heisenberg model, a spin system with direct short-range exchange was modelled, and a study of its competition with the Dzyaloshinskii-Moriya interaction was carried out. Due to the direct exchange interaction, the neighbouring spins of the system are collinearly aligned, and, in turn, the Dzyaloshinskii-Moriya interaction contributes to the deviation of the spins from parallel orientation. As a result, competition results between collinear and noncollinear alignments of spins, which leads to the transition of the system of spins from a ferromagnetic to a spiral ground state. In the presence of an external magnetic field, stable topological structures, i.e., magnetic skyrmions, are generated in such systems.

One of the most effective and popular approaches in statistical physics is Monte Carlo simulation, which consists of a stochastic sample over the state space and an estimate of physical quantities. Monte Carlo methods are not only actively used to study various physical systems, but also continue to actively develop and improve due to the development of supercomputers. The ability of modern machine learning algorithms to classify, identify and interpret large data sets and, on their basis, to predict new properties and states of the systems under study provides an additional paradigm to the above approach for processing the exponentially increasing number of analyzed states in statistical physics.

REFERENCES

1. **Sherrington D., Kirkpatrick S.,** Solvable model of a spin-glass, Phys. Rev. Lett. 35, (1975) 1792−1796

2. **Swendsen R. H., Wang J.-S.,** Nonuniversal critical dynamics in Monte Carlo simulations, Phys. Rev. Lett. 58, (1987), 86

3. **Belokon V. I., D'yachenko O. I., Kapitan V. Y.,** On the possible application of the method of random exchange interaction fields for studying the magnetic properties of the rocks, Izvestiya, Physics of the Solid Earth. 51 (5) (2015) 622−629

4. **Belokon V. I., Kapitan V. Y., Dyachenko O. I.,** Concentration of magnetic transitions in dilute magnetic materials, J. of Physics: Conf. Ser., 490 (1) (2014) 012165

5. **Prudnikov P.V., et al.,** Influence of anisotropy on magnetoresistance in magnetic multilayer structures, Journal of Magnetism and Magnetic Materials. 482 (2019) 201−205

6. **Makarov A.G., et al.,** On the numerical calculation of frustrations in the ising model, JETP Letters, 110 (10) (2019) 702−706

7. **Landau D., Binder K.,** A guide to Monte Carlo simulations in statistical physics, Cambridge University press, Cambridge, 2003

8. **Carleo G. et al.,** Machine learning and the physical sciences, Reviews of Modern Physics, 91 (4) (2019), 045002

9. **Carrasquilla J., Melko R. G.,** Machine learning phases of matter, Nature Physics, 13 (5), (2017) 431−434

10. **Kapitan V., et al.,** Numerical simulation of magnetic skyrmions on flat lattices, AIP Adv. 11, (2021) 015041

11. **Abadi M.et al.,** Tensorflow: Large-scale machine learning on heterogeneous distributed systems // arXiv preprint arXiv:1603.04467. 2016.

12. **Dzyaloshinsky I.,** A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics, J. of Physics and Chemistry of Solids, 4 (1958) 241−255

13. **Moriya T.,** Anisotropic superexchange interaction and weak ferromagnetism, Physical Review 120, (1960) 91

14. **Soldatov K. S., Nefedev K. V., Kapitan V. Yu., Andriushchenko P. D.,** J. Phys. Conf. Ser. 741, (2016) 012199

THE AUTHORS

PERZHU Aleksandr V.
perzhu_av@dvfu.ru
ORCID: 0000-0001-9717-3773

VASILIEV Egor V.
vasilev.eva@dvfu.ru
ORCID: 0000-0001-5209-510X

KOROL Alena O.
korol.ao@dvfu.ru
ORCID: 0000-0001-8527-0752

KAPITAN Dmitrii Yu.
kapitan.diu@dvfu.ru
ORCID: 0000-0001-9815-1891

RYBIN Alexey E.
rybin.ae@dvfu.ru
ORCID: 0000-0002-1055-9217

KAPITAN Vitalii Yu.
kapitan.vyu@dvfu.ru
ORCID: 0000-0002-5068-8910