Learning Renormalization with a Convolutional Neural Network

Alex Nguyen
School of Natural Sciences
Minerva Schools at KGI
San Francisco, CA 94103
alexn@minerva.kgi.edu

Kiel Howe
School of Natural Sciences
Minerva Schools at KGI
San Francisco, CA 94103
khowe@minerva.kgi.edu

Abstract

Renormalization is a technique for studying the scale-dependence of correlations in physical systems through iterative coarse-grainings over longer and longer distance scales. The similarity of this procedure to the iterative layer-by-layer abstraction that occurs in deep learning models has motivated applications of machine learning to modeling physical systems, in particular the Ising model, and has also led to speculation that viewing deep learning through the lens of renormalization may be able to explain some of the effectiveness of deep neural networks in other domains. In this work, we demonstrate a method to find an optimal coarse-graining of the 2D Ising model by training a feedforward convolutional neural network (CNN) with a simple classification objective that incentivizes the network to learn long-distance correlations. We comment on the information theory interpretation of this method and its possible applications to a wider variety of machine learning problems.

1 Introduction

In physics, the correlations of variables in a statistical (or quantum) system are often encoded in a local Hamiltonian, which gives a compact representation of the correlations and makes manifest the underlying interactions of the system's fundamental components. For example, the 2D Ising model is parameterized by a set of 'spin' variables with values $s_i = \pm 1$ and local interactions on a square lattice, and its correlations are completely defined by the temperature T and the local hamiltonian $H = -\sum_{\text{links ij}} s_i s_j$. However, the local Hamiltonian description of a theory does not make manifest the scale-dependence of correlations– for example the emergence of long range correlations at the critical temperature of the phase transition is not at all obvious in the local Hamiltonian description. Renormalization developed as a set of techniques to make this scale-dependence manifest by iteratively coarse-graining the variables of the theory. The 2D Ising model can be renormalized by coarse-graining clusters of spins through a simple averaging or decimation procedure [17], and the 'flow' of the coupling strength of the local interactions between the coarse-grained spins gives a simple theoretical picture of the scale-dependence of the correlations and a way to access the critical behavior of the phase transition. More recently, holographic theories and multiscale tensor networks [25] have also emerged as alternative descriptions of systems originally described by local Hamiltonians. These descriptions make the renormalization of the theory manifest by embedding scale-dependent coarse-grainings of the system into auxiliary variables, and have been proposed for many models including the 2D Ising model [5].

Meanwhile, deep neural networks have developed in the machine learning community as an effective tool for encoding correlations learned from training data in high dimensional spaces. Qualitatively, the layers of the network appear to allow iterative abstraction in the representation of the data [2]. These networks have been shown to be effective in modeling the 2D Ising model [22, 21, 24, 18, 7, 8, 16, 6, 4, 27, 28], and it has been proposed that the deep neural network description of a system may make

Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019), Vancouver, Canada.

manifest the renormalization of the system, as layers of the network can be interpreted as progressive coarse-grainings. Evidence that this is true for deep belief networks and the 2D Ising model has been presented and debated in Refs. [22, 21, 24, 18][1], and leads naturally to connections between deep learning and tensor networks [3, 23, 20, 10] and holography [9, 13, 12, 11, 15, 26, 10, 14]. Conversely, these works also propose that the ability of deep learning models to discover renormalization-like descriptions of physical theories may help explain their effectiveness at modeling more general distributions with highly scale-dependent correlations, such as those found in computer vision problems.

In the literature so far, these connections have primarily been explored in the context of generative RBM and deep belief models. In this work, we instead explore the connections between renormalization of the Ising model and deep neural networks using a simple feedforward CNN and a supervised learning task, which we believe illustrates the generality of this connection to more typical computer vision models. Our approach, similar in spirit to Refs. [18, 19], is to use a neural network to identify an optimal coarse-graining of the system, and our results show that the training procedure naturally identifies a non-trivial coarse-graining in the convolutional layers without any a-priori knowledge of the dynamics of the system. The fact that the network can find such a coarse-graining can be seen both as a practical tool for using machine learning to more efficiently coarse-grain physical theories, and as an insight into how computer vision neural networks may be 'applying renormalization' to develop features that preserve long-distance correlations while discarding short-distance information. In Section 2 we explain the motivation and details for our training procedure, and then in Section 3 we present our results, before returning to these more general connections between renormalization and deep learning in our conclusions.

2 Methods

Since the goal of renormalization is to coarse-grain in a way that preserves long distance correlations in the data, we propose a supervised learning task for our CNN that leads to this behavior. We simply train the network to distinguish between two datasets, the 'correlated' and 'scrambled' set – the 'correlated' set is generated by monte carlo sampling of the model, and the 'scrambled' set is generated from the 'correlated' set by a scrambling procedure that yields identical short range distributions but removes all long-range correlations in the data. The network first filters the short range data through a convolutional layer, generating a coarse-grained representation. Then, a fully connected layers take the coarse-grained representation as input and learn to distinguish the datasets based on the long-range correlations. This process forces the convolutional layer to learn features that extract the long-distance correlations. Refs. [18, 19] have defined an optimal coarse-graining as one that optimizes the mutual information between the coarse-grained region and its environment, and demonstrated a purely information theoretic learning objective with an RBM model that finds optimal features; in this work we show that our simple training scheme, which can be implemented with comparative ease and more closely resembles standard computer vision approaches, can also learn coarse-grainings that approach optimality.

Concretely, we first generated a Monte Carlo data set of 10,000 Ising models configurations on an 81×81 lattice with periodic boundary conditions, sampled at the critical temperature $T_c = 2.269$. From this dataset, we draw 10,000 configurations of a 9×9 square, and call this the 'correlated' data set[2]. We generate the 'scrambled' data set by sampling 3×3 squares from the full set of configurations, and concatenating them at random into 9×9 squares. Samples of the correlated and scrambled class are shown in Figure 3 (extended to the full 81×81 lattice).

We then trained a CNN to distinguish between the scrambled and correlated datasets. The model architecture comprises a convolutional layer of one 3×3 filter with a stride of 3, using a $\tanh$ activation function, followed by a multilayer perceptron with a hidden layer of 10 units. Effectively, the convolutional layer coarse-grains the 9×9 image into a 3×3 image, which is then classified by the fully connected layers. We used the ReLU activation function for the input and hidden layers and used the sigmoid activation function to output a probability that a system is correlated. We used

[1] Refs [4, 27, 28] use similar techniques, but focus on machine learning as a tool to identify phase transitions. Refs. [8, 16, 6] have followed another tack of defining an RBM flow and comparing it to the RG flow, though the generality and consequences of this connection are less clear.

[2] Our 9×9 image size is chosen to be much smaller than the full 81×81 lattice to avoid finite size effects. In future work we will demonstrate this procedure applied iteratively on much larger lattices.

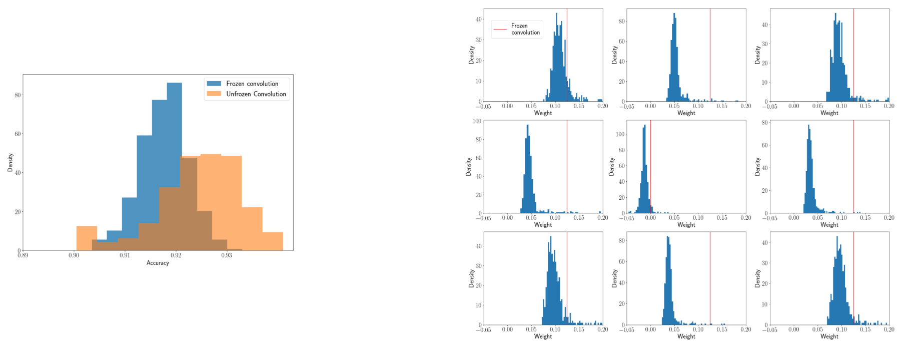

Figure 1: Left: Histogram of test accuracies for the frozen convolution vs. unfrozen convolution training schemes. Right: Histogram of the weights after training. The red vertical line is the weight of the frozen convolution (1/8 except for the center).

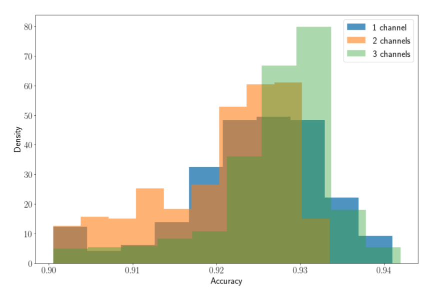

Figure 2: Histogram of test accuracies for different numbers of convolution channels.

the cross-entropy loss function, and trained the network using stochastic gradient descent using the Adam optimizer. The hyperparameters were adaptively tuned using the Adaptive Experimentation platform [1].

We trained the model on 5000 data points, setting aside 2000 data points for testing. Our metric of choice is accuracy on the test set, where we consider a prediction for a data point accurate if the network's output lies on the correct side of the decision boundary (above 0.5 for a correlated system, below 0.5 for a scrambled system). For each model we trained 500 times with different initializations to obtain a distribution of accuracies.

3 Results

To show that our method finds a non-trivial coarse-graining that maintains the long-distance information, we compare two cases. In the first case, we train the fully connected portion of the network on a standard coarse-graining, the average of all the boundary spins (see Fig. 1) – this is called the 'frozen convolution' model, indicating that we have in effect simply fixed the convolutional layer to perform an average over the spins. We compare this to the 'unfrozen convolution', where the weights of the convolutional layer are allowed to adjust during the training period, learning a more effective sets of weights. As shown in Fig. 1 and Table 1, the accuracy of the model with an

Table 1: Test Accuracies

Name	Accuracy $\pm$ Standard Error
Frozen Convolution	$0.918 \pm 2.16 \times 10^{-4}$
Unfrozen Convolution (1 Channel)	$0.924 \pm 3.82 \times 10^{-4}$
Unfrozen Convolution (2 Channels)	$0.921 \pm 3.68 \times 10^{-4}$
Unfrozen Convolution (3 Channels)	$0.927 \pm 3.31 \times 10^{-4}$

Table 2: Mutual Information

System	Mutual Information (bits)
Fine Grained	0.743
Coarse grained (Unfrozen Convolution)	0.618
Coarse grained (Frozen Convolution)	0.602

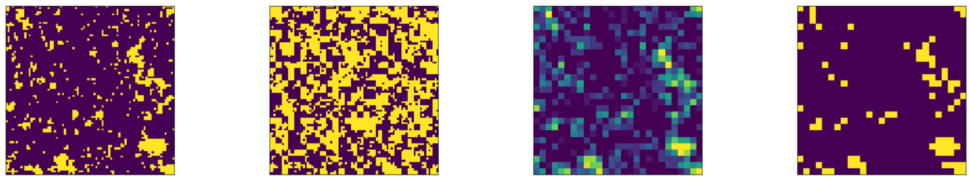

Figure 3: From left to right: A sample of the correlated data class. A sample of the scrambled data class. An undiscretized coarse-graining for the correlated data sample. A discretized coarse-graining for the correlated data sample.

unfrozen convolution is significantly higher, indicating more of the long-range correlations have been preserved by the learned coarse-graining. On the right of Fig. 1 the learned weights are visualized, showing that the non-trivial convolution learned by the network emphasizes the corners of the 3x3 spin block it averages over. This agrees with the intuition that the corners are more strongly coupled to the neighboring spins than the other edge sites. While this remains a fairly simple coarse-graining problem, in more sophisticated physical theories like quantum chromodynamics (QCD) such an approach may have promise to learn highly non-trivial coarse-grainings.

As shown in Table 2, we can use the mutual information to quantify the amount of long-range information preserved by the coarse-graining. We use the Monte Carlo data to estimate the mutual information between two neighboring 3×3 blocks before and after the coarse-graining, where we measure the mutual information in bits defined by

$$I(A,B) = \sum_{A,B} P(A,B) \log_2 \frac{P(A,B)}{P(A)P(B)} \tag{1}$$

As argued in [18, 19], an optimal coarse-graining would preserve all of the mutual information between the block and its environment, and we use the mutual information between two neighboring blocks as a proxy for this. The results show that the learned coarse-graining comes close to the bound, even though unlike the procedure of [18] we have not directly optimized the mutual information in our training scheme.

Another strength of this method is that additional convolutional channels can be added, corresponding to coarse-grainings to multiple variables. As shown in Fig. 2 and Table 1, increasing the number of channels appears to increase the accuracy, though higher statistics samples are needed to verify that the additional channels increase the mutual information.

4 Conclusion

In this work we demonstrated a simple method using a feedforward CNN to learn optimal coarse-grainings that preserve long distance correlations in data, analogous with renormalization approaches in physics. With the 2D Ising model as a test case, we used a mutual information measure to demonstrate that our learned coarse-graining features perform better than a naive averaging used in standard renormalization approaches. In future work, we will demonstrate the flow of the learned coarse-graining parameters under iterative applications of the technique to the Ising model, and intend to show that this procedure can also be applied to significantly more complicated physical theories like QCD, where highly non-trivial coarse-grainings are expected to emerge.

Applications of this technique to other domains are also of great interest– for example, long-distance correlations play an important role in natural image classification, and a promising direction is to apply this technique to *unlabeled* natural image data to hierarchically extract the features that carry this long-distance information. The hierarchical extraction of correlations is also an important issue for developing generative models that accurately capture multi-scale correlations, and our technique for iteratively extracting hierarchical coarse-grainings of data may be useful to train generative models factored over multiple scales to address this difficulty.

References

[1] Eytan Bakshy, Lili Dworkin, Brian Karrer, Konstantin Kashin, Benjamin Letham, Ashwin Murthy, and Shaun Singh. AE: A domain-agnostic platform for adaptive experimentation.

[2] Yoshua Bengio. Learning deep architectures for AI.

[3] Cédric Bény. Deep learning and the renormalization group. *arXiv e-prints*, page arXiv:1301.3124, Jan 2013.

[4] Juan Carrasquilla and Roger G. Melko. Machine learning phases of matter. *Nature Physics*, 13(5):431–434, Feb 2017.

[5] Alejandra Castro, Matthias R. Gaberdiel, Thomas Hartman, Alexander Maloney, and Roberto Volpato. The Gravity Dual of the Ising Model. *Phys. Rev.*, D85:024032, 2012.

[6] Ellen de Mello Koch, Robert de Mello Koch, and Ling Cheng. Is Deep Learning an RG Flow? *arXiv e-prints*, page arXiv:1906.05212, Jun 2019.

[7] Stavros Efthymiou, Matthew J. S. Beach, and Roger G. Melko. Super-resolving the Ising model with convolutional neural networks. *arXiv e-prints*, page arXiv:1810.02372, Oct 2018.

[8] Shotaro Shiba Funai and Dimitrios Giataganas. Thermodynamics and Feature Extraction by Machine Learning. 2018.

[9] Wen-Cong Gan and Fu-Wen Shu. Holography as deep learning. *Int. J. Mod. Phys.*, D26(12):1743020, 2017.

[10] Alexander Ganchev. On bulk/boundary duality and deep networks. *AIP Conference Proceedings*, 2075(1):100002, 2019.

[11] Koji Hashimoto. AdS/CFT correspondence as a deep Boltzmann machine. *Phys. Rev.*, D99(10):106017, 2019.

[12] Koji Hashimoto, Sotaro Sugishita, Akinori Tanaka, and Akio Tomiya. Deep Learning and Holographic QCD. *Phys. Rev.*, D98(10):106014, 2018.

[13] Koji Hashimoto, Sotaro Sugishita, Akinori Tanaka, and Akio Tomiya. Deep learning and the AdS/CFT correspondence. *Phys. Rev.*, D98(4):046019, 2018.

[14] Eric Howard. Holographic Renormalization with Machine learning. *arXiv e-prints*, page arXiv:1803.11056, Mar 2018.

[15] Hong-Ye Hu, Shuo-Hui Li, Lei Wang, and Yi-Zhuang You. Machine Learning Holographic Mapping by Neural Network Renormalization Group. 2019.

[16] Satoshi Iso, Shotaro Shiba, and Sumito Yokoo. Scale-invariant Feature Extraction of Neural Network and Renormalization Group Flow. *Phys. Rev.*, E97(5):053304, 2018.

[17] Leo P. Kadanoff. Scaling laws for ising models near T_c. *Physics Physique Fizika*, 2:263–272, Jun 1966.

[18] Maciej Koch-Janusz and Zohar Ringel. Mutual information, neural networks and the renormalization group. *Nature Physics*, 14(6):578–582, Mar 2018.

[19] Patrick M. Lenggenhager, Zohar Ringel, Sebastian D. Huber, and Maciej Koch-Janusz. Optimal Renormalization Group Transformation from Information Theory. *arXiv e-prints*, page arXiv:1809.09632, Sep 2018.

[20] Shuo-Hui Li and Lei Wang. Neural Network Renormalization Group. , 121(26):260601, Dec 2018.

[21] Henry W. Lin, Max Tegmark, and David Rolnick. Why Does Deep and Cheap Learning Work So Well? *Journal of Statistical Physics*, 168(6):1223–1247, Sep 2017.

[22] Pankaj Mehta and David J. Schwab. An exact mapping between the Variational Renormalization Group and Deep Learning. *arXiv e-prints*, page arXiv:1410.3831, Oct 2014.

[23] E. Miles Stoudenmire and David J. Schwab. Supervised Learning with Quantum-Inspired Tensor Networks. *arXiv e-prints*, page arXiv:1605.05775, May 2016.

[24] David J. Schwab and Pankaj Mehta. Comment on "Why does deep and cheap learning work so well?" [arXiv:1608.08225]. *arXiv e-prints*, page arXiv:1609.03541, Sep 2016.

[25] Brian Swingle. Entanglement Renormalization and Holography. *Phys. Rev.*, D86:065007, 2012.

[26] Jing Tan and Chong-Bin Chen. Deep learning the holographic black hole with charge. 2019.

[27] Evert P. L. van Nieuwenburg, Ye-Hua Liu, and Sebastian D. Huber. Learning phase transitions by confusion. *Nature Physics*, 13(5):435–439, Feb 2017.

[28] Lei Wang. Discovering phase transitions with unsupervised learning. *Phys. Rev. B*, 94:195105, Nov 2016.