Deep Learning Estimation of Critical Temperature of Superconductors: Toward the Search for New Materials

Tomohiko Konno*
National Institute of Information and Communications Technology (NICT), Tokyo Japan

Hodaka Kurokawa,† Fuyuki Nabeshima, and Ryo Ogawa
The University of Tokyo, Tokyo Japan

Michiaki Iwazume and Iwao Hosako
NICT

Atsutaka Maeda
The University of Tokyo

High-temperature superconductors have a lot of promising applications: quantum computers, high-performance classical computers, green energy, safe and high-speed transportation system, medical appliance, and etc. However, to discover new superconductors is very difficult. It is said that only 3% of candidate materials become superconductors. We do not have satisfactory theory of high-temperature superconductors yet, and computational methods do not work either. On the other hand, the data has accumulated. Deep learning suits the situation. We introduce following two methods into deep learning: (1) in order to better represent material information and make good use of our scientific knowledge into deep learning, we introduce the input data form representing "material as image from periodic table with four channels corresponding to s, p, d, and f of electron orbit"; (2) we introduce the method, named "garbage-in", to make use of the non-annotated data, which means the data of materials without any critical temperature T_c in this case.

We show that the critical temperature is well predicted by deep learning. In order to test the ability of the deep neural network it is good if we have the list of materials, which it is hard to tell whether they become superconductors or not from human expert in advance. We use the list made by Hosono's group as such. The test is to predict if a material becomes superconductor beyond 10 Kelvin or not, and the neural network is trained by the data before the list was made. The neural network achieved accuracy 94%, precision 74%, recall 62%, and f1 score 67%. Compared to the baseline precision 9%, which is obtained from a positive prediction to randomly selected material, our deep learning has very good precision. We make a list of candidate materials of superconductors, and we are preparing for the experiments now.

I. INTRODUCTION

Our study is at the intersection between superconductors and deep learning. High-temperature superconductors have a lot of promising applications: quantum computers, high-performance classical computers, green energy, safe and high-speed transportation system, medical appliance, and etc. However, to discover a new superconductor is very difficult. In the paper [1], it was said that only 3% of candidate materials becomes superconductors. Also, it took about ten years until new type of superconductors was discovered. We think deep learning can accelerate the cycle.

Deep learning has achieved great progress. It is sometimes wrongly believed that deep learning can do everything. However, if we have a theory on a subject, deep learning is not necessary. The theory works better than deep learning. In order for deep learning to work, we have the following two conditions; (1) we do not have a theory on the subject; (2) we have enough data. In view of the study of deep learning, we must seek new fields where deep learning works to make use of its power. The ability has been proved in many fields like computer vision, natural language processing, reinforcement learning, and so forth.

If we focus on material science, material data has been accumulated not only for superconductors as summarized in table I. They are useful for material deep learning. Of them, we focus on high-temperature superconductors for the following reasons. High-temperature superconductors have a lot of promising applications as raised in the beginning of the introduction. Therefore, the search for new high temperature superconductors always has been tried extensively. Bardeen-Cooper Schrieffer (BCS) theory explained the mechanism of conventional superconductivity, and initial approaches rely on the prediction of the theory. Along this line, 260 K superconductivity is realized under very high pressure. After the discovery of high-temperature cuprate superconductor, many different semi-empirical approaches were introduced. The most typical one is utilizing the electronic structure calculation by the 1st principle one, provided some crystal structures are gives. However, many of recent superconductors having rather high Tc belong to unconventional category,

* Corresponding author: tomohiko@nict.go.jp
† Second Corresponding author: scottie0018@gmail.com

Database	API	Data Size	Data Provider	Free
Materials Project	✓	70,000		✓
Aflow	✓	1,820,000		✓
MPDS	✓	400,000	Materials Phases data Systems	×
COD		400,000		✓
SuperCon	×	30,000	NIMS	✓
Atomwork	×	80,000	NIMS	✓
ICSD		180,000	FIZ Karlsruhe-NIST	×

TABLE I. Databases for Material Machine Learning

where many body effects play an essential role, it is becoming rather hard to predict new superconductors by these approaches, and rather different kinds of approaches are longed for.

Because of the hardship to discover new superconductors, the lack of satisfactory theory of high-temperature superconductors, abundance of data, and many promising applications, high temperature superconductor is a good theme for deep learning. Here, we introduce a novel approach using deep learning, and succeeded in making a new list of candidate materials of superconductors, which we are currently doing experiments.

Conventional machine learning, by which we mean before deep learning, for superconductors was studied by random forest in a preceding paper [2]. Compared to random forest, the downside of deep learning is its interpretability. In random forest, we know how and what factors contribute. Deep learning does not tell us them. However, all of the recent progress in so-called A.I. and under the name of machine learning has been achieved by deep learning. If deep learning works for discovering superconductors, we expect it really works. Hence, we study deep learning of superconductors in order to find new superconductors.

II. THE CONTRIBUTION

We have the following contributions. The fist two are on the deep learning research itself, and the last two are on the applications of deep learning to superconductors.

- We introduce the input data form that represents "material as images from periodic table with four channels corresponding to s, p, d, and f of electron orbits". This is in an attempt to capture the similarities between elements from limited number of data, and, in particular, to predict new materials out of data. To go beyond the data is an essential problem in deep learning.

- We introduce the semi-supervised method, named "garbage-in", that makes use of non-annotated data in training. In the paper, non-annotated means that the data does not have any T_c at all.

- We predict critical temperature for superconductors by deep learning for the first time to the best of our knowledge.

- We predict if the critical temperature is higher than specific value or not. We used the list of materials in which human expert cannot tell if a material becomes a superconductor or not before the experiment.

III. THE INPUT DATA REPRESENTATION

We will discuss the input data form representing "material as images from periodic table with four channels corresponding to s, p, d, and f of electron orbits"

A. Representing elements of materials as one-hot vectors

Any material consists of elements. We use 118 elements in this paper. Any element can be represented by one-hot vector in the following way. For example, the element He is represented by 118-dimensional vector $(0, 1, 0, \cdots, 0)$, and H is represented by the vector $(1, 0, \cdots, 0)$, and so forth. If we have (unreal) material H_2He_3, then this material is represented by $(2, 3, 0, \cdots, 0)$ or $(2/5, 3/5, 0, \cdots, 0)$. The former is referred to as absolute rep. and the latter is referred to as relative rep, respectively in this paper.

B. Problems in representing the elements by one-hot vectors

The problem in representing elements by one-hot vectors are following two-fold. (1) neural networks cannot learn and tell anything about elements and the combinations that do not appear in the training data. (2) The one-hot reps. cannot capture the similarities between elements, especially if the data is not ample. In principle, any element is treated as quite different entity in one-hot reps, although we know the similarities among elements from quantum mechanics.

C. Material as images from periodic table

To overcome the problems in one-hot rep. of material, we introduce "material as images in periodic table". Basic properties in elements are determined by the number of electrons and their orbits. They are represented in periodic table. Hence, we represent elements and materials as images in periodic tables. The size of periodic table is, 32×7. The data format is with 4 channels corresponding to s, p, d, and f blocks as we will explain in the following section. Although one-hot reps. cannot learn anything about the elements and the combinations that do not present in the training data, the data form "material as image form periodic table" can learn even such elements and combinations because of the way of learning by convolutional neural networks (CNN). CNN has small number of parameters in its filter. If the parameters are determined by training, they can be used for other materials even if they do not appear in the training. We have not included table of nuclides yet, which also determines the basic properties of elements. To do it is not very difficult, and is our future work.

As of now, we do not still have something like word2vec [3] for the elements and the combinations. The word2vec is a distributed representation of words. We can say that our material as image in periodic table is the word2vec in material since periodic table has very rich information about elements from quantum mechanics. Even if we have better distributed rep. of elements like word2vec in the future, we think to start from our data format is better than from reading materials as words. It is often said that we do not have to have feature engineering in deep learning, and deep learning from data without man-made features is better than that with man-made features. Rather, it is often said that man-made feature engineering is harmful. However, it is a quite different story if we have scientific knowledge on the subject. It is waste if we do not include scientific knowledge in the data and the deep neural network structure. We introduce our method, material as image in periodic table, in order to use it. SuperCon has about $13,000$ data with critical temperature. In view of deep learning, the data size is small. It is said that we need tens of millions of images for computer vision, and that Google has data set named JFT that has 300 million images. We can have deep neural network learn superconductors, because the data includes scientific knowledge, which reduces the required data size.

D. Channels as electron orbits: s-block, p-block, d-block, and f-block

In computer vision, images are divided into three channels: red, green, and blue. We have the images representing materials in periodic table divided into four channels that correspond to s-block, p-block, d-block, and f-block of electron orbits.

As we already know from quantum mechanics that electrons reside different orbits, and elements are classified according to which orbit the out-most electron resides. If the out-most electron resides s orbit, then the element is classified as in s-block. We also have p-block, d-block, f-block in the same manner. In view of physics, we know that the difference in the blocks plays very significant role. We must reflect the scientific knowledge to deep learning.

E. Relative and absolute reps. of the numbers in composition formula

As is explained in one-hot reps. in Sec. III A, the numbers in the composition formula is reflected in the data as either of absolute values or relative values. The periodic table and its blocks are shown in Fig. 1. A schematic explanation of our data format is given by Fig. 2.

IV. PREDICTION OF THE CRITICAL TEMPERATURE OF SUPERCONDUCTORS

A. The Three Dataset: SuperCon, COD, and Hosono's Group Data

We use three datasets in this paper. The one is the list of superconductors named "SuperCon". The list is made by National Institute of Materials Science [4]. The data has about $13,000$ superconductors with critical temperatures T_c. We used the data with critical temperature only. Since it is hard to tell for us that whether the data without T_c is the material that does not transit to superconductor or just the critical temperature is not recorded. The other problem is that we cannot obtain spatial structures from SuperCon, so that we use only composition formula as of now. To include the spatial structure in the data is next problem. The second dataset is Crystallography Open Database (COD) [5–7]. It has about $400,000$ materials. We select non-organic materials, delete duplication, and delete overlap with other two datasets. Then, about $47,000$ materials remain. We do not get any information about the critical temperature from COD. This is a list of materials. We will use the dataset in order to search for new superconducting materials also. The third dataset is the list made by Hosono's group. The list has about 300 superconductors and non-superconductors, both of which were expected to transit to superconductors before experiments. As is with COD, we deleted overlap with SuperCon and COD from Hosono's group data.

B. The Neural Networks

We use a convolutional neural network (CNN) for the data format of images from periodic table with four channels corresponding to s-block, p-block, d-block, and f-block. The numbers in composition formula are in relative reps. in the input. In the training, the task is to predict

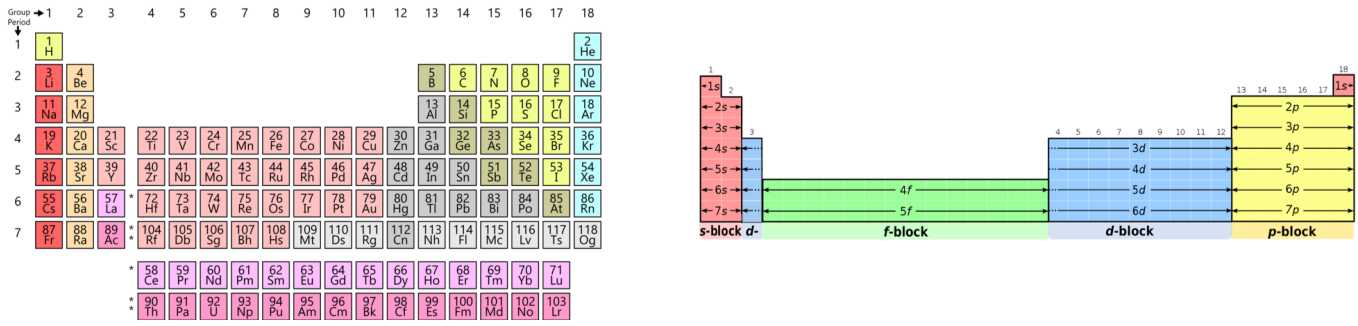

FIG. 1. **The periodic table and its blocks; s-block, p-block, d-block, and f-block** from Wikipedia

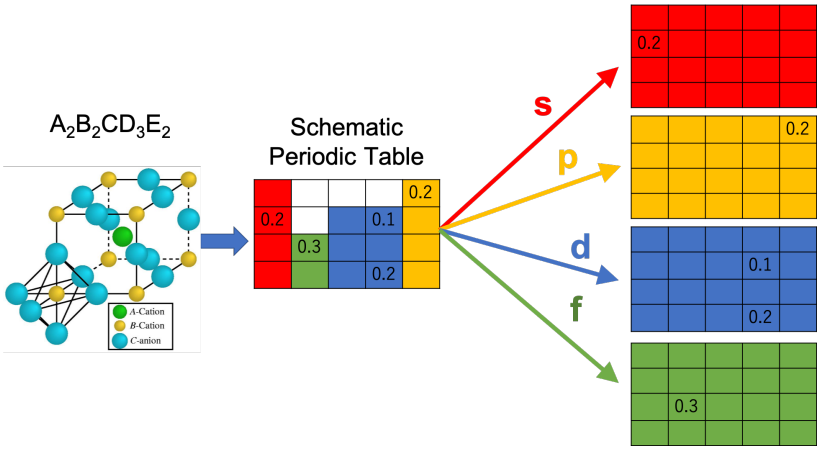

FIG. 2. **A scheme of input data form of material as image from periodic table with four channels corresponding to s, p, d, and f.** Suppose we have a hypothetical material, and the composition formula of which is given by $A_2B_2CD_3E_2$. The relative value of the number of each element in the composition formula is written on each place of element in (schematic) periodic table. The 0.2 in the right-most is for A, the 0.2 in the left-most is for B, the 0.1 in the upper d-block is for C, the 0.3 in the f-block is for D, and the 0.2 in the lower d-block is for E. In the figure, the values are in relative representation, but the absolute representation is also possible. Empty boxes are filled with 0. The original (schematic) periodic table will be divided into four channels. The four channels correspond to s, p, d, and f of electron orbit. In the actual case, the periodic table is 32×7, so one material is represented by the numbers in the form of $4 \times 32 \times 7$. The material image in the figure is taken from Wikipedia.

the critical temperature. We have either take *log* or not on the critical temperature. If we have *log* on the critical temperature, we have another hyper parameters ϵ ranging from 10^{-1} to 10^{-5} added to the temperature.We observed the relative representation of values in composition formula works better than absolute representation in CNN with period table as images. We suppose this is because the input values increases if the number of elements in composition formula increases in absolute representation. On the other hand, it seems that absolute representation works better in one-hot reps. in training and testing within SuperCon. In the following results, we use CNN with the input of material data as images from periodic table.

C. The First Step: Training and Predicting The Critical Temperature in SuperCon

In the first step, we split SuperCon dataset 95% for training and 5% for test. We do not see big difference if we take *log* on the temperature or not. The scatter plots of predicted critical temperature and true Tc in the test data is presented in Fig. 3. The achieved R^2 of the critical temperature in the test data is 0.93. It is very high. Note that we do not classify the materials into copper oxide, iron-based, nor conventional in advance. We do not classify the materials according to the critical temperature either. Contrary to these facts, our deep learning works very well. When it comes to new material search, we think it is better to include as much data as possible in the training. This is the reason why we do not classify before hand.

In addition, we have another task. It is to have deep neural network predict if the critical temperature is larger than 20 Kelvin or not in the test data. Only about 2%

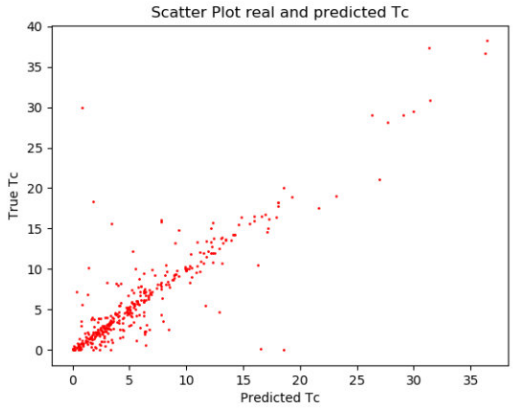

FIG. 3. The scatter plots of predicted critical temperature and true Tc in the test data. The achieved R^2 is 0.93.

of the materials has more than 20 K. The explanation on accuracy, precision, recall, and f1 score is given in Appendix A for unfamiliar readers. For the task, the achieved accuracy is 95%, the precision is 92%, the recall is 94%, and the f1 score is 93%. The result is summarized in the table II. We also have surprisingly good result. Prediction that a randomly selected material has more than 20 K results in about only 2% precision.

Accuracy	Precision	Recall	F1 score
95%	92%	94%	93%

TABLE II. The achieved accuracy, precision, recall, and f1 score for the prediction if the critical temperature is beyond 20 Kelvin or not in the test data in SuperCon.

D. The Second Step: Training Deep Neural Network with SuperCon and COD, and Predicting if T_C is Larger than 10 Kelvin in Hosono's Group Data.

In order to test the deep neural network, it is good if we have the list of materials which human expert cannot tell if the material transit to superconductors or not in advance. We use Hosono's group data as such. In the second step, we try to predict if the critical temperature is higher than 10 Kelvin or not in the list.

1. *The Problem in Training with SuperCon only*

However, if we train deep neural network with SuperCon only, then we have a problem. The problem arises from the fact that SuperCon has only the materials that become superconductors. SuperCon does not have the materials that do not become superconductors. The training task is to predict the critical temperature given that a material becomes superconductor. On the other hand, what we do in the test on Hosono's group data is to predict if a material becomes superconductor (beyond some temperature) or not. They are different tasks. If we try to predict the critical temperature by the deep neural networks trained with SuperCon only, then the predicted critical temperatures are so much high. If it were true, almost half of the materials would become superconductors in Hosono's group data.

2. *"Garbage-in": Usage of Material Data Without Any Critical Temperature*

The problem is that we do not have the data of materials that do not become superconductors. This is called publication bias. Only successful data are reported. In order to overcome the problem, we use COD into training data with SuperCon, though COD does not tell us any T_c. We input material data in COD with supposing as if all critical temperatures were 0 Kelvin. We suppose that most of the materials do not transit to superconductors with finite temperature, which seems reasonable assumption. Although it is very dare to do so, we will see it works as follows. This is the second main contribution of our paper in terms of deep learning method. The lack of data is always the problem, and to collect data is highly costly. We saw that training sometimes fails if *log* is not taken on the critical temperature when we train with SuperCon and COD. A scheme of "garbage-in" is shown in Fig. 4

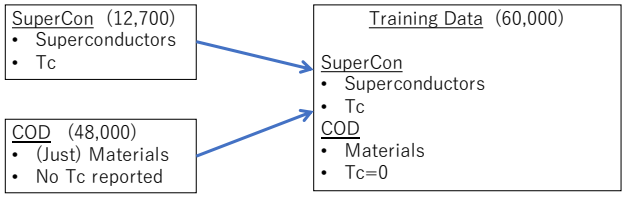

FIG. 4. **A scheme of "garbage-in".** COD is not the database for superconductors.

3. *The Result of Prediction If The Critical Temperature is Beyond 10 K or Not in Hosono's Group Data*

The list was made by the project starting from May 2010. Hence, we used the data from SuperCon before 2010 only. About 9% of materials in the list become superconductors beyond 10 Kelvin. It means that it is the baseline score. The baseline precision is 9%. This is the score you get if you say that a material transits beyond 10 K given that the material is randomly selected from the list. The list after the deletion of overlaps with SuperCon and COD has 291 materials left, and 27 of them have more than (or equal to) 10K. (One material has just 10K.) Our deep neural network achieved accuracy 94%, precision 74%, recall 62%, and f1 score 67%. Compared to the baseline precision of 9%, the precision by deep convolutional neural network

of 74% is surprisingly good. The result is summarized in table IV D 3. The blanks mean the scores cannot be defined without further assumption. We think when it comes to the search for new materials, precision is more important since it takes time to test in the experiment. *The data without Fe*: we delete Fe-based materials from the list, then 245 materials remain. Only five materials transit beyond 10K. Thus we changed the threshold to 1K, and 54 materials transit superconductors beyond 1K. The baseline precision is 22%. Our deep learning has achieved accuracy 80%, precision 65%, recall 24%, and f1 score 35% for the dataset without Fe-based materials. We make a new list of candidate materials of superconductors, and we are currently doing experiments as the third step.

	Accuracy	Precision	Recall	F1 score
Baseline		9%		
Our Deep Learning	94%	74%	62%	67%

TABLE III. The scores in predicting if the critical temperature is beyond 10 Kelvin in Hosono's group data by the deep neural network trained not only with SuperCon but also with COD that has no T_c reported. The baseline is obtained from a positive prediction to randomly selected material from the list.

V. CONCLUDING REMARKS

Deep neural networks, the input data form representing materials as images from periodic table with four channels corresponding to s-block, p-block, d-block, and f-block of electron orbits, and the method named "garbage-in" to make use of non-annotated data, which means only material data without any critical temperature in this case, worked surprisingly well in the two tasks. We succeeded in making a new list of candidate materials of superconductors, which we are currently doing experiments. Crystal structure is not uniquely determined by chemical composite only. We have not included crystal structure in the data yet in this paper because we cannot obtain them from SuperCon. To include the spatial strictures into the data is what we must do if we push for the research project. Also, we have not included table of nuclides yet.

We suppose the input data form of material as image from periodic table will work when we leave the data behind, since our data form can capture the similarities between the elements and the combinations by construction. Also, to read from periodic table is what human expert do. Our input data form allows for telling something about the elements and the combination that do not appear in the training data. To go beyond the data is an essential problem in deep learning not only limited to material informatics. We think the method will work in bioinformatics, drug discovery, material design in both organic and non-organic, and any field related to the elements. Also, the usage of non-annotated data is another contribution in terms of deep learning.

Appendix A: On Accuracy, Precision, Recall, and F1

We are going to explain accuracy, precision, recall, and f1 score quickly for unfamiliar readers. Let's say we have the data of 10,000 samples, and 100 out them have some disease. If the sample has the disease, it is said it is positive, and negative vice versa. The task is to predict if a given sample is positive or not. Accuracy is the rate whether the estimation is right irrespective of the estimate is positive or negative. Precision is how much percentage of the estimates of the positive is correct. If you say that the sample is positive to only one sample which is very obviously positive to anyone, and say that all the others are negative, then you get 100% precision, but you missed the rest of 99 positive samples, which is a problem. To measure such is recall. Recall is how much percentage you find positive samples from all true positive samples. If you say positive to all the10,000 samples, you will get 100 % recall because you find all the 100 positives, but you have only 1% accuracy and precision. So, which of precision or recall is better measure? How do we evaluate the classifier? In an attempt to answer the question, we have f1 score. The f1 score is a kind of mean between precision and recall.

[1] H. Hosono, K. Tanabe, E. Takayama-Muromachi, H. Kageyama, S. Yamanaka, H. Kumakura, M. Nohara, H. Hiramatsu, and S. Fujitsu, Science and Technology of Advanced Materials **16**, 033503 (2015).

[2] V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo, and I. Takeuchi, npj Computational Materials **4**, 29 (2018).

[3] T. Mikolov, K. Chen, G. Corrado, and J. Dean, arXiv preprint arXiv:1301.3781 (2013).

[4] N. I. of Materials Science, .

[5] S. Gražulis, A. Daškevič, A. Merkys, D. Chateigner, L. Lutterotti, M. Quiros, N. R. Serebryanaya, P. Moeck, R. T. Downs, and A. Le Bail, Nucleic acids research **40**, D420 (2011).

[6] S. Grazulis, D. Chateigner, R. T. Downs, A. Yokochi, M. Quirós, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck, and A. Le Bail, Journal of Applied Crystallography **42**, 726 (2009).

[7] R. T. Downs and M. Hall-Wallace, American Mineralogist **88**, 247 (2003).