

Article

Transfer Learning in Inorganic Compounds' Crystal Structure Classification

Hanan Ahmed Hosni Mahmoud 

Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; hahosni@pnu.edu.sa

Abstract: Deep learning consists of deep convolutional layers and an unsupervised feature selection phase. The feature selection of deep learning on a large size dataset can be employed in correlated prediction models with small size datasets. This methodology is titled deep transfer learning model and enhances prediction model generalization. In this research, we proposed a prediction model for the crystal structure classification of inorganic compounds. Deep learning models in structure classification are usually trained using a large size dataset of 300 K compounds from different quantum compounds dataset (DS1). The feature selection of the deep learning models is reused for selecting features in a small size dataset (with 30 K inorganic compounds and containing 150 different crystal structures) and three alloy classes. The selected features are then fed into a random decision forest prediction model as input. The proposed convolutional neural network (CNN) with transfer learning realizes an accuracy of 98.5%. The experiment results display the CPU time consumed by our model, comparing the time required by similar models. The CPU classification time of the proposed model is 21 s on average.

Keywords: deep learning; transfer learning; small datasets; crystal structures



Citation: Mahmoud, H.A.H. Transfer Learning in Inorganic Compounds' Crystal Structure Classification. *Crystals* **2023**, *13*, 87. <https://doi.org/10.3390/cryst13010087>

Academic Editor: Zongyou Yin

Received: 20 October 2022

Revised: 18 November 2022

Accepted: 7 December 2022

Published: 2 January 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Deep Learning and Transfer Learning

The fourth paradigm of science has emerged, and it is data-driven [1]. In research, deep learning has grown to be a crucial addition to experiments, theory, and simulation [2–5]. Deep learning has recently become a popular tool in substance research and development for forecasting the intricate connections between a crystal's substances, processes, structure, and features [6–9].

Typically, structured data is dealt with using conventional deep learning, such as support map deeps, random decision forests, and neural networks. It is usual for the effectiveness of traditional deep learning models to be influenced by the caliber of manually created features (descriptors) and the size of datasets [10–14]. When creating new substances, nevertheless, it can be difficult to find sufficient data and features that are well-built [15,16]. As a result, there are many circumstances where standard deep learning cannot be applied.

Unstructured data (such as images and sounds) can be immediately entered into a deep learning model without the need for manual feature engineering [17,18]. Deep learning combines an automatic selection feature and has been extensively employed in image-related applications, such as defect detection and microstructure recognition [19–21].

Transfer learning could be a valuable deep learning procedure that empowers knowledge sharing between models from related spaces [22–26]. Deep learning has special interest in realizing transfer learning because of its chain of command structure: convolution layers (i.e., programmed, including the selection phase) extricate highlights from the low to the high level in the arrangement. A portion of extricated highlights is common and transferable and has 15 million pictures and over 20 thousand categorical names), as

opposed to building a modern model from scratch utilizing an irregular start. Moreover, the selection of a well-trained highlight from an expansive dataset empowers the exactness of forecast from a small-size dataset that would otherwise be troublesome to realize [27–30].

In any case, deep learning models coupled with exchange learning have not been sufficiently used in fabric investigations. Routine deep learning models are still the most used instruments in understanding fabric-related issues. A critical step in applying customary deep learning to substance investigation is to construct a set of manually highlighted features. This is because substance inquiry involves information on crystals, forms, and fabric features, which comprise large, organized information, and datasets in substance inquiry are usually small-sized (deep learning commonly require large size dataset). However, when we outline organized information to 2D pseudo-images, deep learning can be used to extricate highlights [31–34]. Hence, the extreme and dubious highlight designing can be dodged with deep learning, and the troubles with restricted-size benchmarks might be somewhat overcome by exchange learning.

1.2. Classification of Inorganic Crystal Substances

Crystal patterns in any substance can identify the features of any substance. Classification of crystals and phase formations of any substance are fundamental issues facing substance detection. Substance density theory and force field computation are employed to identify and classify crystals. Nevertheless, their utilizations are hindered by CPU time and computational load [35]. Classification of the crystal substance from the first principle requires high-accuracy computations of entropies for thousands of putative crystals [33–36]. Many problems still occur in substance density theory, such as computation for composite compounds with high entropy and principal compounds [35–37]. Deep learning models are training-driven models which are highly effective compared with statistical methods. To speed up new structures' classification, they can be employed as another model for simulation [35–40]. Deep learning methods [40–44] can predict crystal substance structures. Deep models are used to predict entropy in substances [41–46].

Different deep learning models for crystal classification are depicted in Table 1.

Table 1. Summary of different machine learning and deep learning models to detect the crystal structures in different datasets.

Reference	Model	Dataset	Implementation	Training Time	Classification Time (Seconds)	Accuracy %	Limitation
[15]	Crystal classification using support vector machines	Crystal images	Support vector machines	32 h	127 s	77.5%	Low precision images yielded a high false positive rate
[16]	Crystal structure classification using neural learning	Infrared crystal photos	Convolution network trained with	56 h	212 s	84.4%	Lengthy training time
[17]	Classification of crystal structure using region of interest	Inorganic image dataset	Statistical study	No training process; it is not machine learning	Stochastic process	81%	Small dataset
[18]	Feature fusion	75,000 incidences of inorganic compound crystallization and organic crystals	Decision tree	120 h	220 s	88.2%	Binary classification (inorganic crystal or organic crystal)
[19]	Deep learning CNN model	Inorganic crystal structure data	Deep CNN	79 h	170 s	87.3%	Unbalanced dataset
[20]	3D CNN	3D crystal images	3D deep learning model	240 h	90 s	90.9%	Long training time
[21]	A crystal structure classification intelligent model	Induced dataset	Artificial intelligence method	25 h (small-size datasets)	78 s	77.4%	Low accuracy because of the size of the dataset

Table 1. Cont.

Reference	Model	Dataset	Implementation	Training Time	Classification Time (Seconds)	Accuracy %	Limitation
[22]	Crystal classification using image segmentation	Inorganic crystal dataset	Encoder-decoder model	120 h	160 s	90.6%	Training time increases with the data size
[23]	Crystal detection in videos	Crystal in video frames	Object recognition model	Feature mining	400 s	89.4%	Lengthy recognition time

In this research, we propose a crystal structure classifier that employs transfer learning to predict different compounds of inorganic substances without any prior knowledge.

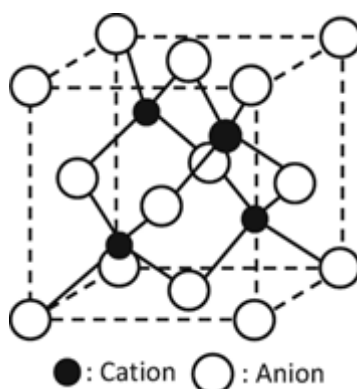
2. Materials and Methods

2.1. Crystal Structure Mapping

Deep learning models such as CNN require one or two-dimensional data maps as the training phase input. They are fed 2D images for training and prediction. Visual features and associations among various 1 or 2 values are computed via the feature selection process, which comprises several convolutional and pooling layers followed by the rectified linear activation function (ReLU). The authors in [30–35] established crystal substance mapping to two-dimensional maps, using 2D maps and atom representation, which allows the CNN to learn from composite data. In the proposed research, we represent the crystal structure of inorganic substances in 2D maps with cation and anion structures. Figure 1a depicts the proposed mapping procedure. A crystal structure is uniquely mapped to a 2D matrix. Each cell of the matrix corresponds to a cation or anion (the cation is 1, and the anion is denoted by 2). The information on the crystal is denoted by a value 1 or 2 located at the same site of the crystal cation or anion in the 2D map (e.g., the information on an example crystal is defined by the 1 or 2 value in its corresponding cell in the matrix. Other cells are occupied by zeros. Figure 1a shows the 2D representation of the crystal compound. Zero values represent the empty areas in the 2D map of a crystal.

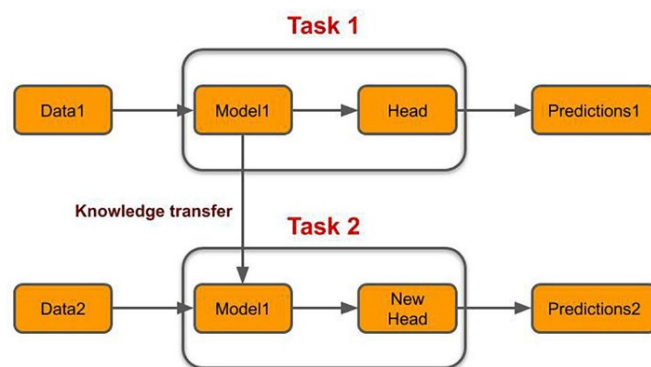
2.2. Deep Learning Models

Many CNN networks, such as AlexNet and GoogLeNet, are well-established. Nevertheless, those networks have to go through layer reduction. This reduction decreases the overfitting problem (the DS1 [38] set utilized in this research is 2% size of ImageNet). A CNN network is utilized in this research because of its feature extraction capability. The used CNN has three modules. The first is a transferable feature selection module using the first five convolutional layers, and the second module involves the ReLU activation function, comprising a nonlinear function $f(x) = \max(0, x)$ and subsequent max pooling. The third module is the regression process with a fully connected (FC) and a ReLU. The CNN has a 5 * 5 convolutional kernel and 3 * 3, as depicted in detail in Table 2.



(a)

Transfer Learning



(b)

Figure 1. (a) 2D mapping of the crystal structure of a substance. (b) The transfer learning model is trained to extract the transferable features using the open quantum compounds dataset.

Table 2. Proposed CNN model layers and hyperparameters.

Layer #	Layer Name	Filter	Activation Nonlinear Function
1	Input	$128 \times 128 \times 3$	-
2	CL	$256/5 \times 5$	-
3	Max pooling	3×3	ReLU Activation function
4	CL	$64/3 \times 3$	-
5	Max Pooling	2×2	ReLU Activation function
6	Dropout layer	0.45	-
8	Parameter	32	ReLU Activation function
9	CL	$32/3 \times 3$	-
10	Dropout layer	0.3	-
11	Classifier	-	Softmax
12	Output	-	Crystal class

Zero-padding is utilized for the data compounds in the CL by embedding zeros at the boundaries to attain as much data as possible. Filters are induced during the learning phase. The proposed CNN has 17,281 parameters, only a fraction of the other neural networks, such as VGG-16, which possesses 120 million parameters. This can hinder the overfitting problem.

2.3. Transfer Learning

Figure 1b depicts the flow diagram of transfer learning. Two input datasets are utilized, i.e., the large-sized input dataset DS1 [16] and the small-sized target dataset (DS2 dataset in our research). The DS1 dataset is utilized for training the CNN to attain the transfer learning features. The transfer learning features are produced feature maps employed for the DS2 dataset. Prediction is performed on the produced maps by a shallow, random decision forest technique. The feature selection process employs the target DS2. The training process of the new neural model from scratch using new target datasets is not compulsory. The new classifiers alone need to be constructed and trained.

2.4. Training Process

CNNs are constructed and tested by applying the open source Python Keras with Tensorflow libraries for employing neural machine learning and execution of deep learning as the backend [39]. Here, 70% of the dataset is utilized for training the CNN, 15% for validation, and 15% for testing. In the training process, the final output of the CNN is compared to the ground truth. The mean square error (MAE) is used as the mean error displacement to compute the fitness. The number of epochs is fixed to 1800 (where loss function values converge).

Random decision forest is embedded in the Python and mat lab Scikit library [39]. Random decision forest is utilized in transfer learning for the classification of the DS2 dataset because of its hyperparameter minimization to the minimal. We controlled the highest number of decision trees in the random decision forest (200 decision trees in our research) with a depth of a decision tree equal to that in our research. Random data splitting in a circular way is employed in the used dataset to guarantee that training and validation subsets have comparable distribution.

The proposed model is trained on a 3.6 GHz i7 CPU and 32 G RAM. The CNN is trained employing random initial parameters on the DS1 dataset with 300 K compounds for 13 h, with batch size equal to 128 with 2200 epochs. The training process of a random decision forest on our DS2 dataset, which has 30 K compounds and 150 classes, takes 0.5 h. The transfer learning technique gains substantial speed over training deep learning models from scratch.

3. Experimental Results

3.1. Datasets Description

3.1.1. DS1

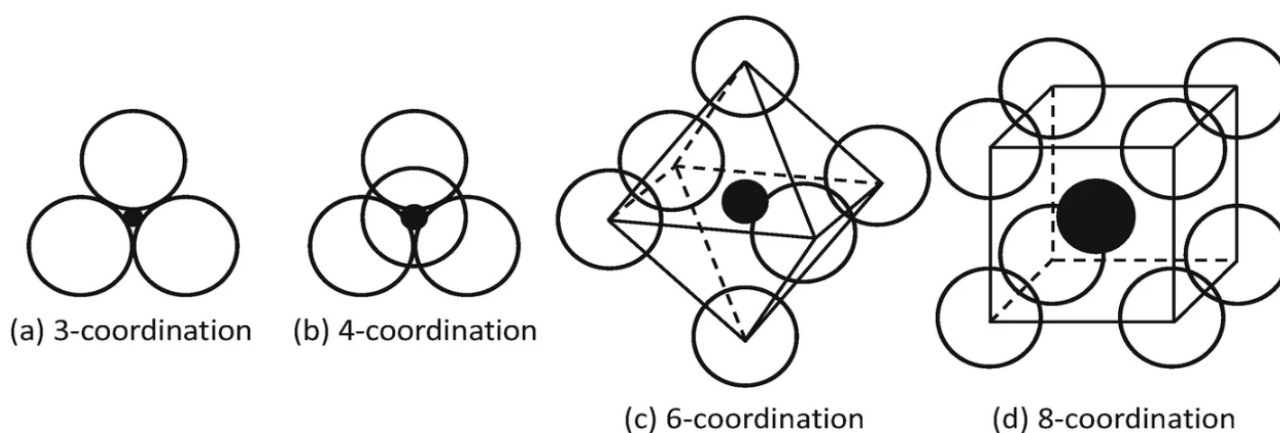
Material science research and practice have made the highest effort in collecting large-sized datasets of substance features, such as the open quantum substances dataset (DS1) [38]. In addition, other datasets include the automatic flow for substances dataset (AFLOW) [40], the substances project [41], and the crystallography dataset (COD) [42]. These large-sized databases can be used as input for transfer learning to produce general, reusable features. In this article, we utilize a dataset of 30 K crystal compounds from DS1. The DS1 dataset comprises compound crystals and their physical features computed by substance density theory. The CNN is trained on the DS1 training subset to identify the formation entropy (EE) and volume (VOL) of these crystals.

In crystals, atoms are circled by several atoms. Coordination is defined by atomic size and bonding axis. Several coordination systems with ionic radius and bonding orbital are depicted in Table 3. Coordination numbers are tri (3), tetra (4), octal (6), and cubic (8). The number 12 defines the close packing arrangements.

Table 3. Configuration of ionic radius.

Coordination	Cation/Anion Radius Ratio (Range)	Configuration
3	0.143–0.223	Tri
4	0.224–0.414	Tetra
6	0.414–0.852	Octal
8	0.852–1.000	Cubic

In crystals, atoms are bounded by atoms in a coordination arrangement which is identified by ionic radius and bonding link. Coordination parameters with ionic radius and surrounding orbit are depicted in Figure 2 and comprise tri, tetra, octal, and cube configuration.

**Figure 2.** Coordination configuration of different ionic crystals.

3.1.2. DS2 Dataset

The DS2 dataset of crystal structures is defined in three features:

- Compounds such as Cu and NaCl;
- Pearson numbers, such as cF4 and cF8;
- Space group symbol.

For compounds, the phase sample of solid alloys, such as $\text{Fe}_{0.4}\text{Co}_{0.2}\text{Ni}_{0.4}$, is Cu, cF4, 225. We chose a dataset with 30 K inorganic substances. Phase samples with more than 30 items are included. The DS2 dataset has 30,000 inorganic materials and comprises 150 phases. The crystal distribution is depicted in Figure 2. Most crystals of inorganic substances exist in the DS2 dataset, and the distribution of crystal existence is nonuniform. Figure 3 depicts the distribution of the compounds in each phase in the dataset. Of 150 compounds, 80 compounds have a frequency of less than 56, and only a few compounds have a frequency higher than 900 (including MgCu_2 , NaCl, and CeAl_2Ga_2). A sample from the 150 compounds is depicted in Figure 3.

There are 14 infrequent compounds with a restricted count (20–40) in our data. Table 4 describes the count of the instances of different compounds.

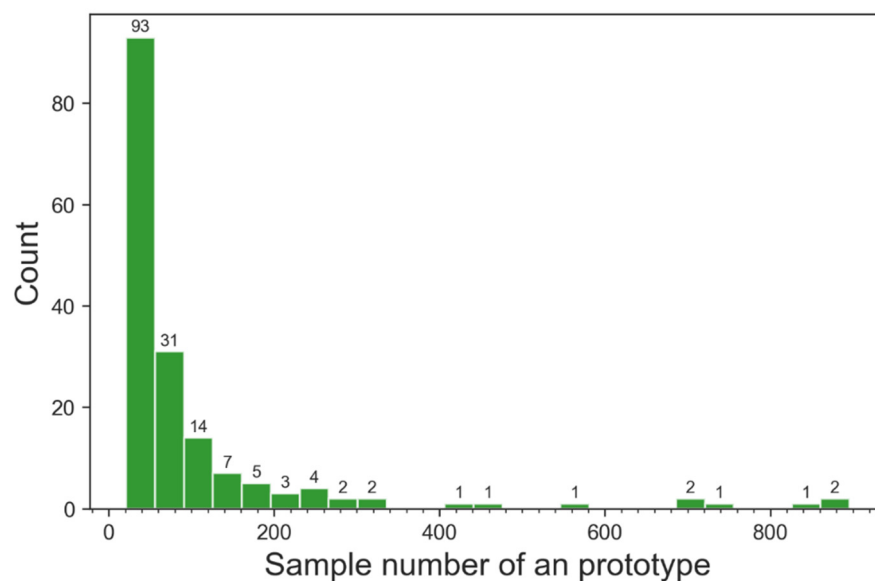


Figure 3. Distribution of each phase compound in the dataset.

Table 4. Frequency of different inorganic compounds.

Compound	Frequency
MgCu ₂ , cF24	227–229
Ca ₂ Nb ₂ O ₇ , cF88	226–231
Cu, cF4	225
ZnS, cF8	216
CaB ₆ , cP7	221
YbFe ₂ Al ₁₀ , oS52	63
Pr ₃ WCl ₃ O ₆ , hP26	176
Y ₄ PdGa ₁₂ , cI34	229
YCo ₅ P ₃ , oP36	62
K ₃ Nb ₈ O ₂₁ , hP64	193
Y ₆ RuI ₁₀ , aP17	22
KGdNb ₆ Cl ₁₈ , hR81	148
KAsF ₆ , hR24	148
Cs ₃ Tl ₂ Cl ₉ , hR84	25
Ba ₂ Cu ₄ YO ₈ , oS30	23
Hf ₉ Mo ₄ B, hP28	32
Er ₃ CrB ₇ , oS44	37
CsMn ₂ P ₆ O ₁₈ , mS56	12
U ₃ Ni ₄ Si ₄ , oI22	37

3.1.3. The Used Datasets

The public datasets (DS1 and DS2) used in this research can be found in [46,47]. DS1 has 345 compounds (binary): 61 single BCC compounds, 31 single FCC compounds, 15 single HCP compounds, 61 amorphous, and the rest are multiphase compounds. The compounds contain seven to nine crystals. DS1 has 49 crystals and their counts. DS2 dataset has 2335 quardary compounds: 600 single BCC compounds, 534 FCC phase compounds,

222 amorphous compounds, and 1254 multiphase compounds. The dataset has only eight crystals.

3.2. Experimental Results

Figure 4 depicts the CNN accuracy of Pearson correlation value between the classified values and ground truth values in both datasets (we used 20% of the dataset for testing). Using compound structures, the CNN network realizes an accuracy of 98.7% in the classification of energy entropy (EE) and 98.3% in crystal volume (VOL). The experiments depict that the CNN regression does not utilize manual feature computation. Moreover, it enhances testing precision. CNNs can select better features (which are employed for transfer learning). CNNs are pretrained on a large-sized data of 231 K compounds and 92 crystals.

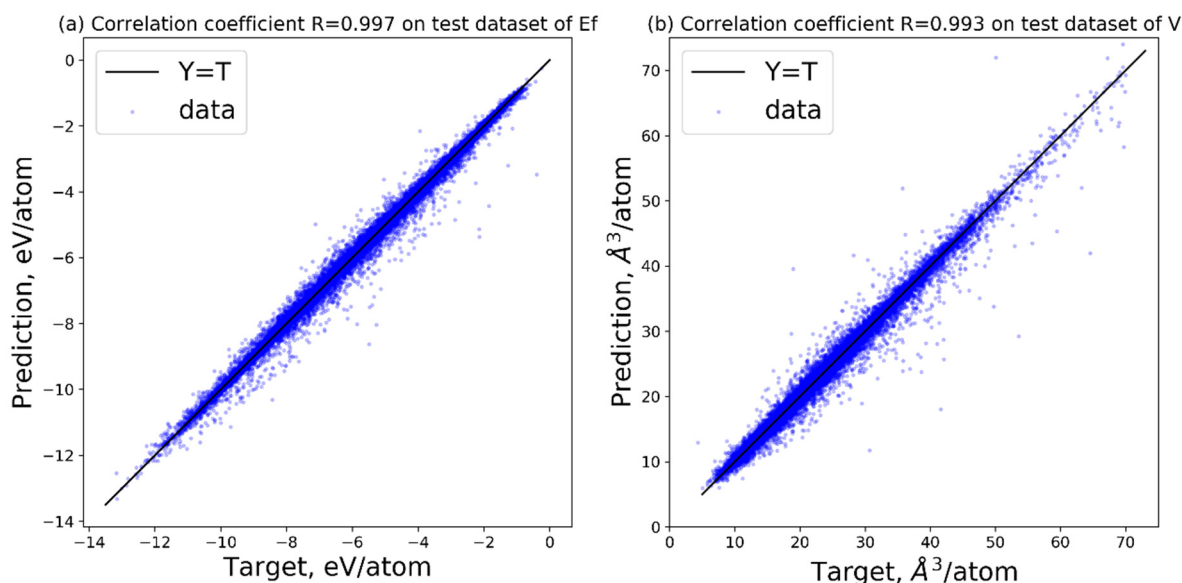


Figure 4. Ground truth on the x-axis, T, and predicted value on the y-axis; (a) energy entropy; (b) volume.

3.3. Experimental Results on Phase Compounds

The feature selection in the DS1 dataset is reused in differentiating between 180 phase compounds. In the deep learning model, the crystal structure is represented by maps and used as input for the feature selection process (EE and VOL), and two feature maps of 188 feature dimensions are generated. These feature maps are fed to the training phase classification scheme (forest trees in our research). We employed three random forest tree models. The first classifier utilizes the feature map produced by the EE feature selection phase (Map^{EE}) as input. The second classifier utilizes the feature map produced by the VOL feature selection phase (Map^{VOL}). The third classifier utilizes both Map^{EE} and Map^{VOL} ($\text{Map}^{\text{EE\&VOL}}$). The comparative study also utilizes a random forest tree model on a structure map of a 112-dimension map representing 112 crystals and counts the percentages in an inorganic compound (Map^{comp}).

Figure 5 depicts the compared models' accuracy versus the ratio of the test dataset to the whole dataset. The performance metrics utilized are recall, precision, and F-score, and, compared to the test ratio, depict trends similar to the classification accuracy. All models are constructed on crystal structures. Nevertheless, the accuracy of the transfer learning models is higher than the model without transfer learning. With 90% of the dataset used for training and 10% for testing, the model using $\text{Map}^{\text{EE\&VOL}}$ can attain an accuracy of 90%, while the model without transfer learning using Map^{comp} only attains an accuracy of 55%. The accuracy percentage of the transfer learning is unaffected by the training/testing ratio. When the test data rise from 10% to 50%, the accuracy declines from 90% to 86%. The

model using $\text{Map}^{\text{EE\&VOL}}$ depicts the highest accuracy percentage because it uses features from both selection phases.

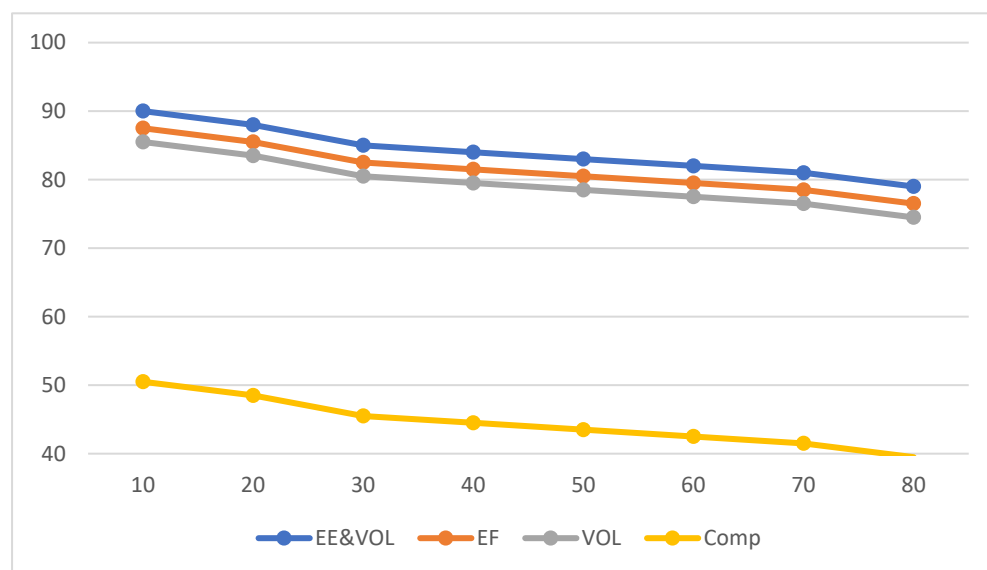


Figure 5. Accuracy versus the testing ratio of the models utilizing feature maps $\text{Map}^{\text{EE\&VOL}}$, Map^{EE} , Map^{VOL} , and the learning model without transfer learning using Map^{Comp} .

3.4. Testing Performance Metrics on High-Entropy Compounds

Table 5 depicts the mean performance metrics and mean square errors of the proposed models on high-entropy benchmarks using an eightfold testing process. The transfer learning models can differentiate between BCC, amorphous, multiple-phase mixture, FCC, and HCP, with an eightfold validation process using accuracy, recall, precision, and F-score on the testing datasets over 94%.

Table 5. Performance metrics of transfer learning models on high-entropy compounds using eightfold validation.

	Accuracy	Precision	Recall	F-Score
EE metric	0.940 ± 0.024	0.945 ± 0.025	0.940 ± 0.024	0.929 ± 0.024
VOL metric	0.949 ± 0.009	0.940 ± 0.008	0.940 ± 0.007	0.949 ± 0.006
EE&VOL	0.986 ± 0.007	0.980 ± 0.004	0.970 ± 0.008	0.986 ± 0.003

3.5. Comparative Study

We performed a comparative study comparing our proposed model with transfer learning against the state-of-the-art models. We utilized the same dataset as depicted in Table 6 and Figure 6.

Table 6. Comparative study on the same datasets with transfer learning.

	Accuracy	Precision	Recall	F-Score
Model1 in [21]	89.12%	88.1%	90.1%	88.1%
Model2 in [34]	92.5%	94.7%	93.55	91.2%
Our proposed model with EE&VOL	98.6%	98.0	97.0	98.6

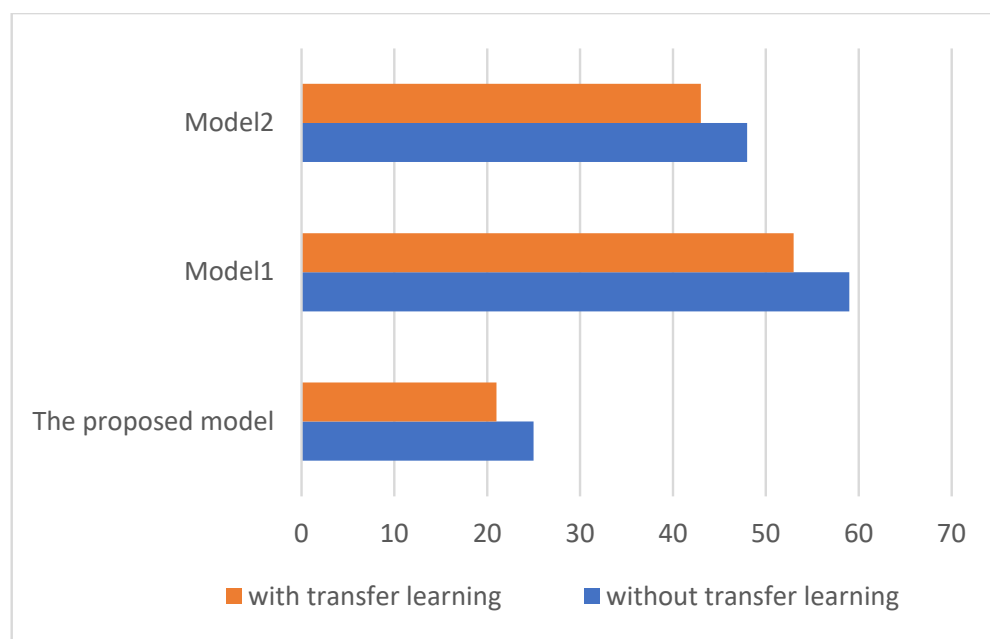


Figure 6. Prediction time of the model with and without transfer learning versus other models.

4. Discussions

Difficulties in the phase compounds' classification using structure alone involve small-sized datasets for the compounds in the dataset, imbalanced data, and a large number of classes (180 classes in our research). The utilized dataset in this research has several inorganic substances such as solid solution compounds, metals, halides, etc. It is problematic in terms of time and accuracy to select features manually. In addition, some physical measures are unobtainable for many crystals in the utilized dataset. For these problems, we used transfer learning.

When using deep learning for a novel compound classification, feature selection engineering transfers the source domain into the model. This decides the performance scores of the deep learning models with insufficient data. Collecting proper descriptors involves deep knowledge of the tools, which is very difficult in classifying new substances. The transfer learning model can be utilized to construct a high-quality standard. As we show, the proposed models in this research can attain high performance using structure without any manual engineering.

Four experiments are performed using 85% of the dataset DS2 (8500 instances) for training and 15% for testing (1500). The subsets are designated randomly. Performance is shown in the confusion matrix for Experiment 1 using the proposed CNN with and without transfer learning, as shown in Tables 7–11. The accuracy and recall are depicted in Tables 12 and 13. It is proven from these metrics that transfer learning increased performance by 40%.

Table 7. Experiment 1 confusion matrix for the proposed CNN without the transfer learning using DS1 eightfold testing and (Map^{EE}) as input for training.

Experiment 1 Testing Phase $8 \times 1500 = 12,000$	Actual Compound		Total Cases
	Correct	Incorrect	
Predicted inorganic compound	5780	6220	12,000

Table 8. Experiment 1 confusion matrix for the proposed CNN with the transfer learning using DS1 eightfold testing and (Map^{EE}) as input for training.

Experiment 1 Testing Phase $8 \times 1500 = 12,000$	Actual Compound		Total Cases
	Correct	Incorrect	
Predicted inorganic Compound Value	11,743	257	12,000

Table 9. Experiment 2 confusion matrix for the proposed CNN without the transfer learning using DS1 eightfold testing and (Map^{VOL}) as input for training.

Experiment 1 Testing phase $8 \times 1500 = 12,000$	Actual Compound		Total Cases
	Correct	Incorrect	
Predicted inorganic compound	6231	5769	12,000

Table 10. Experiment 2 confusion matrix for the proposed CNN with the transfer learning using DS1 eightfold testing and (Map^{VOL}) as input for training.

Experiment 1 Testing Phase $8 \times 1500 = 12,000$	Actual Compound		Total Cases
	Correct	Incorrect	
Predicted inorganic Compound Value	11,661	339	12,000

Table 11. Experiment 3 confusion matrix for the proposed CNN without the transfer learning using DS1 eightfold testing and (Map^{EE&VOL}) as input for training.

Experiment 1 Testing Phase $8 \times 1500 = 12,000$	Actual Compound		Total Cases
	Correct	Incorrect	
Predicted inorganic compound	7023	4977	12,000

Table 12. Experiment 3 confusion matrix for the proposed CNN with the transfer learning using DS1 eightfold testing and (Map^{EE&VOL}) as input for training.

Experiment 1 Testing Phase $8 \times 1500 = 12,000$	Actual Compound		Total Cases
	Correct	Incorrect	
Predicted inorganic Compound Value	11,922	78	12,000

Table 13. Experimental results for the three experiments with transfer learning.

Model	Accuracy %	Average Results		Error Rate
		Sensitivity %	Specificity %	
Experiment 1	95.7	95.9	95.1	0.0321
Experiment 2	97.73	96.7	96.79	0.0219
Experiment 3	98.93	98.8	98.77	0.0019

The correct and incorrect prediction cases of the third experiment using transfer learning and (Map^{EE&VOL}) have the highest count of correctly predicted compounds.

5. Conclusions

CNN with transfer learning is highly successful in classifying phase compounds and accurately classifies 170 phase inorganic compounds based on crystal structures. Representing the crystal structure of inorganic compounds with 2D images allows CNN to learn from structured data. Transfer learning using feature extraction phases of CNN uses transfer

learning on large datasets, such as DS1, and learning outcomes can then be reapplied in new learning processes to produce rich features. Transfer learning reduces the training time for new classification models but also enhances the performance and the process generalization for other models with small-sized datasets.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R113), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Data Availability Statement: The data presented in this study are available in the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Agrawal, A.A.; Choudhary, Perspective: Materials informatics and large-sized data: Realization of the “fourth paradigm” of science in substances science. *APL Mater.* **2016**, *4*, 053208. [CrossRef]
2. Mater, A.C.; Coote, M.L. Deep Learning in Crystal. *J. Chem. Inf. Model.* **2019**, *59*, 2545–2559. [CrossRef] [PubMed]
3. Butler, K.T.; Davies, D.W.; Cartwright, H.; Isayev, O.; Walsh, A. Deep learning for molecular and substances science. *Nature* **2018**, *559*, 547–555. [CrossRef] [PubMed]
4. Vasudevan, R.K.; Choudhary, K.; Mehta, A.; Smith, R.; Kusne, G.; Tavazza, F.; Vlcek, L.; Ziatdinov, M.; Kalinin, S.V.O.L.; Hatrick-Simpers, J. Materials science in the artificial intelligence age: High-throughput library generation, deep learning, and a pathway from correlations to the underpinning physics. *MRS Commun.* **2019**, *9*, 821–838. [CrossRef]
5. Wei, J.; Chu, X.; Sun, X.; Xu, K.; Deng, H.; Chen, J.; Wei, Z.; Lei, M. Deep learning in substances science. *InfoMat* **2019**, *1*, 338–358. [CrossRef]
6. Ward, L.; Agrawal, A.; Choudhary, A.; Wolverton, C. A general-purpose deep learning framework for classification of features of inorganic substances. *Npj Comput. Mater.* **2016**, *2*, 16028. [CrossRef]
7. Stanev, V.O.L.; Oses, C.; Kusne, A.G.; Rodriguez, E.; Paglione, J.; Curtarolo, S.; Takeuchi, I. Deep learning modeling of superconducting critical temperature. *Npj Comput. Mater.* **2018**, *4*, 29. [CrossRef]
8. Li, Y.; Guo, W. Deep-learning model for classification of phase formations of high-entropy compounds. *Phys. Rev. Mater.* **2019**, *3*, 095005. [CrossRef]
9. Islam, N.; Huang, W.; Zhuang, H.L. Deep learning for phase selection in multi-principal crystal compounds. *Comput. Mater. Sci.* **2018**, *150*, 230–235. [CrossRef]
10. Schwarze, C.; Kamachali, R.D.; Kühbach, M.; Mießen, C.; Tegeler, M.; Barrales-Mora, L.; Steinbach, I.; Gottstein, G. Computationally Efficient Phase-field Simulation Studies Using RVE Sampling and Statistical Analysis. *Comput. Mater. Sci.* **2018**, *147*, 204–216. [CrossRef]
11. Sun, Y.T.; Bai, H.Y.; Li, M.Z.; Wang, W.H. Deep learning approach for prediction and understanding of glass-energy ability. *J. Phys. Chem. Lett.* **2017**, *8*, 3434–3439. [CrossRef] [PubMed]
12. Isayev, O.; Oses, C.; Toher, C.; Gossett, E.; Curtarolo, S.; Tropsha, A. Universal fragment descriptors for classification of features of inorganic crystals. *Nat. Commun.* **2017**, *8*, 15679. [CrossRef] [PubMed]
13. Ghiringhelli, L.M.; Vybiral, J.; Levchenko, S.V.O.L.; Draxl, C.; Scheffler, M. Large-sized data of substances science: Critical role of the descriptor. *Phys. Rev. Lett.* **2015**, *114*, 105503. [CrossRef] [PubMed]
14. Zhang, Y.; Wen, C.; Wang, C.; Antonov, S.; Xue, D.; Bai, Y.; Su, Y. Phase prediction in high entropy compounds with a rational selection of substances descriptors and deep learning models. *Acta Mater.* **2019**, *185*, 528–539. [CrossRef]
15. Feng, S.; Zhou, H.; Dong, H. Using deep neural network with small dataset to predict substance defects. *Mater. Des.* **2019**, *162*, 300–310. [CrossRef]
16. Zhang, Y.; Ling, C. A strategy to apply deep learning to small datasets in substances science. *Npj Comput. Mater.* **2018**, *4*, 25. [CrossRef]
17. Goodfellow, Y.I.; Bengio, A.; Courville, *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
18. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
19. Kondo, R.; Yamakawa, S.; Masuoka, Y.; Tajima, S.; Asahi, R. Microstructure Recognition Using Convolutional Neural Networks for Prediction of Ionic Conductivity in Ceramics. *Acta Mater.* **2017**, *141*, 29–38. Available online: <https://www.sciencedirect.com/science/article/abs/pii/S1359645417307383?via%3Dihub> (accessed on 21 July 2019). [CrossRef]
20. Azimi, S.M.; Britz, D.; Engstler, M.; Fritz, M.; Mücklich, F. Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* **2018**, *8*, 2128. [CrossRef]
21. Agbozo, R.; Jin, W. Quantitative Metallographic Analysis of GCr15 Microstructure Using Mask R-CNN. *J. Korean Soc. Precis. Eng.* **2020**, *37*, 361–369. [CrossRef]
22. Ferguson, M.; Ak, R.; Lee, Y.-T.T.; Law, K.H. Detection and Segmentation of Manufacturing Defects with Convolutional Neural Networks and Transfer Learning. *Smart Sustain. Manuf. Syst.* **2018**, *2*, 20180033. [CrossRef] [PubMed]
23. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]
24. Torrey, L.; Shavlik, J. Transfer learning. In *Handbook Machine Learning Application Trends, Algorithms, Methods, Tech.*; IGI Global: Hershey, PA, USA, 2010; pp. 242–264.

25. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *arXiv* **2014**, arXiv:1411.1792.
26. Hutchinson, M.L.; Antono, E.; Gibbons, B.M.; Paradiso, S.; Ling, J.; Meredig, B. Overcoming Data Scarcity with Transfer Learning. *arXiv* **2017**, arXiv:1711.05099.
27. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*; San Diego, CA, USA, 7–9 May 2015; Bengio, Y., LeCun, Y., Eds.; Conference Track Proceedings; IEEE: Piscataway, NJ, USA, 2015.
28. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.O.L.; Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
29. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. 2012, pp. 1097–1105. Available online: <http://code.google.com/p/cuda-convnet/> (accessed on 1 November 2019).
30. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image dataset. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*; Miami, FL, USA, 20–25 June 2009; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2010; pp. 248–255. [[CrossRef](#)]
31. Zheng, X.; Zheng, P.; Zhang, R. Deep learning substance features from the 2D map using convolutional neural networks. *Chem. Sci.* **2018**, *9*, 8426–8432. [[CrossRef](#)]
32. Zeng, S.; Zhao, Y.; Li, G.; Wang, R.; Wang, X.; Ni, J. Atom table convolutional neural networks for an accurate prediction of compounds features. *Npj Comput. Mater.* **2019**, *5*, 84. [[CrossRef](#)]
33. Zheng, X.; Zheng, P.; Zheng, L.; Zhang, Y.; Zhang, R.Z. Multi-channel convolutional neural networks for substances features prediction. *Comput. Mater. Sci.* **2020**, *173*, 109436. [[CrossRef](#)]
34. Feng, S.; Fu, H.; Zhou, H.; Wu, Y.; Lu, Z.; Dong, H. A general and transferable deep learning framework for classification of phase formation in substances. *Npj Comput. Mater.* **2021**, *7*, 10. [[CrossRef](#)]
35. Graser, J.; Kauwe, S.K.; Sparks, T.D. Deep Learning and Entropy Minimization Approaches for Crystal Structure Predictions: A Review and New Horizons. *Chem. Mater.* **2018**, *30*, 3601–3612. [[CrossRef](#)]
36. Egorova, O.; Hafizi, R.; Woods, D.C.; Day, G.M. Multifidelity Statistical Deep Learning for Molecular Crystal Structure Prediction. *J. Phys. Chem. A* **2020**, *124*, 8065–8078. [[CrossRef](#)]
37. Ikeda, Y.; Grabowski, B.; Körmann, F. Ab initio phase stabilities and mechanical features of multicomponent compounds: A comprehensive review for high entropy compounds and structurally complex compounds. *Mater. Charact.* **2019**, *147*, 464–511. [[CrossRef](#)]
38. DS1 Dataset. Available online: <https://icsd.fiz-karlsruhe.de/index.xhtml?sessionId=89D93FA2EF68FBAC54E7DA55C479A1C4> (accessed on 1 June 2022).
39. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M. Tensorflow: A system for large-scale deep learning. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
40. Curtarolo, S.; Setyawan, W.; Hart, G.L.W.; Jahnatek, M.; Chepulskii, R.V.O.L.; Taylor, R.H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; et al. AFLOW: An automatic framework for high-throughput substances discovery. *Comput. Mater. Sci.* **2012**, *58*, 218–226. [[CrossRef](#)]
41. Jain, A.; Ong, S.P.; Hautier, G.; Chen, W.; Richards, W.D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; et al. Commentary: The substances project: A substances genome approach to accelerating substances innovation. *APL Mater.* **2013**, *1*, 011002. [[CrossRef](#)]
42. Graulis, S.; Chateigner, D.; Downs, R.T.; Yokochi, A.F.T.; Quirós, M.; Lutterotti, L.; Manakova, E.; Butkus, J.; Moeck, P.; Le Bail, A. Crystallography Open Dataset-An open-access collection of crystal structures. *J. Appl. Crystallogr.* **2009**, *42*, 726–729. [[CrossRef](#)]
43. Ward, L.; Keeffe, S.C.O.; Stevick, J.; Jelbert, G.R.; Aykol, M.; Wolverton, C. A deep learning approach for engineering bulk metallic glass compounds. *Acta Mater.* **2018**, *159*, 102–111. [[CrossRef](#)]
44. Prashun, G.; Vladan, S.; Eric, S.T. Computationally guided discovery of thermoelectric substances. *Nat. Rev. Mater.* **2017**, *2*, 17053. [[CrossRef](#)]
45. Rahnama, A.; Zepon, G.; Sridhar, S. Deep learning based prediction of metal hydrides for hydrogen storage, part II: Prediction of substance class. *Int. J. Hydrogen Energy* **2019**, *44*, 7345–7353. [[CrossRef](#)]
46. Zhang, Y.; Xu, X. Deep learning the magnetocaloric effect in manganites from lattice parameters. *Appl. Phys. A Mater. Sci. Process.* **2020**, *126*, 341. [[CrossRef](#)]
47. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.O.L.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.O.L.; et al. Scikit-learn: Deep Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. Available online: <http://jmlr.org/papers/v12/pedregosa11a.html> (accessed on 10 January 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.