

Article

Microstructure Representation Knowledge Graph to Explore the Twinning Formation

Cheng Xie , Ziwen Pan  and Chao Shu 

School of Software, Yunnan University, Kunming 650000, China; panzw@mail.ynu.edu.cn (Z.P.); shuchao_shanty@mail.ynu.edu.cn (C.S.)

* Correspondence: xiecheng@ynu.edu.cn

Abstract: Deformation twinning is an important mechanism of the plastic deformation of materials. The density of twins also affects the properties of the material. At present, the research methods of deformation twinning mainly depend on in situ EBSD, numerically investigated analysis and the finite element method. The application of machine learning methods to material microstructure research can shorten the time taken for material analysis. Machine learning methods are faced with the problem of the effective representation of the microstructure. We present a deformation twinning research method based on the representation of grain morphology features in a knowledge graph. We construct an autoencoder to extract grain morphology characteristics for building a grain knowledge graph. Then, a graph convolutional network (GCN) and fully connected network are developed to extract grain knowledge graph features and predict the twin density of materials subjected to specific tensile deformation. We use Mg-2Zn-3Li alloy as an experimental example to predict the twin density on three indexes of average grain size, twin boundaries density and average grain surface. The R^2 score of the prediction result on the twin boundaries density is up to 0.510, and the R^2 score of the average grain size and average grain surface is over 0.750. Therefore, the proposed method for deformation twinning research is effective and feasible.

Keywords: deformation twinning; knowledge graph; autoencoder; material genome; representation learning



Citation: Cheng, X.; Ziwen, P.; Chao, S. Microstructure Representation Knowledge Graph to Explore the Twinning Formation. *Crystals* **2022**, *12*, 466. <https://doi.org/10.3390/cryst12040466>

Academic Editor: Pavel Lukáč

Received: 25 February 2022

Accepted: 25 March 2022

Published: 27 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. 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

Researching the microstructure of materials is of great significance to modern materials science [1,2]. Twinning deformation is one of the main forms of microstructure change in materials [3]. The level of critical stress is responsible for deformation initiation by twinning. The conditions under which twins are produced are also complex. Not all grains with similar orientation and grain size will form twins [4]. The deformation twinning of the microstructure of materials can affect the tensile strength, yield strength and other important physical properties of materials [5,6]. Some studies induce deformation twinning in the material to obtain fine grains, in order to improve the strength properties of the material [7,8].

At present, there are three main methods for analyzing twinning deformation: artificial experimental analysis based on in situ EBSD (Electron Back-Scattered Diffraction) [9–12], numerically investigated analysis [13,14] and finite element method (FEM) analysis [15–18]. The artificial experimental method based on in situ EBSD can effectively analyze the twinning deformation process of a local microstructure and obtain a reasonable analysis [19]. However, this method not only requires the support of precision instruments, but also requires the experimentalists to have rich physical knowledge and experimental experience. The numerical investigated analysis is based on the analysis of experimental data and mathematical modeling to analyze the behavior of material deformation twinning. The finite element method is a more accurate twinning deformation simulation method than

numerical investigated analysis. However, both involve a high degree of knowledge of material science, physics and so on. The finite element method also needs a great deal of computing resources and a long simulation time.

Considering that the finite element method is slow to simulate the microstructure with a mass of grains, and the potential knowledge of material microstructure is not fully studied at present, the application of machine learning in material microstructure research is feasible. This method first requires the effective representation and learning of the microstructure. There are three mainstream methods of material microstructure representation, and this research also predicts the properties of materials to prove the effectiveness of the methods. 1. Material representation methods based on numerical statistics use various statistical properties of the microstructure as inputs to machine learning models to obtain material properties [20–24]. This method can only use the features of a single dimension as input, but the information contained in the microscopic organization is multi-dimensional, so the representation of this method is not comprehensive enough. 2. The representation method based on image features can directly take the scanning data of the microstructure as input to find the characteristics of the microstructure [25–32], for example, using the convolutional neural network (CNN) [33,34] to extract features. However, the defect of this method is that the extracted features lack the structural information of the microstructure. 3. Dai et al. [35] and Shu et al. [36] used the knowledge graph representation learning [37,38] method to represent 3D EBSD and 2D EBSD, respectively, and both showed better results than other methods in the prediction of material physical properties. The morphology of each grain is an important part of the microstructure information. The machine learning method also shows high efficiency and accuracy in morphology information representation. The CNN can be used to extract the morphology features of images [39], but this method requires a large number of labeled images for training. As a kind of unsupervised machine learning model, autoencoder can effectively perform self-supervised training and extract the morphology features of graphics [40–43]. Similarly, in the field of materials science, autoencoder is also used for the representation of material microstructure [44] and the feature extraction of material morphology [45]. However, the present knowledge graph methods only focus on the representation of the structural information of the microstructure and ignore the morphology information of grains. This results in the microstructure information producing an insufficient representation of the model, which leads to the drift of the prediction results. Furthermore, previous research only applied machine learning methods to the prediction of static material properties, and did not apply them to the prediction of dynamic microstructure changes.

In this work, we propose an EBSD knowledge graph representation learning method based on grain morphology features to explore the twin density change in Mg-2Zn-3Li alloy after applying a specified tensile deformation. We first extract the grain adjacency relationship from EBSD data to construct an EBSD knowledge graph. Secondly, we construct an autoencoder to extract the morphology features of each grain. Then, the grain morphology features, grain orientation codes and grain size codes together constitute the node embedding of the EBSD knowledge graph. Finally, a graph convolution network (GCN) is constructed to extract the graph features of the EBSD knowledge graph, and a fully connected network is used to predict the twin density of materials by adding a specified tensile deformation factor to the graph features.

2. Materials and Methods

2.1. Dataset

2.1.1. Polycrystal Sample Preparation

We prepare Mg-2Zn-3Li alloy as an experimental sample. The composition of the alloy is shown in Table 1. Using pure Mg (99.99%), pure Zn (99.99%) and Mg-30Li intermediate alloy as raw materials, a Mg-2Zn-3Li alloy ingot with a diameter no less than 85 mm was prepared by vacuum melting (heated to 750 °C in a vacuum furnace protected by argon and kept for 20 min) and the semi-automatic casting process. The ingot was homogenized

by annealing (350 °C for 12 h), followed by a large plastic deformation extrusion (extrusion temperature of 350 °C) to obtain a thin plate with a thickness of 2–3 mm and stable microstructure. The thin plates were divided into the training group and test group. The two combinations of Mg-2Zn-3Li alloy ingot were produced by vacuum melting, homogenization heat treatment and large plastic deformation hot extrusion. Both the training group samples and the test group samples contained 11 thin plates. We cut each thin plate to prepare a dog-bone-shaped tensile specimen. Tensile deformation tests were carried out at room temperature, 20 °C. Each prestretching and stretching sample was 60 mm in length and the gauge length was 26 mm. The tensile deformation experiment was carried out along the TD direction, and the direction was perpendicular to the extrusion direction of the thin plates. After that, tensile deformation was applied to the specified tensile specimen, and the microstructure of the sample was characterized by the EBSD technique. The acceleration voltage of the EBSD data acquisition procedure is 20.00 kV. The sample tilt angle is 70°. Experiments were carried out at room temperature. We select the scanning area on the TD-ED surface. The final EBSD scanning data had a step size of 1 μm and the scanning area size was $500 \times 500 \mu\text{m}^2$.

Table 1. The composition of the alloy used in the experiment (wt%).

Alloy	Mg	Zn	Li
Mg-2Zn-3Li	Bal	1.82	3.08

Figure 1 shows the inverse pole figure and the pole figure of the initial microstructure prior to deformation. The inverse pole figure on the left was drawn by Channel 5 software, and the inverse pole figure on the right was smoothed by MTEX. The extrusion direction is the ED direction. The TD-ED surface was selected for EBSD data acquisition, which is perpendicular to the extrusion direction of the plates, and IPF Z was selected for observation. Twins were not observed in the inverse pole figure of the initial microstructure prior to deformation. In this work, we predict the twin density of plates ranging from the initial microstructure to 20% tensile deformation, so we draw the pre-tensile stress–strain curve of plates reaching 20% tensile deformation, as shown in Figure 1.

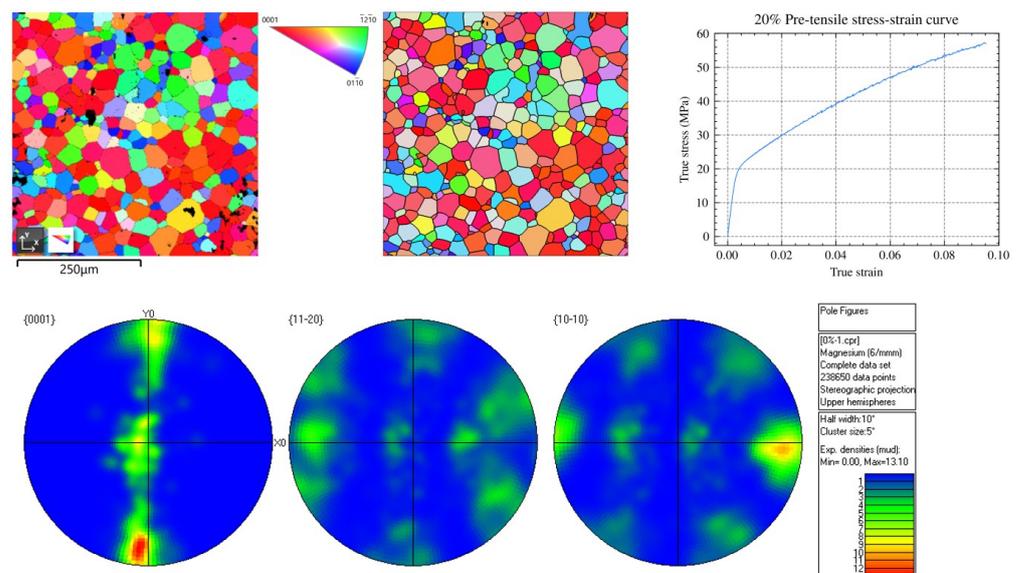


Figure 1. The inverse pole figure and the pole figure of the initial microstructure prior to deformation. Furthermore, the pre-tensile stress–strain curve under 20% tensile deformation.

Figure 2 shows the inverse pole figures of the sample under 2%, 6% and 10% tensile deformations. The inverse pole figures at the top were drawn by Channel 5 software and the figures at the bottom were smoothed using MTEX. We used the MTEX smooth inverse pole figures for twin statistics. It can be seen that a few twins appeared in the microstructure after 2% tensile deformation. Under 6% tensile deformation, a large number of twins appeared on the right and upper left side of the inverse pole figure. After 10% tensile deformation was applied, deformation twinning occurred almost everywhere in the microstructure and the grains tended to be finer.

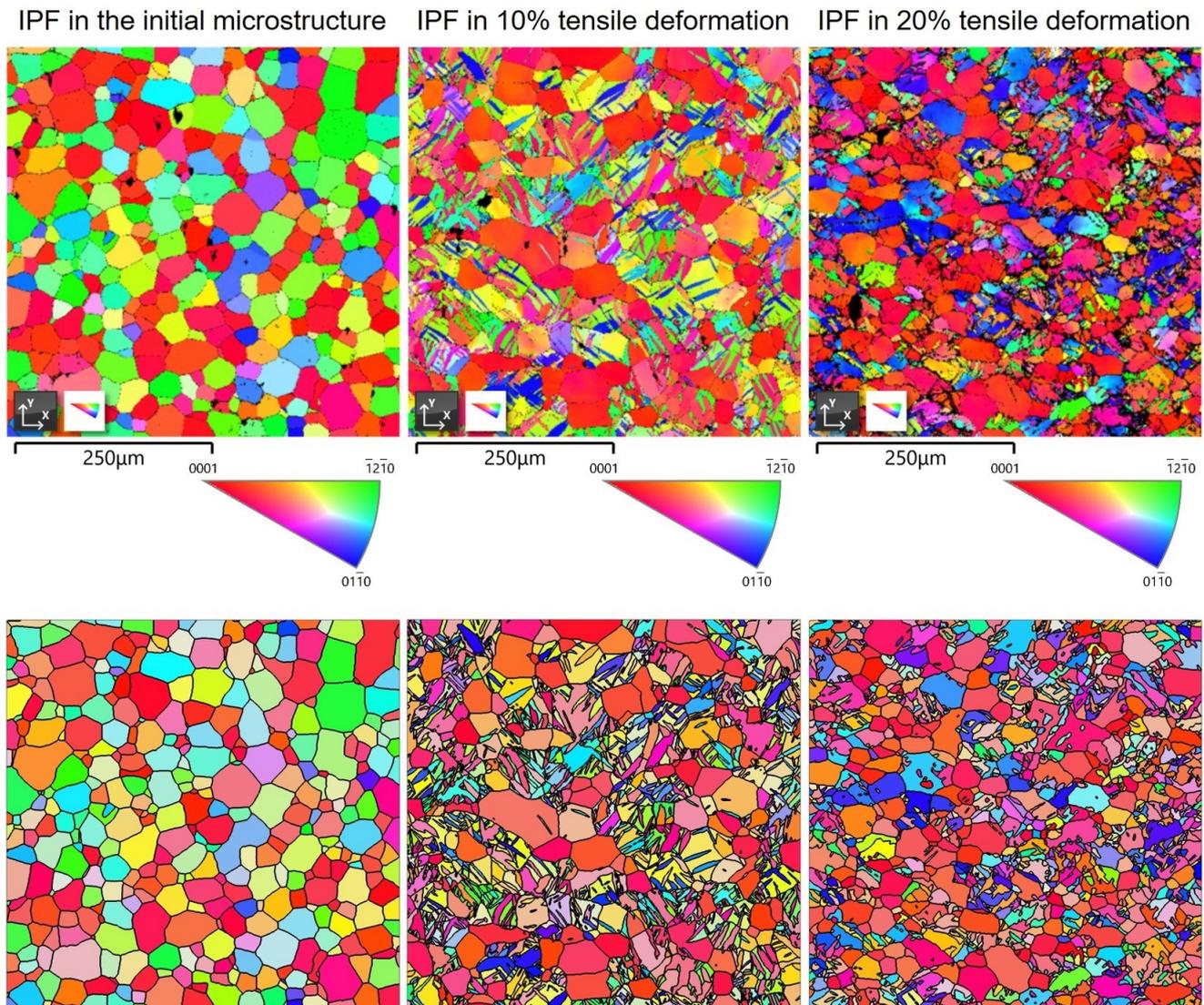


Figure 2. Inverse pole figures of microstructure under 2%, 6%, and 10% tensile deformation.

2.1.2. Dataset Generation

The sample was divided into two groups, and each group was subjected to tensile deformation at a gradient of 2%. Starting from 0% deformation, the sample was stretched to 20% deformation. EBSD scanning was performed on the sample after each stretch. Finally, we obtained 2 sets of data, each of which contained 11 points of EBSD scanning data. The first set of scanning data was used for the training model and the other set of data was used for testing the model.

The training data set and test data set contained 22 EBSD points of scanning data, including a total of 47,418 grains. The number of grains contained in each EBSD scanning data ranged from 410 to 3448. All EBSD scanning data were represented using our proposed

knowledge graph method for material microstructure representation, and the obtained graph contained 47,418 nodes and 219,472 edges in total.

We used a variety of machine learning and CNN (Convolutional Neural Network) models to compare with our proposed model on three twin density indexes. The traditional machine learning model can only accept one-dimensional inputs, so we spliced the attributes of grain angle, boundary length, ellipticity and orientation and added the tensile deformation as the inputs of the traditional machine learning model. The CNN model accepts input in the form of images. Therefore, the inverse pole figure map generated by EBSD scanning data was used as the input of the convolutional neural networks. After extracting image features in the convolutional networks, tensile deformation was added and a fully connected network was used to predict twin density. See “Methods” for the specific parameter settings of the traditional machine learning models and CNN model.

2.2. Grain Knowledge Graphs Predict Twin Density

In this work, an isomorphism graph is proposed to represent the grains and the adjacency relationship between grains in EBSD scanning data. Each node in the graph corresponds to a grain segmented from the EBSD scanning data. The embedding of each grain node consists of the state information of the grain itself. The edges in the graph represent the adjacency between grains, and the edges are undirected. Then, a graph convolutional network (GCN) is constructed to extract the graph features of the grain knowledge graph. The tensile deformation is one-hot coded and added to the graph features extracted by GCN. Finally, the spliced eigenvectors are input into a fully connected network to obtain the twin density of the sample in the current state after the specified tensile deformation is applied. Figure 3 shows the framework of our proposed method.

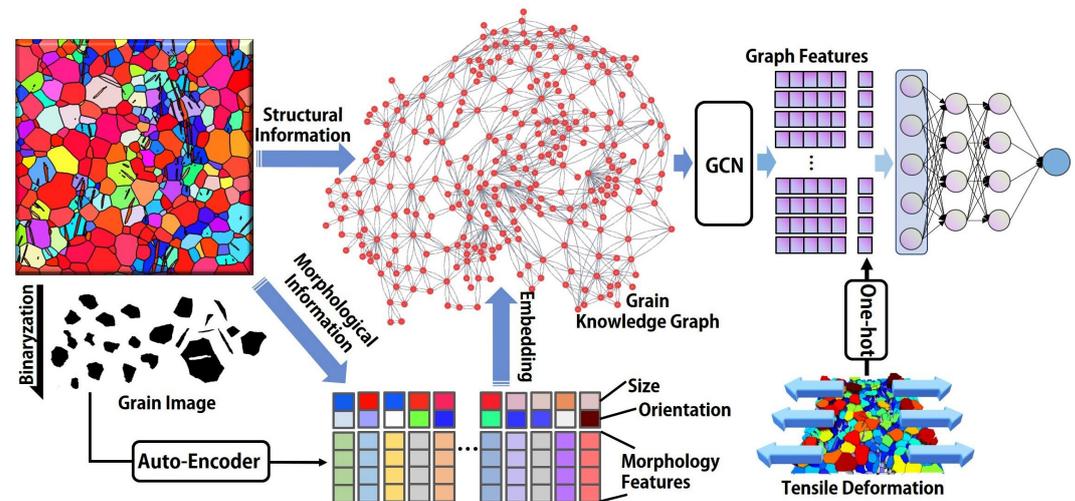


Figure 3. Firstly, each grain is numbered, and the adjacency relationship of each grain is detected to construct the grain knowledge graph. Secondly, an autoencoder is constructed to extract the morphology features of each grain. After that, the grain size and orientation are One-hot coded, which together with the morphology features of the grain is used as the embedding features of the grain. Then, the embedded grain knowledge graph is input into GCN network to extract the graph features. Finally, the tensile deformation One-hot coding applied to the material is input into the prediction network together with graph features, and the twin density of the material after the specified tensile deformation is obtained.

2.2.1. Constructing a Grain Knowledge Graph

Each grain node in the graph is embedded with the morphology features of the grain, Euler angle coding and grain size coding. The edge between each node represents the adjacency between grains. The node embedding features of the grain knowledge graph are composed of the morphology features of the grain, the grain size coding and the

grain orientation coding. The morphology features of the grain were extracted by the constructed autoencoder with 40 dimensions. The grain size features of the grains were one-hot coded. We discretized the grain size using the method shown in Equation (1), where C_{grain} represents the category to which the grain size belongs, and S_{grain} represents the real size of the grain. S_{max} and S_{min} , respectively, represent the maximum and minimum grain size. N_c represents the total number of categories, and we set this value to 11 in our method. Similarly, we divided the three Euler angles Φ_1 , Φ , and Φ_2 into 3 categories of grain orientation. Therefore, the total embedding features of grain nodes in the knowledge graph have 60 dimensions.

$$C_{grain} = \left\lceil \frac{S_{grain}}{\lceil (S_{max} - S_{min}) / N_c \rceil} \right\rceil \quad (1)$$

2.2.2. Autoencoder Design

An autoencoder is constructed to extract the morphology features of each grain. The autoencoder consists of an encoder and a decoder. The encoder is composed of a multi-layer convolutional neural network (CNN), which can extract the morphology features of each grain image. The decoder is composed of multiple deconvolutional layers, which receives the morphology features extracted by the encoder to restore the grain image. The closer the grain image restored by decoder is to the original image, the more effectively the encoder extracts the grain morphology features. The input image is a single-channel 8-bit deep binary image of each grain. Each grain image has a size of 128×128 pixels. The grain image is input into the encoder and the morphology feature vector of 40 dimension is obtained. The decoder receives the morphology features extracted by the encoder and restores the grain image. Equation (2) describes the process of convolutional layer updating, where x_j^l represents the feature map on the j th filter of the l th convolutional layer. k_j^{l+1} represents the j th filter of the $l + 1$ th convolutional layer. The symbol \star represents the convolution operation. The bias is b_j^{l+1} , and $ReLU()$ is the activation function ReLU. The MSE loss is calculated by comparing the difference in each pixel between the restored image and the original grain image. The objective function of the autoencoder is shown in Equation (3), where $X_{i,j}$ is the (i, j) pixel of grain image X , h and w are the height and width of the grain image, respectively, $E(X)$ represents the morphology features of the grain image extracted by the encoder and $D(E(X))$ represents the grain image restored by the decoder.

$$x_j^{l+1} = ReLU\left(\sum_{i \in M_j} x_i^l \star k_{ij}^{l+1} + b_j^{l+1}\right) \quad (2)$$

$$l_{ae} = \sum_{i=0}^{h-1} \sum_{j=0}^{w-1} \frac{(D(E(X)))_{i,j} - X_{i,j})^2}{h \cdot w} \quad (3)$$

2.2.3. GCN Build upon Grain Knowledge Graph

The features of the embedded grain knowledge graph are extracted by GCN layers. Let the grain knowledge graph be $G(V, E)$, where V represents the set of grain nodes and E represents the set of adjacency relationships between grains. Let the output graph feature matrix of the n th GCN layer be X^n , then the output feature matrix of the next GCN layer is X^{n+1} . The transformation from X^n to X^{n+1} can be obtained by Equation (4).

$$X^{n+1} = \sigma(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} X^n W^n + b^n) \quad (4)$$

where $\sigma()$ is the nonlinear transformation, W^n is the weight of the n th layer, and b^n is the intercept of the n th layer. A is the adjacency matrix of G . \hat{A} is the summation of matrix A and identity matrix I , meaning to add the self-loop to the graph. \hat{D} is the degree matrix of \hat{A} .

The amount of tensile deformation applied to the sample is coded. The features extracted from the grain knowledge graph are connected with the code of the amount of tensile deformation, and both are input into the multi-layer fully connected neural network. The output of these multiple fully connected layers is the predicted twin density of the material in its current state with a specified tensile deformation applied.

3. Results

3.1. Model Performance

The prediction target of the model is the twin density of the sample under the specified tensile deformation applied to the current state. We chose three indexes to measure the twin density, namely the average grain size, the average grain surface and the twin boundary density. The first two measurement indexes were obtained from the average grain size and average surface in the EBSD scanning data of the sample under each tensile deformation. The twin boundary density is the ratio of twin boundary length to total grain boundary length in EBSD scanning data. These three measurement indexes can reflect the density of twins in the sample to a certain extent.

Traditional machine learning methods such as Ridge (Ridge Regression), SVR (Support Vector Regression), KNN (K-Nearest Neighbor), ExtraTree, RF (Random Forest), GBDT (Gradient-Boosting Decision Tree), XGBoost, and MLP (Multi-layer Perceptron) are used as comparison methods. The CNN model in the field of deep learning is also used as a comparison method for our proposed methods. In addition, the knowledge graph representation method for EBSD data proposed by Shu et al. [36] is also the method we need to compare. The comparison results are shown in Table 2.

Table 2. The results of the models are compared in terms of average grain size, twin boundaries density and average grain surface.

	Average Grain Size			Twin Boundaries Density			Average Grain Surface		
	MSE	MAE	R ²	MSE	MAE	R ²	MSE	MAE	R ²
Ridge	0.548 ^{+60.8}	0.489 ^{+40.1}	0.452 ^{+73.7}	0.558 ^{+31.0}	0.591 ^{+16.4}	0.442 ^{+39.1}	0.545 ^{+54.5}	0.490 ^{+39.0}	0.455 ^{+65.3}
SVR	0.580 ^{+62.9}	0.518 ^{+43.4}	0.420 ^{+86.9}	0.660 ^{+41.7}	0.614 ^{+19.5}	0.340 ^{+80.9}	0.577 ^{+57.0}	0.523 ^{+42.8}	0.423 ^{+77.8}
KNN	0.688 ^{+68.8}	0.507 ^{+42.2}	0.312 ^{+151.6}	0.724 ^{+46.8}	0.651 ^{+24.1}	0.276 ^{+122.8}	0.687 ^{+63.9}	0.514 ^{+41.8}	0.313 ^{+140.3}
ExtraTree	0.470 ^{+54.3}	0.410 ^{+28.5}	0.530 ^{+48.1}	0.518 ^{+25.7}	0.569 ^{+13.2}	0.482 ^{+27.6}	0.468 ^{+47.0}	0.411 ^{+27.3}	0.532 ^{+41.4}
RF	0.545 ^{+60.6}	0.420 ^{+30.2}	0.455 ^{+72.5}	0.608 ^{+36.7}	0.601 ^{+17.8}	0.392 ^{+56.9}	0.539 ^{+54.0}	0.423 ^{+29.3}	0.461 ^{+63.1}
GBDT	0.631 ^{+65.9}	0.462 ^{+36.6}	0.369 ^{+112.7}	0.614 ^{+37.3}	0.599 ^{+17.5}	0.386 ^{+59.3}	0.615 ^{+59.7}	0.461 ^{+35.1}	0.385 ^{+95.3}
XGBoost	0.463 ^{+53.6}	0.389 ^{+24.7}	0.537 ^{+46.2}	0.538 ^{+28.4}	0.561 ^{+11.9}	0.462 ^{+33.1}	0.461 ^{+46.2}	0.387 ^{+22.7}	0.539 ^{+39.5}
MLP	0.592 ^{+63.7}	0.463 ^{+36.7}	0.408 ^{+92.4}	0.685 ^{+43.8}	0.631 ^{+21.7}	0.315 ^{+95.2}	0.592 ^{+58.1}	0.463 ^{+35.4}	0.408 ^{+84.3}
CNN	1.060	0.659	−0.060	0.748	0.682	0.252	1.015	0.643	−0.015
HGGAT [36]	0.314 ^{+31.5}	0.359 ^{+18.4}	0.686 ^{+14.4}	0.490 ^{+21.4}	0.552 ^{+10.5}	0.510 ^{+20.6}	0.329 ^{+24.6}	0.405 ^{+26.2}	0.671 ^{+12.1}
Our	0.215	0.293	0.785	0.385	0.494	0.615	0.248	0.299	0.752

The closer the MSE (Mean Squared Error) and MAE (Mean Absolute Error) scores are to 0, the better the prediction results are. The closer the R² score is to 1, the better the prediction effect of the model will be.

As shown in Table 2, our proposed method achieves optimal results in most indexes. In terms of average grain size and average grain surface, our proposed method achieves the best improvement, with more than 68% improvement in R² score excluding the CNN model. The CNN model achieved poor effects. The proposed method also achieves a significant improvement in the twin boundaries density index. Compared with other methods, our proposed method improves the R² score of this index by 59% on average. However, the effect of our method on this index is slightly inferior to that on the other two indexes. In our analysis, errors occur in the process of twin labeling and the statistics of twin boundary length, leading to the deviation of prediction results. Other methods also performed differently on each measure. The HGGAT method achieves sub-optimal

prediction results. Compared with the HGGAT method, the R^2 score of our method in terms of average grain size and average grain surface is improved by more than 10%, and the twin boundary density is improved by 20%. The R^2 score of our proposed method on average grain size and average grain surface are 46.2% and 39.5% higher than those of XGBoost, respectively. The R^2 score of the predicted result of twin boundary density by our method is 27.6% higher than that of ExtraTree. XGBoost and ExtraTree methods have achieved excellent results in traditional machine learning methods, but there is still a gap in the effectiveness of HGGAT and our method based on knowledge graph representation learning.

As shown in Figure 4, the predicted twin density curve obtained by our proposed method is basically consistent with the actual material under different tensile deformation. Our method can better fit the change in the twin density of materials under different tensile deformation. The method proposed by us can achieve better results because the knowledge graph representation with multi-dimensional features and spatial relations can better reflect the microstructure of materials compared with one-dimensional features. In addition, the grain morphology extracted by the autoencoder can effectively reflect the physical features of grain shape, grain boundary and so on.

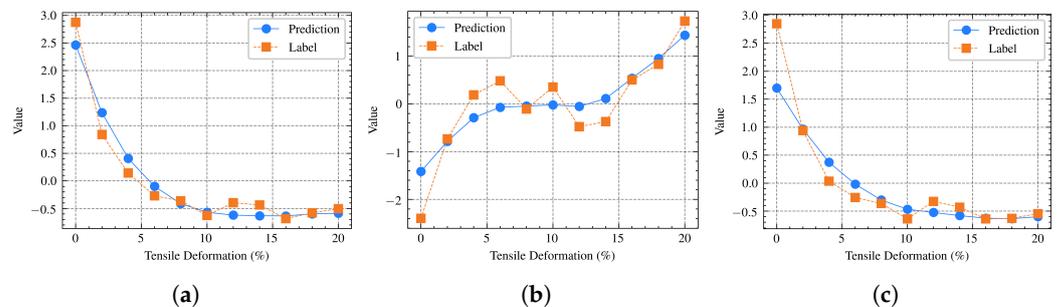


Figure 4. The line charts of the comparison between the prediction results obtained by our proposed method and the actual twin density of the material under different tensile deformations. (a) is the prediction result on average grain size. (b) is the prediction result on the twin boundaries density. (c) is the prediction result on average grain surface.

3.2. Model Analysis

It can be seen from the scatter diagram and the best fit regression line of the scatter diagram that our method achieves the best prediction result. Figure 5 is the best fit regression line of the scatter diagram of the prediction results.

As shown in Figure 5, the results of our proposed method in a best fit regression line most closely approximate the Label–Prediction $x = y$ curves. The second is the best fit regression line of the results obtained by the HGGAT method. Other models have different effects on different indicators. In general, except for the method proposed by us, the prediction results of other models all have large deviation. The prediction result of the CNN model is poor, precisely because the CNN model only learned the image features of the inverse pole figure map, and did not understand the structural relationship features of the microstructure presented by the EBSD scanning data.

As shown in Figure 6, the PCA (Principal Component Analysis) method was used to reduce the dimensions of the grain morphology features extracted from the trained autoencoder to obtain the one-dimensional distribution of grain morphology features. The morphology features distribution after dimensionality reduction was compared with the grain size distribution of EBSD scanning data. Figure 6 clearly shows that the two distributions are similar. This result indicates that the proposed autoencoder can extract the grain features effectively, and the extracted features can reflect the original size, shape and other physical information of the grain to a certain extent.

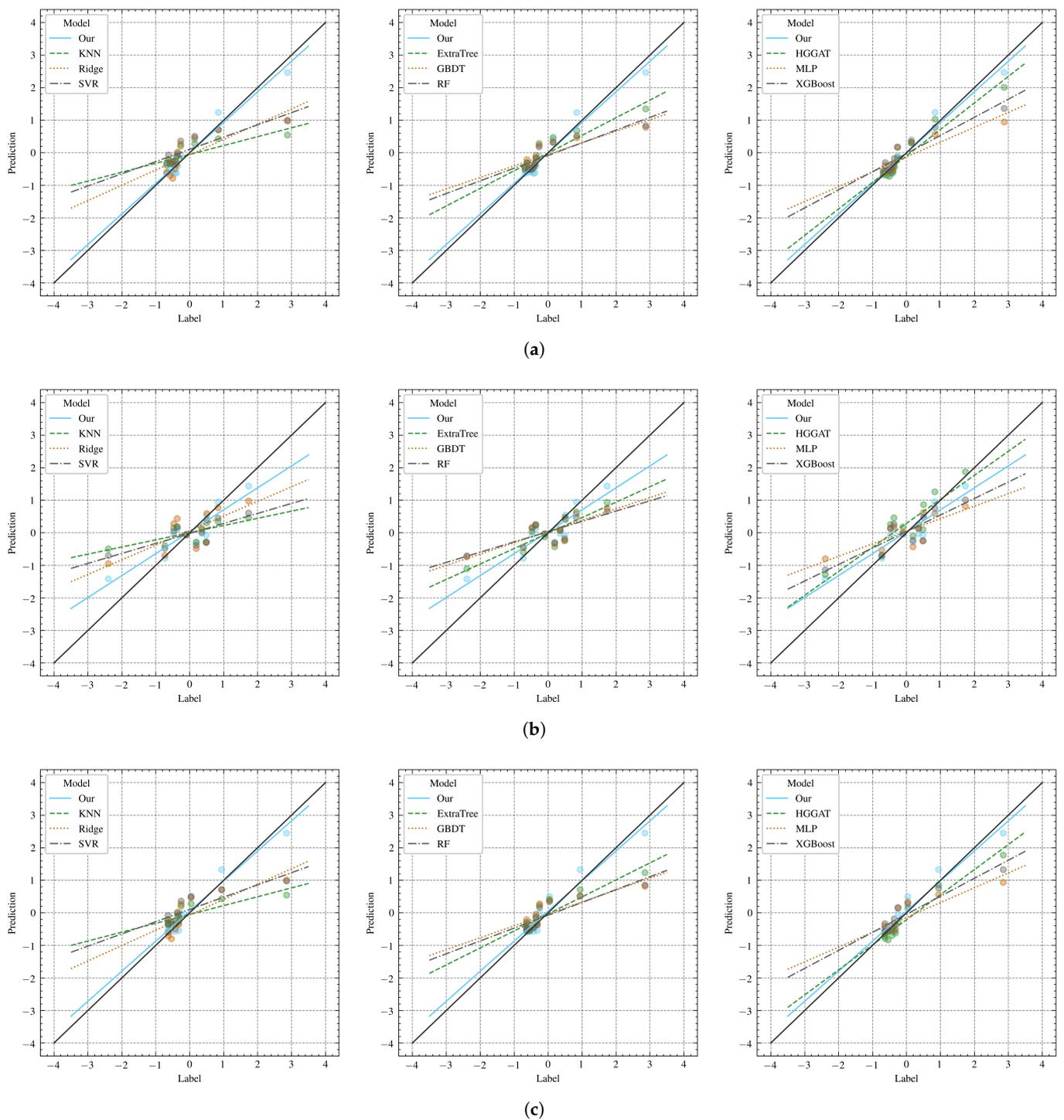


Figure 5. Image structure and graph structure. (a) The best fit regression line of the prediction results under the average grain size index. (b) The best fit regression line of the prediction results under the twin boundaries density index. (c) The best fit regression line of the prediction results under the average surface index.

The model was trained in training rounds, and the loss curve of the three indicators is shown in Figure 7. It can be seen that our proposed method can achieve convergence after limited epochs of training. In the training of the three indexes, the declining trend of loss became slow after 100 epochs of training. After 300 epochs of training, the training loss of the model tended to be stable. From the performance of the model in the test set, the model after 100 epochs of training can achieve better results. This is because the increase in

training epochs leads to the over-fitting of the model, which leads to the deterioration of the prediction effect of the model on the test set. Even though the model after 300 epochs of training had a lower loss on the training set, the generalization of the model became worse.

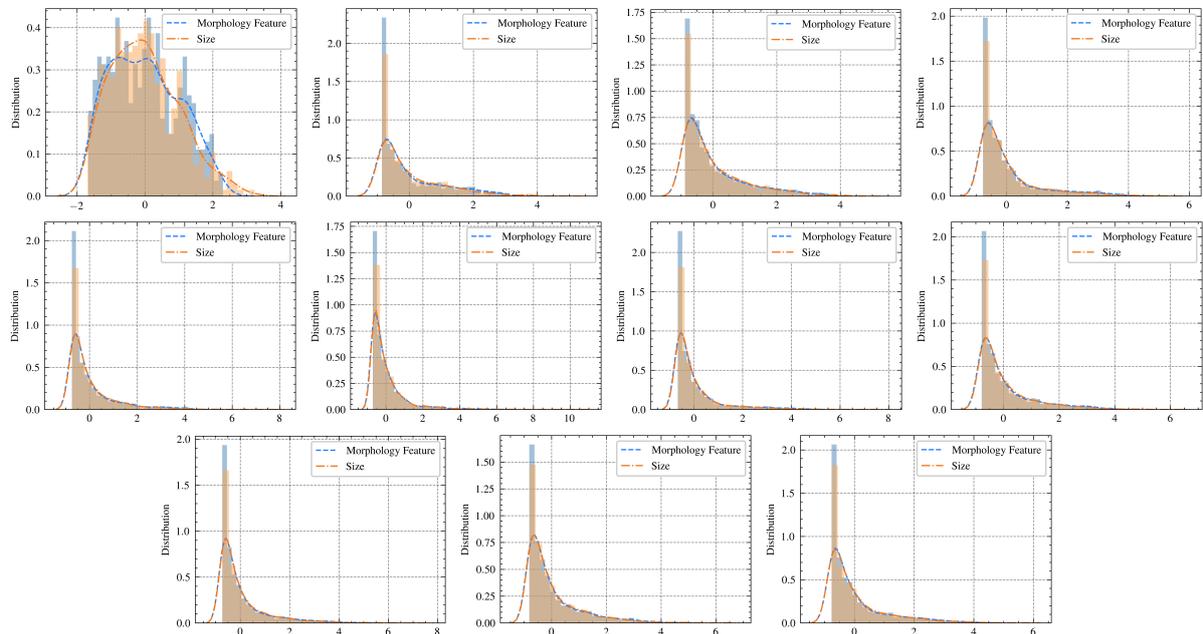


Figure 6. The distribution of grain morphology features and average grain size at each tensile deformation.

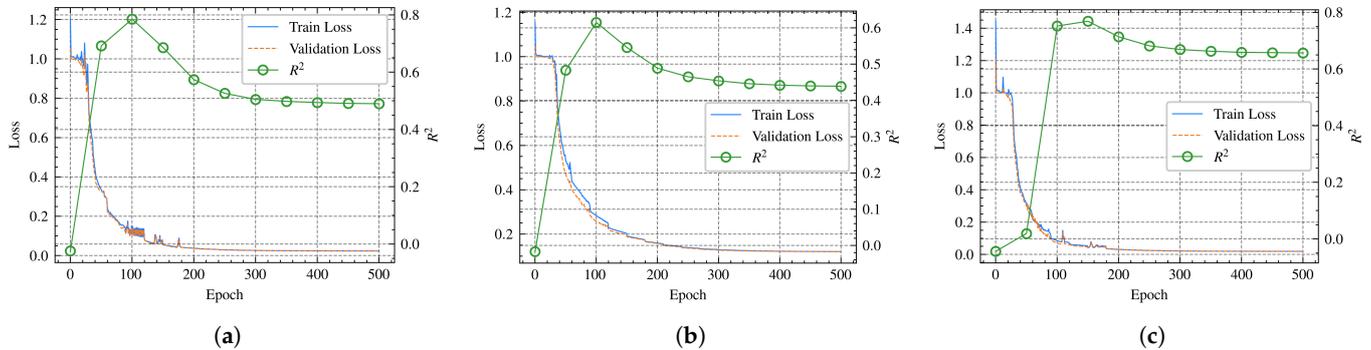


Figure 7. Curve of loss during model training and R^2 changing on test set. (a) The changes in the loss of the model on average grain size. (b) The changes in the loss of the model on twin boundaries density. (c) The loss curve of the model on an average surface. The blue curve in the figure represents the loss curve on the training data set and the orange curve represents the loss curve on the validation data set.

4. Discussion and Conclusions

Deformation twinning is one of the main forms of the microstructure deformation of materials. The twin density also has a certain effect on the macroscopic properties of materials. Artificial experimental analysis based on in situ EBSD, numerically investigated analysis and FEM analysis are the main methods for analyzing deformation twinning. These methods are based on polycrystalline physics and can effectively observe, analyze and simulate deformation twinning. However, these methods are slow and costly.

In this work, the proposed method can represent the structure information and morphology information contained in the microstructure of materials through a knowledge graph, and directly predict the change in the material's microscopic twin density through the EBSD data of the material and the tensile deformation applied to the material. The experimental results show that the proposed method is efficient and feasible. In terms of

time consumption, our method can predict the twin density after deformation twinning on EBSD data with an average grain number of 2100 with only a few hours of training.

We compare other machine learning methods and the recently proposed knowledge graph method to represent the microstructure of materials. The experimental results show that the proposed method achieves better results. Compared with other methods of representing material microstructure by knowledge graph, the amount of microstructure information contained in the knowledge graph is different, and the structure of the graph is also different. We use an autoencoder to extract the morphology features of the grains and embed them as part of the grain nodes. The experimental results show that this method can improve the prediction more effectively, and the morphology features extracted can reflect the physical characteristics of grains to a certain extent. This indicates that to more effectively characterize the microstructure and construct a more rational EBSD knowledge graph are necessary tasks.

Despite the successful results of this work, there are still areas for research and improvement. Firstly, although the machine learning method can learn material microstructure information quickly, the hidden structure inside the model is still not explicable. If the hidden layer interpretation of the model is connected with the real polycrystalline physics, the study of polycrystalline physics will be further advanced. Secondly, approaches to represent the microstructure of materials more effectively is the key link in the application of machine learning methods in materials science research, such as describing the adjacency relationship between grains in more detail, and taking the position of grains in EBSD into account in the grain features. This is an issue for future research to explore.

Author Contributions: Conceptualization, C.X. and Z.P.; methodology, C.X. and Z.P.; software, Z.P.; validation, Z.P.; formal analysis, Z.P., C.S. and C.X.; investigation, Z.P. and C.X.; resources, C.S. and C.X.; data curation, Z.P. and C.S.; writing—original draft preparation, Z.P.; writing—review and editing, C.X.; supervision, C.X.; project administration, C.X. and Z.P.; funding acquisition, C.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China and the Major Project Fund of the Department of Science and Technology of Yunnan Province, China. The specific fund names and numbers are (1) National Natural Science Foundation of China regional project, research on the discovery method of non-intelligent IoT devices driven by data and knowledge, 62162064; (2) National Natural Science Foundation of China Youth Project, Zero Learning Research for New Visual Feature Synthesis and Related Knowledge Generation, 62106216; (3) Yunnan Province Rare and Precious Metal Materials Genetic Engineering (Phase I 2021), R&D and construction of special database for copper materials and engineering demonstration (CN), 202102AB080019-1; and (4) Yunnan Province Science and Technology Major Project, construction of a special database for liquid metal (CN), 202002AB080001-5.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and code are available from <https://github.com/cdoadtea/MRKG-Data-Code> (accessed on 20 February 2022).

Acknowledgments: The authors acknowledge the National Natural Science Foundation of China and the Major Project Fund of Yunnan Provincial Science and Technology Department (CN).

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

References

1. Oganov, A.R.; Pickard, C.J.; Zhu, Q.; Needs, R.J. Structure prediction drives materials discovery. *Nat. Rev. Mater.* **2019**, *4*, 331–348. [[CrossRef](#)]
2. Rekha, S.; Raja, V.B. Review on microstructure analysis of metals and alloys using image analysis techniques. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Chennai, India, 7–9 July 2017; Volume 197, p. 012010.
3. Mahajan, S.; Williams, D. Deformation twinning in metals and alloys. *Int. Metall. Rev.* **1973**, *18*, 43–61. [[CrossRef](#)]

4. Beyerlein, I.; Capolungo, L.; Marshall, P.; McCabe, R.; Tomé, C. Statistical analyses of deformation twinning in magnesium. *Philos. Mag.* **2010**, *90*, 2161–2190. [[CrossRef](#)]
5. Lei, C.; Li, X.; Deng, X.; Wang, Z.; Wang, G. Deformation mechanism and ductile fracture behavior in high strength high ductility nano/ultrafine grained Fe-17Cr-6Ni austenitic steel. *Mater. Sci. Eng. A* **2018**, *709*, 72–81. [[CrossRef](#)]
6. Zhang, A.; Kang, R.; Wu, L.; Pan, H.; Xie, H.; Huang, Q.; Liu, Y.; Ai, Z.; Ma, L.; Ren, Y.; et al. A new rare-earth-free Mg-Sn-Ca-Mn wrought alloy with ultra-high strength and good ductility. *Mater. Sci. Eng. A* **2019**, *754*, 269–274. [[CrossRef](#)]
7. Pan, F.; Mao, J.; Zhang, G.; Tang, A.; She, J. Development of high-strength, low-cost wrought Mg–2.0 mass% Zn alloy with high Mn content. *Prog. Nat. Sci. Mater. Int.* **2016**, *26*, 630–635. [[CrossRef](#)]
8. Praveen, S.; Bae, J.W.; Asghari-Rad, P.; Park, J.M.; Kim, H.S. Ultra-high tensile strength nanocrystalline CoCrNi equi-atomic medium entropy alloy processed by high-pressure torsion. *Mater. Sci. Eng. A* **2018**, *735*, 394–397. [[CrossRef](#)]
9. Zhang, H.J.; Salvati, E.; Papadaki, C.; Fong, K.S.; Song, X.; Korsunsky, A.M. Grain Rotation during Twin-Detwin Deformation of Mg AZ31 Alloy Using In Situ XRD and EBSD. In Proceedings of the Key Engineering Materials, Trans Tech Publ, Oxford, UK, 29–31 March 2019; Volume 793, pp. 17–22.
10. Liu, J.; Chen, H. Nucleation and Grain Boundary Evolution in Dynamic Recrystallization of 316LN Steel During Hot Deformation. *Front. Mater.* **2019**, *6*, 209. [[CrossRef](#)]
11. Tiarniyu, A.; Tari, V.; Szpunar, J.; Odeshi, A.; Khan, A. Effects of grain refinement on the quasi-static compressive behavior of AISI 321 austenitic stainless steel: EBSD, TEM, and XRD studies. *Int. J. Plast.* **2018**, *107*, 79–99. [[CrossRef](#)]
12. Song, L.; Wu, B.; Zhang, L.; Du, X.; Wang, Y.; Esling, C. Twinning characterization of fiber-textured AZ31B magnesium alloy during tensile deformation. *Mater. Sci. Eng. A* **2018**, *710*, 57–65. [[CrossRef](#)]
13. Qiao, H.; Guo, X.; Hong, S.; Wu, P. Modeling of {10-12}- $\{10-12\}$ secondary twinning in pre-compressed Mg alloy AZ31. *J. Alloy. Compd.* **2017**, *725*, 96–107. [[CrossRef](#)]
14. Zhang, F.; Liu, Z.; Yang, M.; Su, G.; Zhao, R.; Mao, P.; Wang, F.; Sun, S. Microscopic mechanism exploration and constitutive equation construction for compression characteristics of AZ31-TD magnesium alloy at high strain rate. *Mater. Sci. Eng. A* **2020**, *771*, 138571. [[CrossRef](#)]
15. Liu, Z.; Wang, H.; Haché, M.J.; Chu, X.; Irissou, E.; Zou, Y. Prediction of heterogeneous microstructural evolution in cold sprayed copper coatings using local Zener-Hollomon parameter and strain. *Acta Mater.* **2020**, *193*, 191–201. [[CrossRef](#)]
16. Chen, S.; Song, H.; Zhang, S.; Cheng, M.; Lee, M. Effect of shear deformation on plasticity, recrystallization mechanism and texture evolution of Mg–3Al–1Zn alloy sheet: Experiment and coupled finite element-VPSC simulation. *J. Alloy. Compd.* **2019**, *805*, 138–152. [[CrossRef](#)]
17. Qayyum, F.; Guk, S.; Prüger, S.; Schmidtchen, M.; Saenko, I.; Kiefer, B.; Kawalla, R.; Prah, U. Investigating the local deformation and transformation behavior of sintered X3CrMnNi16-7-6 TRIP steel using a calibrated crystal plasticity-based numerical simulation model. *Int. J. Mater. Res.* **2020**, *111*, 392–404. [[CrossRef](#)]
18. Choi, W.M.; Jo, Y.H.; Sohn, S.S.; Lee, S.; Lee, B.J. Understanding the physical metallurgy of the CoCrFeMnNi high-entropy alloy: An atomistic simulation study. *NPJ Comput. Mater.* **2018**, *4*, 1–9. [[CrossRef](#)]
19. Liu, F.; Guo, C.; Xin, R.; Wu, G.; Liu, Q. Evaluation of the reliability of twin variant analysis in Mg alloys by in situ EBSD technique. *J. Magnes. Alloy.* **2019**, *7*, 258–263. [[CrossRef](#)]
20. Khosravani, A.; Cecen, A.; Kalidindi, S.R. Development of high throughput assays for establishing process-structure-property linkages in multiphase polycrystalline metals: Application to dual-phase steels. *Acta Mater.* **2017**, *123*, 55–69. [[CrossRef](#)]
21. Yabansu, Y.C.; Patel, D.K.; Kalidindi, S.R. Calibrated localization relationships for elastic response of polycrystalline aggregates. *Acta Mater.* **2014**, *81*, 151–160. [[CrossRef](#)]
22. Yabansu, Y.C.; Kalidindi, S.R. Representation and calibration of elastic localization kernels for a broad class of cubic polycrystals. *Acta Mater.* **2015**, *94*, 26–35. [[CrossRef](#)]
23. Latypov, M.I.; Kühbach, M.; Beyerlein, I.J.; Stinville, J.C.; Toth, L.S.; Pollock, T.M.; Kalidindi, S.R. Application of chord length distributions and principal component analysis for quantification and representation of diverse polycrystalline microstructures. *Mater. Charact.* **2018**, *145*, 671–685. [[CrossRef](#)]
24. Hart, G.L.; Mueller, T.; Toher, C.; Curtarolo, S. Machine learning for alloys. *Nat. Rev. Mater.* **2021**, *6*, 730–755. [[CrossRef](#)]
25. 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. [[CrossRef](#)]
26. Yang, Z.; Yabansu, Y.C.; Al-Bahrani, R.; Liao, W.k.; Choudhary, A.N.; Kalidindi, S.R.; Agrawal, A. Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets. *Comput. Mater. Sci.* **2018**, *151*, 278–287. [[CrossRef](#)]
27. Li, X.; Zhang, Y.; Zhao, H.; Burkhart, C.; Brinson, L.C.; Chen, W. A transfer learning approach for microstructure reconstruction and structure-property predictions. *Sci. Rep.* **2018**, *8*, 1–13. [[CrossRef](#)]
28. Cang, R.; Li, H.; Yao, H.; Jiao, Y.; Ren, Y. Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model. *Comput. Mater. Sci.* **2018**, *150*, 212–221. [[CrossRef](#)]
29. Chan, H.; Cherukara, M.; Loeffler, T.D.; Narayanan, B.; Sankaranarayanan, S.K. Machine learning enabled autonomous microstructural characterization in 3D samples. *NPJ Comput. Mater.* **2020**, *6*, 1. [[CrossRef](#)]

30. Herriott, C.; Spear, A.D. Predicting microstructure-dependent mechanical properties in additively manufactured metals with machine-and deep-learning methods. *Comput. Mater. Sci.* **2020**, *175*, 109599. [[CrossRef](#)]
31. Yang, C.; Kim, Y.; Ryu, S.; Gu, G.X. Prediction of composite microstructure stress–strain curves using convolutional neural networks. *Mater. Des.* **2020**, *189*, 108509. [[CrossRef](#)]
32. Rong, Q.; Wei, H.; Huang, X.; Bao, H. Predicting the effective thermal conductivity of composites from cross sections images using deep learning methods. *Compos. Sci. Technol.* **2019**, *184*, 107861. [[CrossRef](#)]
33. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
34. Sainath, T.N.; Vinyals, O.; Senior, A.; Sak, H. Convolutional, long short-term memory, fully connected deep neural networks. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, Australia, 19–24 April 2015; pp. 4580–4584.
35. Dai, M.; Demirel, M.F.; Liang, Y.; Hu, J.M. Graph neural networks for an accurate and interpretable prediction of the properties of polycrystalline materials. *NPJ Comput. Mater.* **2021**, *7*, 103. [[CrossRef](#)]
36. Shu, C.; He, J.; Xue, G.; Xie, C. Grain Knowledge Graph Representation Learning: A New Paradigm for Microstructure-Property Prediction. *Crystals* **2022**, *12*, 280. [[CrossRef](#)]
37. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81. [[CrossRef](#)]
38. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Philip, S.Y. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4–24. [[CrossRef](#)] [[PubMed](#)]
39. Warmuzek, M.; Żelawski, M.; Jałocha, T. Application of the convolutional neural network for recognition of the metal alloys microstructure constituents based on their morphological characteristics. *Comput. Mater. Sci.* **2021**, *199*, 110722. [[CrossRef](#)]
40. Dong, J.; Li, L.; Han, D. New Quantitative Approach for the Morphological Similarity Analysis of Urban Fabrics Based on a Convolutional Autoencoder. *IEEE Access* **2019**, *7*, 138162–138174. [[CrossRef](#)]
41. Rundo, F.; Conoci, S.; Banna, G.L.; Ortis, A.; Stanco, F.; Battiato, S. Evaluation of Levenberg–Marquardt neural networks and stacked autoencoders clustering for skin lesion analysis, screening and follow-up. *IET Comput. Vis.* **2018**, *12*, 957–962. [[CrossRef](#)]
42. Zhou, C.; Gu, Y.; Fang, G.; Lin, Z. Automatic morphological classification of galaxies: Convolutional autoencoder and bagging-based multiclustering model. *arXiv* **2021**, arXiv:2112.13957.
43. Luo, H.; Tang, Y.Y.; Biuk-Aghai, R.P.; Yang, X.; Yang, L.; Wang, Y. Wavelet-based extended morphological profile and deep autoencoder for hyperspectral image classification. *Int. J. Wavelets Multiresolut. Inf. Process.* **2018**, *16*, 1850016. [[CrossRef](#)]
44. Kim, Y.; Park, H.K.; Jung, J.; Asghari-Rad, P.; Lee, S.; Kim, J.Y.; Jung, H.G.; Kim, H.S. Exploration of optimal microstructure and mechanical properties in continuous microstructure space using a variational autoencoder. *Mater. Des.* **2021**, *202*, 109544. [[CrossRef](#)]
45. Shi, Z.; Mandal, S.; Harimkar, S.; Liu, C. Surface Morphology Analysis Using Convolutional Autoencoder in Additive Manufacturing with Laser Engineered Net Shaping. *Procedia Manuf.* **2021**, *53*, 16–23. [[CrossRef](#)]