On the use of transfer modeling to design new steels with excellent rotating bending fatigue resistance even in the case of very small calibration datasets

Wei, Xiaolu; van der Zwaag, Sybrand; Jia, Zixi; Wang, Chenchong; Xu, Wei

DOI
10.1016/j.actamat.2022.118103

Publication date
2022

Document Version
Final published version

Published in
Acta Materialia

Acta Materialia 235 (2022) 118103

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat

Full length article

On the use of transfer modeling to design new steels with excellent rotating bending fatigue resistance even in the case of very small calibration datasets

Xiaolu Wei[a], Sybrand van der Zwaag[b], Zixi Jia[c], Chenchong Wang[a], Wei Xu[a,*]

[a] State Key Laboratory of Rolling and Automation, Northeastern University, Shenyang, Liaoning 110819, China
[b] Novel Aerospace Materials Group, Faculty of Aerospace Engineering, Delft University of Technology, Delft 2629 HS, the Netherlands
[c] Faculty of Robot Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, China

ARTICLE INFO

Article history:
Received 7 February 2022
Revised 7 June 2022
Accepted 14 June 2022
Available online 17 June 2022

Keywords:
Fatigue strength
Small sample problem
Machine learning
Transfer learning
Alloy design

ABSTRACT

In this research a machine learning model for predicting the rotating bending fatigue strength and the high-throughput design of fatigue resistant steels is proposed. In this transfer prediction framework, machine learning models are first trained to estimate tensile properties (yield strength, tensile strength and elongation) on the basis of composition and critical process conditions. Then, based on the predicted tensile properties, transfer models are trained to estimate fatigue strength. The results are compared with those of a similar model not having such a transfer layer. The transfer prediction framework shows high accuracy for fatigue strength prediction with a remarkably high tolerance to limitations in the amount of calibration data available for training. By combining the transfer prediction framework with evolutionary algorithms, a robust high-throughput alloy design model is achieved requiring only tens of fatigue data points to get a decent reliability. The newly designed steel showed the predicted high fatigue strength. The method as presented here might also be applicable to other alloy design challenges in which only a limited database for the property to be optimized is available.

1. Introduction

Fatigue strength, which is usually defined as the maximum stress amplitude without failure for a given high number of stress cycles (usually 10^7~10^9) under set loading conditions, is commonly used to represent the endurance limit of steels [1–4]. While the actual fatigue strength in use depends on the loading conditions (R-value, frequency, stress states, loading pattern, etc.) the rotating bending stress is a good test to estimate the relative fatigue performance of different engineering steels. While the test is attractive because it requires only small samples, relatively cheap testing machines and can test at high frequencies, measuring the fatigue strength by experimental testing remains a time consuming and hence costly operation even when optimal sample loading strategies to minimize the number of samples to be tested such as the staircase method [5] are used. Due to this high cost the number of fully worked cases in which fatigue strength is coupled to steel composition and heat treatment conditions is relatively small and this greatly complicates the design of new fatigue resistant engineering steels. Recently, with the establishment and development of the Material Genome Initiative (MGI), various studies have tried to greatly reduce the time and monetary cost necessary for alloy design by replacing experimental trial-and-error approaches with advanced computational and statistical methods, e.g., high-throughput computing. Although high-throughput computing has brought about great progress in the efficiency of designing new copper alloys with a high ultimate tensile strength and a low electrical conductivity [6], for titanium alloys with high strength and ductility [7], ultrahigh-strength stainless steels with high hardness [8], and RAFM steels with high yield strength and impact toughness [9], few high-throughput alloy design studies have focused on fatigue strength optimization primarily because of limitation in the number of composition-fatigue strength data available.

To design new steels with a high fatigue strength via high-throughput computing, a proper model that can accurately represent the complex and non-linear relationships between composition and processing conditions on the one hand and the fatigue strength on the other hand should first be established. Most traditional empirical models in the field relate fatigue strength to quasi-static mechanical properties, such as hardness [10], tensile properties [10,11] or impact toughness [12]. The use of these quasi-static properties, which are easier to obtain and hence carry a lower

* Corresponding author.
E-mail address: xuwei@ral.neu.edu.cn (W. Xu).

https://doi.org/10.1016/j.actamat.2022.118103
1359-6454/

cost penalty, to predict fatigue properties can be a cost-effective strategy. In early research [2], a linear relationship between fatigue strength and hardness or tensile strength has been proposed and imposed, and this definition has been widely applied to various alloy systems, such as steels, copper and aluminum alloys. However, in later studies, this linear relationship was found to be too simple and not to apply to engineering steel grades in which the hardness or tensile strength exceeded a critical value [13]. For example, when the tensile strength exceeded the value of approximately 1800 MPa, the fatigue strength of SAE 4340-like steel even decreased with increasing tensile strength value [14]. To describe this nonlinear relationship in the high tensile-strength area, a modified model based on hardness data was built by Yukitaka and Masahiro [15–17]. A more versatile formula that relates fatigue strength to tensile strength in the form of a quadratic expression was proposed by Pang et al. [12,14] and performed well for several materials, such as alloyed steels and copper alloys. However, the model is strictly phenomenological, and only links fatigue strength to tensile strength, and offered no explanation for the non-linearity. Wang proposed that yield strength is also related to fatigue strength as yielding is involved in fatigue crack initiation [11]. It has also been argued that total elongation at break is also relevant for fatigue strength optimization as a high ductility can promote fatigue damage resistance [11,18], and also the work-hardening ability relates to the fatigue strength to some extent [19]. However, no empirical formula has been proposed which quantifies the effect of failure strain and work-hardening on fatigue strength.

However, irrespective of the empirical relationship between the conventional mechanical properties and the fatigue strength selected, such relationships do not directly link the fatigue strength to the chemical composition of the steel and the processing conditions (together defining the microstructure).

Artificial intelligence (AI) strategies may provide a new way to directly define the relationships between the combination of chemical composition and heat treatment conditions and the fatigue properties without considering mechanical relationships or mechanisms. From traditional machine learning (ML) methods, e.g., neural networks (NNs) and support vector regression (SVR), to more recent methods, e.g., extreme gradient boosting (XGB) and convolutional neural networks (CNNs), these approaches exhibit good predictive ability for various material properties [8,20–25]. AI strategies were also used for various fatigue property predictions in previous studies, such as the prediction of fatigue live [26,27], fatigue crack-driving forces [28], fatigue strengths [29–37], etc. However, with limited time and funding for fatigue testing, most previous studies mentioned above were based on small sample databases with only hundreds of samples and hence the predictions lacked reliability and accuracy. To overcome the problem of limited accuracy due to the calibration database being too small (i.e., less than 100 cases), models suitable for sparse data have been used, e.g., the adaptive neuro-fuzzy inference system, which was used to predict the high cycle fatigue life of laser powder bed fusion stainless steel 316 L based on a dataset consisting of 139 experimental fatigue data points [26]. In addition, data preprocessing methods such as feature engineering were considered in Agrawal's work [36]. Although the number of samples nearly reached the minimum amount of data required for traditional ML methods, nevertheless hundreds of samples were required, meaning that the data accumulation still took decades, which is an unacceptable long-term cycle for alloy design. Usually, for most relatively new alloy systems, only tens of existing data points might be available, an amount that is far from sufficient for training a stable or reliable ML model. Therefore, although the traditional AI strategy provides promising prospects for fatigue strength-oriented alloy design by directly building the relationship between composition/processing and fatigue strength, the lack of a sufficient amount of complete datasets and the high cost for additional data accumulation significantly inhibited its further application in efficient and reliable alloy design.

In the present work, a fatigue strength prediction and high-throughput alloy design framework based on the transfer learning (TR) concept [38–45] is proposed. The first layer of the transfer model predicts the quasi-static mechanical properties (yield strength, tensile strength and elongation) using either a convolutional neural network (CNN) framework based on deep learning concepts or a simplified machine learning (SML) framework based on traditional machine learning algorithms. The inputs for both models are the steel composition and the processing parameters (in particular critical heat treatment parameters) for which a large database is available. The second layer links the predicted quasi-static properties to the high cycle fatigue strength for which only a small validation set is available. The predictive power of the transfer model is compared to that of models in which chemical composition and heat treatment parameters are directly linked to the fatigue strength, i.e., a non-transfer model. For the subsequent high-throughput alloy design, a genetic algorithm (GA), which used the transfer prediction model as the objective function, was applied to search for the new composition and processing solutions leading to high fatigue strength values while taking into account as many as twenty dimensional features. This TR framework enabled high-efficiency alloy designs exploring an extremely large range of potential combinations but does rely on only a small dataset of fatigue strength values for training.

2. Methods

2.1. Dataset and data preprocessing

In the present study the publicly available Matnavi fatigue dataset published by the National Institute of Material Science (NIMS) of Japan was used [46]. In this database actual steel compositions, heat treatment conditions and final quasi-static and dynamic mechanical properties are recorded. The complete database consists of datasets for carbon steels (113), low-alloy steels (258), spring steels (18) and stainless steels (22). Fatigue data for carburized steels were not used in the present study since their processing involves additional processing parameters and leads to intentional compositional gradients not to be encountered in the other steel grades. In the end a dataset containing 411 samples was compiled where each datapoint has twenty features (composition and processing details), three tensile properties (yield strength, tensile strength and elongation), hardness, Charpy impact toughness and the target property (rotating bending fatigue strength at 10^7 cycles). The Matnavi dataset as published by NIMS is a high-quality dataset with the advantage of small scattering within the data because all fatigue tests were performed at a single institution known for its accuracy [47]. The global overview of the dataset ranges is reported in Table 1, and histograms of the composition ranges and mechanical properties are shown in Fig. S1 and Fig. S2, respectively.

For the tensile properties, 83, 262 and 66 samples were selected randomly as the validation set, training set and testing set, respectively. To avoid the 'lucky split' problems caused by random partitions, the data partitioning procedure using this fixed splitting ratio was carried out 100 times to obtain an insight into the uncertainty of the model. For the fatigue strength, from the initial dataset with 411 samples, 83 samples were selected randomly as the validation set, and the remaining 328 samples were used for model training and testing. From the remaining 328 samples, various subsets containing ranging from 32 to 328 samples were randomly selected for fatigue strength prediction. In all cases, a ratio of 4:1 was used to create training and testing sets. Two types of

Table 1
Input and output ranges of the various parameters in the total database.

	Inputs and outputs	Minimum	Maximum	Mean	Standard deviation
Inputs	Carbon (wt.%)	0.09	0.63	0.396	0.098
	Silicon (wt.%)	0.16	2.05	0.306	0.254
	Manganese (wt.%)	0.32	1.6	0.825	0.289
	Phosphorus (wt.%)	0.004	0.031	0.017	0.005
	Sulfur (wt.%)	0.002	0.03	0.014	0.006
	Nickel (wt.%)	0.01	2.78	0.493	0.853
	Chromium (wt.%)	0.01	12.7	1.154	2.61
	Copper (wt.%)	0	0.26	0.061	0.049
	Molybdenum (wt.%)	0	0.24	0.061	0.085
	Normalizing Temperature (°C)	30	900	820.47	188.76
	Through Hardening Temperature (°C)	30	975	833	136.5
	Through Hardening Time (min)	0	30	29.2	4.84
	Cooling Rate for Through Hardening (°C/s)	0	24	11.76	7.15
	Tempering Temperature (°C)	30	750	589.76	109.29
	Tempering Time (min)	0	60	58.39	9.68
	Cooling Rate for Tempering (°C/s)	0	24	23.36	3.87
	Reduction Ratio (Ingot to Bar)	289	5530	964.1	576.77
	Area Proportion of the Inclusions Deformed by Plastic Work (dA)	0	0.13	0.047	0.032
	Area Proportion of the Inclusions Occurring in Discontinuous Array (dB)	0	0.05	0.004	0.009
	Area Proportion of the Isolated Inclusions (dC)	0	0.06	0.009	0.012
Outputs	Yield Strength (MPa)	290	1636	789.41	219.81
	Tensile Strength (MPa)	455	1756	911.14	198.65
	Elongation (%)	9	40	21.31	4.75
	Fatigue Strength (MPa)	225	906	492.42	98.71

models were created: (i) a TR (Transfer) model in which the relevant dataset of composition and processing parameters was used to predict the quasi-static mechanical parameters and then a second model to predict the fatigue strength on the basis of the predicted (quasi-static) mechanical properties and (ii) a NonTR (non-Transfer) model in which the composition and processing parameters were directly coupled to the fatigue strength data during training. As for the predictions of the tensile properties, the above fatigue data partitioning procedure was carried out 100 times.

In the both routes some data cleaning was imposed to handle non-reported values for process parameters. Following a published protocol for dealing with non-reported input process parameters [34] the values for austenitization and tempering temperatures were set to room temperature unless properly documented, and related also-not-reported features, such as holding time and cooling rate, were set to zero. For data normalization, the inputs and outputs were normalized with z-score, a standard method for eliminating dimensional differences between feature ranges [48]. The normalization expression is given by Eq. (1):

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

where z denotes the normalized value, x is the original value from the dataset, and μ and σ represent the mean and standard deviation of the original values for a certain dimensional feature, respectively.

Feature relevance for the parameters in the dataset was evaluated to better understand the influence of the compositions and process parameters on the fatigue strength. This evaluation was carried out using the mean decrease accuracy (MDA) values for a random forest (RF) model with 1000 random partitions. The results prove conclusively that all features have positive, though not equal, contributions to the fatigue strength as shown in Fig. S3. Therefore, further feature selection was considered not to be required.

2.2. Construction of TR framework

In the transfer model approach, which treats the tensile properties as intermediate steps towards obtaining an estimate of the fatigue strength, we constructed two transfer frameworks named the CNN TR model and SML TR model. Their implementation process is shown schematically in Fig. 1(a). At first, the source models for tensile properties prediction were constructed via the CNN and SML methods using a relatively big dataset. Then, based on the high correlation between tensile properties and fatigue strength, the TR models for fatigue strength prediction were constructed based on smaller datasets containing the fatigue data, again using the CNN and SML methods.

The arguments for the construction of the unconventional CNN (rather than NN) TR framework are as follows. CNN methods can be applied to the prediction of steel properties because of the strong correlation between physical metallurgical features, which can be regarded as spatial information. The applicability of CNN has been demonstrated in other works, such as the martensite start temperature prediction by numerical input features [49], wherein CNN shows better performance than traditional NN. We also compared the performance of CNN and common neural network (NN) for the fatigue strength prediction, confirming the effectiveness of CNN, and the results are shown in Fig. S5. For the construction of the CNN TR, we followed the transfer learning approach mentioned by Yosinski et al. [50]. First, a source CNN tensile property model was trained, as shown in Fig. 1(b). For this model, 20 dimensional features reshaped into a matrix of size 5 were used as inputs (the 20 dimensional input values are filled into a 5 × 5 matrix in turn, and the value of the last 5 elements in the matrix is set to zero), and the properties of yield strength (YS), ultimate tensile strength (UTS), and elongation (EL) formed the 3-dimensional output. For the reshape method for inputs, different input matrix shapes were compared and the model with input of 5 × 5 performed best compared to models with input matrix shapes of 5 × 4 and 4 × 5. Meanwhile, a square input of 5 × 5 leads to more convenient data processing in the present work. To adapt to the characteristics of the small data (sub-)set of fatigue strength valued used in this research, the complexity of the CNN model was reduced by removing the pool layer and simplifying the architecture. The CNN parameter settings were as follows: convolutional layer 1 was 5 × 5 × 8, convolutional layer 2 was 5 × 5 × 16, the fully connected layer was 1 × 1 × 128, and the filter size was 2 × 2. Then, the related target CNN TR model for fatigue strength was trained as follows: (1) Convolutional layers and fully connected layers of the source CNN model were copied to the corresponding layers of the target CNN TR model, and were called transferred feature layers. Once created the transferred fea-

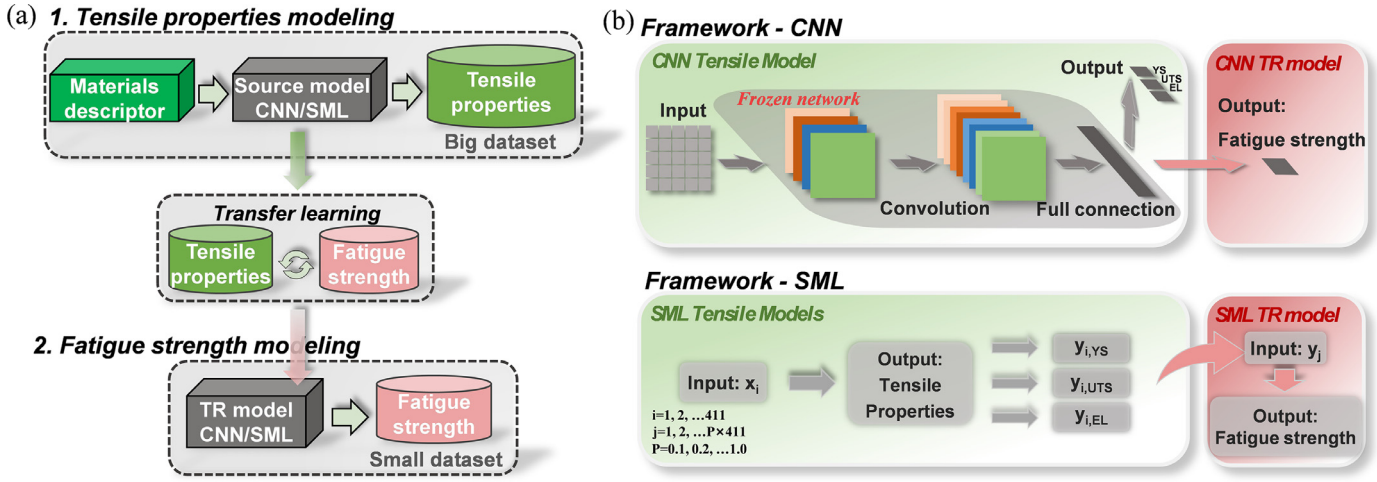

Fig. 1. Schematic diagram of (a) the transfer prediction framework for fatigue strength utilizing tensile properties and (b) the architectures of CNN and SML frameworks.

ture layers remained frozen and did not participate in further training. (2) The remaining layer of the target CNN TR model was then randomly initialized and trained for fatigue strength prediction. For the above CNN training, the model was obtained after 1500 iterations, during which the loss function of mean square error (MSE), a learning rate of 0.005, and the Adam optimizer were adopted.

For the creation of the SML transfer framework, five traditional algorithms were employed to model the tensile properties, including gradient boosting regression (GBR), extreme gradient boosting (XGB), random forest (RF), multilayer perceptron (MLP) and support vector regression with a radial basis function kernel (rbf-SVR). To model the fatigue properties, two algorithms namely rbf-SVR and support vector regression with a linear kernel (linear-SVR) were employed. Three source SML models for YS, UTS and EL, respectively were trained first. The outputs of the three models were used as the three-dimensional input for the SML TR model, which has the fatigue strength as its final output. An internal evaluation of the above models showed that the GBR and linear-SVR were the optimal algorithms for tensile properties and fatigue strength prediction, respectively. The results are shown in Fig. 6. Model parameters for all algorithms in the SML framework were optimized by grid search.

For comparison, corresponding NonTR models were trained as reference models in a similar manner, also taking the 20-dimensional features as inputs and fatigue strength as the (only) output. For both NonTR and TR, the fatigue strength error was used as the loss to train the models. Besides, different loss functions were used for the NonTR models, and results are shown in Section 4.3.

For all the above modeling of CNN and SML, data preprocessing and model training were implemented using Tensorflow and Scikit-learn in Python. The metrics used to evaluate the predictive performance of TR and NonTR models in the present work include the squared correlation coefficient (R^2) and mean absolute error (MAE). The formulas for these metrics are as follows:

$$R^2 = \frac{\left(n\sum_{i=1}^{n} f(x_i)y_i - \sum_{i=1}^{n} f(x_i)\sum_{i=1}^{n} y_i\right)^2}{\left(n\sum_{i=1}^{n} f(x_i)^2 - \left(\sum_{i=1}^{n} f(x_i)\right)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |f(x_i) - y_i| \quad (3)$$

2.3. High-throughput optimization algorithm for alloy design

With the TR and NonTR models for fatigue strength established, an elitist reservation genetic algorithm (GA) was used to redesign steels in order to obtain a high fatigue strength. Considering the large differences in compositions and heat treatment processes between different steel grades in the Matnavi dataset, the search space was limited to low-carbon steel and low-alloy steel systems. For the GA, there are multiple settings to be considered. Reasonable GA parameters need to be set to ensure that it can converge and obtain the optimal design solution. For example, the determination of mutation rate requires careful handling as an excessive value may lead to problems for GA to find a good solution, while a small value may cause GA to easily get stuck in a local optimal solution. Therefore, the value is usually set between 0.001 and 0.1. In addition, the selection of operators such as mutation operator is also critical. In the present work, various GA parameters were compared and a suitable set of GA parameters was applied to alloy design for all models. The design results demonstrated the reliability of GA under this set of parameters, as shown in Supplementary 5. The more details of GA parameters used are shown in Table S4. The GA was run ten times to ensure that for the parameters used local optima were not dictating the final outcome.

To ensure the reliability and effectiveness of the design produced by the TR an NonTR framework, the models with MAE $<$20 MPa for both the training and testing sets were used as the objective functions of GA.

Finally, the combination of the GA model and the TR model with MAE $<$10 MPa and $R^2>$90% for both training and testing sets was used to design a new steel with a fatigue strength $>$643 MPa and this steel was fabricated and properly processed samples were rotating-bending fatigue tested.

It should be mentioned that the influence of factors related to inclusions needs to be considered during the continued design process, and it is well known that inclusions have a significant impact on fatigue performance. The importance of inclusions for fatigue strength prediction was evaluated via RF, and the results show their negligible impact due to low content (S, P) and proportion (dA, dB, dC) levels, as shown in Fig. S4. Hence, solutions leading to steels with a higher amount of inclusions were removed during design processes.

2.4. Validation methods

The ultimately designed alloy was produced as a 50 kg ingot using smelting in and casting from a vacuum induction furnace.

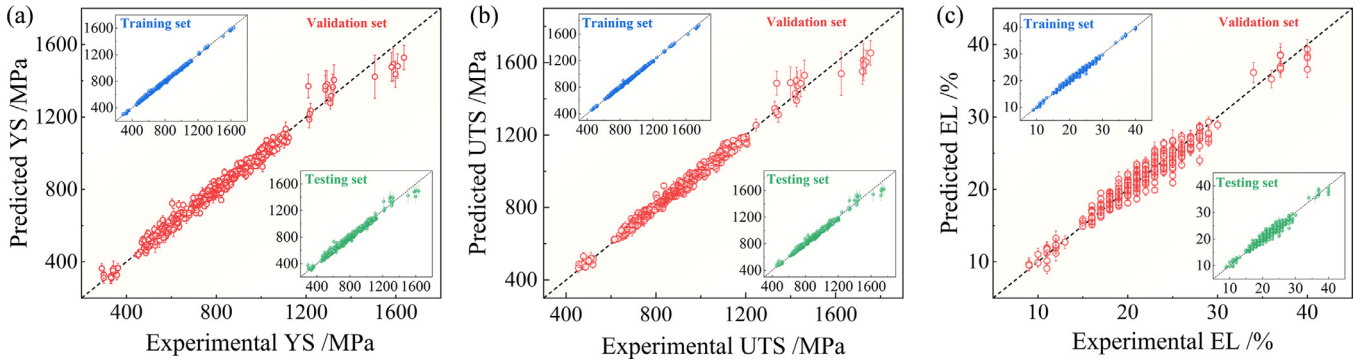

Fig. 2. Experimental values vs. values predicted by the CNN tensile model of (a) YS, (b) UTS and (c) EL (including the mean results for 100 different partitions of training, testing and validation sets).

Table 2
Mean results of the CNN and SML tensile models from 100 partitions of training, testing and validation sets.

Algorithm		Yield Strength		Ultimate tensile strength		Elongation	
		R^2/%	MAE/MPa	R^2/%	MAE/MPa	R^2/%	MAE/MPa
CNN	Training	99.5 ± 0.1	13.1 ± 1.6	99.5 ± 0.1	10.7 ± 1.2	98.7 ± 0.2	0.43 ± 0.04
	Testing	97.2 ± 1.3	27.3 ± 3.6	97.5 ± 1.3	22.5 ± 3.3	94.0 ± 1.9	0.89 ± 0.09
	Validation	97.3 ± 1.0	27.1 ± 3.8	97.6 ± 1.0	22.4 ± 3.4	94.1 ± 1.8	0.88 ± 0.09
SML	Training	99.8 ± 0.1	6.8 ± 2.3	99.8 ± 0.11	6.2 ± 1.9	98.8 ± 0.4	0.39 ± 0.08
	Testing	96.8 ± 2.0	26.3 ± 4.9	97.5 ± 1.5	21.7 ± 4.0	94.3 ± 1.9	0.88 ± 0.09
	Validation	96.9 ± 1.7	26 ± 4.6	97.5 ± 1.3	21.6 ± 3.7	94.3 ± 1.8	0.87 ± 0.08

The ingot was forged at 1100 °C into a billet with cross-sectional dimensions of 200 × 135 mm^2. Then, the billet was reheated to 1200 °C for 3 h and hot rolled to a thickness of 20 mm, with subsequent laminar cooling. Subsequent heat treatments, including normalization, hardening and tempering, were conducted at the temperatures and times suggested by the designed solution. Rotating bending fatigue tests were carried out using the same sample dimensions and test conditions as reported for Matnavi dataset.

3. Results

3.1. Prediction of tensile properties

As mentioned in Section 2.2, firstly 100 CNN models to predict the quasi-static tensile properties based on the chemical composition and the processing conditions were built. Fig. 2 presents scatter plots of the mean predictive values for YS, UTS and EL, including error bars. The prediction results in training set for YS, UTS and EL are highly consistent with the experimental values both having small error bars, indicating the excellent performance of all 100 CNN tensile models yielding high R^2 values with extremely low error bars, 99.5% (±0.1%), 99.5% (±0.1%) and 98.7% (±0.2%) for YS, UTS and EL, respectively. For the validation sets the R^2 values are 97.3% (±1%), 97.6% (±1%) and 94.1% (±1.8%) for YS, UTS and EL, respectively. Therefore, because all the R^2 values are higher than 90% with error bars <2%, the CNN models trained in this research were proven to be suitable and stable for tensile property prediction within the composition range of the original database. The error analysis further indicated the model might be slightly overfitted if an unjustified partitioning of training and testing sets was coincidentally made in the random splitting, Therefore, to avoid the overfitting caused by unjustified partitioning, the most accurate CNN model from the 100 models trained has been selected and was used for building the CNN TR model.

In addition to CNN models, also an SML framework for tensile property prediction was built, as mentioned in Section 2.2. For this SML framework, three predictive models for YS, UTS and EL, respectively were built using the GBR algorithm. Table 2 includes the mean R^2 and MAE results, including error bars, for YS, UTS and EL generated from 100 partitions of training, testing and validation sets. The mean results of the CNN model are also listed for comparison. Similar to the CNN model, all SML tensile models exhibit excellent performance but show minor differences : (i) compared with CNN models, SML models show relatively low MAE values for validation sets, indicating that SML models probably have stronger predictive abilities within the composition ranges of small sample databases; (ii) compared with SML models, CNN models show relatively small R^2 error bars for validation sets, indicating that the CNN models used in this research are more tolerant to unjustified partitioning and are therefore less inclined to result in overfitting.

3.2. Prediction of fatigue strength

The CNN and SML TR models were further trained to predict the fatigue strength initially utilizing an extremely small dataset (32 datapoints), as mentioned in Section 2.2. In addition, NonTR models were trained similarly to show how TR models advance the prediction of fatigue strength. Fig. 3 presents the comparison between the mean predictive values of the validation sets from 100 TR and NonTR models, including error bars. The mean R^2 and MAE are also given in this figure. The figure shows clearly that the TR model yields far more accurate predictions than the NonTR model, in particular in the low and the high fatigue strength domain. For the TR models there is not much difference in accuracy between the CNN and the SML models. The limited accuracy for the NonTR model in the low and high strength region is obviously due to the small dataset used for training. The use of the TR model greatly extends the range in which accurate predictions can be made on the basis of the same small database. The influence of the number of samples in the fatigue database on the accuracy will be analyzed in more detail in the Discussion.

3.3. High-throughput alloy design and experimental validation

With the help of GA, three alloys likely to have the desired high fatigue strength were selected from solutions evaluated by the TR

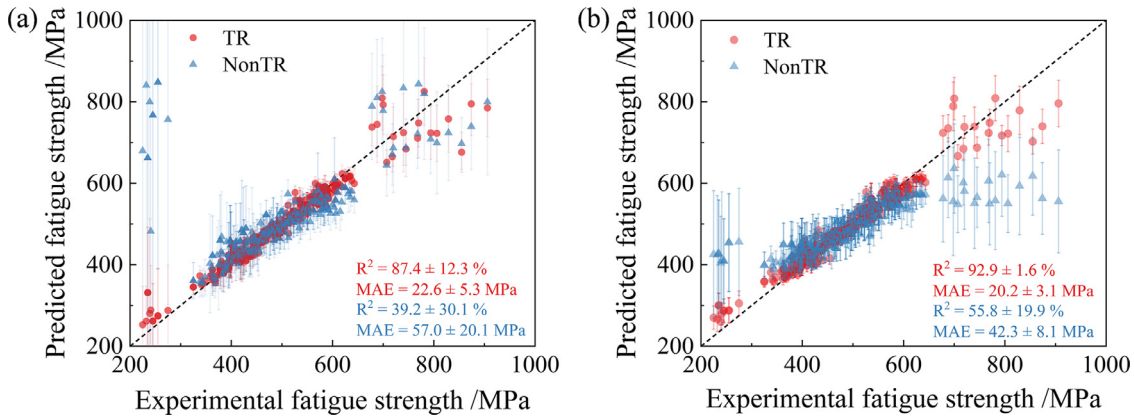

Fig. 3. Experimental values vs. predicted values in a validation set generated by (a) a CNN framework and (b) an SML framework (including the mean results for 100 different dataset partitions).

Table 3
Composition and processing parameters of three potential new alloys designed by TR models and existing alloys. Alloy R is the new steel produced on the basis of the specifications for target steel D1. Elemental compositions are given in weight percentages. Temperatures are in Celsius. Predicted (Actual) fatigue strength is in MPa.

	C	Si	Mn	Ni	Cr	Cu	Mo	NT	THT	TT	Predicted (Actual) Fatigue Strength
Alloy D1	0.48	0.30	0.75	2.76	1.08	0.19	0.18	876	852	556	692
Alloy D2	0.49	0.25	0.80	1.79	1.07	0.12	0.23	891	850	554	715
Alloy D3	0.49	0.26	0.76	0.05	1.07	0.14	0.24	888	852	560	682
Alloy E1	0.37	0.29	0.76	1.88	0.90	0.03	0.24	870	845	580	(643)
Alloy E2	0.39	0.28	0.77	0.09	1.08	0.15	0.19	870	855	550	(638)
Alloy E3	0.41	0.24	0.8	0.03	1.02	0.1	0.18	870	855	550	(635)
Alloy E4	0.42	0.28	0.75	0.07	1.04	0.12	0.17	870	855	550	(634)
Alloy E5	0.4	0.25	0.74	0.24	0.96	0.1	0.18	870	855	550	(631)
Alloy E6	0.43	0.26	0.74	0.02	1.07	0.02	0.22	870	855	550	(625)
Alloy R	0.47	0.32	0.74	2.80	1.06	0.20	0.19	875	850	555	–

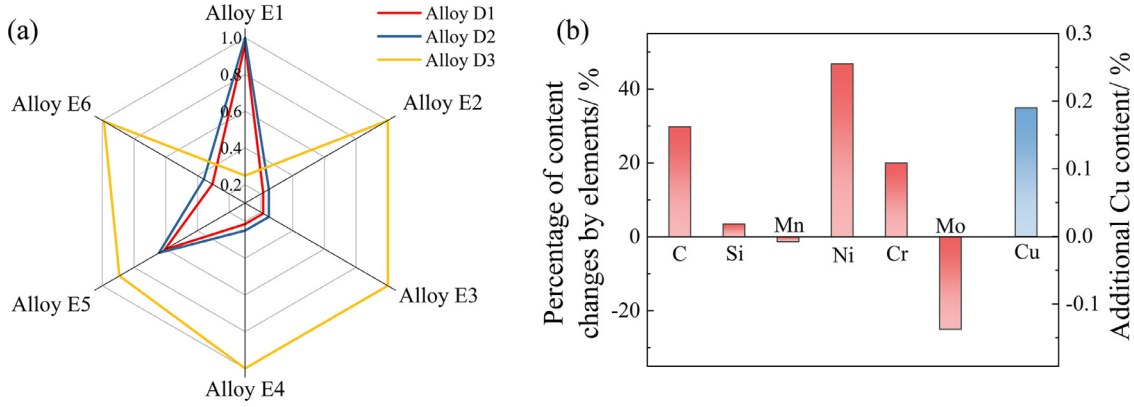

Fig. 4. Compositional comparison results between the designed alloys and original alloys in the dataset. (a) Spearman correlation coefficient. (b) Elemental content changes in Alloy D1 compared to the optimal alloy in the original dataset.

model, predicted upon only 1000 iteration cycles, a process which is considered efficient and high-throughput. Table 3 shows the designed composition and treatment parameters, named Alloy D1, D2 and D3, where the prefix D stands for Designed. The uniqueness of the compositions of the three newly designed alloys with respect to those of six existing steels with a comparably high fatigue strength (Alloys E1-E6, wherein Alloy E1 exhibits the highest fatigue strength and the prefix E stands for Existing) was assessed by calculating the Spearman correlation coefficient, and the results are shown in Fig. 4(a). Alloy D3 was found to have a composition (and processing route) comparable to an existing steel, but alloys D1 and D2 showed a relatively low correlation with existing steels, indicating the novelty of the design results. Alloy D1 presents the widest difference from closest existing steel grade (alloy E1) with real changes in the C, Ni and Mo contents, as shown in Fig. 4(b). In addition, it is interesting to note that more Cu was introduced in Alloy D1 compared to the nearest existing steel grade Alloy E1. Hence, Alloy D1 was chosen for experimental validation. The resulting experimental steel grade is called Alloy R, where the prefix R refers to Realization.

Fig. 5 shows the S-N curves for Alloy R as well as the reported data for the closest reference steel Alloy E1. It can be clearly observed that most S-N points for Alloy 1 are located above those of the original optimal alloy, indicating the excellent fatigue performance of the newly designed alloy. The fatigue property is highly sensitive to the sample and testing conditions. Therefore, the fatigue strength of Alloy R is 645~690 MPa according to the results of five run-out specimens. The upper limit of the experimental

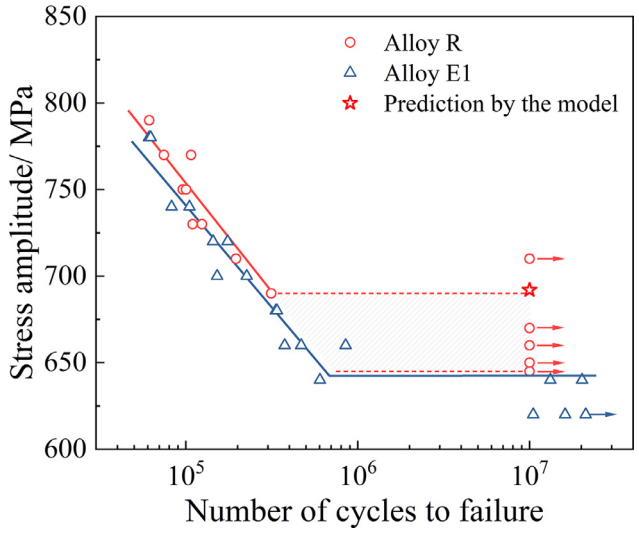

Fig. 5. Rotating bending S-N data for Alloy R and the corresponding closest reference streel Alloy E1. The fatigue data for Alloy E1 were taken from the NIMS database.

results (690 MPa) is consistent with the predicted results, which proves the high accuracy of the model. The lower limit of the experimental results (645 MPa) is still slightly higher than the best performance of the alloys from the original dataset (643 MPa for Alloy E1). This validation result robustly indicates that the SML TR&GA model has the ability to efficiently and accurately design new steel compositions and treatment conditions leading to high rotating bending fatigue strength values.

4. Discussion

4.1. Different prediction models for the SML framework

Selecting the optimal algorithm that adapts to the characteristics of the database used in this research is an important aspect of building high-performance AI models. Before GBR was used for tensile property prediction, several algorithms were tested, including GBR, rbf-SVR, XGB, RF and MLP. Most of these models perform very well for tensile property prediction and can be further used to build TR models. The mean results from 100 partitions in these five models are shown in Table S3. Given three-dimensional small data, two algorithms were employed for building TR models, including rbf-SVR and linear-SVR. The rbf-SVR shows a relatively low R^2 (86.5% ($\pm$11.3%)) for the validation set, which probably represents model overfitting to some extent. Although the relationship between fatigue strength and the combination of three tensile properties is difficult to quantify with a general equation, it is probably not as complicated as quantifying its internal mechanism, so a complex algorithm may not be an appropriate choice. The linear-SVR shows a high R^2 (92.9% ($\pm$1.6%)) for the validation set. Therefore, it was used to map the relationship between tensile properties and fatigue strength for small datasets. It is evident that linear-SVR is an optimal algorithm that presents little tendency of overfitting.

Fig. 6(a) and (b) show the mean R^2 and MAE of the validation set for different TR models. Given the linear SVR for the TR model, all five algorithms accurately predicted fatigue strength, and their R^2 values exceeded 92% with low standard deviations, wherein GBR performed best. Therefore, based on quantitative analysis, GBR and linear SVR were selected as the optimal algorithms for the tensile and TR models in the present work, respectively. Fig. 6(c) and (d) show the results of different NonTR models accordingly. It is apparent that all algorithms show extremely low R^2 and high MAE values for the validation set, especially MLP, which represents a serious overfitting problem. XGB has a better relative performance than the other four algorithms and was selected as the optimal algorithm for the NonTR models.

4.2. Effect of the amount of fatigue training data on fatigue strength prediction

As shown in the Results section, the predictive power of the TR model, even when trained with only 32 fatigue data sets, is already remarkable. To further analyze the tolerance of the model predictions to the number of fatigue data available for training, this dependence is examined in more detail for both the TR and NonTR models.

Fig. 7(a) and (b) show the effects of the number of fatigue data sets for the validation set on the mean R^2 and MAE of the NonTR and TR models. For the NonTR models, the CNN and SML model results demonstrate an evident sensitivity to the amount of fatigue data. The mean accuracy of the NonTR models increases rapidly with increasing the amounts of training data sets from 32 to 98 samples. As already pointed out in Section 3.2, for 32 fatigue strength training data sets, the CNN and SML models have extremely low R^2 and high MAE values with large standard deviations, especially in the CNN model, which is clearly unstable and has a higher requirement for the amount of data. Doubling the number of datasets for training to 66 leads to a strong increase in R^2 of to approximately 75% and a decrease in the MAE to approximately 30 MPa, but error bars remain large. When the amount of fatigue data is in the range from 147$\sim$328 samples, the SML NonTR model shows increasing advantages compared to the CNN model, as shown in Fig. 7(a) and (b). Based on the results of Fig. 7 is can be concluded that in case of direct modeling of the fatigue strength from the chemical composition and key thermal treatment conditions, i.e., the NonTR models, regardless of whether the model is based on deep learning or machine learning, a large number of data sets (in this study $>$ 148 datasets) is needed. In case of high costs per dataset such a large number of datasets may not be available. From this we may conclude that an insufficient number of fatigue up to now formed the critical issue for fatigue strength-oriented alloy design based on traditional AI methods.

The situation is rather different in case of the TR models for which, R^2 and MAE remain almost constant with relatively small error bars over the entire range of 32$\sim$328 fatigue data sets. It should be noted that there is a significant difference between the SML and CNN TR models due to different transfer mechanisms. First, the CNN framework has slightly lower stability in the range of 32$\sim$66 data points. This means that although the CNN TR model has higher tolerance for extremely small sample databases compared to traditional AI models, 50 sample or less are still insufficient for training a highly stable CNN TR model built on random partitioning of training and testing sets. Even so, the optimal CNN TR model can still adapt to extremely small datasets with 32 fatigue data points, as shown in Fig. 7(c). However, the analysis shows that the TR network is always better than that of the CNN NonTR model for all numbers of training sets used, but in particular for smaller training data sets, which is consistent with a previous study [51]. Clearly the presence of an intermediate model fed by a larger number of (cheaply to acquire) quasi-static mechanical data (i.e., YS, UTS and EL) can make a large and positive contribution to linking the chemical composition and processing conditions to the target rotating bending fatigue strength even trained on a small number of validation data sets.

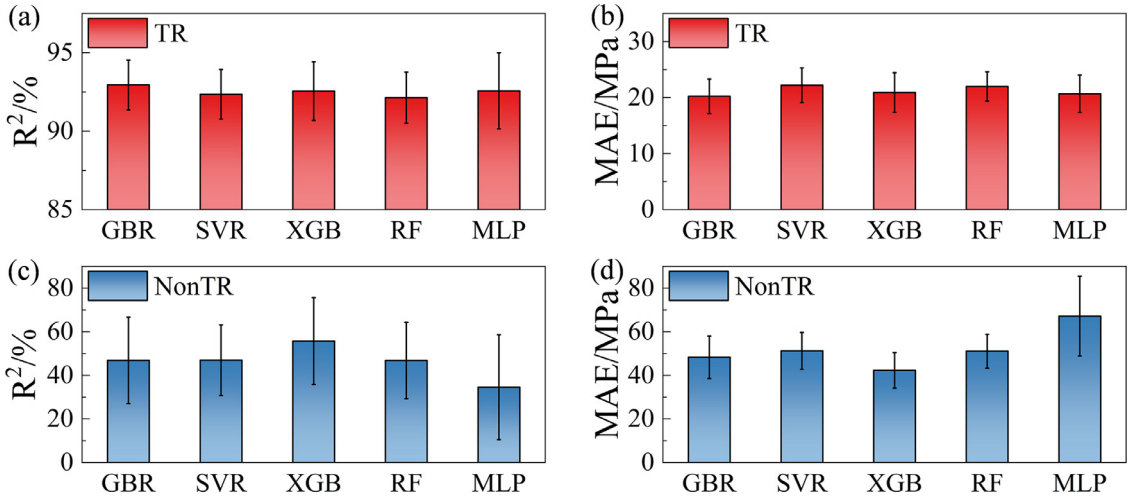

Fig. 6. Mean R^2 and MAE results for (a, b) TR and (c, d) NonTR models using different algorithms.

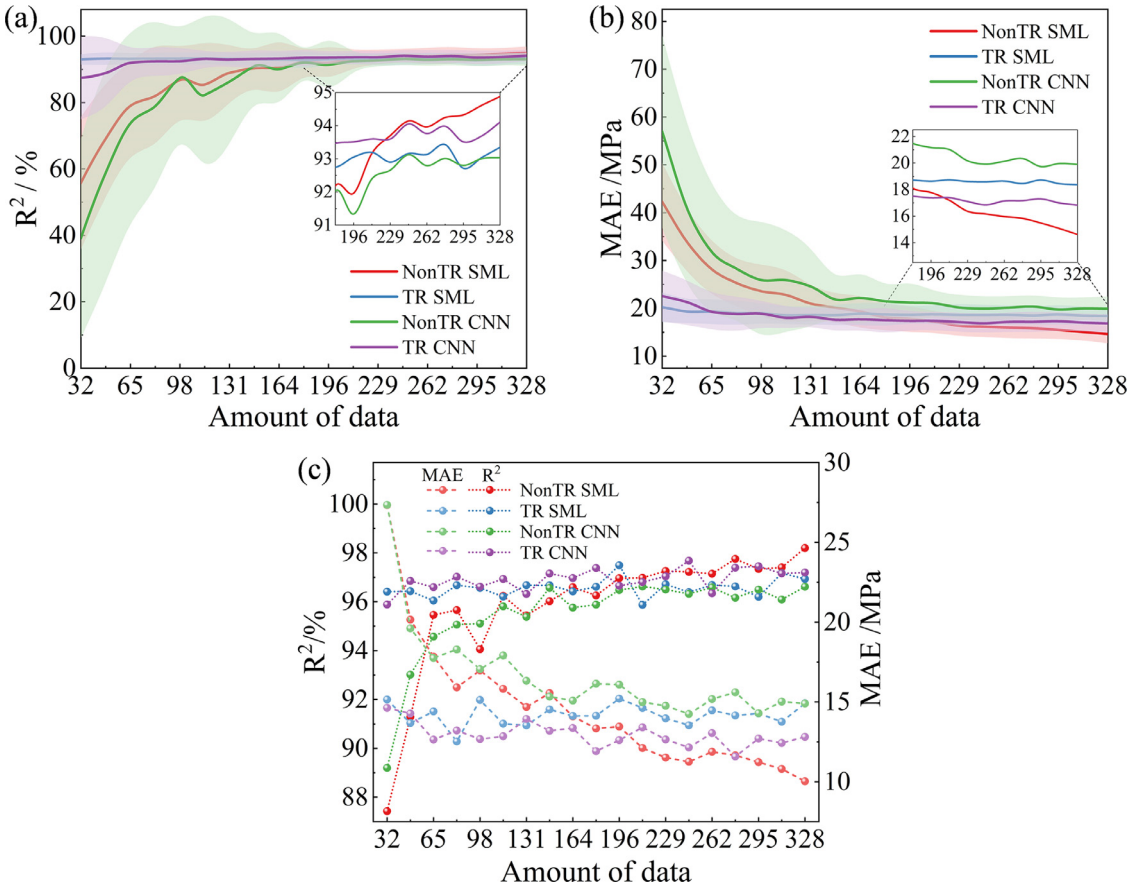

Fig. 7. Comparisons between the TR and NonTR models given different amounts of fatigue data. (a) Mean R^2. (b) Mean MAE. (c) Optimal R^2 and MAE.

4.3. Different loss functions for NonTR models

Changing the loss function may benefit model training and obtain better predictive ability, which approach has been attempted in several studies, such as the work about physics-informed neural networks (PINN) by Zhang et al. [52]. In this section, the performance of NonTR models using different loss functions in SML and CNN framework was investigated. The max-error of 4 output values (YS, UTS, EL and Fatigue strength) was set as the loss, named Tensile+FS. In the SML framework, the NonTR models with different data amount were constructed using the random forest (RF) algorithm. Their performance in the validation set were further compared with the models only using the fatigue strength error (FS) as the loss. Very close performance was found in two cases, as shown in Fig. 8(a) and (b). Besides, the corresponding NonTR models were also constructed using CNN, and the results are shown in Fig. 8(c) and (d). The model parameters and training details of NonTR models using 4 output values error are the same with that of using the fatigue strength error as the loss. Similar with the results of NonTR models using RF, the accuracy of both two types of models is also basically the same when using different loss functions. Meanwhile, a strong dependence of model performance on

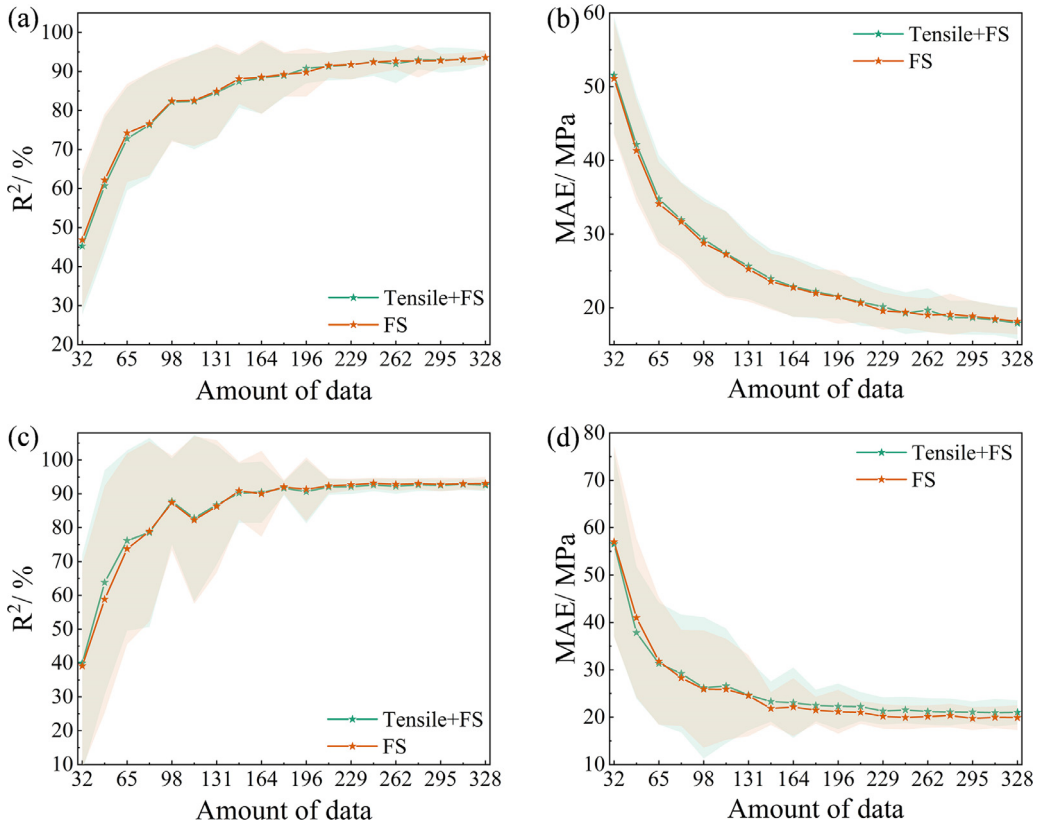

Fig. 8. Comparisons between NonTR models in SML and CNN framework given different amounts of fatigue data using FS and Tensile + FS errors as the loss. (a, b) R^2 and MAE results in SML. (c, d) R^2 and MAE results in CNN.

the amount of data is found for SML and CNN NonTR. NonTR models using modified loss functions are also greatly sensitive to the data amount. Although changing the loss function is a promising approach, the NonTR models performed unsatisfactory as a result of an insufficient amount of data.

4.4. The function of TR on fatigue strength prediction

It should be noted that the architecture of the TR model plays an important but not easily discernable role in establishing a reliable and robust TR model. Therefore, the performance of the TR layer is analyzed in more detail. Fig. 9(a) shows the pair plot corresponding to YS, UTS, EL and fatigue strength for the dataset. For this high quality of the input data, a strong correlation between three quasi-static tensile properties and bending fatigue strength exists. This correlation ultimately forms the basis for accurately predicting the fatigue strength.

The importance of the three tensile properties was further evaluated in the SML TR model by an RF algorithm, and the results are shown in Fig. 9(b). The YS property has the highest contribution to the prediction of fatigue strength in the SML TR model. This YS contribution exceeds 50%, followed by UTS (46.6%) and is only marginally affected by the EL (3.1%).

In order to further clarify the advantages of TR layer, which can consider the comprehensive relationship between tensile properties and fatigue strength, a conventional empirical equation linking fatigue strength to UTS only was used for comparison. The traditional empirical equation used in this research was proposed by Pang [12,14] and named PM-TR in Fig. 9(c) and is given below:

$$\sigma_f = (C - P \cdot UTS) \cdot UTS \tag{4}$$

where C and P are two fitting constants, σ_f is fatigue strength and UTS is the ultimate tensile strength. The comparison of the predictive power of the best fitting PM-TR model linking the fatigue properties and the UTS and that of the TR model is shown in Fig. 9(c) as a function of the number of data sets available. Clearly the TR model taking three quasi-static properties into account performs better than the PM-TR model, which considered only UTS and had a mathematically simple expression. Also, Fig. 9(c) shows an increasing gap for MAE and R^2 between the TR and PM-TR with an increasing amount of data. It indicates the increasing advantages of the TR layer in proposed framework with amount of data increased.

In addition, it should be mentioned that the TR model differs from the above traditional empirical equations. It is a more complex translator from the composition and process parameters to the fatigue strength. The mechanical properties (YS, UTS and EL) are only intermediate information, which can help the models to find the correct relationship from the composition and process to fatigue properties with less data demand. Taken together, the functionality and applicable characteristics of the TR layer are clearly defined. The TR layer in the framework results in better flexibility and robustness for data amount than traditional methods.

4.5. The function of TR on alloy design

The importance of input features on fatigue strength prediction was further investigated by the calculation of MDA, as shown in Fig. 10(a). The MDA results show that the composition/process parameters have similar relevance for the tensile properties and the fatigue strength. Tempering temperature (TT), C and Cr are the three most important features. The importance of TT and C on tensile and fatigue strength was to be expected since the tempering temperature significantly influences both the matrix and precipitates, and C is also the most important strengthening element. The

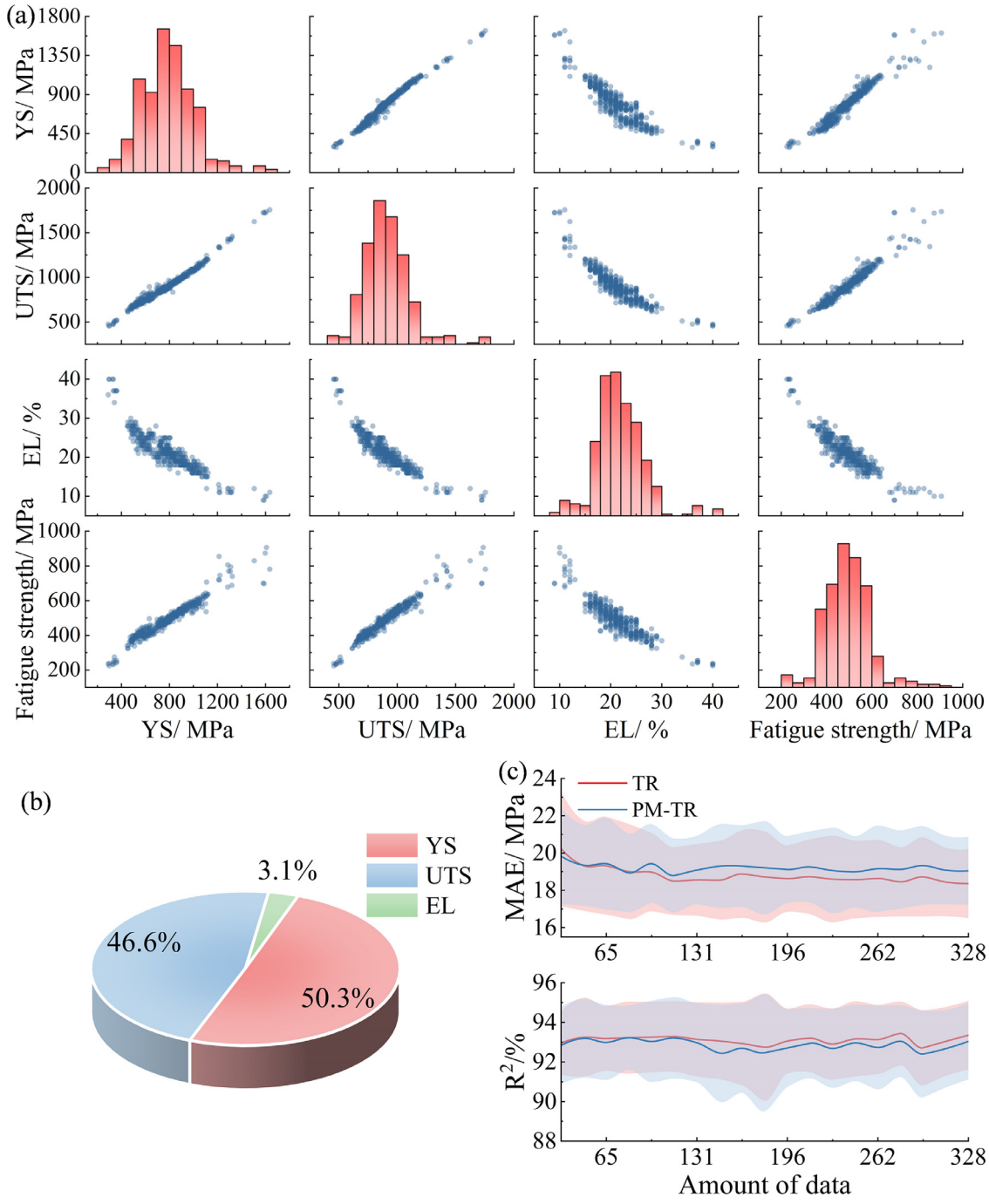

Fig. 9. (a) Distribution of UTS, EL and fatigue strength of the dataset. (b) Importance of three tensile properties to fatigue strength prediction. (c) R^2 and MAE comparisons between the TR and PM-TR models given different amounts of fatigue data.

importance of Cr may have been a little magnified in this analysis because of the inclusion of high-Cr stainless steel data sets in this dataset. The observation confirms the earlier and statistically logical conclusion that dataset should not be used for designing alloys that are extremely different from the composition range of the training dataset.

To further clarify the efficiency of TR models for the alloy design, alloy design was also conducted using NonTR models, which were also trained based on the same small data set with the TR models. In addition, alloys were also designed using ML models trained with the whole original dataset. Based on the relevance analysis results presented in Fig. 10(a), the tempering temperature and C content for the alloys designed by both the NonTR and the TR models are compared in Fig. 10(b) and 10(c), respectively.

As shown in Fig. 10(b), the TT of alloys designed on TR models with extremely small datasets are concentrated in the range of 550~560 MPa, which is basically consistent with the design results of NonTR models with whole data. Additionally, similar C content results are shown for the comparison between the alloys designed via NonTR models with all data and TR models with extremely small datasets, as shown in Fig. 10(c). The results indicate that with the help of TR layer, an effective alloy could be designed based on only tens of fatigue data, and the design results are basically equivalent to the results of traditional AI models with hundreds of fatigue data for training. However, for NonTR models

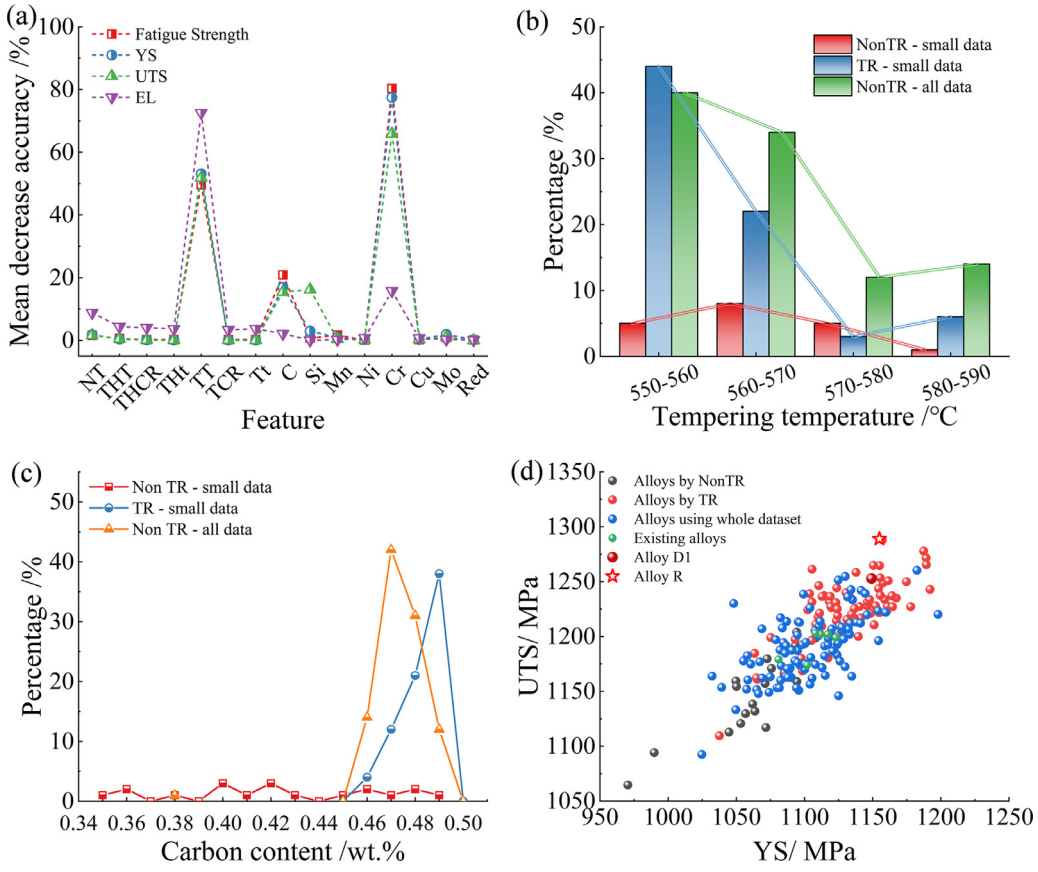

Fig. 10. (a) Effects of input features on fatigue strength and tensile properties prediction. (b-d) Comparisons of the distributions of alloys designed based on NonTR, TR models and models using the whole dataset: (b) tempering temperature. (c) carbon content. (d) Predicted tensile properties.

based on extremely small datasets, the design results show a different distribution of both C content and TT compared to TR models. A much wider variation in C content (0.34~0.49 wt.%) was obtained in the design results from NonTR models, indicating that a limited amount of data could significantly decrease the reliability and efficiency of the designs generated by traditional AI models (Fig. 10(c)). In summary, this comparison fully proved that the TR model used in this research is a more effective method for alloy design with much less data amount requirement than traditional AI methods.

To further analyze the mechanism of TR layer leading to more effective design, distributions of designed alloys by TR/NonTR models are plotted in YS-UTS space in Fig. 10(d). It can be observed that the predicted YS and UTS of alloys designed by TR models are concentrated in the region of high levels, which are similar with that designed by models trained using the whole dataset. According to the general relationship between tensile strength and fatigue strength summarized by previous studies [12], high level of YS and UTS could lead to high fatigue strength. However, the designed results by NonTR models exhibit relatively low predicted YS and UTS. They may not be valuable and reliable for optimal fatigue strength. Additionally, Alloy D1, Alloy R and existing alloys (Alloys E1 to E6) are also plotted in Fig. 10(d). Most of alloys designed by TR models are superior to existing alloys, while the results by NonTR models are generally inferior. Alloy D1 exhibits high predictions for YS (1149 MPa) and UTS (1253 MPa), being confirmed by the experimental validation results of Alloy R (YS: 1155 MPa; UTS: 1289 MPa).

In a word, the discussion above indicates that the TR layer do play an important role in alloy design and lead to more effective alloy design based on only tens of samples.

4.6. The portability analysis for the two transfer learning framework

In fact, lack of data is not a special example only for fatigue. It is a generic problem for various cases, which greatly limited the development of AI strategies in the field of materials science. Although this research only focused on the fatigue strength prediction and related alloy design, the TR framework proposed in this research is expected to be used for other properties with long testing time or high cost, such as creeping, hydrogen embrittlement, etc. For fatigue, the traditional theory provided a clear guidance that fatigue strength has strong relation with tensile properties. So, the TR layers could be directly established between tensile properties and fatigue strength. However, when the TR framework is transfer to other cases with unclear mechanism, it would be hard to find the appropriate traditional properties to transfer from. E.g., it is difficult to decide using YS, fracture toughness or both of them as the transfer source of hydrogen embrittlement. So, in order to clarify the portability of the two TR framework proposed in this research, more analysis was made in this section to evaluate the robustness of the TR models on wrong transfer sources.

Firstly, in addition to YS, UTS and EL, two more mechanical properties, i.e., the hardness and the impact toughness, were further added as the optional transfer source. Compared to the basic model with only YS, UTS and EL as transfer source, Fig. 11 shows the R^2 and MAE changes in the validation set for the TR models using different transfer sources.

It is generally believed that impact toughness (IT) exhibits a relatively lower correlation with fatigue strength compared to other quasi-static mechanical properties used in present work. Meanwhile, the EL contributes the least to the prediction in the basic model. Hence, when the EL and IT were used as the intermedi-

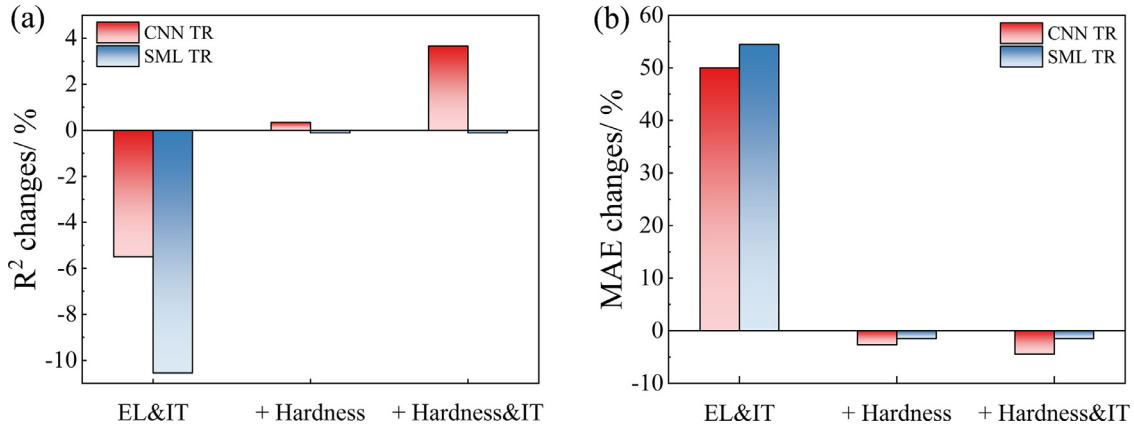

Fig. 11. Comparisons of the changes in (a) R^2 and (b) MAE using different source properties compared to the basic model.

ate properties for the TR models, the accuracy of the TR models markedly decreased, as shown in Fig. 11. Interestingly, there are significant differences between the two TR frameworks. The CNN TR model shows a lower degree of accuracy decline compared to the SML TR, which indicates it is more robust and has a greater relative tolerance when faced to an ambiguous mechanism such as using less relevant intermediate properties.

Based on the basic model, when adding the hardness parameter (intermediate property highly related to fatigue strength) to the transfer layer, the accuracy of two TR frameworks slightly improved. And when IT was further added into the transfer layer, the performance of CNN TR model was further improved significantly, while the SML TR model basically maintained the same performance level as the basic model (or the TR model adding the hardness). The result further demonstrates the robustness of CNN TR.

Summarizing when transferring the TR model for other property predictions or steel designs, the SML TR framework is expected to be more suitable for applications with a relatively clear physical connection between the target and source properties, as it then has a superb strong robustness for the number of datasets available for the model training and validation. In contrast, the CNN TR framework has a stronger tolerance for the unclear connection with the source properties, but it needs relatively more data than SML TR framework to come to accurate predictions.

Conclusion

To provide an efficient and portable method for fatigue strength prediction and fatigue strength oriented alloy design of metal materials, in the present work, a transfer (TR) framework including Conventional Neural Network (CNN) TRs and Simplified Machine Learning (SML) TRs has been proposed which utilizes the correlations between quasi-static tensile properties and fatigue strength.

(1) A comparative study of the TR framework and NonTR models shows that given the high-dimensional features of the compositions and processing parameters of steels, transfer learning requires fewer (expensive) data points than traditional AI strategies for training a reliable model to predicted the fatigue strength, provided that a large yet cheap data base linking the quasi-static mechanical properties to steel composition and processing parameters is available. Guided by the strong correlations between tensile properties and fatigue strength, the TR framework accurately predicted fatigue strength upon training with only tens of fatigue data points. Using easily and cheaply to acquire tensile property data to greatly decrease the demand for expensive fatigue data, the TR framework can greatly reduce both the time and monetary costs of data accumulation that are required for building an accurate fatigue property prediction model.

(2) Combined with the evolutionary algorithm, the SML TR framework was further used for fatigue strength oriented alloy design. Experimental validation shows the high reliability of this TR framework for alloy design, especially for the carbon content and tempering temperature, which is proved to have strength relation with fatigue strength. The designed alloy by TR models was validated for improved fatigue strength.

(3) Although this research only focused on the fatigue strength prediction and related alloy design, the TR framework proposed in this research is expected to be used for other properties with long testing time or high cost. The SML TR framework is more suitable for the application with a relatively clear mechanism between the target and source properties, but it has extremely strong robustness for extremely small data amount. In contrary, the CNN-TR framework has stronger tolerance for the unclear mechanism or source properties, but it needs relatively more data than the SML TR framework.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research was financially supported by National Key R&D Program (No. 2021YFE0204100). The financial support provided by National Natural Science Foundation of China (Nos. U1808208, 51961130389) is gratefully acknowledged.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2022.118103.

References

[1] M.A. Meyers, K.K. Chawla, Mechanical Behavior of Materials, Cambridge University Press, 2008.
[2] L. Tóth, S.Y. Yarema, Formation of the science of fatigue of metals. Part 1. 1825–1870, Mater. Sci. 42 (2006) 673–680.
[3] W. Schütz, A history of fatigue, Eng. Fract. Mech. 54 (1996) 263–300.
[4] W. Xu, M. Huang, J. Wang, C. Shen, T. Zhang, C. Wang, Relations between metastable austenite and fatigue behavior of steels, Acta Metall. Sin. 56 (2020) 459–475.
[5] W.J. Dixon, A.M. Mood, A method for obtaining and analyzing sensitivity data, J. Am. Stat. Assoc. 43 (1948) 109–126.

[6] H. Zhang, H. Fu, X. He, C. Wang, L. Jiang, L. Chen, J. Xie, Dramatically enhanced combination of ultimate tensile strength and electric conductivity of alloys via machine learning screening, Acta Mater. 200 (2020) 803–810.
[7] C. Zou, J. Li, W.Y. Wang, Y. Zhang, D. Lin, R. Yuan, X. Wang, B. Tang, J. Wang, X. Gao, Integrating data mining and machine learning to discover high--strength ductile titanium alloys, Acta Mater. 202 (2021) 211–221.
[8] C. Shen, C. Wang, X. Wei, Y. Li, S. van der Zwaag, W. Xu, Physical metallurgy-guided machine learning and artificial intelligent design of ultrahigh--strength stainless steel, Acta Mater. 179 (2019) 201–214.
[9] C. Wang, C. Shen, X. Huo, C. Zhang, W. Xu, Design of comprehensive mechanical properties by machine learning and high-throughput optimization algorithm in RAFM steels, Nucl. Eng. Technol. 52 (2019) 1008–1012.
[10] S.H. Park, C.S. Lee, Relationship between mechanical properties and high-cycle fatigue strength of medium-carbon steels, Mater. Sci. Eng. A 690 (2017) 185–194.
[11] B. Wang, P. Zhang, Q.Q. Duan, Z.J. Zhang, H.J. Yang, X.W. Li, Z.F. Zhang, Optimizing the fatigue strength of 18Ni maraging steel through ageing treatment, Mater. Sci. Eng. A 707 (2017) 674–688.
[12] J. Pang, S. Li, Z. Wang, Z. Zhang, Relations between fatigue strength and other mechanical properties of metallic materials, Fatigue Fract. Eng. Mater. Struct. 37 (2014) 958–976.
[13] P.G. Forrest, Fatigue of Metals, Elsevier, 2013.
[14] J. Pang, S. Li, Z. Wang, Z. Zhang, General relation between tensile strength and fatigue strength of metallic materials, Mater. Sci. Eng. A 564 (2013) 331–341.
[15] M. Yukitaka, E. Masahiro, Quantitative evaluation of fatigue strength of metals containing various small defects or cracks, Eng. Fract. Mech. 17 (1983) 1–15.
[16] S.X. Li, Effects of inclusions on very high cycle fatigue properties of high strength steels, Int. Mater. Rev. 57 (2012) 92–114.
[17] Y.B. Liu, Z.G. Yang, Y.D. Li, S.M. Chen, S.X. Li, W.J. Hui, Y.Q. Weng, Dependence of fatigue strength on inclusion size for high-strength steels in very high cycle fatigue regime, Mater. Sci. Eng. A 517 (2009) 180–184.
[18] B. Wang, P. Zhang, R. Liu, Q.Q. Duan, Z.J. Zhang, X.W. Li, Z.F. Zhang, An optimization criterion for fatigue strength of metallic materials, Mater. Sci. Eng. A 736 (2018) 105–110.
[19] R. Liu, Z. Zhang, L. Li, X. An, Z. Zhang, Microscopic mechanisms contributing to the synchronous improvement of strength and plasticity (SISP) for TWIP copper alloys, Sci. Rep. 5 (2015) 9550.
[20] C. Wang, H. Fu, L. Jiang, D. Xue, J. Xie, A property-oriented design strategy for high performance copper alloys via machine learning, npj Comput. Mater. 5 (2019) 87.
[21] K. Song, F. Yan, T. Ding, L. Gao, S. Lu, A steel property optimization model based on the XGBoost algorithm and improved PSO, Comput. Mater. Sci. 174 (2020) 109472.
[22] C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, Y. Su, Machine learning assisted design of high entropy alloys with desired property, Acta Mater. 170 (2019) 109–117.
[23] Q. Wu, Z. Wang, X. Hu, T. Zheng, Z. Yang, F. He, J. Li, J. Wang, Uncovering the eutectics design by machine learning in the Al–Co–Cr–Fe–Ni high entropy system, Acta Mater. 182 (2020) 278–286.
[24] Y. Liu, J. Wu, Z. Wang, X. Lu, M. Avdeev, S. Shi, C. Wang, T. Yu, Predicting creep rupture life of Ni-based single crystal superalloys using divide-and-conquer approach based machine learning, Acta Mater. 195 (2020) 454–467.
[25] Y. Wang, Y. Tian, T. Kirk, O. Laris, J.H. Ross, R.D. Noebe, V. Keylin, R. Arróyave, Accelerated design of Fe-based soft magnetic materials using machine learning and stochastic optimization, Acta Mater. 194 (2020) 144–155.
[26] M. Zhang, C. Sun, X. Zhang, P.C. Goh, J. Wei, D. Hardacre, H. Li, High cycle fatigue life prediction of laser additive manufactured stainless steel: a machine learning approach, Int. J. Fatigue 128 (2019) 105194.
[27] L. Zhang, J. Lei, Q. Zhou, Y. Wang, Using genetic algorithm to optimize parameters of support vector machine and its application in material fatigue life prediction, Adv. Nat. Sci. 8 (2015) 21–26.
[28] A. Rovinelli, M.D. Sangid, H. Proudhon, W. Ludwig, Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials, npj Comput. Mater. 4 (2018) 35.
[29] T. Shiraiwa, Y. Miyazawa, M. Enoki, Prediction of fatigue strength in steels by linear regression and neural network, Mater. Trans. 60 (2018) 189–198.
[30] D.K. Choi, Data-driven materials modeling with XGBoost algorithm and statistical inference analysis for prediction of fatigue strength of steels, Int. J. Precis. Eng. Manuf. 20 (2019) 129–138.
[31] A. Keprate, R.C. Ratnayake, Data mining for estimating fatigue strength based on composition and process parameters, in: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering, American Society of Mechanical Engineers, 2019 V004T03A017.
[32] B. Gautham, R. Kumar, S. Bothra, G. Mohapatra, N. Kulkarni, K. Padmanabhan, More efficient ICME through materials informatics and process modeling, in: Proceedings of the 1st World Congress on Integrated Computational Materials Engineering (ICME), Wiley Online Library, 2011, p. 35.
[33] P. Deshpande, B. Gautham, A. Cecen, S. Kalidindi, A. Agrawal, A. Choudhary, Application of statistical and machine learning techniques for correlating properties to composition and manufacturing processes of steels, in: Proceedings of the 2nd World Congress on Integrated Computational Materials Engineering (ICME), Springer, 2013, pp. 155–160.
[34] A. Agrawal, P.D. Deshpande, A. Cecen, G.P. Basavarsu, A.N. Choudhary, S.R. Kalidindi, Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters, Integr. Mater. Manuf. Innov. 3 (2014) 8.
[35] A. Agrawal, A. Choudhary, A fatigue strength predictor for steels using ensemble data mining, in: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, 2016, pp. 2497–2500.
[36] A. Agrawal, A. Choudhary, An online tool for predicting fatigue strength of steel alloys based on ensemble data mining, Int. J. Fatigue 113 (2018) 389–400.
[37] T. Shiraiwa, F. Briffod, Y. Miyazawa, M. Enoki, Fatigue performance prediction of structural materials by multi-scale modeling and machine learning, in: Proceedings of the 4th World Congress on Integrated Computational Materials Engineering (ICME 2017), Springer, 2017, pp. 317–326.
[38] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2009) 1345–1359.
[39] A. Paul, D. Jha, R. Al-Bahrani, W. Liao, A. Choudhary, A. Agrawal, Transfer learning using ensemble neural networks for organic solar cell screening, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN), IEEE, 2019, pp. 1–8.
[40] H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, R. Yoshida, Predicting materials properties with little data using shotgun transfer learning, ACS Cent. Sci. 5 (2019) 1717–1730.
[41] B. Kailkhura, B. Gallagher, S. Kim, A. Hiszpanski, T.Y.J. Han, Reliable and explainable machine-learning methods for accelerated material discovery, npj Comput. Mater. 5 (2019) 1–9.
[42] M. Tsubaki, T. Mizoguchi, Quantum deep descriptor: physically informed transfer learning from small molecules to polymers, J. Chem. Theory Comput. (2021).
[43] S. Feng, H. Fu, H. Zhou, Y. Wu, Z. Lu, H. Dong, A general and transferable deep learning framework for predicting phase formation in materials, npj Comput. Mater. 7 (2021) 10.
[44] M.L. Hutchinson, E. Antono, B.M. Gibbons, S. Paradiso, J. Ling, B. Meredig, Overcoming data scarcity with transfer learning, arXiv preprint arXiv:1711.05099 (2017).
[45] J. Ling, M. Hutchinson, E. Antono, S. Paradiso, B. Meredig, High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates, Integr. Mater. Manuf. Innov. 6 (2017) 207–217.
[46] NIMS (2022) Fatigue data sheet, https://smds.nims.go.jp/fatigue/, [1 November 2019].
[47] Y. Furuya, H. Nishikawa, H. Hirukawa, N. Nagashima, E. Takeuchi, Catalogue of NIMS fatigue data sheets, Sci. Technol. Adv. Mater. 20 (2019) 1055–1072.
[48] A. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, Pattern Recognit. 38 (2005) 2270–2285.
[49] C. Wang, K. Zhu, P. Hedström, Y. Li, W. Xu, A generic and extensible model for the martensite start temperature incorporating thermodynamic data mining and deep learning framework, J. Mater. Sci. Technol. 128 (2022) 31–43.
[50] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks? Adv. Neural Inf. Process. Syst. (2014) 3320–3328.
[51] D. Jha, K. Choudhary, F. Tavazza, W.K. Liao, A. Choudhary, C. Campbell, A. Agrawal, Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning, Nat. Commun. 10 (2019) 5316.
[52] E. Zhang, M. Dao, G.E. Karniadakis, S. Suresh, Analyses of internal structures and defects in materials using physics-informed neural networks, Sci. Adv. 8 (2022) eabk0644.