ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Mécanique

Lei Han, Chunmei Duan and Taochuan Zhang

Predicting stress–strain constitutive relationships: a deep learning approach based on multi-head attention mechanism and long short-term memory networks

Volume 353 (2025), p. 953-988

Online since: 7 August 2025

https://doi.org/10.5802/crmeca.315

The Comptes Rendus. Mécanique *are a member of the Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1873-7234

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mécanique
2025, Vol. 353, p. 953-988
https://doi.org/10.5802/crmeca.315

Research article / *Article de recherche*

Predicting stress–strain constitutive relationships: a deep learning approach based on multi-head attention mechanism and long short-term memory networks

Prédiction des relations constitutives contrainte-déformation : une approche d'apprentissage profond basée sur le mécanisme d'attention multi-tête et les réseaux à mémoire courte et longue (LSTM)

Lei Han [a,*], **Chunmei Duan** [b,c] **and Taochuan Zhang** [b,c,*]

[a] School of Automobile and Transportation Engineering, Guangdong Polytechnic Normal University, Guangzhou, 510665, China

[b] School of Mechanical Engineering, Guangdong Ocean University, Zhanjiang, 524088, China

[c] Guangdong Engineering Technology Research Center of Ocean Equipment and Manufacturing, Zhanjiang, 524088, China

E-mails: hanlei@gpnu.edu.cn (L. Han), zhangtaochuan@hotmail.com (T. Zhang)

Abstract. The stress–strain constitutive relationship of materials is a critical indicator for assessing their mechanical properties and serves as a vital bridge connecting microstructures with macroscopic mechanical behaviors. This study focuses on 106 metal alloy materials, employing the finite element method to establish uniaxial tensile numerical models, thereby simulating and obtaining a comprehensive dataset of stress–strain relationship curves. Innovatively, this research proposes a MA-LSTM-DLnet prediction model that integrates multi-head attention mechanisms with long short-term memory networks, based on deep learning methodologies. The model utilizes conventional key parameters such as material strain, yield strength, tensile strength, elastic modulus, density, strength coefficient, and stress as input features to construct a multi-dimensional training dataset. It particularly explores the learning capabilities and predictive performance of the MA-LSTM-DLnet model under both complete and limited feature datasets. The results demonstrate that the trained MA-LSTM-DLnet model can accurately predict the stress–strain constitutive relationship curves of new materials by merely acquiring conventional physical parameters like yield strength, tensile strength, and elastic modulus, with a similarity exceeding 95% compared to test data. Compared to traditional empirical formulas and numerical simulation methods, the main innovation of this study lies in eliminating key parameters that are difficult to determine in mathematical models (such as the strength coefficient in the Ludwik isotropic hardening model), significantly reducing repetitive work dependent on experiments to obtain constitutive relationships, and achieving efficient and precise material performance evaluation.

* Corresponding authors

Résumé. La relation constitutive contrainte-déformation des matériaux est un indicateur essentiel pour évaluer leurs propriétés mécaniques, servant de lien crucial entre les microstructures et les comportements mécaniques macroscopiques. Cette étude porte sur 106 alliages métalliques, utilisant la méthode des éléments finis pour établir des modèles numériques de traction uniaxiale, permettant ainsi de simuler et d'obtenir un ensemble de données complet de courbes contrainte-déformation. De manière innovante, cette recherche propose un modèle de prédiction MA-LSTM-DLnet intégrant des mécanismes d'attention multi-tête avec des réseaux de mémoire courte et longue (LSTM), basé sur des méthodologies d'apprentissage profond. Le modèle utilise des paramètres clés conventionnels tels que la déformation du matériau, la limite d'élasticité, la résistance à la traction, le module d'élasticité, la densité, le coefficient de résistance et la contrainte comme caractéristiques d'entrée pour construire un ensemble de données d'entraînement multidimensionnel. L'étude explore particulièrement les capacités d'apprentissage et les performances prédictives du modèle MA-LSTM-DLnet, aussi bien avec des ensembles de données complets qu'avec des données limitées. Les résultats démontrent que le modèle entraîné peut prédire avec précision les courbes de relations constitutives contrainte-déformation de nouveaux matériaux, en se basant uniquement sur des paramètres physiques conventionnels comme la limite d'élasticité, la résistance à la traction et le module d'élasticité, avec une similarité dépassant 95 % par rapport aux données de test. Comparé aux formules empiriques traditionnelles et aux méthodes de simulation numérique, la principale innovation de cette étude réside dans l'élimination de paramètres clés difficiles à déterminer dans les modèles mathématiques (comme le coefficient de résistance dans le modèle d'écrouissage isotrope de Ludwik), réduisant ainsi considérablement le travail expérimental répétitif nécessaire pour obtenir les relations constitutives, et permettant une évaluation efficace et précise des performances des matériaux.

Keywords. Deep learning, Stress–strain, Attention mechanism, Long short-term memory network, Material performance prediction.

Mots-clés. Apprentissage profond, Contrainte-déformation, Mécanisme d'attention, Réseau de mémoire courte et longue (LSTM), Prédiction des performances des matériaux.

Funding. Guangdong Polytechnic Normal University (Grant Number 2021SDKYA017), Program for Scientific Research Start-up Funds of Guangdong Ocean University (Grant Numbers 060302062311, 060302062315), Collaborative Education Project of Industry University Cooperation by the Ministry of Education of China (Grant numbers 230802784214645, 230803817223502, 230821384507208).

Manuscript received 16 April 2025, revised 14 June 2025, accepted 8 July 2025.

1. Introduction

The constitutive relationship of materials is a crucial mathematical model in the field of material mechanics, as it precisely describes the correlation between internal stress and strain during the deformation process. By providing a strain value, this model can accurately calculate the corresponding stress, thereby offering a theoretical foundation for predicting and analyzing the mechanical behavior of materials under various working conditions. For instance, the stress–strain relationship enables in-depth analysis of material performance under different temperatures, loads, and time conditions, which is of significant importance for engineering applications and product design [1–5]. Therefore, a thorough investigation of the constitutive relationship not only enhances our understanding of the mechanical responses of materials in practical applications but also significantly improves the reliability and durability of products.

Currently, there are two primary approaches to obtaining the stress–strain constitutive relationship of materials: experimental testing and empirical formulas or numerical simulations. Common experimental methods include tensile testing, which is applicable to a wide range of materials such as metals, plastics, and elastomers [6–10], however, these methods may not be suitable for rare or expensive materials and are often time-consuming. Among the classical empirical formulas, the Ramberg–Osgood formulation is widely used [11]. Under the assumption of isotropic materials, this formula calculates stress and strain by incorporating parameters such as elastic modulus, yield strength, tensile strength, and hardening exponent. However, some key

parameters in the model still require experimental determination based on the material's properties [11]. Additionally, the Johnson-Cook and Ludwik models are extensively employed to simulate the nonlinear constitutive relationships of materials and are embedded in many commercial finite element software packages. Nevertheless, these models also rely on experimentally determined parameters [12,13]. Many scholars have combined experimental and simulation methods to establish or fit constitutive relationships, but these models are often limited to specific materials or working conditions [14–20].

Despite advancements in mathematical and physical theories, it remains challenging to establish a universal mathematical formula or model for the stress–strain relationship that applies to all material types or even within a broad category, such as metal alloys. Currently, the only approach is to conduct experiments for each alloy to obtain relevant parameters, which are then used in empirical formulas or simulations, or to directly perform tensile tests to acquire the corresponding stress–strain curves. With the development of artificial intelligence (AI) technologies, the field of materials science has gradually embraced AI. In recent years, some researchers have utilized machine learning or deep learning methods to predict material properties [21–32]. Unlike traditional mathematical models that establish explicit mapping relationships, deep learning prediction models are based on a generalized, fuzzy mapping relationship between input and output variables, mimicking the neural networks of the brain [33]. This fundamental difference makes it possible to develop a unified deep learning model for predicting the constitutive relationships of materials within a broad category.

Recently, a small number of scholars have begun to explore the use of deep learning methods to predict stress–strain relationship curves [34–40]. For example, El Fallaki Idrissi et al. [39] proposed an advanced meta-modeling framework that integrates machine learning and Model Order Reduction (MOR) techniques. They employed Proper Orthogonal Decomposition (POD) to decompose stress–strain curves and used a Random Forest machine learning model for interpolation. This enables the rapid prediction of macroscopic nonlinear stress–strain curves under varying microstructural parameters. Nikolic et al. [34] developed an automated microstructure generation technique combined with finite element analysis to obtain stress–strain curves under different microstructures. They trained a convolutional neural network (CNN) on microstructure image data to predict the stress–strain response of materials at varying temperatures. Koenuma et al. [35] synthesized material microtextures using numerical methods and computed corresponding stress–strain curves, employing CNNs to capture the features of synthetic textures and predict stress–strain curves under different material textures. Zhang et al. [36] applied long short-term memory (LSTM) deep learning methods to replicate the stress–strain behavior of soils, investigating the ability of LSTM to account for the stress history of soil behavior. Saha et al. [37] utilized a deep learning (DL) framework, specifically a 3D U-Net architecture, to map composite microstructures to corresponding stress and strain fields under predefined loading paths, predicting the evolution of microscale elastoplastic strain and stress in two-phase media. Chen et al. [38] simulated material failure processes using the phase-field method, generating composite microstructures and corresponding stress–strain curves as training and testing data. They employed a conditional generative adversarial network (cGAN) as the deep learning model to predict stress–strain curves. El Fallaki Idrissi et al. [40] presented a novel method named Multiscale Thermodynamics-Informed Neural Networks (MuTINN). It can quickly and cost—effectively compute the inelastic behavior of woven composite structures. On the premise of maintaining thermodynamic consistency, MuTINN can substantially decrease the need for training data. Moreover, it has been successfully incorporated into commercial finite element software, accelerating the multiscale simulation of composite material structures.

In summary, although some researchers have begun to explore the application of deep learning methods to material constitutive relationships, most studies rely on classical and basic neural

network models trained on microstructure images. These approaches essentially perform image classification to derive stress–strain relationship curves. However, predicting material constitutive relationships directly using conventional material properties—such as elastic modulus, yield strength, tensile strength, and density—as inputs to neural network models remains challenging for conventional architectures. To address this limitation, this study proposes a novel deep learning neural network model, MA-LSTM-DLnet, based on long short-term memory (LSTM) networks [41,42] and integrated into a multi-head attention mechanism [43]. This model incorporates a built-in masking and decoding mechanism with a 4-head attention mechanism. By training and validating the model on stress–strain data from 106 alloy materials, we have developed a new and efficient method for calculating material stress–strain relationships. This method requires only conventional material properties—yield strength, tensile strength, and elastic modulus—as input parameters to predict the stress–strain relationship curve. Compared to traditional mathematical models and tensile testing methods, this approach eliminates the need for difficult-to-determine parameters that typically require experimental or empirical derivation. Moreover, it significantly reduces the repetitive work of obtaining constitutive relationships through tensile tests, thereby greatly saving time and labor costs.

2. Method

The research methodology and overall workflow employed in this study are illustrated in Figure 1. First, the finite element method (FEM) was utilized to establish uniaxial tensile simulation models for 106 alloy materials under room temperature conditions. Through computational analysis, the stress–strain relationship curves for each material at room temperature were obtained. These curves, combined with conventional material properties such as yield strength, tensile strength, elastic modulus, and density, were used as characteristic parameters to represent the materials. This process resulted in the creation of a comprehensive dataset comprising 106 time-series entries, each containing complete stress–strain data, which was subsequently used for training and validating the deep learning network model. Second, the established deep learning model was trained on the dataset to develop a high-precision predictive model for stress–strain relationships. The predictive performance of the model was investigated under both complete and limited material feature conditions. Finally, the conventional material properties of the 106 alloy materials used in this study, along with the data and instructions required for finite element simulations, are provided in Appendix A.

2.1. *FEM simulation for stress–strain*

This study selected a total of 106 metal alloy materials from six major categories, as detailed in Appendix A. To obtain the stress–strain relationship curves for all alloy materials, a mathematical model for material tensile behavior was established, and the finite element method (FEM) was employed to simulate the tensile behavior of each metal alloy material under specific heat treatment conditions. Under the assumptions of isotropy, continuity, and small strain rates during tensile deformation, the mathematical model for material tensile behavior can be represented by the stress equilibrium differential equation, geometric equation, and constitutive equation [44]. Since the tensile deformation considers small strain rates, a static tensile simulation was adopted. The stress equilibrium differential equation and geometric equation are as follows:

$$\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{f} = 0 \tag{1}$$

where, $\boldsymbol{\sigma}$ is the stress tensor, and $\boldsymbol{f}$ represents the body force, which can be neglected for conventional static tensile simulations.

$$\boldsymbol{\varepsilon} = \tfrac{1}{2}(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\boldsymbol{T}}) \tag{2}$$

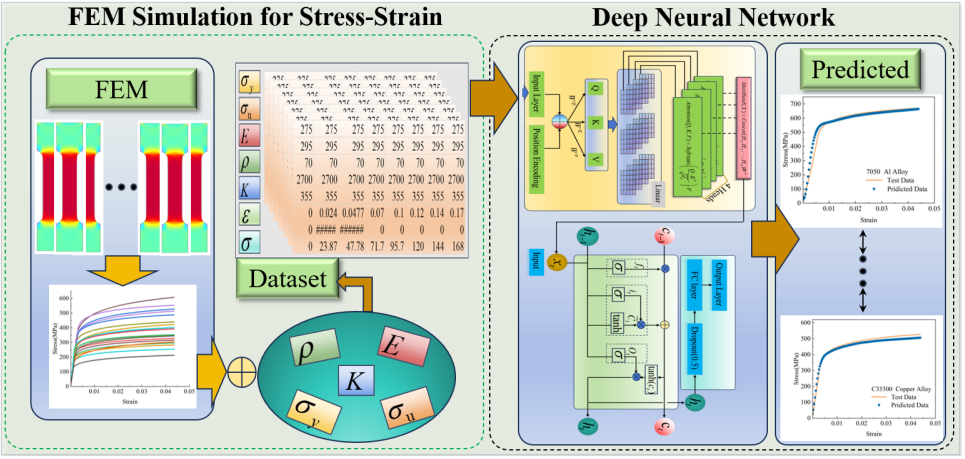


Figure 1. Research method flowchart.

where $\boldsymbol{\varepsilon}$ denotes the strain tensor, $\boldsymbol{u}$ represents the displacement field, and $(\nabla \boldsymbol{u})^{\boldsymbol{T}}$ is the transpose of the displacement gradient tensor. During the tensile process, materials undergo both elastic and plastic deformation. Therefore, the generalized Hooke's law was adopted to describe the constitutive relationship. The Ludwik constitutive equation, an empirical model based on experimental data, effectively describes the relationship between flow stress and plastic strain during the plastic deformation stage [12]. Consequently, the Ludwik hardening model was selected as the constitutive equation for the tensile simulation, expressed as:

$$\sigma = E(\varepsilon_{\mathrm{t}} - \varepsilon_{\mathrm{p}}) + K(\varepsilon_{\mathrm{p}})^{n} \tag{3}$$

where E is the elastic modulus, ε_{t} represents the total strain, ε_{p} denotes the plastic strain, and K and n are parameters required for the Ludwik hardening model, with K representing the strength coefficient and n the hardening exponent. Additionally, the plastic flow rule, yield function, and Karush–Kuhn–Tucker (KKT) conditions are given as follows:

$$\varepsilon_{\mathrm{p}} = \lambda \frac{\partial f}{\partial \sigma} \tag{4}$$

$$f(\sigma, \sigma_y) = \sqrt{\frac{3}{2}\sigma_{\mathrm{eq}}^2} - \sigma_y \tag{5}$$

$$\lambda \geq 0, \quad f \leq 0, \quad \lambda f = 0 \tag{6}$$

where λ is the plastic multiplier, f is the yield function, σ_y represents the yield stress, and σ_{eq} denotes the equivalent stress. Equations (1) through (6) collectively form the mathematical model for the tensile simulation. By incorporating the geometric shape of the tensile specimen, corresponding boundary conditions, and material property parameters, the finite element numerical algorithm was employed to solve these equations, thereby obtaining the stress–strain curve data for the respective materials.

In accordance with the Chinese National Standard (GB/T 228.1-2010), this study employed rectangular cross-section tensile specimens for the simulation of tensile behavior. The specific dimensions of the tensile specimens are illustrated in Figure 2. For the geometric model, boundary conditions were applied as follows: one side of the specimen was fixed, while the other side was subjected to a displacement boundary condition with a uniform tensile length of 15 mm. The remaining boundaries were set as free. By integrating these boundary conditions with the

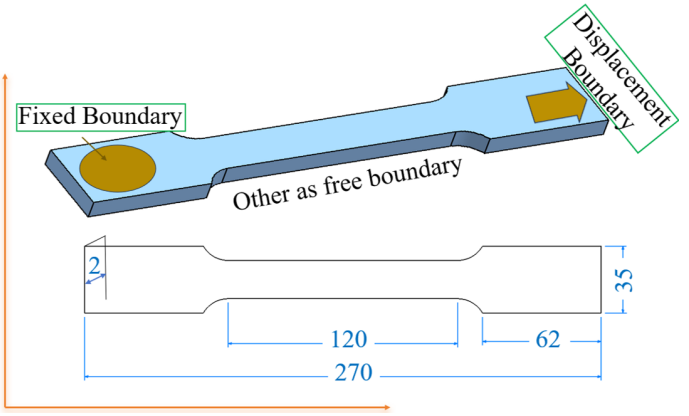


Figure 2. Tensile test specimen geometry according to China national standards.

Table 1. Material parameters for simulation

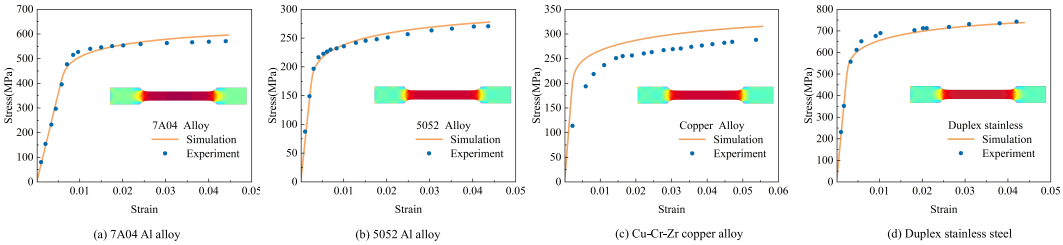
Name	σ_y (MPa)	σ_u (MPa)	E (GPa)	ρ (kg/m^3)	ν	K	n
5052 Al alloy	220	267	70	2680	0.33	350	0.3
7A04 Al alloy	540	570	71	2800	0.33	630	0.3
Duplex stainless steel	652	820	208	7800	0.33	820	0.3
Cu–Cr–Zr copper alloy	254	328	100	8900	0.33	388	0.3

aforementioned mathematical model, the tensile behavior of various materials was simulated, enabling the acquisition of true or equivalent stress–strain relationship curves.

Although the core focus of this study is on the deep learning-based training and prediction of material stress–strain curve data, experimental validation was conducted to ensure that the simulation data closely approximates completeness and realism. Specifically, the feasibility of using the finite element method to solve the aforementioned model was verified. First, stress–strain tensile experimental data for four materials—5052 and 7A04 Al alloys, duplex stainless steel, and copper alloy—were obtained from the literature [45,46]. Subsequently, finite element models for these four materials were established using the methodology described above. The simulation results were then compared with the experimental data to demonstrate the feasibility of obtaining stress–strain curves through material tensile simulation. During the simulation process, seven key material parameters were required, including yield strength σ_y, tensile strength σ_u, elastic modulus E, density ρ, Poisson's ratio ν, strength coefficient K, and hardening exponent n, as listed in Table 1.

Using the parameters listed in Table 1, tensile simulations were conducted for four materials: 5052 and 7A04 Al alloys, duplex stainless steel, and copper alloy. A uniform tensile length of 15 mm was applied to all specimens, a distance sufficient to prevent necking during the tensile process. The simulated stress–strain curves were compared with experimental data, as shown in Figure 3. From the figure, it is evident that the simulated stress–strain data for aluminum alloys and stainless steel exhibit excellent agreement with the experimental data. For the copper alloy, the simulated stress–strain data show slight deviations from the experimental results, but the overall trend of the curves remains consistent. These findings demonstrate the feasibility of using the aforementioned numerical method to obtain stress–strain data from tensile simulations.

When employing the Ludwik hardening model as the plastic constitutive equation, the key parameters influencing the simulation results are the strength coefficient K and the hardening



(a) 7A04 Al alloy (b) 5052 Al alloy (c) Cu-Cr-Zr copper alloy (d) Duplex stainless steel

Figure 3. Stress–strain curves: experimental data and simulation results.

exponent n. For different materials, particularly the strength coefficient K, significant variations can be observed. For the 106 materials selected in this study, the strength coefficient K was determined based on empirical formulas provided by relevant authors [47–49] and fitting analysis of experimental data. Specifically, a compensation value ranging from 0 to 100 MPa was added to the tensile strength of each material to derive its strength coefficient. Furthermore, since the core objective of this study is to obtain stress–strain data from various materials for training a deep learning model to capture the distribution patterns and trends of stress–strain curves, a certain margin of error in the stress–strain data is permissible. For metal alloys, the hardening exponent n exhibits minimal variation across materials, and its impact on simulation results is relatively small. Therefore, a uniform value of 0.3 was adopted as the hardening exponent for all simulations in this study.

Based on the aforementioned model and material parameters, tensile simulations were conducted for 106 alloy materials. The obtained stress–strain data were categorized into six major groups: aluminum alloys, copper alloys, nickel alloys, tool steels, titanium alloys, and other steels, as illustrated in Figure 4. From the figure, it is evident that the stress–strain behavior of all materials initially exhibits a linear relationship during the elastic stage, followed by nonlinear changes during the elastoplastic and plastic hardening stages. Furthermore, a comparison of the stress–strain curves for all materials reveals that the yield points or hardening regions of the 106 alloy materials fall within a range of 100 MPa to 1400 MPa. The majority of the materials exhibit yield points above 200 MPa, with only a few materials below this threshold. Additionally, materials with yield points around 400 MPa are relatively rare. For detailed classification, such as the number of stress–strain curves within the 100 MPa to 300 MPa range, as well as the specific stress–strain curve data for each material, please refer to Appendix B.

2.2. *A deep learning approach*

This study proposes a deep learning model based on a multi-head attention mechanism and a long short-term memory (LSTM) network, referred to as the MA-LSTM-DLnet model. This model belongs to the category of time-series prediction models in neural networks and consists of two main components: the attention mechanism and the LSTM network. The attention mechanism employs a 4-head attention mechanism, as multi-head attention allows the model to learn information in parallel from different representation subspaces. For a detailed explanation of the attention mechanism, refer to the work of Vaswani et al. [50]. For the input sequence data X used for learning or prediction, the attention mechanism first applies three distinct linear transformations to obtain the query (Q), key (K), and value (V) matrices, as described by the following equations:

$$Q = XW^Q \tag{7}$$

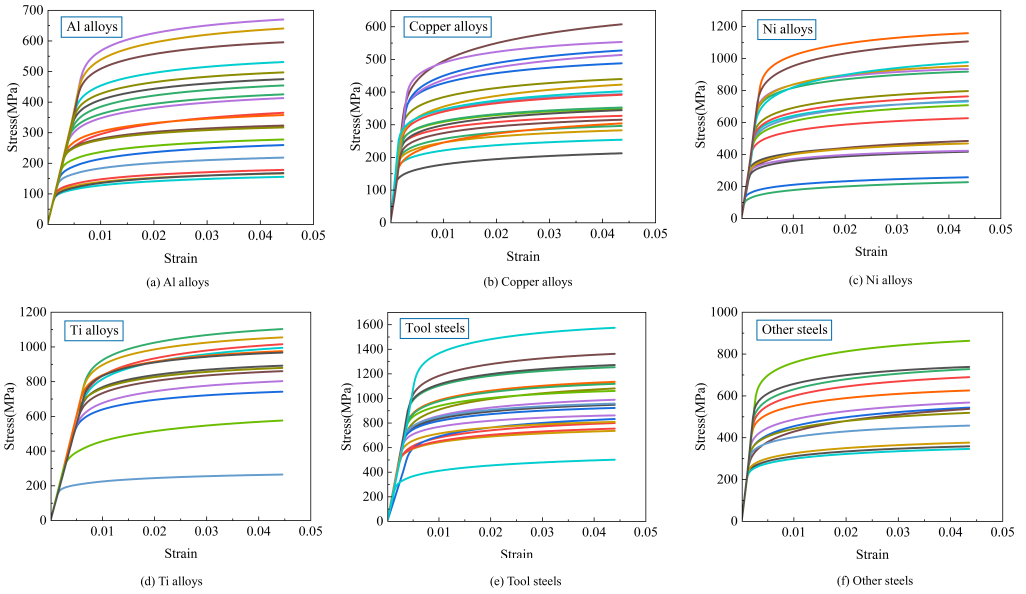


Figure 4. Stress–strain curves of 106 materials.

$$K = XW^K \tag{8}$$

$$V = XW^V \tag{9}$$

Here, W^Q, W^K, W^V are weight matrices learned by the model, and Q, K, and V represent the query, key, and value matrices, respectively. These equations map the input sequence X to the representation spaces of Q, K, and V, enabling the model to capture features and dependencies from different perspectives. In the case of multi-head attention, the transformed vectors are divided into multiple heads, each assigned independent Q, K, and V matrices. For each head, the output is computed using the scaled dot-product attention mechanism, as follows:

$$\text{Attention}\,(Q,K,V) = \text{Soft max}\left(\frac{Q\cdot K^{\mathrm{T}}}{\sqrt{d_{\mathrm{k}}}}\right)\cdot V \tag{10}$$

Here, d_{k} is the dimensionality of the key matrix, and $\sqrt{d_{\mathrm{k}}}$ serves as a scaling factor to prevent excessively large dot-product results, which could lead to gradient vanishing. The term K^{T} denotes the transpose of the key matrix. Equation (10) computes the attention weights for each head, which are then used to calculate a weighted sum of the value matrix V. The outputs of all heads are concatenated and passed through a linear layer to produce the final attention output vector, as shown below:

$$\text{MutiHead}\,(X) = \text{Concat}\,(H_1, H_2, \ldots, H_h)W^o \tag{11}$$

$$H_i = \text{Attention}\,(Q_i, K_i, V_i) \tag{12}$$

Here, W^o is the output projection matrix.

The LSTM network is a specialized type of recurrent neural network (RNN). Its core lies in the gating mechanism, which includes three gates: the forget gate, input gate, and output gate [51]. The forget gate determines which information is retained or discarded from the memory cell, using a sigmoid activation function to produce values between 0 and 1, representing the proportion of information to retain. Its expression is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{13}$$

Here, f_t is the output of the forget gate, σ is the sigmoid activation function, W_f and b_f are the weight matrix and bias, h_{t-1} is the hidden state from the previous time step, and x_t is the current input. The input gate determines which new information is stored in the memory cell. It consists of two parts: a sigmoid function to decide the update portion and a tanh function to generate candidate values. Its expressions are as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{14}$$

$$\widetilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_c) \tag{15}$$

Here, i_t is the output of the input gate, and $\widetilde{C_t}$ is the new candidate value. The memory cell is a critical component of the LSTM, responsible for storing long-term dependency information. Its expression is as follows:

$$\mathrm{c}_t = f_t \cdot c_{t-1} + i_t \cdot \widetilde{C_t} \tag{16}$$

Here, c_t is the current memory cell state, and c_{t-1} is the memory cell state from the previous time step. Finally, the output gate determines the output of the hidden state, using a sigmoid function to decide which information is output. Its expressions are as follows:

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{17}$$

$$h_t = O_t \cdot \tanh(c_t) \tag{18}$$

Here, O_t is the output of the output gate, and h_t is the current hidden state.

Based on the aforementioned attention mechanism and LSTM network model, this study proposes the MA-LSTM-DLnet model for predicting material constitutive relationships. The model consists of three main components, as illustrated in Figure 5. The first component involves feeding preprocessed sequence data into the input layer, accompanied by a positional encoding layer with a maximum dimension of 512. Through Equations (7)–(9), linear transformations are applied to initially obtain the Q, K, and V matrices for the sequence data. Since the model employs four attention heads, the initial Q, K, and V matrices are divided into four parts, each stored in a linear layer and corresponding to one attention head. Using Equation (10), the attention weights and outputs are computed for each attention head based on its respective Q, K, and V matrices. The outputs of all attention heads are then concatenated to produce the final attention output. The second component takes the features extracted and filtered by the attention mechanism as input to the LSTM network. Here, the features undergo further training and analysis through the forget gate, input gate, and output gate. In this study, the selection of key model parameters was based on a combination of empirical optimization and preliminary research. It was found that with 800 hidden neurons, the model demonstrated satisfactory training accuracy and generalization performance on the validation set. Increasing the number beyond 800 did not significantly improve performance but led to longer training times and higher computational resource consumption. Thus, considering the balance between model performance and computational efficiency, we ultimately selected 800 hidden neurons for the LSTM layer. The third component employs a dropout layer to randomly zero out 50% of the neuron outputs. The remaining elements are then fed into a fully connected layer with a single unit, which performs the final regression prediction output.

2.3. *Feature set configuration*

The essence of the prediction model is illustrated in Figure 6, where the input consists of relevant feature vectors or matrices X, and the output is the value Y. For the material stress–strain relationship curve, stress σ represents the output value Y, while strain ε is one of the features in the input feature matrix X. However, since strain data alone is too simplistic, the model cannot distinguish between different materials during training. Therefore, strain data can only

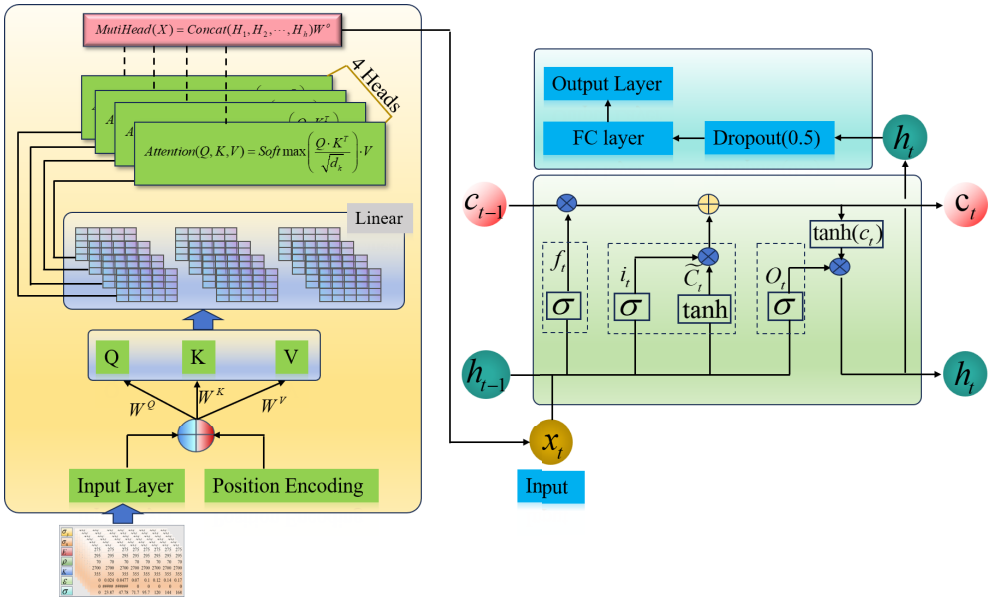


Figure 5. MA-LSTM-DLnet model.

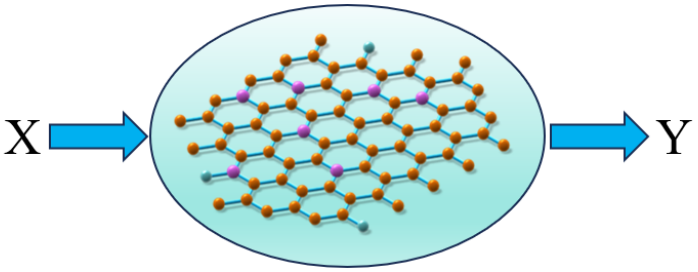


Figure 6. Schematic diagram of the prediction process.

serve as one of the features in the input feature matrix X. To enhance the model's ability to differentiate between materials, additional features are incorporated into the input feature matrix X, including material yield strength σ_y, tensile strength σ_u, elastic modulus E, density ρ, strength coefficient K, and the product of elastic modulus and strain. By including these supplementary features, the model can effectively distinguish between different materials during the training process based on the input feature matrix X.

To better distinguish and study the influence of different features in the input feature matrix X, as well as to evaluate the accuracy of the prediction model under various input feature configurations, the input features were subdivided into different scenarios, as detailed in Table 2.

Five distinct input feature matrices X were established for training based on different environmental configurations, with the corresponding output matrix Y representing stress. In the table, the letter "y" indicates that the corresponding material feature is included in the training, while the letter "n" indicates that the feature is excluded. For example, in *Setting 1*, all seven material feature parameters are included in the training, and the input feature matrix X and output Y can

Table 2. Feature parameter configurations for network model training

Name	σ_y	σ_u	E	ρ	K	$E \times \varepsilon$	ε
Setting 1	y	y	y	y	y	y	y
Setting 2	y	y	y	y	n	y	y
Setting 3	y	y	y	y	n	n	y
Setting 4	y	y	y	n	n	n	y
Setting 5	y	y	n	n	n	n	y

be expressed as follows:

$$\begin{cases} X = [\sigma_y, \sigma_u, E, \rho, K, E \times \varepsilon, \varepsilon]^{\mathrm{T}} \\ Y = [\sigma] \end{cases} \tag{19}$$

For *Setting 2*, as shown in the table, the strength coefficient K is excluded from the input feature matrix X and does not participate in the training. In this case, the input feature matrix X and output Y are expressed as:

$$\begin{cases} X = [\sigma_y, \sigma_u, E, \rho, E \times \varepsilon, \varepsilon]^{\mathrm{T}} \\ Y = [\sigma] \end{cases} \tag{20}$$

For *Setting 3*, *Setting 4*, and *Setting 5*, specific feature parameters are sequentially excluded from the input feature matrix X, as outlined in Table 2.

2.4. *Division of training and testing datasets*

The core dataset of this study consists of stress–strain data for 106 alloy materials obtained through FEM, as described in Section 2.1. As outlined in Section 2.3, these data were combined with material properties such as yield strength σ_y, tensile strength σ_u, elastic modulus E, density ρ, strength coefficient K, and the product of elastic modulus and strain to form a comprehensive dataset of 106 entries, which is suitable for training.

To evaluate the training performance of the MA-LSTM-DLnet model, the 100 data entries were randomly divided into training set, validation set and testing sets. As shown in Figure 7, the dataset was split into 80 entries for the training set, 10 entries for the validation set, and 10 entries for the testing set. Specifically, 80 alloy materials were randomly selected for training to develop the predictive model, while the remaining 10 alloy materials were reserved as the testing set and excluded from training. These testing data were used to compare with the model's predictions. To better demonstrate the predictive accuracy, effectiveness, and generalization capability of the trained model, the above process was repeated three times. This means that the 100 material data entries were randomly divided three times, resulting in a total of 30 alloy materials in the testing set.

2.5. *Hyperparameter configuration*

Based on the data division strategy described in Section 2.4, the deep learning neural network model established in Section 2.2 was employed to train the dataset. To optimize the model's training performance and convergence speed, the following hyperparameters were tuned based on a combination of empirical optimization methods and results from preliminary research:

(1) Optimization Algorithm: Adaptive Moment Estimation (Adam)
(2) Maximum Iterations: 100
(3) Mini-batch Size: 20

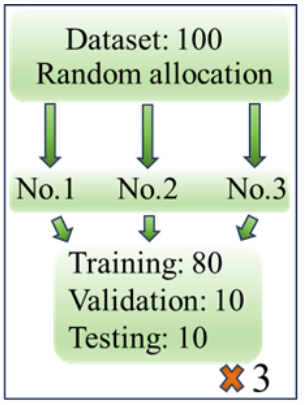


Figure 7. Dataset division: training, validation and testing sets.

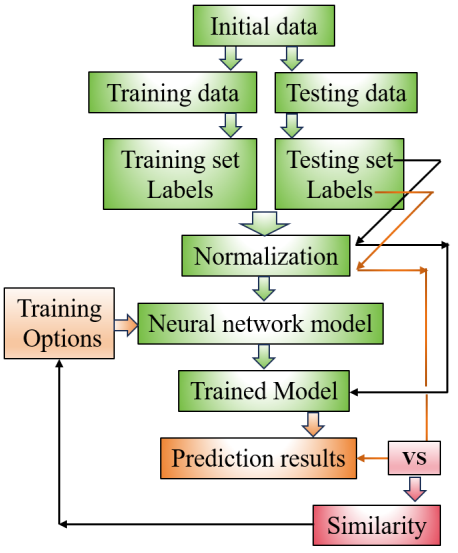


Figure 8. A program framework for MA-LSTM-DLnet training.

(4) Initial Learning Rate: 0.001
(5) Learning Rate Adjustment Strategy: Piecewise constant
(6) Learning Rate Decay Interval: 80
(7) Learning Rate Decay Factor: 0.1
(8) Gradient Threshold: 1

Using the MA-LSTM-DLnet model with the above hyperparameter configuration, the dataset was trained. The overall program execution framework is illustrated in Figure 8.

3. Results and discussion

3.1. *Comprehensive feature analysis: training with complete dataset*

As shown in Equation (19), seven key feature parameters representing material characteristics were selected to form the input feature matrix X. These parameters include yield strength σ_y,

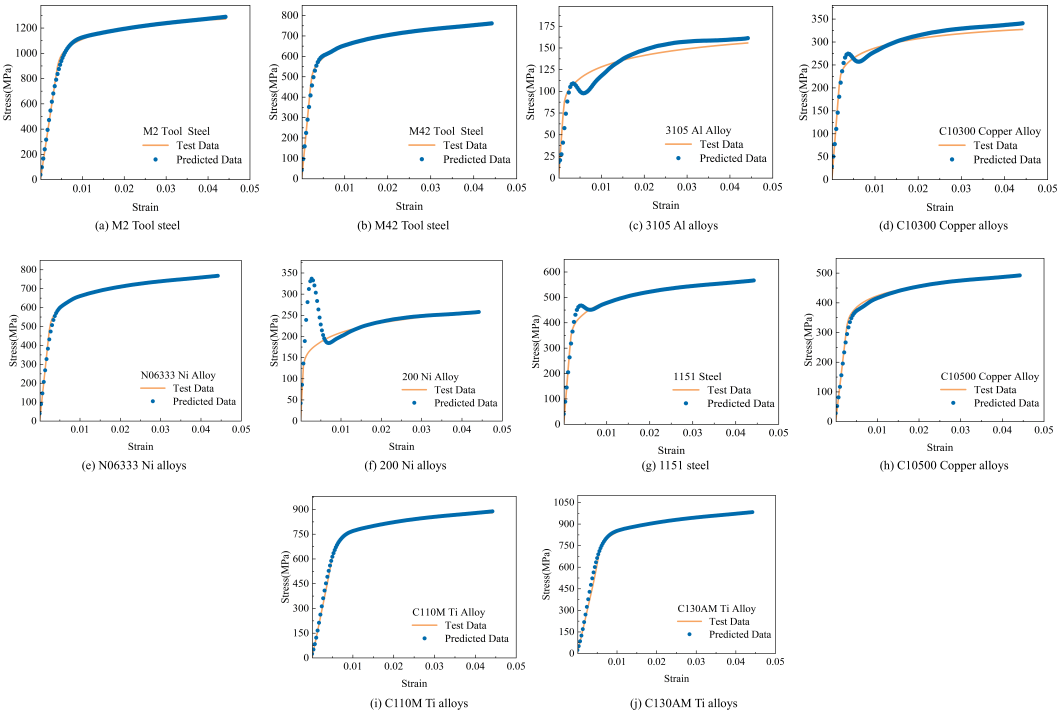


Figure 9. MA-LSTM-DLnet model for predicting stress–strain curves.

tensile strength σ_u, elastic modulus E, density ρ, strength coefficient K, and the product of elastic modulus and strain. Together, they constitute a 7-channel dataset used as input, with stress σ serving as the corresponding labeled output. This mapping relationship dataset was used to train the model. According to the feature configuration scheme presented in Table 2, this section investigates the training performance of the model under *Setting 1*, where all material feature parameters are included in the training.

For the dataset of 100 alloy materials, the training, Validation and testing sets were divided as described in Section 2.4, resulting in three groups with a total of 30 materials reserved for testing and validation. The first group of test materials included 10 randomly selected alloys: M2 tool steel, M42 tool steel, 3105 aluminum alloy, C10300 copper alloy, C10500 copper alloy, N06333 nickel alloy, 200 nickel alloy, 1151 steel, C110M titanium alloy, C130AM titanium alloy. These materials were excluded from training and used to compare with the predictions of the trained model. The remaining 90 alloy materials were processed into the training dataset for the MA-LSTM-DLnet model to develop the stress–strain prediction model. After training the MA-LSTM-DLnet model, the corresponding feature parameters of the first group of test materials were input, and the prediction results are shown in Figure 9. From the figure, it is evident that the MA-LSTM-DLnet model performs well in predicting the stress–strain relationship curves. Notably, the predicted stress–strain curves for eight materials align closely with the test data in both the linear and nonlinear stages, which is a surprising and intriguing result. A further detailed observation of the comparative test data for the 10 materials in the figure reveals that the predicted stress–strain curve of the 200 nickel alloy in Figure 9(f) exhibits an abrupt change near the yield point. Additionally, while the overall trend of the prediction results for the 3105 aluminum alloy in Figure 9(c) is consistent with the test data, the prediction accuracy is slightly lower.

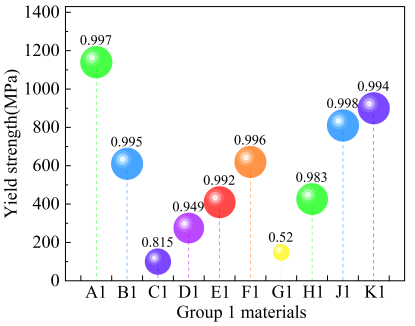


Figure 10. Similarity of predicted and test stress–strain curves.

The deviation in the prediction for 200 nickel alloy and 3105 aluminum alloy, contrasted with the accurate predictions for the other eight alloys, cannot be fully explained by the comparison of these ten materials alone. However, some distinctions can be observed. Comparing Figure 9(a–j), the most notable difference lies in the varying yield strengths of the materials. However, yield strength is not the fundamental reason for prediction accuracy. The primary factor is the size of the training dataset. As shown in Figure 4 and detailed in Appendix B, the number of stress–strain curves for materials with yield strengths of around 100 MPa in the training dataset is relatively small. This scarcity leads to lower prediction accuracy for stress–strain curves within this yield strength range. This observation also indicates that the material's physical properties and types do not significantly affect the model's prediction accuracy.

Since the prediction results are material stress–strain constitutive relationship curves, it is not feasible to compare the accuracy or calculate errors by selecting individual or multiple points on the curves. Instead, statistical calculations are required for the entire curves. To provide a clearer and more scientific comparison between the model's predictions and the test data, this study employs the coefficient of determination method to evaluate the similarity between the two curves. The formula for the coefficient of determination is as follows:

$$R = 1 - \| Y_{\mathrm{P}} - Y_{\mathrm{T}} \|^2 / \| Y_{\mathrm{T}} - \overline{Y_{\mathrm{T}}} \|^2 \tag{21}$$

Here, Y_{T} represents the test values, Y_{P} denotes the predicted values, $\overline{Y_{\mathrm{T}}}$ is the mean of the test data, and $\| \cdot \|^2$ represents the squared norm. Using the above equation, the similarity between the stress–strain curves predicted by the MA-LSTM-DLnet model and the given test curves for the first group of materials is shown in Figure 10. In the figure, the horizontal axis labels A1, B1, C1, D1, E1, F1, G1, H1, J1 and K1 correspond to the ten materials in the first group: M2 tool steel, M42 tool steel, 3105 aluminum alloy, C10300 copper alloy, C10500 copper alloy, N06333 nickel alloy, 200 nickel alloy, 1151 steel, C110M titanium alloy, C130AM titanium alloy. The vertical axis represents the yield strength of the corresponding materials. From Figure 10, it can be observed that the similarity between the stress–strain curves predicted by the MA-LSTM-DLnet model and the test data exceeds 97% for all materials except the 200 nickel alloy.

To further demonstrate the predictive performance of the MA-LSTM-DLnet model, a second group of test materials was randomly selected for prediction and comparison with the original data. Similar to the first group, the second group of test materials consists of ten alloys: S2 Tool steel, T1 Tool steel, 7050 aluminum alloy, C35300 copper alloy, C36200 copper alloy, 718SPF nickel alloy, MA754 nickel alloy, 1110 steel, R54520 titanium alloy and 21S titanium alloy. The feature data representing the second group of test materials were input into the MA-LSTM-DLnet model, and the resulting stress–strain curves are shown in Figure 11.

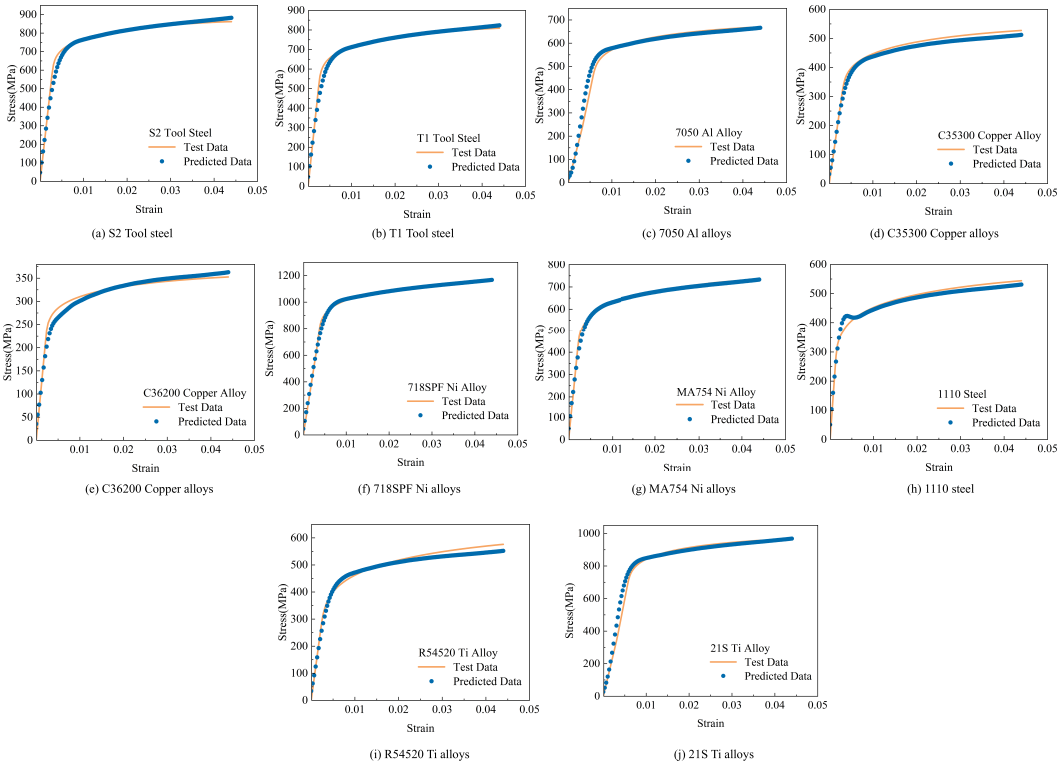


Figure 11. MA-LSTM-DLnet model for predicting stress–strain curves in group 2 materials.

From the figure, it is evident that the ten stress–strain curves predicted by the MA-LSTM-DLnet model align very well with the original test data curves. Compared to the first group of materials, the predictions for the second group are nearly perfect. Upon closer examination, slight discrepancies are observed in the linear regions of the predicted stress–strain curves for 7050 aluminum alloy and 21S titanium alloy. Additionally, the predicted stress–strain curve for C35300 copper alloy and R54520 Ti alloy exhibits minor deviations in the hardening region.

Similarly, using Equation (21), the similarity between the stress–strain curves predicted by the model for the second group of materials and the original test data was calculated and analyzed. As shown in Figure 12, the horizontal axis labels A2, B2, C2, D2, E2, F2, G2, H2, J2 and K2 correspond to the ten materials in the second group: S2 Tool steel, T1 Tool steel, 7050 aluminum alloy, C35300 copper alloy, C36200 copper alloy, 718SPF nickel alloy, MA754 nickel alloy, 1110 steel, R54520 titanium alloy and 21S titanium alloy. The vertical axis represents the yield strength of the corresponding materials. From Figure 12, it can be observed that, compared to the first group of test materials, the similarity between the predicted values and the test values for the second group of materials exceeds 96% in all cases.

The third group of test materials, similar to the previous two groups, consists of ten randomly selected alloys: 5754 aluminum alloy, C43000 copper alloy, C61400 copper alloy, 718 nickel alloy, AF955 nickel alloy, 5115 steel, 15-Mo-0.5Si titanium alloy, 5621 titanium alloy, FX tool steel and H10 tool steel. The MA-LSTM-DLnet model was used to predict the stress–strain curves for the third group of test materials, and the results are shown in Figure 13. From the figure, it can be observed that the predicted stress–strain curves align well with the original test data for all materials except the 5754 aluminum alloy. This further demonstrates that the

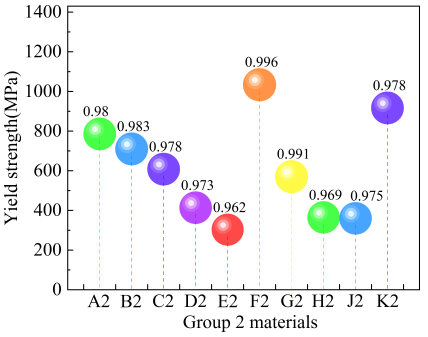


Figure 12. Similarity of predicted and test stress–strain curves in group 2 materials.

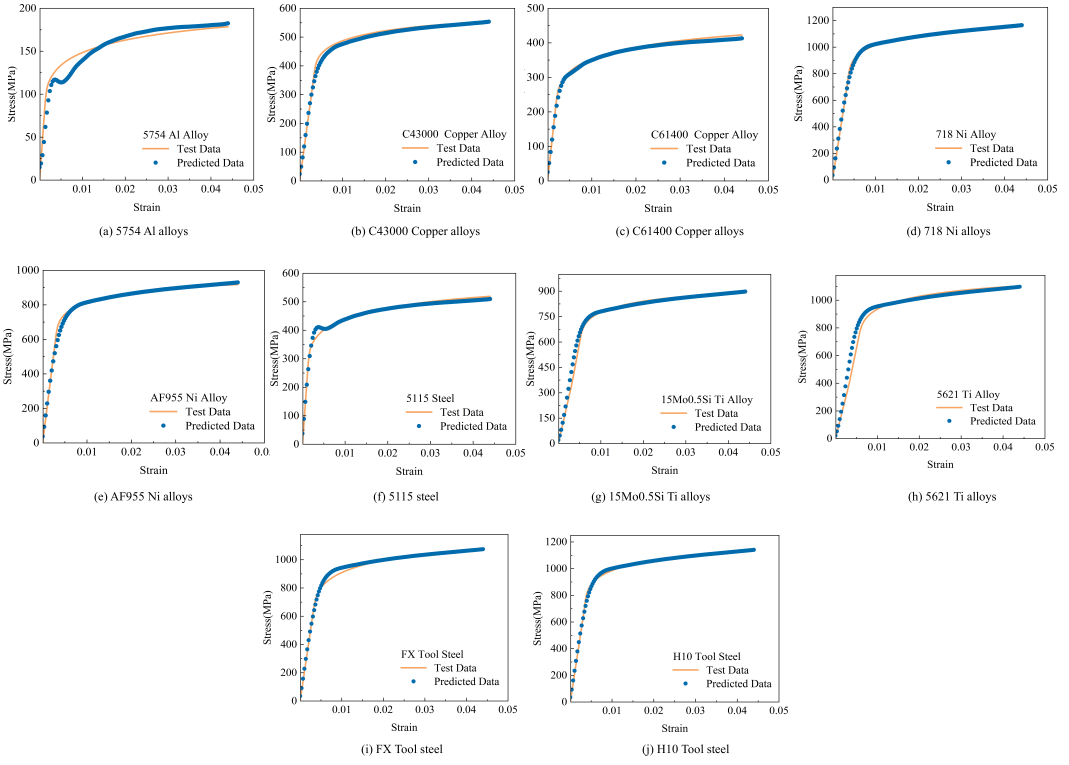


Figure 13. MA-LSTM-DLnet model for predicting stress–strain curves in group 3 materials.

prediction accuracy is highly dependent on the quantity of training data. The yield strength of 5754 aluminum alloy is around 100 MPa, and the scarcity of training data within this range results in lower prediction accuracy for this material.

Using Equation (21), the similarity between the stress–strain curves predicted for the third group of materials and the original test data was calculated and analyzed. As shown in Figure 14, the horizontal axis labels A3, B3, C3, D3, E3, F3, G3, H3, J3 and K3 correspond to the ten materials in the third group: 5754 aluminum alloy, C43000 copper alloy, C61400 copper alloy, 718 nickel alloy, AF955 nickel alloy, 5115 steel, 15-Mo-0.5Si titanium alloy, 5621 titanium alloy, FX tool steel and H10 tool steel. The vertical axis represents the yield strength of the corresponding materials.

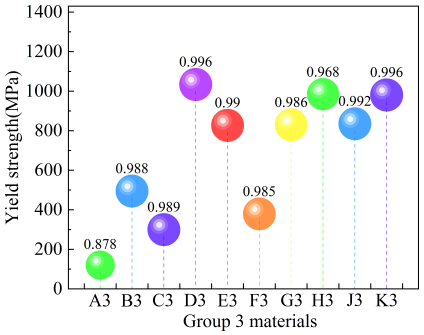


Figure 14. Similarity of predicted and test stress–strain curves in group 3 materials.

From the figure, it can be observed that the similarity for 5754 aluminum alloy is 87.8%, while the similarity for the other five materials exceeds 96%. By comprehensively comparing Figure 10. Figures 12 and 14, it is evident that the similarity is relatively low for two aluminum alloys and 200 nickel alloy with lower yield strengths, resulting in suboptimal prediction performance. As previously discussed, this can be explained by referring to the original stress–strain data in Figure 4 or the training data in Appendix B. The scarcity of original training data for stress–strain curves with yield strengths of around 100 MPa leads to inadequate model training and, consequently, poorer prediction performance. Overall, the prediction results for the three groups of test materials demonstrate that, given sufficient training data, the MA-LSTM-DLnet model exhibits stable performance and strong generalization capabilities in predicting material stress–strain relationships. Under certain conditions, the MA-LSTM-DLnet model can serve as a viable alternative to existing theoretical calculations and tensile testing methods for predicting the constitutive relationships of alloy materials.

3.2. *Exclusion of strength coefficient: impact on training performance*

In Section 3.1, the training dataset under *Setting 1*, which includes all features, was investigated. The analysis results demonstrate that the deep learning approach and the proposed MA-LSTM-DLnet model can effectively predict material constitutive relationships. However, in practical scenarios, such as when using the Ludwik isotropic hardening model to calculate stress–strain relationships, key parameters like the strength coefficient K are often difficult to obtain. Moreover, the representation of K may vary under different working conditions. If alternative models are used, similar challenges arise with other key parameters, which are equally difficult to determine. Therefore, this section aims to explore the predictive performance of the MA-LSTM-DLnet model by removing these hard-to-obtain key parameters and utilizing only conventional material properties to construct the feature training dataset. Following *Setting 2* in Table 2, the strength coefficient K was excluded, and the training dataset was composed of six features: yield strength σ_y, tensile strength σ_u, elastic modulus E, density ρ, the product of elastic modulus and strain, and strain ε.

To facilitate comparison with the results in Section 3.1, the same three groups of 30 materials were used as test materials. The prediction results of the MA-LSTM-DLnet model for the first group of test materials are shown in Figure 15. Here, *Predicted Data1* represents the predictions using the *Setting 1* dataset, where all seven material features were included in the training, as discussed in Section 3.1. *Predicted Data2* represents the predictions using the *Setting 2* configuration, where the strength coefficient K was excluded from the training features.

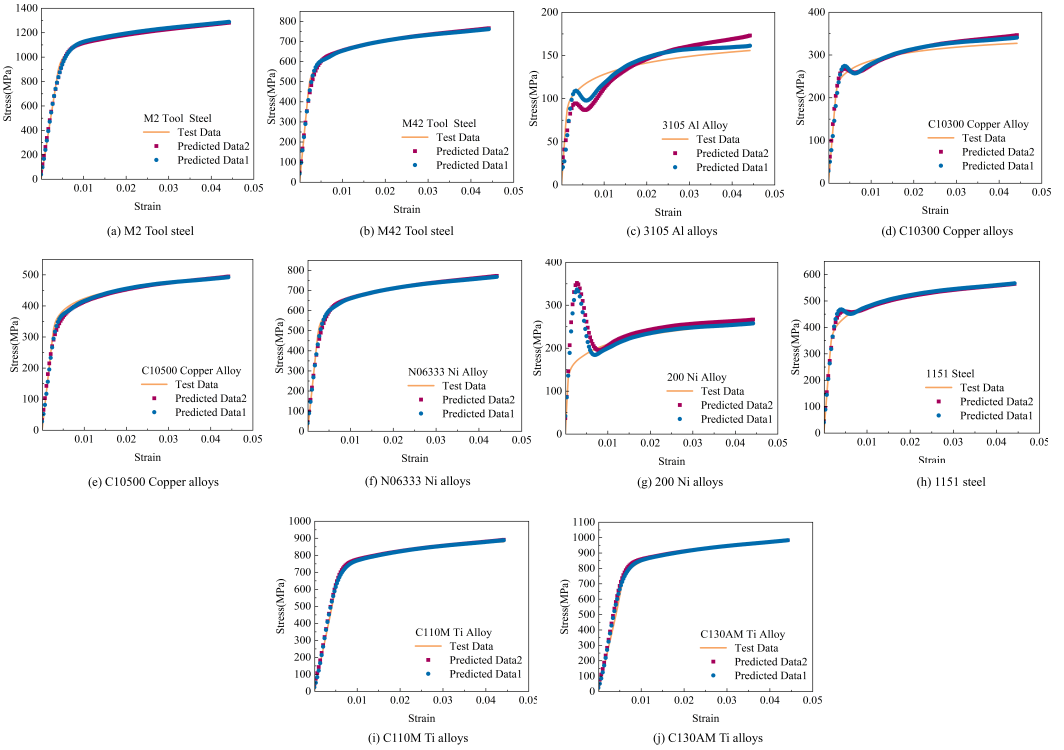


Figure 15. MA-LSTM-DLnet model for predicting stress–strain curves: *Predicted Data1* which incorporates a full feature set and *Predicted Data2* excluding intensity coefficients.

From Figure 15, it is evident that the predictions under both *Setting 1* and *Setting 2* are nearly identical for nine of the ten materials, aligning well with the test data and showing excellent consistency. The only significant deviation occurs for the 200 Ni alloys as in Figure 15(g), as explained in detail in Section 3.1, which is unrelated to the model's predictive performance. These results clearly demonstrate that the proposed MA-LSTM-DLnet model and methodology can replace traditional theoretical and numerical calculation methods. Furthermore, they eliminate the need for hard-to-obtain key parameters, simplifying the process of determining material stress–strain constitutive relationships. Additionally, this study highlights that deep learning methods can leverage simple, conventional material properties to predict stress–strain constitutive relationships, thereby reducing the repetitive work associated with traditional tensile testing methods and significantly lowering labor costs.

Using Equation (21), the similarity between the predicted values (*Predicted Data2*) and the original test data was calculated. The results are shown in Figure 16, where the horizontal axis labels A1, B1, C1, D1, E1, F1, G1, H1, J1 and K1 correspond to the ten materials in the first group: M2 tool steel, M42 tool steel, 3105 aluminum alloy, C10300 copper alloy, C10500 copper alloy, N06333 nickel alloy, 200 nickel alloy, 1151 steel, C110M titanium alloy, C130AM titanium alloy. The vertical axis represents the yield strength of the corresponding materials. From the figure, it can be observed that, except for the 3105 aluminum alloy and the 200 nickel alloy, the similarity between the predicted values and the original test data exceeds 93% for all other materials.

The MA-LSTM-DLnet model was used to predict the stress–strain relationships for the second group of test materials, which includes S2 Tool steel, T1 Tool steel, 7050 aluminum alloy, C35300 copper alloy, C36200 copper alloy, 718SPF nickel alloy, MA754 nickel alloy, 1110 steel, R54520

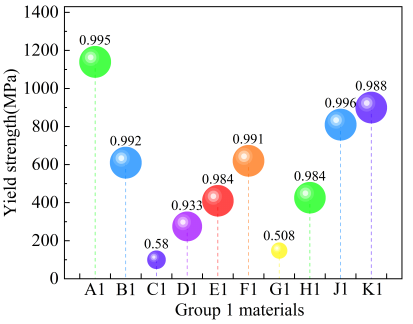


Figure 16. Similarity of predicted data2 and test stress–strain curves.

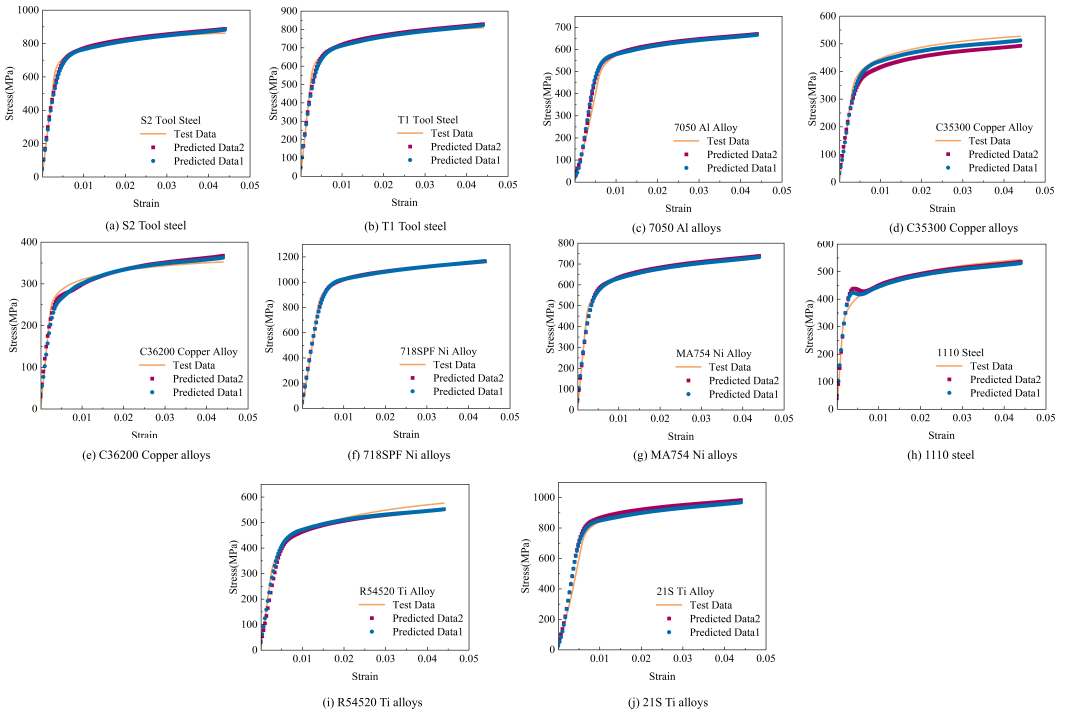


Figure 17. MA-LSTM-DLnet model for predicting stress–strain curves in group 2 materials: *Predicted Data1* which incorporates a full feature set and *Predicted Data2* excluding intensity coefficients.

titanium alloy and 21S titanium alloy. The prediction results are shown in Figure 17. When the strength coefficient K is excluded from the material features, the model's predictions for this group are even more accurate than those for the first group of test materials. As shown in Figure 17, the stress–strain curves of *Predicted Data1* (with K included) and *Predicted Data2* (with K excluded) are nearly identical for five of the six materials, with only minor deviations observed for C35300 copper alloy. This further demonstrates that, when using deep learning methods to predict material constitutive relationships, it is unnecessary to consider key parameters that are difficult to obtain in traditional methods.

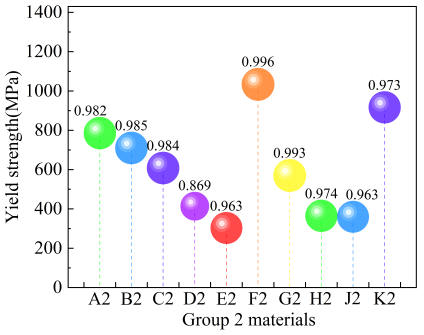


Figure 18. Similarity of predicted data2 and test stress–strain curves in group 2 materials.

Using Equation (21), the similarity between *Predicted Data2* and the original test data for the second group of materials was calculated. The results are shown in Figure 18, where the horizontal axis labels A2, B2, C2, D2, E2, F2, G2, H2, J2 and K2 correspond to the ten materials in the second group: S2 Tool steel, T1 Tool steel, 7050 aluminum alloy, C35300 copper alloy, C36200 copper alloy, 718SPF nickel alloy, MA754 nickel alloy, 1110 steel, R54520 titanium alloy and 21S titanium alloy. The vertical axis represents the yield strength of the corresponding materials. From the figure, it can be observed that the similarity to C35300 copper alloy is 86.9%, while the similarity to the other materials exceeds 96%. This aligns with the trends demonstrated in Figure 17.

The third group of test materials includes 5754 aluminum alloy, C43000 copper alloy, C61400 copper alloy, 718 nickel alloy, AF955 nickel alloy, 5115 steel, 15-Mo-0.5Si titanium alloy, 5621 titanium alloy, FX tool steel and H10 tool steel. The MA-LSTM-DLnet model was used to predict their stress–strain relationships, and the results are shown in Figure 19. From the figure, it is evident that *Predicted Data1*, *Predicted Data2*, and the original test data align well for all materials except 5754 aluminum alloy, which exhibits some deviation. By comprehensively comparing the prediction results for the three groups of test materials in Figures 15, 17 and 19, it is further demonstrated that the predictive performance of the deep learning model is not highly sensitive to the strength coefficient K. Removing the strength coefficient as a feature has a negligible impact on the predicted stress–strain curves.

Using Equation (21), the similarity between *Predicted Data2* and the original test data for the third group of materials was calculated. The results are shown in Figure 20, where the horizontal axis labels A3, B3, C3, D3, E3, F3, G3, H3, J3 and K3 correspond to the ten materials in the third group: 5754 aluminum alloy, C43000 copper alloy, C61400 copper alloy, 718 nickel alloy, AF955 nickel alloy, 5115 steel, 15-Mo-0.5Si titanium alloy, 5621 titanium alloy, FX tool steel and H10 tool steel. The vertical axis represents the yield strength of the corresponding materials. From the figure, it is evident that the similarity for 5754 aluminum alloy is 64.4%, while the similarity to the other materials exceeds 95%. By comparing the similarity calculations in Figures 16, 18 and 20, it is clear that the lower similarity for aluminum alloys is not related to the deep learning model itself. As explained in Section 3.1, the primary reason is the scarcity of training data for materials with yield strengths of around 100 MPa.

3.3. *Sequential feature omission: evaluating training outcomes*

As shown in Table 2, Sections 3.1 and 3.2 correspond totraining performance under *Setting 1* and *Setting 2*, respectively. The analysis above demonstrates that the MA-LSTM-DLnet model is fully

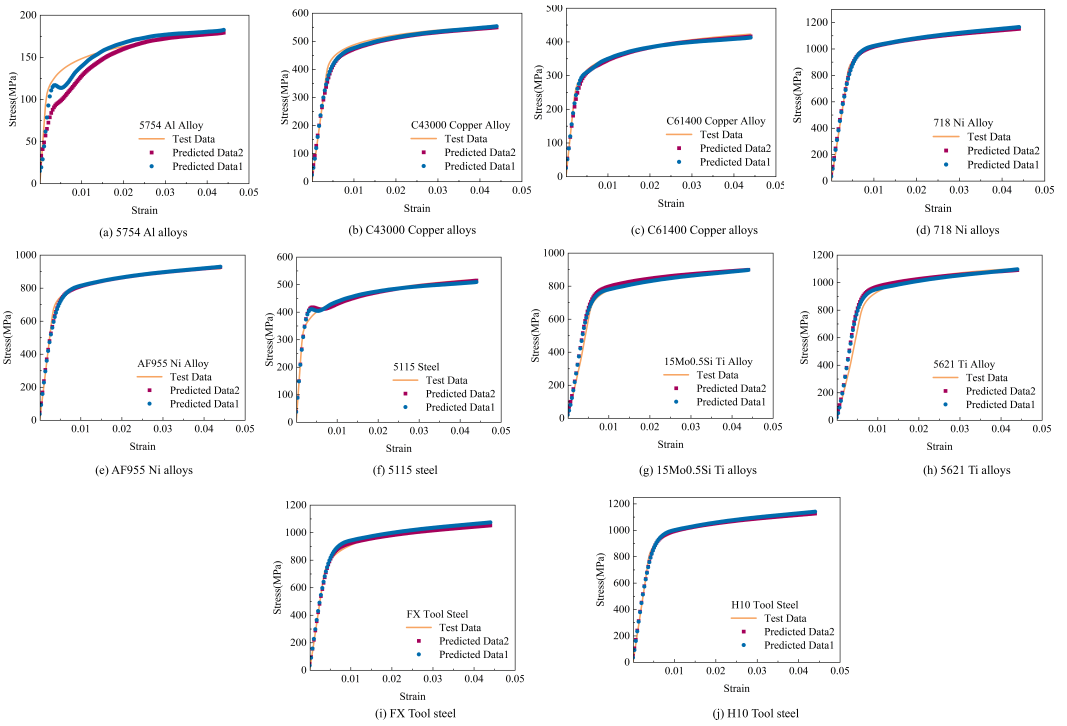


Figure 19. MA-LSTM-DLnet model for predicting stress–strain curves in group 3 materials: *Predicted Data1* which incorporates a full feature set and *Predicted Data2* excluding intensity coefficients.

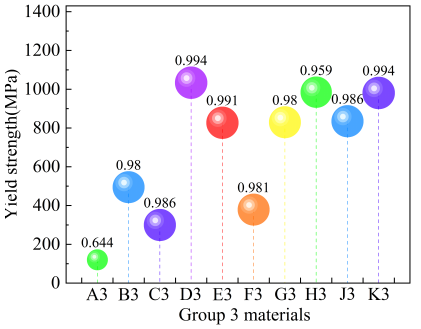


Figure 20. Similarity of predicted data2 and test stress–strain curves in group 3 materials.

capable of being adopted for studying or predicting material constitutive relationships. To further investigate the influence of material features, this section analyzes the predictive performance of the MA-LSTM-DLnet model under *Setting 3*, *Setting 4*, and *Setting 5* as outlined in Table 2. Given the extensive scope and the training performance observed under *Setting 1* and *Setting 2*, the second group of test materials—S2 Tool steel, T1 Tool steel, 7050 aluminum alloy, C35300 copper alloy, C36200 copper alloy, 718SPF nickel alloy, MA754 nickel alloy, 1110 steel, R54520 titanium alloy and 21S titanium alloy—was selected for prediction. For *Setting 3*, the feature configuration

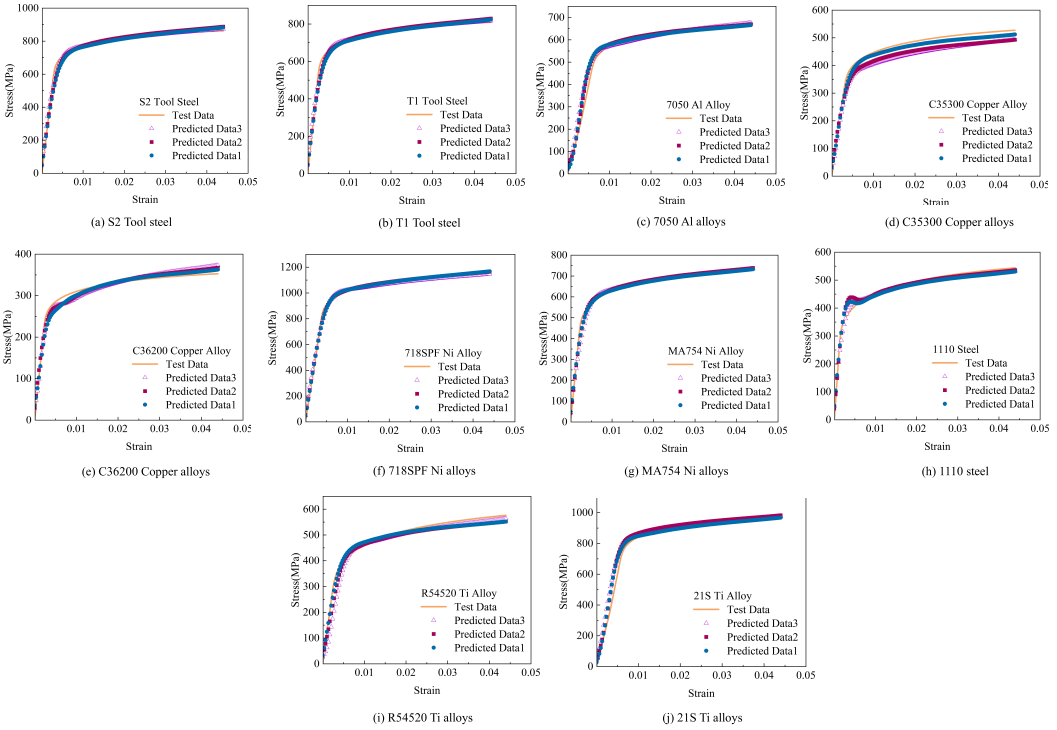


Figure 21. MA-LSTM-DLnet model for predicting stress–strain curves in group 2 materials: *Predicted Data1* which incorporates a full feature set, *Predicted Data2* excluding intensity coefficients, *Predicted Data3* excluding intensity coefficients and $E \times \varepsilon$.

includes yield strength σ_y, tensile strength σ_u, elastic modulus E, density ρ, and strain ε, forming a 5-channel feature dataset as the input matrix X. The strength coefficient K and the product of elastic modulus and strain were excluded. As shown in Figure 21, *Predicted Data1* corresponds to *Setting 1*, *Predicted Data2* corresponds to *Setting 2*, and *Predicted Data3* corresponds to *Setting 3*. From the figure, it is evident that the predictions under *Setting 3* exhibit slightly lower agreement with the test results compared to *Setting 1* and *Setting 2*.

For *Setting 4*, the feature configuration includes yield strength σ_y, tensile strength σ_u, elastic modulus E, and strain ε, forming a 4-channel feature dataset as the input matrix X. The density ρ, strength coefficient K, and the product of elastic modulus and strain were excluded. As shown in Figure 22, *Predicted Data4* corresponds to *Setting 4*. The overall prediction results are slightly lower than the actual test results, but no significant fluctuations are observed.

For *Setting 5*, the feature configuration includes yield strength σ_y, tensile strength σ_u, and strain ε, forming a 3-channel dataset. The elastic modulus E, density ρ, strength coefficient K, and the product of elastic modulus and strain were excluded. As shown in Figure 23, *Predicted Data5* corresponds to *Setting 5*. The overall prediction performance is similar to that of *Setting 4*.

In summary, the results indicate that reducing the number of feature parameters representing the material has a minor impact on the model's prediction performance. However, the most significant factor influencing the results is the size of the training dataset. Based on a comprehensive analysis of the model's predictions under all conditions, it is evident that while reducing the number of features or the absence of key parameters may affect prediction accuracy, the impact is not substantial if the error range remains stable. This can be mitigated by increasing the size of the original training dataset.

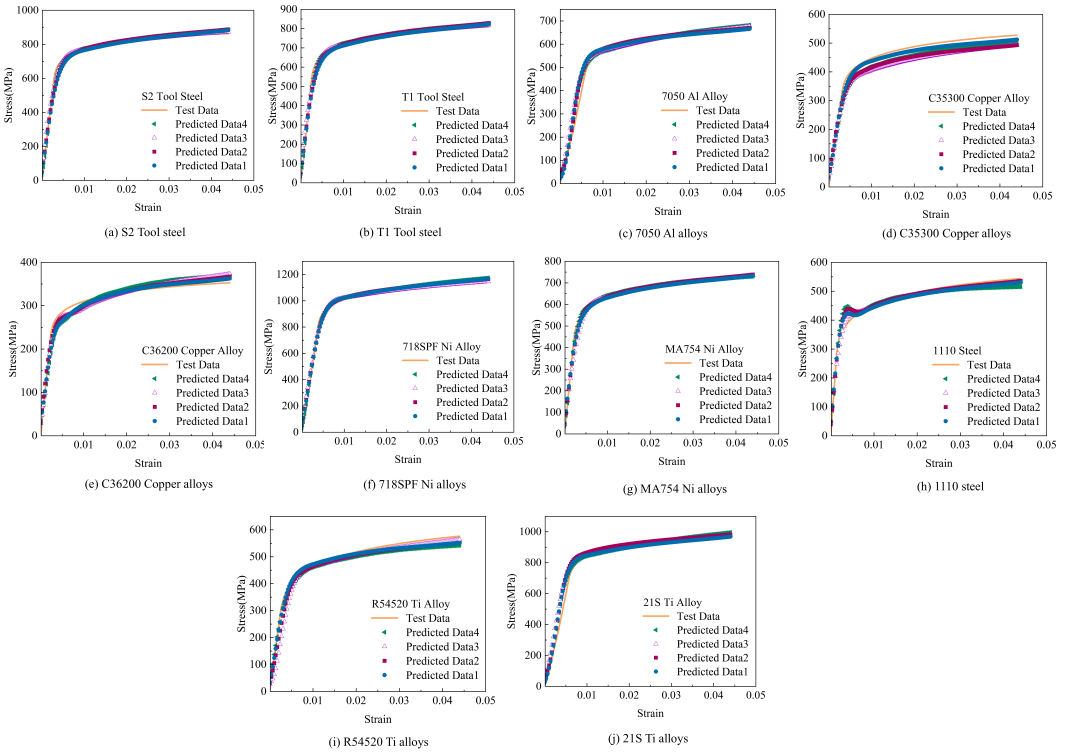


Figure 22. MA-LSTM-DLnet model for predicting stress–strain curves in group 2 materials: *Predicted Data1* which incorporates a full feature set, *Predicted Data2* excluding intensity coefficients, *Predicted Data3* excluding intensity coefficients and $E \times \varepsilon$, *Predicted Data4* excluding density, intensity coefficients and $E \times \varepsilon$.

Using Equation (21), the similarity between *Predicted Data3*, *Predicted Data4*, *Predicted Data5* (under *Setting 3*, *Setting 4*, and *Setting 5*, respectively) and the original test data for the second group of materials was calculated. This allows for a clearer observation of the errors between the model's predictions and the test data. As shown in Figure 24, the horizontal axis labels A2, B2, C2, D2, E2, F2, G2, H2, J2 and K2 correspond to the ten materials in the second group: S2 Tool steel, T1 Tool steel, 7050 aluminum alloy, C35300 copper alloy, C36200 copper alloy, 718SPF nickel alloy, MA754 nickel alloy, 1110 steel, R54520 titanium alloy and 21S titanium alloy. The vertical axis represents the yield strength of the corresponding materials. From the figure, it can be observed that, under different feature input configurations, the similarity to C35300 copper alloy ranges from 80% to 93%, which is the lowest among the materials. For the other materials, the similarity exceeds 91%. Comparing the similarities under different conditions, it is evident that the influence of different feature parameters on prediction accuracy is minimal and lacks a clear pattern. This contrasts with traditional mathematical methods, where feature parameters are critical, and the absence of any single parameter can significantly impact the calculation results.

Therefore, given a sufficient amount of material data, it is possible to train a prediction model with limited parameter dependency. Such a model can not only replace traditional complex mathematical calculations but also establish a unified predictive framework for material constitutive relationships, eliminating the need for repetitive tensile experiments when switching materials.

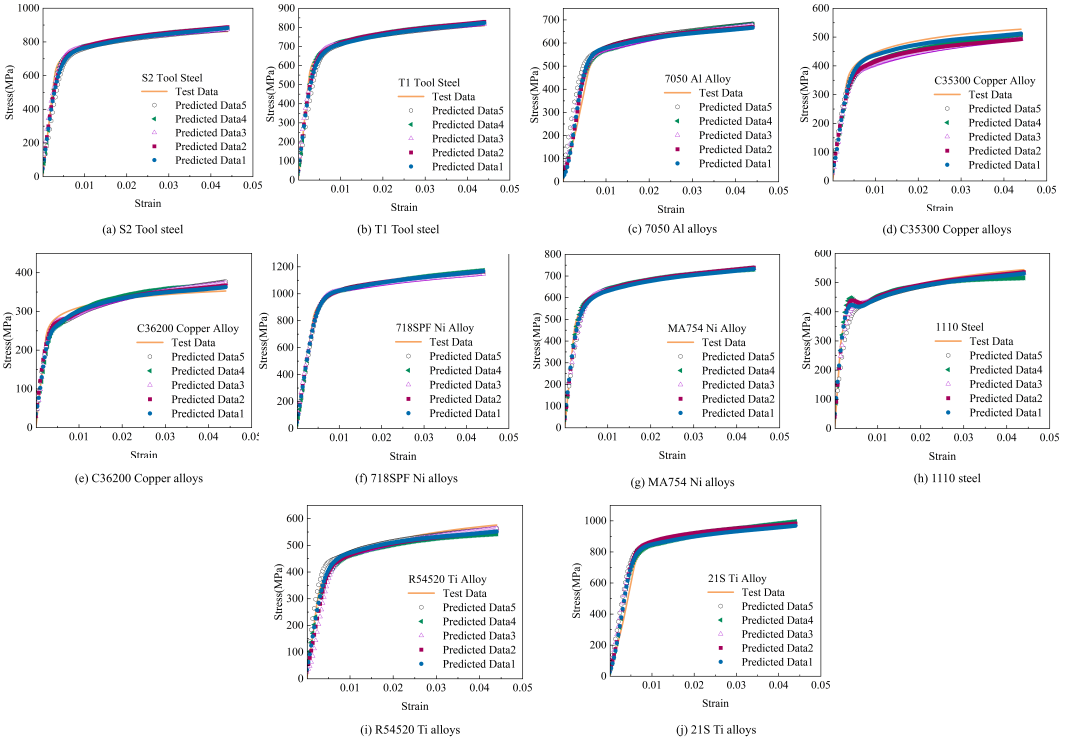


Figure 23. MA-LSTM-DLnet model for predicting stress–strain curves in group 2 materials: *Predicted Data1* which incorporates a full feature set, *Predicted Data2* excluding intensity coefficients, *Predicted Data3* excluding intensity coefficients and $E \times \varepsilon$, *Predicted Data4* excluding density, intensity coefficients and $E \times \varepsilon$, *Predicted Data5* excluding elastic modulus, density, intensity coefficients and $E \times \varepsilon$.

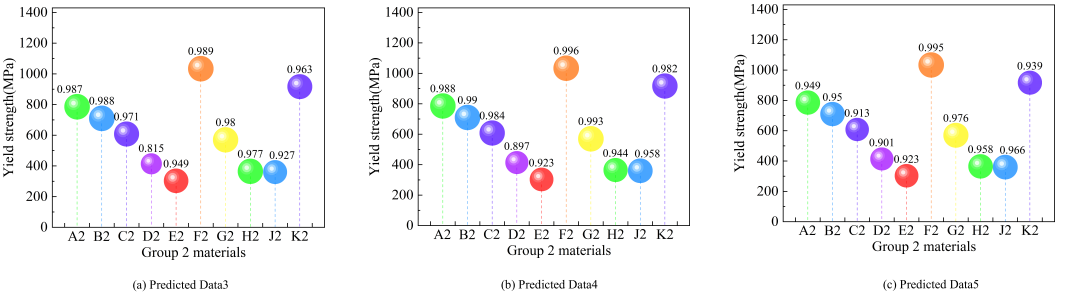


Figure 24. Similarity of predicted data3, predicted data4, predicted data5 and test stress–strain curves in group 2 materials.

4. Conclusions

This study systematically investigated the stress–strain constitutive relationships of metal alloy materials using deep learning methods, achieving the following innovative results:

(1) **High-Quality Dataset Construction**: The finite element method (FEM) was employed to numerically simulate the uniaxial tensile behavior of 106 common alloy materials, successfully constructing a multi-feature dataset containing complete stress–strain

curves. This dataset integrates fundamental mechanical properties and various physical characteristics, providing a high-quality data foundation for training deep learning models.

(2) **Innovative Deep Learning Model**: A novel deep learning time-series model, MA-LSTM-DLnet, was proposed, combining multi-head attention mechanisms and long short-term memory (LSTM) networks. By incorporating the attention mechanism, the model effectively captures key features in stress–strain curves, while leveraging the LSTM network's robust capability for processing sequential data to achieve precise predictions of material mechanical behavior. The results demonstrate that the similarity between the stress–strain curves predicted by the MA-LSTM-DLnet model and the test data exceeds 95%, significantly outperforming traditional methods.

(3) **Practical Advantages of MA-LSTM-DLnet**: A key advantage of the MA-LSTM-DLnet model is its reliance solely on conventional material properties (e.g., yield strength, tensile strength, and elastic modulus) for high-precision predictions, eliminating the need for hard-to-obtain parameters such as strength coefficients and hardening exponents required in traditional elastoplastic theories. This feature not only simplifies the computational process but also significantly enhances the model's practicality and reliability.

(4) **Limitations**: This study combines FEM with deep learning to predict the stress–strain constitutive relationships of materials, achieving good results. However, there are some limitations. Firstly, the model's accuracy depends on the training dataset's quality and quantity. Insufficient or unrepresentative data may lead to poor prediction. Secondly, the model has no physical significance and cannot provide physical explanations for the stress–strain relationship. It merely offers empirical predictions based on data. Lastly, the model's current applicability is limited to the 106 metal alloys studied and may not generalise to other materials.

(5) **Future Directions**: Future research should focus on expanding and diversifying the training dataset, collecting data for more materials with different properties and behaviours to enhance accuracy and generalisation. Also, hybrid models combining deep learning with physical models could be explored to integrate the advantages of both. Furthermore, the model can be improved by adding more relevant features and optimizing its structure and hyperparameters. This would make it a more powerful tool for material constitutive relationship prediction.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on a reasonable request. All data used in this study are fully available without any restrictions. The data that support the findings of this study are available from the corresponding author upon request.

Acknowledgments

We extend our sincere appreciation to the reviewers for their constructive comments and suggestions, which have significantly contributed to the improvement of this manuscript.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

This work was supported by the Start-up Funding Project from Guangdong Polytechnic Normal University (Grant Number 2021SDKYA017), the Program for Scientific Research Start-up Funds of Guangdong Ocean University (Grant Numbers 060302062311 and 060302062315), the Collaborative Education Project of Industry University Cooperation by the Ministry of Education of China (Grant numbers 230802784214645, 230803817223502 and 230821384507208). The authors are grateful for the financial support provided by these funding bodies, which made this research possible.

Appendix A.

Appendix A provides a list of the 106 materials required for this study, along with the corresponding computational parameters needed for tensile simulations. Since the core focus of this study is on exploring the predictive capabilities of deep learning methods—specifically, whether deep learning can learn and predict material stress–strain relationship curves—we prioritized the overall trend of the stress–strain curves over achieving zero-error alignment with experimental data for each material. Given the time constraints, our primary concern was whether the stress–strain curve trends could be accurately predicted using deep learning methods, rather than achieving perfect experimental accuracy.

For this reason, we did not conduct extensive experiments to obtain stress–strain curves for each material. Instead, we utilized the finite element method (FEM) to simulate and generate the material data. Consequently, the material parameters provided in the appendix are primarily sourced from online databases, literature, and other references. These parameters have not been rigorously verified for every material. If you intend to use the material parameters provided in the appendix, please evaluate their suitability for your specific research needs and determine whether further rigorous verification is necessary. The materials are categorized as follows:

Entries 1–21: Aluminum alloys
Entries 22–41: Copper alloys
Entries 42–60: Tool steels
Entries 61–80: Nickel alloys
Entries 81–93: Titanium alloys
Entries 94–106: Other alloy steels

No.	Name	Treatment	σ_y	σ_u	E	ρ	ν	K
1	5052	H36	220	267	70	2680	0.33	350
2	7A04	Solid	540	570	71	2800	0.33	630
3	2219	T81 temper	350	455	73	2839	0.33	515
4	2224	T351tempr	330	460	71	2800	0.33	520
5	2017	T4 temper	245	390	73	2800	0.33	550
6	2020	T6 temper	275	310	73	2780	0.32	370
7	3003	Solid, F temper	115	142	69	2730	0.33	242
8	3105	Solid	100	140	70	2700	0.33	240
9	4032	T6 temper	290	380	70	2700	0.33	440
10	5005	H18 temper	160	200	69	2700	0.32	300
11	5182	Annealed	100	175	70	2700	0.32	275
12	5754	O temper	120	165	69	2670	0.32	265
13	6013	T4 temper	180	317	70	2700	0.33	377
14	6110	T6 temper	380	480	69	2700	0.32	540

15	7050	T73 temper	608	650	73	2820	0.33	710
16	7249	T76511 temper	562	656	73	2820	0.33	716
17	7475	T61 temper	469	524	71	2810	0.33	584
18	201.0	T7 castings	415	470	70	2710	0.33	530
19	332.0	T5 temper	276	310	70	2700	0.33	370
20	355.0	T6 temper	275	295	70	2700	0.32	355
21	A201.0	T7 temper	442	480	70	2700	0.32	540
22	C10100	R220	140	220	130	8930	0.34	320
23	C10300	H01 temper	275	320	120	8930	0.35	380
24	C10500	Cold-rolled strip	410	510	120	8900	0.35	570
25	C10700	Cold-rolled strip	230	320	120	9000	0.33	380
26	C12210	Annealed	392	617	120	8900	0.35	677
27	C13000	Annealed	221	300	120	8900	0.34	360
28	C14310	Annealed	200	260	120	8930	0.34	320
29	C14520	Annealed	397	936	120	8930	0.33	936
30	C15720	Annealed	365	462	117	8940	0.34	522
31	C18070	TR02 temper	170	460	120	8930	0.34	520
32	C19720	HR02 temper	330	400	120	8840	0.34	460
33	C22600	HR01 temper	295	345	120	8780	0.34	405
34	C26800	HR01 temper	275	369	103	8470	0.34	429
35	C33500	Annealed	317	420	105	8470	0.35	480
36	C35300	Hard	414	515	105	8470	0.35	675
37	C36200	Annealed	303	338	100	8500	0.35	398
38	C43000	Hard	495	530	110	8400	0.24	590
39	C61400	Annealed	300	540	117	7890	0.34	600
40	C71500	H01 temper	310	460	152	8490	0.34	520
41	Cu–Cr–Zr	Solid, aging	254	328	100	8900	0.33	388
42	A8	Q&DT	620	1035	200	7870	0.3	1035
43	A9	HDT	635	1110	131	7870	0.3	1110
44	D2	HRC 37	320	780	180	7700	0.28	780
45	H10	QT	981	1280	206	7800	0.3	1280
46	CBS-223	Solid, annealed	1138	1310	216	7750	0.3	1310
47	H12	QT	868	1000	207	7840	0.32	1000
48	H21	QT	966	1100	207	7800	0.3	1100
49	M2	HRC 65	1140	1356	220	8100	0.3	1356
50	M42	Cast, annealed	610	925	220	8100	0.3	925
51	S2	Annealed	786	890	210	7800	0.3	890
52	T1	Annealed	710	886	210	7850	0.3	886
53	6F3	Solid, annealed	835	1035	210	7800	0.3	1035
54	CX	Solid, annealed	835	1151	210	7800	0.3	1151
55	MD Xtra	Annealed	827	979	210	7680	0.3	979
56	FX	Annealed	835	1414	200	7800	0.3	1414
57	PX	HRC 48	986	1225	200	7800	0.3	1225
58	WF	H, 444-477 BHN	1448	1621	200	7800	0.3	1621
59	DIEVAR	HRC 44-46	1210	1480	210	7700	0.32	1480
60	2101	Solid, annealed	617	860	215	7850	0.3	860
61	Nickel200	Annealed	148	370	206	8890	0.3	430
62	Nickel201	Annealed	100	380	206	8890	0.3	440
63	N03301	Age hardened	621	780	193	8030	0.3	840

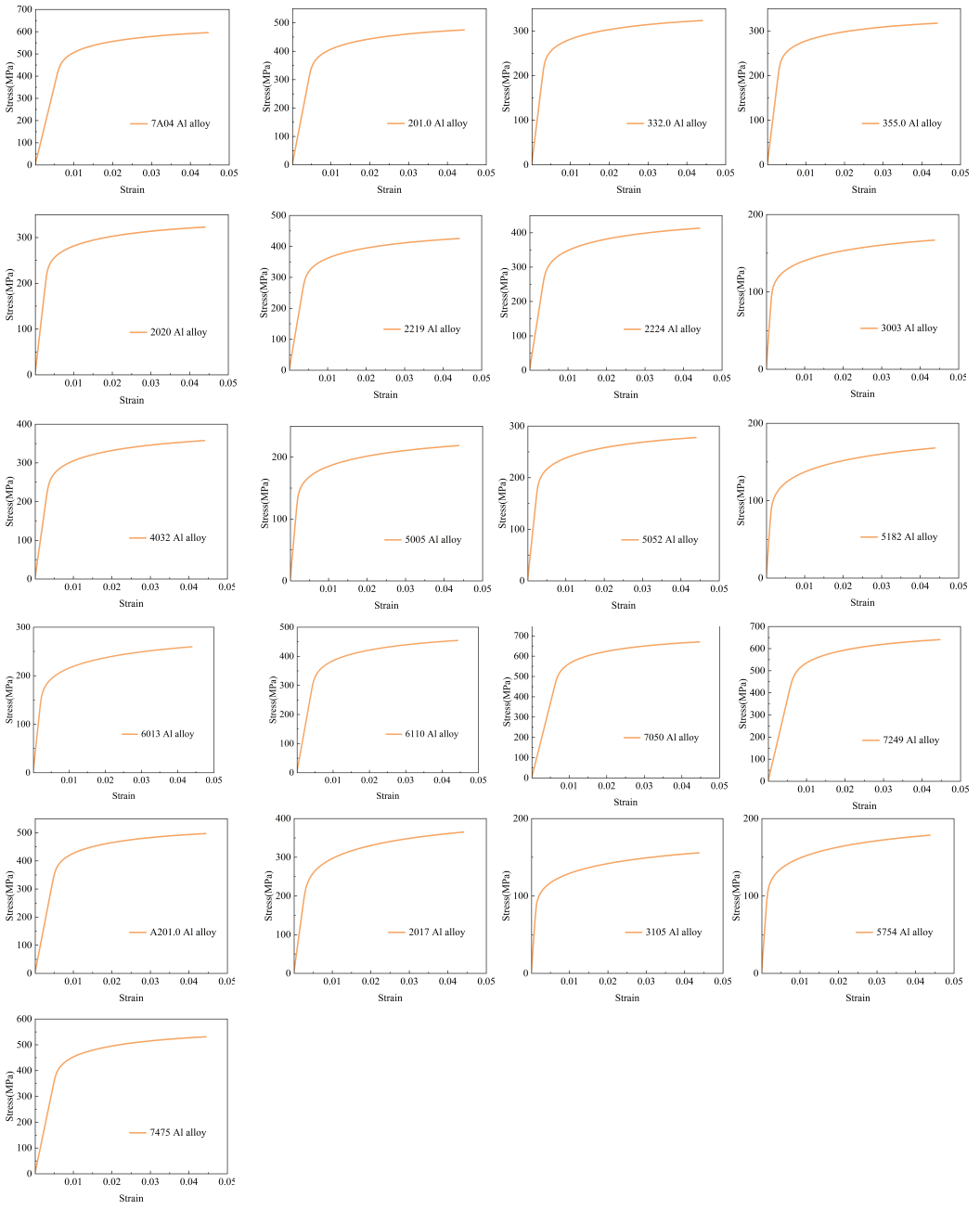
64	PM 1000	CR-A	345	450	210	8900	0.3	510
65	Monel 400	Annealed	320	482	176	8800	0.32	542
66	N06333	HRA	620	930	201	8200	0.3	930
67	20Cb-3	Annealed	380	510	193	8080	0.31	570
68	20Mo-4	Annealed	510	700	206	8050	0.3	760
69	N08031	SAA	347	573	206	8100	0.3	633
70	625 Plus	Double-aged	827	1065	206	8840	0.3	1065
71	Inconel 22	SHT sheet	310	690	207	8700	0.3	750
72	Inconel706	AR20	650	965	210	8050	0.32	965
73	718SPF	AA20	1034	1241	206	8840	0.3	1241
74	N07750	AHRT20	550	910	206	8250	0.3	910
75	MA754	AS20	570	950	203	8300	0.3	950
76	CMSX-4	HTA400	860	950	206	8700	0.33	950
77	Udimet 500	Aged, annealed	720	1340	206	8050	0.3	1340
78	718 Plus	Aged, annealed	1034	1241	206	8250	0.3	1241
79	AF955	Aged, annealed	827	965	206	8750	0.3	965
80	Waspaloy	Mill annealed	910	1335	213	8200	0.3	1335
81	Ti Grade 1	Solid, annealed	200	290	110	4510	0.33	350
82	TC4	Solid, annealed	830	1200	110	4440	0.34	1200
83	5 Al - 2.5 Sn	Solid	685	930	110	4420	0.33	930
84	GR9	Mill annealed	665	735	110	4470	0.39	795
85	ATI 425	Mill annealed	772	924	110	4450	0.33	924
86	UNS R54520	Annealed	360	860	120	4480	0.33	920
87	C-110M	Solid, longitudinal	810	900	120	4730	0.33	900
88	C-130AM	Mill annealed	900	1000	120	4510	0.33	1000
89	5621-S	SSTPF	985	1200	120	4510	0.33	1200
90	Ti-6242	AP-T-L	965	1100	120	4510	0.33	1100
91	Timetal 21S	AP-T-L-SA	917	946	110	4510	0.33	946
92	Ti - 1 Ge	Solid	860	1200	110	4510	0.33	1200
93	Ti15 Mo0.5 Si	Solid, 5 K/min	830	900	110	4510	0.33	900
94	1008 steel	Solid, oxidized	275	405	200	7872	0.33	465
95	1018 steel	Solid, oxidized	560	780	200	7872	0.33	840
96	EN08 steel	QT	530	660	200	7850	0.33	720
97	LZ50 steel	Solid, annealed	744	973	200	7850	0.33	973
98	G4802	Solid	375	489	200	7850	0.33	549
99	22MnB5 steel	Solid	262	395	210	7850	0.33	455
100	1110 steel	Solid	365	749	210	7850	0.33	809
101	1141 steel	Solid	583	843	210	7850	0.33	903
102	1151 steel	Solid	426	698	210	7850	0.33	758
103	1522 steel	Solid	286	433	200	7850	0.33	493
104	3140 steel	Solid	284	953	210	7850	0.33	953
105	5115 steel	Solid	379	651	210	7850	0.33	711
106	Stainless steel	Solid, annealed	652	820	208	7800	0.33	820

Appendix B.

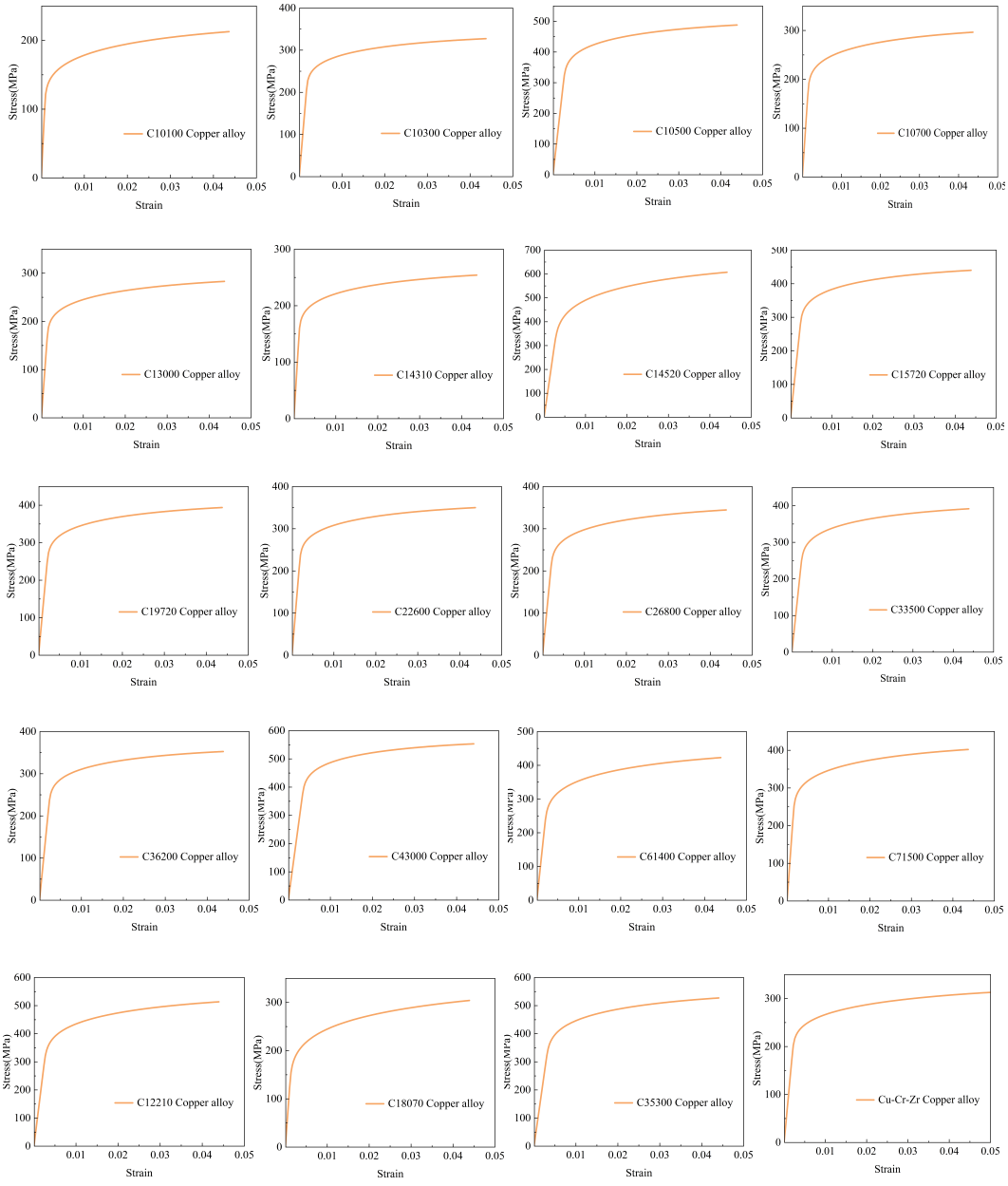
In this study, the finite element method (FEM) was employed to simulate the tensile behavior of 106 alloy materials. The resulting data were primarily used for model training to demonstrate the feasibility of deep learning methods for predicting material constitutive relationships. It is

important to note that the stress–strain curves for each alloy material provided in the appendix have not been rigorously validated. If you intend to use the data provided in the appendix, please assess its suitability for your specific research needs and determine whether further rigorous verification is necessary.

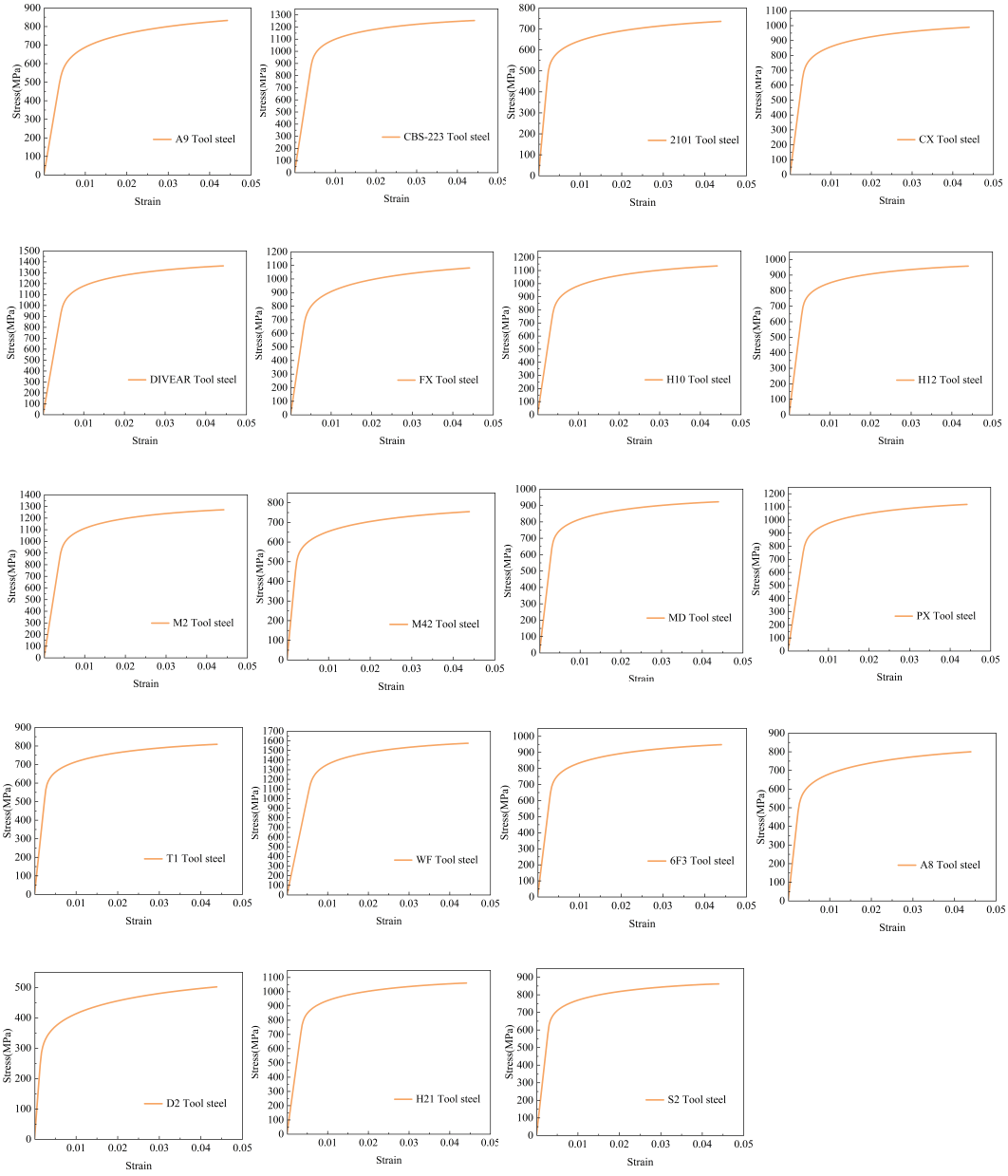
For ease of presentation, the stress–strain curve data for each alloy material are displayed in image format as follows:



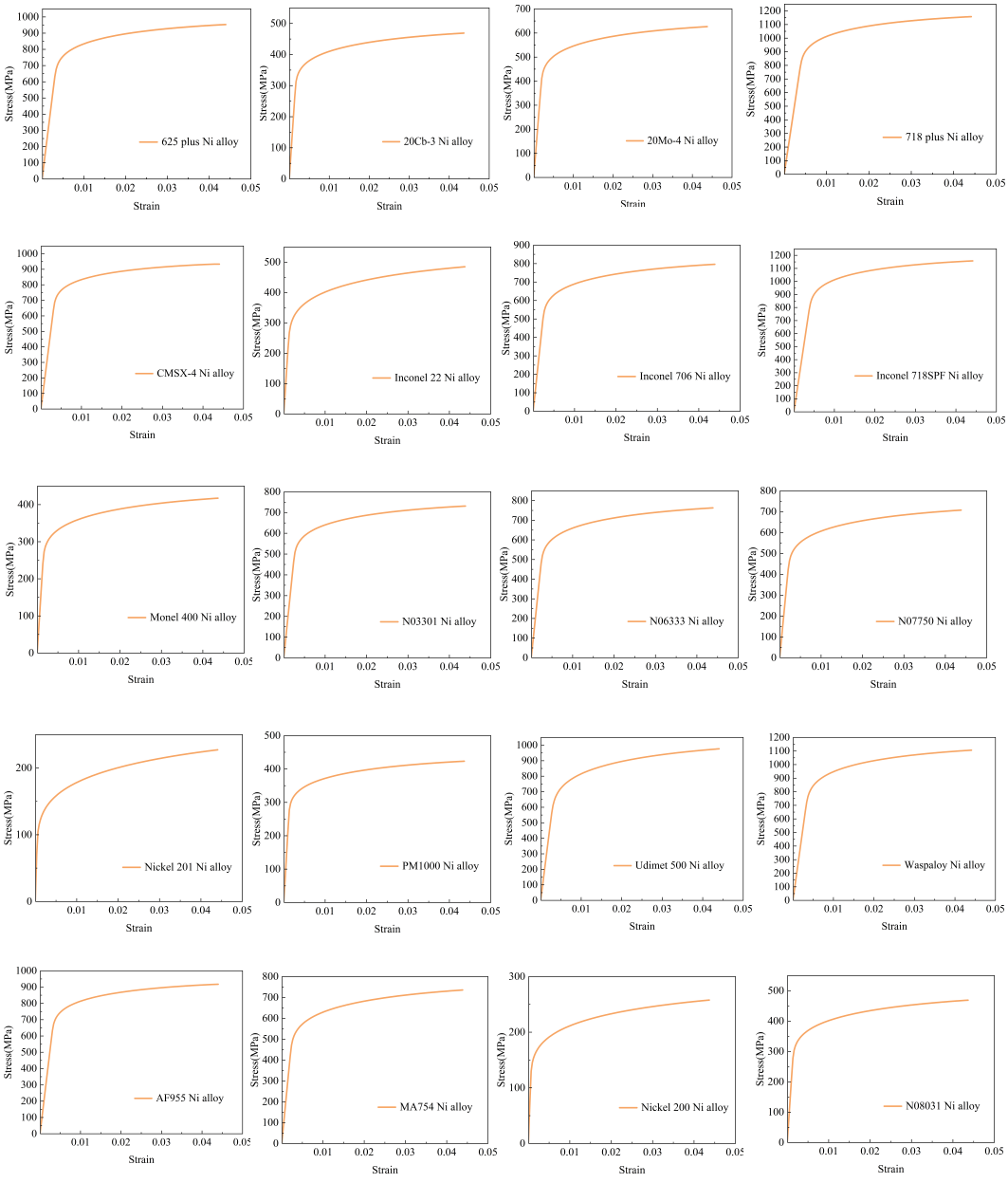
Figures 1–21 correspond to aluminum alloys



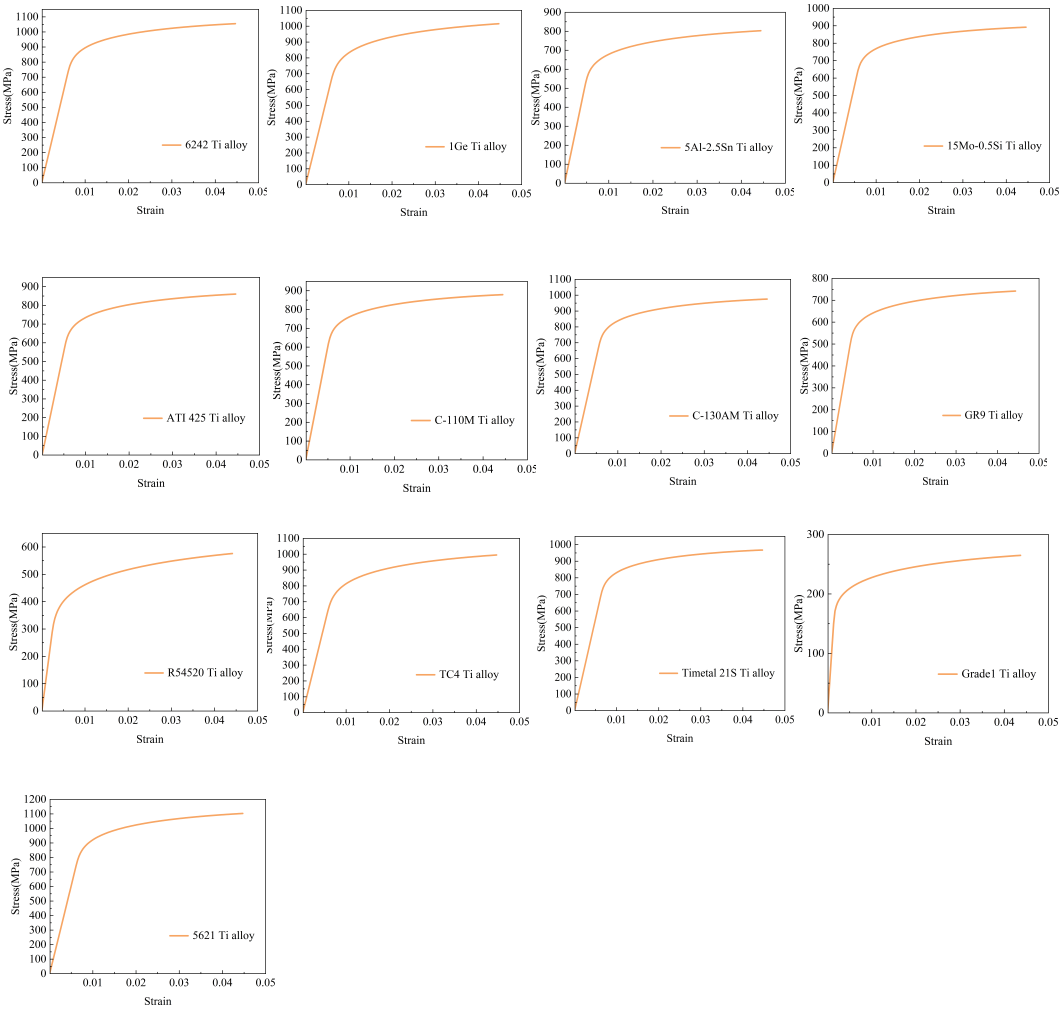
Figures 22–41 correspond to copper alloys



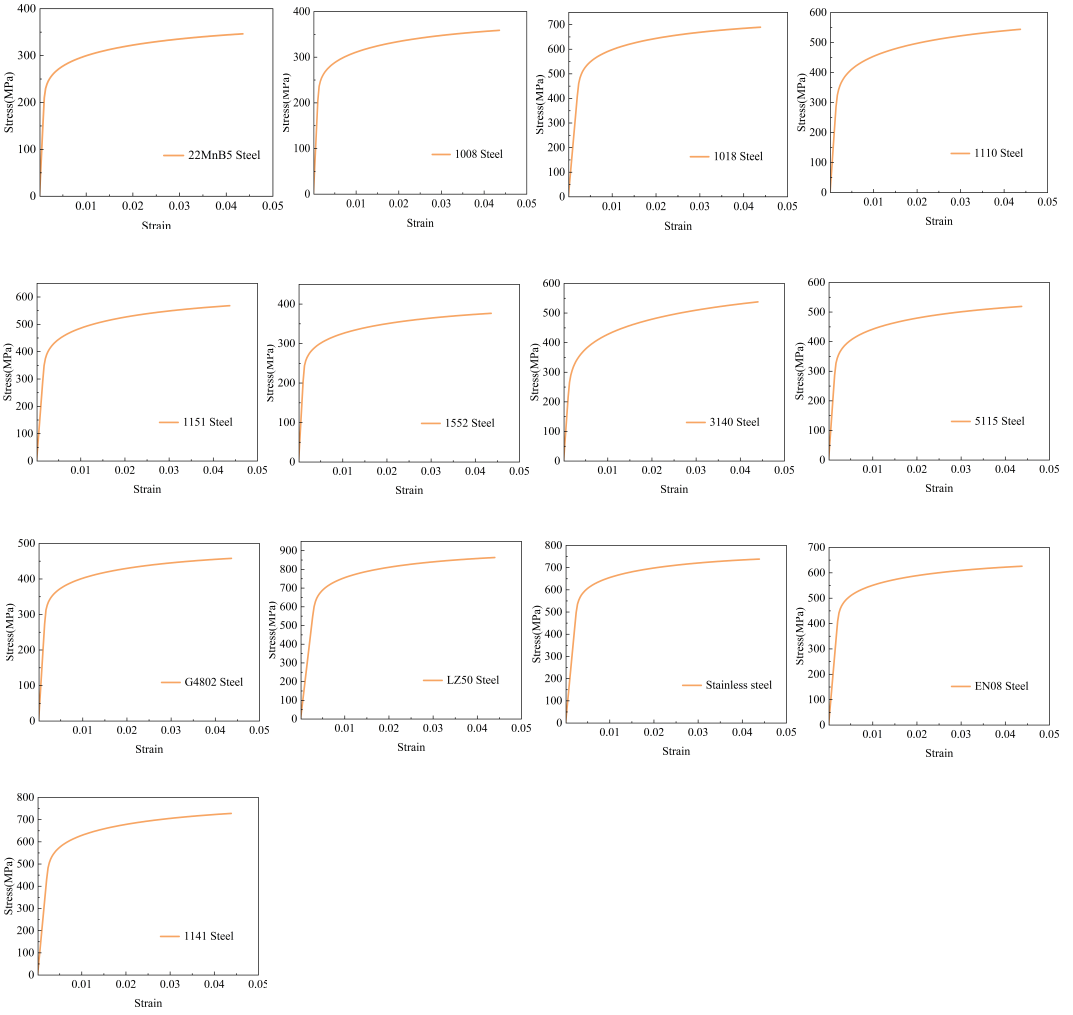
Figures 42–60 correspond to tool steels



Figures 61–80 correspond to nickel alloys



Figures 81–93 correspond to titanium alloys



Figures 94–106 correspond to other alloy steels

References

[1] S. Guchhait, B. Banerjee and J. Alla, "Thermo-elastic material parameters identification using modified error in constitutive equation approach", *Inverse Probl. Sci. Eng.* **26** (2018), pp. 1356–1382.

[2] C. Mareau, "Constitutive equations for thermo-elasto-plastic metallic materials undergoing large temperature variations", *Mech. Mater.* **181** (2023), article no. 104637.

[3] H. Yamada, T. Sakai, N. Ogasawara, M. Kokudai, D. Shiomitsu and T. Fukui, "Determination of a material constitutive equation that includes the effects of strain rate and temperature using an optimization method", *Ocean Eng.* **287** (2023), article no. 115719.

[4] L. Shi-Feng, S. Jia-Min, Y. Xiao-Kang, C. Jun and W. Qing-Juan, "High-temperature flow behaviour and constitutive equations for a TC17 titanium alloy", *High Temp. Mater. Processes* **38** (2019), pp. 168–177.

[5] M. Luo, J. Zhang, X. Liu and D. Xu, "Critical state elastoplastic constitutive model of angular-shaped and fragile granular materials", *Mar. Georesour. Geotechnol.* **39** (2020), pp. 1–14.

[6] R. Schwab and A. Harter, "Extracting true stresses and strains from nominal stresses and strains in tensile testing", *Strain* **57** (2021), article no. e12396.

[7] T. J. F. Southern, J. E. Campbell, C. Fang, A. Nemcova, A. Bannister and T. W. Clyne, "Use of hardness, PIP and tensile testing to obtain stress–strain relationships for metals", *Mech. Mater.* **187** (2023), article no. 104846.

[8] X. Qin, L. Ma and H. Wang, "Comparison analysis of dynamic modulus of asphalt mixture: indirect tension and uniaxial compression test", *Transportmetrica A: Transp. Sci.* **15** (2019), pp. 165–178.

[9] H. Lee, J. Yoo, Y. Kwon and J. Kang, "Calculation method for cold flow stress of Al6082 based on tensile test and compression test results", *Int. J Precis. Eng. Man.* **22** (2021), pp. 1337–1344.

[10] A. Ito, M. Yamato and S. Torizuka, "Measurement of true stress-true strain curve up to large strain extent at elevated temperature in Ti-6Al-4V alloy with image analysis tensile test method", *J. Jpn. Soc. Technol. Plast.* **62** (2021), pp. 177–182.

[11] W. Ramberg, *Description of stress–strain curves by three parameters*, Technical Note, National Advisory Committee for Aeronautics, 1943.

[12] K. B. Park, Y. T. Cho and Y. G. Jung, "Determination of Johnson-Cook constitutive equation for Inconel 601", *J. Mech. Sci. Technol.* **32** (2018), pp. 1569–1574.

[13] B. Si, Z. Li, G. Xiao and X. Shu, "Determination of mechanical properties from sharp dynamic indentation", *J. Strain Anal. Eng. Des.* **57** (2021), pp. 607–613.

[14] Y. Xiang, D. Zhong, P. Wang, G. Mao, H. Yu and S. Qu, "A general constitutive model of soft elastomers", *J. Mech. Phys. Solids* **117** (2018), pp. 110–122.

[15] S. W. Yoo, C. S. Lee, W. S. Park, M. H. Kim and J. M. Lee, "Temperature and strain rate dependent constitutive model of TRIP steels for low-temperature applications", *Comput. Mater Sci.* **50** (2011), pp. 2014–2027.

[16] C. Jo, J. Fu and H. E. Naguib, "Constitutive modeling for intercalated PMMA/clay nanocomposite foams", *Polym. Eng. Sci.* **46** (2010), pp. 1787–1796.

[17] J. D. Robson, T. F. Flint and D. C. J. P. Armstrong, "Calibration of constitutive models using genetic algorithms", *Mech. Mater.* **189** (2024), pp. 1.1–1.10.

[18] X. Luo, J. Y. Xu and Y. Liu, "The static constitutive model of fiber reinforced cellular materials", *Constr. Build Mater.* **187** (2018), pp. 903–906.

[19] Y. H. Xiang, D. M. Zhong, S. Rudykh, H. F. Zhou, S. X. Qu and W. Yang, "A review of physically based and thermodynamically based constitutive models for soft materials", *J. Appl. Mech.: Transac. ASME* **87** (2020), article no. 110801.

[20] M. Pouragha and R. Wan, "μ-GM: a purely micromechanical constitutive model for granular materials", *Mech. Mater.* **126** (2018), pp. 57–74.

[21] K. Zheng, Z. He, L. Che, H. Cheng, M. Ge, T. Si and X. Xu, "Deep alloys: metal materials empowered by deep learning", *Mat. Sci. Semicon. Proc.* **179** (2024), article no. 108514.

[22] W. Zheng, H. Liu, B. Wang and F. Sun, "Cross-modal learning for material perception using deep extreme learning machine", *Int. J. Mach. Learn. Cyb.* **11** (2020), pp. 813–823.

[23] C. Ruijin, L. Hechao, Y. Hope, J. Yang and R. Yi, "Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model", *Comput. Mater. Sci.* **150** (2017), pp. 212–221.

[24] Y. Zuo, M. Qin, C. Chen, W. Ye, X. Li, J. Luo and S. P. Ong, "Accelerating materials discovery with bayesian optimization and graph deep learning", *Mater. Today* **51** (2021), pp. 126–135.

[25] N. Karathanasopoulos and P. Hadjidoukas, "Deep learning based automated fracture identification in material characterization experiments", *Adv. Eng. Inform.* **60** (2024), article no. 102402.

[26] X. Meng, C. Qin, X. Liang, et al., "Deep learning in two-dimensional materials: characterization, prediction, and design", *Front. Phys.-Beijing* **19** (2024), article no. 53601.

[27] R. Kopp, J. Joseph, X. Ni, N. Roy and B. L. Wardle, "Deep learning unlocks X-ray microtomography segmentation of multiclass microdamage in heterogeneous materials", *Adv. Mater.* **34** (2022), article no. 2107817.

[28] B. W. Xu, S. Ye, M. Li, H. P. Zhao and X. Q. Feng, "Deep learning method for predicting the strengths of microcracked brittle materials", *Eng. Fract. Mech.* **271** (2022), article no. 108600.

[29] G. Fuping, L. Wei and F. Y. Chang, "Deep learning for time series-based acoustic emission damage classification in composite materials", *Russ. J. Nondestruct.+* **59** (2023), pp. 665–676.

[30] F. Jubair, A. Alhamayel, R. Aljaiose and K. A. Darabkh, "Deep learning model for early prediction of material fracture in tensile testing", *Neural Comput. Appl.* **36** (2024), pp. 10461–10474.

[31] Z. Tang, L. Guo, T. Zheng, Z. Li, R. Sun and K. Huang, "A combined machine learning and numerical approach for evaluating the uncertainty of 3D angle-interlock woven composites", *Compos. Struct.* **294** (2022), article no. 115726.

[32] J. Wu, Q. Chen, J. Jiang, G. Chatzigeorgiou and F. Meraghni, "Adaptive deep homogenization theory for periodic heterogeneous materials", *Compos. Struct.* **340** (2024), article no. 118171.

[33] Y. Matsuo, Y. Lecun, M. Sahani, D. Precup, D. Silver, M. Sugiyama, E. Uchibe and J. Morimoto, "Deep learning, reinforcement learning, and world models", *Neural Netw.: Official J. Int. Neural Netw. Soc.* **152** (2022), pp. 267–275.

[34] F. Nikolić and M. Čanaensuremathđija, "Deep learning of temperature–dependent stress–strain hardening curves", *C. R. Mec.* **351** (2023), pp. 151–170.

[35] K. Koenuma, "Estimation of texture-dependent stress–strain curve and r-value of aluminum alloy sheet using deep learning+1", *Mater. Trans.* **61** (2020), pp. 2276–2283.

[36] N. Zhang, S. Shen, A. Zhou and Y. Jin, "Application of LSTM approach for modelling stress–strain behaviour of soil", *Appl. Soft Comput.* **100** (2021), article no. 106959.

[37] I. Saha, A. Gupta and L. Graham-Brady, "Prediction of local elasto-plastic stress and strain fields in a two-phase composite microstructure using a deep convolutional neural network", *Comput. Method Appl. M* **421** (2024), article no. 116816.

[38] C. Chen, K. Chen and Y. Shu, "Data-driven bio-mimetic composite design: direct prediction of stress–strain curves from structures using cGANs", *J. Mech Phys. Solids* **193** (2024), article no. 105857.

[39] M. El Fallaki Idrissi, A. Pasquale, F. Meraghni, F. Praud and F. Chinesta, "Advanced meta-modeling framework combining machine learning and model order reduction towards real-time virtual testing of woven composite laminates in nonlinear regime", *Compos. Sci. Technol.* **262** (2025), article no. 111055.

[40] M. El Fallaki Idrissi, F. Praud, F. Meraghni, F. Chinesta and G. Chatzigeorgiou, "Multiscale thermodynamics-informed neural networks (MuTINN) towards fast and frugal inelastic computation of woven composite structures", *J. Mech. Phys. Solids* **186** (2024), article no. 105604.

[41] K. Smagulova and A. P. James, "A survey on LSTM memristive neural network architectures and applications", *Eur. Phys. J. Spec. Top.* **228** (2019), pp. 2313–2324.

[42] G. Van Houdt, C. Mosquera and G. Napoles, "A review on the long short-term memory model", *Artif. Intell. Rev.: An Int. Sci. Eng. J.* **53** (2020), pp. 5929–5955.

[43] Y. Xu, H. Wei, M. Lin, et al., "Transformers in computational visual media: a survey", *Comput. Vis. Media* **8** (2022), pp. 33–62.

[44] Y. Guitong, *Introduction to Elasticity and Plasticity*, Tsinghua University Press: Beijing, 2013.

[45] Y. Q. Wang, Z. X. Wang, X. G. Hu, J. K. Han and H. J. Xing, "Experimental study and parametric analysis on the stability behavior of 7A04 high-strength aluminum alloy angle columns under axial compression", *Thin Wall. Struct.* **108** (2016), pp. 305–320.

[46] Z. Ma, D. Z. Zhang, F. Liu, J. Jiang, M. Zhao and T. Zhang, "Lattice structures of Cu–Cr–Zr copper alloy by selective laser melting: microstructures, mechanical properties and energy absorption", *Mater. Design* **187** (2020), article no. 108406.

[47] X. Yun, Z. Wang and L. Gardner, "Full-range stress–strain curves for aluminum alloys", *J. Struct. Eng.* **147** (2021), article no. 04021060.

[48] R. Rajendran, M. Venkateshwarlu, V. Petley and S. Verma, "Strain hardening exponents and strength coefficients for aeroengine isotropic metallic materials—a reverse engineering approach", *J. Mech. Behav. Mater.* **23** (2014), pp. 101–106.

[49] J. Li, Y. Qiu, H. Wang and Z. Wang, "Estimation of the strength coefficient and strain hardening exponent from monotonic tensile properties of steels", *Int. J. Steel Struct.* **19** (2019), pp. 1951–1968.

[50] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin, "Attention is all you need", in *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA*, 2017.

[51] J. S. S. Hochreiter, "Long short-term memory", *Neural Comput.* **9** (1997), pp. 1735–1780.