



Attention-based multi-head feature-fusion network: A generalised method for hot deformation behaviour prediction in low-alloy steels

Lichao Yang , Arijit Lodh , Jiayang Qiu, Gustavo M. Castelluccio, Yifan Zhao *

Faculty of Engineering & Applied Sciences, Cranfield University, Cranfield, MK43 0AL, United Kingdom

ARTICLE INFO

Keywords:

Hot deformation
Low alloy steel
Shapley additive exPlanations
Work hardening
Tabular data

ABSTRACT

Modelling hot deformation of low alloy steels is important for optimising processing efficiency and cost reduction. Existing approaches lack generalisation as they primarily focus on single steel grades, ignoring chemical composition. To address this, a dataset comprising 58 distinct low-alloy steels and an Attention-Based Multi-Head Feature Fusion Network (AMHFnet) has been established and proposed. AMHFnet uses multi-head residual modules and a head weighting mechanism to adaptively learn high-dimensional representations of both chemical composition and processing variables, which are then fused and processed via a feature filtering and repeat step decision mechanism to predict hot deformation response. It demonstrated superior accuracy and generalisability in 10-fold cross-validation compared to existing decision tree-based models and state-of-the-art deep learning models. Ablation studies demonstrate the effectiveness of individual components of the network. An investigation of the effect of carbon demonstrated that AMHFnet could reliably reflect elemental effects on hot deformation behaviour, further validating its reliability and applicability. The model successfully captures work hardening behaviour and dynamic recrystallisation for most compositions. By accurately predicting hot deformation behaviour across diverse low-alloy steels, this work can simplify the alloy-composition design process and target the experimental testing.

1. Introduction

Low-alloy steels have been widely employed in automotive body structures, aerospace components, machinery manufacturing, and railway tracks (Herian and Anioek, 2008), (Tomita, 2000), (Lu et al., 2021), due to their superior mechanical properties. Investigation of their hot deformation behaviour facilitates the efficient selection and optimisation of hot working processing parameters during the design phase (Kumar et al., 2020), (Zhang et al., 2023a) and precise control of rolling, forging, and extrusion operations to achieve the desired mechanical performance (Dear et al., 2025), (Zhang et al., 2024). Moreover, the modelling of the stress-strain-strain rate-temperature behaviour in hot deformation can be used to construct processing maps. This can identify the safe and efficient strain rate-temperature windows for high temperature thermomechanical treatments and prevent stress concentrations that may lead to failure, such as cracking (Yang et al., 2015a). Hot deformation modelling also aids in tool design, e.g. dies, rollers, and other tooling, by predicting forces and stresses during deformation, reducing the risk of tool wear and failure. Therefore, the development of

high-accuracy models for hot deformation behaviour is of great importance for enhancing the economic efficiency and reliability of low-alloy steel production.

The common approach for modelling hot deformation is constitutive modelling (Savaedi et al., 2022) by either phenomenological approaches (e.g. Arrhenius law, Johnson-Cook model) (Johnson and Cook, 1983), (Molinari and Ravichandran, 2005), (Lin and Liu, 2010), (Ghosh et al., 2021), (Kingkam et al., 2020) or physics-based modelling (e.g. Zerilli-Armstrong model) (He et al., 2014), (Mirzaie et al., 2016), (Gurusamy and Rao, 2022), (Kanrar et al., 2023). Phenomenological approaches, such as Johnson-Cook, are simple to use and popular for simulating manufacturing processes like hot forging and hot rolling. However, the complex interrelationship between the process parameters and deformation mechanisms is generally ignored. It does not account for strain softening due to microstructural changes and is better suited for predicting only a characteristic stress (e.g. peak stress (Chen et al., 2019)) rather than a complete deformation response. On the other hand, physics-based approaches are based on thermally activated dislocation motion, capturing microstructure dependent plastic deformation

* Corresponding author.

E-mail address: Yifan.Zhao@cranfield.ac.uk (Y. Zhao).

<https://doi.org/10.1016/j.engappai.2025.113464>

Received 9 July 2025; Received in revised form 24 October 2025; Accepted 3 December 2025

Available online 8 December 2025

0952-1976/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

behaviour. They provide more accurate prediction of complete deformation response and onset of strain softening, compared to phenomenological approaches. However, the constants are material dependent and extrapolation of one model is limited to a narrow range of process parameters and alloy chemistry (Lee and Liu, 2006).

Modelling the hot deformation behaviour traditionally relies on vast amounts of laboratory experimental data (Li et al., 2013), (Li et al., 2012). Constitutive modelling struggles to process such high-dimensional, large-scale datasets, while neural network-based constitutive models (Churyumov et al., 2022a), (Mandal et al., 2006), (Ji et al., 2010), (Churyumov and Kazakova, 2023), (Mandal et al., 2009) are able to capture nonlinear relationships and multivariate interactions between multiple process parameters more efficiently, and there is no prior knowledge about the relationship between the parameters is required. Recent studies have shown the effectiveness of ML techniques, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and ensemble methods like XGBoost in capturing the material behaviour (Xu et al., 2025), (Xu et al., 2023), (Jain et al., 2024). Opêla et al. (2022) made a thorough comparison between shallow learning methods and deep learning methods, based on Restricted Boltzmann Machine (RBM) and an Auto-Encoder (AE). They observed both RBM and AE-based deep learning methods outperformed shallow learning for medium carbon steels. Huang et al. (2023) compared random forest (RF), support vector machine (SVM), radial basis function (RBF) ANN and backpropagation (BP) ANN for predicting the high-temperature deformation behaviour of GH3536 superalloy. They concluded prediction power of RBF ANN is superior to RF. However, Murugesan et al. (2023) also compared the use of SVM, single decision tree (SDT) and random forest regression (RFR) to characterize the flow behaviour of AISI 1045 steel. They found RFR performed the most accurately and efficiently. This is contrary to Huang's research results (Huang et al., 2023), which could be due to the complexity of the microstructure and alloy chemistry of different materials. For predicting hot rolling behaviour of micro-alloyed steels, Thakur et al. proposed a multivariate supervised machine learning approach that incorporates the effect of process parameters such as roll dimensions and their wear, temperature variation, which cannot be captured in mathematical models (Thakur et al., 2023).

In modelling the hot deformation behaviour of low-alloy steels, the phenomenological approaches, physically-based approaches, or data-driven approaches have a common limitation – the developed models are typically tailored to specific steel grades and lack generalisability to other compositions. These models are not applicable to predicting the deformation behaviour of different steel types. Moreover, the modelling process heavily relies on extensive and costly experimental data, which significantly prolongs the development cycle of new materials. Currently, there is a lack of a generalised modelling approach capable of predicting the high-temperature flow behaviour of low-alloy steels across a certain range of chemical compositions (Churyumov et al., 2022b).

To address this gap and achieve low-cost and generalised low-alloy steel hot deformation prediction, we proposed an attention-based multi-head feature-fusion network (AMHFnet) that effectively extracts and fuses features from both chemical compositions and physical conditions. Then, the feature filtering and the repeat step decision mechanism have been employed to achieve accurate behaviour prediction. The main contributions of this paper are as follows:

- Assembled a database encompassing 58 literature-based distinct steel compositions for further research.
- Proposed a feature encoding module that leverages multi-head residual encoders and attention mechanisms to adaptively learn the complex and high-dimensional representations for tabular data. Ablation studies have been conducted and have validated the effectiveness of this module and its individual components.

- Visualised the importance of the input features by SHapley Additive exPlanations (SHAP). By combining the research on the effects of alloying elements in the existing literature, we analysed and verified the interpretability of the proposed model.
- Analysed the prediction for multiple low-alloy steel grades and found that the data-driven methods that only rely on chemical composition and physical processing conditions are difficult to capture the behaviour when dynamic recrystallisation takes place, highlighting the importance of incorporating microstructural evolution descriptors such as recrystallised volume fraction, peak strain, steady state strain as function of alloy chemistry to data-driven models in the future.

The remainder of this paper is organised as follows: In Section 2, we present the methodology in detail, including the data acquisition process and its composition, the proposed network and designed modules, and the employed feature importance analysis method. Section 3 compares the performance of the proposed model with state-of-the-art approaches on the constructed dataset and conducts a comprehensive evaluation of the effectiveness of each module through ablation studies. Furthermore, Section 4 provides an in-depth analysis of model performance from three perspectives: fitting performance under varying processing conditions, model interpretability via feature importance, and a case of the effect of single element (carbon) content variations on the model performance. Finally, Section 5 concludes the paper, highlights the key advantages of the proposed model and its limitations, and provides future research directions.

2. Methodology

This study followed the workflow illustrated in Fig. 1. First, a data collection tool was developed to mine the literature for hot-deformation behaviour data on various low-alloy materials, which were compiled into a comprehensive dataset. Next, we designed a novel attention-based multi-head feature-fusion network (AMHFnet) for stress prediction. The performance of the model was evaluated by comparison with the commonly used and state-of-the-art methods. Then, the stress-strain curves have been visualised for the discussion of the hot deformation behaviour fitting across different physical conditions and materials. Finally, SHAP (SHapley Additive exPlanations) analysis has been utilised to quantify each input feature's contribution to the network's predictions, thereby revealing the relative importance of chemical and physical parameters in controlling hot-deformation behaviour. The training and evaluation were conducted using a laptop with a 1.2 GHz Core Ultra 5 125H CPU and 32 GB RAM, and an NVIDIA GeForce RTX 3090 GPU.

2.1. Dataset generation

2.1.1. Data acquisition

All the publications involved in this research for data acquisition were retrieved from the Scopus database based on two keywords, "Hot Deformation" and "Steel". The retrieved publications are further filtered and selected by the following requirements: (1) Since high temperature compression test is more reliable than tensile or torsion tests, only high temperature compression test data were shortlisted. (2) The alloy chemistry was restricted to low alloy steels as wide variations in steels will introduce a broader range of chemical composition and different factors active at hot deformation conditions (e.g. dynamic recrystallisation, dynamic recovery, mechanism of work hardening, etc). After filtering out, 56 publications have been identified (Zhao et al., 2020), (ping Xi et al., 2022), (Zhang et al., 2023b), (Li et al., 2020a), (Jia et al., 2022), (Dou and Liu, 2020), (Li et al., 2019), (Menapace et al., 2018), (lian Wei et al., 2023), (Qiao et al., 2020), (Zhao et al., 2023a), (Kishor et al., 2018), (Kawulok et al., 2022), (Gao et al., 2016), (Kim et al., 2020), (Yang et al., 2017), (He et al., 2024), (Li et al., 2020b), (Xiang

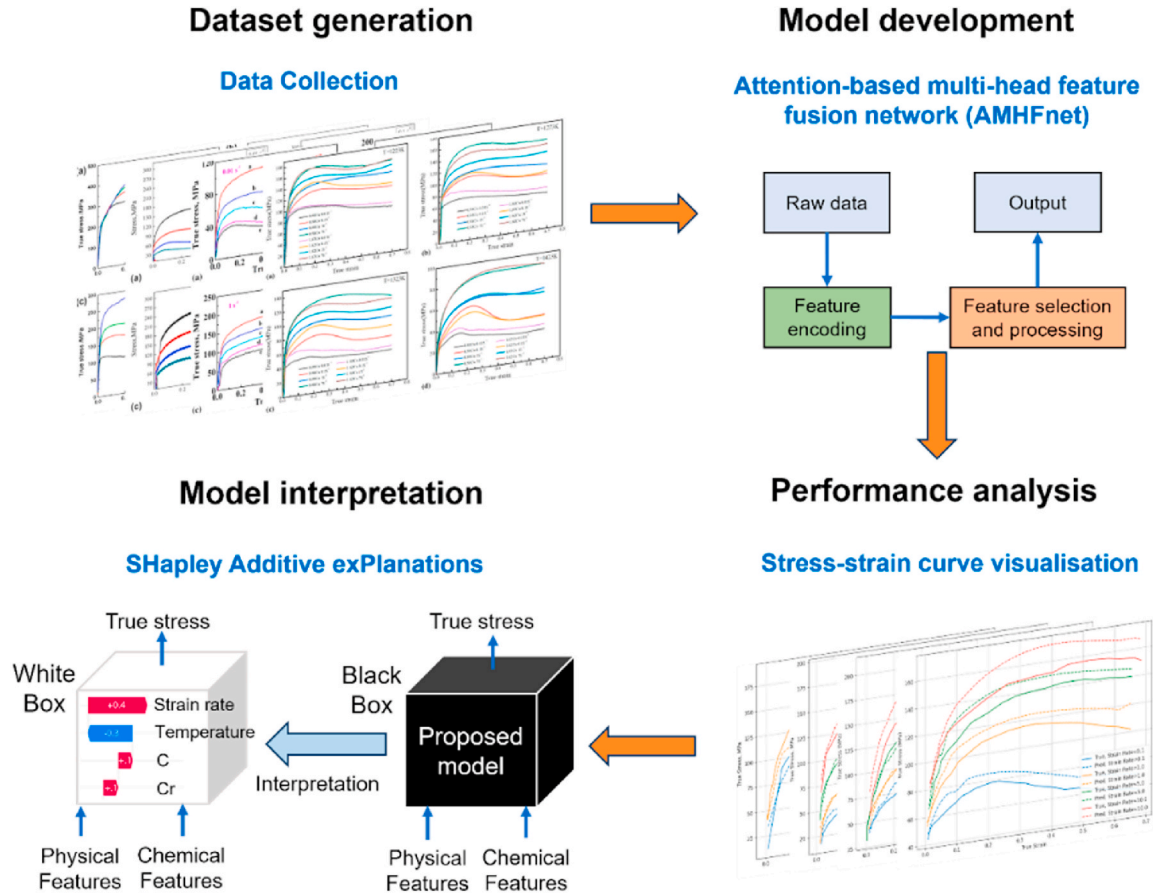


Fig. 1. Overall workflow of the study.

et al., 2023), (Jiang et al., 2023), (Huang et al., 2018), (Liu et al., 2018), (Ye et al., 2020), (Boroumand et al., 2021), (Li et al., 2022), (Kingkam et al., 2017), (Zhao et al., 2023b), (Alcelay et al., 2014), (Lian Wei et al., 2013), (Dong and Liu, 2019), (Chen and Nash, 2018), (Yin et al., 2013), (Thakur et al., 2020), (Rakhshkhorshid and Hashemi, 2013), (Wen et al., 2015), (Xu et al., 2016), (Dong et al., 2017), (Zhang et al., 2016), (Sajadifar et al., 2012), (Lin and Chen, 2010), (Ren et al., 2018), (XU et al., 2009), (Lin et al., 2009), (Wang et al., 2018), (Qi et al., 2011), (Singh et al., 2016), (feng Dong et al., 2009), (Zhao et al., 2023c), (Qian and Peng, 2015), (Sasaki and Katsumura, 2019), (Yang et al., 2015b), (Lan et al., 2019), (Duan et al., 2021), (Li et al., 2020c), (Jang et al., 2017), (Deng et al., 2024).

Based on these publications, 58 low-alloy steels with different chemical compositions were selected. For each stress-strain plot in the publications, at least 30 data points were extracted to cover the complete hot deformation behaviour. The features used in the dataset can be divided into two types: (1) chemical features (14 alloying elements) and (2) physical features (temperature, strain rate, true strain). The output feature is true stress. A detailed statistical analysis of the dataset is shown in Table 1. The data distribution for all steel grades can be found as a heatmap in the supplementary file.

The histogram distribution of all the input features is shown in Fig. 2. The temperature and true strain are more evenly distributed in their respective range than the alloying elements, while strain rate distribution is concentrated primarily below 10 s^{-1} , frequently used in experiments. As we considered only low-alloy steel, carbide or precipitate forming elements (Ti, Nb, V, Cr) are present in lower concentrations. Carbon concentration is considered for both low-carbon and high-carbon steels with a more uniform distribution in the range of 0–0.5 wt%. In contrast, Mn has a uniform distribution in the considered range. In total, the constructed dataset has 17653 sampling points for different

Table 1

Range and statistical information of the features in the dataset.

	Features	Max	Min	Average	Standard deviation
Chemical features	C(wt.%)	1.030	0.013	0.275	0.245
	Mn(wt.%)	1.780	0.140	1.087	0.442
	Si(wt.%)	1.720	0.020	0.363	0.326
	Cr(wt.%)	2.680	0	0.471	0.611
	Cu(wt.%)	0.500	0	0.045	0.109
	Mo(wt.%)	2.100	0	0.142	0.322
	Ni(wt.%)	1.800	0	0.22	0.436
	Nb(wt.%)	0.250	0	0.011	0.038
	Ti(wt.%)	0.083	0	0.009	0.019
	V(wt.%)	0.630	0	0.049	0.093
	Al(wt.%)	0.970	0	0.025	0.111
	P(wt.%)	0.033	0	0.009	0.006
	S(wt.%)	0.086	0	0.009	0.013
	N(wt.%)	0.018	0	0.004	0.005
Physical features	Strain rate (1/s)	30	0.001	2.577	5.574
	Temperature (K)	1473	1023	1271.705	102.92
	True strain	1.111	0	0.327	0.260
	True stress (MPa)	343.52	2.987	96.389	56.67

steel grades. The data is cleaned by deleting outliers, and the final dataset has 17444 data points.

2.1.2. Data pre-processing

Data pre-processing was applied to both physical and chemical features before feeding them into the network. As seen in Fig. 2, the physical features have a wide range of values. For example, deformation

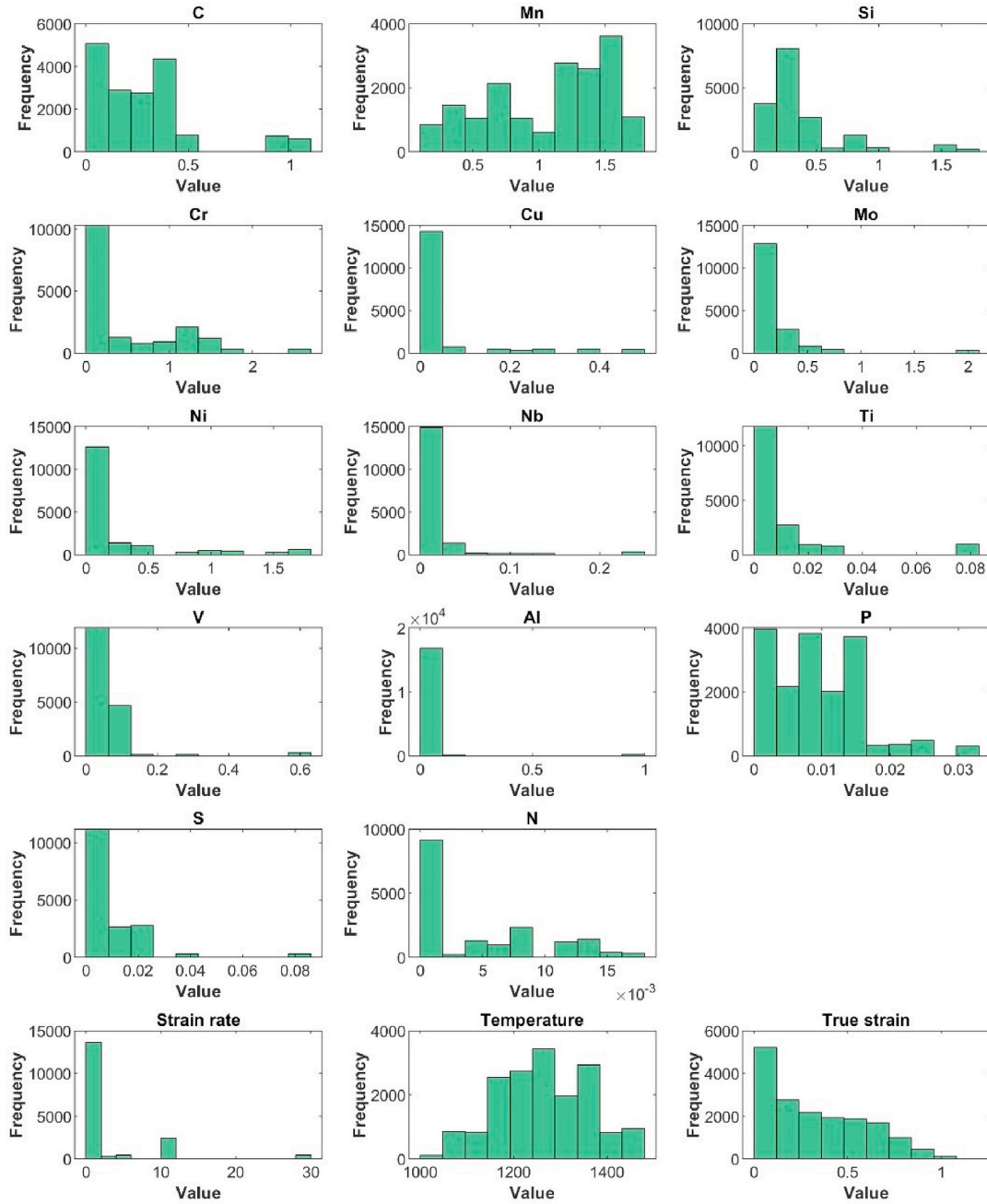


Fig. 2. Histogram distribution of the chemical and physical features.

temperature is in the range of 1023–1473K while the amount of true strain is 0–1.11. To mitigate the influence of differing scales and enhance model convergence and accuracy, Z-score standardisation was applied. This transformed the physical features in the dataset into a standard normal distribution with a mean of 0 and a standard deviation of 1 using Equation (1).

$$S = \frac{X - \mu}{\sigma}, \quad (1)$$

where X represents the original data, μ is the mean of the data, and σ is the standard deviation of the data.

For the chemical features, Fig. 3 illustrates the Pearson correlation analysis of 14 chemical elements. It can be observed that the correlations between multiple chemical elements are high, especially the correlation between Nb and Mo is 0.75, between V and Mo is 0.68, and between V

and Nb is also 0.69. This high correlation indicates that these elements have significant collinearity. The high-dimensional feature space created by all of these features could introduce difficulties to the feature extraction and lead to the overfitting problem, which poses a challenge to the generalisation of the model. Hence, principal component analysis (PCA) has been employed to reduce the dimensionality of the chemical features (from 14 to 8).

2.2. Attention-based multi-head feature-fusion network

The diverse alloy compositions of low-alloy steels create a sparse feature space. Accurate prediction of their hot deformation behaviour requires large amounts of experimental data, which is costly and time-consuming to obtain. This poses a key challenge for applying data-driven or machine learning approaches. Therefore, under such limited data conditions, it is of great importance to develop an efficient and

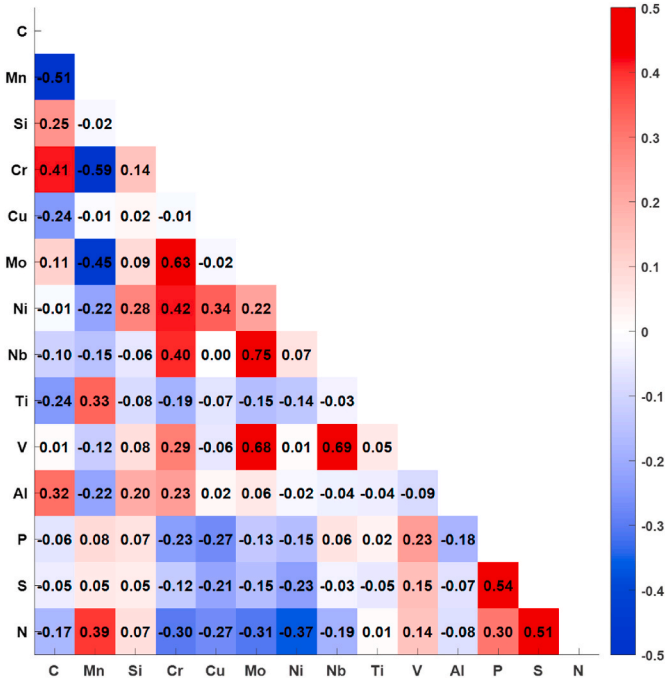


Fig. 3. Pearson correlation coefficients heatmap of chemical features.

generalisable representation model. To address this, this work proposed an AMHFnet. The data used in the conventional regression task comprises multiple diverse features that correspond to meaningful variables. Decision tree-based methods or feedforward networks are commonly used methods for the regression problem. However, for this task, there are some characteristics of the input data that pose challenges for the direct adoption of the mentioned methods, where chemically relevant features and physically relevant features are independent of one another, while the chemical features exhibit strong intercorrelations among constituent elements. The idea of the proposed AMHFnet is to split the raw inputs into two streams, which enables the model to independently learn higher-dimensional representations for both chemical and physical features, preventing interference between heterogeneous feature types. The learned representations were then integrated and merged through efficient feature selection mechanisms to produce the prediction. As presented in Fig. 4(c), AMHFnet includes two parts, which are feature encoding, and feature selection and processing. The feature encoding learns a high-dimensional representation of the materials' chemical and physical features. A feature selection and processing module has been employed to obtain the final prediction from the representations.

The feature encoding section consists of three modules, which are multi-head encoder Fig. 4(a), head weighting Fig. 4(b) and feature fusion. The multi-head encoder takes the input $X \in R^{n \times f}$, where n is the number of the sample in one batch, f is the number of the input features, and generates a multi-head representation $M \in R^{n \times h \times f}$, where h is the number of the head. It employs multiple heads to operate on low-dimensional projections of the input, which allows the module to learn diverse and complex representations. The calculation of each head is independent and leads to different results. For each head, the encoder is a 3-layer residual fully connected (FC) layer. The residual connection has been widely used in convolutional neural networks and other deep learning architectures, due to its capabilities, such as facilitation of identity mappings (allowing layers to skip transformations), mitigation of the vanishing gradient, etc. However, in the context of the data addressed in this study, the demand to learn high-dimensional representations and the difference in feature dimensions prevent the direct application of residual connections. To address this, we propose a novel

multi-head fully connected residual module, which is presented in Fig. 4 (a). This module leverages the advantages of residual learning to learn high-level representations efficiently. Specifically, we designed an inverted bottleneck structure, which expands the number of neurons in the first 2 FC layer four times the input features, and narrows the number of neurons in the last FC layer to the same as the input features. Such an architecture could keep the dimension while improving the representation power of the network. The head can be expressed as:

$$head(X) = (W_3^h \cdot \sigma(W_2^h \cdot \sigma(W_1^h X + b_1^h) + b_2^h) + b_3^h) + X, \quad (2)$$

where W^h and b^h are the weight and bias of the fully connected layer for the head h . σ is the ReLU activation function. $W_1^h \in R^{f \times f}$, $W_2^h \in R^{4f \times 4f}$ and $W_3^h \in R^{f \times f}$.

Then, a multi-head mechanism has been employed to enhance the diversity of the learned representations and increase the module's capability of capturing complex patterns and dependencies. The outputs of each head are concatenated to obtain the representation, which can be expressed as

$$MultiHead = Concat(head_1, \dots, head_h). \quad (3)$$

To further model the interdependencies between heads and extract the importance of each head, a head weighting module has been developed, which can adaptively recalibrate head-wise representations and is presented in Fig. 4 (b). In this module, the multi-head representation M is squeezed in the feature domain by a 1D adaptive average pooling (AAP_{1D}) to integrate the feature information. Then 2 FC layers are applied to generate the head weighting function H_w , which can be calculated as:

$$H_w = \delta(W_{h2} \sigma(W_{h1} AAP_{1D}(M))), \quad (4)$$

where δ is the sigmoid activation function, $W_{h1} \in R^{h \times h}$ and $W_{h2} \in R^{h \times \frac{h}{2}}$.

The output of the head weighting module $M_H = (M_{H1}, M_{H2}, \dots, M_{Hh})$ can be expressed as:

$$(M_{H1}, M_{H2}, \dots, M_{Hh}) = H_w(M_1, M_2, \dots, M_h). \quad (5)$$

In the feature fusion module, a 1d convolutional layer is employed to integrate the features in the weighted multi-head representation M_H in the head domain and further expand it as channels for the output $M_O \in R^{n \times c}$. The convolutional operation can be expressed as:

$$M_{Oj} = \sum_i^h w_{Cj,i} * M_{Hi} + b_{Cj}, \quad (6)$$

where w_C is the kernel of the convolutional operation, whose size is f . b_C is the bias of the convolutional operation. i and j are the indicators for the input number of the head h and the output number of channels c . In this study, h is 12 and c is 48.

In the last part of the feature encoding section, the feature map used to feed into the feature selection and processing section is generated by a concatenation operation for the fusion of both chemical-related representation and physical-related representation, written as

$$M_F = Concat(M_O(X_{chemical}), M_O(X_{physical})). \quad (7)$$

The feature selection and processing section starts with a batch normalisation layer, which is used to enhance training stability. Then the merged features are fed into sequential decision steps for representation learning. Given the superior capability of feature selection, feature transformer and attention transformer modules from TabNet (Arik and Pfister, 2021) have been employed in the decision steps for the efficient feature filtering. In each decision step, there are two main components, the feature transformer and the attentive transformer, which have been presented in Fig. 4(d) and (e). The attentive transformer is designed to learn the importance of the proceed features in the

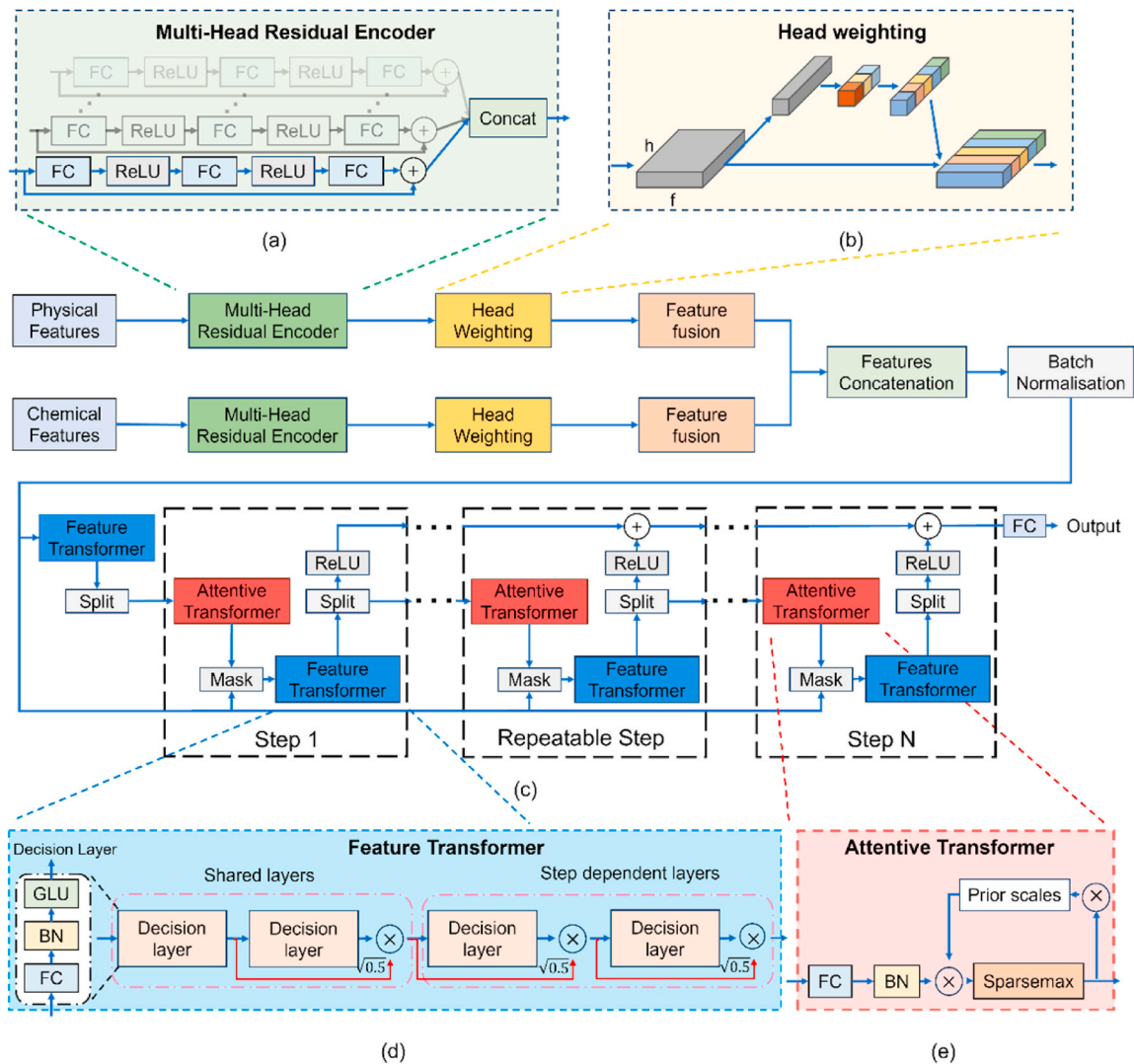


Fig. 4. The architecture of the proposed AMHFnet in this paper. (a) and (b) are the structure of the modules (multi-head residual encoder, head weighting) used in the feature encoding section. (c) is the flowchart for the structure of the AMHFnet. (d) and (e) are the structure of the modules (feature transformer and attentive transformer) used in the feature selection and process.

preceding step, a_{s-1} , by generating a mask $Mask_s$. It regularises the model by penalising over-diffuse attention and encouraging it to focus on the most relevant features. The $Mask_s$ can be expressed as:

$$Mask_s = \text{sparsemax}(P_s \cdot l_{As}(a_{s-1})), \quad (8)$$

where l_s is a trainable function, which is a FC layer, followed by BN layer and is presented in Fig. 4 (e), P_s is the prior scale term, which is used to record how much each feature is applied in the previous step and can be expressed as:

$$P_s = \prod_{j=1}^s (\gamma - Mask_j), \quad (9)$$

where γ is defined as a relaxation parameter to control the flexibility of feature reuse in multiple decision steps. P_0 is set as all ones, where there is no prior on the masked features. With the value of γ increase, the model has more flexibility in using features in different steps.

The attentive transformer utilises this masking mechanism to select significant steps features based on the current features and outputs from previous steps, ensuring that the model can focus on different important features at various steps. The filtered features are further processed by the feature transformer module, which is split into two components d_s

and a_s . Where d_s is the output of the current decision step s and a_s is the features that will be used for the attentive transformer in the next decision step $s + 1$. The whole process can be explicit as below.

$$d_s, a_s = l_{Ds}\{Mask_s \cdot l_F\} \quad (10)$$

The structure of the learnable functions, shared layers l_F and step-dependent layers l_{Ds} in the feature transformer module are presented in Fig. 4(d).

After the last decision step, the output features of all steps are integrated through the fully connected layer to finally generate the prediction result of the model. In this study, the number of steps is 3.

2.3. Evaluation

To evaluate the generalisation of the hot deformation behaviour prediction, the dataset is split based on the chemistry. Different chemical compositions were divided into two sets. The 10-fold cross-validation method has been used to evaluate the robustness of the model. Around 90 % of the chemical compositions (52 chemical compositions) and their corresponding data were used as the training dataset, while the remaining 10 % were used as the testing dataset (6 chemical compositions) in each fold.

The evaluation metrics used in this paper are the coefficient of determination (R^2), mean absolute error (MAE), and root mean squared error (RMSE). The equations are provided below:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (13)$$

Where $\hat{y}_i$ is the prediction, y_i is the actual value, $\bar{y}$ is the mean value of y_i

Furthermore, we compare the model's performance with several well-used Decision Tree (DT) based machine learning methods and state-of-the-art deep learning methods, which are commonly used in regression tasks.

2.4. Feature importance analysis

SHAP (SHapley Additive exPlanations) is an advanced method for explaining machine learning and deep learning models based on game theory (Lundberg et al., 2019). The SHAP value considers all possible feature combinations, evaluates the marginal contribution of each feature under different feature combinations, and averages these marginal contributions to finally obtain the SHAP value of each feature. Specifically, for each input feature of the model, SHAP will traverse all possible feature subsets and calculate the changes in model output before and after adding a certain feature. In this way, the marginal contribution of each feature to the prediction result can be quantified. Then, these marginal contributions are weighted averaged to ensure that the importance assessment of each feature is fair and explanatory. Given a feature, the formula for its SHAP value ϕ_i is as follows:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(N - |S| - 1)!}{N!} [f_x(S \cup \{i\}) - f_x(S)], \quad (14)$$

where N is the set of all features, S is a feature subset in N that does not contain i , $f_x(S)$ represents the output of the model using only the feature subset S , $f_x(S \cup \{i\})$ represents the output of the model using the feature subset S plus feature i , $|S|$ is the size of the feature subset S (i.e. the number of features), N is the total number of all features.

3. Results

3.1. Training and methods comparison

All models were trained on a workstation equipped with an Intel(R) Core(TM) Ultra 5 125H CPU, 32 GB of RAM, and an NVIDIA GeForce RTX 3090 GPU. Training was conducted for 150 epochs. The Adam optimisation algorithm was used with an initial learning rate of 2×10^{-3} , which was decreased at the 80th and 120th epochs.

Table 2 presents the comparison of the proposed AMHFnet with other decision tree-based and state-of-the-art deep learning based methods in terms of mean and standard deviation values calculated from the test dataset in 10 folds across R^2 , MAE, and RMSE. It can be observed that the conventional decision tree-based methods, such as CatBoost, ExtraTrees, LightGBM, and their variants, show a similar performance. The RMSE values of CatBoost, ExtraTreeMSE, and LightGBMXT all hover around 16.5, which demonstrates the strength of such methods in capturing nonlinear patterns effectively. In contrast, most deep neural network models did not demonstrate a competitive performance for this task. For instance, the radar chart of performance as plotted in Fig. 5 shows TabNet and TabKANet are concentrated in the middle, which achieved RMSE values of 19.089 and 19.497, respectively, with

Table 2

The comparison of the evaluated methods on the evaluation metrics. The results are calculated based on the 10-fold cross-validation.

Models	R^2		MAE		RMSE	
	Mean	Std dev	Mean	Std dev	Mean	Std dev
CatBoost	0.907	0.014	11.797	1.442	16.185	2.613
ExtraTreeMSE	0.901	0.026	11.924	1.323	16.489	2.280
LightGBM	0.893	0.016	12.662	1.727	17.492	3.294
LightGBMXT	0.902	0.015	12.119	1.662	16.651	2.871
RandomForestMSE	0.875	0.030	13.466	1.528	18.645	2.780
Xgboost	0.887	0.018	12.886	2.129	18.096	4.395
Tabnet (Arik and Pfister, 2021)	0.866	0.052	14.313	2.437	19.089	3.339
AutoGluon-NN (Erickson et al., 2020)	0.881	0.029	13.507	2.438	18.200	3.014
AutoGluon-MLSE (Erickson et al., 2020)	0.899	0.023	12.159	1.559	16.757	2.706
TabKANet (Eslamian et al., 2025)	0.851	0.071	15.777	4.102	19.497	5.142
TabPFN (Hollmann et al., 2022)	0.912	0.028	11.635	1.661	15.553	2.351
AMHFnet – no PCA	0.899	0.030	12.522	2.008	16.565	2.487
AMHFnet	0.915	0.018	11.506	1.298	15.346	1.934

standard deviations as high as 3.339 and 5.142. This indicates a high sensitivity to data distribution and suggests a significant risk of overfitting. The complex architectures of these deep models struggle to generalise on such a scale of feature space. Moreover, a comparison between the proposed models with and without PCA has been shown in Table 2. The results indicate that removing PCA and directly using the full 14-dimensional chemical composition features leads to a more complex feature space and challenges the model's representation learning capability.

In comparison, it can be seen from Fig. 5, the proposed AMHFnet outperforms all other models across all metrics. It achieves the highest R^2 value of 0.915 and the lowest MAE and RMSE values, at 11.506 and 15.346, respectively, with superior fitting accuracy. Moreover, it also demonstrates the lowest standard deviations for MAE (1.298) and RMSE (1.934), indicating the robustness across different data splits, and thus excellent generalisation capability.

Table 3 presents a comparison of model and computational complexity among the evaluated deep neural network-based models. The proposed model contains 0.09M parameters and requires 1.52 MFLOPs, which has slightly more parameters than Tabnet and TabKANet, but significantly outperforms them in both prediction accuracy and stability. Compared with the next best model, TabPFN, our model achieves higher performance with fewer parameters and lower computational cost.

3.2. Ablation studies

In addition, an ablation study was conducted to investigate the impact of each module within the proposed model and to verify the effectiveness of the design. The proposed AMHFnet consists of four main components: the Feature Processing (FP) module, Multi-Head Residual Encoder (MH), Head Weighting (HW), and Feature Fusion (FF) modules. The FP module is responsible for predicting the final output from the learned feature representations and serves as the key regression component. The MH module constructs multiple feature subspaces from the input representations, while the HW and FF modules apply attention-based weighting and feature fusion on these subspaces to enhance the representation quality. In Table 4, different module combinations were analysed to reveal both individual and synergistic effects. As shown in Table 4, when using only the feature processing (FP) module, which takes raw data as input, the model achieves a mean

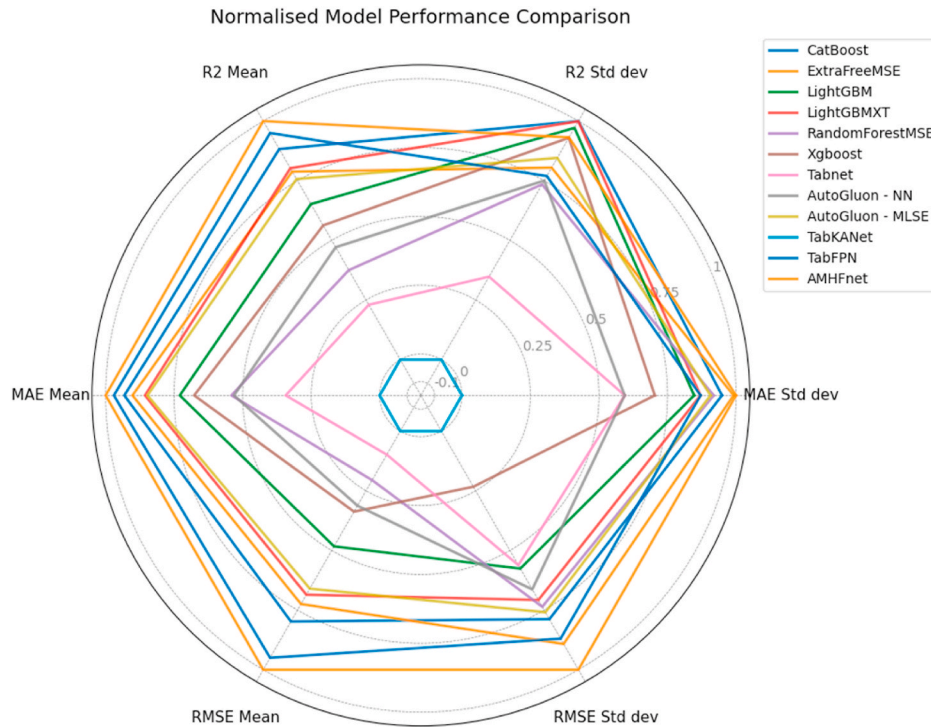


Fig. 5. Radar chart of the performance for evaluation metrics. Higher values indicate better model performance. MAE and RMSE have been reversed and normalised accordingly to ensure a consistent interpretation.

Table 3
The comparison of the model and computational complexity.

Models	RMSE		Model size (M)	FLOPs (MFLOPs)
	Mean	Std dev		
Tabnet (Arik and Pfister, 2021)	19.089	3.339	0.04	0.73
TabKANet (Eslamian et al., 2025)	19.497	5.142	0.05	0.06
TabPFN (Hollmann et al., 2022)	15.553	2.351	60.29	120.59
AMHFnet	15.346	1.934	0.09	1.52

Table 4
The ablation study of the proposed AMHFnet. FP: Feature processing module, MH: multi-head residual encoder, HW: head weighting, FF: Feature fusion. The performance of the model with only the FP module is used as the baseline. The “+” sign denotes the combination.

Models	RMSE		Reduction	
	Mean	Std dev	Mean	Std dev
FP	20.553	5.111	/	/
MH + FP	19.089	3.339	7.12 %	34.67 %
MH + HW + FP	16.188	1.477	21.23 %	71.10 %
MH + FF + FP	15.993	2.014	22.18 %	60.59 %
AMHFnet – single stream	17.799	4.276	/	/
AMHFnet	15.346	1.934	25.34 %	62.17 %

RMSE of 20.553 with a standard deviation of 5.111. The multi-head residual encoder (MH) module enhances the diversity of learned representations, resulting in improved model performance, particularly reducing the standard deviation of RMSE to 3.339. The head weighting (HW) module further incorporates an attention mechanism to improve the performance, which emphasises more important representations. This refinement reduces the mean RMSE to 16.188 and the standard

deviation to 1.477. The reduction of 71.10 % in the standard deviation demonstrates a substantial improvement in model robustness. Similarly, the feature fusion (FF) module fuses the representations by head-wise convolutional operations and increases the number of feature maps to enhance expressiveness. The reduction of the RMSE is 22.18 %, reflecting a significant improvement in the model’s reliability. Overall, the proposed AMHFnet effectively integrates the advantages of each individual module, compared to the FP module, achieving notable improvements in both accuracy (25.34 %) and robustness (62.17 %). Furthermore, an early-fusion version of the AMHFnet model has been tested to justify the effectiveness of separately learning representations from chemical and physical features before fusion. The results show that the early-fusion approach yielded a mean RMSE of 17.799 and a standard deviation of 4.276. Since the physical parameters such as temperature and strain rate are state variables that describe deformation conditions rather than intrinsic material properties. Early fusion of these features with chemical compositions tends to produce a sparse feature space, thereby reducing the efficiency of the representation learning process.

3.3. Predictive performance of the model

Fig. 6 shows the overall performance in RMSE of our model for a wide range of low-alloy steels at different temperatures and strain rates. The residual plot shows statistical performance. The residual plot reveals a higher frequency and less wide distribution of predicted stress close to ± 20 MPa (Lodh et al., 2014) confirming the accuracy.

4. Discussion

4.1. Softening mechanisms in different physical conditions

Typically, hot deformation behaviour shows three distinctive regions based on processing conditions – work hardening, dynamic recovery (DRV) and dynamic recrystallisation (DRX). Work hardening shows a continuous increment in stress with deformation, where a steady state

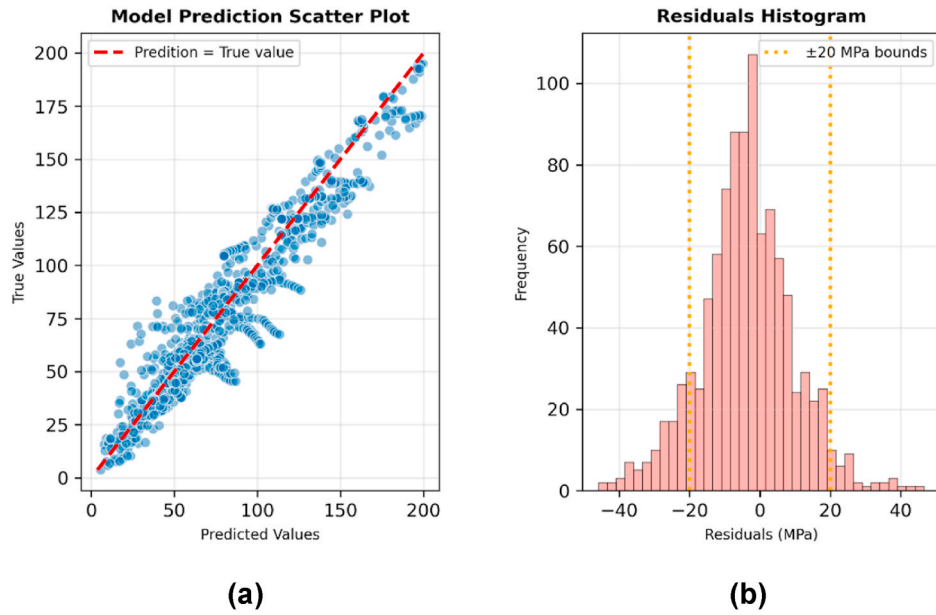


Fig. 6. Fitting performance for AMHFnet. (a) Scatter plot of prediction results. (b) histograms of the residuals.

stress is reached when dynamic recovery occurs. On the other hand, dynamic recrystallisation exhibits a peak stress followed by a decline as newly formed recrystallised grains reduce dislocation density,

eventually leading to a steady-state stress that reflects significant softening behaviour. For dynamic recrystallisation, a critical strain is necessary to initiate the process, and the steady state stress is much

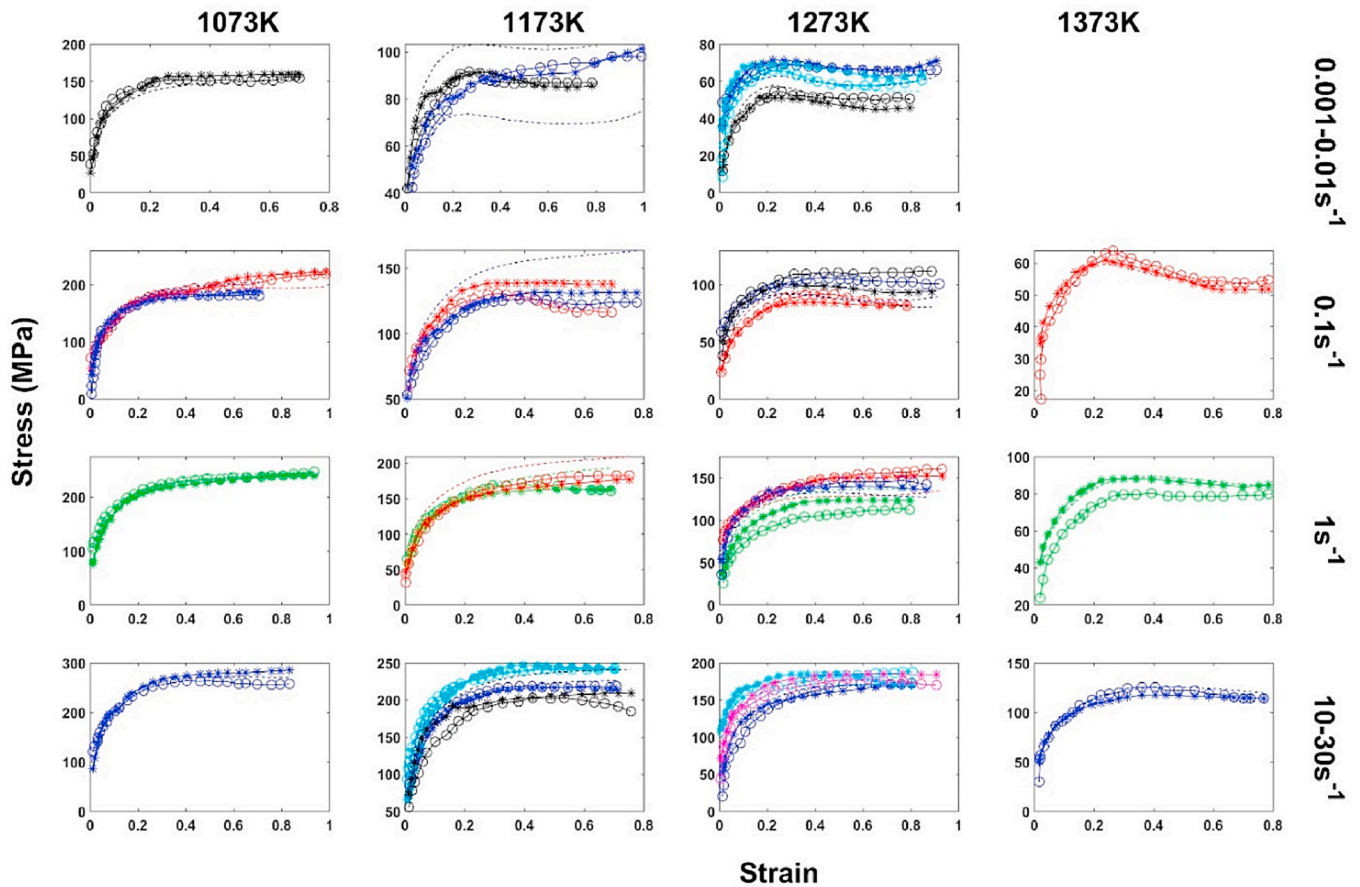


Fig. 7. Predicted flow stress for different steels at various combinations of strain rate and temperatures showing the predictive power of the developed model. For comparison, TabFPN predictions are also included in the plot. Circles represent experimental data from literature, while asterisk (*) symbol and dashed line show predictions from AMHFnet and TabFPN respectively. Different colours represent different steel compositions for a particular condition, shown in the supplementary file. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

lower than dynamic recovery (Humphreys and Hatherly, 2012).

Fig. 7 illustrates the fitting performance of the hot strength behaviour of various steel chemistries across different strain rates and temperatures. For comparison, the predictions of AMHFnet and the next best model, TabPFN, are plotted together with the experimental data from the literature. It can be observed that, in most cases, the predictions of TabPFN are comparable to those of the proposed model and remain close to the true values. However, at 1173 K, noticeable deviations are observed at certain strain rates (e.g., 0.001–0.01, 0.1, and 1 s^{-1}), indicating potential challenges for TabPFN in generalising to unseen materials. The proposed AMHFnet can capture the flow behaviour, particularly at lower (1073 K) and higher (1373 K and above) temperatures for all strain rates. We observed that the flow behaviour of steels, characterised by initial work hardening followed by dynamic recovery, is well-fitted where the stress value increases and holds to a steady state and has no obvious peak value in the curve. Our model captures this occurrence for different steel grades. On the other hand, in the low range of strain rates (0.1 s^{-1}) and the medium and high range of temperature (1173K, 1273K), the predicted behaviour shows a mismatch at the higher strain values for some of the steels. Typically, dynamic recrystallisation is observed in those instances. It can be seen that the accuracy of the model's prediction is high till peak stress (or strain) and drops if the experimental behaviour shows dynamic recrystallisation at intermediate temperatures. In most of the physical conditions, our model achieves good fitting performance in the work hardening stage and dynamic recovery-dominated behaviour. However, at intermediate temperatures where the occurrence of dynamic recrystallisation is more dependent on steel chemistry than physical features, our model cannot capture the behaviour. This limitation comes as DRX is influenced by a coupled effect of processing parameters and alloying elements. For a single steel composition, activation energy remains constant, making the model dependent on processing parameters only, with improved performance. For instance, studies by Medina and Hernandez (1996) and Lin & Chen (Lin and Chen, 2011) demonstrated that activation energy varies widely, from ~ 300 to 700 kJ/mol , depending on alloying additions such as Mn, Si, Cr, and Ni. These variations influence the diffusion of atoms and dislocation movements during high-temperature deformation, leading to the activation of DRX and softening kinetics in flow stress. In other words, steels with similar processing parameters (or physical features) but different compositions may exhibit widely different hot deformation responses due to the effect of alloying elements. Since our model is not physical metallurgy-based, post-dynamic recrystallisation behaviour is not accurately predicted in this range.

4.2. Importance analysis of input features based on SHAP

Fig. 8 shows the SHAP beeswarm plot based on the proposed AMHFnet. It displays an information-dense summary of how the features in the dataset impact the model's output, while the colours refer to the SHAP value of a feature. It shows that strain rate and temperature are the most important features. From the colour scale, it is observed that higher values of strain rate and lower values of temperature impact the flow stress significantly. During hot deformation, dynamic recovery happens while defects such as dislocations rapidly move and rearrange under the applied force, decreasing overall dislocation density. In contrast, dynamic recrystallisation refers to the formation of new, defect-free grains through nucleation and growth under heavy plastic deformation. At higher strain rates or lower temperatures, there is not enough time or activation energy for these processes. Hence, strain hardening dominates, showing an increase in stress as a deformation response, which is reflected in the SHAP importance plot. In the intermediate range, a competition between dynamic recovery and dynamic recrystallisation happens, which is physical metallurgy dependent. Since it is not considered in our model, the SHAP importance in that range is low as well. Also, at low to moderate strains, stress is less likely to be influenced. Overall, the effect of chemical composition on hot

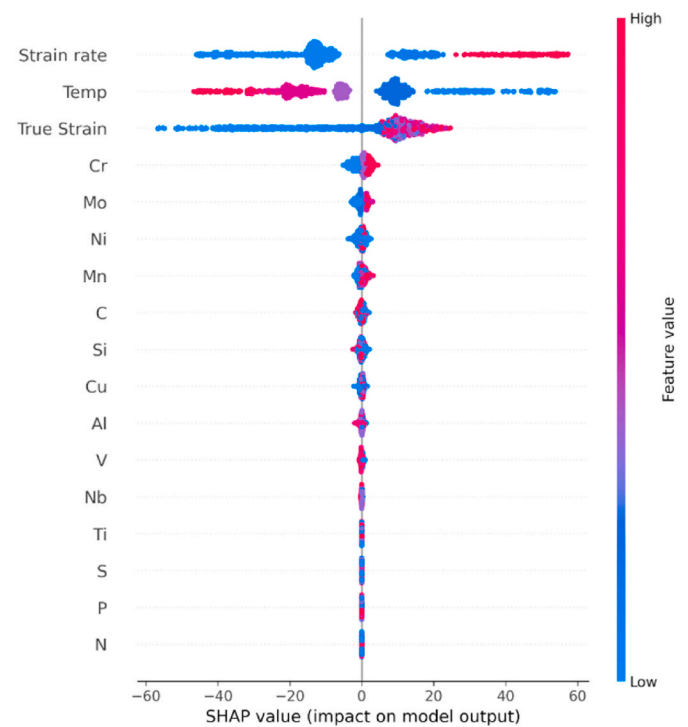


Fig. 8. The impact of each feature on prediction results based on SHAP.

strength can be interpreted from activation energy (or Z). Fig. 8 shows increase in alloying elements has a positive impact on hot strength, while for carbon, a decrease in amount increases the stress. This is consistent with the observation that activation energy is the function of chemical composition, where all elements except carbon have a positive influence (Medina and Hernandez, 1996). Alloying elements like niobium or titanium have a smaller range in our dataset, which restricts in capturing the in-depth impact on stress.

The absolute value of SHAP ignores the direction of impact (positive or negative) and focuses on the magnitude of impact only. Averaging these values across all measurement points out how important each feature is. Fig. 9 indicates that the relative effect of strain rate on the hot strength behaviour is greater than temperature, as temperature dominates the shape of the stress-strain behaviour rather than peak stress.

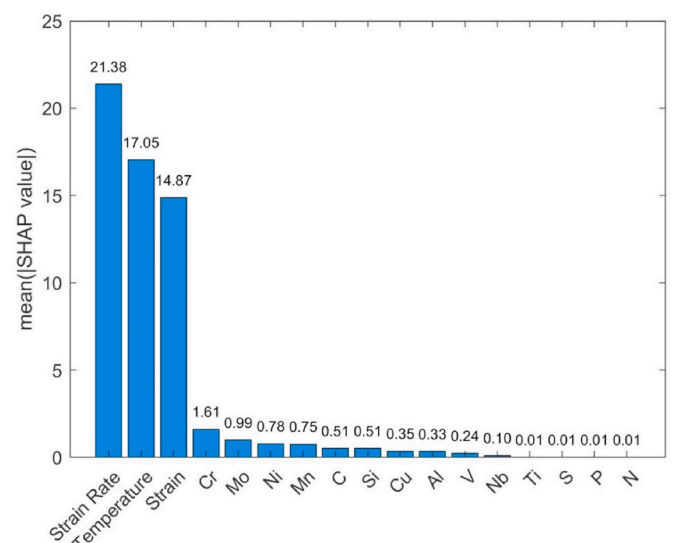


Fig. 9. Global measure of feature importance for the current AMHFnet model.

The effect of true strain and temperature is comparable, as dynamic recrystallisation is favoured. As dynamic recrystallisation produces a mixed microstructure of fine recrystallised grains and large non-recrystallised grains, the overall plasticity response deteriorates.

The effect of alloying elements on hot strength is counterintuitive (Liu et al., 2017). For austenitic stainless steels, artificial neural network modelling revealed the relative importance of different alloying elements (Medina and Hernandez, 1996). Our model considered much higher compositional variation, although the results are comparable. The impact of carbon on the hot strength strongly depends on the deformation condition. Below a critical strain and temperature combination for dynamic recrystallisation, carbon increases the drag tension needed for dislocation migration, promoting the strain hardening. However, as dynamic recrystallisation happens, increasing diffusivity of vacancies lowers the hardening response and the flow stress. From Fig. 9, quantitatively, carbon's impact is approximately three times less than chromium's. Since we considered only the hot deformation temperature regime, where the strengthening effect of carbon as the interstitial element is less pronounced due to higher diffusivity.

Amongst carbide-forming elements (e.g. Cr, Mo, Mn, Nb, V), higher impact on hot strength has been observed from chromium and molybdenum. Typically, precipitation hardening is associated with their size and stability. At higher temperatures ($>700^\circ\text{C}$), thermal stability of vanadium carbide (M_{23}C_6 type) decreases (Liu et al., 2017). The effect of niobium is underestimated as it not only improves precipitation hardening but refines the microstructure of micro-alloyed steels (Tan et al., 1997). Since recrystallised grain size is not considered in our model, the contribution to flow stress from microstructural refinement due to alloying is absent. Nickel, being a strong austenite stabiliser, dissolves into austenite and contributes to solid solution hardening at high temperatures. Manganese acts in a similar way to nickel. Additionally, it suppresses dynamic recrystallisation, leading to higher flow stresses at higher temperatures.

Furthermore, Fig. 10 shows how the SHAP value (model attribution) changes with carbon content and how that relationship interacts with temperature or strain rate (indicated by colour). Between 0 and 0.4 wt% carbon, SHAP values vary widely (-10 to $+5$), indicating a strong but inconsistent effect of carbon on the model output. This suggests that other alloying elements or processing parameters exhibit strong interactions in this range. In contrast, a further increase in carbon shows a narrower cluster in SHAP (near zero or slightly positive), indicating that adding more carbon beyond a threshold does not increase strengthening.

On the other hand, at higher temperatures (shown as red points) SHAP values tend to cluster near zero or positive, implying carbon promotes softening. Therefore, the SHAP dependence analysis reveals that carbon shows strong interaction with both temperature and strain rate. At low strain rates and temperatures, carbon contributes positively to the flow stress through solid solution and carbide strengthening, as seen by high SHAP values. However, its contribution reduces at higher temperatures and strain rates, indicating the dominance of diffusion-controlled recovery and carbide dissolution. These trends together denote the metallurgical interaction between alloying elements, deformation kinetics and thermal activation during hot deformation. This also validates that the model aligns with the physical metallurgy of deformation behaviour during hot deformation.

4.3. Effect of carbon on high temperature flow behaviour

A case study on the effect of carbon has been employed to reflect the model's generalisability, since carbon is the primary alloying element in any steel and has a contradictory effect (Zhang et al., 2021) on high deformation behaviour.

Fig. 11 shows the prediction for two different steels (Tan et al., 1997) at two different temperatures and strain rates. Our model predicts the overall flow behaviour with reasonable accuracy, particularly for higher strain rates and low temperatures. Under these conditions, carbon increases the atomic diffusion resistance of the grain boundaries. Therefore, high-angle grain boundaries are difficult to move, suppressing nucleation of new grains and increasing flow stress (Fig. 11(b)). However, with an increase in temperature, atomic bonds weaken and grain boundary movement is accelerated due to additional activation energy. This minimises the pinning effect of carbon, diminishing the gap in flow stress between 0.055 wt% and 0.37 wt% carbon (Fig. 11(d)). But at lower strain rates, it does not capture the carbon effect, though it predicts the occurrence of dynamic recrystallisation (Fig. 11(a,c)).

Since dynamic recrystallisation is associated with microstructural modification (e.g. grain size) and our model is not physics-informed, we further investigate peak stress rather than the complete flow behaviour. Fig. 12 shows peak stress prediction for plain C-Mn-Si steels at 1373K for different strain rates (Lodh et al., 2014). The prediction has good accuracy (± 20 MPa). Also, the influence of carbon, as shown by the slopes at different strain rates, follows the same trend in predictions and experimental observations and further validates the reliability of the model.

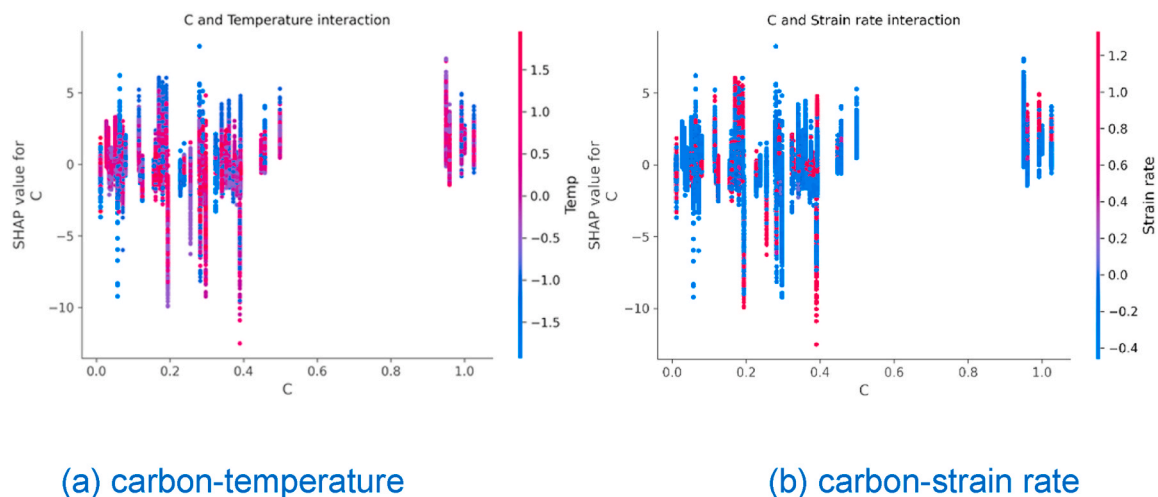


Fig. 10. SHAP dependence plots showing key interaction effects on predicted stress (a) carbon-temperature and (b) carbon-strain rate. In both plots, feature magnitude (carbon content in wt%) is shown on the horizontal axis, and its SHAP value for flow stress prediction is shown on the vertical axis. The colour scale denotes the interacting variable. A positive SHAP value indicates an increase in predicted flow stress. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

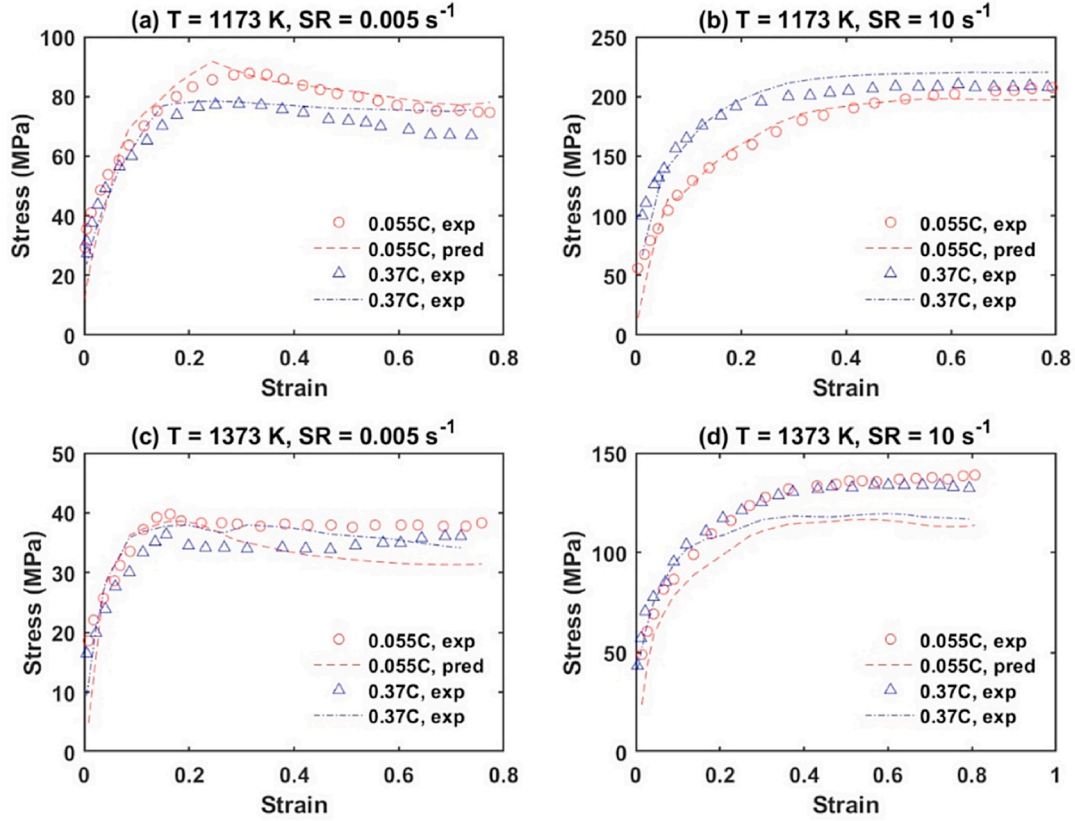


Fig. 11. Effect of Carbon for 0.005 C-1.45 Mn-0.38 Si-0.08 V and 0.37 C-1.45 Mn-0.38 Si-0.08 V steels. T and SR represents temperature and strain rate respectively.

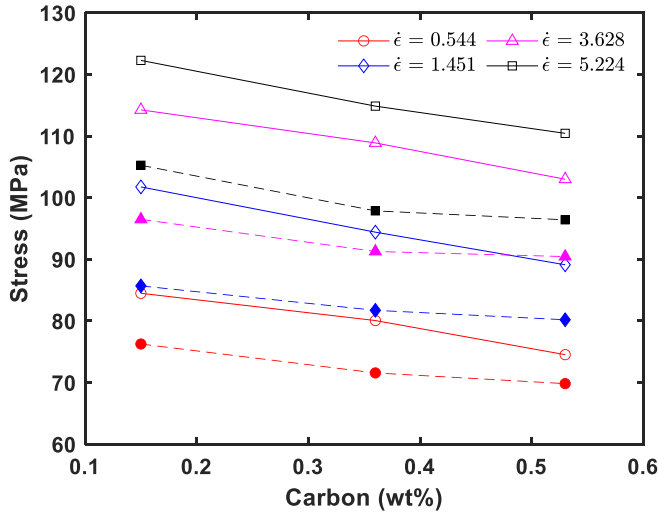


Fig. 12. Effect of Carbon on peak stress for C-Mn-Si steels at different strain rates and 1373K. Open symbols are experimental values, whereas filled-in symbols are predictions.

5. Conclusions

This study proposes a novel AMHFnet capable of predicting the high-temperature deformation behaviour of low-alloy steels. Unlike traditional methods that focus on a specific grade of low-alloy steel, by explicitly including alloying elements, the proposed model quantifies composition-dependent contributions to flow stress, strain hardening, and dynamic softening as functions of temperature and strain rate. The dedicated designed modules provide the capability of capturing the

complex dependencies and enhancing the diversity of the features. By integrating weighted feature maps, the model effectively learns high-dimensional representations of both physical features and chemical compositions, resulting in more accurate predictions of deformation behaviour. A high-temperature deformation dataset comprising 58 low-alloy steels was established for this research. The superiority of the proposed network is demonstrated through comparisons with commonly used and state-of-the-art methods on this dataset. In 10-fold cross-validation, the model exhibits low variance and high prediction accuracy, demonstrating strong robustness and generalisation capabilities across different low-alloy steels. The comprehensive analysis of predicted stress-strain curves for various low-alloy steels shows that the model is able to capture deformation behaviour during work hardening and dynamic recovery. However, it faces challenges in accurately modelling the dynamic recrystallisation process. A SHAP-based feature importance visualisation incorporating elemental characteristics analysis was conducted, which validates the interpretability of the model's predictions. Moreover, the observed dependence between flow stress predictions and alloying elements (carbon) further presented the reliability of the model's decision-making process.

Accurate prediction of hot deformation behaviour provides a powerful tool for the rapid design and optimisation of hot-forming processes in advanced steel manufacturing. In the meantime, experimental results demonstrate that the proposed AMHFnet not only predicts deformation responses across diverse low-alloy steels but also captures the influence of compositional variations on these responses. It offers a robust predictive framework for chemical composition impact on low-alloy steel to guide and target experiments in alloy-composition design, significantly shortening the development cycle. Ablations study shows that learning representations separately for processing variables and alloy chemical composition before fusion improves accuracy and stability over early fusion. The separation reduces feature interference and improves representation efficiency. The findings also highlight that

accurately capturing behaviour during the dynamic recrystallisation process remains a major limitation of the current model. The existing models of hot deformation behaviour typically focus on single materials and can achieve good fits based on physical features or chemical compositions. However, when the same approaches are extended to different low-alloy steels, the limited information from physical and chemical features makes it difficult to fully capture the flow behaviour change in the dynamic recrystallisation phase. As future work, integrating microstructural information, such as recrystallised grain size, crystallographic texture, and recrystallised volume fraction, into the existing framework is expected to enable more accurate and generalisable predictions of hot deformation behaviour in low-alloy steels. Furthermore, as an inherent limitation of data-driven approaches, model performance may degrade under distributional or domain shifts. Expanding the dataset through continuous data acquisition or applying knowledge transfer from related material systems could help mitigate distributional shifts and further improve the prediction of hot deformation behaviour in low-alloy steels in future work.

CRedit authorship contribution statement

Lichao Yang: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Arijit Lodh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Investigation, Formal analysis, Data curation, Conceptualization. **Jiayang Qiu:** Writing – original draft, Visualization, Investigation, Data curation. **Gustavo M. Castelluccio:** Writing – review & editing, Supervision. **Yifan Zhao:** Writing – review & editing, Supervision, Project administration.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.engappai.2025.113464>.

Data availability

The datasets and code generated in this study are available in the GitHub repository, <https://github.com/LY-613/AMHFnet>.

References

- Alcelay, I., Al Omar, A., Prado, J.M., 2014. Characterization of hot flow behaviour and deformation stability of medium carbon microalloyed steel using artificial neural networks and dynamic material model. *Int. J. Mater. Res.* 105 (8), 743–754 [Online]. Available: www.hanser-elibrary.com.
- Arik, S., Pfister, T., 2021. TabNet: attentive interpretable tabular learning. In: 35th AAAI Conference on Artificial Intelligence, AAAI 2021. <https://doi.org/10.1609/aaai.v35i8.16826>.
- Boroumand, Khalil, Hadi, M., Vafaei, R., 2021. Hot deformation behavior and constitutive modelling of a medium-carbon structural steel. *Phys. Met. Metallogr.* 122 (14), 1611–1620. <https://doi.org/10.1134/S0031918X21140052>.
- Chen, Z., Nash, P., 2018. Hot deformation behavior and processing maps for a large marine crankshaft S34MnV steel. *Steel Res. Int.* 89 (3). <https://doi.org/10.1002/srin.201700321>.
- Chen, X., et al., 2019. A constitutive relation of AZ80 magnesium alloy during hot deformation based on arrhenius and johnson-cook model. *J. Mater. Res. Technol.* 8 (2). <https://doi.org/10.1016/j.jmrt.2019.01.003>.
- Churyumov, A.Y., Kazakova, A.A., 2023. Prediction of true stress at hot deformation of high manganese steel by artificial neural network modeling. *Materials* 16 (3). <https://doi.org/10.3390/ma16031083>.
- Churyumov, A., Kazakova, A., Churyumova, T., 2022a. Modelling of the steel high-temperature deformation behaviour using artificial neural network. *Metals* 12 (3). <https://doi.org/10.3390/met12030447>.
- Churyumov, A., Kazakova, A., Churyumova, T., 2022b. Modelling of the steel high-temperature deformation behaviour using artificial neural network. *Metals* 12 (3). <https://doi.org/10.3390/met12030447>.
- Dear, J., Zhang, R., Shi, Z., Lin, J., 2025. A novel data-driven machine learning approach for improved strain rate control in thermomechanical testing of sheet metals. *Eng. Appl. Artif. Intell.* 151 (Jul). <https://doi.org/10.1016/j.engappai.2025.110746>.
- Deng, J., Chen, L.W., Cui, B.Y., Li, Y.P., Ji, H.C., 2024. Arrhenius constitutive model of FV520B steel. *Metalurgija* 63, 93–96.
- Dong, Q., Liu, H., 2019. Deformation behaviors of non-oriented silicon steel plate in hot rolling process. In: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing. <https://doi.org/10.1088/1757-899X/635/1/012027>.
- Dong, J., et al., 2017. Hot deformation behavior and microstructural evolution of Nb-V-Ti microalloyed ultra-high strength steel. *J. Mater. Res.* 32 (19), 3777–3787. <https://doi.org/10.1557/jmr.2017.365>.
- Dou, K., Liu, Q., 2020. Hot deformation behavior and constitutive modelling of low carbon micro-alloyed steel YQ450NQR1 during isothermal compression. *Mech. Mater.* 148 (Sep). <https://doi.org/10.1016/j.mechmat.2020.103430>.
- Duan, X.W., Liu, J.J., Gong, B., Li, P., Liu, J.S., 2021. Experimental study and numerical simulation of dynamic recrystallization behavior of a high-strength steel. *Met. Mater. Int.* 27 (5), 1044–1059. <https://doi.org/10.1007/s12540-019-00433-w>.
- Erickson, N., et al., 2020. AutoGluon-Tabular: robust and accurate AutoML for structured data [Online]. Available: <http://arxiv.org/abs/2003.06505>.
- Eslamian, A., Aghaei, A.A., Cheng, Q., 2025. TabKAN: advancing tabular data analysis using kolmogorov-arnold network [Online]. Available: <http://arxiv.org/abs/2504.06559>.
- feng Dong, H., yong Cai, D., xiang Yang, Q., Zhang, Y., Liao, B., 2009. Hot deformation behavior of a Cr-containing low carbon steel in the ferrite range. *Int. J. Miner. Metall. Mater.* 16 (5), 549–553. [https://doi.org/10.1016/S1674-4799\(09\)60094-7](https://doi.org/10.1016/S1674-4799(09)60094-7).
- Gao, Y.L., Xue, X.X., Yang, H., 2016. Effect of titanium on hot deformation behaviors of boron microalloyed steel. *ISIJ Int.* 56 (4), 619–627. <https://doi.org/10.2355/isijinternational.ISIJINT-2015-605>.
- Ghosh, S., Somani, M.C., Setman, D., Mula, S., 2021. Hot deformation characteristic and strain dependent constitutive flow stress modelling of Ti + Nb stabilized interstitial free steel. *Met. Mater. Int.* 27 (8). <https://doi.org/10.1007/s12540-020-00827-1>.
- Gurusamy, M., Rao, B.C., 2022. A modified Zerilli-Armstrong constitutive model for simulating severe plastic deformation of a steel alloy. *Proc. Inst. Mech. Eng. B J. Eng. Manuf.* 236 (8). <https://doi.org/10.1177/09544054211060914>.
- He, A., Xie, G., Zhang, H., Wang, X., 2014. A modified Zerilli-Armstrong constitutive model to predict hot deformation behavior of 20CrMo alloy steel. *Mater. Des.* 56. <https://doi.org/10.1016/j.matdes.2013.10.080>.
- He, G.N., Wan, S.Q., Jiang, B., Zhang, C.L., Liu, Y.Z., Wu, C.J., 2024. Adverse effect of niobium and boron on hot deformation behavior of sulfur-containing steel. *J. Iron Steel Res. Int.* 31 (1), 252–263. <https://doi.org/10.1007/s42243-023-01002-7>.
- Herian, J., Anioek, K., 2008. The structure and properties of steel with different pearlite morphology and its resistance to abrasive wear. *Reading* 31 (2).
- Hollmann, N., Müller, S., Eggensperger, K., Hutter, F., 2022. TabPFN: a transformer that solves small tabular classification problems in a second [Online]. Available: <http://arxiv.org/abs/2207.01848>.
- Huang, H.Q., et al., 2018. Hot Deformation Behavior and Processing Maps of a High Al-low Si Transformation-Induced Plasticity Steel: microstructural Evolution and Flow Stress Behavior. *Acta Metall. Sin.* 31 (5), 503–514. <https://doi.org/10.1007/s40195-017-0676-2>.
- Huang, M., Jiang, J., Wang, Y., Liu, Y., Zhang, Y., Dong, J., 2023. Comparison and error evaluation of Arrhenius model and typical machine learning algorithms for high-temperature flow stress prediction of GH3536 superalloy. *Mater. Lett.* 349. <https://doi.org/10.1016/j.matlet.2023.134754>.
- Humphreys, F.J., Hatherly, M., 2012. Recrystallization and Related Annealing Phenomena. <https://doi.org/10.1016/C2009-0-07986-0>.
- Jain, S., Jain, R., Rao, K.R., Bhowmik, A., 2024. Leveraging machine learning to minimize experimental trials and predict hot deformation behaviour in dual phase high entropy alloys. *Mater. Today Commun.* 41. <https://doi.org/10.1016/j.mtcomm.2024.110813>.
- Jang, M.H., Kang, J.Y., Jang, J.H., Lee, T.H., Lee, C., 2017. Hot deformation behavior and microstructural evolution of alumina-forming austenitic heat-resistant steels during hot compression. *Mater. Char.* 123, 207–217. <https://doi.org/10.1016/j.matchar.2016.11.038>.
- Ji, G., Li, F., Li, Q., Li, H., Li, Z., 2010. Prediction of the hot deformation behavior for Aermet100 steel using an artificial neural network. *Comput. Mater. Sci.* 48 (3). <https://doi.org/10.1016/j.commatsci.2010.02.031>.
- Jia, D.S., et al., 2022. Hot-compression deformation behavior and constitutive equations of LZ50 axle steel. *Mater. Technol.* 56 (2), 87–95. <https://doi.org/10.17222/mit.2021.331>.
- Jiang, H., et al., 2023. Hot deformation behavior and processing maps of 26CrMo7S steel used in oil exploration. *Materials* 16 (21). <https://doi.org/10.3390/ma16217056>.
- Johnson, G.R., Cook, W.H., 1983. A computational constitutive model and data for metals subjected to large strain, high strain rates and high pressures. *Proceedings 7th International Symposium on Ballistics*.
- Kanrar, B., Mukherjee, S., Das, D., 2023. Flow stress modeling of a new HSLA steel by Zerilli-Armstrong model. In: *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2023.05.671>.
- Kawulok, P., et al., 2022. Hot deformation behavior of non-alloyed carbon steels. *Materials* 15 (2). <https://doi.org/10.3390/ma15020595>.
- Kim, K.S., Du, L.X., sung Choe, H., Lee, T.H., Lee, G.C., 2020. Influence of vanadium content on hot deformation behavior of low-carbon boron microalloyed steel. *Acta Metall. Sin.* 33 (5), 705–715. <https://doi.org/10.1007/s40195-020-01005-3>.

- Kingkam, W., Li, N., Zhang, H.X., Zhao, C.Z., 2017. Hot deformation behavior of high strength low alloy steel by thermo mechanical simulator and finite element method. In: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing. <https://doi.org/10.1088/1757-899X/205/1/012001>.
- Kingkam, W., Li, N., Pang, H., Zhao, C.Z., He-Xin, Z., Li, Z.M., 2020. Constitutive modelling of high strength low alloy steel under hot compression test. In: IOP Conference Series: Materials Science and Engineering. <https://doi.org/10.1088/1757-899X/715/1/012018>.
- Kishor, B., Chaudhari, G.P., Nath, S.K., 2018. Hot workability of 16Cr-5Ni stainless steel using constitutive equation and processing map [Online]. Available: www.science-direct.com/www.materialstoday.com/proceedings2214-7853.
- Kumar, D., Kumar, S., Nath, S.K., 2020. Study on hot deformation behavior of high carbon low alloy steel by constitutive and ann modeling and development of processing maps. *Mater. Perform. Charact.* 9 (2). <https://doi.org/10.1520/MPC20190036>.
- Lan, L., Zhou, W., Misra, R.D.K., 2019. Effect of hot deformation parameters on flow stress and microstructure in a low carbon microalloyed steel. *Mater. Sci. Eng., A* 756, 18–26. <https://doi.org/10.1016/j.msea.2019.04.039>.
- Lee, W.S., Liu, C.Y., 2006. The effects of temperature and strain rate on the dynamic flow behaviour of different steels. *Mater. Sci. Eng., A* 426 (1–2). <https://doi.org/10.1016/j.msea.2006.03.087>.
- Li, H.Y., Wang, X.F., Wei, D.D., Hu, J.D., Li, Y.H., 2012. A comparative study on modified Zerilli-Armstrong, Arrhenius-type and artificial neural network models to predict high-temperature deformation behavior in T24 steel. *Mater. Sci. Eng., A* 536. <https://doi.org/10.1016/j.msea.2011.12.108>.
- Li, H.Y., Li, Y.H., Wang, X.F., Liu, J.J., Wu, Y., 2013. A comparative study on modified Johnson Cook, modified Zerilli-Armstrong and Arrhenius-type constitutive models to predict the hot deformation behavior in 28CrMnMoV steel. *Mater. Des.* 49. <https://doi.org/10.1016/j.matdes.2012.12.083>.
- Li, N., Zhao, C., Jiang, Z., Zhang, H., 2019. Flow behavior and processing maps of high-strength low-alloy steel during hot compression. *Mater. Char.* 153, 224–233. <https://doi.org/10.1016/j.matchar.2019.05.009>.
- Li, C., Huang, L., Zhao, M., Zhang, X., Li, J., Li, P., 2020a. Influence of hot deformation on dynamic recrystallization behavior of 300M steel: rules and modeling. *Mater. Sci. Eng., A* 797. <https://doi.org/10.1016/j.msea.2020.139925>.
- Li, M., Duan, Y., Yao, D., Guan, Y., Yang, L., 2020b. Constitutive model and flow behavior of B1500hs high-strength steel during the hot deformation process. *Metals* 10 (1). <https://doi.org/10.3390/met10010064>.
- Li, R., et al., 2020c. Hot deformation behavior and processing maps of a 9Ni590B steel. *J. Mater. Eng. Perform.* 29 (6), 3858–3867. <https://doi.org/10.1007/s11665-020-04907-6>.
- Li, F., et al., 2022. A comparative study on modified and optimized Zerilli-Armstrong and arrhenius-type constitutive models to predict the hot deformation behavior in 30Si2MnCrMoVE steel. *J. Mater. Res. Technol.* 20, 3918–3929. <https://doi.org/10.1016/j.jmrt.2022.08.134>.
- Lian Wei, H., quan Liu, G., tao Zhao, H., mu Kang, R., 2013. Hot deformation behavior of two C-Mn-Si based and C-Mn-Al based microalloyed high-strength steels: a comparative study. *Mater. Des.* 50, 484–490. <https://doi.org/10.1016/j.matdes.2013.03.043>.
- Lian Wei, H., ju Deng, X., shuang Chen, S., you Zhou, Z., bo Pan, H., 2023. Study on hot flow behaviors and Deformation/diffusion mechanisms of V and V-Ti microalloyed steels by physical constitutive modeling. *Steel Res. Int.* 94 (7). <https://doi.org/10.1002/srin.202200847>.
- Lin, Y.C., Chen, X.M., 2010. A combined Johnson-Cook and Zerilli-Armstrong model for hot compressed typical high-strength alloy steel. *Comput. Mater. Sci.* 49 (3), 628–633. <https://doi.org/10.1016/j.commatsci.2010.06.004>.
- Lin, Y.C., Chen, X.M., 2011. A Critical Review of Experimental Results and Constitutive Descriptions for Metals and Alloys in Hot Working. <https://doi.org/10.1016/j.matdes.2010.11.048>.
- Lin, Y.C., Liu, G., 2010. A new mathematical model for predicting flow stress of typical high-strength alloy steel at elevated high temperature. *Comput. Mater. Sci.* 48 (1). <https://doi.org/10.1016/j.commatsci.2009.06.026>.
- Lin, Y.C., Chen, M.S., Zhang, J., 2009. Modeling of flow stress of 42CrMo steel under hot compression. *Mater. Sci. Eng., A* 499 (1–2), 88–92. <https://doi.org/10.1016/j.msea.2007.11.119>.
- Liu, Y., Zhu, J.C., Cao, Y., 2017. Modeling effects of alloying elements and heat treatment parameters on mechanical properties of hot die steel with back-propagation artificial neural network. *J. Iron Steel Res. Int.* 24 (12). [https://doi.org/10.1016/S1006-706X\(18\)30025-6](https://doi.org/10.1016/S1006-706X(18)30025-6).
- Liu, M., Ma, Q.X., Bin Luo, J., 2018. Microstructure evolution and flow stress model of a 20Mn5 hollow steel ingot during hot compression. *Materials* 11 (4). <https://doi.org/10.3390/ma11040463>.
- Lodh, A., Biswas, A., Das, S., 2014. Modelling hot strength behaviour of steel. *Ironmak. Steelmak.* 42 (4). <https://doi.org/10.1179/1743281214Y.00000000230>.
- Lu, Q., et al., 2021. Revolutionizing car body manufacturing using a unified steel metallurgy concept. *Sci. Adv.* 7 (49). <https://doi.org/10.1126/sciadv.abk0176>.
- Lundberg, S.M., Erion, G.G., Lee, S.-I., 2019. Consistent individualized feature attribution for tree ensembles. In: KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631 [Online]. Available: <http://arxiv.org/abs/1802.03888>.
- Mandal, S., Sivaprasad, P.V., Venugopal, S., Murthy, K.P.N., 2006. Constitutive flow behaviour of austenitic stainless steels under hot deformation: artificial neural network modelling to understand, evaluate and predict. *Model. Simulat. Mater. Sci. Eng.* 14 (6). <https://doi.org/10.1088/0965-0393/14/6/012>.
- Mandal, S., Sivaprasad, P.V., Venugopal, S., Murthy, K.P.N., 2009. Artificial neural network modeling to evaluate and predict the deformation behavior of stainless steel type AISI 304L during hot torsion. *Appl. Soft Comput. J.* 9 (1). <https://doi.org/10.1016/j.asoc.2008.03.016>.
- Medina, S.F., Hernandez, C.A., 1996. General expression of the Zener-Hollomon parameter as a function of the chemical composition of low alloy and microalloyed steels. *Acta Mater.* 44 (1). [https://doi.org/10.1016/1359-6454\(95\)00151-0](https://doi.org/10.1016/1359-6454(95)00151-0).
- Menapace, C., Sartori, N., Pellizzari, M., Straffellini, G., 2018. Hot deformation behavior of four steels: a comparative study. *J. Eng. Mater. Technol., Transact. ASME* 140 (2). <https://doi.org/10.1115/1.4038670>.
- Mirzaie, T., Mirzadeh, H., Cabrera, J.M., 2016. A simple Zerilli-Armstrong constitutive equation for modeling and prediction of hot deformation flow stress of steels. *Mech. Mater.* 94. <https://doi.org/10.1016/j.mechmat.2015.11.013>.
- Molinari, A., Ravichandran, G., 2005. Constitutive modeling of high-strain-rate deformation in metals based on the evolution of an effective microstructural length. *Mech. Mater.* 37 (7). <https://doi.org/10.1016/j.mechmat.2004.07.005>.
- Murugesan, M., et al., 2023. Supervised machine learning approach for modeling hot deformation behavior of medium carbon steel. *Steel Res. Int.* 94 (2). <https://doi.org/10.1002/srin.202200188>.
- Opéla, P., Schindler, I., Kawulok, P., Kawulok, R., Rusz, S., Sauer, M., 2022. Shallow and deep learning of an artificial neural network model describing a hot flow stress evolution: a comparative study. *Mater. Des.* 220 (Aug). <https://doi.org/10.1016/j.matdes.2022.110880>.
- ping Xi, S., et al., 2022. Hot deformation behavior and processing map of low-alloy offshore steel. *J. Iron Steel Res. Int.* 29 (3), 474–483. <https://doi.org/10.1007/s42243-021-00603-4>.
- Qi, H., Li, Y., Fu, J., Liu, Z., 2011. Flow stress behavior of the normalized and tempered as-cast 42CrMo steel during hot deformation. *Adv. Mater. Res.* 2560–2564. <http://doi.org/10.4028/www.scientific.net/AMR.314-316.2560>.
- Qian, D., Peng, Y., 2015. Mathematical modeling for microstructural evolution in multipass hot compression of Q345E alloy steel. *J. Mater. Eng. Perform.* 24 (5), 1906–1917. <https://doi.org/10.1007/s11665-015-1473-6>.
- Qiao, L., Deng, Y., Wang, Y., Zhu, J., 2020. A comparative study on arrhenius equations and BP neural network models to predict hot deformation behaviors of a hypereutectoid steel. *IEEE Access* 8, 68083–68090. <https://doi.org/10.1109/ACCESS.2020.2986389>.
- Rakhshkhorshid, M., Hashemi, S.H., 2013. Experimental study of hot deformation behavior in API X65 steel. *Mater. Sci. Eng., A* 573, 37–44. <https://doi.org/10.1016/j.msea.2013.02.045>.
- Ren, F., Li, B., Tang, X., 2018. Hot deformation behavior of 2Cr13 stainless steel. In: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing. <https://doi.org/10.1088/1757-899X/394/3/032048>.
- Sajadifar, S.V., Ketabchi, M., Bemanizadeh, B., 2012. Dynamic recrystallization behavior and hot deformation characteristics in 4340 steel. *Metallurgist* 56, 310–320.
- Sasaki, S., Katsumura, T., 2019. Influence of thermal history before hot working on flow stress of multiphase steel. In: METAL 2019 - 28th International Conference on Metallurgy and Materials, Conference Proceedings. TANGER Ltd., pp. 248–253. <https://doi.org/10.37904/metal.2019.697>.
- Savaedi, Z., Motalebi, R., Mirzadeh, H., 2022. A Review of Hot Deformation Behavior and Constitutive Models to Predict Flow Stress of high-entropy Alloys. <https://doi.org/10.1016/j.jallcom.2022.163964>.
- Singh, K., Rajput, S.K., Mehta, Y., 2016. Modeling of the hot deformation behavior of a high phosphorus steel using artificial neural networks. *Materials Discovery* 6, 1–8. <https://doi.org/10.1016/j.md.2017.03.001>.
- Tan, K.E., et al., 1997. Modelling and simulation in materials science and engineering. *Model. Simulat. Mater. Sci. Eng.* 5 (1).
- Thakur, S.K., Harish, L., Das, A.K., Rath, S., Pathak, P., Jha, B.K., 2020. Hot deformation behavior and processing map of Nb-V-Ti micro-alloyed steel. In: Materials Today: Proceedings. Elsevier Ltd, pp. 1973–1979. <https://doi.org/10.1016/j.matpr.2020.05.566>.
- Thakur, S.K., Kumar Das, A., Jha, B.K., 2023. Application of machine learning methods for the prediction of roll force and torque during plate rolling of micro-alloyed steel. *J. Alloys and Metallurgical Syst.* 4. <https://doi.org/10.1016/j.jalmes.2023.100044>.
- Tomita, Y., 2000. Development of fracture toughness of ultrahigh strength, medium carbon, low alloy steels for aerospace applications. *Int. Mater. Rev.* 45 (1). <https://doi.org/10.1179/095066000771048791>.
- Wang, Z., Liu, X., Xie, F., Lai, C., Li, H., Zhang, Q., 2018. Dynamic recrystallization behavior and critical strain of 51CrV4 high-strength spring steel during hot deformation. *JOM* 70 (10), 2385–2391. <https://doi.org/10.1007/s11837-018-3054-2>.
- Wen, E., et al., 2015. Hot deformation behavior of NM550 wear-resistant steel. In: HSLA Steels 2015. Microalloying 2015 & Offshore Engineering Steels 2015, pp. 873–879.
- Xiang, Y., Qian, P., Gu, X., Wang, H., 2023. Effect of strain on hot deformation behavior and 3-D processing maps of a vanadium and titanium microalloyed steel. *J. Mater. Eng. Perform.* 32 (4), 1747–1759. <https://doi.org/10.1007/s11665-022-07229-x>.
- Xu, L., yang Wang, C., quan Liu, G., zhe Bai, B., 2009. Hot deformation behavior of medium carbon V-N microalloyed steel. *Trans. Nonferrous Metals Soc. China* 19 (6), 1389–1394. [https://doi.org/10.1016/S1003-6326\(09\)60038-8](https://doi.org/10.1016/S1003-6326(09)60038-8).
- Xu, J., Gao, X., Jiang, Z., Wei, D., 2016. A comparison of hot deformation behavior of High-Cr white cast iron and High-Cr white cast iron/low carbon steel laminate. *Steel Res. Int.* 87 (6), 780–788. <https://doi.org/10.1002/srin.201500234>.
- Xu, L., Wang, Y., Mo, L., Tang, Y., Wang, F., Li, C., 2023. The Research Progress and Prospect of Data Mining Methods on Corrosion Prediction of Oil and Gas Pipelines. <https://doi.org/10.1016/j.engfailanal.2022.106951>.
- Xu, L., Wen, S., Huang, H., Tang, Y., Wang, Y., Pan, C., 2025. Corrosion failure prediction in natural gas pipelines using an interpretable XGBoost model: insights and applications. *Energy* 325 (Jun). <https://doi.org/10.1016/j.energy.2025.136157>.

- Yang, Z., Zhang, F., Zheng, C., Zhang, M., Lv, B., Qu, L., 2015a. Study on hot deformation behaviour and processing maps of low carbon bainitic steel. *Mater. Des.* 66 (PA). <https://doi.org/10.1016/j.matdes.2014.10.068>.
- Yang, Z., Zhang, F., Zheng, C., Zhang, M., Lv, B., Qu, L., 2015b. Study on hot deformation behaviour and processing maps of low carbon bainitic steel. *Mater. Des.* 66 (PA), 258–266. <https://doi.org/10.1016/j.matdes.2014.10.068>.
- Yang, Z.N., Dai, L.Q., Chu, C.H., Zhang, F.C., Wang, L.W., Xiao, A.P., 2017. Effect of aluminum alloying on the hot deformation behavior of nano-bainite bearing steel. *J. Mater. Eng. Perform.* 26 (12), 5954–5962. <https://doi.org/10.1007/s11665-017-3018-7>.
- Ye, L., Zhai, Y., Zhou, L., Wang, H., Jiang, P., 2020. The hot deformation behavior and 3D processing maps of 25Cr2Ni4MoV steel for a super-large nuclear-power rotor. *J. Manuf. Process.* 59, 535–544. <https://doi.org/10.1016/j.jmapro.2020.09.062>.
- Yin, F., Hua, L., Mao, H., Han, X., 2013. Constitutive modeling for flow behavior of GCr15 steel under hot compression experiments. *Mater. Des.* 43, 393–401. <https://doi.org/10.1016/j.matdes.2012.07.009>.
- Zhang, C., Zhang, L., Shen, W., Liu, C., Xia, Y., Li, R., 2016. Study on constitutive modeling and processing maps for hot deformation of medium carbon cr-ni-mo alloyed steel. *Mater. Des.* 90, 804–814. <https://doi.org/10.1016/j.matdes.2015.11.036>.
- Zhang, Z., et al., 2021. Effects of vanadium content on the carbides transformation and strengthening mechanism of MPS700V hot-work die steel at room and elevated temperatures. *Mater. Sci. Eng., A* 813. <https://doi.org/10.1016/j.msea.2021.141091>.
- Zhang, K., et al., 2023a. Hot deformation behavior, dynamic recrystallization mechanism and processing maps of Ti–V microalloyed high strength steel. *J. Mater. Res. Technol.* 25. <https://doi.org/10.1016/j.jmrt.2023.06.195>.
- Zhang, K., et al., 2023b. Hot deformation behavior, dynamic recrystallization mechanism and processing maps of Ti–V microalloyed high strength steel. *J. Mater. Res. Technol.* 25, 4201–4215. <https://doi.org/10.1016/j.jmrt.2023.06.195>.
- Zhang, R., Qi, Y., Kong, S., Wang, X., Li, M., 2024. A hybrid artificial intelligence algorithm for fault diagnosis of hot rolled strip crown imbalance. *Eng. Appl. Artif. Intell.* 130. <https://doi.org/10.1016/j.engappai.2023.107763>.
- Zhao, H., Qi, J., Liu, G., Su, R., Sun, Z., 2020. A comparative study on hot deformation behaviours of low-carbon and medium-carbon vanadium microalloyed steels. *J. Mater. Res. Technol.* 9 (5), 11319–11331. <https://doi.org/10.1016/j.jmrt.2020.08.016>.
- Zhao, C., et al., 2023a. Hot Deformation Behavior of a Novel Cr–Mo–V Hot Work Die Steel. *Steel Res. Int.* 94 (8). <https://doi.org/10.1002/srin.202200698>.
- Zhao, M., Xu, Y., Guo, J., Liao, X., Zhang, D., Han, P., 2023b. Hot deformation behavior and processing map of Cu-Bearing ship plate steels. *J. Mater. Eng. Perform.* 32 (21), 9918–9925. <https://doi.org/10.1007/s11665-023-07803-x>.
- Zhao, C., et al., 2023c. Hot Deformation Behavior of a Novel Cr–Mo–V Hot Work Die Steel. *Steel Res. Int.* 94 (8). <https://doi.org/10.1002/srin.202200698>.



DSPACE

<https://dspace.org/>

Attention-based multi-head feature-fusion network: a generalised method for hot deformation behaviour prediction in low-alloy steels

Yang, Lichao; Lodh, Arijit; Qiu, Jiayang; Castelluccio, Gustavo M.; Zhao, Yifan

2026-02-01

Elsevier

<https://doi.org/10.1016/j.engappai.2025.113464>

Yang L, Lodh A, Qiu J, et al., (2026) Attention-based multi-head feature-fusion network: a generalised method for hot deformation behaviour prediction in low-alloy steels. Engineering Applications of Artificial Intelligence, Volume 165, Part B, February 2026, Article number 113464