



A mechanics-informed deep learning constitutive model for sequential prediction of strain rate-dependent behavior and microstructural evolution

Wei Dai^{1,2}, Xiaochuan Sun³, Kecheng Zhou⁴, Yongkai Wang⁵, Huamiao Wang⁶, Yue Liu³

Keywords:

Deep learning constitutive model, strain rate sensitivity, mechanics-informed design, multi-element alloys

Citation:

Dai, W.; Sun, X.; Zhou, K.; Wang, Y.; Wang, H.; Liu, Y. A mechanics-informed deep learning constitutive model for sequential prediction of strain rate-dependent behavior and microstructural evolution. *Microstructures* 2026, 6, 2026081. <https://dx.doi.org/10.20517/microstructures.2026.17>

Received: 31 Jan 2026

First Decision: 3 Mar 2026

Revised: 13 Apr 2026

Accepted: 24 Apr 2026

Published: 12 Jun 2026

Academic Editor:

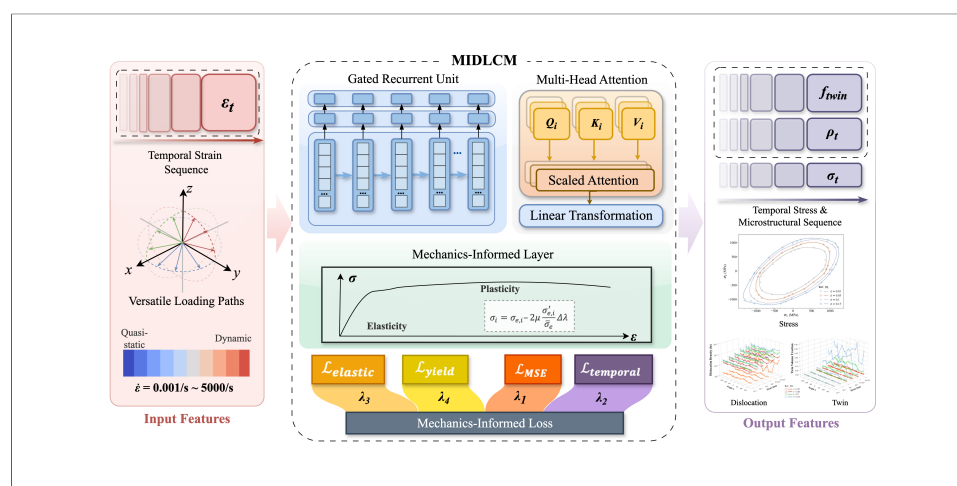
Liangzhi Kou

Copy Editor:

Shu-Yuan Duan

Production Editor:

Shu-Yuan Duan



Abstract

Classical constitutive models explicitly couple macroscopic mechanical responses with underlying microstructural evolution, which is crucial for capturing complex deformation mechanisms across varying strain rates. However, current deep learning (DL) constitutive models predominantly focus on macroscopic stress-strain mapping, often neglecting these critical microstructural transitions. To bridge this gap, this work proposes a mechanics-informed deep learning constitutive model (MIDLCM) that integrates gated recurrent units and multi-head attention with a mechanics-informed layer and a mechanics-informed loss, enabling simultaneous prediction of stress response and microstructural descriptors. Trained on a CrFeNi FCC alloy dataset spanning strain rates from 10^{-4} to $5,000 \text{ s}^{-1}$, MIDLCM accurately reproduces strain-rate-dependent stress-strain behavior and captures the associated evolution of dislocation density and twin volume

¹State Key Laboratory of Deep-sea Manned Vehicles, China Ship Scientific Research Center, Wuxi 214082, Jiangsu, China.

²Taihu Laboratory of Deepsea Technological Science, Wuxi 214082, Jiangsu, China.

³State Key Lab of Metal Matrix Composites, School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China.

⁴State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China.

⁵Shanghai Key Laboratory of Digital Manufacture for Thin-walled Structures, Shanghai Jiao Tong University, Shanghai 200240, China.

⁶Shanghai Institute of Applied Mathematics and Mechanics, School of Mechanics and Engineering Science, Shanghai University, Shanghai 200072, China.

Correspondence to: Dr. Xiaochuan Sun, Prof. Yue Liu, State Key Lab of Metal Matrix Composites, School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. E-mail: garaireen@sjtu.edu.cn; yliu23@sjtu.edu.cn

fraction. Crucially, the model successfully represents the distinct dislocation accumulation regimes and the dynamic transition of plasticity mechanisms - from dislocation-dominated to twinning-assisted - across extreme dynamic loading, consistent with experimental trends and crystal-plasticity-based references. Ablation studies show that attention-based temporal encoding and mechanics-informed constraints contribute complementary improvements while preserving inference efficiency. By explicitly tracking these internal state variables, the proposed framework provides a mechanism-level interpretable and computationally efficient microstructure-mechanics coupled alternative for rate-dependent constitutive modeling and is readily extendable to other alloy systems and loading paths.

INTRODUCTION

Mechanical constitutive models establish fundamental relationships between loading conditions and material responses^[1]. In metallurgy, conventional phenomenological models typically describe stress-strain behavior, whereas physics-based formulations such as crystal plasticity (CP) explicitly connect macroscopic responses to microstructural evolution, proving essential for both fundamental research and industrial applications. Recently, deep learning (DL) constitutive models (or data-driven constitutive models in some works) have emerged, leveraging large datasets to predict mechanical responses through efficient inference^[2,3].

The proliferation of advanced materials, including multi-principal element alloys (MEAs), functionally graded materials, and nanostructured alloys^[4], demands constitutive frameworks that capture microstructure-property relationships with both theoretical rigor and interpretability^[5,6]. Industrial processes such as high-speed machining, additive manufacturing, and cold spray impose strain rates exceeding 10^4 s^{-1} , driving dynamic deformation behaviors fundamentally different from quasi-static responses^[7]. These dynamic responses are intrinsically coupled to microstructural evolution, including dislocation and twinning activities, necessitating constitutive models capable of representing complex strain-rate-dependent behavior across broad deformation ranges^[8-10].

Existing constitutive modeling paradigms face distinct challenges in addressing these requirements. Conventional physics-based models often require complex formulations to capture microstructure-linked, strain-rate-sensitive behavior, potentially impairing solver convergence and substantially increasing computational costs^[11,12]. DL approaches using recurrent neural networks (RNNs) enable efficient sequence-to-sequence inference, as illustrated in [Figure 1A](#), but lack interpretability in connecting microstructure to macroscopic properties, particularly regarding strain-rate sensitivity. This limitation constrains their deployment in complex material systems such as MEAs^[13,14]. Furthermore, standard RNN architectures like gated recurrent units (GRUs)^[15], despite their enhanced memory capabilities for complex deformation histories^[16,17], can encounter vanishing-gradient issues with long sequences and high-dimensional inputs. This may result in inadequate capture of cooperative interactions between multiple deformation mechanisms that govern complex mechanical responses.

To address the aforementioned challenges, we propose integrating the advantages of different constitutive modeling paradigms through a mechanics-informed deep learning framework. The proposed framework employs an RNN structure to efficiently map loading sequences to sequential mechanical responses while capturing latent patterns that serve as inherent state variables. Specifically, GRUs are adopted as the backbone architecture due to their enhanced long-term memory capabilities under complex deformation histories. Recent applications in constitutive modeling have demonstrated GRUs' effectiveness in capturing path-dependent material behavior^[18,19].

However, GRUs may encounter vanishing-gradient issues when processing extended loading histories and high-dimensional inputs, potentially causing information loss from early loading stages^[14,20]. This limitation

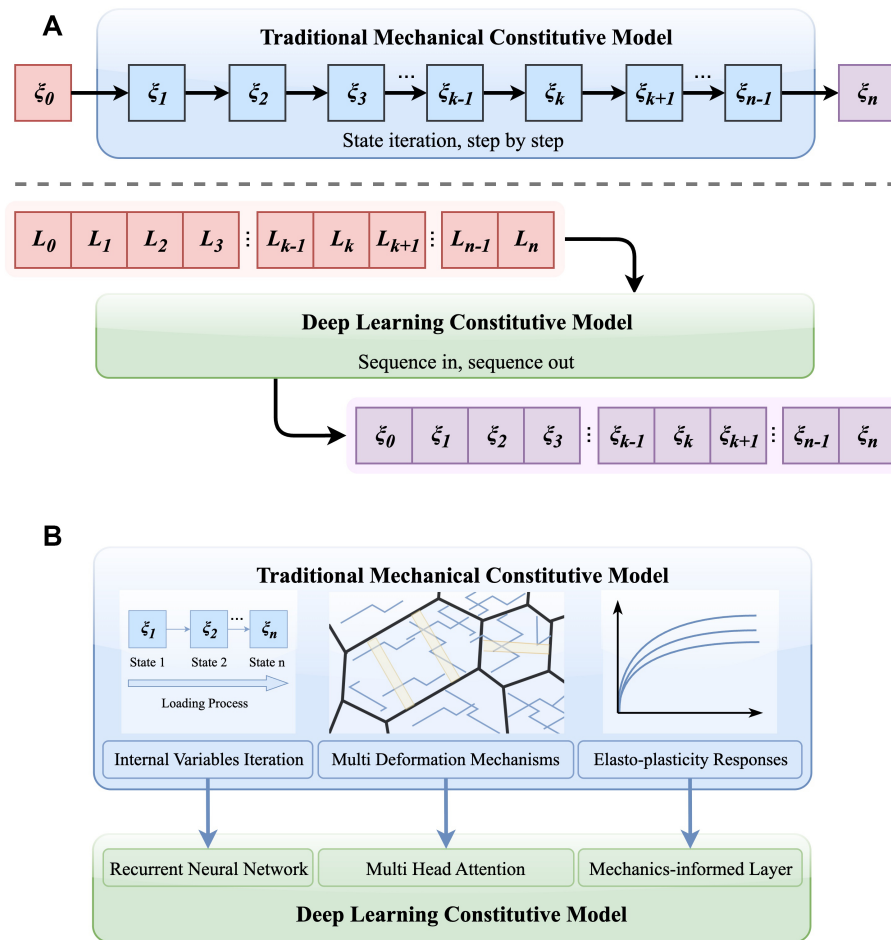


Figure 1. (A) The concept difference between the traditional constitutive model and the DL constitutive model; (B) The design of the DL constitutive model with a conventional modeling insight. DL: Deep learning.

may compromise the model's ability to capture the cooperative interactions among multiple deformation mechanisms that govern complex mechanical responses. To represent this interplay, multi-head attention (MHA) is introduced that simultaneously processes features from different representation subspaces, thereby capturing diverse relational patterns in parallel^[21]. The MHA mechanism functions as a robust temporal encoder, enabling long-range dependency modeling and processing information across multiple scales to manage concurrent deformation mechanisms. Furthermore, cross-attention within the MHA layer efficiently couples the complex loading history with microstructural variables. Ultimately, by explicitly predicting these internal microstructural descriptors alongside the macroscopic stress, the proposed framework can achieve enhanced mechanism-level interpretability.

To further overcome the 'black-box' nature of purely data-driven models, recent studies have increasingly explored mechanics-constrained learning strategies^[22-24], demonstrating immense potential in complex engineering materials - such as enforcing monotonic damage evolution in high-strength concrete^[25]. Inspired by these advancements, our framework integrates deep learning with conventional constitutive modeling insights. While not a physics-informed neural network (PINN) in the strict PDE-residual sense, this framework serves as a rigorous mechanics-informed surrogate. As illustrated in [Figure 1B](#), the proposed framework integrates: (i) a GRU layer for history-dependent behavior representation, (ii) an MHA encoding cooperative deformation mechanisms, and (iii) a mechanics-informed layer with specialized loss functions constraining elasto-viscoplastic behavior. The model inputs comprise deformation history and strain rate,

while outputs include stress response and microstructural descriptors quantifying dislocation and twinning activity. Unlike existing GRU-based surrogates that primarily map macroscopic responses, this explicit microstructure-mechanics coupling enables the present framework to capture distinct rate-dependent dislocation evolutions and the dynamic slip-twinning competition. Furthermore, the differentiable mechanics-informed layer provides a universal elastoplastic baseline, anchoring the data-driven residual learning in fundamental physical principles. To rigorously validate the framework, a dataset of the complex FCC CrFeNi medium-entropy alloy serves as the material testbed. This alloy was selected because its pronounced rate-dependent responses are governed by the dynamic slip-twinning competition across strain rates from 10^{-4} to $5,000 \text{ s}^{-1}$ ^[26]. Accurately predicting this alloy's microstructural evolution demonstrates the framework's generalizable capability in handling complex multi-mechanism constitutive behaviors.

The remainder of this paper is structured as follows. Section **MATERIALS AND METHODS** presents the architecture and implementation of the proposed deep learning constitutive model. This section then details the construction of a comprehensive CrFeNi mechanical response dataset incorporating strain rate sensitivity effects. Section **RESULTS AND DISCUSSION** evaluates model performance through comprehensive validation, including ablation studies and comparison with experimental data, while examining the accuracy of predicted strain rate-dependent mechanical behavior. This section further elucidates the underlying mechanisms by which the MHA and mechanics-informed layers effectively capture complex deformation phenomena. The concluding remarks are provided as the last segment of this paper.

MATERIALS AND METHODS

The mechanics-informed deep learning constitutive model

The mechanics-informed deep learning constitutive model (MIDLCM) architecture, illustrated in [Figure 2](#), comprises five key components: GRUs, an embedding layer, MHA, feed-forward networks (FFN), and a mechanics-informed layer. This end-to-end framework maps input strain histories (ϵ_t) and strain rates ($\dot{\epsilon}$) to predictions of stress evolution (σ_t), dislocation densities (ρ), and twin volume fractions (TVFs) (f_{twin}). The model is optimized through a mechanics-informed loss function that enforces fundamental plasticity principles, thereby enhancing physical consistency and predictive accuracy beyond conventional data-driven methodologies.

MIDLCM operates through three stages: feature preprocessing, feature interaction, and predictive output. The scalar strain rate ($\dot{\epsilon}$) undergoes dimensional expansion through an embedding layer, transforming it into a high-dimensional representation that captures rate-dependent effects more effectively than direct input. This embedded strain rate is concatenated with the strain history along the feature dimension, enabling the model to extract higher-order interactions between deformation rates and strain states. This approach is particularly crucial for accurately modeling rate-dependent phenomena such as viscoplastic responses in medium-entropy alloys.

The feature interaction stage exploits the complementary capabilities of GRU and MHA mechanisms to capture both long-term temporal dependencies and complex inter-feature relationships. The GRU layers maintain memory of deformation history, while the MHA mechanism models cooperative interactions among multiple deformation mechanisms throughout the loading process. The synergistic combination enables comprehensive representation of path-dependent material behavior across extended deformation sequences.

In the MIDLCM architecture, high-dimensional latent features are initially mapped to microstructural descriptors [dislocation density (DD) and TVF] via dedicated FFNs. Subsequently, these microstructural predictions are concatenated with strain components and processed through additional GRU and FFN layers

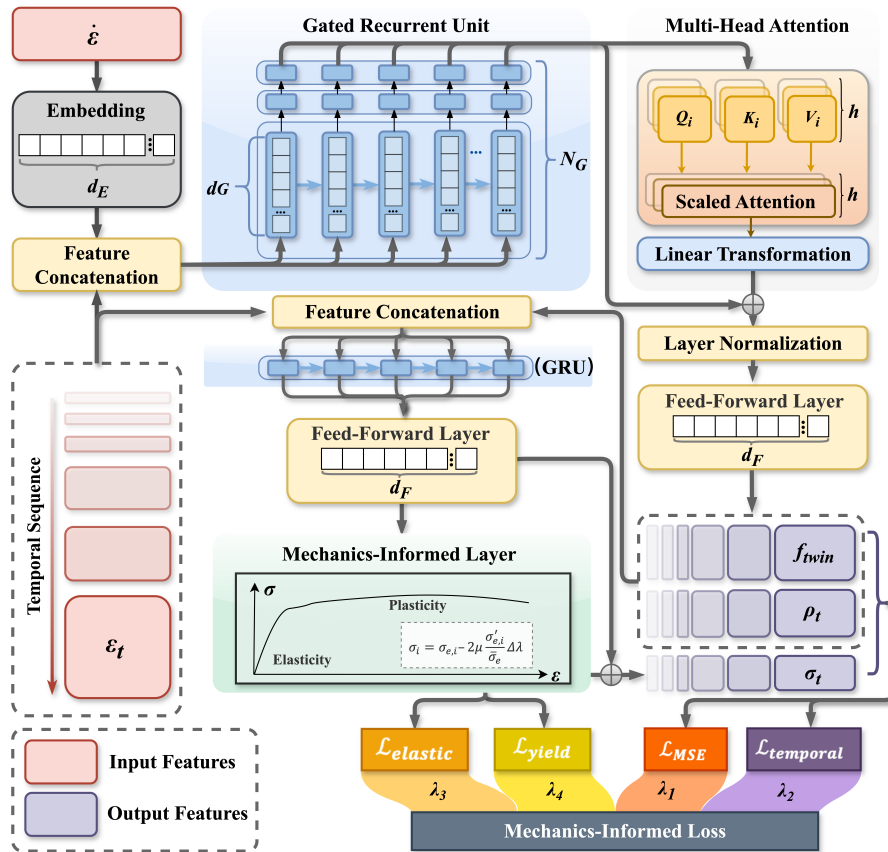


Figure 2. The architecture of the mechanics-informed deep learning constitutive model. The overall architecture illustrates input/output layers, the mechanics-informed layer integration, and the data flow from strain inputs through GRU, MHA, and FFN modules to stress and microstructural outputs. $\dot{\epsilon}$: strain rate; ϵ_t : strain; σ_t : stress; ρ_t : dislocation density; f_{twin} : twin volume fractions; GRU: gated recurrent unit; MHA: multi-head attention; FFN: feed-forward network.

to predict stress evolution. Importantly, stress predictions also incorporate constraints from the mechanics-informed layer, which explicitly enforces plasticity principles to enhance physical consistency and predictive accuracy.

Gated recurrent unit

GRUs, as efficient RNN variants, have proven effective for constitutive modeling in CP simulations. Their gating mechanisms effectively address gradient vanishing issues in traditional RNNs while preserving long-term dependencies with computational efficiency. For input sequence $x = (x_1, x_2, \dots, x_T)$, the GRU cell regulates information flow through update and reset gates:

$$z_t = \text{sigmoid}(W_z [h_{t-1}, x_t] + b_z) \tag{1}$$

$$r_t = \text{sigmoid}(W_r [h_{t-1}, x_t] + b_r) \tag{2}$$

$$\tilde{h}_t = \tanh(W_h [r_t \odot h_{t-1}, x_t] + b_h) \tag{3}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{4}$$

where z_t and r_t represent update and reset gates, h_t denotes the hidden state, x_t is the current input, and \odot indicates the Hadamard product. The update gate controls previous state retention, while the reset gate

modulates historical information's influence on the current candidate state. Here, $W_z, W_r, W_h \in R^{d_h \times (d_h + d_x)}$ are learnable weight matrices, $b_z, b_r, b_h \in R^{d_h}$ are bias vectors, d_h denotes hidden dimension, d_x denotes input dimension, and \tilde{h}_t represents the candidate hidden state.

In classical constitutive formulations, loading paths induce discrete updates of microstructure-dependent internal variables. GRUs function as learned, discrete-time state-space models of this evolution, combining sequential interpretability with intrinsic state-update operators. By capturing temporal dependencies, they recover mechanical states governed by both instantaneous loading and accumulated deformation history. The selective retention of critical temporal features through gating mechanisms enables accurate predictions across diverse loading trajectories. Compared to traditional RNNs, GRUs achieve comparable accuracy with reduced computational cost, making them particularly suitable for the present study.

Multi-head attention mechanism

MHA, a cornerstone of modern DL architectures^[21], captures cooperative interactions among deformation mechanisms through parallel processing across multiple attention heads. Each head focuses on distinct aspects of the input sequence, extracting mechanism-specific representations across multiple timescales and loading conditions. For input sequence $X \in R^{T \times d}$, where T denotes timesteps and d represents mechanical feature dimensions, single-head self-attention is formulated as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

where $Q = XW^Q \in R^{T \times d_k}$, $K = XW^K \in R^{T \times d_k}$, $V = XW^V \in R^{T \times d_v}$ are query, key, and value matrices, obtained through learnable projection matrices $W^Q, W^K \in R^{d \times d_k}$ and $W^V \in R^{d \times d_v}$, respectively. The scaling factor $\sqrt{d_k}$ controls dot product magnitudes to prevent gradient vanishing.

To enhance expressivity and capture inherent features tied to dislocation and twinning, MHA projects inputs into h subspaces, computes self-attention in each, then concatenates results with linear transformation:

$$\text{MHA}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W_o \quad (6)$$

where $\text{head}_i = \text{Attention}(Q_i, K_i, V_i)$, W_o denotes the output transformation matrix, and h represents the number of attention heads. The input feature dimension must be divisible by h ; if not, a 1D convolution can adjust it. Individual heads may specialize in specific mechanisms (dislocation dynamics, strain-rate sensitivity, or coupled effects), enabling simultaneous focus on distinct representation subspaces and enhancing feature diversity and multiscale characterization.

Residual connections enhance information flow and mitigate gradient vanishing:

$$X' = \text{MHA}(X) + X \quad (7)$$

Layer normalization stabilizes training through feature normalization:

$$\mu_{LN} = \frac{1}{d} \sum_{x_i} \quad (8)$$

$$\sigma_{LN} = \sqrt{\frac{1}{d} \sum (x_i - \mu_{LN})^2 + \epsilon_s} \quad (9)$$

$$\text{LayerNorm}(x) = \gamma \left(\frac{x_i - \mu_{LN}}{\sigma_{LN}} \right) + \beta \quad (10)$$

where d represents feature dimensions, ϵ_s ensures numerical stability, μ_{LN} and σ_{LN} denote the mean and standard deviation of the input features used in the layer-normalization operation, ϵ_s is a small constant for numerical stability, and γ, β are learnable scaling and bias parameters. The normalization suppresses internal covariate shifts and accelerates convergence. The integration of MHA, residual connections, and layer normalization constructs a stable and efficient network architecture.

Mechanics-informed layer

The proposed mechanics-informed layer integrates classical elasto-plastic theory with deep learning to regularize predicted stress-strain responses. Derived from conventional J₂ flow theory, this layer is implemented through differentiable neural network formulations that preserve end-to-end differentiability while incorporating fundamental mechanics principles. By assuming initially random crystallographic texture, it enables differentiable mapping between principal strains and stresses, with tensors expressed in principal form.

Since crystal plasticity involves continuous texture evolution and lattice rotation, fixed-axis anisotropic criteria (e.g., Hill48) are unsuitable. Instead, the evolving anisotropy is captured by GRU-predicted residual terms, effectively enabling the model to learn a dynamic yield surface. It should be noted that the J₂ plasticity formulation within the mechanics-informed layer serves strictly as a universal isotropic elastoplastic baseline, rather than a complete description of evolving anisotropy. This architectural decoupling leverages the principle of residual learning: the physics layer governs the fundamental isotropic transition, freeing the data-driven components to fully dedicate their representational capacity to extracting higher-order anisotropic residuals. The effectiveness of this approach in breaking the isotropic assumption is explicitly evidenced by the model's ability to capture the non-linear distortion of the yield surface under varying multiaxial loading ratios, demonstrating its capability in representing pronounced directional dependencies. While the J₂ baseline inherently assumes tension-compression symmetry, the data-driven residual terms effectively compensate for this limitation, allowing the model to capture the distinct tension-compression asymmetry induced by the polarity of deformation twinning at high strain rates.

The layer processes a 3D principal strain tensor $\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3]$ as input. The initial estimate of the corresponding stress tensor, $\sigma_{e,i}$ is computed using a simplified isotropic elasticity model:

$$\sigma_{e,i} = 2\mu\epsilon_i + \lambda(\epsilon_1 + \epsilon_2 + \epsilon_3) \quad (11)$$

where λ and μ are Lamé constants, treated as learnable parameters within the model. To evaluate potential plastic deformation, the von Mises yield criterion is implemented. The equivalent trial stress, $\bar{\sigma}_e$, is calculated as:

$$\bar{\sigma}_e = \sqrt{\frac{3}{2} \sum_{i=1}^3 \sigma_{e,i}'^2} \quad (12)$$

where $\sigma_{e,i}' = \sigma_{e,i} - \frac{1}{3}(\sigma_{e,1} + \sigma_{e,2} + \sigma_{e,3})$ is the deviatoric tensor component of the trial stress. The yield function ϕ is defined as $\phi = \bar{\sigma}_e - \sigma_y$, where σ_y is implemented as a learnable parameter. To ensure differentiability across the elastic-plastic transition, a smooth transition is achieved by using a sigmoid function:

$$s = \text{sigmoid}(k\phi) \quad (13)$$

with parameter k controlling the transition steepness, effectively regularizing the yield surface sharpness.

During plastic deformation ($\phi > 0$), the plastic strain increment $\Delta\lambda$ is determined as:

$$\Delta\lambda = \frac{s\phi}{3\mu + H} \quad (14)$$

where H is a learnable hardening parameter accounting for strength evolution due to plastic strain. The corresponding stress is subsequently updated to include plastic contributions:

$$\sigma_i = \sigma_{e,i} - 2\mu \frac{\sigma'_{e,i}}{\bar{\sigma}_e} \Delta\lambda \quad (15)$$

Consequently, the layer outputs three-dimensional (3D) principal stress tensor $\sigma = [\sigma_1, \sigma_2, \sigma_3]$. This mechanics-informed approach preserves classical plasticity foundations while providing adaptive capabilities through physics-constrained learnable parameters, simultaneously offering physical interpretability and data-driven flexibility for material behavior simulation.

Mechanics-informed loss function

To address the mechanics-informed layer requirements, a mechanics-informed loss function is developed, designed to ensure both data fitting accuracy and physical fidelity. The multi-term loss function is formulated as:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{MSE} + \lambda_2 \mathcal{L}_{\text{temporal}} + \lambda_3 \mathcal{L}_{\text{elastic}} + \lambda_4 \mathcal{L}_{\text{yield}} \quad (16)$$

The foundational mean squared error (MSE) loss quantifies the deviation between predicted and ground truth values:

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_i (y_{i,pred} - y_{i,true})^2 \quad (17)$$

where N denotes the total sample count, and $y_{i,pred}$, $y_{i,true}$ represent predicted and actual mechanical quantities (stress components, DD, TVFs, ensuring accuracy across strain-rate dependent responses and microstructural features).

To capture temporal evolution with physical continuity, a temporal consistency constraint loss $\mathcal{L}_{\text{temporal}}$ is implemented:

$$\mathcal{L}_{\text{temporal}} = \frac{1}{N} \sum_t^{N-1} \left\| \frac{d}{dt} (y_{t+1} - y_t) \right\|^2 \quad (18)$$

where y_t and y_{t+1} represent consecutive time-step quantities. This term penalizes discontinuities in predicted mechanical quantity evolution, enforcing continuous deformation laws.

For optimized performance in the elastic regime, prevalent in the initial stages of crystal plasticity simulations, an elastic region constraint loss is implemented. This specialized term ensures that the model exhibits linear elastic behavior under small strain conditions, adhering to Hooke's law:

$$\mathcal{L}_{\text{elastic}} = \frac{1}{N} \sum_i \left((\sigma_{i,pred} - E(\epsilon_i + \epsilon_s))^2 \cdot I\left(|\epsilon_i| \leq \frac{\sigma_y}{E}\right) \right) \quad (19)$$

where $I\left(|\varepsilon_i| \leq \frac{\sigma_y}{E}\right)$ is an indicator function equaling 1 within the elastic regime, and 0 otherwise. The term ε_s prevents numerical instability. Here, the elastic constraint is explicitly defined using a linear stress-strain proportionality to enforce Hooke-type behavior in the small-strain regime.

The yield criterion loss enforces physically consistent elastic-plastic transition behavior, constraining stress predictions to remain within physically meaningful boundaries:

$$\mathcal{L}_{\text{yield}} = \frac{1}{N} \sum_i (\max(0, \sigma_{i, \text{pred}} - \sigma_y - H \cdot \varepsilon_{p,i}))^2 \quad (20)$$

By incorporating the yield stress σ_y , hardening coefficients H , and plastic strain $\varepsilon_{p,i}$, the model accurately simulates the nonlinear mechanical responses. This component ensures that predicted stress trajectories respect fundamental plasticity principles, preventing unphysical extrapolations.

The comprehensive loss function is formulated as a weighted sum of individual components, with coefficients $\lambda_1 - \lambda_4$ adjustable to specific applications and optimization objectives. This parameterization allows precise modulation of model focus, increasing λ_3 enhances elastic regime performance, while amplifying λ_2 strengthens physical plausibility constraints.

These embedded mechanical constraints provide intrinsic regularization, reducing training data requirements, which is critical for data-scarce materials research. The MIDLCM transforms traditional data-intensive paradigms into knowledge-enhanced learning processes, leveraging established physical principles for accurate, physically consistent, and empirically effective regularization. Note that in the present study, the specific weight allocation ($\lambda_1 = 1.0, \lambda_{2-4} = 0.1$) was determined based on physical intuition rather than an exhaustive sensitivity analysis. This 10:1 ratio establishes a hierarchical regularization strategy: the data-fitting term (λ_1) acts as the primary optimization objective, while the mechanics-informed terms (λ_{2-4}) serve as auxiliary soft constraints. As validated by the subsequent ablation study, this empirically established ratio effectively enhances physical consistency without overwhelming the data-driven learning process, representing a reasonable baseline setting rather than a mathematically optimal choice. These current weights establish a stable hierarchical relationship; nevertheless, precisely optimized values may be further determined using adaptive loss weighting techniques.

Data preparation and training configuration

To validate the proposed MIDLCM's capability and efficiency in capturing strain-rate sensitive, multi-mechanism-coupled mechanical behavior, this study employs CrFeNi MEA as the model material, exhibiting well-documented strain-rate dependent mechanical behavior^[26]. Specifically, the strain rate can influence the strain rate sensitivity value defined at yielding points, the dislocation multiplication, as well as the activation of twinning. This strain rate dependent, mechanism coupled complexity is hard to capture using the NN-based models already published^[27]. As a predominantly FCC solid solution with lattice distortion from equiatomic chemical disorder, CrFeNi provides a rigorous test case encompassing fundamental FCC metal physics and inherent complexity^[28,29].

DL-based models require a curated dataset for training. To fully capture the mechanical and microstructural responses of the CrFeNi alloy, experimental measurements are augmented with simulations from a fully calibrated polycrystalline CP model, Elasto-Visco-Plastic Self-Consistent model (EVPSC)^[30]. Accordingly, a large training dataset is generated with the EVPSC model and used to train the MIDLCM and assess its capability. This data-preparation strategy is widely used in data-driven constitutive modeling. [Figure 3](#) summarizes the workflow; detailed procedures are provided in the following subsections.

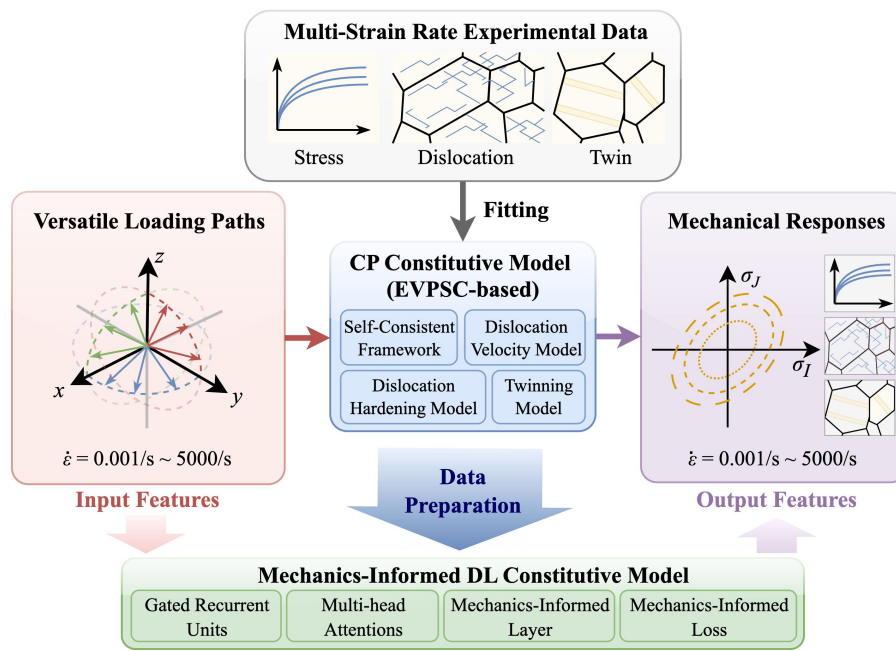


Figure 3. Data-preparation workflow illustrating dataset construction, considered mechanisms, and input-output feature composition. EVPSC: Elasto-Visco-Plastic Self-Consistent model; DL: deep learning.

A brief introduction of training dataset generator (EVPSC model)

EVPSC model^[30] is employed for its comprehensive integration of multiple deformation mechanisms with strain-rate sensitivity while maintaining computational efficiency superior to conventional crystal plasticity models, enabling practical training dataset generation within reasonable computational costs. The EVPSC model incorporates dislocation and twinning mechanisms within a self-consistent crystal plasticity formulation^[1,31]. The constitutive formulation describes the elastic-plastic deformation of individual grains through the deformation gradient tensor, where elasticity is characterized by the Jaumann rate of Cauchy stress coupled with anisotropic stiffness tensors^[30].

We incorporated multiple mechanisms for a reliable description of the strain rate sensitive mechanisms and mechanical behaviors. Dislocation-mediated plasticity follows Orowan's equation^[32], with dislocation dynamics (nucleation, multiplication, annihilation) governed by differential equations contributing to strain-rate sensitivity^[33,34]. Concurrently, the TDT model captures twin nucleation and evolution^[30]. Grain boundaries influence plasticity through the Hall-Petch effect^[35-39]. Dislocation pile-up and boundary nucleation are incorporated into the DD equations, dynamically updating slip resistances and enabling strain-rate-dependent hardening^[40,41]. The self-consistent framework homogenizes grain interactions by iteratively solving stress equilibrium and compatibility across the polycrystal, simulating uniaxial tension at different strain rates^[31]. This approach captures macroscopic stress-strain responses, texture evolution, and TVF changes, with time-stepping algorithms resolving the coupling between grain-scale mechanisms and polycrystalline homogenization.

The polycrystalline system comprises 5,000 grains with 90% face-centered cubic (FCC) matrix strengthened by 10% body-centered cubic (BCC) particles. BCC particles are incorporated by increasing slip resistances on 12 $\{111\}\langle 110 \rangle$ slip systems in the FCC phase, approximating their barrier effect on dislocation motion without explicit mesoscale modeling. Strain rates range from 0.001/s to 4,000/s. $\{111\}\langle 112 \rangle$ twinning also contributes to plasticity, becoming more significant at high strain rates, consistent with experimental observations. A more detailed formulation, the calibrated model parameters, and the nomenclature are

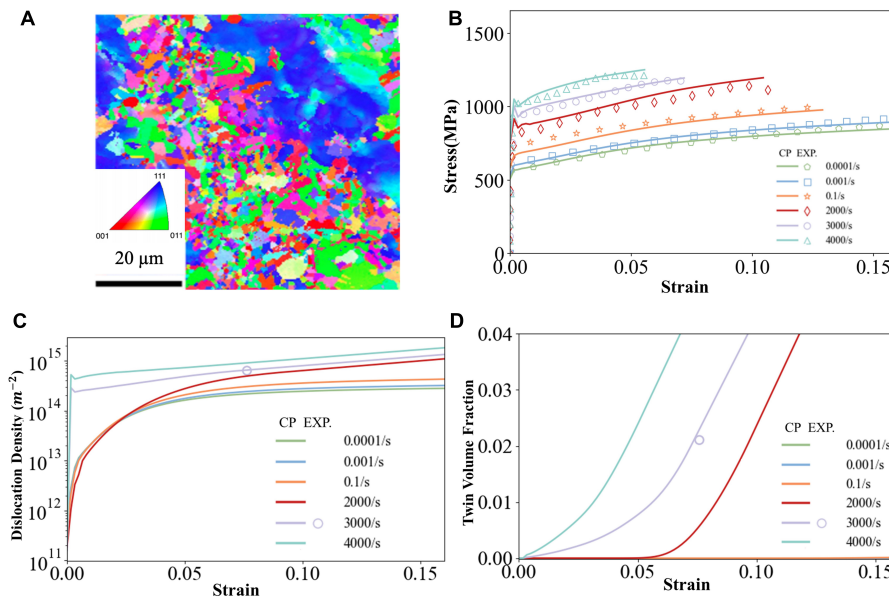


Figure 4. (A) Electron backscatter diffraction map illustrating the initial texture of the CrFeNi alloy prior to deformation, depicted through an inverse pole figure color scheme^[26]; (B-D) Comparison of the mechanical response predictions including stress, DD and twin volume fraction derived from crystal plasticity (CP) simulations to experimental (Exp.) measurements obtained at varying strain rates. Experimental results adapted from Wang et al.'s study^[26].

provided in the [Supplementary Materials \[Supplementary Tables 1 and 2\]](#). The EVPSC dataset assumes linear elasticity, justifying the Hooke's Law constraint. For extensions into shock physics, this constraint requires modification to accommodate adiabatic heating and shock-induced nonlinear elastic effects.

Parameter calibration of EVPSC for CrFeNi alloys

To establish a reliable dataset for training the constitutive model, numerical simulations are performed using the aforementioned EVPSC model, with subsequent validation against experimental CrFeNi alloy observations. Model parameters are calibrated using tensile deformation data across multiple strain rates, with experimental validation encompassing four strain rates (0.001/s, 0.1/s, 3,000/s, and 4,000/s) along the rolling direction^[26]. The initial microstructure features a random crystallographic texture (see [Figure 4A](#)). Parameter optimization yields an optimal constitutive parameter set (see [Supplementary Table 1](#)), enabling high-fidelity dataset generation. The calibrated model accurately reproduces CrFeNi's mechanical response across the deformation regime, with predictions showing excellent agreement with experimental stress-strain data [[Figure 4B](#)], as well as the corresponding DD evolution [[Figure 4C](#)] and TVF evolution [[Figure 4D](#)] at different strain rates. This consistency reflects the model's accurate representation of fundamental mechanisms, including dislocation nucleation, multiplication, grain boundary accumulation, and strain hardening.

The EVPSC predictions exhibit clear strain-rate-sensitive behavior. The strain rate sensitivity of the yielding stress is markedly higher near 3,000/s than under quasi-static conditions, and only the dynamic cases show a rapid initial rise in DD and activation of twinning. Accurately reproducing these strain rate sensitive features is essential yet challenging for deep learning constitutive models, and our proposed MIDLCM is designed to capture them. This reproduction will underscore the advantages of the present model design.

High-fidelity dataset generation and pre-processing

High-fidelity training data are generated using the calibrated CrFeNi polycrystalline model under plane stress conditions, with specimens subjected to biaxial tension at loading angle $\alpha = \arctan\left(\frac{\varepsilon_2}{\varepsilon_1}\right)$, as shown in

Figure 3. Remaining strain components satisfy stress-free boundary conditions. As shown in **Figure 4A**, a single microstructure with random crystallographic orientations is employed across multiple loading directions to optimize computational efficiency while ensuring dataset representativeness. Simulations span six strain rates (0.001/s, 0.1/s, 10/s, 1,000/s, 3,000/s, and 5,000/s) covering quasi-static to high rate regimes, with maximum equivalent strain limited to 0.16. Critical micromechanical features, including stress-strain tensors, dislocation densities, and TVFs, are recorded at each increment. The dataset employs a single random-texture microstructure, with 756 cases spanning six strain rates, and three crystallographic planes. Rather than an algorithmic limitation, this approach serves as a deliberate scientific control. The primary objective of this framework is to isolate and project the complex rate-dependent kinetics - specifically, the transition from dislocation-dominated plasticity to twinning-assisted mechanisms under extreme dynamic loading - into the deep learning architecture. Simultaneously introducing variations in initial grain size, texture, and phase fractions would heavily convolute these geometric and crystallographic variables with the kinetic rate-dependent mechanisms, thereby obscuring the underlying physical causality. By maintaining a controlled initial microstructural state, we ensure that the framework's capability to accurately capture explicit macro-micro coupled physics across varying strain rates and loading paths can be rigorously and unambiguously evaluated.

Traditional constitutive models employ sequential incremental integration, computing stress and microstructural states step-by-step through iterative time-stepping procedures, as shown in **Figure 1B**. Conversely, DL constitutive models adopt parallel computation, accepting complete strain histories as input and simultaneously predicting full stress responses and microstructural evolution in a single forward pass, achieving orders of magnitude computational acceleration while eliminating incremental procedures.

Data preprocessing enhances model convergence and performance. Logarithmic transformation is applied to dislocation densities and strain rates to compress multi-order magnitude variations and improve model sensitivity across scales. Subsequently, all features, including strain-stress histories, strain rates, dislocation densities, and TVFs, undergo standardization to zero mean and unit variance:

$$\tilde{x} = \frac{x - \mu}{\sigma} \quad (21)$$

where μ and σ represent the mean and standard deviation of the original data.

The dataset is partitioned into training and testing sets at a 4:1 ratio, with representative sampling ensuring balanced distributions across strain paths and rates. The case-based splitting assigns each case, defined by strain rate, loading angle, and crystallographic plane, atomically to training or test sets, preventing temporal correlation leakage while ensuring test cases represent genuinely unseen loading paths.

Model training configuration

All experiments are conducted on a computer equipped with an NVIDIA RTX 3090 GPU, utilizing PyTorch 2.0.0 and CUDA 11.8 to leverage GPU-accelerated computation. The optimization strategy is meticulously designed to ensure model convergence, stability, and performance. The AdamW optimizer^[42] is employed, combined with a cosine annealing learning rate scheduler. AdamW, an advanced variant of the Adam, decouples weight decay from gradient updates, enhancing generalization through explicit weight regularization.

The hyperparameters for AdamW are configured with $\beta_1 = 0.9$, $\beta_2 = 0.999$, a weight decay coefficient of 0.01, an epsilon value of 10^{-8} , and a learning rate of 0.001. A cosine annealing scheduling dynamically adjusted the learning rate throughout training. This scheduler gradually increases the learning rate during initial training

to prevent parameter instability, subsequently modulating the learning rate through a cosine annealing function:

$$\eta_t = \eta_{\min} + 0.5 * (\eta_{\max} - \eta_{\min}) * \left(1 + \cos\left(\frac{t}{T}\pi\right)\right) \quad (22)$$

where η_{\min} and η_{\max} represent minimum and maximum learning rates, t denotes the current step, and T represents the total steps. The first 10% of steps are allocated for warm-up, with $\eta_{\max} = 0.001$, $\eta_{\min} = 0$, and $T = 1,000$. Compared to traditional linear decay methods, this scheduler enables effective parameter space exploration while mitigating convergence to suboptimal local minima.

Model weights are initialized using Xavier initialization to establish appropriate initial parameter distributions. MIDLCM is trained from scratch for 1,000 epochs with a batch size of 64. Gradient clipping with a maximum norm threshold of 1.0 is implemented to enhance training stability and prevent gradient explosion. Validation performance is continuously monitored during training, with optimal model parameters preserved based on validation loss minimization. This protocol ensured robust convergence while maintaining physical consistency in the learned representations of material behavior. Training stability was promoted through Xavier initialization, cosine-annealed AdamW optimization, gradient clipping, and validation-based checkpoint selection.

The present MIDLCM is formulated as a deterministic constitutive framework, and no formal uncertainty quantification is considered in this study. Variance-based sensitivity analysis and probabilistic treatment of uncertain material parameters or loading conditions remain important directions for future work when extending the model to reliability-oriented engineering applications.

RESULTS AND DISCUSSION

Model performance analysis

Understanding the optimal configuration and performance characteristics of MIDLCM requires a systematic evaluation of its architectural components and computational behavior. This section presents hyperparameter sensitivity analysis, ablation studies, and computational efficiency analysis to elucidate the structural characteristics and performance mechanisms of MIDLCM. The hyperparameter analysis examines critical architectural parameters through sensitivity evaluation, quantifying the model's learning capacity and generalization performance across the parameter space. Ablation studies progressively integrate key components into the baseline architecture, quantifying each component's marginal contribution to predictive accuracy. Additionally, computational efficiency analysis benchmarks MIDLCM's inference speed and resource utilization against conventional CP simulations. These analyses elucidate the model's structural mechanisms, providing insights and guidance for the applications of DL constitutive model.

Hyperparameter sensitivity analysis

This study employs Bayesian optimization with Tree-structured Parzen Estimator^[43] for efficient hyperparameter tuning of MIDLCM. Unlike traditional grid search methods, Bayesian optimization constructs a probabilistic model of the objective function (validation MSE), and enables directed exploration of the high-dimensional parameter space through sequential sampling with acquisition functions that balance exploitation and exploration. As shown in Figure 2, the hyperparameter search space encompassed: GRU hidden dimensions, GRU layer number $N_G \in \{1, 2, 3, 4\}$, embedding dimensions $d_E \in \{1, 2, 3, 4\}$, MHA heads $h \in \{1, 2, 4, 8\}$, and FFN hidden dimensions $d_F \in \{64, 128, 256, 512\}$. Optimization converged after 35 iterations out of 50 total iterations, yielding an optimal configuration of a two-layer GRU with 256 hidden dimensions, three embedding dimensions, four attention heads, and an FFN hidden dimension of

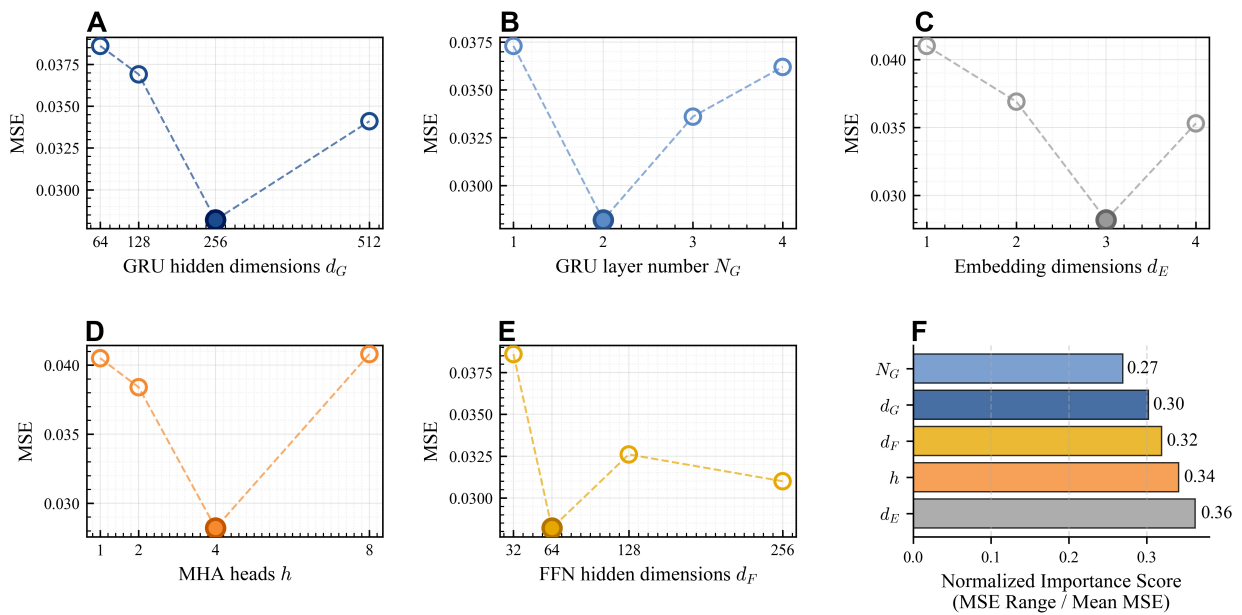


Figure 5. Experimental results of hyperparameter sensitivity. (A-E) MSE variations across key hyperparameters: GRU hidden dimensions, GRU layers, embedding dimensions, MHA heads, and FFN hidden dimensions; (F) Normalized importance scores of investigated hyperparameters. MSE: Mean squared error; GRU: gated recurrent unit; MHA: multi-head attention; FFN: feed-forward network.

128. This configuration balances representational capacity with computational efficiency while minimizing overfitting risk.

As shown in Figure 5, sensitivity analysis results reveal distinct performance patterns across architectural configurations. GRU hidden dimensions exhibit a non-monotonic relationship with prediction accuracy, with MSE decreasing from 0.0386 to 0.0282 as dimensions increase from 64 to 256, then deteriorating at 512 dimensions (MSE = 0.0341), indicating an optimal representational capacity threshold beyond which overfitting occurs. Network depth analysis confirms optimal performance at two GRU layers, with both shallower and deeper architectures showing reduced efficacy. Embedding dimensionality achieves peak performance at three dimensions, sufficiently capturing the rate feature space complexity. MHA mechanism peaks at four heads, indicating that moderate attention complexity best facilitates feature integration. FFN performance maximizes at 128 hidden dimensions, with larger configurations degrading due to overparameterization.

These findings demonstrate that model performance maximizes at intermediate complexity levels across all hyperparameters, emphasizing the critical balance between representational capacity and overfitting in crystal plasticity constitutive modeling. The narrow MSE variation range (0.028–0.041) indicates robust architectural performance with limited hyperparameter sensitivity, suggesting inherent stability in capturing crystal plasticity mechanisms. Normalized importance analysis reveals that for the rate-dependent modeling task, embedding dimension exhibits the highest relative importance, followed sequentially by MHA heads, FFN hidden dimensions, GRU hidden dimensions, and GRU layer number.

Ablation experiments

Ablation experiments of MIDLCM quantify individual component contributions by progressively incorporating MHA, mechanics-informed layer, and mechanics-informed loss function into a baseline GRU network. Performance metrics include MSE, mean absolute error (MAE), parameter count, and inference time, with results summarized in Figure 6.

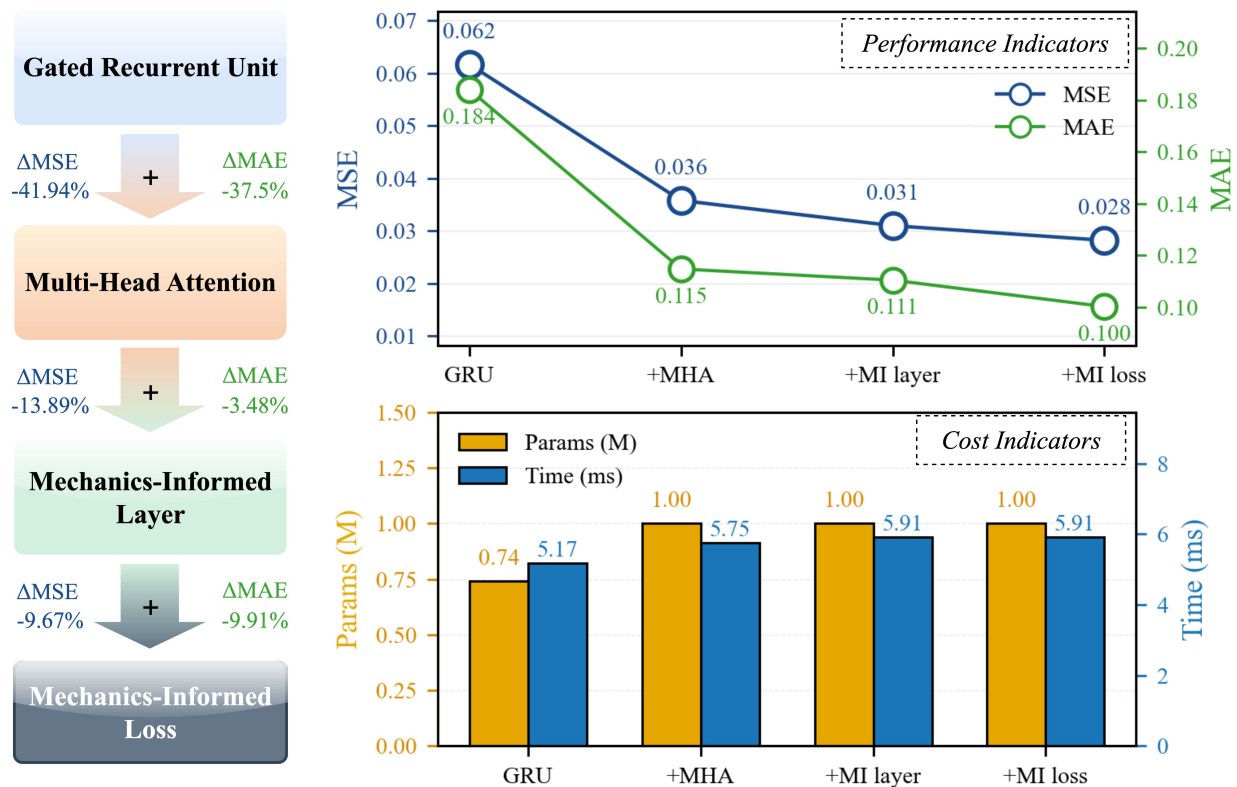


Figure 6. Performance and computational cost results of model ablation experiments examining the four principal components of MIDLCM architecture. MSE: Mean squared error; MAE: mean absolute error; GRU: gated recurrent unit; MHA: multi-head attention; MIDLCM: mechanics-informed deep learning constitutive model.

The baseline GRU network achieved moderate accuracy (MSE = 0.062, MAE = 0.184) with computational efficiency (5.17 ms inference, 0.74M parameters). Incorporating MHA substantially improved performance, reducing MSE and MAE by 42.0% and 37.5% respectively, with minimal computational overhead (1.00M parameters, 5.75 ms). This improvement demonstrates MHA's effectiveness in capturing temporal dependencies inherent in mechanical and microstructural evolution. By mitigating the vanishing-gradient issues of standard GRUs, MHA ensures that critical historical shifts across the extreme strain-rate spectrum are accurately preserved and transmitted to the subsequent mechanics-informed layer, laying the mathematical foundation to enforce physical consistency. Sequential integration of the mechanics-informed layer and loss function yielded progressive enhancements, ultimately achieving optimal performance (MSE = 0.028, MAE = 0.100) while maintaining computational efficiency (5.91 ms). The mechanics-informed components contributed incremental but consistent improvements of 13.9% and 9.7% in MSE reduction, respectively. These incremental gains confirm that the improvement arises from the powerful synergy of explicitly tracking rate-dependent mechanism shifts, mechanics-informed elastoplastic regularization, and sequence interaction learning, rather than from using the GRU architecture alone. To further evaluate algorithmic robustness against data partitioning and local minima, comprehensive five-fold cross-validation and multi-seed repeated runs were conducted on the optimized MIDLCM framework. The five-fold case-based cross-validation yielded a mean test MSE of 0.0288 ± 0.0010 and a mean test MAE of 0.1016 ± 0.0024 , indicating stable generalization under different train-test splits. Furthermore, five independent training runs using different random initialization seeds resulted in a mean MSE of 0.0282 ± 0.0005 and MAE of 0.1002 ± 0.0013 . These results, detailed in [Supplementary Materials \[Supplementary Table 3\]](#), explicitly demonstrate that the framework is highly stable, maintaining consistent predictive performance without acute sensitivity to random stochasticity.

These results demonstrate the synergistic effect of combining deep learning with mechanics-informed constraints. Each component addresses distinct aspects of crystal plasticity modeling: MHA captures complex temporal patterns, while mechanics-informed components embed domain knowledge into both architecture and optimization. The progressive improvement indicates that the mechanics-informed components improve predictive accuracy and physical consistency relative to the purely data-driven baseline for constitutive modeling, achieving superior accuracy without computational penalty through principled integration of mechanical principles.

Computational efficiency analysis

Evaluating computational efficiency requires careful context. While EVPSC serves as the high-fidelity ground truth resolving comprehensive grain-level details, MIDLCM significantly accelerates computations by capturing the overall microstructural evolution. For single-case predictions, MIDLCM processes 100-timestep strain paths in 0.005 s, while EVPSC requires approximately 120 s - a 24,000-fold speed-up. This computational advantage is further enhanced in batch processing. For parallel predictions across 100 distinct loading conditions, MIDLCM completed analyses in 0.07 s compared to approximately 200 min for EVPSC. Furthermore, comparing MIDLCM directly to purely macroscopic deep learning models would be unequal, as it overlooks the inherent computational complexity of explicitly tracking internal state variables like DD and TVF. Ultimately, this acceleration preserves critical mechanism-level fidelity, establishing the framework as highly efficient for rapid material behavior assessment and multi-scale simulations.

While empirical macroscopic models (e.g., Johnson-Cook) offer extreme computational efficiency for industrial FEM simulations, they inherently trade microstructural fidelity for speed and cannot spontaneously capture dynamic state-driven mechanism transitions. Achieving comparable microstructural fidelity requires traditional constitutive frameworks with multiple internal state variables and implicit integration of highly nonlinear, stiff evolution equations (ODEs), substantially increasing computational costs. In contrast, MIDLCM avoids these expensive iterative updates through a unified deep learning architecture, providing an efficient surrogate while retaining explicit microstructural descriptors. Because the present study focuses on mechanism-coupled predictions rather than purely macroscopic stress mapping, a direct benchmark against calibrated traditional phenomenological models is not included. Therefore, the current computational results should be strictly interpreted as a highly favorable fidelity-efficiency tradeoff relative to the EVPSC reference model, rather than a general superiority claim over all macroscopic constitutive formulations.

Prediction competence of MIDLCM

After describing the overall prediction loss, the design capability for performance enhancement, and the inference speed, the intuitive performance in describing the strain-rate-sensitive, multi-mechanism coupling mechanical response in metals (with the CrFeNi alloy in this work as an example) will be discussed in detail in this section. This is to show that the current design of MIDLCM can effectively capture complex yet relatively meaningful behaviors, a capability not yet shown in other deep learning constitutive models.

Uniaxial tension prediction competence compared to experimental data

The proposed MIDLCM predictions for uniaxial tension at strain rates of 0.001/s, 0.1/s, and 3,000/s are validated against experimental data from Wang *et al.*^[26]. As shown in [Figure 7A](#), MIDLCM demonstrates excellent agreement with the experimental stress-strain curves across the examined strain rates. The model accurately captures microstructural evolution, including initial dislocation proliferation during early plasticity and subsequent saturation due to dynamic recovery processes, as shown in [Figure 7B](#). TVF predictions effectively reproduce strain rate-dependent activation and growth patterns, demonstrating the

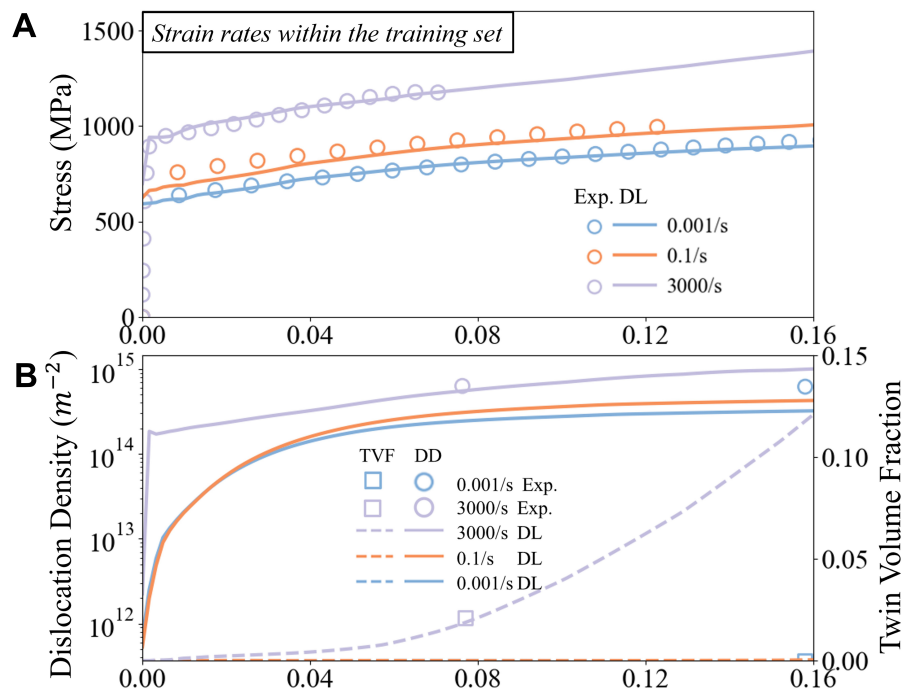


Figure 7. Evolution of stress, DD, and TVF with deformation in polycrystalline CrFeNi samples with strain rates of 0.001/s, 0.1/s, and 3,000/s. TVF: Twin volume fraction; DD: dislocation density; DL: deep learning.

model's capability to represent underlying deformation mechanisms across multiple orders of magnitude in strain rate.

To evaluate extrapolation performance, MIDLCM is tested on CrFeNi alloys under uniaxial tension at strain rates of 0.0001/s, 2,000/s, and 4,000/s - conditions outside the training dataset. Figure 8 illustrates remarkable consistency between predicted stress-strain curves and experimental measurements across these untrained conditions. The model accurately captures strain rate-dependent behavior, including elastic response, yield phenomena, and strain hardening characteristics. Microstructural predictions successfully reproduce DD and TVF evolution from quasi-static (0.0001/s) to ultra-high strain rate conditions (4,000/s), demonstrating robust representation of rate-dependent mechanisms beyond training boundaries. This extrapolation capability validates the physical consistency of the mechanics-informed architecture, indicating a genuine understanding of underlying physics rather than data interpolation. The demonstrated accuracy and computational efficiency across diverse loading conditions establish the model as an effective alternative to computationally intensive crystal plasticity simulations, particularly valuable for engineering applications with limited experimental data availability.

Prediction competence under biaxial loading conditions

Quantitative validation demonstrates excellent agreement between MIDLCM predictions and crystal plasticity simulations. Figure 9 presents parity plots for timestep-averaged equivalent stress, dislocation density, and TVF under biaxial loading conditions, where predicted values closely align with reference solutions along the $y = x$ diagonal across all samples. This strong correlation validates the model's capability to simultaneously capture macroscopic mechanical response and underlying microstructural evolution with high fidelity.

To demonstrate MIDLCM's predictive accuracy across varying performance levels, three representative samples with different MSE magnitudes and loading angle α across different planes are examined in

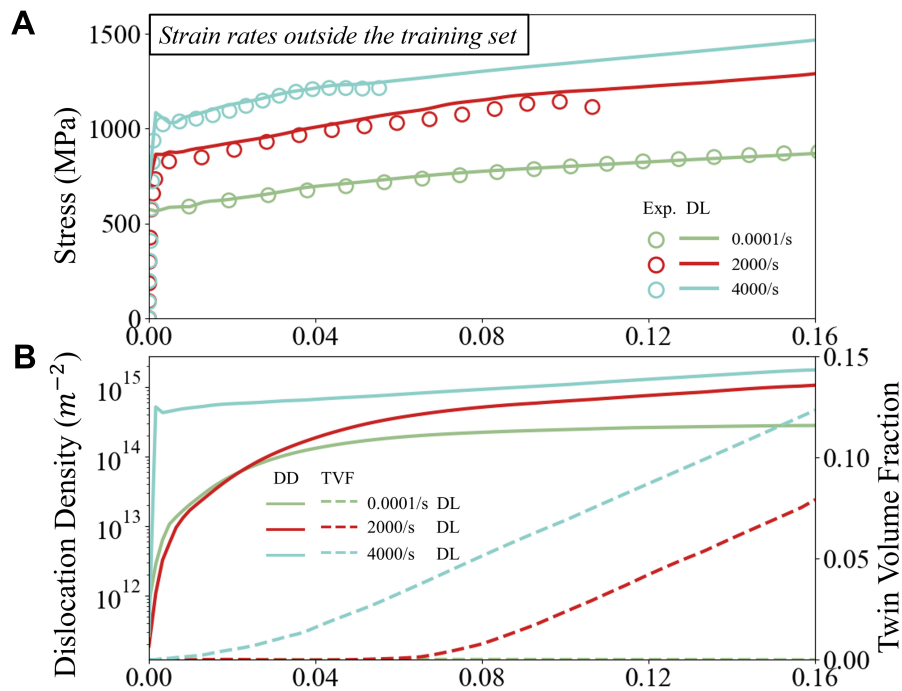


Figure 8. Evolution of stress, DD, and TVF with deformation in polycrystalline CrFeNi samples with strain rates of 0.0001/s, 2,000/s, and 4,000/s. TVF: Twin volume fraction; DD: dislocation density; DL: deep learning.

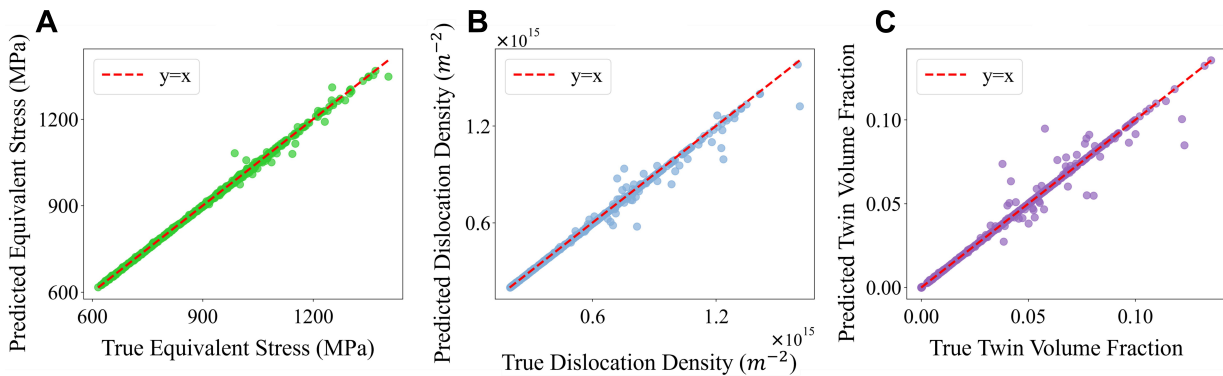


Figure 9. Parity plots comparing MIDLCM predictions with reference dataset under biaxial loading conditions for (A) time-averaged equivalent stress; (B) time-averaged DD; and (C) time-averaged twin volume fraction. MIDLCM: Mechanics-informed deep learning constitutive model; DD: dislocation density.

Figure 10. These samples are strategically selected to showcase predictions spanning from high-accuracy cases to moderate-precision scenarios and more challenging predictions, collectively illustrating the model's capability range across diverse loading conditions. Compared to reference dataset, MIDLCM accurately reproduces elastic behavior, yield transition, and strain hardening characteristics. Low strain rates exhibit gradually declining hardening rates, while high rates induce initial stress fluctuations followed by rapid stabilization and elevated hardening. These variations reflect strain rate-dependent activation of dislocation and twinning mechanisms.

DD evolution demonstrates the model's microscale accuracy. The model captures nonlinear behavior, including initial rapid growth and subsequent saturation. Below 2,000/s, DD increases steadily; above this threshold, accelerated proliferation occurs at plastic flow onset. This behavior, governed by grain boundary

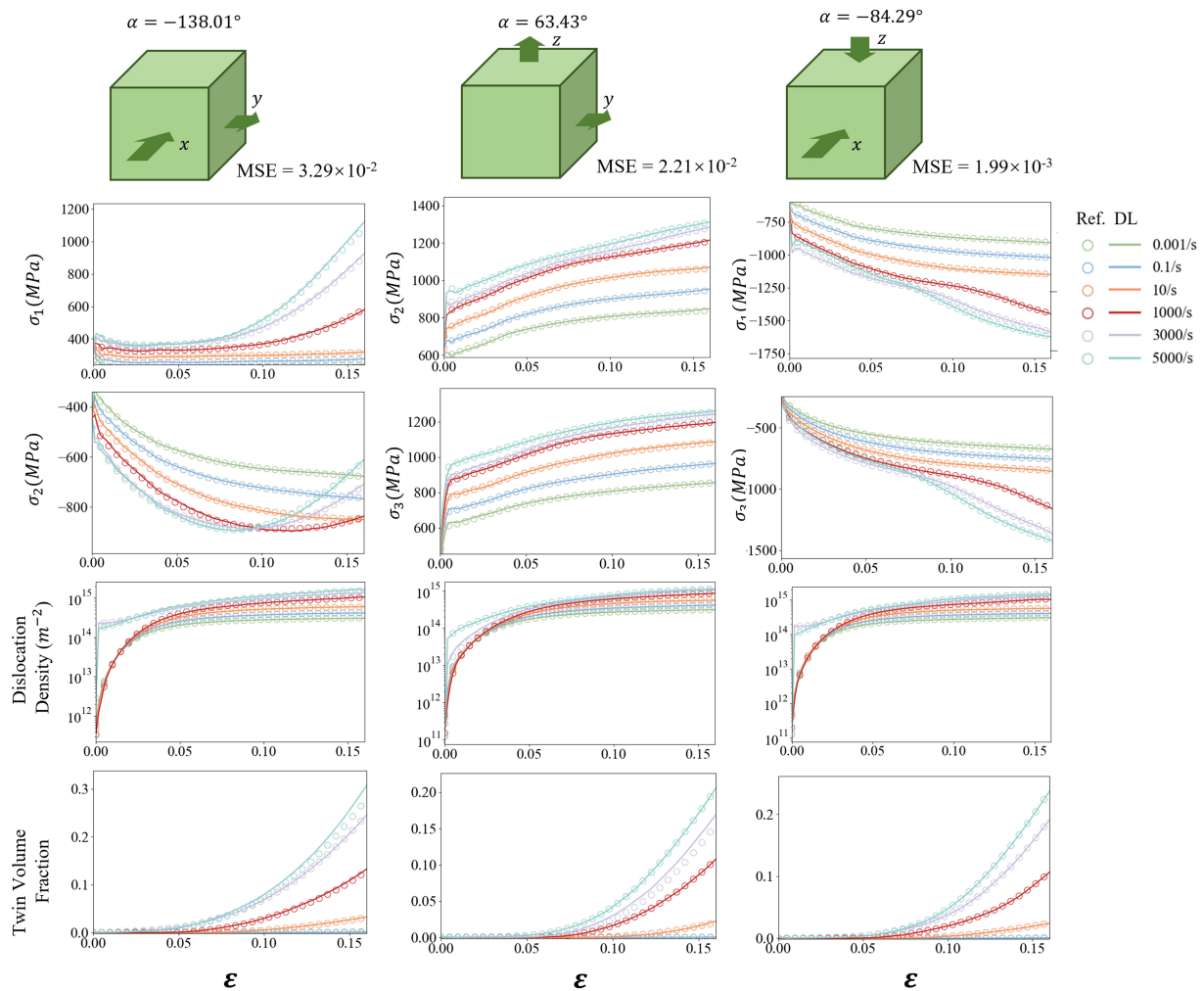


Figure 10. MIDLCM (labeled as DL) and reference dataset (labeled as Ref.) results for the mechanical response of CrFeNi alloys with strain rates from 0.001/s to 5,000/s. MIDLCM: Mechanics-informed deep learning constitutive model; DL: deep learning.

nucleation mechanisms, is essential for representing dislocation hardening under high strain rates. Higher saturation densities at elevated rates result in increased final dislocation densities.

For twinning behavior, MIDLCM correctly predicts negligible twin nucleation under quasi-static loading and significant activation at medium-to-high strain rates. Twin proliferation rates vary with strain rate due to stress-dependent growth kinetics: below 10/s, twinning propagation remains minimal, while higher rates induce substantial acceleration. MIDLCM successfully reproduces these strain rate-dependent characteristics, capturing the complex interplay between coupled deformation mechanisms and validating its capability for strain rate-sensitive phenomena simulation.

Twinning mechanisms significantly influence mechanical behavior through orientation-dependent deformation that induces asymmetric responses. To investigate this phenomenon, yield surfaces are predicted under biaxial loading at 0.001/s and 3,000/s across equivalent plastic strains of 0.01, 0.05, 0.1, and 0.15. Data are extracted from xy , yz , and xz principal strain planes in the sample coordinate system (x : rolling direction, z : normal direction) as shown in [Figure 11](#).

Under quasi-static conditions (0.001/s), yield surfaces maintain symmetry throughout deformation, exhibiting a uniform expansion characteristic of isotropic hardening dominated by dislocation-mediated

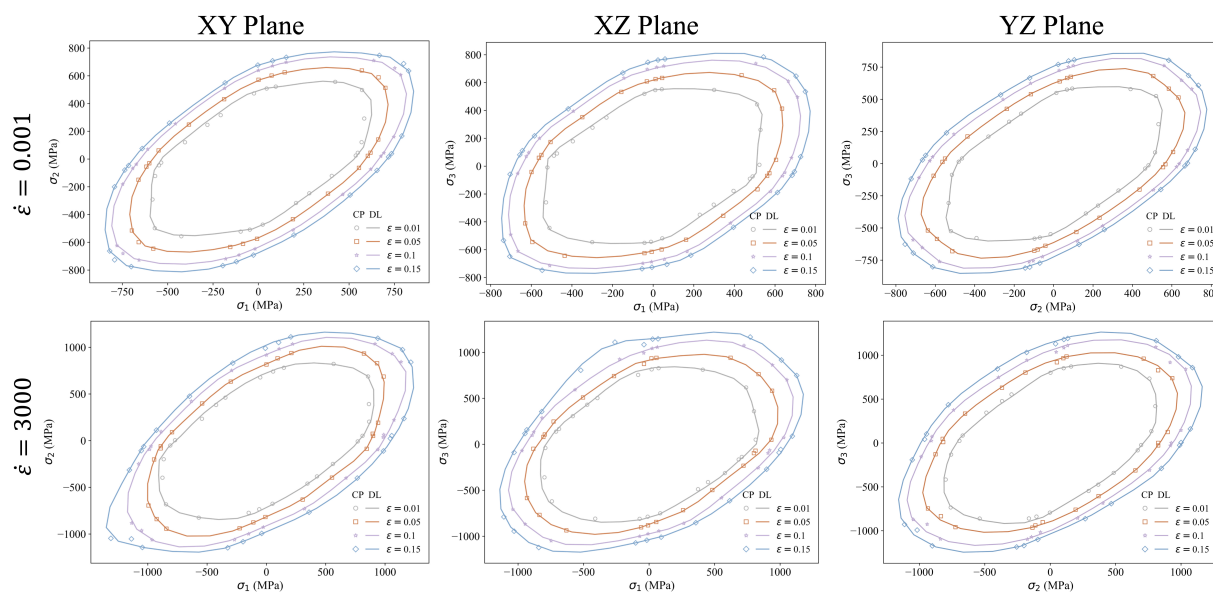


Figure 11. Prediction of yield surface shape with different plastic strains for CrFeNi polycrystalline system with strain rates of 0.001/s and 3,000/s. $\dot{\epsilon}$: strain rate.

plasticity. At dynamic rates (3,000/s), initial yield surfaces remain symmetric, but subsequent deformation induces pronounced asymmetry due to twinning activation. This demonstrates how the data-driven residual terms compensate for the symmetry of the J_2 baseline to capture the polar nature of dynamic twinning. Comparative analysis reveals distinct morphological differences and substantial stress magnitude variations between strain rates. These results demonstrate strain rate sensitivity in various aspects, encompassing concurrent changes in flow stress, yield surface morphology, mechanical anisotropy, and tension-compression asymmetry - all accurately captured by MIDLCM. Due to the extreme challenges of dynamic multiaxial experiments, these multiaxial predictions currently constitute a simulation-to-simulation benchmark. However, the rigorously calibrated EVPSC model provides a physically reasonable reference, with its true accuracy awaiting future experimental confirmation.

Furthermore, MIDLCM predictions are compared with reference dataset to assess dislocation density and TVF evolution under varying strain rates and loading orientations [Figure 12]. Analysis is performed at equivalent strains of 0.01, 0.05, 0.1, and 0.15 to capture progressive microstructural evolution. Mechanistically, the predicted rate sensitivity originates from the dynamic competition between dislocation and twinning. Under quasi-static conditions, hardening is dominated by progressive dislocation accumulation. Conversely, at dynamic rates ($3,000 \text{ s}^{-1}$), conventional dislocation mobility is severely restricted, driving a rapid, drastic defect accumulation near yielding. This physical phenomenon is consistent with dynamic mechanisms observed in highly rate-sensitive metallic systems like SAC305 solders^[44]. Explicitly linking this dislocation structure evolution to macroscopic strength ensures robust dynamic hardening predictions, aligning with recent mechanistic studies on Ni-based superalloys^[45]. Subsequently, this restricted dislocation mobility triggers extensive deformation twinning to accommodate the dynamic strain, intrinsically governing the non-linear evolution and pronounced asymmetry of the yield surfaces (as observed in Figure 11). Capturing such profound mechanism mutations across critical rate thresholds is universally essential for dynamically loaded materials, a necessity similarly highlighted in extreme thermo-mechanical studies of biomimetic polymers^[46]. Ultimately, this quantitative agreement validates the framework's reliability in capturing complex dislocation-twinning interplays.

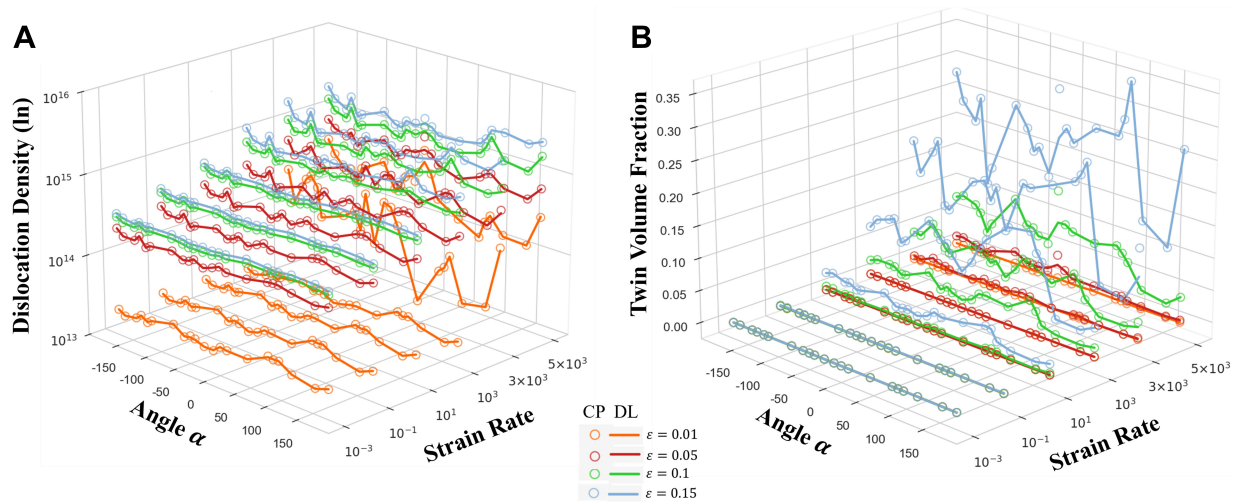


Figure 12. Prediction of dislocation density and TVF with different plastic strains for the CrFeNi polycrystalline system under varying strain rates and load conditions. CP: Crystal plasticity; DL: deep learning; TVF: twin volume fraction.

Visualization analysis of MHA under different strain rates

Deep learning models often encode mechanisms in ways that are hard to interpret. As shown previously, MHA substantially improves predictive performance by tracking temporal trends, visualized via average attention heatmaps and diagonal evolution curves in Figure 13. MIDLCM's interpretability fundamentally derives from its mechanics-informed layer enforcing plasticity principles, alongside explicit predictions of microstructural descriptors (dislocation density, TVF). Therefore, these MHA visualizations serve strictly as diagnostic tools revealing the network's temporal focus across loading stages.

In the heatmaps, color intensity denotes attention weight, with bright regions highlighting historical strain states referenced for predicting current stress. The maps are clearly strain-rate sensitive. Elevated attention around the elastic-plastic transition, particularly at lower rates, suggests yielding dictates the macroscopic response and history dependence. Moreover, early attention re-concentration mitigates GRU vanishing-gradient issues for long sequences, improving prediction fidelity. This evolving attention distribution aligns broadly with yielding and hardening behaviors. However, it must be emphasized that these learned components provide [Supplementary Materials](#) about how the model distributes attention across deformation stages; they should be interpreted only as model-level temporal attribution, rather than direct evidence of mechanism-specific causality.

An additional insight from the attention maps is that diagonal attention evolution highlights critical physical events and the model's focus on instantaneous states during deformation. Attention variations near the yield point reflect emphasis on plastic deformation initiation, while patterns during uniform plastic deformation capture work hardening behavior. Significant attention changes at failure points validate adaptive attention adjustments during crucial physical transitions. Overall, diagonal attention values increase with strain rate, indicating prioritization of instantaneous states at high strain rates. This behavior aligns with physical intuition that high strain rate responses are more rate-dependent than history-dependent. Results demonstrate adaptive attention adjustments based on loading conditions: emphasizing instantaneous effects at high strain rates while comprehensively considering historical deformation paths at low strain rates. This attention distribution provides unique insights in MIDLCM, distinct from conventional models.

While averaged attention patterns demonstrate physically meaningful behavior, individual attention heads do not form one-to-one mappings with dislocation or twinning processes, rather, MHA primarily improves temporal modeling, and the model's physical interpretability fundamentally derives from its explicit

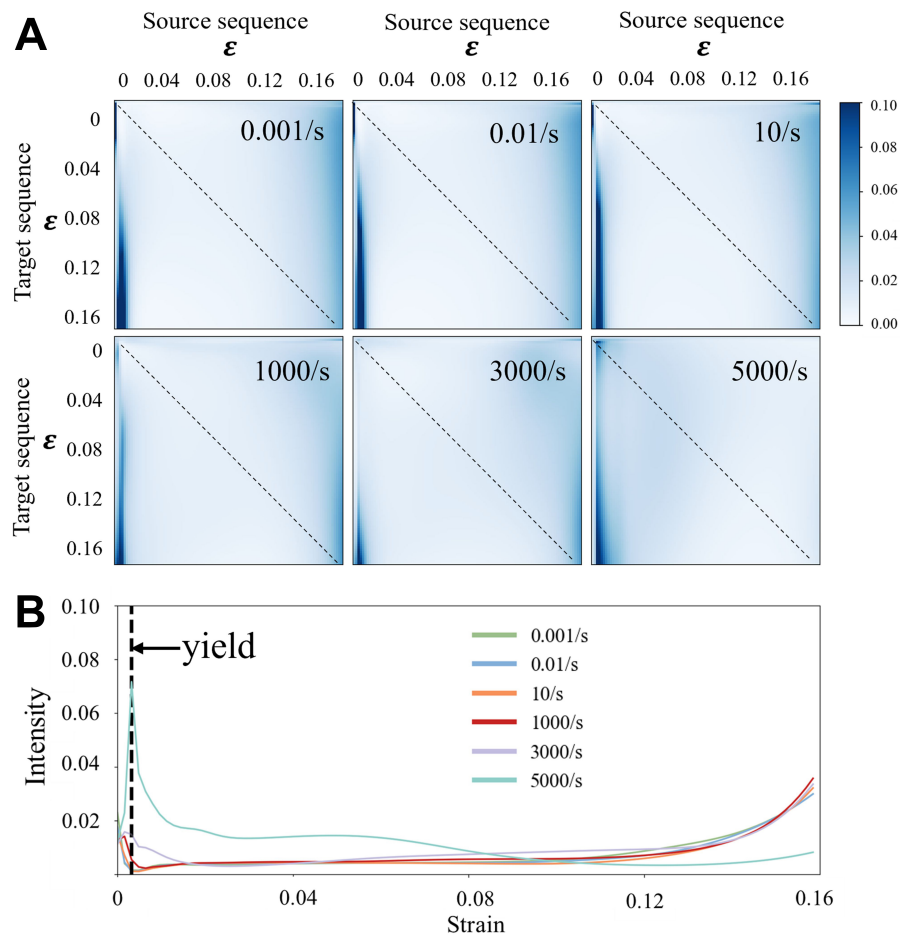


Figure 13. MHA-based attention visualization under different strain rates: (A) attention heatmaps; and (B) diagonal attention evolution. MHA: Multi-head attention.

microstructural outputs. MHA operates as a distributed representation learning mechanism capturing complex temporal dependencies, with explicit physical constraints encoded through the mechanics-informed layer rather than pre-assigned mechanism-specific attention roles. We note that attention weights provide supplementary insight into model behavior but should be interpreted cautiously, as attention patterns do not necessarily reflect causal relationships. The primary interpretability of MIDLCM derives from its explicit microstructural outputs and mechanics-informed architecture.

The analysis of mechanics-informed components

Given that the mechanics components (mechanics-informed layer and loss) clearly enhance the model's competence in the ablation experiments (see Figure 6), we should further emphasize their importance, particularly how they influence the outcomes. To highlight this, we built a model without the mechanics-informed components to show that this ablated model's performance lacks detailed accuracy. The ablated model retains identical hyperparameters as the complete MIDLCM to ensure fair comparison of the mechanics-informed components' contribution.

The results demonstrate the critical importance of mechanics-informed guidance. Compared to the complete model [Figure 9], the parity plots in Figure 14A reveal substantial performance degradation, characterized by increased scatter and deviations from the ideal diagonal. To make this degradation more intuitive, we

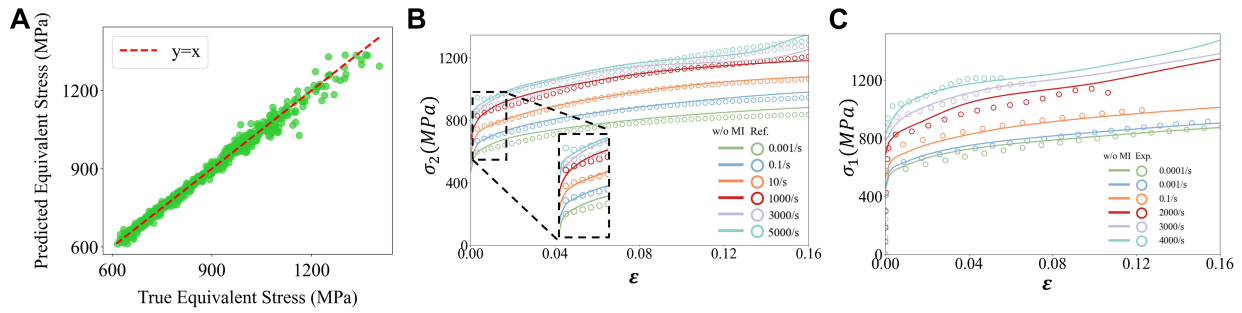


Figure 14. Performance of MIDLCM without mechanics-informed component: (A) The parity plot of the time-averaged equivalent stress; (B) results from a biaxial loading case (yz , $\alpha = 63.43^\circ$) conducted by the ablated model compared to the reference CP results; (C) uniaxial tension cases prediction from the ablated model compared with experiments results. MIDLCM: Mechanics-informed deep learning constitutive model; CP: crystal plasticity.

selected a biaxial case (yz , $\alpha = 63.43^\circ$). The ablated model yields superficially plausible predictions, but the details are insufficient for quantitative, theory-driven use. Most notably, the ablated model exhibits significant deficiencies in capturing the elastic-plastic transition regime [Figure 14B], demonstrating poor accuracy at the yield point where complex dislocation mechanisms initiate. In addition, the predicted flow stress fluctuates at higher strains, which is a prevailing phenomenon across all cases. These deficiencies also impair predictions for experimental uniaxial tension over the wide range of strain rates [Figure 14C].

The mechanics-informed components are embedded in this framework to mitigate these deficiencies. The poor prediction near the yield point largely stems from limited elastic data, with the elastic mechanical responses available only up to 0.2% strain. To address this, the mechanics-informed layer explicitly encodes the elastic-plastic transition, and the loss includes the yield point and the elastic part. We also observe flow-stress fluctuations at high strain from the ablated model. To mitigate this, we incorporate dislocation density and twinning fraction into the final stress prediction immediately before the mechanics-informed layer. This design brings microstructural information into the stress response and yields a more stable flow-stress prediction.

As shown in Figure 10, comparing with MIDLCM indicates that the mechanics-informed components successfully overcome these deficiencies, indicating that the mechanics-informed layers effectively incorporate fundamental physical principles rather than merely overfitting training data. The mechanics-informed architecture enables accurate prediction of complex mechanical behavior while maintaining physical consistency across diverse loading conditions, demonstrating its essential role in physics-informed machine learning frameworks for materials modeling. The improvement validates that the proposed mechanics-informed neural architecture, whose principles are detailed in Sections *Mechanics-informed layer* and *Mechanics-informed loss function*, significantly outperforms purely data-driven approaches for constitutive modeling.

From mechanistic modeling to engineering applications

Transitioning MIDLCM to engineering applications is an important next step. Currently, the mechanics-informed layer and corresponding validations are restricted to principal stress and strain spaces (e.g., uniaxial and proportional biaxial loading). This deliberately decouples 3D kinematic complexities to strictly isolate the rate-dependent kinetic competition between dislocation and twinning mechanisms. Generalizing this into a fully tensorial framework - incorporating frame-indifferent objective stress rates and 3D differentiable elastoplastic updates - remains a critical future goal. To handle non-proportional loading with rotating principal axes, the model can be extended by integrating a rotation-mapping layer that projects

global tensors into the material's local frame. By utilizing objective stress rates (e.g., Jaumann rates) for tensor integration, the framework can ensure frame-indifference while preserving its ability to track microstructural evolution. With multiaxial training data, this will yield a versatile surrogate for general 3D loading scenarios.

Another consideration is the framework's reliance on EVPSC-generated training data. Since capturing high-density microstructural labels (dislocation density, TVF) is experimentally unattainable across extreme strain rates, a calibrated crystal-plasticity generator is indispensable. To mitigate the risk of a pure simulation-to-simulation validation loop, MIDLCM was rigorously evaluated against independent extrapolative experimental measurements (Section *Uniaxial tension prediction competence compared to experimental data*). The strong agreement confirms the model captures physical trends rather than merely memorizing generator assumptions. Nevertheless, transitioning from this 'simulation-trained, experiment-validated' paradigm to multi-fidelity training with direct experimental data represents a vital next step.

Furthermore, while this study demonstrates material-point computational acceleration and microstructural fidelity, solving full boundary value problems (BVPs) remains the ultimate engineering metric. Currently, physical consistency is introduced through differentiable constitutive regularization. However, physics-informed networks using soft constraints face limitations for inelastic solids characterized by internal state variables and non-smooth microstructural evolutions. To ensure robust global convergence, formulating the constitutive update as the minimizer of an explicit incremental energy-dissipation potential is critical. Specifically, integrating the macro-micro coupled architecture with the Deep Energy Method (DEM)^[47], which has proven highly effective for non-smooth evolutions like phase-field fracture^[48], offers a theoretically superior route. By formulating inelastic history dependence at the level of an incremental potential, global equilibrium and thermodynamic consistency can be naturally achieved without complex return-mapping algorithms. While this transition to a fully variational BVP solver remains a subject for future multiscale implementations, the current framework's ability to circumvent stiff ODE integrations while preserving explicit microstructural descriptors establishes a vital foundation for such endeavors.

Finally, MIDLCM is currently deterministic to rigorously isolate its intrinsic capability in tracking complex coupled mechanisms. While the current scope focuses on fundamental framework feasibility, we acknowledge that the absence of confidence intervals remains a limitation for safety-critical deployment. Transitioning to probabilistic assessments - such as Bayesian Neural Networks or Deep Ensembles - requires addressing significant complexities in stochastic modeling and represents a major expansion for future reliability-oriented engineering applications.

CONCLUSIONS

This study established a mechanics-informed deep learning constitutive model (MIDLCM) for the efficient, sequential prediction of strain-rate-dependent macroscopic mechanical responses coupled with explicit microstructural evolution. By combining GRU-based history encoding, MHA, and a differentiable elasto-plastic regularization layer, MIDLCM provides accurate stress-strain prediction across a wide strain-rate range while maintaining physical consistency. Crucially, the model achieves genuine mechanism-level interpretability by explicitly predicting the internal state variables (dislocation density and TVF), successfully capturing the distinct dislocation accumulation regimes and the dynamic transition of plasticity mechanisms - from dislocation-dominated to twinning-assisted - under varying loading conditions. The proposed framework is readily extendable to other alloy systems and loading paths, offering a highly efficient microstructure-mechanics coupled alternative for computationally intensive crystal-plasticity simulations in multi-scale modeling and process design.

DECLARATIONS

Authors' contributions

Writing - original draft: Dai, W.; Sun, X.; Liu, Y.

Writing - review & editing: Dai, W.; Sun, X.

Visualization: Dai, W.; Sun, X.

Validation: Dai, W.

Software: Dai, W.; Sun, X.; Zhou, K.; Wang, H.

Data curation: Dai, W.; Wang, Y.

Methodology: Dai, W.; Zhou, K.

Investigation: Dai, W.; Sun, X.; Wang, Y.

Formal analysis: Dai, W.

Conceptualization: Dai, W.; Sun, X.; Wang, H.; Liu, Y.

Supervision: Sun, X.; Liu, Y.

Resources: Liu, Y.

Project administration: Liu, Y.

Funding acquisition: Liu, Y.

Availability of data and materials

The raw data supporting the conclusions of this article will be made available by the authors on request. The core source code for the proposed MIDLCM framework is openly available in GitHub at <https://github.com/cwcaiken/MIDL-CM-Mechanics-Informed-Deep-Learning-Constitutive-Model>.

AI and AI-assisted tools Statement

During the preparation of this work the authors used Gemini 3 Pro just in order to rectify latent grammar mistakes and enhance the readability of this article. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Financial support and sponsorship

This work is supported by the 2020-JCJQ project (GFJQ2126-007). Yue Liu also acknowledges the support from the State Key Laboratory of Metal Matrix Composites at Shanghai Jiao Tong University, and the State Key Laboratory for Mechanical Behavior of Materials at Xi'an Jiao Tong University.

Conflicts of interest

Liu, Y. is Editor in Editorial Board Member of the journal *Microstructures*. Liu, Y. was not involved in any steps of editorial processing, notably including reviewers' selection, manuscript handling and decision making. The other authors declare that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

Supplementary Materials

[Supplementary Materials](#)

REFERENCES

1. Tomé, C. N.; Lebensohn, R. A. Material modeling with the visco-plastic self-consistent (VPSC) approach: theory and practical applications; Elsevier, 2023. DOI

2. He, L.; Li, Y.; Torrent, D.; Zhuang, X.; Rabczuk, T.; Jin, Y. Machine learning assisted intelligent design of meta structures: a review. *Microstructures* **2023**, *3*, 2023034. DOI
3. Miao, B.; Lin, G.; Zhang, Y.; Cheng, Y.; Yang, H. Machine learning applications in metallic materials: Recent advances and future perspectives. *J. Alloys. Compd.* **2026**, *1062*, 187486. DOI
4. Li, W.; Xie, D.; Li, D.; Zhang, Y.; Gao, Y.; Liaw, P. K. Mechanical behavior of high-entropy alloys. *Prog. Mater. Sci.* **2021**, *118*, 100777. DOI
5. Wang, Y.; Jiao, Z.; Bian, G.; et al. Dynamic tension and constitutive model in Fe₄₀Mn₂₀Cr₂₀Ni₂₀ high-entropy alloys with a heterogeneous structure. *Mat. Sci. Eng. A-Struct.* **2022**, *839*, 142837. DOI
6. Zhang, Z.; Sheng, H.; Wang, Z.; et al. Dislocation mechanisms and 3D twin architectures generate exceptional strength-ductility-toughness combination in CrCoNi medium-entropy alloy. *Nat. Commun.* **2017**, *8*, 14390. DOI PubMed PMC
7. Dowding, I.; Schuh, C. A. Metals strengthen with increasing temperature at extreme strain rates. *Nature* **2024**, *630*, 91-5. DOI PubMed PMC
8. Borodin, E. N.; Gruzdkov, A. A.; Mayer, A. E.; Selyutina, N. S. Physical nature of strain rate sensitivity of metals and alloys at high strain rates. *J. Phys.: Conf. Ser.* **2018**, *991*, 012012. DOI
9. Fan, H.; Wang, Q.; El-awady, J. A.; Raabe, D.; Zaiser, M. Strain rate dependency of dislocation plasticity. *Nat. Commun.* **2021**, *12*, 1845. DOI PubMed PMC
10. Liang, Z.; Wang, X.; Huang, W.; Huang, M. Strain rate sensitivity and evolution of dislocations and twins in a twinning-induced plasticity steel. *Acta. Mater.* **2015**, *88*, 170-9. DOI
11. Li, C.; Cao, F.; Chen, Y.; Wang, H.; Dai, L. Crystal plasticity model analysis of the effect of short-range order on strength-plasticity of medium entropy alloys. *Metals* **2022**, *12*, 1757. DOI
12. Ma, E. Unusual dislocation behavior in high-entropy alloys. *Scr. Mater.* **2020**, *181*, 127-33. DOI
13. Chandan, A. K.; Tripathy, S.; Ghosh, M.; Chowdhury, S. G. Evolution of substructure of a non-equiatomic FeMnCrCo high entropy alloy deformed at ambient temperature. *Metall. Mater. Trans. A.* **2019**, *50*, 5079-90. DOI
14. Choudhuri, D.; Gwalani, B.; Gorse, S.; et al. Enhancing strength and strain hardenability via deformation twinning in fcc-based high entropy alloys reinforced with intermetallic compounds. *Acta. Mater.* **2019**, *165*, 420-30. DOI
15. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, September, 2014; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014, pp 1724-34. DOI
16. Mozaffar, M.; Bostanabad, R.; Chen, W.; Ehmann, K.; Cao, J.; Bessa, M. A. Deep learning predicts path-dependent plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 26414-20. DOI PubMed PMC
17. Hu, C.; Martin, S.; Dingreville, R. Accelerating phase-field predictions via recurrent neural networks learning the microstructure evolution in latent space. *Comput. Methods. Appl. Mech. Eng.* **2022**, *397*, 115128. DOI
18. Bonatti, C.; Mohr, D. On the importance of self-consistency in recurrent neural network models representing elasto-plastic solids. *J. Mech. Phys. Solids.* **2022**, *158*, 104697. DOI
19. Heidenreich, J. N.; Mohr, D. Recurrent neural network plasticity models: Unveiling their common core through multi-task learning. *Comput. Methods. Appl. Mech. Eng.* **2024**, *426*, 116991. DOI
20. Lim, B.; Zohren, S. Time-series forecasting with deep learning: a survey. *Phil. Trans. R. Soc. A.* **2021**, *379*, 20200209. DOI
21. Vaswani, A.; Shazeer, N.; Parmar, N.; et al. Attention is All you Need. In *Advances in Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, 2017; Vol. 30, pp 5998-6008. DOI
22. Borkowski, L.; Sorini, C.; Chattopadhyay, A. Recurrent neural network-based multiaxial plasticity model with regularization for physics-informed constraints. *Comput. Struct.* **2022**, *258*, 106678. DOI
23. Liu, X.; He, J.; Huang, S. Mechanistically informed artificial neural network model for discovering anisotropic path-dependent plasticity of metals. *Mater. Des.* **2023**, *226*, 111697. DOI
24. Eghtesad, A.; Tan, J.; Fuhg, J. N.; Bouklas, N. NN-EVP: A physics informed neural network-based elasto-viscoplastic framework for predictions of grain size-aware flow response. *Int. J. Plast.* **2024**, *181*, 104072. DOI
25. Long, X.; Steve, N. T. N. A.; Wahaajuddin, K. K. Physics-constrained machine learning for constitutive modelling of high-strength concrete across strain rates. *Eng. Appl. Artif. Intell.* **2026**, *171*, 114314. DOI
26. Wang, K.; Jin, X.; Zhang, Y.; Liaw, P. K.; Qiao, J. Dynamic tensile mechanisms and constitutive relationship in CrFeNi medium entropy alloys at room and cryogenic temperatures. *Phys. Rev. Mater.* **2021**, *5*, 113608. DOI
27. Ibragimova, O.; Brahme, A.; Muhammad, W.; Lévesque, J.; Inal, K. A new ANN based crystal plasticity model for FCC materials and its application to non-monotonic strain paths. *Int. J. Plast.* **2021**, *144*, 103059. DOI
28. Hou, J.; Qiao, J.; Lian, J.; Liaw, P. K. Revealing the relationship between microstructures, textures, and mechanical behaviors of cold-rolled Al_{0.1}CoCrFeNi high-entropy alloys. *Mat. Sci. Eng. A-Struct.* **2021**, *804*, 140752. DOI

29. Zecevic, M.; Cawkwell, M.; Ramos, K.; Luscher, D. Crystal plasticity including a phase-field deformation twinning model for the high-rate deformation of cyclotetramethylene tetranitramine. *J. Mech. Phys. Solids*. **2022**, *163*, 104872. DOI
30. Wu, K.; Sun, X.; Liu, B.; et al. A multiscale investigation into the electroplastic effects in copper: Experiments and crystal plasticity modeling. *J. Mech. Phys. Solids*. **2026**, *212*, 106597. DOI
31. Sun, X.; Zhou, K.; Liu, C.; et al. A crystal plasticity based strain rate dependent model across an ultra-wide range. *Int. J. Plast.* **2024**, *180*, 104056. DOI
32. Orowan, E. Problems of plastic gliding. *Proc. Phys. Soc.* **1940**, *52*, 8-22. DOI
33. Beyerlein, I.; Tomé, C. A dislocation-based constitutive law for pure Zr including temperature effects. *Int. J. Plast.* **2008**, *24*, 867-95. DOI
34. Zinovev, A.; Terentyev, D.; Dubinko, A.; Delannay, L. Constitutive law for thermally-activated plasticity of recrystallized tungsten. *J. Nucl. Mater.* **2017**, *496*, 325-32. DOI
35. Hall, E. O. The deformation and ageing of mild steel: III discussion of results. *Proc. Phys. Soc. B*. **1951**, *64*, 747-53. DOI
36. Petch, N. J. The cleavage strength of polycrystals. *J. Iron. Steel. Inst.* **1953**, *174*, 25-28. DOI
37. Conrad, H.; Feuerstein, S.; Rice, L. Effects of grain size on the dislocation density and flow stress of niobium. *Mater. Sci. Eng.* **1967**, *2*, 157-68. DOI
38. Ashby, M. F. The deformation of plastically non-homogeneous materials. *Philos. Mag.* **2006**, *21*, 399-424. DOI
39. Cordero, Z. C.; Knight, B. E.; Schuh, C. A. Six decades of the Hall-Petch effect - a survey of grain-size strengthening studies on pure metals. *Int. Mater. Rev.* **2016**, *61*, 495-512. DOI
40. Zhu, T.; Li, J.; Samanta, A.; Kim, H. G.; Suresh, S. Interfacial plasticity governs strain rate sensitivity and ductility in nanostructured metals. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 3031-6. DOI PubMed PMC
41. Beyerlein, I.; McCabe, R.; Tomé, C. Effect of microstructure on the nucleation of deformation twins in polycrystalline high-purity magnesium: A multi-scale modeling study. *J. Mech. Phys. Solids*. **2011**, *59*, 988-1003. DOI
42. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. In *Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)*, New Orleans, LA, USA, May 6-9, 2019; Curran Associates, Inc.: Red Hook, NY, USA, 2019, pp. 4061-78. Available from: <https://openreview.net/pdf?id=Bkg6RiCqY7> [accessed on 2026-6-2].
43. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna. In *KDD '19: The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Anchorage AK USA; ACM: New York, NY, USA, 2019, pp 2623-31. DOI
44. Long, X.; Su, T.; Lu, C.; Wang, S.; Huang, J.; Chang, C. An insight into dynamic properties of SAC305 lead-free solder under high strain rates and high temperatures. *Int. J. Impact. Eng.* **2023**, *175*, 104542. DOI
45. Long, X.; Shen, Z.; Li, J.; et al. Size effect of nickel-based single crystal superalloy revealed by nanoindentation with low strain rates. *J. Mater. Res. Technol.* **2024**, *29*, 2437-47. DOI
46. Long, X.; Hu, Y.; Su, T.; et al. Thermomechanical constitutive behaviour of 3D printed biomimetic polymer material under high strain rates. *Polym. Test.* **2024**, *134*, 108439. DOI
47. Samaniego, E.; Anitescu, C.; Goswami, S.; et al. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Comput. Methods. Appl. Mech. Eng.* **2020**, *362*, 112790. DOI
48. Goswami, S.; Anitescu, C.; Chakraborty, S.; Rabczuk, T. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. *Theor. Appl. Fract. Mech.* **2020**, *106*, 102447. DOI

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.



© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.