Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.079587

ARTICLE

A Graph-Based Spatio-Temporal Attention Network for Stress–Strain Behavior Prediction of Copper-Based Composites

Chuhan Zhang[1], Jinguo You[1,*], Jialin Xu[1], Mingqian Li[1], Xiaofeng Chen[2], Jingmei Tao[2], Caiju Li[2] and Jianhong Yi[2]

[1]Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, China
[2]Faculty of Materials Science and Engineering, Kunming University of Science and Technology, Kunming, China

*Corresponding Author: Jinguo You. Email: jgyou@126.com

Received: 24 January 2026; Accepted: 19 March 2026; Published: 08 May 2026

ABSTRACT: With the rapid development of artificial intelligence and data-driven modeling, deep learning has become an effective tool for analyzing scientific discovery such as predicting material behaviors. Graphene-reinforced copper-based composites, which exhibit excellent mechanical, electrical, and thermal properties, have attracted extensive attention in advanced engineering applications; however, accurate prediction of their stress–strain behavior still relies heavily on computationally expensive molecular dynamics simulations or experiments. In this work, we propose a Graph-based Spatio-Temporal Attention Network, termed GraphSTAN, for stress–strain behavior prediction of copper-based composites. Specifically, atomic-scale initial microstructures are encoded as graphs and integrated with static physical parameters. A topology-aware spatio-temporal feature interaction mechanism is introduced to effectively couple structural representations with stress–strain time-series dynamics, enabling accurate prediction of full stress–strain evolution. Moreover, a multi-features dataset is constructed based on LAMMPS molecular dynamics simulations, consisting of 596 independent simulation samples corresponding to distinct combinations of microstructural configurations, loading conditions, and stress–strain time series. Experimental results demonstrate that GraphSTAN effectively predicts full stress–strain curves and achieves the higher performance of R^2, MAE and RMSE for yield strength and Young's modulus, respectively, significantly outperforming baseline methods.

KEYWORDS: Temporal convolutional network; graph attention network; copper-based composites; mechanical property prediction; stress–strain behavior

1 Introduction

In recent years, the rapid development of artificial intelligence has provided new paradigms for the design and performance prediction of complex engineering materials. By integrating data-driven modeling with physical insights, intelligent models can efficiently uncover intrinsic relationships among material composition, microstructure, and macroscopic properties under limited experimental or simulation data, thereby reducing research and development costs and shortening design cycles. Copper and copper alloys, owing to their excellent plastic formability and reliable service performance, have been widely used in mechanical manufacturing, railway transportation, and aerospace applications [1]. However, as service conditions continue to evolve toward high-temperature and heavy-load environments, the inherent limitations of conventional copper-based materials—particularly in terms of strength and deformation resistance—have become increasingly pronounced. To overcome these limitations, the introduction of high-performance

nanoscale reinforcements has emerged as an effective strategy. Among them, graphene, since its successful isolation in 2004 [2], has demonstrated exceptional potential for strengthening copper-based composites due to its outstanding mechanical properties [3–5].

Despite the promising application prospects of graphene-reinforced copper-based composites in high-end equipment, the evaluation of their mechanical properties still primarily relies on experimental testing or high-fidelity molecular dynamics simulations. These approaches are typically time-consuming and computationally expensive, which to some extent constrains material design efficiency [6]. Therefore, accurate prediction of key mechanical properties prior to material synthesis is of great significance. The core objective of material property prediction lies in establishing effective mappings between material composition, microstructural characteristics, and macroscopic mechanical responses. In this regard, considerable progress has been achieved within machine learning and deep learning frameworks [7–9].

Hasan et al. [10] employed various machine learning models, including artificial neural networks (ANN), k-nearest neighbors (KNN), random forests (RF), support vector machines (SVM), and gradient boosting machines (GBM), to predict the wear rate and friction coefficient of graphene-reinforced aluminum matrix composites. Liu et al. [11] combined molecular dynamics simulation data with ANN, SVM, and AdaBoost models to predict the Young's modulus and ultimate tensile strength of graphene/aluminum nanocomposites. Gao et al. [12] integrated long short-term memory networks (LSTM) and backpropagation neural networks (BP) to learn stress–strain evolution features from molecular dynamics tensile sequences and to predict Young's modulus and yield strength based on static parameters, respectively. In addition, Zhao and Ge [13] proposed a machine learning-assisted multi-objective synergistic optimization framework that simultaneously predicted the mechanical and electrical properties of CNTs/Cu composites under limited experimental data, enabling Pareto optimization and providing a new data-driven approach for the co-design of multiple material properties.

Beyond composition and processing parameters, microstructural characteristics play a decisive role in determining macroscopic mechanical behavior. In recent years, researchers have increasingly incorporated microstructural information—such as grain size, phase distribution, defect types, and their evolution—into performance modeling. Chen et al. [14] developed adaptive material structure descriptors based on graph attention networks (GAT), encoding local atomic environments and topological relationships in a learnable manner for predicting intrinsic material properties. Sun et al. [15] employed convolutional neural networks (CNN) to automatically extract spatial features directly from microstructural images, achieving high-accuracy predictions of local stress fields and mechanical response distributions without explicit manual feature engineering, thereby providing a novel data-driven pathway for microstructure-based mechanical property evaluation. Gao et al. [16] utilized graph neural networks (GNN) to encode crystal structures or atomic graphs, leveraging interatomic topology and local geometric information to predict intrinsic physical properties such as formation energy, elastic constants, and electronic band structures.

Complementing these representation techniques, high-fidelity numerical simulation—specifically the finite element (FE) method—has emerged as a primary pipeline for generating the extensive datasets required by deep learning frameworks. This approach allows for the systematic mapping of microstructural configurations to mechanical responses. For example, Chai et al. [17] utilized FE analysis to generate 2D particulate-reinforced microstructures, which were encoded as binary images to train spatial–temporal deep learning models. Building on this, Saha et al. [18] performed uniaxial tensile simulations using 2D representative volume elements (RVE) under periodic boundary conditions in ABAQUS to construct a multimodal CNN-ANN framework. To further optimize multiscale modeling, Zhou and Semnani [19] leveraged RVE-based FE simulations to generate macroscopic strain sequences for training a Gated Recurrent Unit (GRU)

surrogate model. This model was subsequently integrated into ABAQUS via a FORTRAN UMAT subroutine, enabling multiscale finite element simulations of 2D and 3D boundary value problems.

Although the aforementioned studies have achieved notable progress in predicting the mechanical properties of copper-based composites, several limitations remain. On the one hand, traditional machine learning approaches rely heavily on manually engineered features, making it difficult to fully capture the complex spatial topology of microstructures. On the other hand, while existing deep learning methods are capable of modeling temporal evolution, they still face challenges in jointly capturing microstructural heterogeneity and spatiotemporal feature interactions. Moreover, approaches relying on FE simulations for data generation often encounter prohibitive computational costs in nonlinear path-dependent behaviors, requiring extensive high-fidelity RVE evaluations; furthermore, when surrogate models are integrated into nonlinear FE solvers, the strong dependence of predicted stresses on irregular strain increment sizes—unpredictable during global iterations—can cause convergence difficulties and significant errors. Many approaches also fail to effectively integrate microstructural information with multi-source data such as composition and processing parameters, which limits the accuracy and generalization of the learned mappings between structural features and macroscopic mechanical responses.

To address the aforementioned challenges, this work develops a graph-based spatio-temporal modeling framework, termed GraphSTAN, for predicting the mechanical response of copper-based composites. The proposed approach represents atomic-scale microstructures as standardized material graphs and integrates temporal convolutional networks to capture the synergistic evolution of microstructural features and stress-strain behavior. The scope of this research encompasses the theoretical formulation of the graph-based spatio-temporal attention mechanism, the construction of a comprehensive atomistic dataset via molecular dynamics simulations, and a rigorous evaluation of the model's predictive performance and generalization capability. The remainder of this paper is organized as follows: Section 2 details the architecture, implementation, and theoretical foundations of the GraphSTAN framework; Section 3 describes the construction of the spatio-temporal dataset based on molecular dynamics, followed by an analysis of the model's performance, external validation, and out-of-distribution generalization; finally, Section 4 concludes the paper with a discussion on the implications and future directions of this work.

2 Overall Framework

This section provides an overview of the proposed GraphSTAN architecture, which is designed for atomic-scale mechanical behavior prediction of materials. The model consists of two core modules: a material graph construction and static feature fusion module, and a spatiotemporal feature interaction and stress prediction module.

The material graph construction and static feature fusion module extracts topological information from the initial atomic configuration. In this module, a Graph Attention Network (GAT) [20] is employed to encode atomic-level topological features, which are then fused with multiple static physical parameters to generate conditional features for modulating the temporal prediction network. The spatiotemporal feature interaction and stress prediction module models the strain time series under physical causality constraints and predicts the complete 2000-step stress evolution through a topology-aware spatiotemporal interaction mechanism. Finally, Young's modulus and yield strength are extracted from the predicted stress time series. The overall workflow of GraphSTAN is illustrated in Fig. 1.

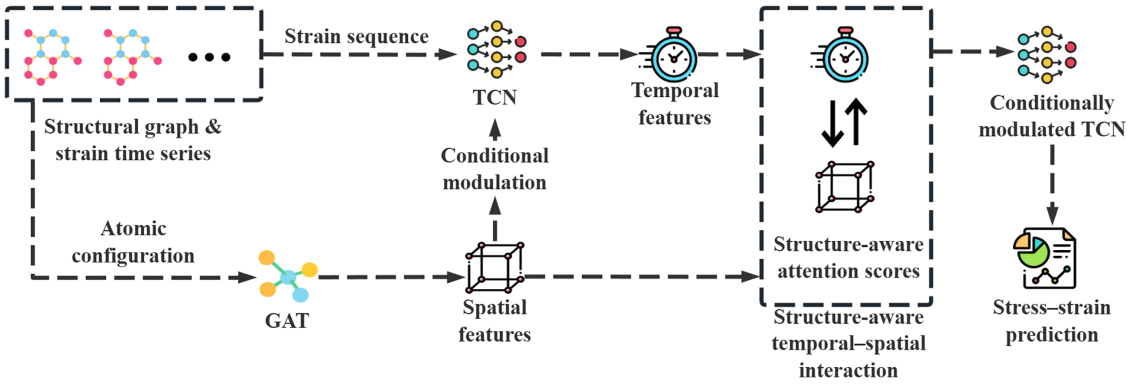

Figure 1: Overview of the proposed GraphSTAN framework. Atomic configurations are encoded as material graphs using a graph attention network, while strain time series are modeled by a conditionally modulated temporal convolutional network. A structure-aware spatiotemporal interaction mechanism enables coupling between structural features and temporal dynamics to predict the full stress–strain response.

2.1 Material Graph Construction and Static Feature Fusion Module

This module is responsible for constructing standardized material graphs from the initial atomic configuration and extracting topological features, while simultaneously fusing static parameters into modulation variables to support subsequent adaptive temporal prediction. The core idea of this module is to employ graph neural networks to process atomic-scale microstructural information and integrate it with global static parameters, thereby forming conditional control signals for the temporal model. This design enhances the adaptability of the prediction framework to diverse microstructures and environmental conditions.

2.1.1 Material Graph Construction and Topological Representation

To accurately characterize the topological features of the initial material microstructure, the atomic configuration is first preprocessed to remove isolated atomic nodes. Interatomic connectivity is then established based on a cutoff radius of 7.736 Å derived from the Lennard–Jones (LJ) potential [21], resulting in a standardized material graph structure. In the constructed graph, nodes correspond to atoms, with node attributes including atomic species and their three-dimensional spatial coordinates. Edges represent interatomic interactions, and the associated edge attributes are composed of interatomic distances and interaction strengths computed from the LJ potential function. Through this procedure, the atomic-scale microstructure is formalized into a topological representation suitable for processing by graph neural networks. Through this procedure, the atomic-scale microstructure is formalized into a topological representation suitable for processing by graph neural networks, as schematically illustrated in Fig. 2.

2.1.2 Topological Feature Encoding

After constructing the standardized material graph, GAT is employed to encode the atomic topological structure. GAT consists of three hidden layers with 256, 128, and 64 neurons, respectively. The attention coefficients are computed using the LeakyReLU activation function, defined as:

$$\mathrm{LeakyReLU}(x) = \begin{cases} x, & x > 0, \\ \alpha x, & x \le 0, \end{cases} \tag{1}$$

where the negative slope is set to $\alpha = 0.2$. The outputs of each hidden layer are further activated by the ReLU function.

$$\mathrm{ReLU}(x) = \max(0, x) \tag{2}$$

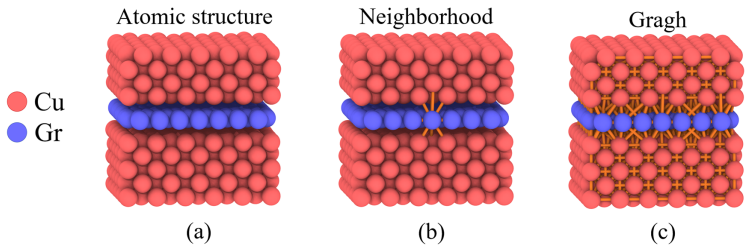

Figure 2: Construction of the material graph from atomic configurations. (**a**) Atomic configuration of the graphene-reinforced copper composite, where Cu and graphene atoms are shown in red and blue, respectively; (**b**) identification of local atomic neighborhoods around the graphene interface based on a cutoff radius; (**c**) graph representation of the material microstructure, where atoms are treated as nodes and interatomic interactions within the cutoff radius are represented as edges.

A dropout layer with a probability of 0.15 is applied after each layer to mitigate overfitting.

Through the self-attention mechanism, the model adaptively assigns different weights to neighboring nodes, enabling effective aggregation of local structural information and extraction of high-dimensional topological features that capture neighbor density, edge-weight distribution, and local structural correlations. It should be noted that, at each layer, GAT first generates node-level topological feature representations, which characterize local atomic environments and their neighborhood relationships. Subsequently, a global average pooling operation is applied to aggregate node-level features into a 64-dimensional graph-level topological representation, which describes the overall microstructural characteristics of the material.

2.1.3 Modulation Vector Generation

After obtaining the topological feature representation of the material graph, global static parameters are further incorporated to enhance the model's adaptability to different material configurations and loading conditions. Specifically, five categories of static parameters—including layer number, temperature, orientation angle, and other material- and loading-related attributes—are mapped through embedding layers and normalized into a 32-dimensional static parameter vector. This vector is then concatenated with the 64-dimensional topological feature output by GAT, resulting in a 96-dimensional fused static feature representation.

To explicitly inject microstructural information into the temporal convolutional network, a weight-modulated dynamic convolution strategy is introduced, enabling structure-aware temporal feature extraction under different atomic topological conditions. Concretely, the 96-dimensional fused static feature is passed through two fully connected layers to generate a 32-dimensional modulation vector, where the first 16 dimensions are used for channel-wise scaling of convolutional kernels, and the remaining 16 dimensions are used for modulation of kernel weight magnitudes.

Let W denote the base convolutional kernel weights of a given convolutional layer in the TCN, and let m represent the corresponding modulation vector. The dynamically modulated convolutional kernel weights are defined as:

$$W' = W \odot (1 + m) \tag{3}$$

where $\odot$ denotes element-wise multiplication along the channel dimension. With this formulation, the proposed method enables continuous and differentiable regulation of the temporal feature extraction process without altering the number of network layers, kernel sizes, or dilation structures. As a result, the receptive field distribution and feature response patterns of the TCN can adaptively adjust to the current material topology and static physical conditions.

2.2 Spatiotemporal Feature Interaction and Stress Prediction Module

2.2.1 Conditional Temporal Modeling

During the temporal modeling stage, a causal mask is applied to the input strain time series to ensure that the prediction process strictly satisfies physical causality constraints, such that the model at any time step can access only the current and historical information. The masked strain sequence is concatenated with the 96-dimensional fused static feature along the temporal dimension to form a conditioned input sequence of length 2000, which is then used for subsequent temporal feature extraction.

Temporal feature extraction is performed using TCN based on one-dimensional convolutions. Under the modulation of the generated conditioning vectors, the convolutional responses of TCN are dynamically adjusted according to the material topology and static conditions, enabling the model to capture the evolution characteristics of stress–strain relationships across multiple temporal scales, including short-, medium-, and long-term dependencies.

2.2.2 Topology-Aware Spatiotemporal Interaction and Stress Prediction

In the spatiotemporal feature interaction stage, both node-level topological features and graph-level structural representations extracted by the GAT are utilized. Specifically, node-level features are employed to localize key spatial regions, while graph-level features are used to modulate the overall temporal modeling behavior. First, the temporal features extracted by the TCN are mapped to node-level spatial attention weights, which are normalized using a softmax function. These attention weights are then multiplied element-wise with the node-level topological features output by the GAT, thereby highlighting key spatial regions that are most relevant to the current dynamic evolution.

Furthermore, the filtered topological features are fed back into the temporal model through a gating mechanism to modulate the activation intensity of temporal features. The gating modulation is defined as:

$$\text{Threshold}_{\text{new}} = \text{Threshold}_{\text{base}} \cdot (1 + \tanh(\phi)) \tag{4}$$

where ϕ denotes the mapped topological features, and $\text{Threshold}_{\text{base}} > 0$ is a base threshold parameter. This mechanism suppresses temporal noise that is irrelevant to the current topological structure while enhancing the model's focus on critical evolution patterns. The spatiotemporal interaction process is repeated twice, and after each interaction, the attention weights and modulation parameters are dynamically optimized via backpropagation.

After completing the spatiotemporal interaction, the model outputs a full 2000-step predicted stress time series. According to classical solid mechanics definitions, Young's modulus is obtained by performing least-squares linear fitting on the predicted stress–strain curve within the small-strain regime ($\varepsilon < 2\%$), while the yield strength is determined using the standard 0.2% offset method. Through this procedure, end-to-end prediction from atomic-scale microstructural information to macroscopic mechanical performance indicators is achieved.

2.3 Theoretical Analysis

Although the GraphSTAN architecture integrates graph attention and spatio-temporal interaction, we provide a theoretical analysis demonstrating that the learned representations are physically consistent with key mechanisms in graphene-reinforced Cu composites, particularly Cu–graphene interfacial stress concentration and yield initiation sites. The material graph is constructed using the same Lennard–Jones (LJ) cutoff radius r_c as in MD simulations [22], where pairwise interactions follow the potential E_{ij}:

$$E_{ij} = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{5}$$

the resulting interatomic forces $\mathbf{F}_{ij}$ and virial stress contributions are inherently strongest at the Cu–graphene interface due to lattice mismatch and misfit dislocations. In the GAT module, attention coefficients α_{ij} characterize the importance of neighboring atoms to a central atom, which are computed based on the self-attention mechanism:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\mathbf{a}^{\top}[\mathbf{W}\mathbf{h}_i|\mathbf{W}\mathbf{h}j]\right)\right)}{\sum k \in \mathcal{N}(i) \exp\left(\text{LeakyReLU}\left(\mathbf{a}^{\top}[\mathbf{W}\mathbf{h}_i|\mathbf{W}\mathbf{h}_k]\right)\right)} \tag{6}$$

gradient-based optimization of the stress-prediction loss drives α_{ij} to correlate with regions of elevated local strain energy density. Consequently, high-attention nodes naturally emphasize **Cu–graphene interfacial regions** and **defect-sensitive zones** where misfit strains induce localized strengthening, consistent with the atomistic deformation observed in the subsequent results (see Fig. 3c,d in Section 3). In the spatio-temporal interaction stage, temporal features $\mathbf{f}_t$ are mapped to node-level spatial attention $\boldsymbol{\beta}^{(t)}$ to yield structure-aware representations $\mathbf{h}_{\text{ST}}^{(t)}$:

$$\mathbf{h}\text{ST}^{(t)} = \boldsymbol{\beta}^{(t)} \odot \mathbf{h}\text{GAT} \tag{7}$$

the temporal sensitivity is further modulated by the gating mechanism $\text{Threshold}_{\text{new}}$ defined in Eq. (4), which adjusts the activation threshold based on the global topology ϕ:

$$\phi = \text{MLP}(\mathbf{h}_{\text{GAT}}) \tag{8}$$

in the elastic regime ($\varepsilon < 2\%$), $\boldsymbol{\beta}^{(t)}$ remains relatively uniform, reflecting a homogeneous matrix response. However, near yield ($\varepsilon \approx 0.12$–0.15), nonlinearity in the strain sequence amplifies the attention weights β_k for interface-adjacent nodes, as the training objective aligns the model with MD trajectories where yield initiation is dominated by interfacial dislocation nucleation. Specifically, by applying the chain rule, we derive the gradient sensitivity of the predicted stress $\hat{\sigma}(t)$ with respect to node features $\mathbf{h}_k$:

$$\frac{\partial\hat{\sigma}(t)}{\partial\mathbf{h}k} \propto \beta_k^{(t)} \cdot \frac{\partial\mathcal{F}(\mathbf{h}\text{ST}^{(t)})}{\partial\mathbf{h}_{\text{GAT}}} \tag{9}$$

this relationship implies that GraphSTAN dynamically prioritizes yield initiation sites at the interface by amplifying $\beta_k^{(t)}$ while suppressing bulk Cu contributions that are irrelevant to the nonlinear regime. This mechanistic sensitivity is validated by the high agreement with MD stress–strain curves presented later in Section 3.

3 Experiment

3.1 Dataset

Molecular dynamics (MD) simulation is a well-established and widely adopted approach for investigating the properties of nanomaterials, particularly the mechanical behavior of graphene/metal matrix composites [23]. In this study, atomic-scale models of graphene-reinforced copper-based composites were constructed using the LAMMPS [24] simulation package. For all simulations, the simulation box size was fixed at 6 nm × 6 nm × 6 nm, and graphene sheets with a square geometry were embedded into the copper matrix.

To construct the dataset, molecular dynamics (MD) simulations were performed to obtain stress–strain responses of graphene-reinforced copper composites under different structural and thermal conditions. Five variables were considered, including graphene distribution mode (layered and random), temperature, number of graphene layers (1–3), orientation angle (0°–30°), and chirality (zigzag and armchair). For layered distributions, these structural parameters yield 36 distinct configurations. A hierarchical temperature sampling strategy was adopted to capture temperature-dependent behavior while maintaining computational feasibility. Seven representative temperatures (1, 100, 200, 300, 400, 500, and 600 K) were selected, and full structural combinations were simulated at these temperatures (252 datasets). For intermediate temperatures within 1–600 K, only 3–5 representative structural configurations were simulated to capture continuous temperature-dependent trends, since structural variability had already been covered at the key temperatures. In addition, a limited number of random distribution configurations were included to provide structural diversity. In total, 630 MD simulations were performed to construct the initial dataset.

The Cu–Cu interactions are described using the embedded atom method (EAM) potential [25], while the C–C interactions are modeled using the AIREBO potential [26] with a cutoff distance of 2.0 Å. The Cu–C interactions are represented by the Lennard–Jones potential with a well depth of ε = 0.0875 eV and a distance parameter of σ = 3.0945 Å.

The thickness of a single graphene layer is set to 0.34 nm, and the integration time step is 0.001 ps. The simulation procedure is as follows. First, the system is relaxed under the NPT ensemble for 50 ps to eliminate initial residual stresses. Subsequently, the ensemble is switched to NVT, and uniaxial tensile loading is applied at a strain rate of 0.001 ps for a duration of 300 ps. All simulation trajectories are visualized and analyzed using OVITO [27], and a representative tensile schematic is shown in Fig. 3.

Fig. 4 presents the stress–strain curves of monolayer graphene and armchair-oriented copper-based composites with a 0° orientation angle at different temperatures. It can be observed that, with increasing temperature, both the yield strength and Young's modulus decrease noticeably.

During data processing, samples with anomalous stress–strain responses were removed based on objective physical and numerical criteria, including (i) numerical instability with temperature deviation beyond ±10 K, (ii) unphysical structural failure during relaxation, and (iii) discontinuous stress–strain curves without a clear elastic regime. A total of 34 samples were excluded, resulting in 596 valid datasets. The remaining data were normalized and randomly divided into training, validation, and test sets with a ratio of 8:1:1.

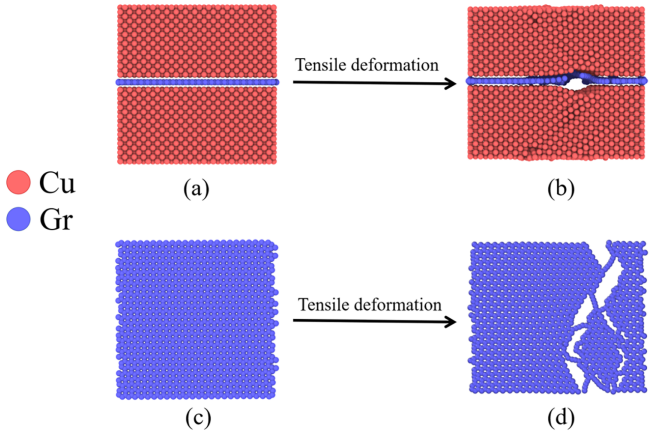

Figure 3: Initial atomic configurations and tensile deformation behavior of the Cu/graphene composite with layered distribution, single-layer graphene, 0° orientation angle, and armchair chirality. (**a**) Initial atomic configuration of the Cu/graphene composite with an embedded graphene layer; (**b**) atomic configuration of pristine monolayer graphene; (**c**) interfacial deformation and damage initiation at the Cu-graphene interface during tensile deformation; (**d**) fracture morphology of graphene after tensile deformation.

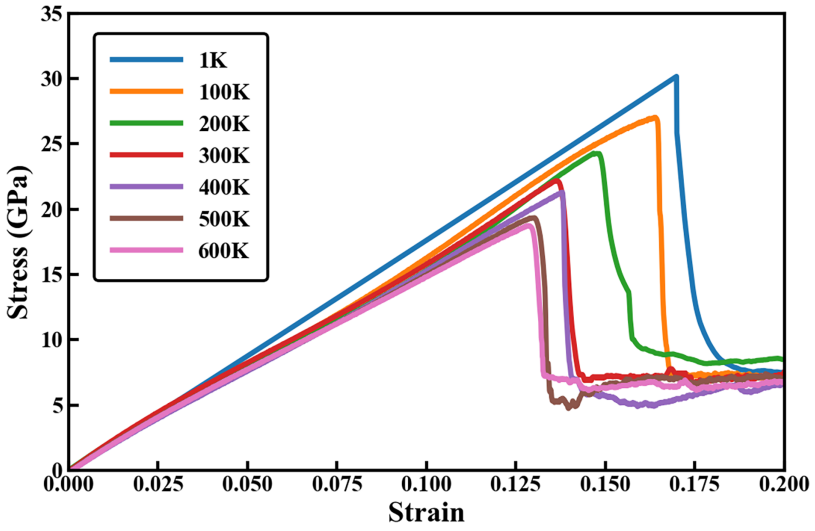

Figure 4: stress–strain responses at different temperatures (1–600 K). The curves exhibit an approximately linear elastic regime followed by a peak stress and subsequent softening behavior. As the temperature increases, both the peak stress and elastic modulus decrease, indicating temperature-induced mechanical degradation. The post-peak stress drop reflects structural failure and plastic deformation, with higher temperatures leading to earlier failure and reduced load-bearing capacity.

3.2 Performance Evaluation Metrics

To reflect the gap between predicted results and actual values, appropriate performance metrics need to be introduced. The coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) are commonly used metrics for evaluating model prediction performance. R^2 provides an overall perspective on the model's fit, while RMSE, MSE, and MAE provide specific quantitative indicators of model prediction error. RMSE and MSE emphasize the impact of large errors, while MAE provides a more intuitive measure of average error. By comprehensively considering these metrics, the performance of the

regression model can be evaluated more comprehensively. The specific calculation formulas are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{12}$$

where n denotes the total number of test samples, y_i and $\hat{y}_i$ represent the ground-truth value and the corresponding predicted value of the i-th sample, respectively, and $\bar{y}$ denotes the mean of all ground-truth values.

3.3 Results

In this study, the overall prediction accuracy of the stress–strain curves is adopted as the primary criterion for model training and selection. Model performance is evaluated on the validation set using the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). After model convergence, Young's modulus and yield strength are further extracted from the predicted stress–strain curves, and their prediction accuracy is assessed on the test set to verify the effectiveness of the model in predicting key mechanical parameters. In addition, all models are trained and tested using the same training and test datasets, and their performances are quantitatively compared using the three evaluation metrics mentioned above.

As shown in Fig. 5, the red curves represent the model predictions, while the black curves correspond to the results obtained from molecular dynamics simulations. A high level of agreement is observed between the predicted and simulated curves, particularly at critical extrema and regions where the trend changes. These results indicate that the proposed model can accurately capture the complex nonlinear relationships between microstructural features and mechanical responses, demonstrating its robustness and accuracy in structure-aware mechanical property prediction.

Table 1 summarizes the performance of the proposed method in comparison with several competitive baseline approaches for yield strength prediction on the copper-based composite dataset. A temporal Transformer model was implemented as an additional baseline to evaluate state-of-the-art sequence modeling capability. The model consists of a two-layer Transformer encoder with multi-head self-attention, followed by a fully connected layer for stress prediction. The model was trained under the same data splits and optimization settings as GraphSTAN to ensure a fair comparison. Compared with the LSTM model, STAGNet achieves reductions of 42.30% and 37.41% in MAE and RMSE, respectively. Although the Transformer captures long-range dependencies in the strain sequence, the absence of explicit structural interaction limits its predictive performance compared with GraphSTAN. In addition, conventional machine learning models generally exhibit inferior performance compared with deep learning–based methods for material property prediction tasks. Among them, the SVM model performs the worst, with an RMSE of 0.9984 and an R^2 value of 0.9037. These results clearly demonstrate that, compared with traditional machine learning models and deep learning approaches that do not adequately account for both temporal and spatial characteristics, the proposed GraphSTAN model achieves substantially higher prediction accuracy in composite material property prediction.

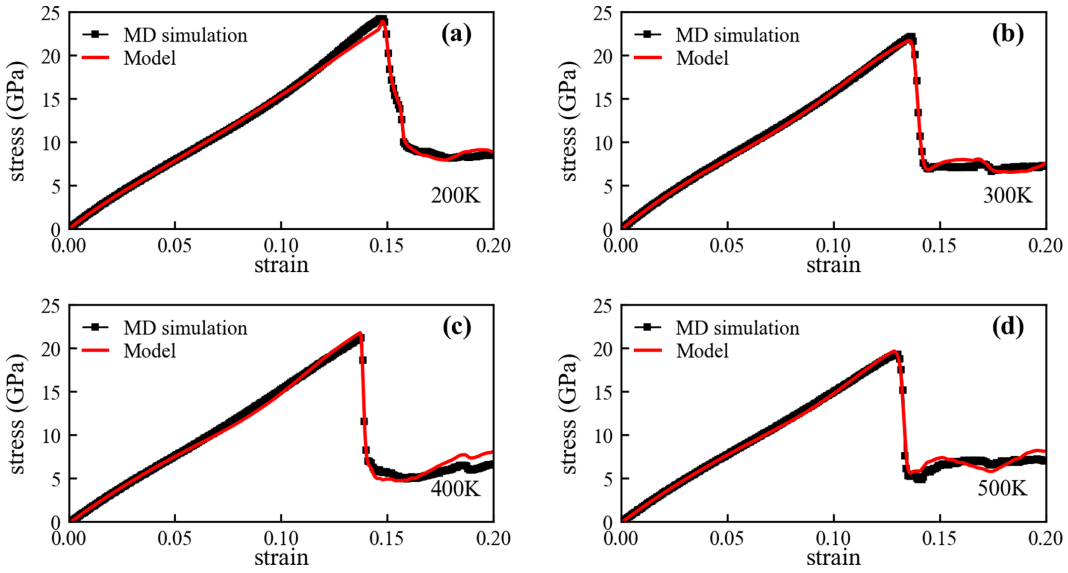

Figure 5: Comparison of predicted and MD-simulated stress–strain curves at different temperatures for a layered distribution with a single-layer structure, 0° orientation angle, and armchair chirality: (**a**) 200 K, (**b**) 300 K, (**c**) 400 K, and (**d**) 500 K. The red curves represent the model predictions, and the black curves denote the MD simulation results.

Table 1: Comparison of performance metrics for yield strength prediction across different models.

ML Model	MAE	RMSE	R^2 Value
RF	0.4683	0.6690	0.9627
SVM	0.7125	0.9984	0.9037
LSTM	0.5645	0.8065	0.9458
BP	0.4677	0.6681	0.9628
Transformer	0.3964	0.5877	0.9698
GraphSTAN	0.3257	0.5048	0.9792

Table 2 reports the performance of the proposed method and various baseline models for Young's modulus prediction on the copper-based composite dataset. Compared with conventional machine learning models and single temporal modeling approaches, the overall prediction error is further reduced. These results indicate that the proposed model is capable of accurately capturing the nonlinear mechanical behavior near the yield stage, while also exhibiting strong modeling capability in describing the stress–strain relationship within the elastic regime. This observation further validates the necessity of jointly modeling atomic-scale microstructural topological information and temporal mechanical responses for reliable mechanical property prediction.

Table 2: Comparison of performance metrics for Young's modulus prediction across different models.

ML Model	MAE	RMSE	R^2 Value
RF	0.9326	1.4348	0.9634
SVM	1.4179	2.1815	0.9154
LSTM	0.9466	1.4562	0.9623

(Continued)

Table 2 (continued)

ML Model	MAE	RMSE	R^2 Value
BP	0.8172	1.2572	0.9719
Transformer	0.7810	1.0258	0.9754
GraphSTAN	0.5473	0.8420	0.9874

3.3.1 Effect of Temperature Sampling Granularity

In this experiment, two temperature sampling strategies are compared. Discrete-7T refers to a temperature sampling strategy in which training data are generated at seven discrete temperature points, namely 1, 100, 200, 300, 400, 500, and 600 K. In contrast, Dense-121T denotes a dense temperature sampling strategy that employs a finer temperature discretization, covering the 1–600 K range with a uniform 5 K interval, resulting in 121 temperature points. Except for the temperature sampling strategy, all other experimental settings, including network architecture, training protocol, and data split, are kept identical.

As reported in Table 3, the model trained with Dense-121T consistently achieves lower MAE and RMSE together with higher R^2 values in the prediction of Young's modulus compared with its Discrete-7T counterpart. This improvement indicates that denser temperature coverage enables more accurate learning of the temperature-dependent elastic response.

Table 3: Performance comparison of LSTM and GraphSTAN under different temperature sampling strategies.

ML Model	MAE	RMSE	R^2 Value
LSTM (Discrete-7T)	1.5257	2.2112	0.9030
LSTM (Dense-121T)	0.8466	1.3562	0.9623
GraphSTAN (Discrete-7T)	0.8927	1.3464	0.9142
GraphSTAN (Dense-121T)	0.5473	0.8420	0.9874

3.4 External Validation

To establish the physical realism and external validity of the GraphSTAN framework, we performed a comparative analysis against independent benchmarks and established theoretical limits. Given the inherent challenges of direct atomic-scale experimental validation, the reliability of our model is evaluated based on its consistency with high-fidelity MD studies and its ability to capture the intrinsic mechanical trends of graphene-reinforced copper composites.

As a prerequisite for interpreting these results, it is essential to distinguish between the idealized atomic-scale response and macroscopic engineering behavior. The yield strengths predicted in this study represent the intrinsic theoretical limits of defect-free graphene/copper nanostructures. Unlike bulk-scale experiments, MD simulations model localized regions devoid of stochastic structural defects such as dislocations, grain boundaries, or impurities. Consequently, these values act as an upper-bound performance limit, naturally exceeding engineering-scale measurements due to the pronounced nanoscale size effect and the absence of macroscopic fracture-initiating sites.

To ensure a rigorous quantitative reliability, our model's predictions were benchmarked against an independent MD study by Huang et al. [28]. Crucially, the validation was conducted under a strictly controlled environment, employing identical interatomic potential functions (EAM, AIREBO, and LJ) as the

benchmark study. Furthermore, all structural and loading parameters, including the number of graphene layers, crystallographic orientation, chirality, and applied strain rate, were maintained consistent with the reference study. As summarized in Table 4, GraphSTAN predicts an intrinsic yield strength of 21.79 GPa at 300 K, which closely aligns with the 23.0 GPa reported in the reference study. Furthermore, the model accurately reproduces the systematic thermal softening observed in benchmark simulations, where the yield strength decreases from 30.69 GPa at 1 K to 17.84 GPa at 600 K. This qualitative agreement confirms that the framework effectively captures the temperature-dependent deformation mechanisms, including thermally activated lattice vibrations and interface weakening.

Table 4: Comparison of predicted yield strength with an independent MD study.

Study	Method	Temperature (K)	Yield Strength (GPa)
Huang et al. (2025) [28]	MD	300	23.0
GraphSTAN	Graph-based ML + MD	300	21.79

The physical fidelity of the underlying dataset is governed by the selected interatomic potentials (EAM, AIREBO, and LJ), which are recognized benchmarks for the Cu–Graphene system. However, a critical limitation of such empirical potentials is their phenomenological nature; they may not fully resolve complex electronic-level interactions, such as interfacial charge redistribution or dynamic bond hybridization, which are more accurately (though at a vastly higher cost) described by density functional theory (DFT).

Additionally, the implications for real-scale mechanical behavior are constrained by the temporal discrepancy between MD simulations and experimental conditions. The high strain rates required in atomistic modeling (typically 10^7 to 10^{11} s^{-1}) may suppress long-term relaxation processes, potentially leading to an overestimation of flow stress compared to quasi-static testing. Therefore, GraphSTAN should be viewed as a high-fidelity surrogate engine designed to identify structure-property trends and reinforcement mechanisms. Its primary utility lies in providing a foundational roadmap for multiscale material design rather than serving as a standalone substitute for macroscopic experimental characterization.

In conclusion, the consistency between GraphSTAN's predictions and established literature suggests that the model learns physically meaningful patterns rather than simple empirical correlations. Future work will bridge these atomistic insights with multiscale modeling and experimental validation to further explore the model's performance under realistic service environments.

3.5 Out-of-Distribution Generalization Analysis

To evaluate the generalization capability of the proposed GraphSTAN model beyond the training distribution, an out-of-distribution (OOD) extrapolation test was conducted. The model was evaluated at an unseen temperature of 700 K, which exceeds the training range (1–600 K). Independent molecular dynamics (MD) simulations were performed at 700 K to provide reference data, and the pre-trained model was used to predict the corresponding mechanical responses.

Table 5 presents the comparison between MD results and model predictions. The model achieves strong predictive performance under this unseen condition, with a coefficient of determination of $R^2 = 0.941$. The predicted Young's modulus and yield strength show good agreement with MD results, and the model correctly captures the temperature-induced softening behavior.

Table 5: Comparison of MD results and GraphSTAN predictions at the unsee temperature of 700 K.

Model	Young's Modulus (GPa)	Yield Strength (GPa)
MD	172.43	16.11
GraphSTAN	161.56	15.48

These results indicate that the proposed model learns physically meaningful thermo-mechanical relationships rather than merely interpolating training data, demonstrating robust predictive capability beyond the training range.

3.6 Ablation Experiments

To verify the effectiveness of individual components in the proposed GraphSTAN framework, ablation experiments are conducted on the copper-based composite dataset for Young's modulus prediction, where the coefficient of determination (R^2) is used to evaluate the contribution of each component to the overall performance. As reported in Table 6, compared with the TCN model, GraphSTAN reduces the MAE and RMSE by 35.7% and 35.6%, respectively. In comparison with the GAT model, the reductions in MAE and RMSE reach 60.05% and 59.01%, respectively.

Table 6: Ablation results for Young's modulus prediction on the copper-based composite dataset.

ML Model	MAE	RMSE	R^2 Value
GAT	1.3701	2.0542	0.9211
TCN	0.8520	1.3108	0.9695
GraphSTAN	0.5473	0.8420	0.9874

These results indicate that the hybrid architecture integrating GAT and TCN achieves substantially better performance than either component alone in predicting the mechanical properties of copper-based composites. The ablation results further demonstrate that the proposed spatiotemporal attention–based interaction mechanism plays a critical role in enhancing predictive accuracy and effectively improves the overall performance of the framework.

4 Conclusion

Accurate prediction of mechanical properties prior to material synthesis is crucial for efficient allocation of simulation and experimental resources. This work addresses the challenges of high computational cost and insufficient structure–temporal coupling through a graph-based spatio-temporal attention network. By encoding atomic-scale microstructures as material graphs, GraphSTAN jointly models microstructural topology and macroscopic tensile behavior. A structure–temporal integrated dataset based on molecular dynamics tensile simulations is constructed for model training and evaluation. Experimental results on 596 molecular dynamics simulation samples demonstrate that GraphSTAN achieves accurate prediction of full stress–strain responses and significantly outperforms baseline methods in predicting Young's modulus and yield strength. This study provides an effective data-driven approach for uncovering structure–property relationships in copper-based composites.

Future work will focus on extending the proposed framework in several key directions. First, the applicability of GraphSTAN to a broader range of composite systems and material classes will be investigated

to evaluate its generalization capability. Second, we plan to explicitly incorporate graphene concentration (volume fraction) as a primary variable. Given its critical role in determining the percolation threshold and load-transfer efficiency in metal matrix composites, exploring the synergistic effects between concentration and structural topology will provide a more comprehensive design space. Third, incorporating additional physical factors, such as strain-rate effects, defect evolution, and damage mechanisms, is expected to further enhance the model's ability to capture complex mechanical responses under diverse loading conditions. Finally, improving model interpretability by analyzing the correlation between learned structural representations and underlying physical mechanisms remains a priority for bridging the gap between machine learning and materials science.

Acknowledgement: Not applicable.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Chuhan Zhang, Jinguo You and Xiaofeng Chen; methodology, Chuhan Zhang, Jinguo You and Xiaofeng Chen; software, Chuhan Zhang and Jialin Xu; validation, Chuhan Zhang, Jinguo You, Jialin Xu, Mingqian Li, Xiaofeng Chen, Jingmei Tao, Caiju Li and Jianhong Yi; formal analysis, Chuhan Zhang; investigation, Jialin Xu and Mingqian Li; data curation, Chuhan Zhang and Mingqian Li; writing—original draft preparation, Chuhan Zhang; writing—review and editing, Chuhan Zhang, Jinguo You, Jialin Xu and Mingqian Li; visualization, Chuhan Zhang; supervision, Jinguo You and Jianhong Yi. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: Data is available upon request from the authors.

Ethics Approval: Not applicable.

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

References

1. Li L, Pan D, Li B, Wu Y, Wang H, Gu Y, et al. Patterns and challenges in the copper industry in China. Resour Conserv Recycl. 2017;127:1–7. doi:10.1016/j.resconrec.2017.07.046.
2. Novoselov KS, Geim AK, Morozov SV, Jiang D, Zhang Y, Dubonos SV, et al. Electric field effect in atomically thin carbon films. Science. 2004;306(5696):666–9. doi:10.1126/science.1102896.
3. Hidalgo-Manrique P, Lei X, Xu R, Zhou M, Kinloch IA, Young RJ. Copper/graphene composites: a review. J Mater Sci. 2019;54(19):12236–89. doi:10.1007/s10853-019-03703-5.
4. Khanna V, Singh K, Kumar S, Bansal SA, Channegowda M, Khalid M, et al. Engineering electrical and thermal attributes of two-dimensional graphene reinforced copper/aluminium metal matrix composites for smart electronics. ECS J Solid State Sci Technol. 2022;11(12):127001. doi:10.1149/2162-8777/aca933.
5. Mutlu A, Çavdar U. Investigation of mechanical properties of Copper-Graphene composites in terms of production methods and additive ratios: a review. Sigma J Eng Nat Sci. 2024;42(2):600–13. doi:10.14744/sigma.2023.00126.
6. Barbhuiya S, Das BB. Molecular dynamics simulation in concrete research: a systematic review of techniques, models and future directions. J Build Eng. 2023;76(3):107267. doi:10.1016/j.jobe.2023.107267.
7. Kimura D, Tajima N, Okazaki T, Muroga S. Explainable multimodal machine learning for revealing structure-property relationships in carbon nanotube fibers. Carbon. 2025;241:120390. doi:10.1016/j.carbon.2025.120390.
8. Liu Y, Kelley KP, Vasudevan RK, Funakubo H, Ziatdinov MA, Kalinin SV. Experimental discovery of structure–property relationships in ferroelectric materials via active learning. Nat Mach Intell. 2022;4(4):341–50. doi:10.1038/s42256-022-00460-0.
9. Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: recent applications and prospects. npj Comput Mater. 2017;3(1):54. doi:10.1038/s41524-017-0056-5.

10. Hasan MS, Wong T, Rohatgi PK, Nosonovsky M. Analysis of the friction and wear of graphene reinforced aluminum metal matrix composites using machine learning models. Tribol Int. 2022;170:107527. doi:10.1016/j.triboint.2022.107527.
11. Liu J, Zhang Y, Zhang Y, Kitipornchai S, Yang J. Machine learning assisted prediction of mechanical properties of graphene/aluminium nanocomposite based on molecular dynamics simulation. Mater Des. 2022;213(3):110334. doi:10.1016/j.matdes.2021.110334.
12. Gao T, Chen L, Wang B, Liu Y, Ma Y, Liang Y. Prediction of the mechanical properties of graphene/copper nanocomposites using machine learning and molecular dynamics simulations. J Phys Condens Matter. 2025;37(23):235701. doi:10.1088/1361-648X/add2c2.
13. Zhao X, Ge C. Machine learning-assisted dual-objective synergistic optimization for mechanical and electrical properties of CNTs/Cu composites. Compos Commun. 2025;59:102583. doi:10.1016/j.coco.2025.102583.
14. Chen J, Zhang J, Wang Z, Han X, Zhang Y. Self-adaptable materials structure descriptor based on graph attention network for machine learning. Mater Des. 2022;223(4):111162. doi:10.1016/j.matdes.2022.111162.
15. Sun Y, Hanhan I, Sangid MD, Lin G. Predicting mechanical properties from microstructure images in fiber-reinforced polymers using convolutional neural networks. J Compos Sci. 2024;8(10):387. doi:10.3390/jcs8100387.
16. Gao H, Guo X-W, Li G, Li C, Yang C. GCPNet: an interpretable generic crystal pattern graph neural network for predicting material properties. Neural Netw. 2025;188(33):107466. doi:10.1016/j.neunet.2025.107466.
17. Chai X, Su Y, Lin Z, Qiu C, Liu X, Zhang X, et al. Mechanical property prediction and configuration effect exploration of particulate reinforced metal matrix composites via an interpretable deep learning approach. Mater Sci Eng A. 2025;925(3):147880. doi:10.1016/j.msea.2025.147880.
18. Saha D, Sun L, Lai CQ. A science directed progressive neural network for multimodal prediction of elastoplastic behavior in composite materials. Compos Part A Appl Sci Manuf. 2025;199:109179. doi:10.1016/j.compositesa.2025.109179.
19. Zhou Y, Semnani SJ. A machine learning based multi-scale finite element framework for nonlinear composite materials. Eng Comput. 2025;41(5):2795–2831. doi:10.1007/s00366-025-02121-3.
20. Veličković P, Cucurull G, Casanova A, Romero A, Lió P, Bengio Y. Graph attention networks. In: Proceedings of the International Conference on Learning Representations; 2018 Apr 30–May 3; Vancouver, BC, Canada.
21. Qanbarian M, Qasemian A, Arab B, Bashiri M, Ebrahiminejad S. Molecular dynamics investigation of argon evaporation on copper surfaces covered with graphene and carbon nanotubes. Mater Today Commun. 2024;38(9):107900. doi:10.1016/j.mtcomm.2023.107900.
22. Nosé S. A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys. 1984;81(1):511–9. doi:10.1063/1.447334.
23. Safina LR, Baimova JA, Krylova KA, Murzaev RT, Mulyukov RR. Simulation of metal-graphene composites by molecular dynamics: a review. Lett Mater. 2020;10(3):351–60. doi:10.22226/2410-3535-2020-3-351-360.
24. Plimpton S. Fast parallel algorithms for short-range molecular dynamics. J Comput Phys. 1995;117(1):1–19. doi:10.1006/jcph.1995.1039.
25. Mishin Y, Mehl MJ, Papaconstantopoulos DA, Voter AF, Kress JD. Structural stability and lattice defects in copper: ab initio, tight-binding, and embedded-atom calculations. Phys Rev B. 2001;63(22):224106. doi:10.1103/PhysRevB.63.224106.
26. Stuart SJ, Tutein AB, Harrison JA. A reactive potential for hydrocarbons with intermolecular interactions. J Chem Phys. 2000;112(14):6472–86. doi:10.1063/1.481208.
27. Stukowski A. Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool. Model Simul Mater Sci Eng. 2010;18(1):015012. doi:10.1088/0965-0393/18/1/015012.
28. Huang M, Fan S, Peng M, Li J, Bu H, Li M. Molecular dynamics study of the effect of temperature and strain rate on the mechanical properties of graphene/copper matrix composites. Mater Today Commun. 2025;43:111716. doi:10.1016/j.mtcomm.2025.111716.