

Revealing structural signatures associated with stress overshoot in two-dimensional Lennard-Jones systems based on interpretable deep learning

Tongtong Shen(申童童), Xinyao Wang(王新尧), Zhuohang Li(李卓杭), Xueyu Liu(刘雪宇), Wen Zheng(郑文)

Citation: Chin. Phys. B. 35, 098301 (2026). DOI: 10.1088/1674-1056/ae843a |
CSTR: 32038.14.CPB.ae843a
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Inverse design of 3D integrated high-efficiency grating couplers using deep learning

Yu Wang(王玉), Yue Wang(王越), Guohui Yang(杨国辉), Kuang Zhang(张狂), Xing Yang(杨星), Chunhui Wang(王春晖), Yu Zhang(张雨)

Chin. Phys. B. 35, 024101 (2026). DOI: 10.1088/1674-1056/adf69c | CSTR: 32038.14.CPB.adf69c

Quantum state estimation based on deep learning

Haowen Xiao(肖皓文), Zhiguang Han(韩枝光)

Chin. Phys. B. 33, 120307 (2024). DOI: 10.1088/1674-1056/ad78d7

Learning physical states of bulk crystalline materials from atomic trajectories in molecular dynamics simulation

Tian-Shou Liang(梁添寿), Peng-Peng Shi(时朋朋), San-Qing Su(苏三庆), Zhi Zeng(曾志)

Chin. Phys. B. 31, 126402 (2022). DOI: 10.1088/1674-1056/ac98a4

SPECIAL TOPIC — AI + Physical Science

Revealing structural signatures associated with stress overshoot in two-dimensional Lennard–Jones systems based on interpretable deep learning

Tongtong Shen(申童童)[1], Xinyao Wang(王新尧)[2,†], Zhuohang Li(李卓杭)[2,3], Xueyu Liu(刘雪宇)[1], and Wen Zheng(郑文)[2,3,‡]

[1] *College of Artificial Intelligence, Taiyuan University of Technology, Jinzhong* 030600, *China*
[2] *Key Laboratory for Satellite Digitalization Technology, Innovation Academy for Microsatellites of Chinese Academy of Sciences, Shanghai* 201306, *China*
[3] *University of Chinese Academy of Sciences, Beijing* 100039, *China*

(Received 1 April 2026; revised manuscript received 12 June 2026; accepted manuscript online 30 June 2026)

Determining whether stress overshoot can be predicted from pre-yield structural information remains a central challenge in understanding amorphous yielding. However, conventional analyses usually require the complete stress–strain response to determine whether stress overshoot occurs. To overcome this limitation, a precise predictive model with hybrid convolutional–temporal encoding is developed, and the temporal–spatial class activation mapping (TS-CAM) method is introduced to interpret stress-overshoot behavior from pre-yield amorphous configuration sequences. Combined with the TS-CAM analysis, the spatiotemporal evolution of model-salient regions associated with stress-overshoot classification is quantitatively characterized. Furthermore, a model-derived scalar descriptor is constructed to characterize predictive signatures correlated with stress-overshoot yielding within the present dataset. This work demonstrates the capability of interpretable deep learning in exploring the plastic deformation mechanism of amorphous solids and provides a data-driven insight for the yielding regulation and structural optimization of amorphous materials.

Keywords: amorphous materials, deep learning, microstructural evolution, stress overshoot

PACS: 83.10.Mj, 83.60.La, 83.50.Ax, 78.55.Qr

DOI: 10.1088/1674-1056/ae843a **CSTR:** 32038.14.CPB.ae843a

1. Introduction

Amorphous materials are a prototypical class of disordered systems that lack the long-range periodicity of crystals. Their structure and dynamics often exhibit pronounced many-body correlations, and they can remain trapped in metastable regions of a rugged energy landscape over extended time scales.[1–3] Over the past decades, amorphous and glassy systems have attracted sustained interest in soft condensed matter, metallic glasses and granular materials, and have shown considerable promise in applications such as high-strength structural materials, functional devices and processing and forming technologies.[4–8] More importantly, as relatively simple model systems that nonetheless embody key mechanisms of disorder, amorphous materials provide a unified platform for exploring complex phenomena including the glass transition, jamming and yielding/plastic flow,[6,9,10] thereby helping to deepen our understanding of the formation, evolution and macroscopic response of disordered matter.

However, the research and application of amorphous materials still face several critical challenges, among which the yielding behavior under external shear is particularly representative.[10–14] Recent molecular-dynamics simulations under oscillatory shear have further shown that the yielding transition in metallic glasses is strongly affected by strain amplitude and thermal history, with aging below the yield strain and irreversible rearrangements or shear-band formation above it.[15] Extensive experiments and numerical simulations[11,13] have shown that, during start-up shear or continuous deformation, the stress–strain response of amorphous solids typically falls into two characteristic categories: (i) stress-overshoot yielding, characterized by a nonmonotonic transient stress response with a distinct stress maximum before relaxation toward the steady-flow state, and (ii) smooth yielding, in which the stress evolves toward steady flow without a pronounced stress maximum. In this study, the distinction is made using an operational criterion: a trajectory is classified as stress-overshoot yielding when the peak stress is more than 20% higher than the steady-state stress, whereas the remaining cases are classified as smooth yielding. Although a number of mechanistic pictures have been proposed — ranging from structural yielding[16,17] and cascades of localized plastic events[18] to the formation of shear bands and stress-relaxation

†Corresponding author. E-mail: wangxinyao@microsate.ac.cn
‡Corresponding author. E-mail: zhengwen@microsate.ac.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

dynamics[19–21] — a unified microscopic explanation for the distinct origins of overshoot versus smooth yielding remains elusive. This difficulty is also reflected in constitutive descriptions: recent homogeneous-rheology measurements on a high-entropy metallic glass indicate that conventional free-volume models struggle to quantitatively reproduce transient stress-overshoot responses.[22] Meanwhile, despite the availability of structural characterization and statistical analyses, such as x-ray diffraction,[23] pair distribution function analysis[24] and various order-parameter-based metrics,[25] it is still difficult in many cases to reliably identify and quantify the local regions that contribute most significantly to overshoot-type yielding. This limitation hampers a mechanistic understanding of yielding and, consequently, the predictive modeling of deformation responses in amorphous systems.

In recent years, driven by significant improvements in computational hardware performance and the explosive growth in data volume, the field of machine learning has undergone rapid development. As a prominent branch, deep learning has exhibited remarkable performance and achieved notable success across a multitude of domains.[26–29] However, the inherent "black-box" nature of machine learning models, particularly those in deep learning, curtails their application in high-stakes decision-making fields such as finance and medicine, and concurrently diminishes public trust in their predictive outcomes. Consequently, research on the interpretability of these models, a field known as explainable artificial intelligence, has emerged as a focal point of intense interest within both academia and industry.[30,31]

Machine learning has been widely used to analyze the structure and study the dynamical properties of amorphous materials.[32–35] More recently, machine-learning workflows have also been used to construct composition–property databases for ternary amorphous alloys, enabling efficient screening of glass-forming ability and property targets.[36] For instance, machine learning has been instrumental in identifying "soft spots" or localized structural features that are prone to rearrangement under shear, predicting the potential energy landscape and even forecasting the onset of plastic failure or yielding from the initial undeformed structure.[37,38] Indeed, while machine learning has proven effective at identifying localized structural motifs, a generalized representation of amorphous structure continues to be a significant hurdle owing to its long-range disordered nature. Moreover, the high accuracy often achieved by these predictive models is paradoxically accompanied by a lack of interpretability. To address this "black-box" problem, the scientific community has increasingly turned to explainable AI techniques.[39,40] The application of such methods is critical for moving beyond mere prediction, offering a pathway to extract meaningful physical principles, quantify structure–property relationships and advance the study of amorphous state dynamics.

In this work, we systematically investigate the mechanical behavior of a two-dimensional Lennard–Jones amorphous solid under a constant shear rate using molecular dynamics simulations. The shear deformation process is examined under both athermal and thermal conditions. Based on the simulation data, a precise predictive model based on EfficientNet[41] is proposed to predict the yielding mode of the system, specifically the occurrence of stress overshoot. Compared with baseline models, the model demonstrates superior classification performance, achieving a maximum prediction accuracy of 98.75%. We note that, because the label fractions vary with temperature, temperature-related correlations may contribute to the observed classification performance. To elucidate the model's decision-making mechanism, we employ the temporal–spatial class activation mapping (TS-CAM) method, an explainability technique, to generate saliency maps. These maps highlight model-salient microstructural regions that are strongly correlated with stress overshoot within the present data distribution. Finally, based on these saliency maps, a structural scalar descriptor is defined and extracted. This parameter effectively quantifies the structural features closely associated with the stress-overshoot behavior. Furthermore, it exhibits excellent consistency across simulations with varying system sizes, thus effectively mitigating the influence of finite-size effects.

2. Methodology

2.1. Simulation details

In this study, we utilize a two-dimensional (2D) bidisperse amorphous model.[42] The system is subject to periodic boundary conditions and consists of N particles in total, with $N/2$ large and $N/2$ small particles, all having the same mass m. The pairwise interaction between particles is described by a truncated Lennard–Jones potential, given by

$$V(r_{ij}) = \begin{cases} 4\varepsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], & \text{for } r_{ij} \le 2.5\sigma_{ij}, \\ 0, & \text{for } r_{ij} > 2.5\sigma_{ij}, \end{cases} \quad (1)$$

where r_{ij} is the distance between particle i and particle j, and σ_{ij} is the effective particle diameter, defined as the sum of their radii, $\sigma_{ij} = R_i + R_j$. The parameter ε represents the characteristic energy scale of the interaction.

In our simulations, the system was initialized by assigning random positions to all particles within the simulation box. We normalized the diameter of the small particles to set 1, resulting in the large particles having a diameter of 1.4. Periodic boundary conditions were applied in both the x and y directions. For each independent sample, particles were randomly placed in the simulation box using different random seeds,

and the resulting configurations were relaxed by conjugate-gradient energy minimization to remove unfavorable overlaps and obtain mechanically stable inherent structures. All simulations were performed in reduced Lennard–Jones units with a time step of $\Delta t = 0.001$. The minimized configurations were subsequently equilibrated at the target temperature for 5×10^4 steps using a Nosé–Hoover NVT thermostat, with a thermostat damping time of 0.05 in reduced Lennard–Jones units. After equilibration, simple shear was imposed using the SLLOD equations of motion together with a continuously deformed simulation box at a constant shear rate of $\dot{\gamma} = 0.001$. During shear deformation, particle configurations and stress data were recorded every 3000 steps, corresponding to a strain interval of $\Delta\gamma = 0.003$. Each shear simulation was run up to a total strain of $\gamma = 0.3$, which is sufficient to cover the initial elastic response, the yielding transition and the subsequent steady-flow regime. To systematically investigate the influence of temperature on the yielding behavior, shear simulations were conducted at five different temperatures: $T = 1.0$, 0.1, 0.01, 0.001 and 0.0001. These temperatures were selected to span several orders of magnitude, from a relatively high-temperature thermal regime to an effectively athermal regime. This broad range allows us to examine how thermal fluctuations affect the yielding response.

To examine potential finite-size effects, we considered systems with total particle numbers $N = 128$, 256, 512, 1024, and 2048. The smallest system, $N = 128$, was used only to illustrate the simulation conditions, as shown in Fig. 2(a), whereas the subsequent model training, quantitative analysis, and finite-size assessment were performed using the larger systems with $N = 256$, 512, 1024, and 2048. Previous studies[43,44] have shown that such model systems are useful for investigating yielding behavior in amorphous solids and have also highlighted the importance of finite-size considerations in this context. Additional finite-size statistics and threshold-sensitivity analysis are provided in the supplementary material.

To probe the mechanical response, the systems were subjected to simple shear at a constant strain rate $\dot{\gamma}$. The total applied strain was sufficient to drive the system from its initial elastic response, through the yielding transition and into the steady-state plastic flow regime. This loading protocol falls within the model of non-equilibrium molecular dynamics (NEMD) simulations. To systematically investigate the influence of temperature on the yielding behavior, shear simulations were conducted at five different temperatures: 1.0, 0.1, 0.01, 0.001 and 0.0001. For each temperature, 1000 independent simulation runs were performed with different initial configurations to ensure robust and representative statistical results.

2.2. Dataset construction and preprocessing

For this study, we constructed a comprehensive dataset of particle configuration images derived from 2D Lennard–Jones amorphous systems of various sizes. The dataset comprises a total of 4000 independent shear deformation sequences. Each sequence consists of five temporally ordered snapshots of the nanoscale particle configurations, capturing their evolution. For each independent shear trajectory, the yielding strain γ_y was first determined as the strain corresponding to the maximum stress in the stress–strain curve. Although γ_y can be determined for each simulated trajectory, its absolute value varies among independent samples. Therefore, we used the normalized strain position γ_t/γ_y, rather than fixed absolute strain values, to define a consistent frame-selection criterion across all trajectories. Five input frames were then selected at evenly spaced relative strain positions within $0.2\gamma_y \leq \gamma_t \leq 0.6\gamma_y$, namely $\gamma_t/\gamma_y = 0.20$, 0.30, 0.40, 0.50, and 0.60. This strain window was chosen to ensure that all selected configurations remain in the pre-yield regime while being sufficiently separated from the stress maximum, thereby minimizing the influence of near-yield deformation features. With the particle trajectory information recorded in the simulations, the corresponding configuration images were then generated directly from the particle positions at the selected strain values. These sequences were generated for four distinct system sizes, with $N = 256$, 512, 1024, and 2048 particles, amounting to 1000 sequences per system size.

Each sequence was manually annotated based on its macroscopic stress–strain profile, classifying it into one of two categories: smooth or overshoot. The entire dataset was then partitioned into training, validation, and test sets following a 7:1:2 ratio. Specifically, for each system size, 700 sequences were allocated for training, 100 for validation, and 200 for testing. This balanced partitioning scheme is designed to ensure robust statistical analysis and to provide an unbiased evaluation of the model's performance.

It should be noted that these labels are operational labels derived from the stress–strain response of the same simulation trajectories that provide the input image sequences. Therefore, the learning task is designed to identify predictive correlations between pre-yield configurations and subsequent yielding responses, rather than to establish a causal structure–property relationship.

2.3. The yielding-mode predictive model

To address the two distinct yielding modes in amorphous materials — namely, yielding with a stress overshoot and smooth yielding — we have developed a precise predictive model based on deep learning. The model utilizes a convolutional neural network (CNN) as its backbone for spatial fea-

ture extraction, integrated with a time-series modeling module to effectively capture the spatiotemporal correlations within the local stress field preceding yielding. The results demonstrate that this approach can accurately classify the final yielding mode using only image sequences from the pre-yielding stage. Furthermore, by employing feature-contribution analysis, the model identifies model-salient regions that are associated with the predicted yielding mode. These regions reflect the information used by the neural network for classification and should not be directly interpreted as microscopic mechanisms of yielding.

Figure 1(a) illustrates a typical pre-yielding sequence in a 2D amorphous system, where each frame is an image of the non-affine displacement field. From left to right, as the simulation temperature increases, the yielding behavior transitions from an overshoot-dominated regime to a smooth yielding regime. The latter is associated with a proliferation of numerous plastic events. The system contains $N = 512$ particles. Each sequence is composed of five frames, with each frame having a resolution of 224×224 pixels.

The input to the proposed predictive model consists of sequences of 2D amorphous image frames, each sequence being associated with a ground truth label. An image sequence, denoted as $I_i = \left\{I_i^1, I_i^2, I_i^3, I_i^4, \ldots, I_i^j\right\}$, where I_i^j represents the jth frame in the ith sequence, is first fed into the CNN backbone. We employ EfficientNet as the backbone network to ensure robust feature extraction capabilities. Specifically, we employed the EfficientNet-B0 model with a width coefficient of 1.0, a depth coefficient of 1.0, and a dropout rate of 0.2. The model was initialized with pre-trained weights provided by the official implementation to ensure efficient feature extraction and strong generalization performance. For each frame, the CNN backbone outputs a spatial feature vector, which is subsequently projected to a uniform 256-dimensional space via a mapping layer. The resulting sequence of frame-level feature vectors is then processed by a two-layer long short-term memory (LSTM) network. In our experiments, the temporal network utilized a two-layer LSTM model with a hidden layer dimension of 256. This was determined as the minimal dimension through experimentation, balancing computational speed and modeling efficiency. The model was trained for 100 epochs using the Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 16. The cross-entropy loss was used as the optimization objective. During training, the validation set was used to monitor the model performance.

The LSTM network outputs a single vector, stF_i, which encodes the comprehensive spatiotemporal information of the entire input sequence. This vector stF_i is then passed through a classifier to yield the predicted probability distribution, $\widehat{Y}_i$, for the ith sequence. Our objective is to train the model by minimizing the loss between the predicted distribution $\widehat{Y}_i$ and the true label Y_i. Here, $I(\cdot)$ acts as an indicator function, taking the value 1 if the condition inside the parentheses is satisfied and 0 otherwise. In this context, $I(Y_i = 1)$ equals 1 when the ith sequence belongs to the positive class and 0 otherwise. The cross-entropy cost function we utilize is defined as

$$L = -\sum_i \left[I(Y_i = 1) \log \widehat{Y}_i + I(Y_i = 0) \log\left(1 - \widehat{Y}_i\right) \right]. \quad (2)$$

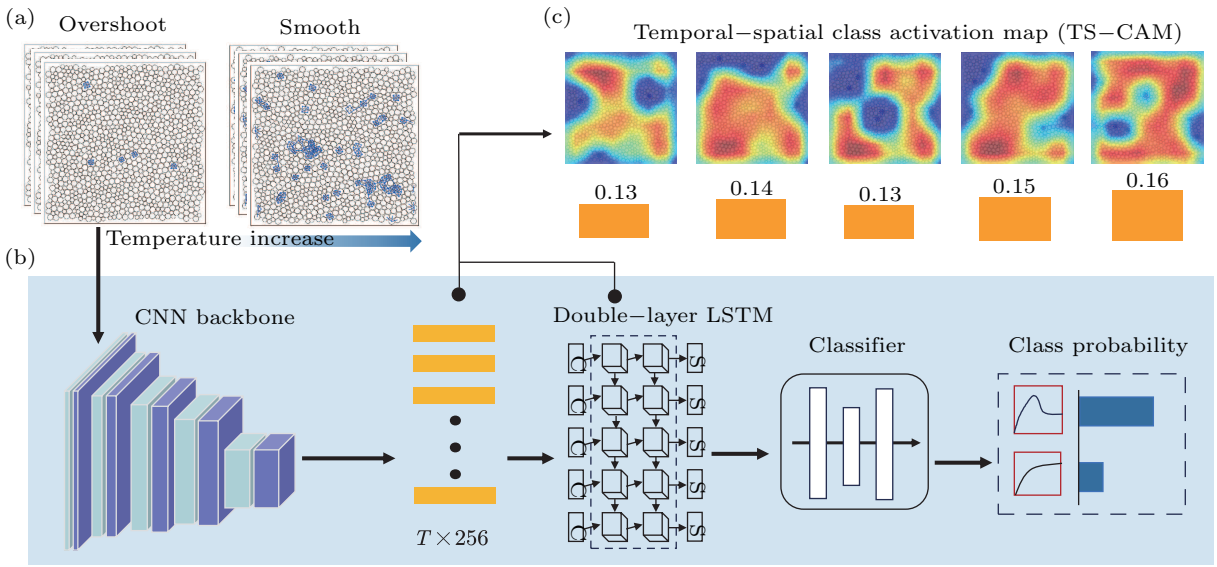


Fig. 1. Amorphous sequence data and the proposed predictive model. (a) Configuration sequence of a 2D amorphous system under shear. From left to right, as the NVT simulation temperature gradually increases, the system transitions from overshoot yielding to smooth yielding. (b) Schematic of the predictive model. Each frame is encoded into a feature vector through a CNN backbone, the sequence features are processed by a two-layer long short-term memory (LSTM) to capture temporal information, and finally the probability of the yielding type is output through a fully connected layer. (c) Spatiotemporal class activation heatmap. The heatmap combines CNN frame-level feature vectors and LSTM outputs to highlight the key frames contributing to the prediction. The numbers below each heatmap (e.g., 0.13, 0.14,...) quantify the contribution of each frame in the sequence to the classification of stress-overshoot yielding.

Figure 1(b) illustrates the training and learning process of the proposed model using the amorphous sequence dataset, which ultimately yields an accurate probability value for the corresponding yielding mode.

As depicted in Fig. 1(c), we visualize the intermediate outputs of the network: the spatial microstructural feature vectors extracted by the CNN backbone and the temporal feature vectors encoded by the LSTM model. This is achieved using the TS-CAM technique to generate contribution maps across both the spatial and temporal dimensions. These maps serve to quantify the respective roles of these features in the discrimination process for the overshoot yielding mode.

2.4. TS-CAM technology

To quantify the contribution of various regions within an image sequence to overshoot yielding, both spatially and temporally, we have developed TS-CAM. This technique, based on Grad-CAM,[45] is designed to visualize the contribution of spatiotemporal regions throughout the image sequence.

Consider a neural network for classification, where the final layer is a softmax over classes. Let y^c denote the output score for class c of interest. For each frame t, let $A^k_{pq}(t)$ be the activation of the kth channel at spatial location (p,q) in a selected convolutional layer. The importance weight of channel k for class c at time t is obtained by globally averaging the gradients over all spatial locations, where Z is the total number of spatial positions (i.e., $Z = H \times W$ for a $H \times W$ feature map)

$$\alpha^c_k(t) = \frac{1}{Z}\sum_p\sum_q \frac{\partial y^c}{\partial A^k_{pq}(t)}. \tag{3}$$

Finally, the class activation map for each frame is obtained by computing a weighted sum of the feature maps from the final convolutional layer, where the weights are the neuron importance scores (i.e., the gradients averaged over spatial positions) calculated previously. A ReLU operation is applied to this linear combination to retain only the spatial locations that positively contribute to the classification of interest, while setting negative contributions to zero. Conceptually, this corresponds to a weighted superposition of feature channels, with each channel scaled by its importance.[45] For a sequence of T frames, the resulting frame-wise class activation maps are denoted as $H_{pq}(t)$, where $t = 1,\ldots,T$. Here, T represents the total number of frames in the input image sequence. The final temporal–spatial class activation map for class c is then defined as

$$H^c_{pq}(t) = \mathrm{ReLU}\left(\sum_k \alpha^c_k(t) A^k_{pq}(t)\right). \tag{4}$$

After obtaining the frame-wise class activation maps $H_{pq}(t)$ from TS-CAM, we compute a scalar score $S(t)$ for each frame to quantify its overall contribution to the classification of the sequence. Specifically, $S(t)$ is defined as the sum of activations over all spatial locations

$$S(t) = \sum_{p,q} H^c_{pq}(t), \tag{5}$$

where the summation runs over all spatial positions (p,q). Here, $t = 1,\ldots,T$ and T denotes the total number of frames in the input sequence. $S(t)$ provides a concise measure of how strongly each frame influences the model's decision.

3. Results and analysis

3.1. Yielding behavior and dynamical signatures

We performed shear simulations on a 2D Lennard–Jones glass to investigate its mechanical behavior and dynamical signatures. Figure 2(a) shows a representative instantaneous configuration of the amorphous solid with $N = 128$ under a fixed shear rate $\dot{\gamma} = 0.001$. It can be seen that the amorphous solid is composed of two particle species, and it exhibits long-range disorder with no stable crystalline lattice structure. To eliminate the influence of simulation-box deformation, we apply a de-affine transformation to all configurations, following Ref. [46]. Figures 2(b) and 2(c) compare a representative instantaneous configuration before and after this procedure: the left panel shows the configuration under the apparent distortion induced by box deformation whereas the right panel presents the corresponding de-affined configuration. To visualize the spatial distribution of particle energy more clearly, particles are rendered with a color fill, where darker colors indicate higher energy. It is evident that, under NEMD shear, the energy distribution in the amorphous solid is spatially heterogeneous. Moreover, the de-affine transformation does not alter the particles' relative positions or their energy values, but only removes the global affine contribution associated with the box deformation.

We record the stress–strain response in a representative NEMD shear run, as shown in Fig. 2(d). The curve exhibits a pronounced multi-step serrated pattern, similar to that reported in Ref. [47]. Such serrations are a typical feature of amorphous solids at small strains and are generally associated with successive, localized rearrangements that continuously induce slight structural remodeling. For a single amorphous system, the measured stress is intrinsically noisy and inevitably affected by thermal and mechanical fluctuations. To mitigate these fluctuation effects, we perform statistical averaging over multiple independent realizations prepared under identical conditions, which reveals a clear and reproducible trend. As summarized in Fig. 2(e), upon decreasing the temperature, the yielding behavior evolves from smooth yielding at high temperature to an overshoot-type yielding at low temperature. To better resolve the elasto-plastic response in the small-strain regime, we further plot the stress–strain data in a

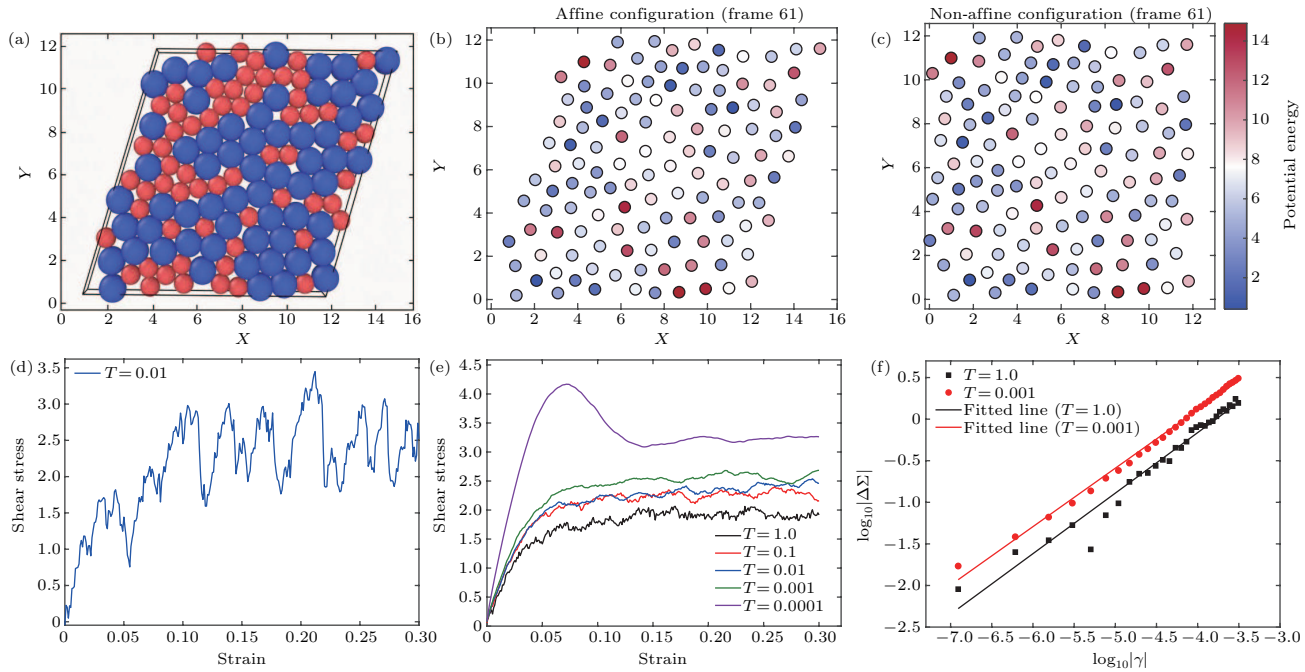


Fig. 2. Mechanical response of two-dimensional amorphous materials under shear. (a) Representative configuration of a 2D Lennard–Jones amorphous system with $N = 128$ under simple shear, shown only to illustrate the simulation geometry and particle arrangement. (b) Affine particle configuration. (c) Corresponding non-affine configuration after removing the global affine deformation. (d) Single-run stress–strain response of a system with $N = 1024$, showing serrated stress fluctuations associated with localized plastic events. (e) Averaged stress–strain curves for systems with $N = 1024$ sheared at $\dot{\gamma} = 0.001$ over a range of temperatures ($T = 1.0, 0.1, 0.01, 0.001, 0.0001$). (f) Log–log stress–strain representation for $T = 1.0$ and $T = 0.001$, emphasizing the distinct initial elastic responses of thermal and athermal-like systems.

log–log representation, as shown in Fig. 2(f). This representation highlights the pronounced elastic response of the amorphous solid during shear start-up.

3.2. Classification results for yielding modes

We constructed a predictive model based on a hybrid CNN-LSTM architecture to distinguish different yielding modes, and systematically evaluated the classification performance using area under the curve (AUC) for different datasets. The AUC is defined as the area under the receiver operating characteristic (ROC) curve, and the ROC curve is drawn from the true positive rate (TPR) and the false positive rate (FPR). An AUC of 0.5 indicates that the model performs no better than random guessing, whereas an AUC of 1.0 corresponds to perfect discrimination. Next, we compared the classification metrics of the proposed model with the results of other sequence classification networks. The networks we chose for comparison are AlexNet,[48] VGG,[49] DenseNet[50] and ResNet.[51] Table 1 presents the AUCs of our classification model under different datasets. Four of these datasets have AUC > 0.81, indicating that the classification model is very efficient for identifying the yielding modes.

To more clearly compare the performance of different models and assess robustness under varying conditions, we performed additional experiments. Figure 3(a) shows the ROC curves and corresponding AUC values for different classifiers on the test set with $N = 512$. Among them, the proposed model using EfficientNet as the backbone achieves the best performance, with an AUC of 0.975, demonstrating a strong ability to discriminate between the two yielding modes. It is worth emphasizing that, although a high AUC provides an intuitive measure of overall discrimination, it does not fully capture class-specific prediction behavior. To this end, Fig. 3(b) presents the confusion matrix of the proposed model, which directly visualizes correct predictions and misclassifications for both classes, thereby offering a more detailed view of class sensitivity and error patterns.

Table 1. Comparison of mean classification AUC for the proposed model with different CNN backbones. The evaluation was performed on datasets with system sizes N ranging from 256 to 2048.

Model	$N = 256$	$N = 512$	$N = 1024$	$N = 2048$
AlexNet	0.852	0.964	0.896	0.868
VGG	0.864	0.980	0.914	0.886
ResNet	0.876	0.980	0.932	0.904
DenseNet	0.888	0.966	0.949	0.922
Ours (EfficientNet)	0.913	0.975	0.973	0.946

Figure 3(c) further reports how the AUC of the proposed model varies with the system size N. The AUC increases markedly with N and then gradually saturates; specifically, the AUC is about 0.92 for $N = 256$, increases rapidly for $N = 512$, and then shows only marginal improvement for larger systems. This trend indicates that larger systems provide more stable macroscopic yielding responses and richer statistics of microstructural features, thereby mitigating finite-size fluctuations that can compromise labeling and representation learn-

ing, and ultimately improving separability and generalization. In Fig. 3(d), we examine the effect of the input sequence length L (i.e., the number of consecutive frames fed into the LSTM) on classification performance. The AUC increases monotonically with L and saturates around $L \approx 5$: when only short sequences are used, $L = 2$ and the AUC is approximately 0.82, whereas increasing to $L = 4$ boosts the AUC to 0.97. Further increasing to $L = 5$ yields only limited improvement, reaching a plateau near 0.98. These results suggest that yielding-mode identification benefits not only from single-frame spatial information but also from short-time dynamical evolution. At the same time, the decisive "dynamical fingerprints" are largely contained within a finite temporal window, beyond which additional frames contribute little marginal information. Balancing accuracy against computational cost, we therefore choose $L \approx 5$ as a reasonable default for subsequent experiments.

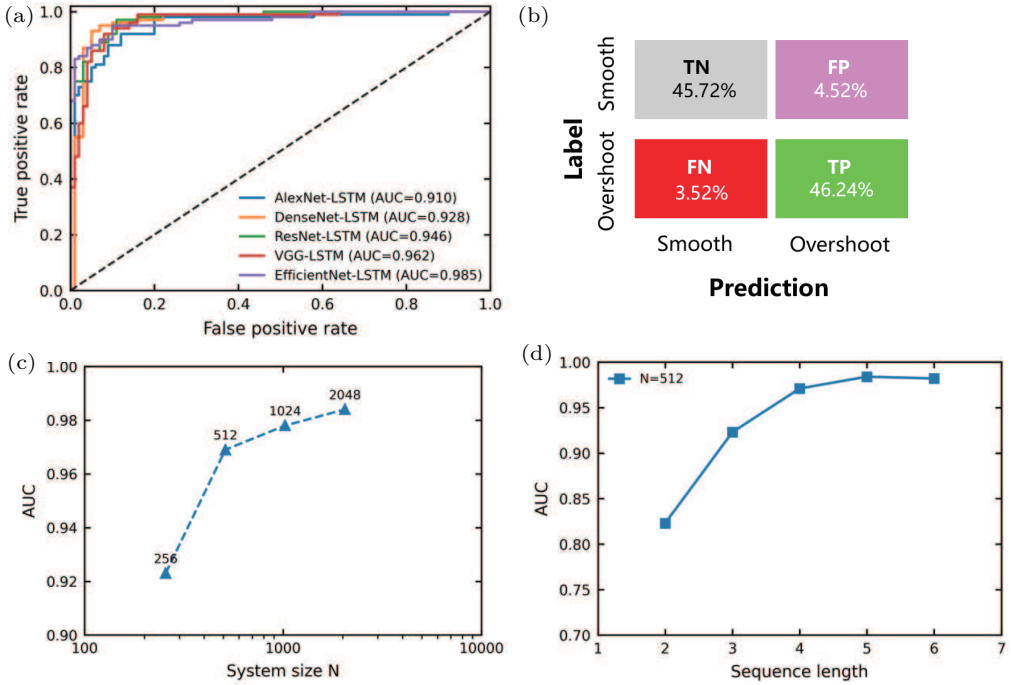


Fig. 3. Comparison of classification performance. (a) ROC curves for the proposed model with different CNN backbones evaluated on the test set at system size $N = 512$. (b) Confusion matrix of the best-performing model on the test set. (c) Finite-size analysis showing the dependence of AUC on system size N. (d) Effect of input sequence length L as measured by the number of consecutive frames fed into the LSTM on the AUC.

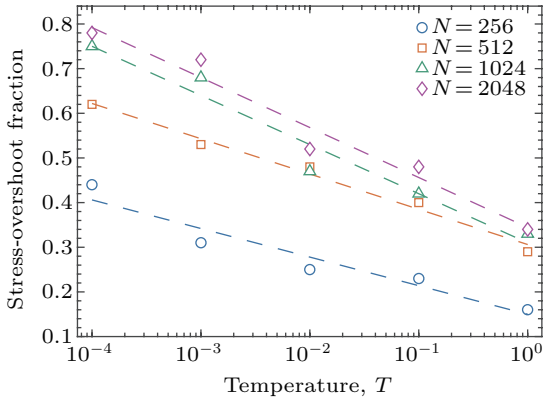


Fig. 4. Temperature dependence of the stress-overshoot fraction for different system sizes. Dashed lines are linear fits highlighting the temperature-dependent trend for each system size.

Figure 4 presents the finite-size dependence of the stress-overshoot fraction under different temperature conditions. The lower overshoot fraction in smaller systems suggests that pronounced stress overshoot can be partly smeared out by stronger finite-size fluctuations in the stress–strain response. As the system size increases, the overshoot fraction becomes more stable, indicating that the yielding statistics are less dominated by sample-to-sample fluctuations. This finite-size trend provides useful support for the stress-overshoot-based label assignment used in the present classification task. More detailed data and related analyses are provided in the supplementary material.

3.3. Visualizations and interpretability analysis

Although deep learning models have achieved impressive performance across a wide range of tasks, their black-box nature limits the verifiability of the predictions and makes it challenging to interpret the relationships between the learned features and the underlying data. To examine whether the discrimination of the two yielding modes by the proposed model is supported by interpretable structural signatures, and to assess the necessity of the CNN-LSTM temporal hybrid architecture, we visualize the learned representations using t-distributed stochastic neighbor embedding (t-SNE). This nonlinear dimensionality-reduction technique embeds high-dimensional features into a 2D space while approximately pre-

serving local neighborhood relationships, thereby providing an intuitive view of class clustering and separability.

Specifically, we extract feature vectors of amorphous configuration sequences from the test set and perform t-SNE mapping for the static features produced by the CNN backbone and the spatiotemporal features further encoded by the LSTM. As shown in Fig. 5(a), the static features exhibit substantial overlap in the 2D embedding, with no clear decision boundary between the two classes; moreover, the distribution evolves continuously from smooth yielding to overshoot yielding. This indicates that single-frame structural information has limited discriminative power and is insufficient for robust classification. In contrast, Fig. 5(b) shows a pronounced separation in the spatiotemporal feature space, where samples from the two yielding modes form relatively distinct clusters. This observation suggests that the LSTM effectively introduces temporal correlations associated with shear-driven structural evolution, capturing key spatiotemporal information that markedly enhances class separability. These results, from the perspective of representation learning, corroborate the effectiveness of the CNN-LSTM hybrid modeling: relying solely on CNN-derived static features does not support stable classification, whereas incorporating spatiotemporal evolution is essential for distinguishing the two yielding modes.

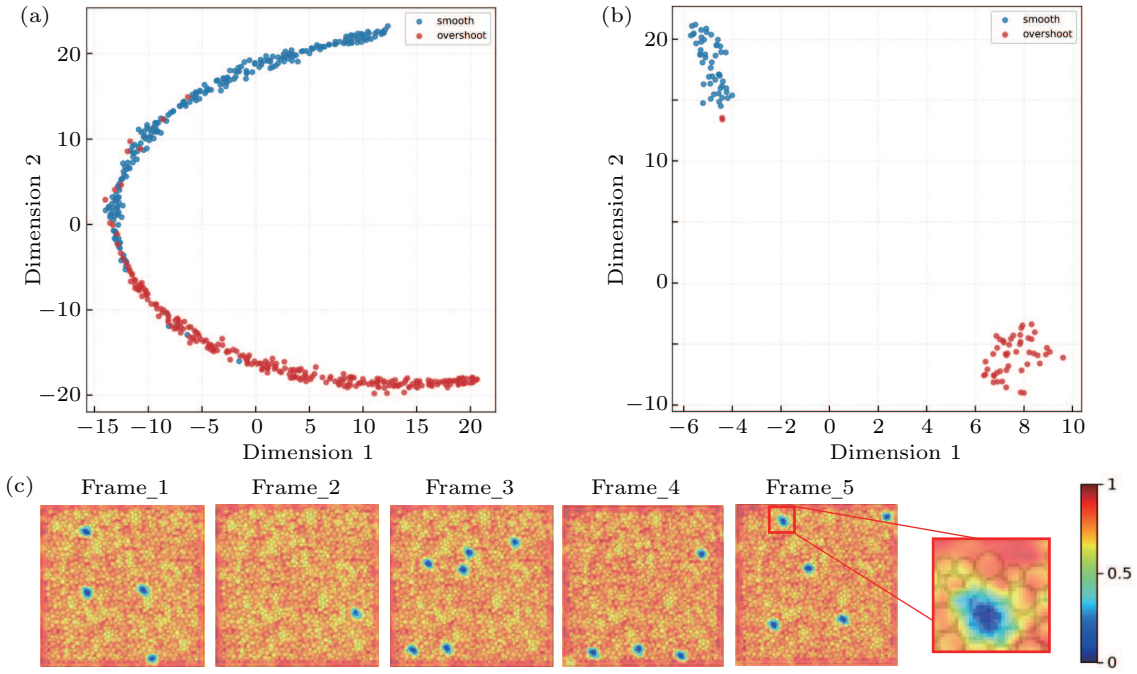


Fig. 5. Visualization and analysis of the test set data. (a) The feature space of static image feature vectors extracted by EfficientNet, visualized using the t-SNE technique. This shows that the image vectors from overshoot to smooth yielding change continuously. (b) The feature vector of the last frame of the LSTM module, which represents a comprehensive encoding of the spatiotemporal structure of the entire time series. This vector exhibits good separation in the feature space, indicating that the proposed model can effectively utilize sequential spatiotemporal features to distinguish between different modes. (c) Visualization of the proposed model's convolutional layers using TS-CAM, showing the spatiotemporal heatmap for a stress-overshoot sample. The color bar represents the normalized TS-CAM activation values from 0 to 1. Warmer colors, especially red regions, indicate image structures with higher activation values and stronger relevance to stress overshoot.

We employ the TS-CAM method introduced in Subsection 2.4 to generate class-activation heat maps for sequences of amorphous configurations, which quantify the contribution of each spatial location at every time step to the final prediction. The workflow for producing these heat maps is illustrated in Fig. 1. Specifically, by visualizing the spatial gradients at the pixel level, we identify the regions in the configuration images that the model relies on when making its decision. The resulting heat maps therefore highlight the image patches that are most strongly activated by the neural network. Figure 5(c) presents TS-CAM visualizations for an amorphous configuration sequence exhibiting overshoot yielding under shear. The class-activation heat map can be regarded as an infrared-like image: darker colors indicate locations that contribute more strongly to the classification decision — here, regions that are more influential for identifying the overshoot behavior. From left to right, the panels correspond to successive frames of the same sequence together with the predictions of the proposed model. The results show that the model's salient regions are not fixed; instead, the locations most relevant to the yielding response evolve in both space and time during deformation.

It should be emphasized that the TS-CAM maps are interpreted here as model-salient regions rather than direct microscopic mechanisms. In the present input representation, these regions indicate where the network extracts the strongest predictive signals from the pre-yield configuration

sequence. Their spatial localization is consistent with the heterogeneous nature of amorphous plasticity, in which localized particle activity may precede larger-scale rearrangements. Notably, the stress-overshoot-related zones evolve spatially over time during shear, whereas their overall statistical state becomes relatively stable after the system enters the steady-flow regime. However, because no quantitative comparison with established descriptors, such as non-affine displacement, local strain, shear transformation zones or soft-spot indicators, is performed in the present work, we do not identify these regions as a specific microscopic plasticity mechanism.

3.4. Localization and quantification of overshoot-related zones

Because the key regions highlighted by the deep learning model vary in their spatial extent, we introduce a unified and quantifiable descriptor for the salient structures identified under different conditions. We call the regions recognized by the proposed model that contribute to the overshoot-yielding mode overshoot-related zones (ORZs).

We introduce a model-derived scalar descriptor h to quantify, from a configuration sequence of an amorphous system, the degree to which the sequence is associated with overshoot yielding. To construct a sequence-level scalar descriptor, we first normalize the frame-wise contribution score $S(t)$ using min–max normalization. The scalar descriptor h is then defined as the average normalized contribution over the T frames. Here, $S_{\min}$ and $S_{\max}$ denote the minimum and maximum values of $S(t)$ within each sequence, respectively, and ε is a small constant introduced to avoid numerical instability. While h serves as a unified, quantitative descriptor generated by the proposed model rather than a conventional physical observable, it effectively captures how the system's structure responds to external shear and reflects the propensity of each test sequence toward overshoot yielding based on $S(t)$ defined in Subsection 2.4

$$\tilde{S}(t) = \frac{S(t) - S_{\min}}{S_{\max} - S_{\min} + \varepsilon}, \tag{6}$$

$$h = \frac{1}{T}\sum_{t=1}^{T} \tilde{S}(t). \tag{7}$$

Figure 6 shows the structural signatures of the ORZs extracted by the proposed model for the two yielding modes. One can see that both $S(t)$ and h clearly discriminate the two modes, with overshoot-yielding samples exhibiting markedly stronger ORZ signatures. In this study, the distinction between smooth and overshoot-dominated responses was mainly obtained by varying temperature, while the observed yielding behavior may also be influenced by the shear rate and the initial configuration. By systematically analyzing the simulation data under different temperatures, shear rates, and initial configurations, we observed that the operational yielding labels of the 2D Lennard–Jones system exhibit a relatively sharp crossover between smooth and overshoot-dominated responses within the present dataset. Accordingly, after computing the scalar descriptor h and sorting the samples by a prescribed rule, h exhibits an S-shaped curve with a sharp variation in the intermediate region. The scalar h thus provides a clear separation between overshoot and smooth yielding in the context of the proposed model. Its S-shaped profile becomes sharper for larger system sizes N, reflecting the model's increasing ability to discriminate yielding modes, although further work is required to establish direct correlations with physically measurable quantities.

Overall, these results demonstrate that the proposed model can accurately predict the yielding mode using only pre-yield configuration sequences of 2D amorphous materials while simultaneously enabling a quantitative characterization of the ORZ in both space and time. This capability highlights the substantial potential of the proposed model for studying shear yielding in amorphous solids.

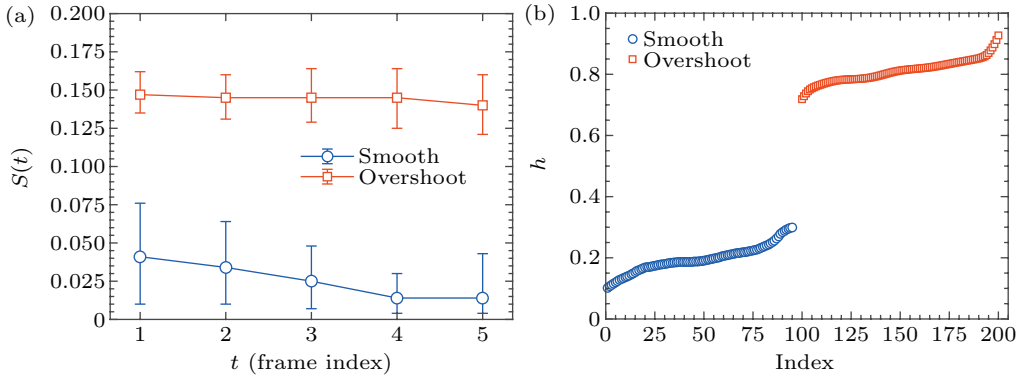


Fig. 6. Evolution and statistics of the scalar descriptor h in amorphous configuration sequences. (a) $S(t)$ denotes the structural signatures of the ORZs identified by the proposed model from the configuration at frame t. As shear loading proceeds, $S(t)$ evolves in time for both yielding modes and exhibits pronounced sample-to-sample variability. (b) A scalar descriptor h is constructed based on $S(t)$. After ranking the samples according to the prescribed rule, h follows a characteristic S-shaped trend. In particular, for sequences corresponding to smooth yielding, the signatures of the ORZs captured by the proposed model remain weak overall.

4. Conclusion

In this work, we develop a precise predictive model based on deep learning to accurately identify yielding modes in 2D amorphous systems under shear loading. Trained on molecular-dynamics simulation data with different particle numbers, the model takes time sequences of amorphous particle configurations as input and achieves high-precision classification of yielding behaviors. To further elucidate the physical basis underlying the model's decisions, we adopt an improved TS-CAM approach. The resulting class-activation heat maps clearly localize the key structural regions that contribute strongly to the model's prediction of overshoot yielding, indicating that the model does not rely on incidental noise but instead captures robust structural signals correlated with the yielding response.

Based on the heatmaps generated by the proposed model, we extracted and constructed a new scalar descriptor that quantitatively characterizes the key structural regions playing a decisive role in stress-overshoot yielding and their spatiotemporal evolution throughout the configuration sequence. These structural signals are derived from the output of the neural network based on sequence images. Future research will further explore the relationship between these signals and known deformation mechanisms, such as shear transformation zones and soft spots. The system size dependence of this scalar descriptor is consistent with the theoretical expectations from finite-size effects, supporting its usefulness as a model-derived descriptor within the present dataset.

In summary, our study demonstrates that deep learning methods, combined with interpretability strategies, can effectively identify key structural information associated with macroscopic mechanical behaviors in complex disordered systems, providing a data-driven approach to understanding yielding mechanisms. While the current results are obtained for 2D Lennard–Jones amorphous systems and cannot be directly generalized to all amorphous materials, they reveal distinct yielding modes and the potential of deep learning to capture physically relevant structural signatures. This work develops a precise predictive model, providing a new approach for investigating yielding behaviors and offering a foundation for future studies to validate and extend the method to three-dimensional or more realistic systems.

Program availability

The code used in this study is available at https://github.com/happytong1/VideoClassification. Due to the large size of the dataset, the data are not publicly available but may be obtained from the authors upon reasonable request.

Acknowledgments

Project supported by the National Natural Science Foundation of China (Grant No. 11702289) and the Science and Technology Commission of Shanghai under the Second Batch of the 2024 Annual Science and Technology Innovation Action Plan in the Field of Aerospace and Marine Science and Technology (Grant No. 24DZ3101402).

References

[1] Debenedetti P G and Stillinger F H 2001 *Nature* **410** 259
[2] Berthier L and Biroli G 2011 *Rev. Mod. Phys.* **83** 587
[3] Shell M S, Debenedetti P G, La Nave E and Sciortino F 2003 *J. Chem. Phys.* **118** 8821
[4] Sollich P, Lequeux F, Hébraud P and Cates M E 1997 *Phys. Rev. Lett.* **78** 2020
[5] Liu A J and Nagel S R 2010 *Annu. Rev. Condens. Matter Phys.* **1** 347
[6] O'Hern C S, Silbert L E, Liu A J and Nagel S R 2003 *Phys. Rev. E* **68** 011306
[7] Bi D, Zhang J, Chakraborty B and Behringer R P 2011 *Nature* **480** 355
[8] Spaepen F 1977 *Acta Metall.* **25** 407
[9] Falk M L and Langer J S 2011 *Annu. Rev. Condens. Matter Phys.* **2** 353
[10] Greer A L, Cheng Y Q and Ma E 2013 *Mater. Sci. Eng. R: Rep.* **74** 71
[11] Nabizadeh M and Jamali S 2021 *Nat. Commun.* **12** 4274
[12] Balogun A, Lazarenko D, Khabaz F and Khare R 2021 *Soft Matter* **17** 7210
[13] Benzi R, Divoux T, Barentin C, Manneville S, Sbragaglia M and Toschi F 2021 *Phys. Rev. E* **104** 034612
[14] Dutta B, Ghosh S, Gurrola A, Julson D, Kamon T and Kumar J 2023 *J. High Energy Phys.* **03** 164
[15] Ren N, Meng T, Huang H, Ma Q, Fang J, Li Q and Li W 2026 *Chin. Phys. B* **35** 016103
[16] Zhou Y N, Yang Z Y and Dai L H 2024 *Phys. Rev. Mater.* **8** 073603
[17] Yang Z Y and Dai L H 2022 *Phys. Rev. Mater.* **6** L100602
[18] Zhang Z, Jiang J, Lu B, Meng X, Herrmann H, Chen J and Li X 2022 *PNAS Nexus* **1** pgac266
[19] Spieckermann F, Şopu D, Soprunyuk V, Kerber M B, Bednarčík J, Schökel A, Rezvan A, Ketov S, Sarac B, Schafler E and Eckert J 2022 *Nat. Commun.* **13** 127
[20] Akamo D O 2023 *Development and Characterization of Stable Low-cost Salt Hydrate-based Phase Change Materials for Thermal Energy Storage Applications* (Ph.D. Dissertation) (Knoxville: University of Tennessee)
[21] Yang Z Y, Miao Q, Dan J K, Liu M T and Wang Y J 2024 *Acta Mater.* **281** 120410
[22] Xing G, Bletry M, Mottelet S and Qiao J 2025 *Chin. Phys. B* **34** 126402
[23] Guinier A 1964 *Phys. Today* **17** 70
[24] Tuckerman M 2010 *Statistical Mechanics: Theory and Molecular Simulation* (New York: Oxford University Press)
[25] Steinhardt P J, Nelson D R and Ronchetti M 1983 *Phys. Rev. B* **28** 784
[26] Silver D, Huang A, Maddison C J, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T and Hassabis D 2016 *Nature* **529** 484
[27] Kim J and Park C 2017 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, July 21–26, 2017, Honolulu, HI, USA, pp. 1194–1202
[28] Tang T T, Zawaski J A, Francis K N, Qutub A A and Gaber M W 2019 *Sci. Rep.* **9** 12529
[29] Zhu Q, Samanta A, Li B, Rudd R E and Frolov T 2018 *Nat. Commun.* **9** 467
[30] Doshi-Velez F and Kim B 2017 arXiv:1702.08608 [stat.ML]
[31] Murdoch W J, Singh C, Kumbier K, Abbasi-Asl R and Yu B 2019 arXiv:1901.04592 [stat.ML]
[32] Dietz C, Kretz T and Thoma M H 2017 *Phys. Rev. E* **96** 011301
[33] Xiong J, Shi S Q and Zhang T Y 2020 *Mater. Des.* **187** 108378
[34] Anker A S, Kjær E T S, Juelsholt M, Christiansen T L, Skjærvø S L, Jørgensen M R V, Kantor I, Sørensen D R, Billinge S J L, Selvan R and Jensen K M Ø 2022 *npj Comput. Mater.* **8** 213

[35] Du T, Liu H, Tang L, Sørensen S S, Bauchy M and Smedskjaer M M 2021 *ACS Nano* **15** 17705

[36] Gong X, Li R, Xiao R, Zhang T and Li H 2025 *Chin. Phys. B* **34** 016101

[37] Schoenholz S S, Cubuk E D, Sussman D M, Kaxiras E and Liu A J 2016 *Nat. Phys.* **12** 469

[38] Batzner S, Musaelian A, Sun L, Geiger M, Mailoa J P, Kornbluth M, Molinari N, Smidt T E and Kozinsky B 2022 *Nat. Commun.* **13** 2453

[39] Gunning D 2017 *Explainable Artificial Intelligence (XAI)* (Defense Advanced Research Projects Agency, DARPA/I2O)

[40] Adadi A and Berrada M 2018 *IEEE Access* **6** 52138

[41] Tan M and Le Q 2019 *Proceedings of the 36th International Conference on Machine Learning*, June 9–15, 2019, Long Beach, CA, USA, pp. 6105–6114

[42] Xu N, O'Hern C S and Kondic L 2005 *Phys. Rev. Lett.* **94** 016001

[43] Pollard J and Fielding S M 2022 *Phys. Rev. Res.* **4** 043037

[44] Liu C, Ferrero E E, Puosi F, Barrat J L and Martens K 2016 *Phys. Rev. Lett.* **116** 065501

[45] Selvaraju R R, Cogswell M, Das A, Vedantam R, Parikh D and Batra D 2017 *Proceedings of the IEEE International Conference on Computer Vision*, October 22–29, 2017, Venice, Italy, pp. 618–626

[46] Utter B and Behringer R P 2008 *Phys. Rev. Lett.* **100** 208302

[47] Dubey A K, Procaccia I, Shor C A B Z and Singh M 2016 *Phys. Rev. Lett.* **116** 085502

[48] Krizhevsky A, Sutskever I and Hinton G E 2017 *Commun. ACM* **60** 84

[49] Simonyan K and Zisserman A 2014 arXiv:1409.1556 [cs.CV]

[50] Huang G, Liu Z, van der Maaten L and Weinberger K Q 2017 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, July 21–26, 2017, Honolulu, HI, USA, pp. 2261–2269

[51] He K, Zhang X, Ren S and Sun J 2016 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 27–30, 2016, Las Vegas, NV, USA, p. 770–778