DOI：10.11883/bzycj-2025-0259

基于深度学习的亚稳态高熵合金高应变率冲击响应预测*

刘传志[1,2]，安稳[1,2]，熊启林[1,2]

（1.华中科技大学航空航天学院，武汉 430074；
2. 工程结构分析与安全评定湖北省重点实验室，武汉 430074；

摘 要：亚稳态高熵合金因其在高应变率下优异的力学性能而受到广泛关注。然而，由于对微观结构与冲击响应关系的认识不足，限制了其在高应变率下的工程应用。本研究采用一种结合晶体塑性有限元方法和卷积神经网络的深度学习框架，阐明了微观结构与冲击响应之间的关系。基于晶体塑性模拟收集数据集，该数据集包含高应变率下亚稳态高熵合金在拉伸、压缩和剪切载荷条件下不同织构的完整应力应变响应和相变体积分数的演变。构建了一个双分支卷积神经网络模型，输入为织构和载荷条件。该模型的两个分支用于预测不同的输出：即应力应变曲线与马氏体体积分数的演变。基于收集的数据集对卷积神经网络模型进行训练。结果表明，该模型能够准确预测高应变率条件下亚稳态高熵合金的冲击响应。该研究进一步证明了深度学习框架在保证预测精度的同时，相比晶体塑性有限元模拟具有显著的计算效率优势。为高效评估高应变率下亚稳态高熵合金的力学行为提供了一种新思路。

关键词：深度学习；冲击响应；晶体塑性；亚稳态高熵合金

中图分类号： O347.3 **国标学科代码：**13015 **文献标识码：**A

Deep learning-based crystal plasticity framework to predict the shock response of metastable high-entropy alloy

LIU Chuan-zhi[1,2]，AN Wen[1,2]，XIONG Qi-lin[1,2]

(*1. Department of Mechanics, School of Aerospace Engineering, Huazhong University of Science & Technology,1037 Luoyu Road, Wuhan 430074, China;*
2.Hubei Key Laboratory of Engineering Structural Analysis and Safety Assessment, Luoyu Road 1037, Wuhan 430074, China)

Abstract: Metastable high-entropy alloys (HEA) have attracted considerable attention due to their exceptional mechanical properties at high strain rates. However, their engineering applications are limited by an incomplete understanding of the relationship between microstructure and shock response. An end-to-end deep learning framework has been implemented, combining the crystal plasticity finite element (CPFE) method with a convolutional neural network (CNN) to elucidate the mapping between microstructure and shock response. A crystal plasticity constitutive model, which couples dislocation slip and martensitic transformation mechanisms, has been developed and validated against experimental results, confirming the model's effectiveness. Based on this constitutive model, a dataset for training the deep learning model is generated, including the complete stress-strain response and martensite volume fraction evolution of metastable HEA with corresponding textures and loading conditions at high strain rates. The two-branch CNN model is used to extract microstructural features and make forward predictions of mechanical responses at high strain rates. Its input is microstructural information in image format, including Euler

*收稿日期：2025-08-11； 修回日期：2025-10-24；
基金项目：国家自然科学基金（12172144）；冲击波物理与爆轰物理全国重点实验室基金(2023JCJQLB05403)
第一作者：刘传志（1997— ），男，博士研究生. E-mail: 137668501@qq.com；
通信作者：熊启林（1987— ），男，教授，博士生导师. E-mail: xiongql@hust.edu.cn。

angles and loading conditions, and its output consists of two branches corresponding to stress-strain curves and the evolution of transformed volume fraction, respectively. Based on shared convolutional features and two parallel regression branches, the model predicts the stress-strain curve and martensite volume fraction evolution for a given input. The results show that the model can accurately predict the shock response of metastable HEA under high strain rate conditions. The predicted yield stress, peak stress, and martensite volume fraction under different loading conditions and microstructures are in good agreement with the CPFE results. This study further demonstrates that the deep learning framework offers significant computational efficiency advantages over crystal plasticity finite element simulations while maintaining prediction accuracy. This work helps to clarify the connection between shock response and microstructural evolution in metastable HEA, and lays the foundation for optimizing the shock response of metastable HEA at different temperatures through customized microstructural design.

Keywords: deep learning; high strain rate; crystal plasticity; metastable high entropy alloy

金属材料高应变率下的剪切局域化失效通常会导致尖端装备的灾难性事故，亚稳态高熵合金由于其优异的强度塑性协同，成为抵抗冲击失效的有效候选材料 [1-3]。系统研究和调控亚稳态高熵合金在高应变率下的冲击响应对于拓展其工程应用具有重要的理论价值和实际意义。

作为一种新型的合金体系，亚稳态高熵合金已成为高性能结构材料研究的重要方向。亚稳态是指合金的状态并非热力学上的全局最低能态，通过成分和组织设计引入亚稳态特征，可以显著改善高熵合金的力学性能。亚稳态高熵合金在高应变率下优异的冲击性能得益于位错滑移、孪生和应变诱发马氏体相变等多种变形机制的协同作用 [4-6]。在亚稳态高熵合金中，其较低的热力学稳定性有利于马氏体相变的发生[7, 8]。Fe50Mn30Co10Cr10 在动态单轴拉伸后的组织表征表明发生了马氏体相变，展现出了优异的强度-塑性协同效应 [2]。在动态压缩载荷下，V10Cr10Fe45Co35 表现出 1.6GPa 的极限抗压强度，这归因于马氏体相变机制的激活，从而显著改善了应变硬化，提高了材料的冲击响应 [9]。Yang 等研究了 Fe40Mn40Co10Cr10 在动态压缩下的强度-塑性协同作用，发现马氏体相变的发生在几乎不损失塑性的前提下显著提高了材料强度 [10]。

Xie 等发现 FeNiCrCoCu 高熵合金的冲击相变严重依赖于晶体取向 [11]。高熵合金不同的织构使材料的变形机制发生从位错滑移到位错滑移与孪生的转变，显著影响了材料强度 [12]。Xu 等通过分子动力学模拟发现，沿[001]方向冲击主要产生 HCP 相，而[110]方向则是发现了低密度位错[13]。通过 1045 钢板冲击实验事后表征发现，球化微观结构的不连续屈服归因于低初始位错密度和低位错成核率[14]。目前已经对亚稳态高熵合金的冲击响应进行了部分研究，但是由于定量机制表征的局限性，很难捕捉微观组织演变或载荷条件与冲击响应之间的定量关系。

为了克服这些挑战，晶体塑性有限元方法可以有效地解释亚稳态高熵合金中的变形机制，并捕捉微观结构和载荷条件对冲击响应的影响 [15-17]。然而晶体塑性有限元方法计算量大，受制于计算资源。随着计算机科学的发展，深度学习为研究微观结构演化与载荷条件及冲击响应之间的关系提供了新的视角[19-21]。近期研究越来越多地探索使用深度学习方法加速晶体塑性模拟[22-24]。全连接神经网络已被用于预测变形后织构以及单晶在不同载荷路径下的应力-应变响应 [25]。此外，Tran 等揭示了不同材料微观结构特征与屈服应力之间的跨尺度关系 [26]。然而，这些研究大多受限于模型架构的简单性，通常局限于低维特征映射。随着微观结构描述符数量的增加，基于人工神经网络的模型往往表现出训练不稳定性，并导致准确性下降。

凭借其局部连通性和参数共享等架构特性，卷积神经网络显著减少了高维特征映射任务所需的参数。liu 等成功地利用卷积神经网络预测了异质多晶体中微孔洞的演变 [27]。Han 等实现了对双相钢中应力-应变响应的有效预测 [28]。Inal 等系统地研究了卷积神经网络模型的架构和超参数配置，并进一步将其应用于预测不同织构下的应力-应变曲线 [29]。Carl 等将卷积神经网络应用于增材制造合金，建立了微观结构与应力-应变响应之间的映射 [30]。然而，这些工作主要集中于微观结构和力学响应之间的单一相关性，并没有考虑载荷条件和变形过程中相变等关键物理变量的动态演变。

本研究提出了一个双分支卷积神经网络框架来预测不同织构和载荷条件下的冲击响应（即应力-应变曲线和相变分数的演变）。该框架如图 1 所示，包含四个部分：数据生成、数据预处理、训练和验证以及最终预测。在第二部分中，利用耦合相变机制的晶体塑性有限元本构模型构建数据集。在训练和预测部分，训练了卷积神经网络来捕捉微观结构与应力-应变曲线以及相变分数演变之间的定量关系（见第三部分和第四部分）。

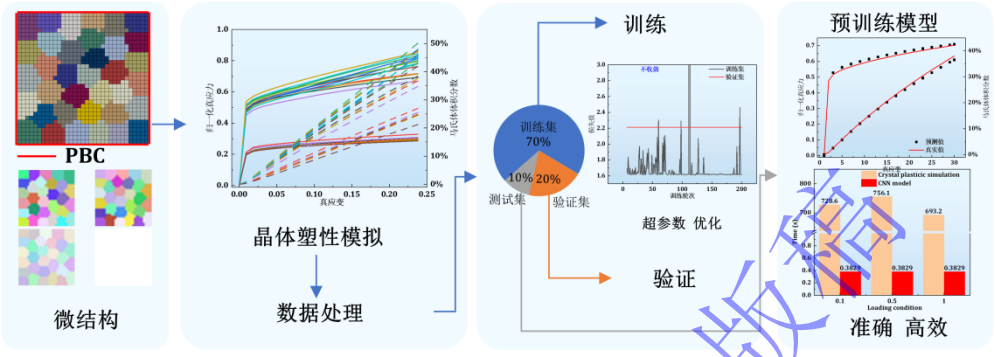

图 1 基于卷积神经网络的冲击响应预测工作流程

Fig.1 Schematic of shock response prediction based on convolutional neural networks

1 方法

1.1 晶体塑性框架

从连续介质力学的角度看，表征从初始构型到当前构型转变的总变形梯度可以乘法分解为弹性分量和塑性分量

$$\mathbf{F}=\frac{\partial x}{\partial X}=\mathbf{F}_e\mathbf{F}_p \#(1)$$

式中： x 和 X 分别代表当前构型和初始构型，$\mathbf{F}_e$ 为变形梯度的弹性部分，$\mathbf{F}_p$ 为变形梯度的塑性部分。

第二 P-K 应力 $\mathbf{S}$ 可以由格林应变通过胡克定律进行计算

$$\mathbf{S}=\mathbb{C}:\mathbf{E}_e \#(2)$$

式中： $\mathbb{C}$ 是四阶弹性张量，对于立方晶体，由三个独立的弹性常数C_{11}、 C_{12}、 C_{44}决定。格林应变张量 $\mathbf{E}_e$ 由弹性变形梯度计算

$$\mathbf{E}_e=\frac{1}{2}\left(\mathbf{F}_e^{\mathrm{T}}\mathbf{F}_e-\mathbf{I}\right) \#(3)$$

塑性变形梯度由塑性速度梯度表示

$$\dot{\mathbf{F}}_p=\mathbf{L}_p\mathbf{F}_p \#(4)$$

塑性速度梯度 $\mathbf{L}_p$ 由所有的塑性变形机制贡献，本文中的塑性变形机制包含位错滑移与马氏体相变。因此塑性速度梯度 $\mathbf{L}_p$ 可以被表示为

$$\mathbf{L}_p=(1-f_{tr})\sum_{\alpha=1}^{N_s}\dot{\gamma}^{\alpha}\mathbf{m}^{\alpha}\otimes\mathbf{n}^{\alpha}+f_{tr}\sum_{\beta=1}^{N_{tr}}\dot{\gamma}_{tr}^{\beta}\mathbf{m}_{tr}^{\beta}\otimes\mathbf{n}_{tr}^{\beta} \#(5)$$

式中： $f_{tr}=\sum_{\beta=1}^{N_{tr}} f^{\beta}$ 为总的相变体积分数， $\dot{\gamma}^{\alpha}$ 为滑移系 α 上的剪切应变率，m 和 n 为滑移方向与滑移面法向，$\dot{\gamma}_{tr}^{\beta}$ 为相变系 β 上的剪切滑移率，$\mathbf{m}_{tr}$ 和 $\mathbf{n}_{tr}$ 为变形方向与惯习面法向，$N^{s}=N^{tr}=12$ 为 FCC 结构中滑移系与相变系的个数。

通过 Orowan 方程得到每个滑移系统的剪切应变率

$$\dot{\gamma}^{\alpha}=\rho^{\alpha}bv^{\alpha}\#(6)$$

式中： ρ^{α} 是滑移系的位错密度，b 为伯氏矢量的大小，位错运动速度 v^{α} 被表示为

$$v^{\alpha}=v_{0}exp\left\{\frac{-Q}{kT}\left\{1-\left\{\frac{\left\langle\tau^{\alpha}-\tau_{LR}^{\alpha}\right\rangle}{\tau_{SR}^{\alpha}}\right\}^{p}\right\}^{q}\right\}\mathrm{sgn}(\tau^{\alpha})\#(7)$$

式中： Q=3.5×10-16mJ 为激活能，k=1.38×10-20 mJ/k 为玻尔兹曼常数，T 为瞬态温度，p 和 q 为热激活参数。τ^{α}、τ_{LR}^{α}、τ_{SR}^{α} 分别代表滑移系 α 上的分切应力、长程阻力、短程阻力 [31]。

位错密度的演变可以表示为

$$\dot{\rho}^{\alpha}=\frac{k_{mult}|\dot{\gamma}^{\alpha}|}{b\Lambda^{\alpha}}-\frac{2k_{anni}}{b}\rho^{\alpha}|\dot{\gamma}^{\alpha}|\#(8)$$

式中： k_{mult}、k_{anni} 代表位错增值系数与湮灭系数，Λ^{α} 为位错运动的平均自由程 [32]

$$\frac{1}{\Lambda^{\alpha}}=\frac{1}{d}+\frac{1}{\lambda_{\mathrm{slip}}^{\alpha}}+\frac{1}{\lambda_{\mathrm{tr}}^{\alpha}}\#(9)$$

式中： d 为平均晶粒尺寸，第二项表示由位错滑移引起的自由程变化 $\frac{1}{\lambda_{\mathrm{slip}}^{\alpha}}=\frac{\varphi}{i_{\mathrm{slip}}}$， φ 是相互作用参数，i_{slip} 代表滑移系之间的相互作用，第三项表示由相变引起的自由程变化，可以被写为 [31]

$$\frac{1}{\lambda_{\mathrm{tr}}^{\alpha}}=\sum_{\beta=1}^{N_{tr}}\xi_{\alpha\beta}f^{\beta}\frac{1}{H_{tr}(1-f_{tr})}\#(10)$$

式中： $\xi_{\alpha\beta}$ 为相互作用系数。

相变对剪切滑移速率的贡献是通过相变体积分数的变化来实现的。将经典相变模型耦合到晶体塑性框架中，每一个相变系上的滑移率 $\dot{\gamma}_{tr}^{\beta}$ 可以被表示为 [32]

$$\dot{\gamma}_{tr}^{\beta}=\gamma_{tr}\dot{f}_{tr}^{\beta}\#(11)$$

式中： $\dot{f}_{tr}^{\beta}$ 是相变体系中相变体积分数的变化率， γ_{tr} 代表特征剪切应变。对于 FCC 金属而言被设置为$\sqrt{2}/2$。

β 相变体系体积分数的演变速率可表示为 [33]

$$\dot{f}_{tr}^{\beta}=\dot{\mathrm{f}}_{max}\tanh\left(\frac{1}{v_{tr}}\frac{\left\langle f_{tr}^{\beta}-f_{cr}^{\beta}\right\rangle}{f_{cr}^{\beta}}\right)\#(12)$$

式中： $\dot{f}_{max}$ 是最大相变速率，用于控制相变演化速率。f_{cr}^{β} 为 β 相变系上的临界相变阻力。

此外，在各相变体系中，随着相变体积分数的增加，相变的发生变得越来越困难。因此，临界相变阻力的演变应该与相变体系中的体积分数有关 [33]

$$\dot{f}_{cr}^{\beta}=\sum_{\beta=1}^{N_{tr}}\Re\left|\dot{f}_{tr}^{\beta}\right|\#(13)$$

式中： $\Re$ 是相变硬化参数，随着相变的发生，相变阻力会增加。这里假设在任何系统中，相变阻力的增加都是相同的。

1.2 深度学习模型

作为深度学习领域的常见架构，卷积神经网络相较于全连接神经网络具有显著优势，尤其是在图像识别和计算机视觉任务中。卷积神经网络通过连续的卷积层和最终的全连接层处理图像，有效地提取层次化特征，并最终实现从特征空间到目标输出的映射关系。本研究将材料的微观结构转换为多通道图像格式，以更直观地表征其空间分布和物理特性。

卷积运算的数学公式由以下方程描述

$$Y_{i,j}^{L}=\sum_{s}\sum_{t}\boldsymbol{X}_{i+s,j+t}^{L-1}\boldsymbol{W}_{s,t}^{L-1}+\boldsymbol{b}^{L-1}\#(14)$$

式中：$\boldsymbol{b}^{L-1}$ 是偏置。 $\boldsymbol{W}_{s,t}^{L-1}$ 与 $\boldsymbol{b}^{L-1}$ 为可训练参数。

卷积层之后通常会使用池化层来保留关键特征，同时降低特征图维度和模型复杂度。池化有两种方式：最大池化和平均池化。最大池化从局部区域选择最大值，而平均池化计算局部区域的平均值。图 2 展示了卷积和最大池化的示意图。

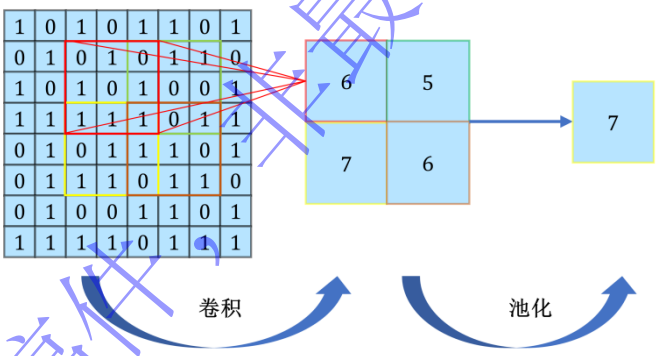

图 2 卷积与池化示意图

Fig.1 Schematic of Convolution and pooling

每一层的输出通常通过激活函数进行非线性变换，这对于增强卷积神经网络模型的表达能力至关重要。本研究选择了两个激活函数：tanh 和 ReLU。tanh 函数用于将输出压缩在[-1, 1]范围内，而 ReLU 函数则有效缓解梯度消失问题，加速网络的训练过程。

$$\tanh(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}},\mathrm{ReLU}(x)=\max(0,x)\#(15)$$

在卷积神经网络训练过程中，损失函数用于衡量预测值与真实值之间的差异。对于回归问题，均方误差（MSE）通常用作损失函数，因为它能够有效地量化预测值与实际值之间的偏差

$$\mathrm{MSE}\left(Y_{pred}^{\ i},Y_{tar}^{\ i}\right)=\frac{1}{n}\sum_{i=1}^{n}\left(Y_{tar}^{\ i}-Y_{pred}^{\ i}\right)^{2}\#(16)$$

在卷积神经网络模型的训练过程中，优化器起着至关重要的作用。它根据损失函数的反馈调整模型参数，最小化损失值，从而提升预测能力和整体性能。常用的优化方法是 Adam 优化器，它结合了动量和自适应学习率机制，在处理复杂问题时表现出色。对于卷积神经网络模型来说，增加训练参数

的数量有助于提高准确率，但也存在过拟合的风险。为了解决这个问题，在模型中引入正则化技术来加速训练过程并有效防止过拟合。

2 结果

2.1 参数标定

参考 Ravi 等 [34]对亚稳态高熵合金 Fe38.5Mn20Co20Cr15Si5Cu1.5 在高应变率下的实验，采用相同的加载条件进行晶体塑性建模。由于在高应变率下生成数据集的计算成本较高，因此采用二维多晶模型高效地构建数据集。二维代表性体积元由 30 个取向随机分配的晶粒组成。对代表体元施加周期性边界条件，如图 3 （a）所示。从侧面施加 2.5 m/s 的速度载荷，实现 2500 s-1 的压缩应变率。为了可靠地评估材料的冲击响应，采用斜坡速度曲线引入速度载荷，如图 3（b） 所示。

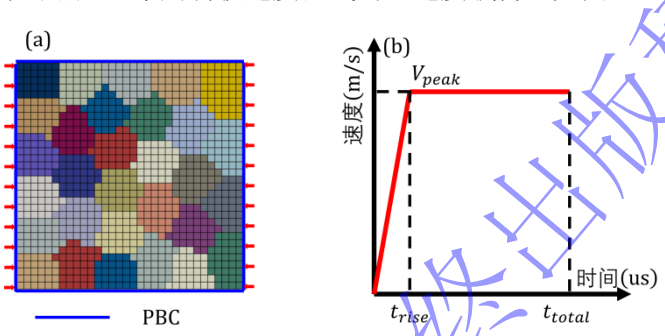

图 3 随机取向多晶模型的 RVE：（a）RVE（32×32）的网格和边界条件，（b）速度载荷曲线

Fig. 3. RVE of the randomly oriented polycrystal model: (a) RVE (32×32) mesh and boundary conditions, (b) velocity load curve

图 4（a）为应力-应变曲线模拟结果与实验结果对比，图 4（b）为马氏体体积分数对比。可以看出，模拟结果与实验数据吻合较好。图 4（a）还给出了不考虑马氏体相变的模拟结果得到的应力-应变响应（红色虚线）。结果表明，在变形初期，模拟曲线与实验数据吻合较好。然而，随着塑性变形的加剧，两条曲线之间的差异逐渐出现并增大。这种偏差主要归因于该模型没有引入马氏体相变诱导的强化机制，从而低估了亚稳态高熵合金在大应变阶段的冲击性能。图 4 （c）进一步展示了马氏体形核的临界阻力对亚稳态高熵合金冲击响应的影响，可以看到随着 f_{cr} 的增加，马氏体相变的发生变得越来越困难。当 $f_{cr}=280\text{MPa}$,马氏体相变受到明显抑制，材料的强度进一步降低。由此可见，临界相变阻力的提升增大了马氏体相变发生的门槛，降低了马氏体体积分数及其演化速率，最终影响了材料强度。在本工作中，晶体塑性有限元模型的参数如表 1 所示

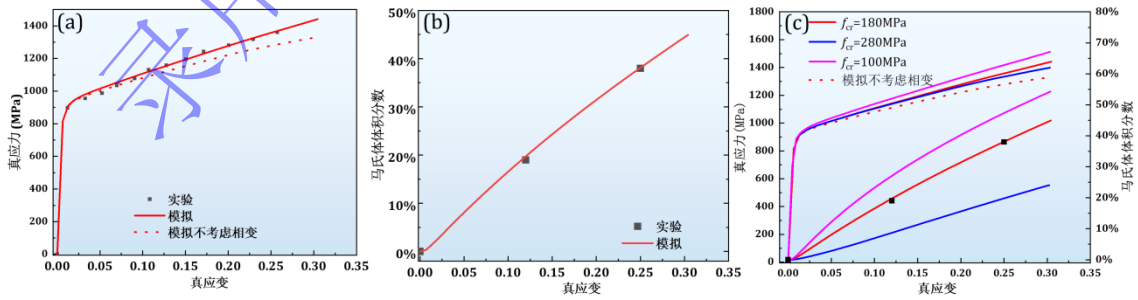

图 4 晶体塑性模拟与实验的比较：（a）有无相变的应力-应变曲线，（b）马氏体体积分数的演变，（c）马氏体的临界阻力对冲击响应的影响

Fig. 4 Comparison of crystal plasticity simulation and experiment: (a) Stress-strain curves with and without phase transformation, (b) Evolution of martensite volume fraction, (c) Effect of critical resistance of martensite on impact response

表 1 晶体塑性模型本构模型参数

Table 1 Constitutive model parameters of crystal plasticity model

参数	名称	值（单位）	来源
ρ	密度	7655 Kg·m^{-3}	[34]
C_p	比热容	430 J/（K·Kg）	[34]
C_{11}	弹性参数	174.2GPa	[35]
C_{12}	弹性参数	97.9GPa	[35]
C_{44}	弹性参数	139.7GPa	[35]
α_{SR}	短程阻力系数	0.29	（本工作）
α_{LR}	长程阻力系数	0.7	（本工作）
$\xi_{\alpha\beta}$	滑移与相变的相互作用系数	0 （共面）	（本工作）
		0.6 （非共面）	（本工作）
α_{mult}	位错增值系数	0.0017	（本工作）
α_{anni}	位错湮灭系数	0.24	（本工作）
f_{cr}	临界相变阻力	180MPa	（本工作）

2.2 数据集

数据集的质量在卷积神经网络模型的训练过程中起着至关重要的作用，因为它直接影响模型的收敛速度和整体性能。高质量的数据集可以提供更相关、更丰富的特征，从而显著加快收敛过程。

本研究生成了 1500 个随机取向，将其分配给有限元模型，采用三种不同的载荷条件：拉伸、压缩和剪切。这种配置能够探索不同微观结构和载荷条件下亚稳态高熵合金在高应变速率下的应力-应变曲线和相体积分数演变曲线。值得注意的是，这 1500 个样本获得了令人满意的结果。虽然增加数据集大小可以提高训练的准确性，但选择 1500 个样本是经过仔细考虑数据集生成所涉及的计算成本和时间限制后做出的。

卷积神经网络模型的每个输入都表示为一个维度为（32，32，4）的张量。前两个维度对应于代表体元的网格，每个像素代表有限网格中的相应位置。第三个维度代表图像的四个通道：前三个通道是三个欧拉角，第四个通道代表代表体元的载荷情况。当载荷为拉伸时，第四个通道的所有值被设置为 0.1，载荷为压缩时，第四个通道的所有值被设置为 0.5，剪切被设置为 1。正如图 5 所示，一个包含 1024 个高斯点的代表体元就被有效地变换成一个 32×32 的彩色图像。这种变换在保留晶粒形貌的同时，还能捕捉到每个晶粒的取向及其周围晶粒的取向信息以及载荷情况，这有利于后续卷积神经网络的训练。

在本研究中，每个微观组织都与一条应力-应变曲线和马氏体体积分数的演变相关联，两者均被离散化为 30 个数据点。这样，微观组织、载荷条件以及相应的宏观力学响应（包括应力-应变曲线和相体积分数的演变）作为输入输出对输入，用于训练卷积神经网络模型。

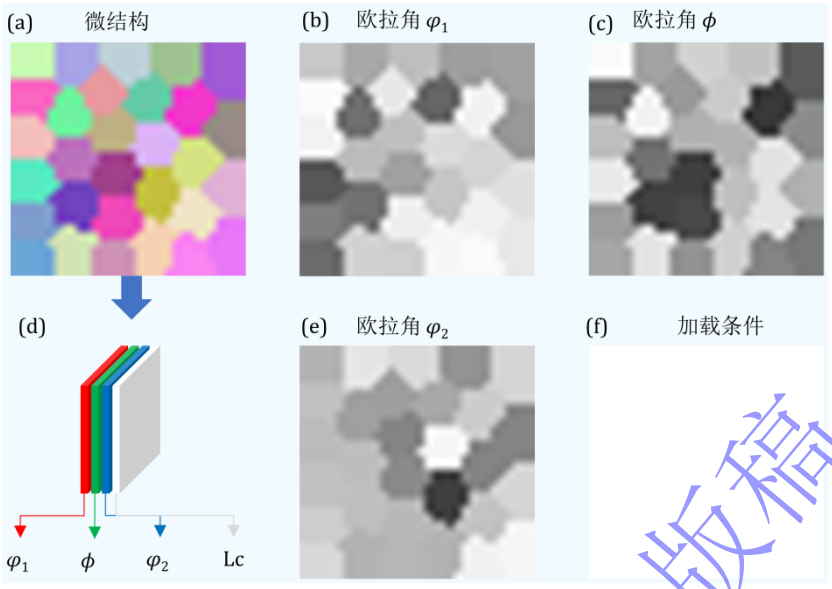

图 5 数据集中代表性样本：（a）四通道微结构，（b）第一通道欧拉角 φ1，（c）第二通道欧拉角 φ，（d）图像预处理示意图，（e）第三通道欧拉角 φ2，（f）第四通道加载条件

Fig. 5 Representative samples in the dataset: (a) Four-channel microstructure, (b) Euler angle φ1 of the first channel, (c) Euler angle φ of the second channel, (d) image preprocessing diagram, (e) Euler angle φ2 of the third channel, and (f) loading conditions of the fourth channel.

如上文所述，通过晶体塑性模拟获得了不同微观结构在拉伸、压缩和剪切载荷条件下的宏观力学响应。然后，这些响应通过特定的表示方法转换为用于卷积神经网络模型训练的数据集。训练前，数据集被随机打乱，并按 7:2:1 的比例划分为训练集、验证集和测试集。图 6 显示了不同数据集中屈服应力、峰值应力和相体积分数演变的分布。屈服应力、峰值应力和相体积分数演变的统计数据表明，各数据集中的比例基本一致。因此，每个数据集中的数据能够充分反映力学响应的统计分布。这种数据集划分对于深度学习模型的训练、验证和测试是合理的。

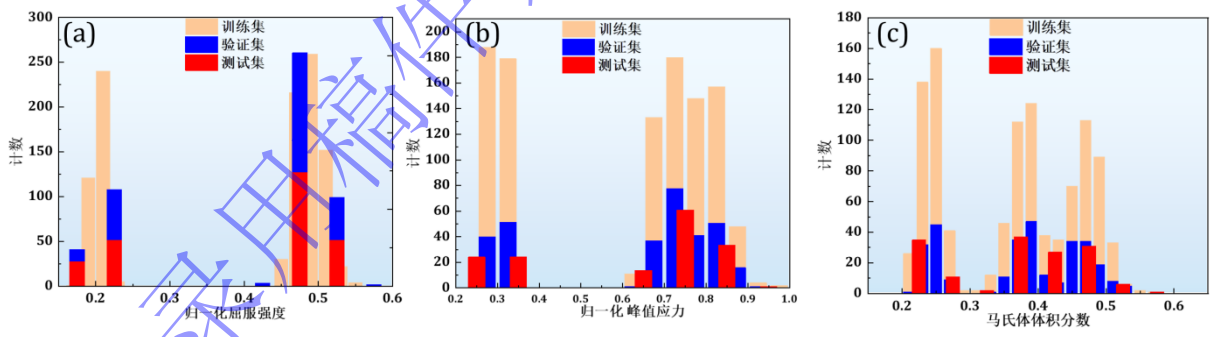

图 6 图 6 训练集、验证集和测试集中的数据分布：（a）屈服强度、（b）峰值应力、（c）相体积分数

Fig. 6 Data distribution in the training, validation, and test sets: (a) yield strength, (b) peak stress, and (c) phase volume fraction

2.3 卷积神经网络模型的训练和测试

卷积神经网络的架构设计和超参数调优对其预测性能和收敛性有显著的影响。在本节中，训练了一个高效的卷积神经网络模型，用于准确预测高应变率下亚稳态高熵合金的应力-应变响应和相变体积分数。本研究使用包含 1050 个样本的数据集进行训练，共进行了 200 个迭代循环。在每次训练迭代中，批次大小设置为 64，学习率为 1e-3，并使用 Adam 优化器。这种配置不仅有助于加速训练过程，还能确保训练过程的稳定性，同时降低模型收敛到局部最优的风险。Adam 优化器根据梯度调整学习率，由于其能够有效地管理参数更新中的动态调整，特别适合训练复杂模型。

本研究采用的卷积神经网络模型由三个卷积层和三个全连接层组成。每个卷积层后接一个最大池化层，该层在保持特征空间不变性的同时，降低了特征图的维度和计算量，从而有效地减少了训练所需的参数数量。在全连接层中，引入了 Dropout 正则化以防止过拟合。Dropout 的工作原理是在训练过程中随机停用一定比例的神经元，从而破坏神经元之间的相互依赖关系，提高卷积神经网络模型的泛化能力。

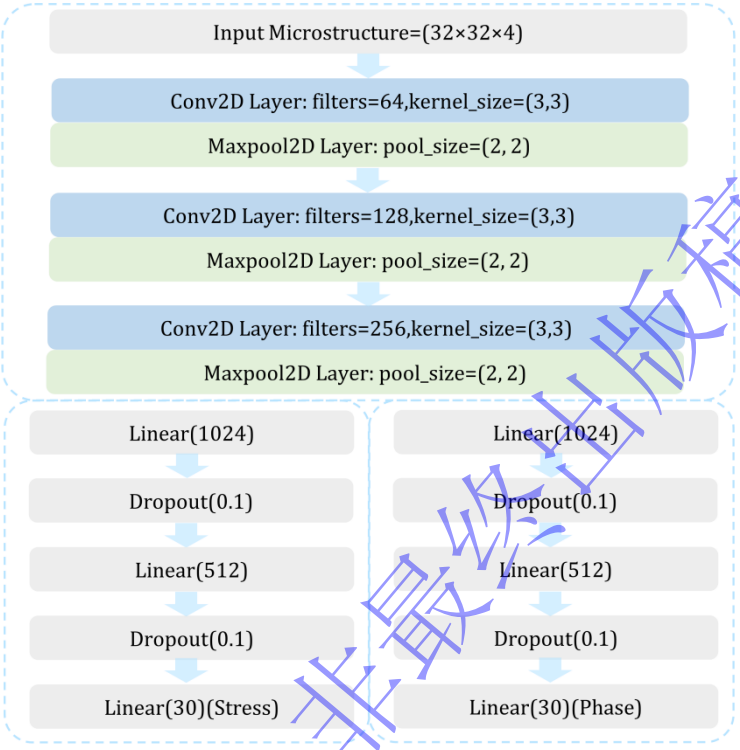

图 7 具有双分支回归器的卷积神经网络模型的详细架构

Fig. 7 Convolutional neural network model architecture with a two-branch regressor

本研究针对微结构-响应映射的单输入双输出特性（即单个微结构同时对应应力-应变曲线和相体积分数演化曲线），设计了一个基于参数共享的多任务学习框架。该网络架构由一个特征编码器和一个双分支回归器组成。该结构能够有效地学习不同微结构在不同载荷条件下的响应模式，从而提升其在多任务学习场景下的性能。卷积神经网络的具体架构如图 7 所示。

经过 200 个轮次的训练，模型达到了稳定的收敛状态。如图 8（a）所示，由于参数初始化的随机性，卷积神经网络模型初始损失值较高，此时模型尚未学习到有效的特征。然而，随着训练的进行，损失值迅速下降并稳定在一个较低的水平，表明模型已经逐步优化，获得了较强的数据拟合能力。此外，在训练过程结束时，训练损失（0.0002）和验证损失（0.0014）均保持在较低水平，表明模型在训练集和验证集上均取得了良好的优化性能。

从图 8（b）中的误差曲线可以看出，训练误差随着训练迭代次数的增加而快速下降，最终稳定在较低水平。这进一步证实了模型能够有效地拟合训练数据。此外，验证误差也呈现出与训练误差类似的下降趋势，既没有出现明显的波动，也没有出现大幅的上升。这表明模型在训练过程中并未出现明显的过拟合现象，在验证集上始终保持较低的误差，展现出良好的泛化能力。

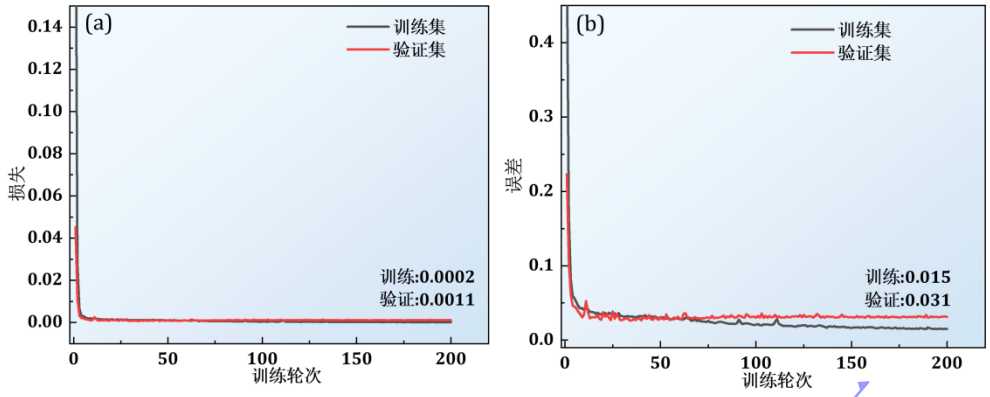

图 8 训练集与测试集上损失函数与误差随训练过程的演化

Fig. 8 Evolution of loss function and error on training and test sets during training

基于训练好的卷积神经网络模型，计算了测试集样品的力学响应和相体积分数演变的预测结果，并将预测值与晶体塑性模型结果的对比如图 9 所示。需要强调的是，测试集的数据没有参与训练过程，因此这些测试样本对于训练好的卷积神经网络模型来说是全新的。

图 9 （a）、（b）和（c）展示了三个关键变量的回归分析结果：归一化屈服强度、归一化峰值应力和最大相体积分数。x 轴和 y 轴分别表示这些变量的预测值和真实值。从回归分析中可以看出，这三个变量的预测值与真实值均表现出较高的相关性，R^2 值分别为 0.9897、0.9886 和 0.9495。其中，归一化屈服强度和归一化极限强度的 R^2 值接近 1，反映出预测值与真实值之间存在极强的线性关系。这表明该模型对这两个变量具有出色的预测精度。虽然最大相体积分数的 R^2 值略低，但仍在一定范围内。各子图右下角插入的柱形图给出了对应变量绝对误差的相对频率分布：误差分布整体呈左偏，绝大多数样本集中在低误差区域内，其中图 9（a）与（b）的误差分布更为集中。在测试集中超过 85%的样本屈服强度的误差值小于 40MPa，75%样本的峰值应力误差值小于 70MPa。图 9（c）的误差分布相对分散，但 90%以上的误差都保持在 4%以内，这也进一步印证了回归分析的结果。

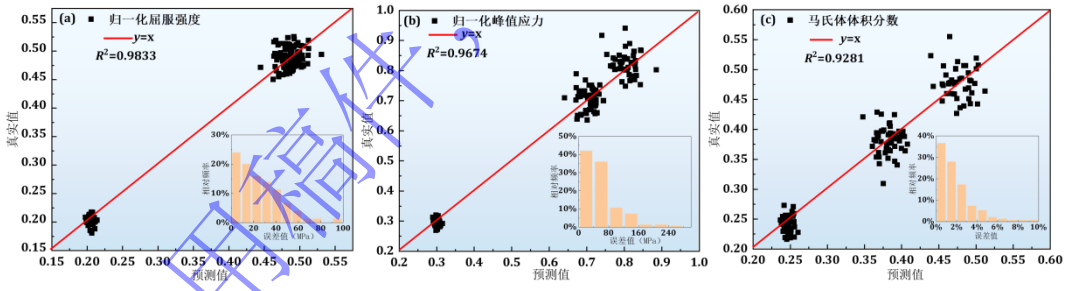

图 9 测试集上的预测结果：（a）归一化屈服强度（b）归一化极限强度和（c）相体积分数

Fig. 9 Prediction results on the test set: (a) normalized yield strength (b) normalized ultimate strength and (c) phase volume fraction

图 10（a）、（b）和（c）展示了卷积神经网络模型预测拉伸、压缩和剪切应力-应变曲线以及相体积分数演变的性能。尽管预测值与真实值之间存在微小偏差，但模型整体表现出较高的准确性和可靠性。这些结果进一步验证了卷积神经网络模型在模拟复杂力学行为和相变过程方面的有效性，并表明该模型能够对不同载荷条件下的材料响应提供高质量的预测。

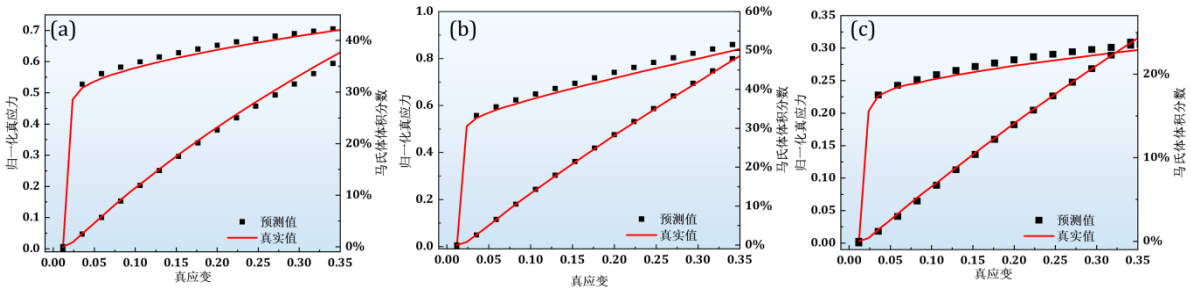

图 10 测试集上预测值与真实值的对比：（a）拉伸、（b）压缩和（c）剪切的真实应力和相体积分数曲线

Fig. 10 Comparison of predicted and true values on the test set: true stress and phase volume fraction curves for (a) tension, (b) compression, and (c) shear

3 讨论

3.1 学习率对训练过程的影响

学习率是卷积神经网络训练过程中最关键的超参数之一，它决定了梯度下降过程中参数更新的步长，直接影响模型的收敛速度、稳定性以及整体的优化性能。通常，较大的学习率会导致参数更新步长过大，模型难以稳定地逼近全局最优，甚至可能引发梯度爆炸或参数振荡，最终导致训练过程无法收敛。相反，较小的学习率通常可以获得更稳定的训练过程，但同时也会减慢收敛速度，降低优化效率，并可能导致模型陷入局部最优。

为了探究不同学习率大小对模型性能的影响，采用不同数量级的学习率训练了卷积神经网络。具体来说，学习率设置为 10-n（n=1，2，3，4），并使用 Adam 优化器。监测训练损失和验证损失，以评估不同学习率下的优化行为。如图 11（a）所示，当学习率设置为 0.1 时，训练损失出现了明显的震荡，甚至出现了极端的数值波动，表明模型完全无法收敛。同时，验证损失一直处于较高水平，没有下降趋势，这表明模型无法有效地学习到有意义的数据特征。这种不稳定性是由于学习率过大导致参数更新过于激进，最终导致模型持续超越最优点而无法稳定下降，最终训练失败。图 11（b）中，当学习率降低至 0.01 时，训练过程开始呈现一定程度的收敛，但仍然出现了明显的震荡。虽然训练损失相比 0.1 的情况有显著降低，但在某些训练阶段仍然出现了损失的突然飙升。这种现象很可能是由于学习率仍然较高，导致梯度更新不稳定，容易出现梯度爆炸或参数波动过大的情况。图 11（c）展示了学习率为 0.001 时的训练场景。此时，训练损失和验证损失在训练初期都迅速下降并最终趋于稳定。在整个训练过程中，损失曲线没有出现明显的震荡，表明模型能够稳定地进行优化，并成功捕捉到数据中潜在的模式。与前两种情况相比，0.001 的学习率能够使模型在更短的训练时间内达到最优收敛状态。最后，图 11（d）展示了学习率为 0.0001 的训练过程。虽然整体收敛效果令人满意，但训练损失和验证损失略高于学习率为 0.001 时的数值。这表明，虽然较小的学习率有利于优化过程更加稳定，但也会减慢收敛速度，需要更多的训练轮数才能达到相当的收敛水平。基于这些观察，本研究选择 0.001 作为最佳学习率，以平衡收敛速度和优化稳定性。

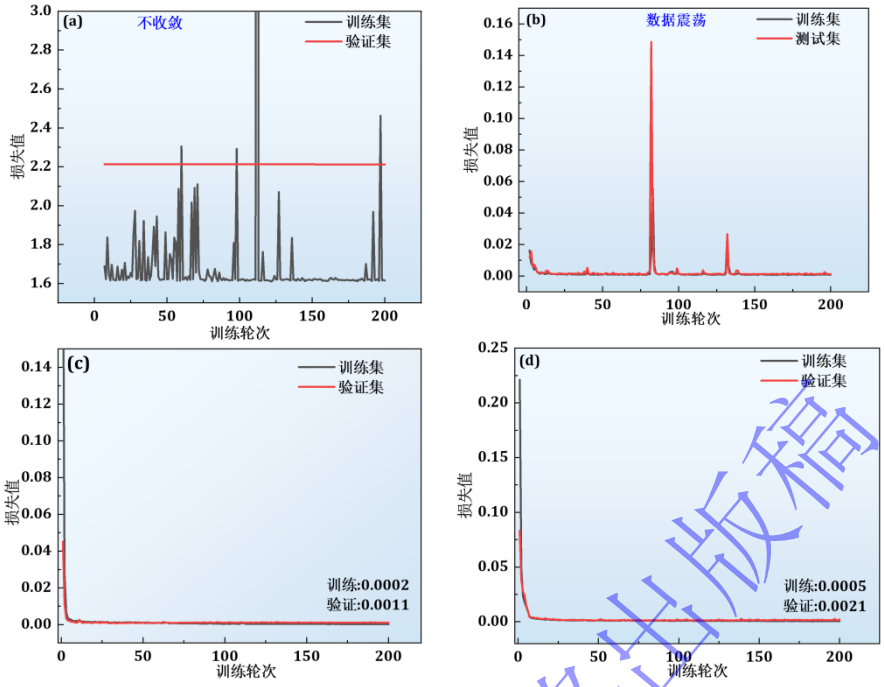

图 11 不同学习率的训练过程：（a）学习率=0.1，（b）学习率=0.01，（c）学习率=0.001，（d）学习率=0.0001

Fig. 11 Training process with different learning rates: (a) learning rate = 0.1, (b) learning rate = 0.01, (c) learning rate = 0.001, (d) learning rate = 0.0001

3.2 优化器对训练过程的影响

在卷积神经网络的训练过程中，优化器的选择对模型的收敛速度、稳定性以及整体优化性能至关重要。本研究比较了四种不同优化器的优化性能：随机梯度下降（SGD）、带动量梯度下降的 SGD+Momentum、RMSprop 和 Adam，并分析了它们在损失收敛趋势上的差异。为了确保公平比较，所有优化器均采用固定的 0.001 学习率进行训练，并保持一致的实验条件。

正如图 12 所示，不同的优化器在收敛速度、稳定性和最终损失方面存在显著差异。SGD 仅根据当前梯度更新参数，导致训练过程相对较慢。虽然最终收敛，但最终损失仍然较高。SGD+momentum 通过引入动量项，在一定程度上加快了收敛速度，并减少了训练过程中的震荡。然而，其最终损失仍然较高，并未达到最优性能。

RMSprop 作为一种自适应学习率优化器，在本研究中展现出更快的收敛趋势，但其损失曲线存在明显的震荡和突变，表明训练过程缺乏稳定性，可能会影响模型的泛化能力。相比之下，Adam 优化器融合了 momentum 和 RMSprop 的优势，训练过程更平滑，收敛速度最快，最终损失最低，在收敛速度和损失最小化方面取得了最佳的综合表现。

总体而言，在给定的实验条件下，Adam 优化器表现出最优性能，它不仅加速了卷积神经网络训练的收敛，还有效降低了损失，从而提升了模型的整体预测能力。因此，对于高应变率条件下不同微观组织的亚稳态高熵合金的力学性能预测，Adam 优化器展现出优异的收敛性和鲁棒性，从而提供更稳定、高效的训练过程。

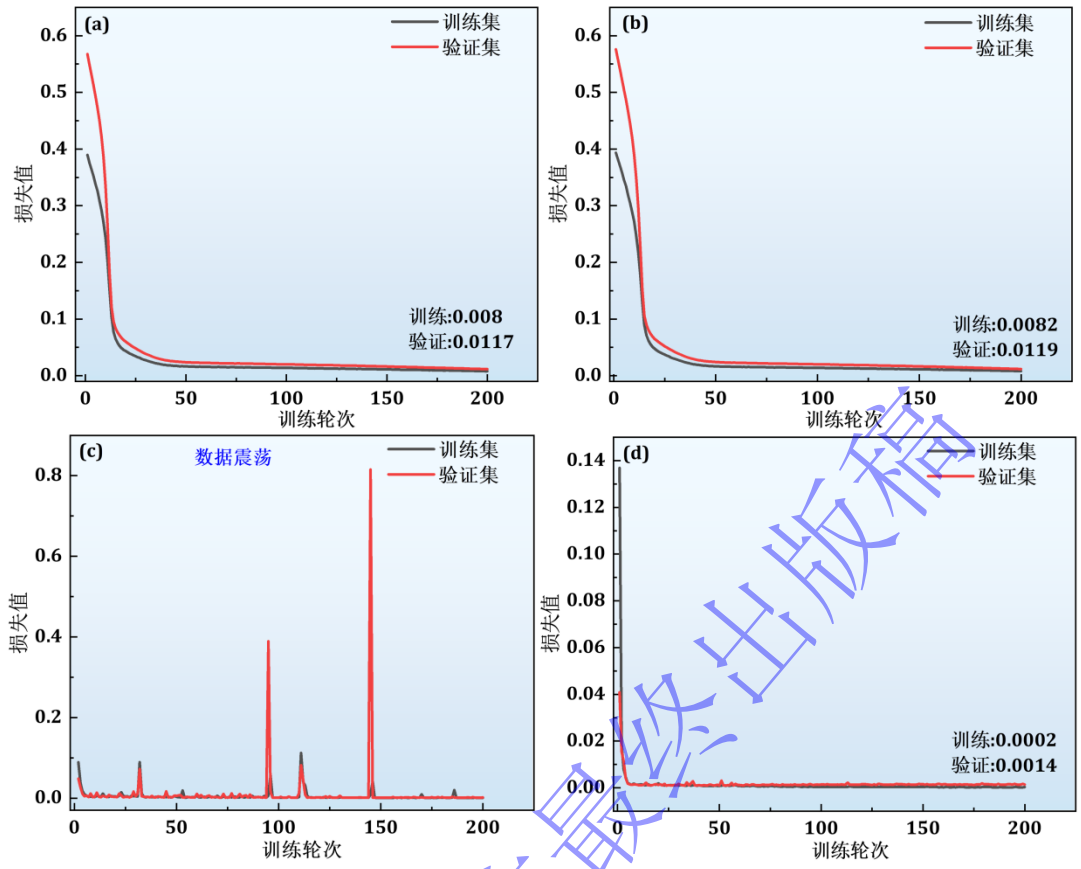

图 12 不同优化器的训练过程：(a) SGD, (b) SGD+moment, (c) RMSprop, (d) Adam

Fig. 12 Training process of different optimizers: (a) SGD, (b) SGD+moment, (c) RMSprop, (d) Adam

3.3 计算效率比较

从图 13 可以看出，基于卷积神经网络的预测模型在计算效率方面具有显著优势：当载荷分别为拉伸（加载条件为 0.1），压缩（加载条件为 0.5），剪切（加载条件为 1）时，它仅需一次快速前向传播即可完成材料力学行为的预测，而无需晶体塑性模拟逐步求解复杂的增量迭代方程。对于晶体塑性模拟，需要对材料内部位错、滑移系激活等微观机制进行高保真计算，往往需要大量的迭代和收敛判断，耗时较长；而卷积神经网络模型经过离线训练后，只需进行简单的张量运算即可在极短的时间内完成推理，尤其在批量处理或在线监测场景中，可以显著提升效率。

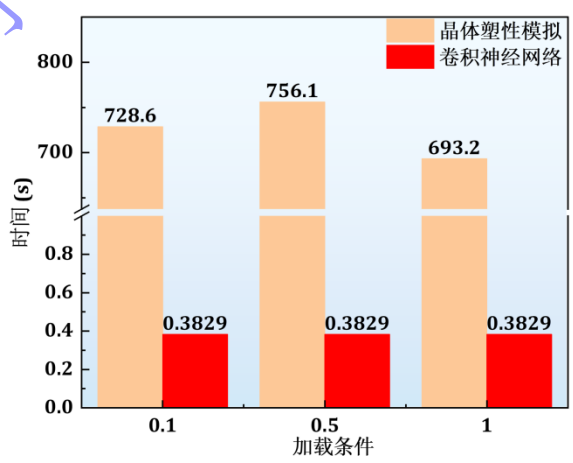

图 13 晶体塑性模拟与卷积神经网络计算效率比较

Fig. 13 Comparison of computational efficiency between crystal plasticity simulation and convolutional neural network

为了充分验证深度学习框架的预测能力，在随机加载条件下下预测了亚稳态高熵合金的冲击响应过程。从图 14 对比了材料在经历压缩、剪切、拉伸、剪切、压缩变形过程后晶体塑性模拟结果与卷积神经网络的预测结果。可以看到，模型在各个加载阶段均能准确捕捉材料的冲击响应，仅在大塑性末段出现轻微高估。总体而言，应力-应变行为被准确地捕获，进一步证实了卷积神经网络框架预测亚稳态高熵合金冲击响应的准确性。

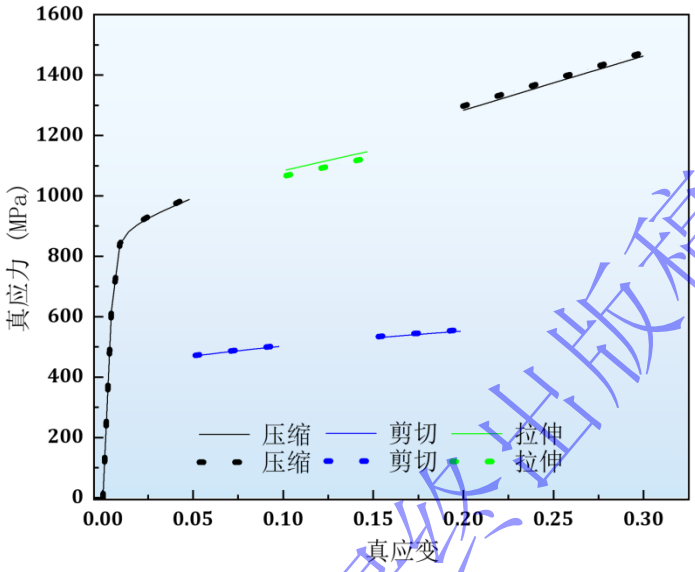

图 14 变加载条件下晶体塑性模拟与卷积神经网络结果对比（实线：晶体塑性有限元模拟；虚线：卷积神经网络）

Fig. 14 Comparison of crystal plasticity simulation and convolutional neural network results under variable loading conditions (solid line: crystal plasticity finite element simulation; dashed line: convolutional neural network)

4 结论

在本研究中，提出了一个将晶体塑性与卷积神经网络耦合的计算框架，以实现对不同微观结构和载荷条件下应力-应变曲线和相体积分数演变的可靠预测。通过系统地比较不同的优化器和学习率，确定了对数据集表现出优异预测性能的最优卷积神经网络模型参数。最后，利用测试数据集验证了训练后的卷积神经网络模型的稳健性，这为研究高应变率条件下亚稳态高熵合金的变形行为提供了宝贵的见解。本研究的主要结论可以概括如下：

1. 随机生成共 1500 个不同的微结构，并将其纳入不同载荷条件下的晶体塑性有限元模型。将生成的数据集划分为训练集、验证集和测试集，以支持卷积神经网络的训练。

2. 所提出的计算框架将晶体塑性与卷积神经网络相结合，在预测高应变率下亚稳态高熵合金的力学行为方面，与晶体塑性有限元模拟表现出高度一致性。这种一致性在测试集中同样得到了证实。

3. 与晶体塑性有限元模拟相比，提出的基于卷积神经网络的框架显著提高了计算效率。基准测试结果表明，计算时间从分钟级缩短至亚秒级，同时保持了较高的预测精度。

参考文献

[1].George E P, Curtin W A, Tasan C C. High entropy alloys: a focused review of mechanical properties and deformation mechanisms[J]. Acta Materialia, 2020. 188: p. 435-474.**DOI:** 10.1016/j.actamat.2019.12.015.

[2].He Z F, Jia N, Wang H W, et al. The effect of strain rate on mechanical properties and microstructure of a metastable FeMnCoCr high entropy alloy[J]. Materials Science and Engineering: A, 2020. 776: p. 138982.**DOI:** 10.1016/j.msea.2020.138982.

[3]. Zhang N B, Zhang C X, Li B, et al. Impact response of metastable dual-phase high-entropy alloy Cr10Mn30Fe50Co10[J]. Journal of Alloys and Compounds, 2023. 965: p. 171341.**DOI:** 10.1016/j.jallcom.2023.171341

[4]. Xu J, Liang L, Tong W, et al. Role of strain rate in phase stability and deformation mechanism of non-equiatomic Fe38-xMn30Co15Cr15Ni2Gdx high-entropy alloy[J]. Materials Characterization, 2022. 194: p. 112356.**DOI:** 10.1016/j.matchar.2022.112356.

[5]. Wang P, Bu Y, Liu J, et al. Atomic deformation mechanism and interface toughening in metastable high entropy alloy[J]. Materials Today, 2020. 37: p. 64-73.**DOI:** 10.1016/j.mattod.2020.02.017.

[6]. Wang K, Cheng Z, Liu C, et al. Deformation behavior and strengthening mechanisms of high-entropy alloys under high strain rate across wide temperature ranges[J]. International Journal of Plasticity, 2025: p. 104321.**DOI:** 10.1016/j.ijplas.2025.104321.

[7]. Wang X, De Vecchis R R, Li C, et al. Design metastability in high-entropy alloys by tailoring unstable fault energies[J]. Science Advances, 2022. 8(36): p. eabo7333.**DOI:** 10.1126/sciadv.abo7333.

[8]. Werner K V, Naeem M, Niessen F, et al. Experimental and computational assessment of the temperature dependency of the stacking fault energy in face-centered cubic high-entropy alloys[J]. Acta Materialia, 2024. 278: p. 120271.**DOI:** 10.1016/j.actamat.2024.120271.

[9]. Song H, Kim D G, Kim D W, et al. Effects of strain rate on room-and cryogenic-temperature compressive properties in metastable V10Cr10Fe45Co35 high-entropy alloy[J]. Scientific Reports, 2019. 9(1): p. 6163.**DOI:** 10.1038/s41598-019-42704-x.

[10]. Yang J, An W, Liu C, et al. Strength-plasticity synergy of metastable fe40mn40co10cr10 high entropy alloy at high strain rate and cryogenic temperature[J]. Materials Science and Engineering: A, 2025: p. 148635.**DOI:** 10.1016/j.msea.2025.148635.

[11]. Xie H, Ma Z, Zhang W, et al. Phase transition in shock compressed high-entropy alloy FeNiCrCoCu[J]. International Journal of Mechanical Sciences, 2023. 238: p. 107855.**DOI:** 10.1016/j.ijmecsci.2022.107855.

[12]. Liu S, Feng G, Xiao L, et al. Shock-induced dynamic response in single and nanocrystalline high-entropy alloy FeNiCrCoCu[J]. International Journal of Mechanical Sciences, 2023. 239: p. 107859.**DOI:** 10.1016/j.ijmecsci.2022.107859.

[13]. Xu W, Geng Y, Zheng H, et al. Microstructural evolution of FeCoNiCrMn high-entropy alloy subjected to laser shock peening: Molecular dynamics simulation study[J]. Next Materials, 2025. 7: p. 100523.**DOI:** 10.1016/j.nxmate.2025.100523.

[14]. Euser V K, Jones D R, Martinez D T, et al. The effect of microstructure on the dynamic shock response of 1045 steel[J]. Acta Materialia. 2023, 250: 118874.**DOI:** 10.1016/j.actamat.2023.118874.

[15]. Yang Y, Yang S, Wang H. Effects of microstructure on the evolution of dynamic damage of Fe50Mn30Co10Cr10 high entropy alloy[J]. Materials Science and Engineering: A, 2021. 802: p. 140440.**DOI:** 10.1016/j.msea.2020.140440.

[16]. Gao T J, Zhao D, Zhang T W, et al. Strain-rate-sensitive mechanical response, twinning, and texture features of NiCoCrFe high-entropy alloy: Experiments, multi-level crystal plasticity and artificial neural networks modeling[J]. Journal of Alloys and Compounds, 2020. 845: p. 155911.**DOI:** 10.1016/j.jallcom.2020.155911.

[17]. Zhou Z, Huo Y, Wang Z, et al. Experimental investigation and crystal plasticity modelling of dynamic recrystallisation in dual-phase high entropy alloy during hot deformation[J]. Materials Science and Engineering: A, 2025. 922: p. 147634.**DOI:** 10.1016/j.msea.2024.147634.

[18]. Connolly D S, Kohar C P, Inal K. A novel crystal plasticity model incorporating transformation induced plasticity for a wide range of strain rates and temperatures[J]. International Journal of Plasticity, 2022. 152: p.

103188.**DOI:** 10.1016/j.ijplas.2021.103188.

[19].Dai W, Wang H, Guan Q, et al. Studying the micromechanical behaviors of a polycrystalline metal by artificial neural networks[J]. Acta Materialia, 2021. 214: p. 117006.**DOI:** 10.1016/j.actamat.2021.117006.

[20]. Liu C, Zhang X, Liu X, et al. Mechanical field guiding structure design strategy for meta-fiber reinforced hydrogel composites by deep learning[J]. Advanced Science, 2024. 11(22): p. 2310141.**DOI:** 10.1002/advs.202310141.

[21]. Wang K, Sun W. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning[J]. Computer Methods in Applied Mechanics and Engineering, 2018. 334: p. 337-380.**DOI:** 10.1016/j.cma.2018.01.036.

[22]. Yuan M, Paradiso S, Meredig B, et al. Machine learning–based reduce order crystal plasticity modeling for ICME applications[J]. Integrating Materials and Manufacturing Innovation, 2018. 7(4): p. 214-230.**DOI:** 10.1007/s40192-018-0123-x.

[23]. Mellbin Y, Hallberg H, Ristinmaa M. Accelerating crystal plasticity simulations using GPU multiprocessors[J]. International journal for numerical methods in engineering, 2014. 100(2): p. 111-135.**DOI:** 10.1002/nme.4724.

[24]. De Oca Zapiain D M, Lim H, Park T, et al. Predicting plastic anisotropy using crystal plasticity and Bayesian neural network surrogate models[J]. Materials Science and Engineering: A, 2022. 833: p. 142472.**DOI:** 10.1016/j.msea.2021.142472.

[25]. Ibragimova O, Brahme A, Muhammad W, et al. A new ANN based crystal plasticity model for FCC materials and its application to non-monotonic strain paths[J]. International Journal of Plasticity, 2021. 144: p. 103059.**DOI:** 10.1016/j.ijplas.2021.103059.

[26].Tran A, Wildey T. Solving stochastic inverse problems for property–structure linkages using data-consistent inversion and machine learning[J]. JOM, 2021. 73(1): p. 72-89.**DOI:** 10.1007/s11837-020-04432-w.

[27]. Liu J, Huang M, Li Z, et al. A deep learning method for predicting microvoid growth in heterogeneous polycrystals[J]. Engineering Fracture Mechanics, 2022. 264: p. 108332.**DOI:** 10.1016/j.engfracmech.2022.108332.

[28]. Han S, Wang C, Lai Q, et al. Fitting-free mechanical response prediction in dual-phase steels by crystal plasticity theory guided deep learning[J]. Acta Materialia, 2025: p. 120936.**DOI:** 10.1016/j.actamat.2025.120936.

[29].Ibragimova O, Brahme A, Muhammad W, et al. A convolutional neural network based crystal plasticity finite element framework to predict localised deformation in metals[J]. International journal of plasticity, 2022. 157: p. 103374.**DOI:** 10.1016/j.ijplas.2022.103374.

[30]. Herriott C, Spear A D. Predicting microstructure-dependent mechanical properties in additively manufactured metals with machine-and deep-learning methods[J]. Computational Materials Science, 2020. 175: p. 109599.**DOI:** 10.1016/j.commatsci.2020.109599.

[31]. Liu C, Xiong Q, An W. Shear localization in gradient high-entropy alloy at high strain rates: Crystal plasticity modeling[J]. Extreme Mechanics Letters, 2024. 70: p. 102194.**DOI:** 10.1016/j.eml.2024.102194.

[32]. Gui Y, An D, Han F, et al. Multiple-mechanism and microstructure-based crystal plasticity modeling for cyclic shear deformation of TRIP steel[J]. International Journal of Mechanical Sciences, 2022. 222: p. 107269.**DOI:** 10.1016/j.ijmecsci.2022.107269.

[33].Alley E S, Neu R W. Neu, A hybrid crystal plasticity and phase transformation model for high carbon steel[J]. Computational Mechanics, 2013. 52: p. 237-255.**DOI:** 10.1007/s00466-012-0810-y.

[34]. Haridas R S, Agrawal P, Thapliyal S, et al. Strain rate sensitive microstructural evolution in a TRIP assisted high entropy alloy: Experiments, microstructure and modeling[J]. Mechanics of Materials, 2021. 156: p. 103798.**DOI:** 10.1016/j.mechmat.2021.103798.

[35]. Soltani M, Ferdousi S, Haridas R S, et al. A crystal plasticity finite element—machine learning combined

approach for phase transformation prediction in high entropy alloy[J]. International Journal of Applied Mechanics, 2024. 16(02): p. 2450024.**DOI:** 10.1142/S1758825124500248.