

# 基于 U-Net 模型的材料动态力学响应预测： 织构表征差异的影响

高翔<sup>1,2</sup>, 赵聃<sup>1,2</sup>, 方慧青<sup>1,2</sup>, 王建军<sup>1,2</sup>, 马胜国<sup>1,2</sup>, 王志华<sup>1,2</sup>

(1.太原理工大学航空航天学院, 山西 太原 030024;

2. 材料强度与结构冲击山西省重点实验室, 山西 太原 030024)

**摘要:** 利用神经网络进行合金性能的预测以及合金微观结构的反向设计已经成为工业界了解合金性能、开发新型合金的一种新兴手段。织构是合金变形过程中微结构演化的重要影响因素, 通常以无空间关联的离散晶粒取向欧拉角、考虑代表性体积单元 (RVE) 空间-取向耦合欧拉角、极图/反极图等形式描述, 然而选取何种形式的织构表示方式作为神经网络模型的输入能发挥模型最优性能还需进一步验证。因此, 以改进后的 U-Net 模型为主体架构, 对比了使用离散欧拉角、空间欧拉角、极图 3 种织构表示方法作为神经网络模型输入对模型性能的影响。采用训练完成的 3 种神经网络模型分别对测试集中的样本进行预测, 结果表明, 使用极图作为织构表示方法给出了最优的效果。除此以外, 利用训练好的神经网络模型, 并采用在输出层添加一维卷积的改进方法, 预测了合金的宏观应力-应变曲线, 相比传统只使用全连接的方法, 该改进显著地提高了应力-应变曲线的预测精度。

**关键词:** 神经网络; 织构; 动态应力-应变曲线; U-Net 模型

**中图分类号:** O344.3; O521.2

**文献标志码:** A

合金属于结构复杂、多晶态的材料, 其力学响应受微观结构特征 (如晶粒取向、形态、尺寸和相分布) 的影响。晶粒取向作为其中最重要的微观结构特征之一, 如何准确表征该特征, 并使用该特征对材料的力学性能进行预测评估、辅助新型合金的设计, 仍然是一个核心挑战。

晶体材料的初始织构对材料后续的屈服应力、织构演化、疲劳预测等诸多方面都有重要的影响。使用晶体塑性有限元 (CPFEM) 的方法对该特征进行研究是最常见的方法之一。杨亚玉<sup>[1]</sup>等采用晶体塑性有限元方法建立了 304 不锈钢箔带轧制有限元模型, 对具有不同初始织构取向的材料进行分析, 证明了晶粒取向显著影响模型轧制过程的接触压力和轧制力; 赵宏伟<sup>[2]</sup>采用均匀化建模和 Voronoi 算法建模两种不同的方法对镍基多晶合金 GH4169 进行有限元建模, 对该材料在拉伸条件下的力学性能和织构演化进行了研究; 郑战光<sup>[3]</sup>等采用晶体塑性有限元的方法针对晶粒取向对过载下疲劳裂尖塑性区的影响进行了系统性的研究。

\*收稿日期: 2025-12-16 修回日期: 2026-02-11

基金项目: 国家自然科学基金 (12225207, 12372364, 12172245)

作者简介: 高翔 (2000-), 男, 硕士研究生, 主要从事机器学习在多晶体有限变形中的应用研究。

E-mail: gx1648862114szzr@163.com

通讯作者: 赵聃 (1985-), 男, 博士, 副教授, 主要从事多晶体有限变形算法及多尺度本构理论

研究. E-mail: zhaodan@tyut.edu.cn

然而, CPFEM 通常面临着两个主要的局限性: 本构参数众多且难以标定<sup>[4]</sup>、处理复杂微观结构(晶粒形貌、织构演化)时所需的计算成本极高<sup>[5]</sup>。为了应对这一问题, 近年来深度神经网络(DNN)、卷积神经网络(CNN)等机器学习方法越来越多地被运用<sup>[6-7]</sup>, 如 U-Net 模型作为经典的卷积神经网络构架, 主要解决图像分割中的小样本数据处理以及高精度定位等问题, 适合应用在材料表征、缺陷检测等需要高精度密集预测的工程领域。在 DNN 的方法当中, 晶体学织构通常都是以离散欧拉角的形式被应用: Ibragimova<sup>[8]</sup>等将 CPFEM 和人工神经网络(ANN)的方法相结合, 将晶粒的欧拉角、材料的初始硬度、初始硬化模量、加载应变率等作为 ANN 模型的输入, 成功预测了多个面心立方(FCC)材料在不同加载条件下的织构演化; M.J. Rezaei<sup>[7]</sup>等将不同扭转数下晶粒的欧拉角和转数当作 ANN 的输入向量, 用以预测演化以后晶粒的欧拉角。还有使用旋转矩阵、空间欧拉角等形式用以表示晶粒取向的方法: Ibragimova<sup>[9]</sup>等将材料的硬化参数、宏观应变值、以旋转矩阵形式表示的晶粒取向作为卷积神经网络模型(CNN)的输入, 成功预测了材料局部应力、应变值的分布。无论是以离散欧拉角形式表示的晶粒取向还是以旋转矩阵-空间欧拉角形式表示的晶粒取向, 其中都存在着一定的局限性: 基于离散欧拉角的方法往往忽略了局部晶粒取向变化对机械响应的影响; 尽管空间欧拉角输入包含了局部邻域效应, 但有限元模型中缺乏标准化的单元划分往往导致体系数据维度不一致, 并且随着有限元模型单元划分密集程度的增加, 有限元模型中会产生巨大的冗余信息, 使用空间欧拉角会极大地增加神经网络模型拟合的复杂性; 除此以外, 工程实际中晶体的欧拉角通过都会伴随着较多的异常角度, 直接将其作为训练数据会导致输入神经网络的数据分布中存在较多的异常值。这些问题给训练好的神经网络模型的直接部署带来了重大挑战, 而使用极图的表达方式则可以有效的避免上述的各种问题。另外, 本研究注意到在应力应变曲线的预测中, 大多数的神经网络模型都直接采用了在输出层连接多层感知机(MLP)的方法<sup>[10-12]</sup>, 作者认为由于材料的变形是连续发生的, 因此应力应变曲线上不同应变点的应力值受其相邻应变点应力数值的影响要大于受其距离较远的应变点应力数值的影响, 因此在输出应力应变曲线的时候应当重点考虑曲线上局部数据点的影响, 而不是全局数据点的影响。

综上所述, 本研究的主要关注点如下:

1. 使用极图作为织构的表示方法, 并直接将其作为神经网络模型的训练数据进行训练。这种表示方法可以既可以有效地考虑晶体变形中的局部邻域变形效应, 还可以避免有限元模型网格划分不一致带来的数据对齐问题以及网格划分过密带来的信息冗余问题, 还可以保证输入神经网络模型数据分布的均匀性, 即不需要重新设计映射函数以保证输入神经网络的欧拉角数据作为训练数据集的合理性。
2. 对于应力应变曲线的预测, 提出的模型在输出层添加了一维卷积层进行优化, 以突出相邻应变点的影响。

## 1. 理论基础

### 1.1 晶体塑性本构模型

率相关的晶体塑性理论依据以滑移系分切应力为控制变量的框架<sup>[13]</sup>。滑移系(面心立方结构滑移系 $\alpha$ 的数目为12)上的剪切应变率 $\dot{\gamma}^\alpha$ 可用幂律形式表示<sup>[14]</sup>:

$$\dot{\gamma}^\alpha = \dot{\gamma}_0 \left| \frac{\tau^\alpha}{g^\alpha} \right|^m \text{sgn}(\tau^\alpha) \quad (1)$$

其中,  $\dot{\gamma}_0$  表示模型的参考滑移应变率;  $\tau^\alpha$  是滑移系上的剪应力;  $m$  是应变率敏

感系数； $g^\alpha$  是滑移阻力。

析出强化的微观物理本质在于弥散分布的析出相对位错运动的有效阻碍。根据析出相尺寸、共格关系及剪切模量等特征差异，位错与析出相间的相互作用主要有切过与绕过机制。当析出相尺寸较小且与基体保持共格/半共格关系时，位错倾向于切割析出相并产生额外的相界面能，在析出相内部形成反相畴界或层错，从而显著增加位错运动的阻力。当位错遇到尺寸较大或高模量析出相时，转而发生 Orowan 绕过机制，会在析出相两侧发生弯曲并绕过后形成位错环。考虑上述位错运动机制，当前滑移系的强度  $\alpha$  可以分解为初始强度  $g_0^\alpha$  和后继强度  $g_{\text{other}}^\alpha$ ：

$$g^\alpha = g_0^\alpha + g_{\text{other}}^\alpha \quad (2)$$

当前强度  $g_0^\alpha$  的初始值分为三个部分<sup>[15]</sup>：

$$g_0^\alpha = g_{0,ss}^\alpha + g_{0, \text{shear}}^\alpha + g_{0, \text{loop}}^\alpha \quad (3)$$

其中  $g_{0,ss}^\alpha$  代表固溶强化引起的阻力； $g_{0, \text{shear}}^\alpha$  和  $g_{0, \text{loop}}^\alpha$  分别表示材料中由于位错运动时剪切析出相和位错绕过析出相对强度的影响。

固溶强化相基于溶质在基体中分布且彼此之间不发生相互作用，其形式如下<sup>[16]</sup>：

$$g_{0,ss}^\alpha = \frac{1}{M} \sum \frac{d\sigma}{d\sqrt{g_i}} \sqrt{g_i} \quad (4)$$

其中， $M$  为泰勒因子， $d\sigma/d\sqrt{g_i}$  表示合金元素的强化系数， $g_i$  为原子分数。可剪切析出物的抗剪切强度如下<sup>[17]</sup>：

$$g_{0, \text{shear}}^\alpha = \alpha \left( \mu_2 \left( \frac{APB_2}{\mu_2 b_2} \right)^{\frac{3}{2}} \left( \frac{f_2 r_2}{b_2} \right)^{\frac{1}{2}} + \mu_3 \left( \frac{APB_3}{\mu_3 b_3} \right)^{\frac{3}{2}} \left( \frac{f_3 r_3}{b_3} \right)^{\frac{1}{2}} \right) \quad (5)$$

其中，脚标 2,3 代表了二级和三级析出相， $\mu$  表示析出相的剪切模量， $b$  是博格斯矢量， $APB$  是析出相的反向畴界面能， $f$  是析出相的体积分数， $r$  是析出物的平均尺寸。

Orowan 绕过机制所需的强度由下式描述<sup>[17]</sup>：

$$g_{0, \text{loop}}^\alpha = \beta \left( \mu_2 b_2 \frac{f_2}{r_2 (1-f_2)} + \mu_3 b_3 \frac{f_3}{r_3 (1-f_3)} \right) \quad (6)$$

$\alpha$  和  $\beta$  是需要进行拟合的参数。关于此模型的较为详细介绍，可参考课题组之前的工作<sup>[18]</sup>。

## 1.2 神经网络算法

### 1.2.1 MLP 算法

MLP 在预测单晶应力-应变曲线和织构演化方面已经得到了验证<sup>[19]</sup>，在预测 FCC 材料非单调加载路径方面也取得了有效的成果<sup>[20]</sup>。MLP 由多层神经网络层相互连接而成，每一

层都由若干人工神经元构成，每一个神经元存储一个数据，并执行一次计算，其相邻两层之间的传播过程可以用以下的公式来表示：

$$\mathbf{x}_l = f(\mathbf{w}_l \mathbf{x}_{l-1} + \mathbf{b}_l) \quad (7)$$

式中， $\mathbf{x}_l$  表示神经网络  $l$  层的输出向量， $\mathbf{w}_l$  表示  $l$  层的权重矩阵， $\mathbf{x}_{l-1}$  表示  $l-1$  层的输出向量，即  $l$  层的输入向量， $\mathbf{b}_l$  表示  $l$  层的偏置向量，函数  $f$  代表了该层神经元执行完线性计算之后执行的非线性操作，通常被称为是激活函数。文章中使用的激活函数为 Swish 激活函数，将在下一节中进行介绍。

### 1.2.2 CNN 算法

CNN 通常应用在计算机视觉领域，其十分擅长处理二维数据和三维数据。近年来其在晶体塑性领域也取得了巨大的成果，体现出了其强大的信息提取和结果预测能力<sup>[21]</sup>。与 MLP 相比，CNN 在每一层中使用了共享参数的卷积核来进行卷积计算，而非神经元的全连接，所以其涉及的参数更少，其卷积计算过程可以用以下的公式来表示：

$$x_{ij}^{lm} = f\left(\sum_{k=1}^C \sum_{r,s=1}^n x_{r+i-1,s+j-1}^{l-1,k} w_{rs}^{lkm} + b_m^l\right) \quad (8)$$

式中， $x_{ij}^{lm}$  表示神经网络  $l$  层输出向量第  $m$  个特征图中的位置为  $(i, j)$  的数据值， $x_{r+i-1,s+j-1}^{l-1,k}$  表示神经网络  $l-1$  层输出向量，即  $l-1$  层输入向量的第  $k$  个特征图中位置为  $(r+i-1, s+j-1)$  的数据值， $w_{rs}^{lkm}$  表示神经网络  $l$  层第  $m$  个卷积核的第  $k$  个子层中位置为  $(r, s)$  的数据值， $b_m^l$  表示神经网络  $l$  层第  $m$  个卷积核的偏置值，与 MLP 一致， $f$  依然表示激活函数。

### 1.2.3 激活函数

激活函数被用来增加神经网络的非线性。如上述公式所示，如果没有激活函数，神经网络只是一个线性函数，其表达能力十分有限，因此需要加入非线性的激活函数以增加神经网络对任何函数的逼近能力。本研究使用 Swish 激活函数，在深层神经网络中，其已被证明拥有强大的效果<sup>[22]</sup>，数学公式表述如下：

$$x = \frac{x}{1 + e^{-x}} \quad (9)$$

### 1.2.4 损失函数和优化器

损失函数是衡量神经网络每个样本预测值与真实值之间的相似程度的一个度量函数。在训练阶段，将每个训练批次所有样本的损失值求平均值，便得到其代价函数值。本研究的代价函数即为损失函数 Huber。Huber 计算公式如下所示：

$$L_{Huber} = \begin{cases} \frac{1}{2}(y_{true} - y_{pre})^2, & |y_{true} - y_{pre}| \leq \delta \\ \delta |y_{true} - y_{pre}| - \frac{1}{2}\delta^2, & |y_{true} - y_{pre}| > \delta \end{cases} \quad (10)$$

式中， $y_{true}$  和  $y_{pre}$  分别代表样本真实值和预测值， $\delta$  为 Huber 损失函数的超参数。

确定了代价函数之后，便可以使用优化器对神经网络的可训练参数进行优化调整，从而最小化样本真实值和预测值之间的差异，即代价函数取得最小值。本研究的优化器选择使用 Adamw 随机梯度下降优化算法，采用自适应机制自动调整梯度下降的学习率：

$$\alpha_{t+1} = \alpha_t - \frac{A}{\sqrt{\hat{v}_t + B}} \hat{m}_t - \chi \alpha_t \quad (11)$$

式中, A 为学习率, B 为一个常数, 防止分母为 0 的情况,  $\chi$  为权重衰减系数,  $\hat{m}_t$  和  $\hat{v}_t$  分别为一阶和二阶动量矩,  $\alpha$  为待更新的神经网络参数。

## 2. 数据集生成

### 2.1 有限元模型

使用 ABAQUS 建立了一个  $30 \times 30 \times 30$  的 3D RVE 模型。该模型包含 216 个晶粒, 平均晶粒尺寸为  $10 \mu\text{m}$ 。模型的初始取向通过 JTEX 生成, 并将其赋予到整个模型中。模型中包含两种属性不同的 FCC 结构的晶粒, 分别对应  $\text{CoCrNiSi}_{0.3}\text{C}_{0.048}$  的 CoCrNi 基体相和一级  $\text{Cr}_{23}\text{C}_6$  析出相。模型的详细结构如图 1 所示。

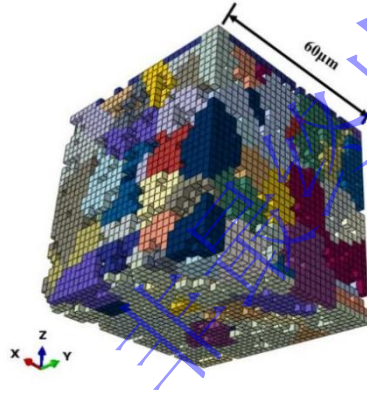


图 1 3D RVE 模型微观结构

Figure 1 Microstructure of 3D RVE model

图 1 给出了 3D RVE 模型的微观结构, 空缺的单元示意一级析出相的位置, 其分布在晶内和晶界上, 并通过析出相的单元数量变化实现析出相体积分数的改变。通过改变 Z 方向的加载位移和时间实现不同应变率的动态加载。模型共 27000 个独立单元, 每个独立单元的尺寸为  $2 \mu\text{m}$ , 单元类型为三维实体减缩积分单元(C3D8R)。CPFEM 模型的本构参数中各相的弹性常数, 剪切模量等取自参考文献<sup>[18]</sup>, 其余关键参数如滑移系初始剪切应力, 临界分切应力等采用试错-校正的方式和实验的应力-应变数据对比获得。具体的数值见表 1。

表 1 CPFEM 本构参数

Table 1 Constitutive parameters of CPFEM model

Variable	Value	Variable	Value
$C_{11}^{matrix}$	245 GPa	$h_0^{matrix}$	51 MPa
$C_{44}^{matrix}$	142 GPa	$\tau_0^{matrix}$	92.7 MPa
$C_{12}^{matrix}$	156 GPa	$\tau_s^{matrix}$	215 MPa
	0.001 /s		10

$\dot{\gamma}_0$		$m$	
$C_{11}^1$	477 GPa	$h_0^1$	51 MPa
	137 GPa		183 MPa
$C_{44}^1$	198 GPa	$\tau_0^1$	1150 MPa
$C_{12}^1$		$\tau_s^1$	
$APB_2, APB_3$	0.5 J/m <sup>2</sup>	$b_2, b_3$	0.43 nm
$\mu_2, \mu_3$	299 GPa	$\alpha$	0.0008
$\beta$	0.143	$r_3$	50 nm
$r_2$	500 nm	$f_2$	0.1
$f_3$	0.1		

注：上标和脚标的数字 1,2,3 分别为 1,2,3 级析出相。

神经网络的数据集是由检验后的 CPFEM 模型模拟样本在不同应变率(1000 s<sup>-1</sup>, 6000 s<sup>-1</sup>), 增量为 500; 不同的析出相体积分数(0, 35), 增量为 5%; 不同初始织构(全部丝织构取向、一半丝织构取向一半随机织构取向、全部随机织构取向); 不同加载方式(压缩、拉伸、压缩)下的变形演化生成的。对于使用极图为训练数据的神经网络模型来说, 从每个有限元模型样本中选取 4 张变形演化之后的极图(分别对应有有限元模型的平均宏观应变 0.1, 0.2, 0.3, 0.35)作为神经网络模型训练的标签, 图 2 展示了数据集中的三种初始极图, 极图通过 MTEX 5.8.1 绘制的, 极图取向均为(1,0,1)方向; 对于使用欧拉角(离散/空间)作为训练数据的神经网络模型来说, 依然选取对应的平均宏观应变下的欧拉角作为神经网络模型的训练标签。

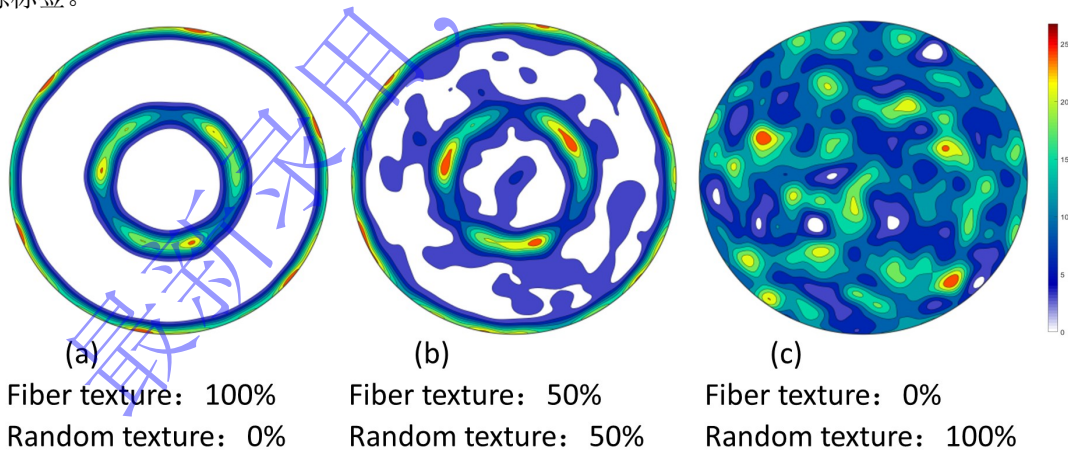


图 2 三种初始极图

Figure 2 Three Initial Polar Graphs

## 2.2 数据集的预处理

通过上述方法绘制的极图为 RGB 三通道形式, 其尺寸为(1200, 1200, 3), 训练该规模的神经网络需要昂贵的计算资源, 因此有必要对其进行降维操作, 极图的降维不是本研究的重点内容, 因此本研究没有设计一个专用的降维模型, 而是通过 stability AI 团队推出的

开源预训练 VAE 模型进行的, 该 VAE 模型经 stability 团队验证拥有强大的性能<sup>[23]</sup>, 适合各种图片的降维压缩, 同时保持很小的信息损失。降维后的极图潜空间特征尺寸为(128, 128, 4)。三种初始极图潜空间特征图的第一个通道可视化如图 3 表示。降维之后的样本数据值落在[-30,30]之间, 在边界值稍微有所浮动, 因此使用最大值归一化的方法将样本数据缩放到[0,1]之间, 归一化公式如下所示:

$$x^* = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (12)$$

公式中的  $x_{\min}$  表示全部样本的最小值,  $x_{\max}$  表示全部样本中的最大值。

对于使用欧拉角的数据类型, 本文采取的归一化方式如下所示:

$$x^* = x / 360.0 \quad (13)$$

此外, 引导条件中的析出相占比、加载应变率也分别乘以系数[0.02, 0.0001], 使其归一化到[0,1]之间。

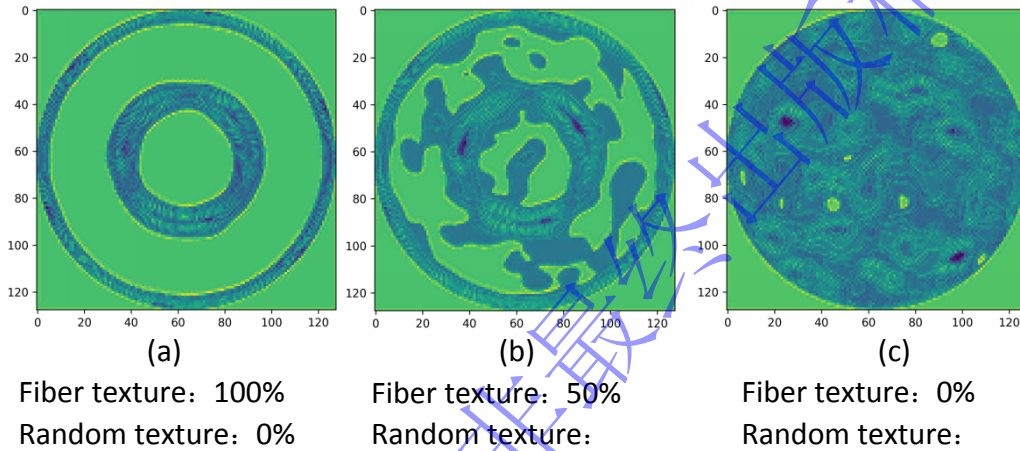


图 3 初始极图潜空间特征图的第一个通道可视化表示

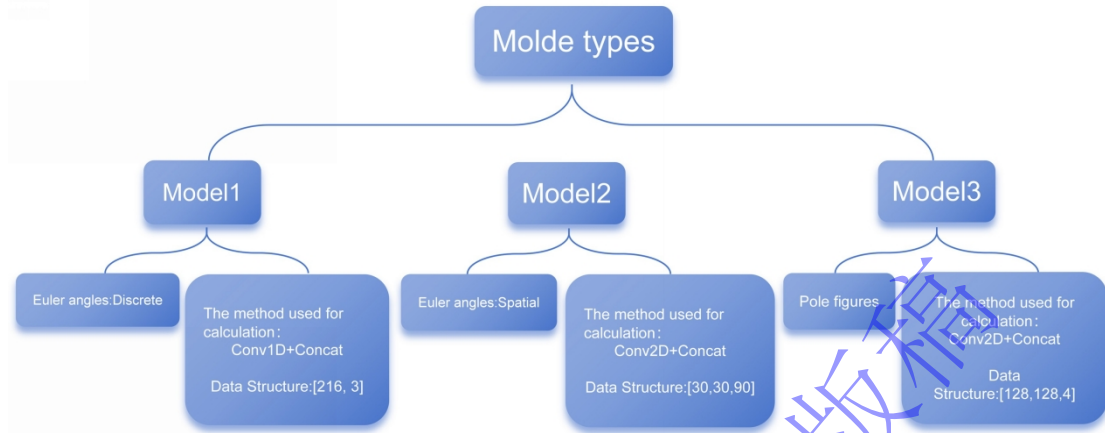
Figure 3 Visualization of the first channel of the latent space feature map of the initial extreme map

### 3. 结果和讨论

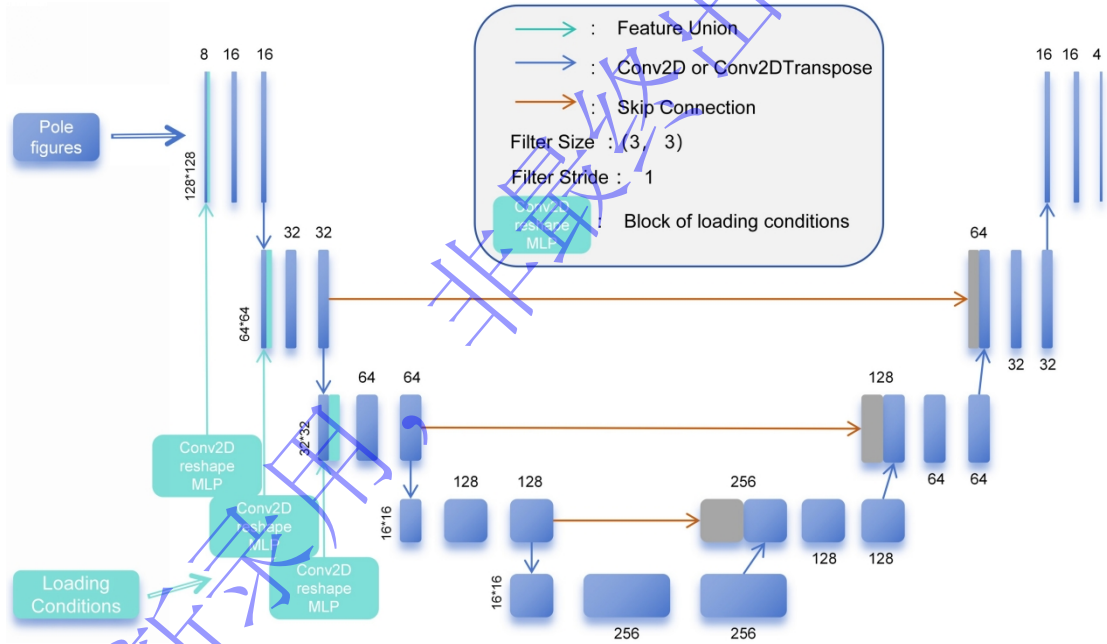
#### 3.1 神经网络模型的建立

根据织构在神经网络模型输入中不同的表达形式, 本文建立了三种神经网络模型, 如图 4 (a) 所示。Model1 直接使用离散取向欧拉角做为神经网络模型的输入, 其最终输入的数据格式为(216, 3), 其中包括了 216 个晶粒的全部欧拉角; Model2 将空间-取向耦合的欧拉角作为输入数据, 其最终输入格式为(30,30,90), 其中包括了 27000 个单元的全部欧拉角数据; Model3 则使用第 3 节中介绍的方法, 使用潜空间极图做为神经网络模型的输入, 其最终的输入格式为(128, 128, 4)。三种模型在整体架构上均采用 U-Net 的架构, 但由于输入数据格式上的差异, 三种模型在数据融合、神经网络算法方面略有不同。图 4 (b) 以使用极图为例说明了本研究使用的神经网络模型主体架构。三种模型均采用将加载条件特征数据和特征图在最后一个维度相连的信息融合方式, 不过在具体算法方面略有不同: Model1 使用一维卷积的算法进行离散欧拉角数据的处理、Model2 为了与 Model3 使用基本相同的模型参数, 首先将空间欧拉角数据进行反卷积、填充处理, 以匹配其数据尺寸 (具体操作为: 先使用两层 Transpose2D 层, 卷积核大小为(3, 3), 卷积步幅为 2; 再进行指定

大小的填充操作，使其最终大小为(128, 128))，之后使用二维卷积的算法进行特征数据的处理、Model3 也使用二维卷积的算法进行潜空间极图数据的处理。除此以外，加载条件和极图未采取在输入层直接相连的融合方式，而是采用了图 4 (b) 所示的早期-多通道融合  
的改进策略。



(a) Illustration of three different models



(b) Main architecture diagram of the neural network model (illustrated by using the polar graph as an example)

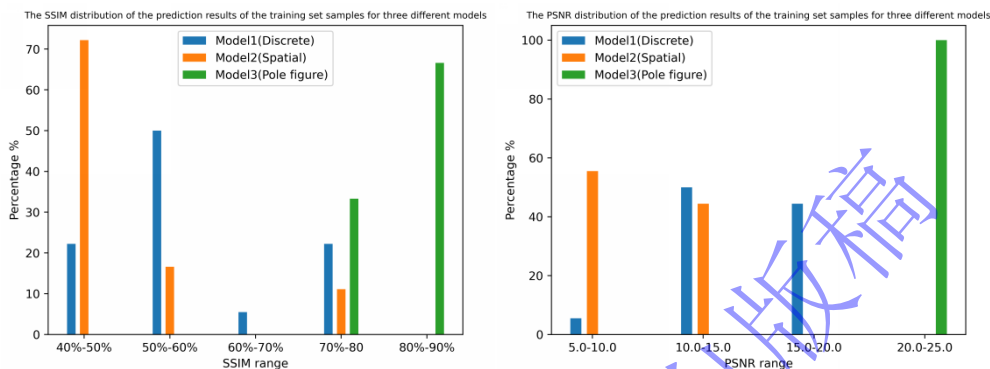
图 4 (a) 三种不同模型的说明，(b) U-Net 模型主体架构图 (以使用极图为例进行说明)

Figure 4 (a) Illustration of three different models; (b) Main architecture diagram of the neural network model (illustrated by using the polar graph as an example)

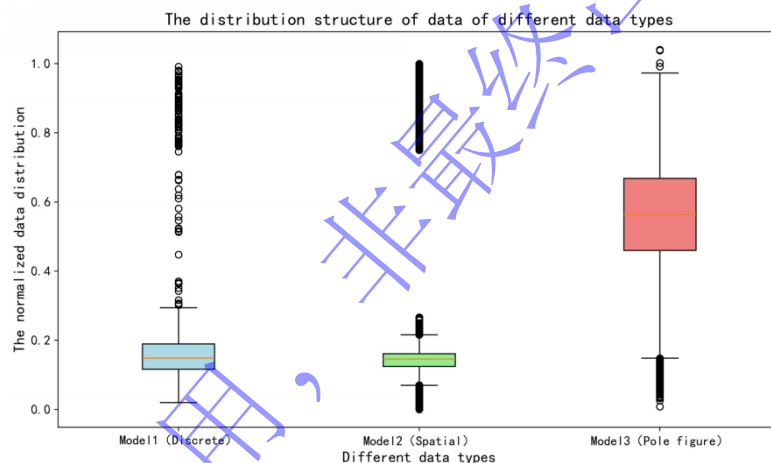
### 3.2 神经网络模型训练过程

本次训练数据集中共有 3456 个样本，并且按照 8: 1: 1 的比例将其随机划分为训练集、验证集和测试集。训练过程中，每个训练批次由 5 个样本组成，总共训练 30 轮，学习率设置为指数衰减型，初始学习率为 0.01，衰减步数为 3000，衰减率为 0.5。图 5 (a) 展示了完成训练的三种神经网络模型分别对训练集样本的预测结果，使用结构相似性指标 (SSIM)

和峰值信噪比指标（PSNR）作为预测结果与真实值的衡量指标。从图中可以清晰地看出，使用极图作为神经网络模型的训练数据，可以获得三种情况中最好的效果。除了引言中提到的局部邻域效应和信息冗余以外，作者认为使用极图获得更好的效果还与不同数据结构的数据分布有关：图 5（b）以数据集中其中一个样本的数据分布为例说明了三种不同数据结构的数据分布情况，从图中显然可知极图类型的数据分布更加均匀，没有太多的异常值；而离散欧拉角和空间欧拉角的数据分布非常不均匀，拥有较多异常值。作者认为这也是导致使用欧拉角作为训练数据产生较差效果一个重要因素。



(a) The SSIM distribution of the prediction results of the three models for the test set samples



(b) The three different types of data distribution structures

图 5 (a) 三种模型对测试集样本预测结果的 SSIM 和 PSNR 的分布情况，  
(b) 三种不同类型的数据分布结构

Figure 5 (a) The SSIM distribution of the prediction results of the three models for the test set samples; (b) the three different types of data distribution structures

### 3.3 织构演化结果对比

图 6 展示了训练完成之后三种模型的织构演化预测结果，预测精度由结构相似性指标（SSIM）来衡量，对测试集不同样本的织构演化结果做了可视化处理。样本的初始织构分别为丝织构占比 100%、随机织构占比 0%以及丝织构占比 50%，随机织构占比 50%；样本的加载条件分别为一级析出相体积分数：0%，加载应变率：5000/s，宏观应变：-0.1；一级析出相体积分数：5%，加载应变率：6000/s，宏观应变：-0.2。从图中可以看出，直接使用离散欧拉角作为神经网络模型的输入则给出相对较差的结果，其预测结果的结构相似性指标只有 0.74 和 0.65，这与其缺乏局部邻域信息是有直接关系的；使用空间欧拉角作为神经网络模型的输入给出了最差的预测结果，其预测结果的结构相似性指标分别为 0.71 和

0.56, 从图中也可以直接看出, 其预测结果与真实值之间有着较大的偏差, 产生该结果的原因主要是由于有限元模型单元划分带来的冗余信息增加的神经网络模型拟合的复杂性; 而使用极图作为神经网络模型的输入则给出了最好的预测结果, 在两个加载条件之下的预测结果的结构相似性指标分别达到了 0.89 和 0.85。以上预测结果的对比再次证明了极图作为织构表达形式, 用作神经网络模型训练数据集的优越性。

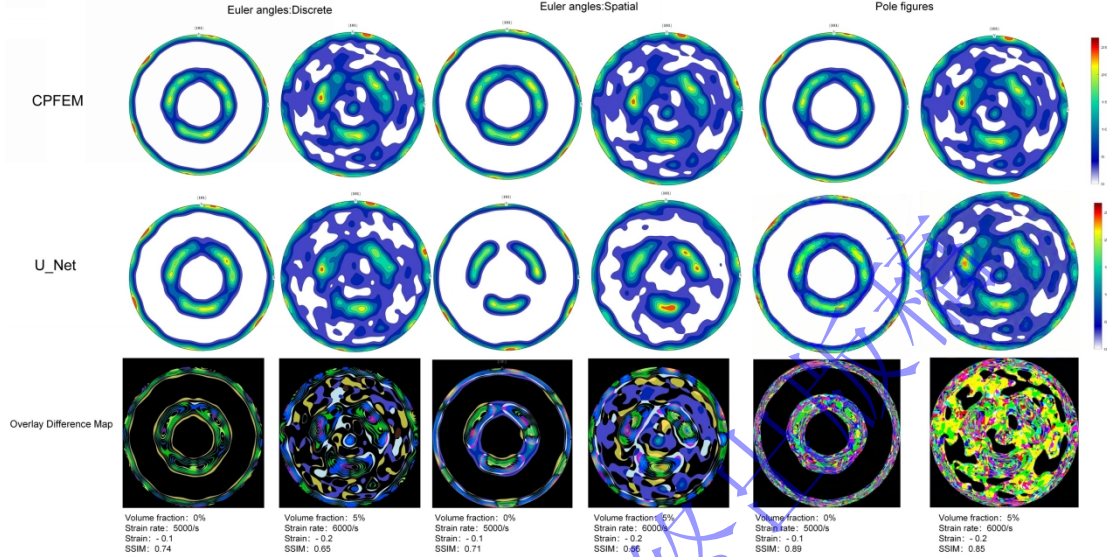
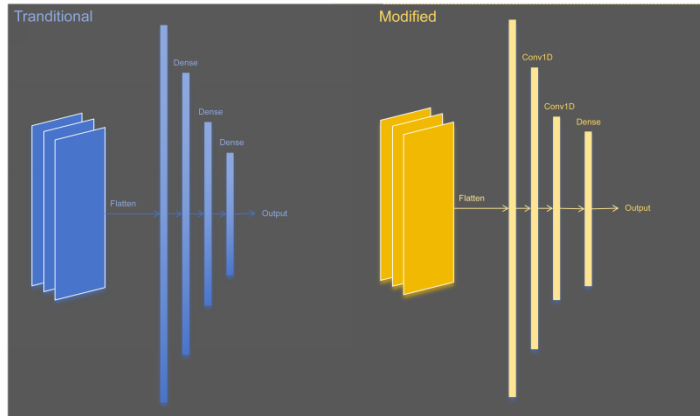


图 6 分别采用不同织构表达形式作为神经网络模型输入的预测结果

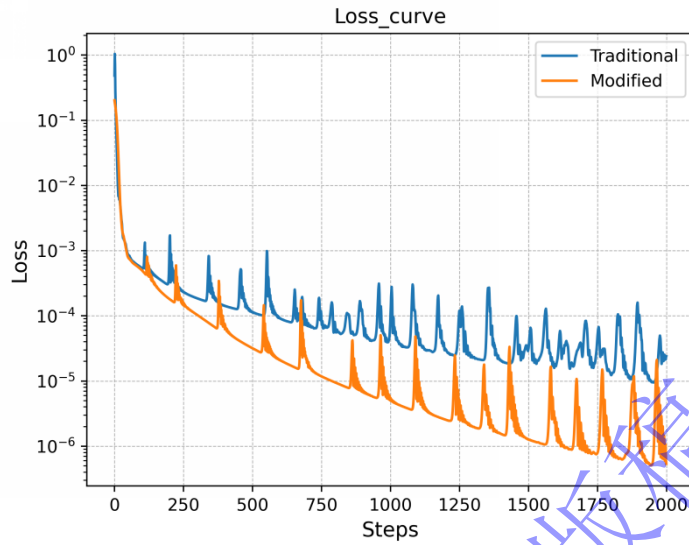
Figure 6 Prediction results obtained by using different texture expression forms as the input of the neural network model

### 3.4 应力-应变曲线结果对比

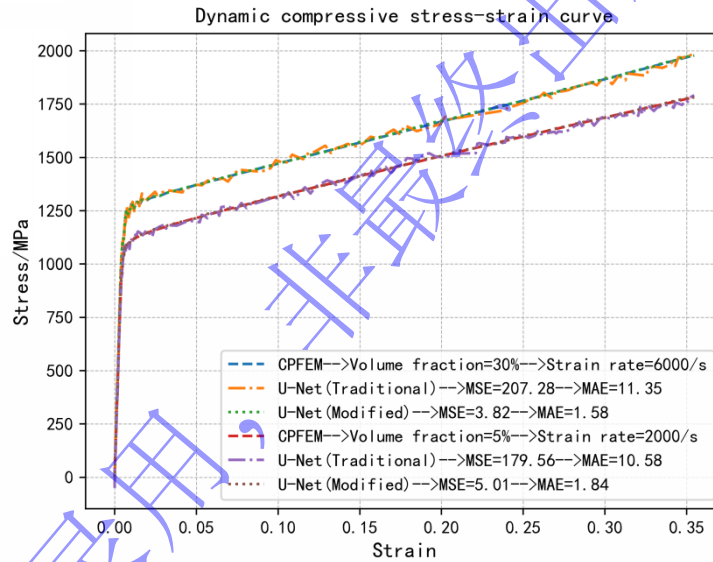
图 7 (a) 可视化了本研究对输出层添加的一维卷积层。改进的方法主要在于将两个全连接层用两个一维卷积层代替, 以重点考虑曲线上局部相邻应变点之间的相互影响。图 7 (b) 中的两条曲线分别代表传统方法和改进方法在训练过程中的损失函数曲线。在相同的训练方法和训练时间下, 传统方法损失值下降的速度以及最后的精度都远远不及改进之后的方法, 两条曲线的对比再次有力地证明了添加一维卷积层之后的神经网络架构在提升应力应变曲线拟合精度方面的合理性。图 7 (c) 展示了使用传统方法和改进方法预测的两组应力应变曲线, 两条曲线的预测误差由均方误差 (MSE) 和平均绝对误差 (MAE) 来衡量, 可以看到, 传统方法的 MSE 分别为 207.28 和 179.56, MAE 分别为 11.35 和 10.58。但是改进方法显然获得了更高的预测精度, MSE 分别为 3.82 和 5.01, MAE 分别为 1.58 和 1.84。再次证明了重点考虑局部应变点应力数值对应力应变曲线预测精度提升的重要意义。



(a) The improved method of the output layer



(b) The loss function curve



(c) The dynamic compressive stress-strain curve

图 7 (a) 输出层的改进方法 (b) 损失函数曲线 (c) 动态压缩应力-应变曲线  
Figure 7 (a) The improved method of the output layer (b) The loss function curve (c) The dynamic compressive stress-strain curve

## 4. 结论

(1) 通过对不同织构表示方法的系统性研究, 得到极图作为织构表示方法对 U-Net 架构的网络模型训练性能提升有一定增强的原因是: 其有效地考虑晶体变形中的局部邻域变形效应, 避免了有限元模型网格划分不一致带来的数据对齐问题以及网格划分过密带来的信息冗余, 还避免了直接使用欧拉角作为 U-Net 输入数据可能存在的数据分布不合理问题。其为合金性能预测和合金微观结构反向设计提供了全新思路。

(2) 计算结果显示, 在 U-Net 模型输出层添加一维卷积层, 即重点考虑应变点局部邻域信息, 能显著提升合金应力-应变曲线预测的精度。

(3) 后续的研究在织构演化层面: 将探索包含更多沉淀物的复杂多相材料、建立欧拉角合理映射函数以适应神经网络模型、欧拉角与极图的混合使用策略; 在应力应变曲线预测层面: 将深入研究输出层中各个神经网络层的合理设置、输出层和隐藏层数据的混合使用策略。

上述结论, 为后续研究人员提供了一个普适性的架构设计策略, 同时为材料性能预测以及材料逆向设计提供了一个较为独特的视角。

## 参考文献:

- [1] 杨亚玉,范婉婉,刘红艳,等.不同晶粒取向的 304 不锈钢箔带介观尺度轧制变形[J]. 锻压技术, 2025, 50 (05): 188-196.DOI:10.13330/j.issn.1000-3940.2025.05.023.  
YANG Y Y, FAN W W, LIU H Y, et al.Mesoscale rolling deformation of 304 stainless steel foil with different grain orientations [J]. Forging & Stamping Technology, 2025, 50 (05): 188-196. DOI:10.13330/j.issn.1000-3940.2025.05.023.
- [2] 赵宏伟.基于 CPFEM 的镍基合金力学性能及织构演化行为研究[D].太原科技大学, 2024.DOI:10.27721/d.cnki.gyzjc.2024.000819.  
ZHAO H W.Study on the mechanical properties and texture evolution behavior of nickel-based alloys based on CPFEM.[D] Taiyuan University of Science and Technology,2024.DOI:10.27721/d.cnki.gyzjc.2024.000819.
- [3] 郑战光,周滋,孙腾,等.晶粒取向对过载下疲劳裂纹塑性区的影响[J].塑性工程报, 2023,30(11):194-202.  
ZHAN Z G, ZHOU Z, SUN T, et al.Effect of grain orientation on plastic zone of fatigue crack tip under overload [J]. Journal of Plasticity Engineering, 2023,30(11):194-202.
- [4] Graff S, Brocks W, Steglich D. Yielding of magnesium: From single crystal to polycrystalline aggregates[J]. International Journal of Plasticity, 2007, 23(12): 1957-1978.
- [5] Priddy M W, Paulson N H, Kalidindi S R, et al. Strategies for rapid parametric assessment of microstructure-sensitive fatigue for HCP polycrystals[J]. International Journal of Fatigue, 2017, 104:231-242.
- [6] 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): 214-230.
- [7] Rezaei M J, Sedighi M, Poletti M C, et al. Crystallographic texture prediction of torsioned aluminum wire using hybrid of machine learning and multi-scale crystal plasticity[J]. Materials Characterization, 2025, 224: 115000.
- [8] 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: 103059.
- [9] 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: 103374.
- [10] Flávio P G M, Muniz L R C, Doca T. ANN strategies for the stress-strain analysis of metallic materials: Modeling, database, supervised learning, validation and performance analysis[J]. Finite Elements in Analysis and Design, 2024, 230: 104097.
- [11] Liu Q, Wu D. Machine learning and feature representation approaches to predict stress-strain curves of additively manufactured metamaterials with varying structure and process parameters[J]. Materials &

- Design, 2024, 241: 112932.
- [12] Milad R, Hegazy H, Zhang J, et al. Estimating the stress distribution within MERO joint using (FEM-ANN) hybrid technique[J]. *Journal of Computational Science*, 2024, 79: 102294.
- [13] Huang Y. A user-material subroutine incorporating single crystal plasticity in the ABAQUS finite element program[M]. Cambridge, UK: Harvard Univ., 1991.
- [14] Hutchinson J W. Bounds and self-consistent estimates for creep of polycrystalline materials[J]. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 1976, 348(1652): 101-127.
- [15] Lu P, Jin X, Li P, et al. Crystal plasticity constitutive model and thermodynamics informed creep-fatigue life prediction model for Ni-based single crystal superalloy [J]. *International Journal of Fatigue*, 2023, 176: 107829.
- [16] Gypen L A, Deruyttere A. Multi-component solid solution hardening: Part 1 Proposed model[J]. *Journal of Materials Science*, 1977, 12(5): 1028-1033.
- [17] Eghtesad A, Knezevic M. A full-field crystal plasticity model including the effects of precipitates: Application to monotonic, load reversal, and low-cycle fatigue behavior of Inconel 718[J]. *Materials Science and Engineering: A*, 2021, 803: 140478.
- [18] Zhao W T, Wang Q, Zhao D, et al. Numerical and experimental investigation of the dynamic mechanical behavior of precipitation-strengthened NiCoCrSi0.3C0.048 medium-entropy alloy[J]. *Journal of Materials Research and Technology*, 2024, 30: 5826-5841.
- [19] Ali U, Muhammad W, Brahme A, et al. Application of artificial neural networks in micromechanics for polycrystalline metals[J]. *International Journal of Plasticity*, 2019, 120: 205-219.
- [20] 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: 103059.
- [21] Beniwal A, Dadhich R, Alankar A. Deep learning based predictive modeling for structure-property linkages[J]. *Materialia*, 2019, 8: 100435.
- [22] Ramachandran P, Zoph B, Le Q V. Searching for activation functions[J]. *arXiv preprint arXiv:1710.05941*, 2017.
- [23] Meng C, He Y, Song Y, et al. Sdedit: Guided image synthesis and editing with stochastic differential equations[J]. *arXiv preprint arXiv:2108.01073*, 2021.

# Prediction of Dynamic Mechanical Response of Materials Based on U-Net Model: Influence of Texture Representation Differences

GAO Xiang<sup>1,2</sup>, ZHAO Dan<sup>1,2</sup>, FANG Huiqing<sup>1,2</sup>, WANG Jianjun<sup>1,2</sup>, MA Shengguo<sup>1,2</sup>, WANG Zhihua<sup>1,2</sup>

(1. College of Aeronautics and Astronautics, Taiyuan University of Technology, Taiyuan 030024, Shanxi, China;  
2. Shanxi Key Laboratory of Material Strength & Structural Impact, Taiyuan 030024, Shanxi, China)

**Abstract:** The utilization of neural networks for the prediction of alloy properties and the inverse design of alloy microstructures has emerged as a novel approach in the industry for understanding material performance and developing new alloys. Texture acts as a critical factor influencing microstructural evolution during alloy deformation. It is typically characterized by spatially uncorrelated discrete grain orientation Euler angles, spatial-orientation coupled Euler angles within a Representative Volume Element (RVE), or pole figures/inverse pole figures. However, identifying which texture representation method serves as the optimal input to maximize the performance of neural network models requires further investigation. Consequently, employing a modified U-Net model as the backbone architecture, this study evaluates and compares the impact of three texture representation methods - discrete Euler angles, spatial Euler angles, and pole figures-as model inputs on the overall performance of the neural network. The three trained neural network models were individually deployed to predict samples within the test set. The results demonstrate that employing pole figures as the texture representation method yields the optimal performance. Furthermore, the trained neural network models were utilized to predict the macroscopic stress-strain curves of the alloys by incorporating a one-dimensional (1D) convolutional layer at the output stage. Compared to traditional methods relying solely on fully connected layers, this modification significantly enhances the prediction accuracy of the stress-strain curves.

**Key words:** neural network; texture; dynamic stress-strain curve; U-net model