

DOI: 10.13718/j.cnki.xdzk.2025.08.003

李仲玉, 高阳, 袁绍阳, 等. 基于深度学习的 CFRP 蜂窝铝三明治结构冲击能量预测研究 [J]. 西南大学学报(自然科学版), 2025, 47(8): 26-37.

基于深度学习的 CFRP 蜂窝铝三明治结构冲击能量预测研究

李仲玉^{1,2,3}, 高阳¹, 袁绍阳³, 王彦翔¹, 王兵²

1. 中车长春轨道客车股份有限公司 工程研究中心, 长春 130062;

2. 哈尔滨工业大学 特种环境复合材料技术国家级重点实验室, 哈尔滨 150001;

3. 哈尔滨工业大学 汽车工程学院, 山东 威海 264209

摘要: 针对碳纤维复合材料(Carbon Fiber Reinforced Polymer, CFRP)蜂窝铝三明治结构, 提出一种使用有限元仿真作为损伤数据来源, 并基于改进后的深度学习模型对三明治结构冲击能量进行预测的方法。该方法在 ResNet50 主干网络的全局平均池化层之前引入一种多头注意力(Multi-Head Attention)模块, 对卷积层提取的特征进行整合和筛选, 使最终输入到全连接层中的特征向量更具有代表性。在 RegNet 主干网络的初始卷积层之后引入相同的注意力机制模块, 使网络在早期就聚焦于图像中的关键区域和特征, 增强底层特征的表达。研究结果表明: 针对预测准确率, ResNet50 网络最高达到 93.7%, RegNet 网络最高达到 93.1%; 在引入注意力模块之后, ResNet50 网络能够达到 98.9%, RegNet 网络能够达到 97.1%。2 种改进的深度学习网络能够有效地预测出损伤图像中的冲击能量, 为 CFRP 蜂窝铝三明治结构冲击能量的预测研究提供了一种新的参考。

关键词: CFRP 蜂窝铝三明治结构; 低速冲击; 深度学习模型;

多头注意力模块; 冲击能量预测

中图分类号: TP183; TP391.9

文献标识码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2025)08-0026-12

Research on Impact Energy Prediction of CFRP Aluminum Honeycomb Sandwich Structures Based on Deep Learning

LI Zhongyu^{1,2,3}, GAO Yang¹, YUAN Shaoyang³,
WANG Yanxiang¹, WANG Bing²

1. Engineering Research Center, CRRC Changchun Railway Vehicles Co. Ltd., Changchun 130062, China;

2. National Key Laboratory of Science and Technology on Advanced Composites in Special Environments, Harbin Institute of Technology, Harbin 150001, China;

3. School of Automotive Engineering, Harbin Institute of Technology, Weihai Shandong 264209, China

收稿日期: 2025-04-30

基金项目: 山东省高等学校“青创团队计划”项目(2023KJ078); 山东省自然科学基金项目(ZR2023QE33)。

作者简介: 李仲玉, 博士, 副教授, 主要从事轻量化设计研究。

通信作者: 高阳, 博士, 教授级高级工程师。

Abstract: This paper proposes a method using finite element simulations as a source of damage data and an improved deep learning model to predict the impact energy of carbon fiber reinforced polymer (CFRP) honeycomb aluminum sandwich structures. The method introduces a multi-head attention module in the ResNet50 backbone network before the global average pooling layer to integrate and filter the features extracted by the convolutional layers, making the feature vectors finally input into the fully connected layers more representative. The same attention mechanism module is introduced after the initial convolutional layer in the RegNet backbone network, enabling the network to focus on key regions and features in the images at early stage, thereby enhancing the expression of low-level features. Research results show that for prediction accuracy, the ResNet50 network reaches up to 93.7%, and the RegNet network reaches up to 93.1%. After introducing the attention modules, the ResNet50 network can achieve 98.9%, and the RegNet network can achieve 97.1%. The two improved deep learning networks can effectively predict the impact energy of damage images, providing a new reference for the study of impact energy prediction in CFRP aluminum honeycomb sandwich structures.

Key words: CFRP aluminum honeycomb sandwich structure; low-velocity impact; deep learning module; multi-head attention module; impact energy prediction

碳纤维增强复合材料(Carbon Fiber Reinforced Polymer, CFRP)是一种高比强度、高比刚度的轻质材料,目前被广泛应用于航空航天、汽车制造和轨道交通等众多领域。由 CFRP 蒙皮和蜂窝铝夹心组成的三明治结构在使用过程中常会受到外界物体的冲击作用,产生不同程度、多种类型的损伤。这些损伤往往出现在 CFRP 内部,可能会导致 CFRP 力学性能大幅下降。对于 CFRP 三明治结构受到低速冲击后所产生的损伤,研究人员往往通过构建物理模型进行研究^[1-4],然而 CFRP 三明治结构在冲击载荷作用下,仿真计算效率较低,因此基于数据挖掘和深度学习等人工智能技术为 CFRP 三明治结构受到低速冲击后的损伤预测提供了一种新的思路^[5]。

近年来,国内外学者将深度学习技术广泛应用于复合材料研究领域,并取得显著进展。Karsandik 等^[6]总结了影响复合材料夹层结构冲击力学行为的主要参数;Li 等^[7]提出将弹性体引入 CFRP 铺层形成新型三明治结构,研究表明其在低速冲击下具有优异性能;Liao 等^[8]开发出一种结合主成分分析与卷积神经网络(Convolutional Neural Networks, CNN)的特征提取方法,在 ResNet50 架构下显著提升了模型效能;Chen 等^[9]通过改进 RegNet 模型中的激活函数,使分类准确率达到 99.5%;Sumit 等^[10]提出轻量级 CNN 模型,将知识蒸馏与超参数优化相结合,实现模型尺寸显著缩减的同时保持预测性能。在图像处理领域,Oliveira 等^[11]采用 U 型网络对 CFRP 低速冲击损伤进行分割,在 4 种传统方法中表现最优;Liu 等^[12]提出基于深度学习的缺陷剖面重建方法,在细节表征上超过传统热成像后处理;Wei 等^[13]将人工智能和红外热成像相结合,以检测和分割弯曲层压板的冲击损伤;Oliver 等^[14]开发出一种模态数据驱动的人工神经网络复合材料损伤检测系统,准确率达 95%;Deng 等^[15]基于深度学习方法提出了一个新的框架,以研究 CFRP 中出现的几乎不可见的冲击损伤(Barely Visible Impact Damage, BVID)特征与测试冲击能量之间的关系;Cao 等^[16]设计双流 CNN 用于一维热信号分类,实现缺陷区域精确识别。在深度学习模型优化方面,Saeed 等^[17]结合 CNN 和 DFF-NN 算法开发智能后处理器,对缺陷进行检测与深度估计;Dong 等^[18]提出三维 CNN 模型,在单个缺陷检测任务中超越传统方法;Hu 等^[19]创新性地融合时空特征,提出了基于注意力模块的新型热成像缺陷检测模型,显著提升了结果的精度和鲁棒性;Deng 等^[20]构建结合注意力机制的时空深度学习框架,成功建立了 BVID 与能级影响的关系模型;徐杰^[21]提出了一种基于深度学习和损伤图像来预测车用 CFRP 冲击能量的方法;Hasebe 等^[22]结合特征工程,使用机器学习模型确认了从 CFRP 表面损伤剖面推断低速冲击信息的可能性;端木正等^[23]使用改进后的深度学习模型进行材料设计;杜禹樵等^[24]基于人工神经网络深度学习算法,设计了一套 CFRP 损伤监测系统;Wang 等^[25]通过深度学习方法对复合材料的宏观等效特性进行预测,预测结果与有限元仿真误差不超过 0.5%;王敏等^[26]对碳纳米管/碳纤维增强型复合材料(Carbon Nanotube/Carbon Fibre Reinforced

Polymer, CNT/CFRP)在低速冲击下的响应和破坏进行了数值模拟研究,结果表明在相同的冲击能量下,更大的冲击速度会造成更多的拉伸破坏;Huang 等^[27]使用机器学习高效准确地预测了 CNT 增强水泥复合材料的力学性能;Yang 等^[28]提出了一种基于迁移学习的逆方法,通过吸能结构的变形图片提取复合材料的结构参数和耐撞特性;Loutas 等^[29]提出了一种将数值模拟和主动学习相结合的方法,能够快速有效地搜索仿生多级不连续纤维复合材料几何参数的设计空间;钱奇伟等^[30]采用深度学习算法,对大量三维编织碳/碳(Carbon Fiber Reinforced Carbon, C/C)复合材料的微观 X 射线计算机断层扫描(Micro X-ray Computed Tomography, XCT)图像进行训练,结果表明深度学习算法能够有效过滤噪声与伪影并自动精准分割各组分和缺陷;Fonseca 等^[31]结合有限元、人工智能和进化搜索方法,实现了混合动力汽车结构及其塑料注塑成型工艺的增强。

综上所述,目前深度学习在复合材料冲击损伤检测领域中的应用,其数据集大多来源于红外热成像等试验技术,然而相关技术需要大量的试验和制备成本,并且可获得的数据集样本有限。因此,本研究使用有限元仿真得到的 CFRP 蜂窝铝三明治结构低速冲击损伤图像作为数据来源,提出一种融合多头注意力机制的深度学习方法,对三明治结构冲击能量进行预测,采用 ResNet50 和 RegNet 2 种神经网络对数据集进行训练,并分别在 ResNet50 网络的全局平均池化层(Global Average Pooling, GAP)之前和 RegNet 网络的初始卷积层之后添加注意力模块,使 2 种网络的准确率得到提升,为 CFRP 蜂窝铝三明治结构冲击能量预测研究提供一种新的参考,同时也为车辆碰撞性能的研究开拓一种新的思路。

1 CFRP 蜂窝铝三明治结构低速冲击仿真

本研究基于深度学习模型对 CFRP 蜂窝铝三明治结构的冲击能量进行预测。深度学习模型构建与训练过程中需要大量的三明治结构冲击行为和损伤信息数据集,仅依靠冲击试验获得的数据集面临着试验成本高、周期长以及信息表征局限等问题,因此本研究采用有限元仿真模型对 CFRP 蜂窝铝三明治结构进行低速冲击模拟,可以快速、大量获取深度学习模型所需数据集。

1.1 有限元模型

采用 Abaqus/Explicit 进行低速冲击模拟,该模型由具有初速度的刚性冲击器以及 CFRP 蜂窝铝三明治结构模型组成,如图 1 所示。按照实际低速冲击试验夹具约束方式对三明治结构夹持区域所有自由度进行约束,图 1b 中 U_z 表示 z 方向平移自由度, UR_x 、 UR_y 、 UR_z 分别表示绕 x 、 y 、 z 轴的旋转自由度。

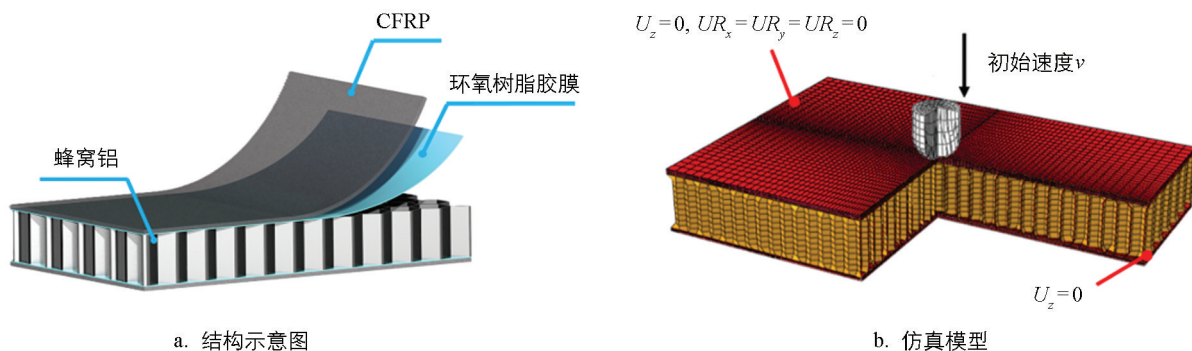


图 1 CFRP 蜂窝铝三明治结构模型

采用孤立网格偏置方法对 CFRP 蜂窝铝三明治结构进行建模^[7],为确保模型精度,提高计算效率,将上、下蒙皮的网格划分为渐进式网格,使得靠近冲击中心的网格更密。按照实际结构进行偏置得到上、下蒙皮模型,其中 CFRP 层内采用 SC8R 单元、层间采用 COH3D8 单元进行建模。蜂窝铝模型采用 S4R 单元进行建模,网格尺寸为 1 mm。蜂窝和上、下蒙皮之间使用 Tie 连接进行绑定,同时,由于蜂窝和蒙皮之间的连接是较为脆弱的线连接,在蒙皮和蜂窝之间设置层间单元以模拟蒙皮和蜂窝的脱黏行为。采用 surface to surface 接触算法定义刚性冲击器和三明治结构中心区域的接触以及蜂窝面板和蜂窝之间的自接触行为,摩擦因数 $\mu=0.3$ 。蜂窝结构的材料选用 5052 系铝合金,并采用 Johnson-Cook(J-C)模型对蜂窝铝的塑性变形和失效进行模拟。

1.2 仿真验证

为验证所建仿真模型的准确性,开展 9 J 和 14 J 能量下的 CFRP 蜂窝铝三明治结构低速冲击试验,并将试验数据与仿真模型进行对比,本研究以冲击后 CFRP 蜂窝铝冲击侧的形变程度作为对标指标。图 2 分别展示了 9 J 和 14 J 冲击后试验损伤形貌和仿真模型形变云图的对比,可以发现仿真模型形变云图与对应低速冲击试验中 CFRP 蜂窝铝三明治结构的损伤形貌展现出较高的一致性。

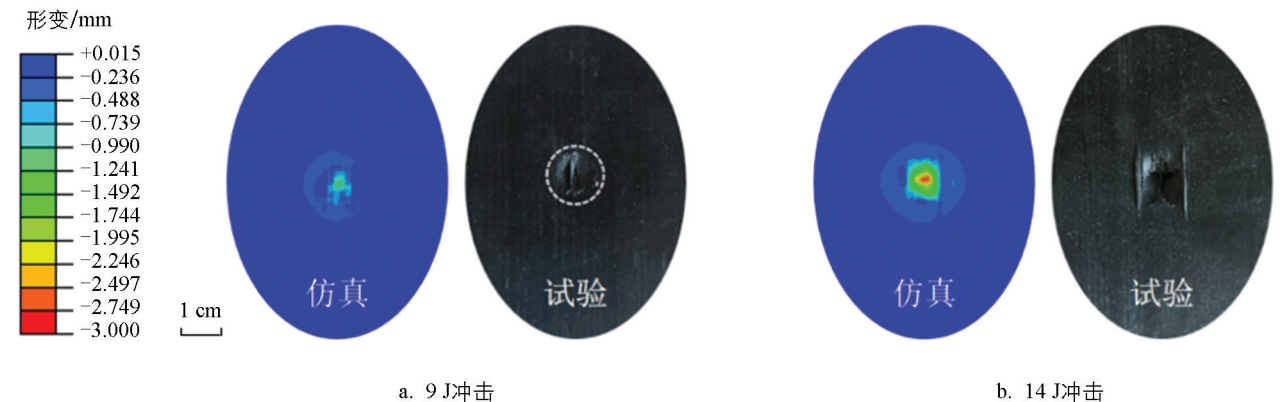


图 2 样件冲击试验和仿真对比

如图 2a 所示,在 9 J 冲击下,上蒙皮冲击侧出现了较明显的凹陷,主要损伤形式为基体开裂,冲击凹坑与上蒙皮平面连接较为平滑,这一点在对应的 9 J 冲击仿真模型形变云图中的表现相同。如图 2b 所示,在 14 J 冲击下,试验样件出现了剧烈的纤维断裂损伤,冲击中心形变出现大幅增加,冲击坑与面板之间的界限明显;在 14 J 冲击仿真模型形变云图中,冲击中心出现了由纤维断裂造成的较大变形,并且冲击中心与蒙皮平面存在明显界限。因此,仿真模型能够较好地模拟 CFRP 蜂窝铝三明治结构在 2 种冲击能量下的冲击行为和损伤形式,具有较高的准确性和可信度。此外,CFRP 蜂窝铝三明治结构受落锤冲击后的载荷-时间曲线和载荷-位移曲线如图 3 所示,说明该仿真模型在三明治结构的冲击行为和损伤模拟方面满足研究要求。

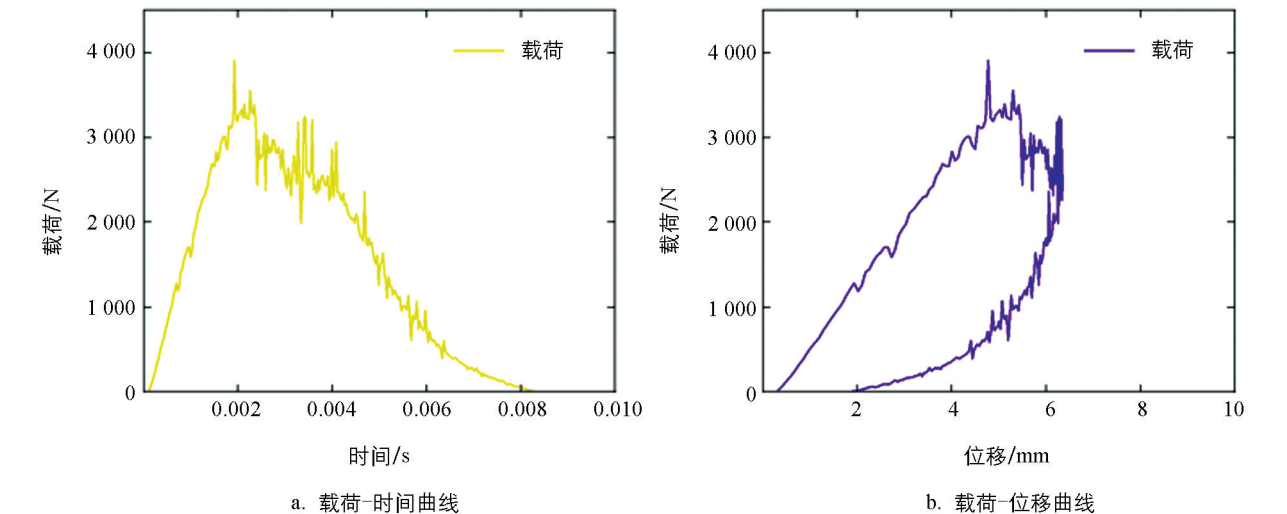


图 3 落锤冲击三明治结构的力学响应

2 基于深度学习的冲击能量预测

2.1 数据集构建

本研究建立了 5 组低速冲击仿真模型,冲击能量分别为 4 J、7 J、9 J、12 J 和 14 J。在每组冲击能量下,通过修改冲击锤头质量与速度,进行多个不同组合的仿真,以提高可获得数据的体量。由仿真结果可知,蜂窝铝产生的塑性变形在冲击过程中起到了一定的缓冲作用。CFRP 蜂窝铝夹层合板的下蒙皮

只受到了很小的损伤, 损伤程度并不明显。图 4 分别展示了上、下蒙皮的基体拉伸损伤情况。在创建数据集时, 仅获取上蒙皮各层的基体拉伸损伤图像, 然后通过 Abaqus 运行 Python 脚本快速得到仿真结果中的原始损伤图像。

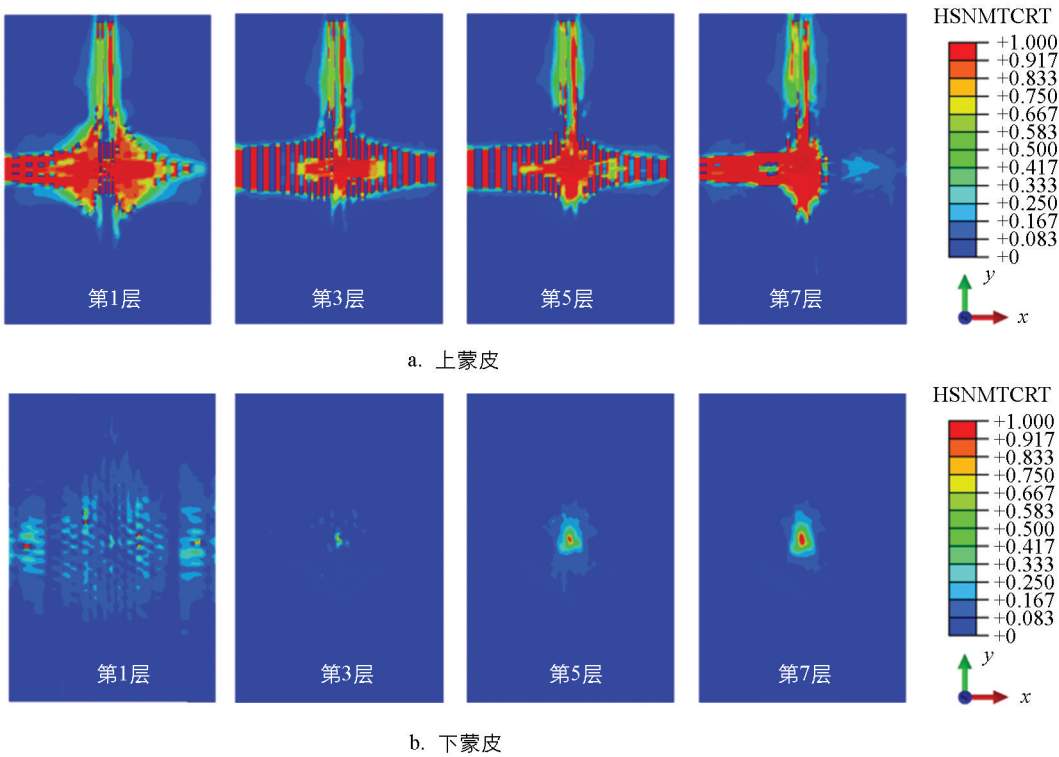


图 4 CFRP 基体拉伸损伤图像

在构建数据集时, 本研究采用了一种数据扩充的方法, 即在 1 个仿真模型的第 300 帧到 500 帧中每 20 帧取 1 组损伤图像构建数据集, 这样可以有效减少时间成本, 所得到的图像数据集满足深度学习模型训练的要求。此外, 通过仿真结果得到的原始数据集不能直接用来训练模型, 还需要对图像进行预处理来防止模型的过拟合。在原始图像中, 有部分是图例信息, 可视作损伤图像的噪声, 为了简化网络计算以及去除背景和其他噪点的干扰, 将原始损伤图像像素裁剪为 310×412 。此外, 为了增强神经网络的鲁棒性, 对数据进行扩充及转换。对于同一张图片, 结果应具有旋转不变性, 因此对图像进行随机旋转。图 5 展示了旋转 180° 前后的损伤图像, 最终得到处理好的损伤图像数据集。

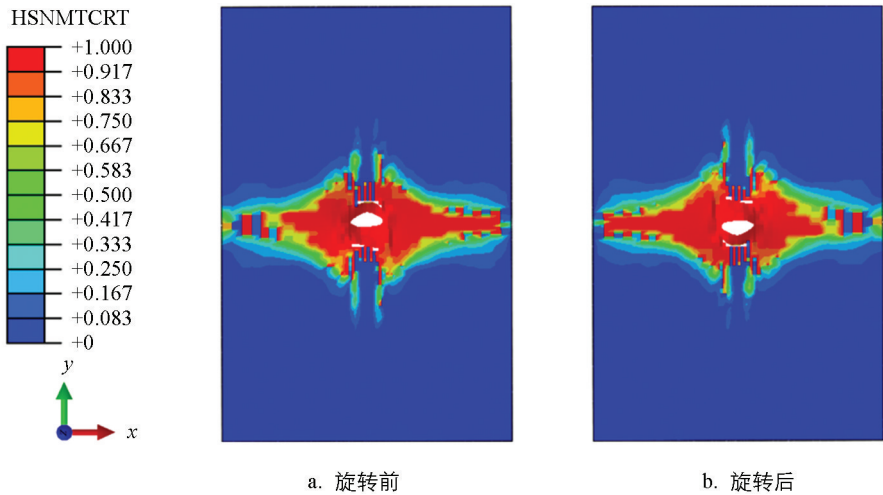


图 5 旋转前后的基体损伤图像

2.2 注意力机制算法

注意力机制模块在深度学习神经网络中的作用类似于人类视觉注意力的机制,它允许模型在处理数据信息时能够聚焦于当前任务的重要部分,从而提高模型的性能和效率。本研究采用的注意力机制是一种多头注意力机制(Multi-Head Attention),其主要思想是分别聚焦数据集中的多个方面,多个“头”采取并行的方式,对数据集中不同的特点进行注意力分数的计算,最终将所有“头”的注意力分数拼接起来,通过 1 个线性变换得到多头注意力机制的最终输出。多头注意力模块的算法流程如图 6 所示,其中头数 H 为 16。当模型的输入是一批图像时,输入矩阵为 $(batch_size, channel, height, width)$,通过模型将矩阵展平为 $(batch_size, sequence_length, inner_dim * 3)$ 作为输入 $\mathbf{X}$ 进行计算。

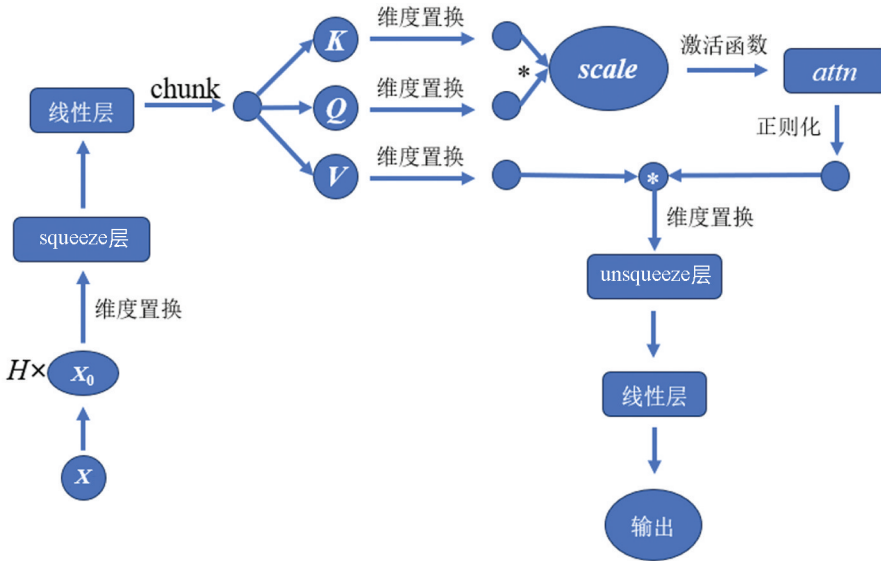


图 6 多头注意力模块的算法流程

首先将输入矩阵 $\mathbf{X}$ 映射到多个“头”中,得到多个子矩阵 $\mathbf{X}_0$,再将输入的子矩阵进行维度置换并经过 1 个 squeeze 层,随后再经过 1 个线性层进行线性变换,通过 chunk 函数输出 3 个矩阵,分别为键矩阵 $\mathbf{K}$ 、查询矩阵 $\mathbf{Q}$ 、值矩阵 $\mathbf{V}$,公式为:

$$\begin{bmatrix} \mathbf{K} \\ \mathbf{Q} \\ \mathbf{V} \end{bmatrix} = \mathbf{W} \cdot \mathbf{X}_0 \quad (1)$$

式中: $\mathbf{W}$ 为线性层的权重矩阵。

将 $\mathbf{K}$ 、 $\mathbf{Q}$ 、 $\mathbf{V}$ 矩阵维度置换之后对 $\mathbf{K}$ 和 $\mathbf{Q}$ 做点积运算求得注意力分数,公式为:

$$dots_{ij} = \frac{\mathbf{Q}_i \mathbf{K}_j^T}{\sqrt{d_k}} \quad (2)$$

式中: $dots_{ij}$ 为注意力分数,衡量了第 i 个查询和第 j 个键的相关性; $\mathbf{Q}_i$ 为查询矩阵的第 i 行; $\mathbf{K}_j$ 为键矩阵的第 j 列; d_k 为键矩阵的维度,这里等于 $\dim_head$; $\frac{1}{\sqrt{d_k}}$ 为缩放点积注意力中的缩放因子。

将运算结果通过 Softmax 归一化为概率分布得到注意力权重 A_{ij} ,公式为:

$$A_{ij} = \frac{\exp(dots_{ij})}{\sum_j \exp(dots_{ij})} \quad (3)$$

随后,通过 Dropout 操作,对 A_{ij} 进行随机失活,防止过拟合,将 A_{ij} 与 $\mathbf{V}$ 进行加权求和,得到每个“头”的输出,公式为:

$$\mathbf{o}_i^h = \sum_j A_{ij} \mathbf{v}_j$$

(4)

式中: $\mathbf{o}_i^h$ 为第 h 个“头”的第 i 个输出; $\mathbf{v}_j$ 为值矩阵的第 j 行。

对得到的结果矩阵通过 unsqueeze 层进行增维重排, 并将所有“头”的输出通过线性层 nn.Linear 拼接之后得到最终的注意力输出。

由上述注意力模块的计算过程可知, 注意力机制的输入数据与其输出加权和向量的大小是一致的, 即注意力机制并不会改变输入矩阵的形状。因此, 通过 torch 生成 1 个形状与输入数据相同的随机矩阵, 将函数输入该随机矩阵, 观察函数输出矩阵的形状。若输出矩阵的形状与输入随机矩阵的形状一致, 则该注意力机制便可以使用。由于本研究使用的图像数据是每次输入大小为 (64, 3, 224, 224) 的矩阵, 数组中的 4 个参数代表 (batch_size, channel, height, width), 而该注意力模块需要的输入为三维矩阵, 因此将矩阵中的 height 和 width 相乘, 并做换序处理, 作为模块的输入。随后生成 1 个大小为 (64, 224×224, 3) 的随机矩阵作为试验, 输出的矩阵形状与输入的随机矩阵形状一致, 注意力模块即可以使用。

注意力机制的本质是神经网络中前向传播中的加权环节, 因此在向模型中添加注意力机制时, 核心目标是通过动态计算权重来突出输入数据中与当前任务更相关或重要的部分, 从而使模型能够更有效地聚焦于关键信息, 提升性能和效果。将注意力机制函数添加到模型当中, 在模型前向传播 forward 函数中调用注意力机制函数, 对输入数据进行注意力加权, 输出后继续进行前向传播的剩余计算。不同的模型结构不同, 对应的前向传播步骤也不同, 因此对于不同的模型, 注意力机制函数添加在 forward 函数中的具体位置也不同, 必要时需要将卷积、循环等操作进行添加。从模型优化的角度来看, 注意力模块的添加位置存在多种选择策略, 包括层级间嵌入和残差连接嵌入。

2.3 深度学习模型配置与训练

本研究的神经网络模型是基于 PyTorch 2.1.2 版本实现的。为了训练模型, 采用分类交叉熵损失函数, 初始的训练学习率为 0.001, 模型迭代步数为 100。交叉熵损失函数是一种常用于分类任务的损失函数, 它通过衡量模型预测的概率分布与真实分布之间的差异来优化模型参数, 其核心思想是让模型预测的正确类别概率尽可能接近 1, 从而指导模型的训练方向。交叉熵损失函数的优点在于能够明确地反映预测与真实标签之间的差异, 并且在与 Softmax 函数结合后具有良好的数值稳定性。然而, 其对预测错误的惩罚较大, 容易受到异常值的影响, 并且需要模型输出概率分布。

深度学习中神经网络是一个大规模的、非线性的、自适应的动态系统, 由许多相互连接的单元组成。神经网络结构中的基本单元是人工神经单元或神经元, 它通过激活函数将上一层的线性加权和转换为非线性激活值传递到下一层。激活函数有多种类型, 最常用的有 Sigmoid、tanh、整流线性单元 (ReLU)、Softmax。表 1 展示了几种深度学习中常用的激活函数表达式和取值范围。与非线性激活函数 (Sigmoid 和 tanh) 相比, ReLU 不存在梯度消失问题, 其非负区间的梯度是恒定的, 因此该模型的收敛率保持在稳定状态; Softmax 函数一般只用于网络的最后一层进行分类和归一化。

表 1 激活函数

激活函数名称	表达式	取值范围
Sigmoid	$\frac{1}{1 + e^{-x}}$	(0, 1)
tanh	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	(-1, 1)
ReLU	$\max(0, x)$	(0, +∞)
Softmax	$\frac{e^i}{\sum_j e^j}$	(0, 1)

随机梯度下降(Stochastic Gradient Descent, SGD)是一种常用的深度学习优化方法,在处理大型数据集时具有优势,其主要思想是每次迭代过程中随机选择 1 个样本或小批量样本,然后再计算这个样本或小批量样本的梯度,并用其更新模型的参数达到训练模型的目的。它的优点是计算非常简单,每次计算只需要 1 个样本或小批量的梯度即可完成,收敛速度较快,调整学习率也可以使其更加优秀地适应不同的图像特征。

ResNet50 是一种经典的深度卷积神经网络,由多个层构成,包括卷积、批量归一化、激活函数、池化和全连接等,其核心在于通过残差块解决梯度爆炸问题,使网络训练更深层,从而提取更复杂的特征。本研究将注意力机制添加到 ResNet50 网络的全局平均池化层之前,即在模型将特征图转化为特征向量之前添加注意力。在网络接近输出的位置添加注意力机制,能够对前面所有阶段提取的特征进行最后的整合和筛选,聚焦于最具判别性的特征,使得最终输入到全连接层的特征向量更具有代表性,从而提高模型的准确率。通过关注主要特征,减少无关信息的干扰,有助于模型更好地泛化到不同的数据集合任务中。

RegNet 是一种基于神经架构搜索(Neural Architecture Search, NAS)技术的卷积神经网络,以系统方式探索高效网络结构,在有限计算资源下实现优异的表现。RegNet 采用层次化设计,分为多个阶段,每个阶段由相同模块组成,结构规则化,便于训练和部署。每个模块使用瓶颈结构,包括降维、特征提取和升维 3 个步骤,与 ResNet50 类似,但更注重宽度、深度的平衡,通过优化参数将特征表达能力最大化。本研究将注意力机制添加到 RegNet 网络的初始卷积层之后,初始卷积层负责提取图像的底层特征,如边缘和纹理信息。在此处添加多头注意力机制,模型可以在早期就聚焦于图像中的关键区域和特征,增强底层特征的表达。同时通过引入全局信息,让模型初始便对图像整体布局有所了解,有助于后续网络模块更好地理解图像,提升后续深层特征提取的有效性。

本研究将准备好的数据集按照 70%训练集、15%验证集和 15%测试集的比例将图像数据进行划分,其中测试集的数据由未包含在训练过程中的损伤图像组成,确保了对模型泛化到未见损伤图像样本能力的严格评估。训练模型时,以 CFRP 蜂窝铝夹心层合板基体拉伸损伤图像为输入,层合板受到的冲击能量为输出进行训练,实现图像识别的分类功能。训练过程中采用小批量学习方法,当损失值减小到一定值、准确率趋于稳定不再上升或者模型无法收敛时结束模型的训练,超参数的值取训练迭代过程中准确率最高的训练参数数据。训练和测试过程在每个模型上重复 3 次,每次重复会再次重新随机分配训练集、验证集和测试集,取训练结果最优的模型。

3 结果与分析

本研究采用准确率 A (Accuracy)对模型进行评估,计算公式如下:

$$A = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

式中: TP (True Positive)为做出 Positive 的判定,且判定是正确的次数; FP (False Positive)为做出 Positive 的判定,但判定是错误的次数; TN (True Negative)为做出 Negative 的判定,但判定是正确的次数; FN (False Negative)为做出 Negative 的判定,但判定是错误的次数。

表现优异的模型,其准确率在 90%以上,并且不会出现梯度爆炸以及梯度消失的情况。损失值的大小随着迭代过程而不断降低,当接近于 0 时,模型中的超参数几乎不再更替变化,模型趋于稳定,准确率也不再有明显的变化,此时不需要对模型进行再训练。因此,在训练的时候若出现迭代次数未达到训练步数的设定值就停止的情况,这并不会影响试验数据和结果,反而提高了训练效率,节省了时间成本。

3.1 不同深度学习模型效果对比

将改进前后的 ResNet50 和 RegNet 网络模型训练得到的准确率与损失曲线进行对比,如图 7 所示。由

图 7a、7c 可以看出, 改进前 2 种模型的训练曲线在训练动态与性能表现上具有显著差异。ResNet50 的损失曲线在训练初期快速下降, 体现出模型对数据特征的初期学习效率, 可迅速缩小预测值与真实值之间的差距; 进入训练中期后损失曲线下下降平滑, 反映出训练稳定, 网络稳步收敛, 其准确率曲线平均值也趋于稳定, 但中期准确率不断震荡, 说明模型对数据特征提取模式未稳固; 训练后期, ResNet50 的损失曲线基本收敛于一个较小的值, 准确率增长趋缓, 稳定在 93.7% 左右。而 RegNet 的损失曲线在训练初期虽有大幅下降, 但相较 ResNet50 更保守, 无剧烈震荡; 训练中期损失曲线持续稳步降低, 波动幅度小, 这得益于其网络结构对梯度传播的优化, 参数更新更加有效, 而准确率曲线虽呈持续上升趋势, 但模型整体收敛缓慢, 最终可能并未完全收敛; 最后 RegNet 的准确率达到 93.1% 左右, 损失曲线收敛水平比 ResNet50 低, 说明 RegNet 收敛缓慢, 对信息的提取能力也不如 ResNet50。

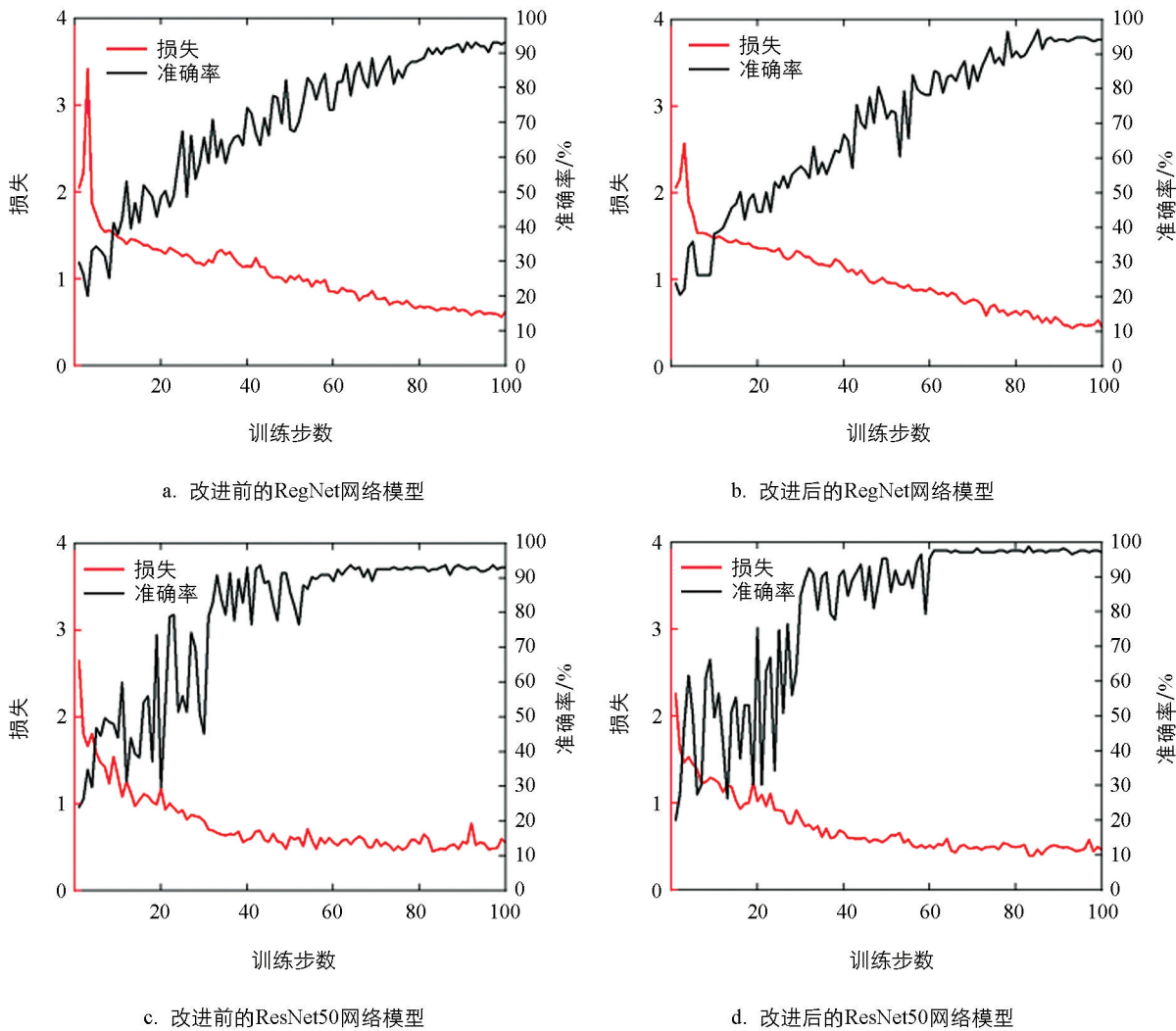


图 7 深度学习模型训练结果

综合对比, ResNet50 能够更加高效地提取图像信息, 而 RegNet 收敛缓慢, 并且准确率相较于 ResNet50 更低, 对信息的特征提取能力不如 ResNet50。这一差异凸显了不同网络结构处理训练动态能力的差异, 即 RegNet 的设计性能更加保守, 而 ResNet50 的学习能力更强。

3.2 改进后的深度学习模型效果对比

ResNet50 网络模型在添加多头注意力机制之后, 训练得到的准确率与损失曲线如图 7d 所示。对比图 7c、7d 可知, ResNet50 在添加多头注意力机制前后的训练曲线出现明显的变化。在未添加多头注意力机制时, 模型的训练过程表现出较大的不稳定性, 主要表现为损失曲线波动剧烈, 尤其是训练步数在 20~40 阶

段, 准确率曲线也相应出现高频震荡。这表明模型对数据噪声和复杂样本较为敏感, 特征提取模式不够稳定。在引入多头注意力机制后, 训练的稳定性得到了显著提升, 主要表现为损失曲线的波动幅度大幅减小, 下降趋势更加平滑, 准确率曲线的震荡频率和幅度也明显降低。这种改进源于多头注意力机制的作用, 它使模型能够更好地关注关键特征, 同时抑制无关信息, 从而帮助模型更稳定地捕捉数据的内在模式并逐步固化特征提取方式。

模型训练效果如表 2 所示, 从收敛效果来看, 在未添加多头注意力机制时, ResNet50 网络模型的损失曲线虽然在后期持续下降, 但最终仍停留在相对较高的水平, 而准确率增长也在后期趋于平缓, 这表明模型存在一定的拟合瓶颈, 难以深入挖掘数据中的深层特征。相比之下, 添加多头注意力机制后, 损失曲线收敛到更低的值, 说明模型对任务的拟合能力得到了显著提升, 同时准确率也达到了更高水平, 稳定在 98.9% 左右, 这表明多头注意力机制有效增强了模型的特征表征能力, 其核心作用在于通过并行计算多个子空间的注意力分布, 使模型能够从多个角度捕捉数据特征之间的关联关系, 尤其是在处理图像中的长距离依赖和关键信息区域时表现出明显优势。这种机制弥补了原始 ResNet50 在特征筛选和权重分配上的局限, 使模型能够更高效地利用数据信息。不仅如此, 它同时还优化了训练的稳定性、收敛性能以及特征学习能力, 实现了模型综合表现的全面提升。

表 2 模型训练效果

模型类型	准确率/%	损失
ResNet50	93.7	0.45
ResNet50+多头注意力机制	98.9	0.39
RegNet	93.1	0.56
RegNet+多头注意力机制	97.1	0.44

RegNet 网络模型在添加多头注意力机制之后, 训练得到的准确率与损失曲线如图 7b 所示。对比图 7a、7b 可知, RegNet 在添加多头注意力机制前后的训练曲线出现明显的变化。在未添加多头注意力机制时, 模型在训练初期表现出快速下降的损失曲线, 随后在训练中期出现了较大的波动, 这表明模型在训练过程中的稳定性不足, 容易受到数据特征的干扰影响。同时, 准确率曲线虽然总体呈上升趋势, 但伴随着频繁的振荡, 反映出模型在特征提取过程中缺乏稳定性, 导致对样本的判别能力无法持续提升。在引入多头注意力机制后, 模型的训练稳定性得到了显著改善, 损失曲线的波动幅度明显减小, 下降趋势更加平缓, 这说明模型在优化过程中能够更稳定地逼近最优解。而准确率曲线也表现出更低的振荡频率和幅度, 上升趋势更加平稳, 表明模型能够持续且稳定地学习数据特征, 从而减少了因特征学习不充分导致的预测波动。从收敛效果来看, 在未添加注意力机制的情况下, 损失曲线在后期虽然继续下降, 但仍然存在一定的优化空间, 准确率曲线的增长也显得乏力。而通过引入多头注意力机制后, 损失曲线进一步收敛到更低的水平, 表明模型对任务的拟合更加充分, 同时准确率达到更高的水平, 并稳定在 97.1%, 这充分体现了多头注意力机制对模型特征表征能力的显著提升。

多头注意力机制通过专注于关键区域和重要特征, 抑制无关干扰信息, 从而提高了特征提取的质量。同时, 其多子空间并行计算的特性能够捕捉不同维度的特征关联, 这弥补了原始 RegNet 在特征权重分配上的不足, 使得模型能够更高效地利用数据中的关键信息, 不仅优化了训练过程, 还显著提升了最终的分类性能。通过这种双重提升, RegNet 在处理复杂特征时展现出更强的优势, 并能更充分地挖掘数据的潜在价值, 表明引入多头注意力机制是 RegNet 网络模型优化和性能提升的一种有效方法。

综上所述, 通过多头注意力机制的添加, 大幅提升了 ResNet50 网络模型的训练稳定性和特征学习能力, 而 RegNet 网络模型在特征学习能力方面也得到提升, 但模型收敛仍然较为缓慢。

4 结论

本研究针对 CFRP 蜂窝铝三明治结构, 利用深度学习方法构建了材料内部损伤和所受冲击能量之间的关系, 借助有限元仿真, 避免了获取数据集时浪费过多的试验和制备成本。分别在深度学习模型 ResNet50 和 RegNet 中引入一种多头注意力机制模块进行改进, 改进后的 2 种网络模型相较于原模型, 其准确率和效率都有不同程度的提升。最终, 改进后 ResNet50 和 RegNet 的准确率分别达到了 98.9% 和 97.1%, 在通过损伤图像预测冲击能量的任务中达到了理想的效果。同时, 本研究还引入了一种在有限数据体量情况下高效扩充数据集的方法, 提高了数据准备的效率。本研究在冲击能量预测的方法上具有普适性, 可应用到其他复合材料损伤图像和损伤指标预测中, 为复合材料冲击力学性能研究提供了一种新的思路和参考。

参考文献:

- [1] SHI Y, PINNA C, SOUTIS C. Low-Velocity Impact of Composite Laminates: Damage Evolution [M] //Dynamic Deformation, Damage and Fracture in Composite Materials and Structures. Amsterdam: Elsevier, 2023: 89-119.
- [2] ZHANG X F, WU X, HE Y Z, et al. CFRP Barely Visible Impact Damage Inspection Based on an Ultrasound Wave Distortion Indicator [J]. Composites Part B: Engineering, 2019, 168: 152-158.
- [3] LIANG T, REN W W, TIAN G Y, et al. Low Energy Impact Damage Detection in CFRP Using Eddy Current Pulsed Thermography [J]. Composite Structures, 2016, 143: 352-361.
- [4] BALAGEAS D L, ROCHE J M, LEROY F H, et al. The Thermographic Signal Reconstruction Method: A Powerful Tool for the Enhancement of Transient Thermographic Images [J]. Biocybernetics and Biomedical Engineering, 2015, 35(1): 1-9.
- [5] 张峻铭, 杨伟东, 李岩. 人工智能在复合材料研究中的应用 [J]. 力学进展, 2021, 51(4): 865-900.
- [6] KARSANDIK Y, SABUNCUOGLU B, YILDIRIM B, et al. Impact Behavior of Sandwich Composites for Aviation Applications: A Review [J]. Composite Structures, 2023, 314: 116941.
- [7] LI Z Y, MA Z, WANG J F, et al. Low-Velocity Impact Behavior and Damage Mechanisms of Honeycomb Sandwich Structures with Elastomeric Interlayers in CFRP Skins [J]. Thin-Walled Structures, 2024, 205: 112482.
- [8] LIAO J, GUO L P, JIANG L, et al. A Machine Learning-Based Feature Extraction Method for Image Classification Using ResNet Architecture [J]. Digital Signal Processing, 2025, 160: 105036.
- [9] CHEN Y, SHARIFUZZAMAN S A S M, WANG H X, et al. Deep Learning Based Underground Sewer Defect Classification Using a Modified RegNet [J]. Computers, Materials & Continua, 2023, 75(3): 5455-5473.
- [10] SUMIT S S, ANAVATTI S, TAHTALI M, et al. ResNet-Lite: On Improving Image Classification with a Lightweight Network [J]. Procedia Computer Science, 2024, 246: 1488-1497.
- [11] OLIVEIRA B C F, SEIBERT A A, BORGES V K, et al. Employing a U-Net Convolutional Neural Network for Segmenting Impact Damages in Optical Lock-in Thermography Images of CFRP Plates [J]. Nondestructive Testing and Evaluation, 2021, 36(4): 440-458.
- [12] LIU H C, LI W H, YANG L C, et al. Automatic Reconstruction of Irregular Shape Defects in Pulsed Thermography Using Deep Learning Neural Network [J]. Neural Computing and Applications, 2022, 34(24): 21701-21714.
- [13] WEI Z A, FERNANDES H, HERRMANN H G, et al. A Deep Learning Method for the Impact Damage Segmentation of Curve-Shaped CFRP Specimens Inspected by Infrared Thermography [J]. Sensors, 2021, 21(2): 395.
- [14] OLIVER G A, ANCELOTTI A C, GOMES G F. Neural Network-Based Damage Identification in Composite Laminated Plates Using Frequency Shifts [J]. Neural Computing and Applications, 2021, 33(8): 3183-3194.
- [15] DENG K L, LIU H C, YANG L C, et al. Classification of Barely Visible Impact Damage in Composite Laminates Using Deep Learning and Pulsed Thermographic Inspection [J]. Neural Computing and Applications, 2023, 35(15): 11207-11221.

[16] CAO Y P, DONG Y F, CAO Y L, et al. Two-Stream Convolutional Neural Network for Non-Destructive Subsurface Defect Detection via Similarity Comparison of Lock-in Thermography Signals [J]. NDT & E International, 2020, 112: 102246.

[17] SAEED N, KING N, SAID Z, et al. Automatic Defects Detection in CFRP Thermograms, Using Convolutional Neural Networks and Transfer Learning [J]. Infrared Physics & Technology, 2019, 102: 103048.

[18] DONG Y F, XIA C J, YANG J X, et al. Spatio-Temporal 3-D Residual Networks for Simultaneous Detection and Depth Estimation of CFRP Subsurface Defects in Lock-in Thermography [J]. IEEE Transactions on Industrial Informatics, 2022, 18(4): 2571-2581.

[19] HU B Z, GAO B, WOO W L, et al. A Lightweight Spatial and Temporal Multi-Feature Fusion Network for Defect Detection [J]. IEEE Transactions on Image Processing, 2020, 30: 472-486.

[20] DENG K L, LIU H C, CAO J, et al. Attention Mechanism Enhanced Spatiotemporal-Based Deep Learning Approach for Classifying Barely Visible Impact Damages in CFRP Materials [J]. Composite Structures, 2024, 337: 118030.

[21] 徐杰. 基于深度学习和损伤图像预测车用 CFRP 冲击能量的研究 [D]. 哈尔滨: 哈尔滨工业大学, 2024.

[22] HASEBE S, HIGUCHI R, YOKOZEKI T, et al. Internal Low-Velocity Impact Damage Prediction in CFRP Laminates Using Surface Profiles and Machine Learning [J]. Composites Part B: Engineering, 2022, 237: 109844.

[23] 端木正, 张晨, 肖航, 等. 一种基于多头自注意力深度学习的巨阻效应纳米线材料设计方法: 202411284915. X [P]. 2025-04-01.

[24] 杜禹樵, 马成坤, 王柏涛, 等. 基于碳纳米纸传感器和深度学习的碳纤维复合材料损伤监测 [J]. 沈阳航空航天大学学报, 2024, 41(3): 43-52.

[25] WANG Z F, ZHAO C C, YANG Z Y, et al. Multi-Scale Collaborative Prediction of Optimal Configuration for Carbon Fiber Woven Composites Based on Deep Learning Neural Networks [J]. Composite Structures, 2024, 339: 118165.

[26] 王敏, 文鹤鸣. 碳纳米管/碳纤维增强复合材料层合板低速冲击响应和破坏的数值模拟 [J]. 爆炸与冲击, 2022, 42(3): 25-36.

[27] HUANG J S, LIEW J X, LIEW K M. Data-Driven Machine Learning Approach for Exploring and Assessing Mechanical Properties of Carbon Nanotube-Reinforced Cement Composites [J]. Composite Structures, 2021, 267: 113917.

[28] YANG C X, MENG K P, YANG L T, et al. Transfer Learning-Based Crashworthiness Prediction for the Composite Structure of a Subway Vehicle [J]. International Journal of Mechanical Sciences, 2023, 248: 108244.

[29] LOUTAS T, OIKONOMOU A, REKATSINAS C. Bio-Inspired Discontinuous Composite Materials with a Machine Learning Optimized Architecture [J]. Composite Structures, 2025, 351: 118597.

[30] 钱奇伟, 张昕, 杨贞军, 等. 基于 CT 图像深度学习的三维编织 C/C 复合材料微观组分与缺陷智能识别 [J]. 复合材料学报, 2024, 41(7): 3536-3543.

[31] FONSECA J H, JANG W, HAN D, et al. Strength and Manufacturability Enhancement of a Composite Automotive Component via an Integrated Finite Element/Artificial Neural Network Multi-Objective Optimization Approach [J]. Composite Structures, 2024, 327: 117694.

责任编辑 柳剑