

基于“AM-GoogLeNet + BP”联合数据驱动的混凝土细观模型压缩应力应变曲线预测

刘溢凡, 张 杰, 张新宇, 王志勇, 王志华

PREDICTION OF CONCRETE MESO-MODEL COMPRESSION STRESS-STRAIN CURVE BASED ON “AM-GOOGLENET + BP” COMBINED DATA-DRIVEN METHODS

Liu Yifan, Zhang Jie, Zhang Xinyu, Wang Zhiyong, and Wang Zhihua

在线阅读 View online: <https://doi.org/10.6052/0459-1879-22-506>

您可能感兴趣的其他文章

Articles you may be interested in

基于数据驱动的流场控制方程的稀疏识别

DATA-DRIVEN SPARSE IDENTIFICATION OF GOVERNING EQUATIONS FOR FLUID DYNAMICS

力学学报. 2021, 53(6): 1543-1551

基于深度学习和细观力学的颗粒材料本构关系研究

CONSTITUTIVE RELATIONS OF GRANULAR MATERIALS BY INTEGRATING MICROMECHANICAL KNOWLEDGE WITH DEEP LEARNING

力学学报. 2021, 53(9): 2404-2415

基于细观混凝土模型的时间逆转损伤成像方法

DAMAGE DETECTION OF MESO-SCALE CONCRETE STRUCTURES BASED ON TIME REVERSAL METHOD

力学学报. 2017, 49(4): 953-960

基于深度强化学习算法的颗粒材料应力应变关系数据驱动模拟研究

DATA-DRIVEN STRESS-STRAIN MODELING FOR GRANULAR MATERIALS THROUGH DEEP REINFORCEMENT LEARNING

力学学报. 2021, 53(10): 2712-2723

基于多变量小样本的渗流代理模型及产量预测方法

SEEPAGE PROXY MODEL AND PRODUCTION FORECAST METHOD BASED ON MULTIVARIATE AND SMALL SAMPLE

力学学报. 2021, 53(8): 2345-2354

基于一类非局部宏-微观损伤模型的混凝土典型试件力学行为模拟

SIMULATION OF BEHAVIOUR OF TYPICAL CONCRETE SPECIMENS BASED ON A NONLOCAL MACRO-MESO-SCALE CONSISTENT DAMAGE MODEL

力学学报. 2021, 53(4): 1196-1121



关注微信公众号, 获得更多资讯信息

基于“AM-GoogLeNet + BP”联合数据驱动的混凝土 细观模型压缩应力-应变曲线预测¹⁾

刘溢凡^{*,†} 张 杰^{*,†} 张新宇^{*,†} 王志勇^{*,2)} 王志华^{*,†}

^{*}(太原理工大学机械与运载工程学院应用力学研究所, 太原 030024)

[†](材料强度与结构冲击山西省重点实验室, 太原 030024)

摘要 本文结合 GoogLeNet 卷积神经网络和 BP 神经网络分别在图像数据挖掘和数据分析方面的良好性能, 采用“AM-GoogLeNet + BP”联合数据驱动方法, 对混凝土细观模型(含砂浆、骨料及孔隙)的单轴压缩应力-应变曲线进行了有效预测. 通过引入力学参量对图像数据驱动的训练结果进行优化, 从而提升了神经网络的物理可解释性. 基于 Python 语言实现混凝土细观模型在 Abaqus 中的自动建模及细观图像生成过程, 并将生成的细观图像数据库与相应的压缩应力-应变曲线作为训练数据集. 在 GoogLeNet 中分别引入 SENet, ECANet 和 CBAM 三种代表性注意力机制并对三种注意力机制的性能进行对比和分析, 以自适应方式提升神经网络对混凝土各相组分的分析能力, 并以此得到混凝土细观模型的初步应力-应变预测曲线; 将骨料体积分数、孔隙率及初步峰值应力等物理参量作为输入引入 BP 神经网络以改善峰值应力的预测精度, 并与将物理参量直接引入卷积神经网络输入层的方法进行了对比, 最后定量给出了骨料体积分数和孔隙率对峰值应力的影响权重. 结果表明, 对于不同骨料体积分数及孔隙率的混凝土细观模型, 该方法均展现了较高的预测精度. 本文采用的“AM-GoogLeNet + BP”联合数据驱动预测模型从统计角度解决了传统方法对细观尺度参量分析的复杂性, 为复合材料的跨尺度力学行为研究提供了新思路.

关键词 混凝土, 神经网络, 数据驱动, 细观模型, 单轴压缩

中图分类号: TU3T, TP39, O34 文献标识码: A doi: 10.6052/0459-1879-22-506

PREDICTION OF CONCRETE MESO-MODEL COMPRESSION STRESS-STRAIN CURVE BASED ON “AM-GOOGLENET + BP” COMBINED DATA-DRIVEN METHODS¹⁾

Liu Yifan^{*,†} Zhang Jie^{*,†} Zhang Xinyu^{*,†} Wang Zhiyong^{*,2)} Wang Zhihua^{*,†}

^{*}(Institute of Applied Mechanics, College of Mechanical and Vehicle Engineering, Taiyuan University of Technology,
Taiyuan 030024, China)

[†](Shanxi Key Laboratory of Material Strength and Structural Impact, Taiyuan 030024, China)

Abstract This paper uses the “AM-GoogLeNet + BP” combined data-driven methods to predict the uniaxial

2022-10-19 收稿, 2023-01-17 录用, 2023-01-19 网络版发表.

1) 国家自然科学基金(12272257, 12102292)和山西省基础研究计划(20210302124083)资助项目.

2) 通讯作者: 王志勇, 副教授, 主要研究方向为断裂力学. E-mail: wangzhiyong@tyut.edu.cn

引用格式: 刘溢凡, 张杰, 张新宇, 王志勇, 王志华. 基于“AM-GoogLeNet + BP”联合数据驱动的混凝土细观模型压缩应力-应变曲线预测. 力学学报, 2023, 55(4): 925-938

Liu Yifan, Zhang Jie, Zhang Xinyu, Wang Zhiyong, Wang Zhihua. Prediction of concrete meso-model compression stress-strain curve based on “AM-GoogLeNet + BP” combined data-driven methods. *Chinese Journal of Theoretical and Applied Mechanics*, 2023, 55(4): 925-938

compression stress-strain curve of the concrete meso-model (including mortar, aggregates, porosity) effectively by combining the good performance of GoogLeNet convolutional neural network and BP neural network in image data mining and data analysis, respectively. The physical interpretability of the neural network is improved by introducing the mechanical parameters to optimize the image data-driven training results. The automated modeling of the concrete meso-model in Abaqus and microscopic image generation process are realized by Python language, and the generated mesoscopic image database and the corresponding compression stress-strain curves are used as the training dataset. Three typical attention mechanisms, SENet, ECANet and CBAM, are introduced into GoogLeNet respectively to enhance the analysis ability of neural network for each phase of concrete in an adaptive manner and the performance of the three attention mechanisms is compared and analyzed. The initial stress-strain prediction curves of the concrete meso-model are obtained with this method; In order to improve the prediction accuracy of the peak stress, the physical parameters such as aggregate volume fraction, porosity and initial peak stress are introduced into BP neural network as inputs. It is also compared with the method of introducing the physical parameters directly into the convolutional neural network input layer. At the same time, the weight of influence of aggregate volume fraction and porosity on peak stress is given quantitatively. The results show that this method has high prediction accuracy for the concrete meso-model with different aggregate volume fraction and porosity. In this paper, the “AM-GoogLeNet + BP” combined data-driven prediction model is used to solve the complexity of the traditional method in the analysis of mesoscale parameters from the statistical point of view, which provides a new idea for the study of the cross-scale mechanical behavior of composite materials.

Key words concrete, neural network, data-driven, meso-model, uniaxial compression

混凝土是一种具有代表性的复合材料,一般认为其细观组分包括砂浆、骨料、孔隙及界面过渡层,混凝土的细观组分对其宏观力学性能有决定性影响^[1-2]。但是由于细观层次一般具有力学参量和界面的复杂性、增强相和初始缺陷分布的随机性等特点,使得用传统的力学方法建立混凝土宏观细观联系的表达式极为困难^[3-4]。近年来,随着计算机算力的不断增长,深度学习技术为解决这些问题提供了新的范式,强大的数据驱动技术可以自动建立大量数据“端到端”的高精度映射,进而从统计学的角度研究混凝土的力学行为^[5]。

越来越多的研究人员将深度学习应用于复合材料的结构设计和优化、裂缝识别、力学性能预测等领域^[6-7]。卷积神经网络作为深度学习的代表性网络之一,可以有效提取图像从低维到高维的特征,从而输出数据标签。Yang 等^[8]提出了一种主成分分析加卷积神经网络的组合方法,用于预测二元复合材料在整个失效路径上的应力-应变曲线,结果表明预测的平均绝对误差小于 10%。增加神经网络层数是进一步提升卷积神经网络学习能力的常见方法,但同时也会带来梯度消失等使训练过程无法收敛的问题^[9]。GoogLeNet 的出现有效避免了这一难题,其核心架

构“Inception”模块的提出使神经网络变得“轻便且快速”。梁雪慧等^[10]使用改进的 GoogLeNet 对桥梁裂缝特征图像进行了训练,利用滑动窗口定位裂缝并结合骨架提取算法得到了裂缝的长度和宽度,相比于传统 GoogLeNet,识别精度提升了 3.13%。另一方面,注意力机制的引入也从模块化插入的角度提升了卷积神经网络的性能^[11-12]。王晓玲等^[13]基于耦合的 ALO-LSTM 神经网络,通过在特征维度上引入注意力机制,以较高的精度预测了土石坝的渗压效应量。

在实际工程中,尽管不断优化的神经网络结构大大提升了研究目标的预测效率及精度,但其较差的物理可解释性越来越受到诟病,神经网络的数据映射构建过程有时还可能会违背现有的部分物理原则。因此,在神经网络的训练过程中引入主要物理参量,使预测结果在满足预测效率的同时兼顾物理意义层面的准确度,已成为当前神经网络发展的重要趋势。Ye 等^[14]将杨氏模量及泊松比引入神经网络的全连接层,对具有任意分布夹杂物复合材料的有效弹性性能进行了预测。需要指出的是,将物理参量和其他神经元共同作为神经网络某一层合理性仍值得讨论,且作为固定值的物理参量能否在神经网络各层神经元的连接中保持其“物理属性”影响梯度下

降的结果并且在数据流动过程中不被“淡化”, 目前也存在争议^[15]. Hu 等^[16]提出了一种结合卷积神经网络和 BP 神经网络算法的棉纱强度预测方法 (CNN-BP neural network), 试验结果表明, 该方法的预测误差为 0.9%, 为物理背景在神经网络中的引入提供了新的思路.

应力-应变曲线是反映混凝土性能的重要表达形式, 从应力-应变曲线中可以获取强度、等效弹性模量等重要宏观性能. 因此, 本文使用引入注意力机制的改进 GoogLeNet, 即 AM-GoogLeNet, 对混凝土细观模型的单轴压缩应力-应变曲线进行预测, 然后提取预测的峰值应力, 引入骨料体积分数和孔隙率, 使用 BP 神经网络进一步提升预测精度, 并与将物理参量直接引入卷积神经网络输入层的方法进行了对比. 同时, 详细说明了 AM-GoogLeNet 和 BP 的网络架构、数据集的建立过程、各阶段的训练过程, 并对不同注意力机制下网络的预测精度定量评估和比较. 最后, 定量给出了骨料体积分数和孔隙率对混凝土应力-应变曲线峰值应力的影响权重.

1 注意力机制

1.1 注意力机制原理

Bahdanau 等^[17]提出的注意力机制 (attention mechanism), 是一种在神经网络中对数据进行再处

理的思想, 也是神经网络自适应学习的一种方式. 在神经网络中引入注意力机制后通常可以在不显著提升参数规模和训练时间的前提下有效提高预测精度, 即通过赋予输入特征层权重值, 使得神经网络在关注图像中感兴趣区域的同时抑制其他区域. 同时, 注意力模块可以根据需要嵌入到主流卷积神经网络的任意位置. 根据赋予权重方式和嵌入位置的不同, 注意力机制可以分为空间注意力机制、通道注意力机制和混合注意力机制. 本文对 SENet (squeeze-excitation net), ECANet (efficient channel attention net), CBAM (convolutional block attention module) 三种具有代表性的注意力机制进行了对比和分析.

1.2 SENet

SENet 作为一种具有代表性的通道注意力机制, 在 2017 年被 Hu 等^[18]提出. 图 1 展示了 SENet 的架构, F_{tr} 代表对输入的特征层进行 2D 卷积操作, 得到 X 后首先采用全局平均池化, 获取维数为 $1 \times 1 \times c$ 的数据, “挤压”将每个通道平面的信息压缩为一个数据点, 即去除了空间分布的影响, 强化通道分布的相关性, 同时在获得 X 的全部信息后增加感受野. 随后使用两个全连接层, 第一个全连接层将通道数 C 压缩以降低参数规模, 采用 ReLU 激活函数, 第二个全连接层将通道数恢复为 C , 采用“Sigmoid”激活函数生成 $[0,1]$ 的权重值. 最后将权重值与 X 的每个通道相乘得到 $\hat{X}$ 最为下一层的输入.

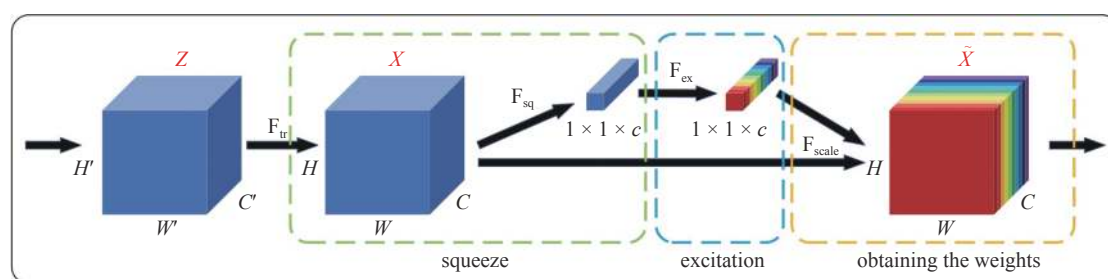


图 1 SENet 模块

Fig. 1 SENet module

1.3 ECANet

Wang 等^[19]认为, SENet 中“激励”过程的降低通道维数的做法对模型的预测结果带来了负面影响, 同时获得全局数据的非线性跨通道交互信息的效率较低. ECANet 在不降低通道维数的前提下采用局部跨通道交互的方法, 同时可以自适应地决定交互覆

盖率 k , 即卷积核的大小. 图 2 展示了 ECANet 的架构, 首先对输入特征层全局平均池化, 将数据的维数转化为 $1 \times 1 \times c$, 随后用 1D 卷积操作代替 SENet 中的全连接层. 同时, k 与通道维数成正比. 最后通过“Sigmoid”激活函数生成权重值并与原特征层的各个通道相乘.

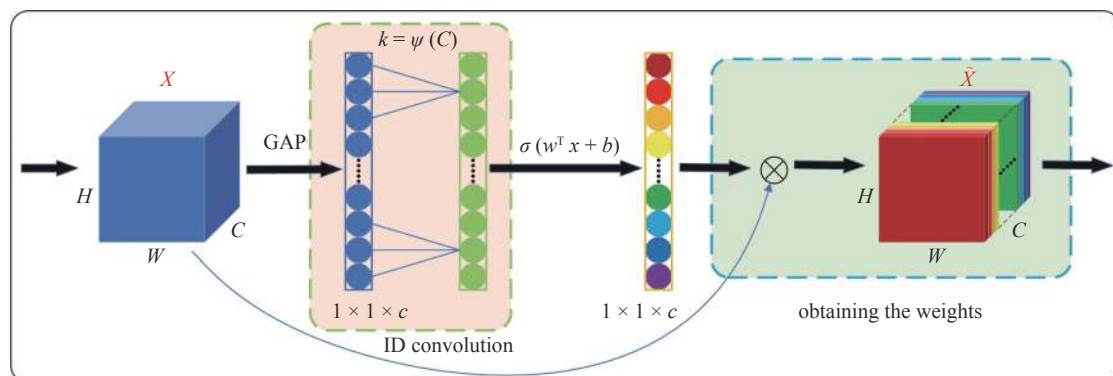


图2 ECANet 模块

Fig. 2 ECANet module

1.4 CBAM

CBAM 由 Woo 等^[20]在 2018 年提出,它结合了通道注意力机制和空间注意力机制,相比于 SENet 和 ECANet 往往可以取得更好的效果.图 3 展示了 CBAM 的架构,第一个部分为“channel attention module (CAM)”,首先对输入特征层分别进行最大池化和全局平均池化,将数据维数转为 $1 \times 1 \times c$,之后在共享的多层感知器中先压缩通道维数,再恢复通

道维数为 c .然后将处理后的两组数据逐个元素求和,通过“Sigmoid”激活函数产生权重值.最后用权重值与原特征层相乘.第二个部分为“spatial attention module (SAM)”,首先对输入特征层的每个元素的通道分别进行最大池化和平均池化,将数据维数转为 $H \times W \times 1$.随后通过“concat”操作对两组数据在通道上连接,再使用 2D 卷积将合并后的数据通道数压缩为 1,通过“Sigmoid”激活函数得到权重值.最后用权重值与原特征层相乘.

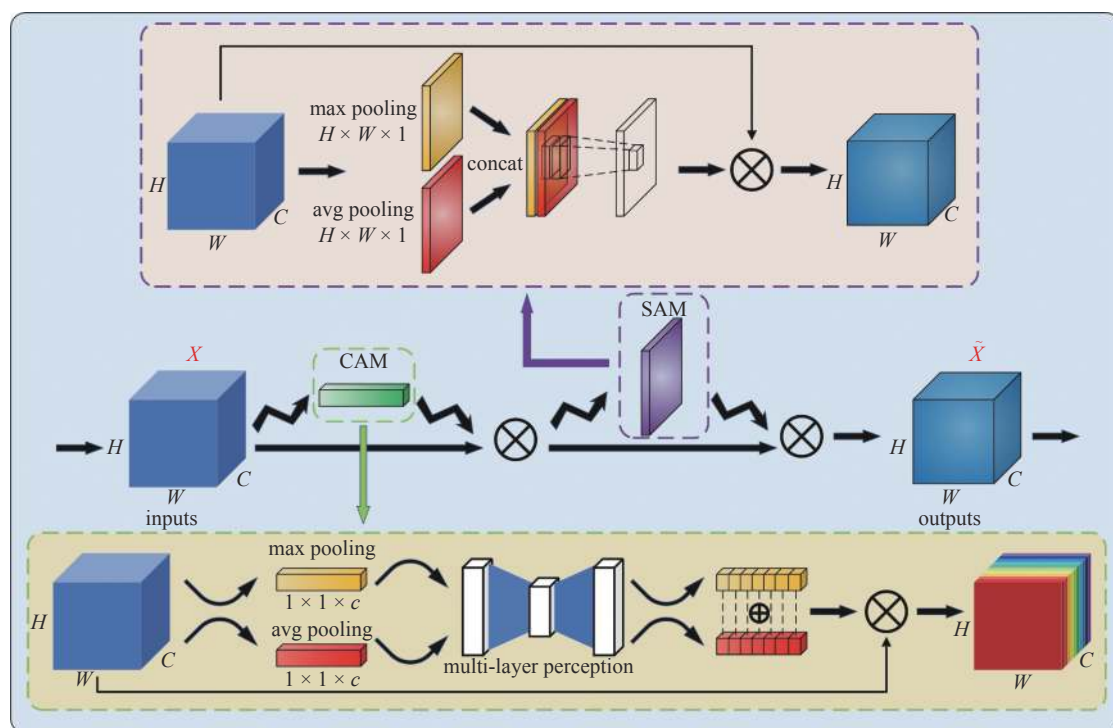


图3 CBAM 模块

Fig. 3 CBAM module

2 神经网络架构

2.1 GoogLeNet

为了解决小数据集训练中的过拟合问题和神经

网络变宽变深后的参数规模过大问题,谷歌团队提出了“network in network”的思想^[21],通过在主体神经网络中加入子网络的方式来增加网络的深度,使用“inception”模块代替传统卷积神经网络中“卷

积层+池化层”的连接方式,该模块使用 1×1 的卷积核进行降维以减小参数规模,同时避免了经过“concatenate”操作后特征层深度过大的问题.实践证明“GoogLeNet”通过改善数据连接的稀疏性保证了神经网络的性能.

2.2 AM-GoogLeNet

由于直接使用 GoogLeNet 得到从混凝土细观图像到应力-应变曲线映射的难度较高,尤其是软化段的预测精度较低^[22-23].故本文采用参数规模小且预测精度较高的 GoogLeNet 作为神经网络的主体,将 SE, ECA 和 CBAM 三种注意力模块分别嵌入以搭建 AM-GoogLeNet,利用注意力机制获得混凝土细观图像的通道权重和空间权重,有效读取骨料和孔隙随机分布的信息,使得神经网络可以自适应地、有针对性地提升对应力-应变特定阶段的预测精度(线弹性阶段,峰值应力,软化阶段).另外,为了保证加入注意力模块后减少对神经网络整体数据流动的影响^[24],嵌入位置选择位于神经网络末端的“inception”模块和全局平均池化层之间的位置.本文最终使用的神经网络架构如图 4 所示.

注意力模块的参数设置对其性能有较大的影响,下面是本文的注意力模块的参数设置^[18-20,25].

(1) SENet: 通道压缩系数采用 4,即将输入特征层的通道数由 C 压缩为 $C/4$ 后再恢复为 C .

(2) ECA Net: 本文设置一维卷积操作中的交互覆盖率 k 的计算公式如下

$$k = \lfloor (\log_2 C + 1) / 2 \rfloor \quad (1)$$

其中, C 为特征层的通道数.

同时通过程序控制保证此公式计算出的 k 为 2 的整数倍.再使用“the same padding”保持输出通道和输入通道的一致性.

(3) CBAM: 在 CAM 模块的共享多层感知器中,通道压缩系数采用 8,第一层感知器由于采用线性整流激活函数 ReLu,故采用“he_normal”权重初始化方法,使得正向传播过程中状态值的方差保持不变,反向传播过程中激活值的梯度的方差保持不变,即保证了信息流动过程中的有效性.同时偏置值采用“zeros”,即全零初始化方法.在 SAM 模块的 2D 卷积操作中,卷积核的大小为 7×7 ,同样采用“he_normal”权重初始化方法.

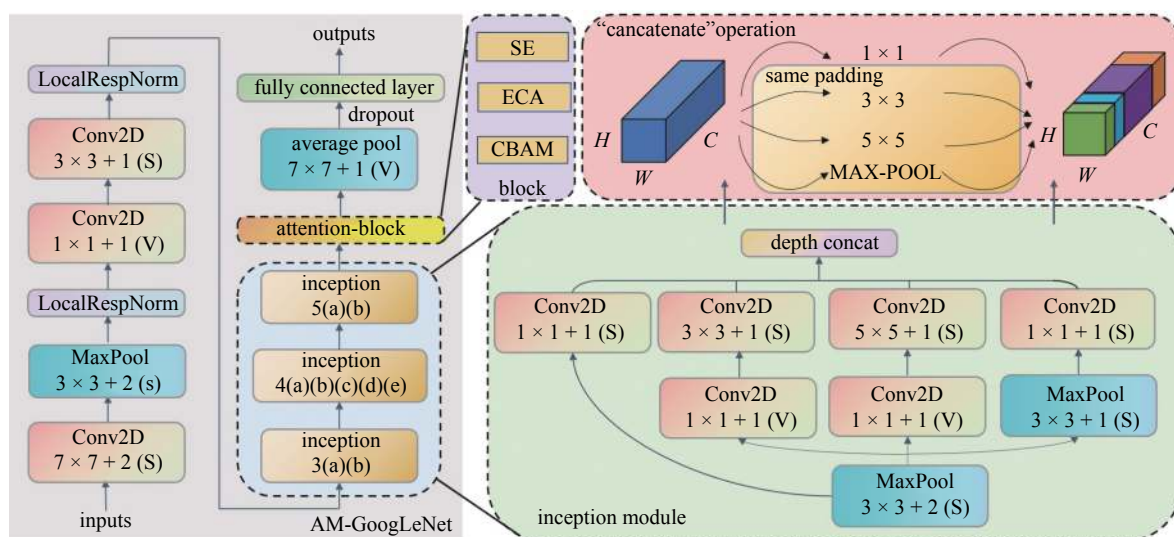


图 4 AM-GoogLeNet 架构

Fig. 4 The architecture of AM-GoogLeNet

3 建立数据集

3.1 混凝土细观有限元模型建立

在我们现有工作的基础上^[26],使用搭建深度学习框架常用的 Python 语言编写建立混凝土细观模型的脚本,导入数值模拟软件 Abaqus 进行二次开

发,实现自动化建模、赋予材料属性、建立分析步、施加载荷和边界条件、网格划分等操作以提高数据集的建立效率.本文采用蒙特卡罗方法进行混凝土二维细观模型建模,生成满足均匀分布的随机变量,并基于富勒曲线的 Walraven 公式计算二维平面中随机骨料的分布, Walraven 公式的具体表达如下

$$P_c(D < D_0) = P_k \left[1.065 \left(\frac{D_0}{D_{\max}} \right)^{0.5} - 0.053 \left(\frac{D_0}{D_{\max}} \right)^4 - 0.012 \left(\frac{D_0}{D_{\max}} \right)^6 - 0.0045 \left(\frac{D_0}{D_{\max}} \right)^8 + 0.0025 \left(\frac{D_0}{D_{\max}} \right)^{10} \right] \quad (2)$$

式中, $P_c(D < D_0)$ 为混凝土样本截面内颗粒粒径 D 小于筛孔 D_0 范围内的概率; P_k 为混凝土样本中粗细骨料体积占总体积的百分比; D_0 为筛孔直径; $D_{\max}$ 为最大骨料粒径。

在该数据集中, 混凝土样本的尺寸为 $150\text{ mm} \times 150\text{ mm}$, 骨料采用三级配, 骨料粒径分别为 6 mm ,

12 mm 和 18 mm , 孔隙直径均为 2 mm , 随机骨料和孔隙均为圆形, 网格尺寸均为 1 mm 。其中, 骨料、砂浆采用三角形平面应变网格。为了提高数值模拟的效率, 采用线弹性模型模拟骨料的力学行为, 砂浆则采用混凝土塑性损伤模型 (concrete damage plasticity)^[27]。其他具体的材料参数如表 1 所示^[28]。最后, 使用“Abacus/Standard”显式模块进行数值模拟计算, 采用准静态单轴压缩实验, 即在试件顶部施加均匀位移载荷, 底部采用固定约束。由脚本自动建模生成的有限元模型如图 5(a) 所示。

表 1 两种细观组分的力学参量

Table 1 Mechanical parameters of the two meso-components							
	Elasticity modulus	Poisson's ratio	Compressive strength	Density	Dilatancy angle	Eccentricity	Stress ratio
	E/GPa	ν	F_c/MPa	$\rho/(\text{t}\cdot\text{m}^{-3})$	$\psi/(\text{^\circ})$	$\eta/\%$	σ_{b0}/σ_{c0}
aggregate	43	0.23	—	2.67	—	—	—
mortar	25	0.2	35	2.40	38	0.1	1.16

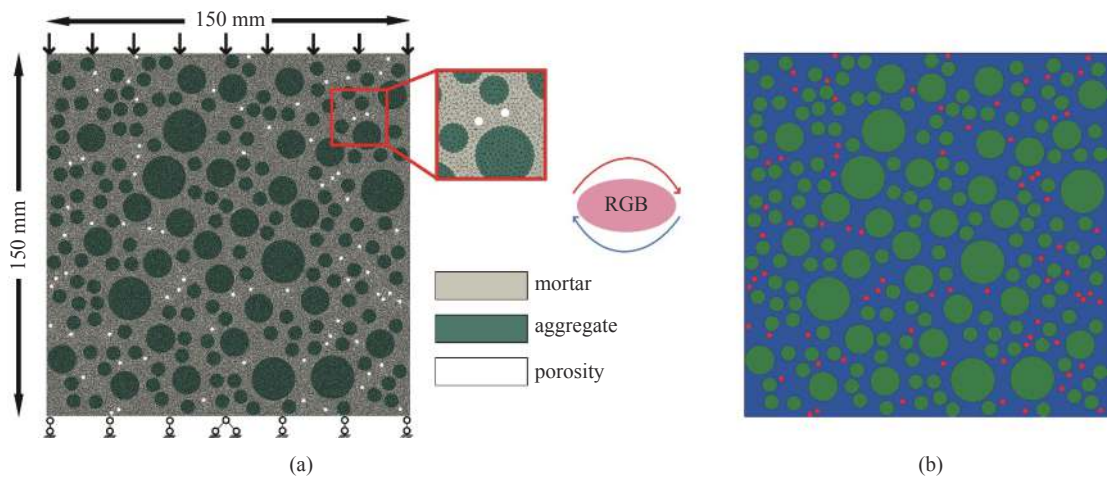


图 5 混凝土 (a) 细观有限元模型和 (b) RGB 格式图像

Fig. 5 (a) The mesoscopic finite element model of concrete and (b) the image in RGB format

3.2 混凝土细观模型图像及应力-应变曲线预处理

数据集中图像的质量直接影响神经网络的训练效果, 直接采用 Abaqus 软件导出的混凝土图像分辨率较低且存在放大图片后细节失真的问题, 所以本文将每次自动建模生成的骨料和孔隙的几何信息进行存储, 导入集成开发环境 PyCharm 中使用 Python 语言进行二次建模, 并且调用 Python 中的“Matplotlib”库绘制图形, 最后将图像格式转换为 RGB, 分辨率

设置为 200 dpi, 重建的 RGB 格式图像如图 5(b) 所示, 其中, 绿色、蓝色、红色分别代表骨料、砂浆、孔隙。图 6 展示了本文最终建立的混凝土细观模型数据集, 骨料体积分数和孔隙率范围分别为 20% ~ 40% 和 0 ~ 10%。此时的图像像素值为 700×700 , 为了减小输入数据的取值范围 (将所有数据映射到同一尺度) 以降低神经网络的学习难度, 在输入神经网络后, 均将维度改为 (224, 224, 3), 然后进行最值归

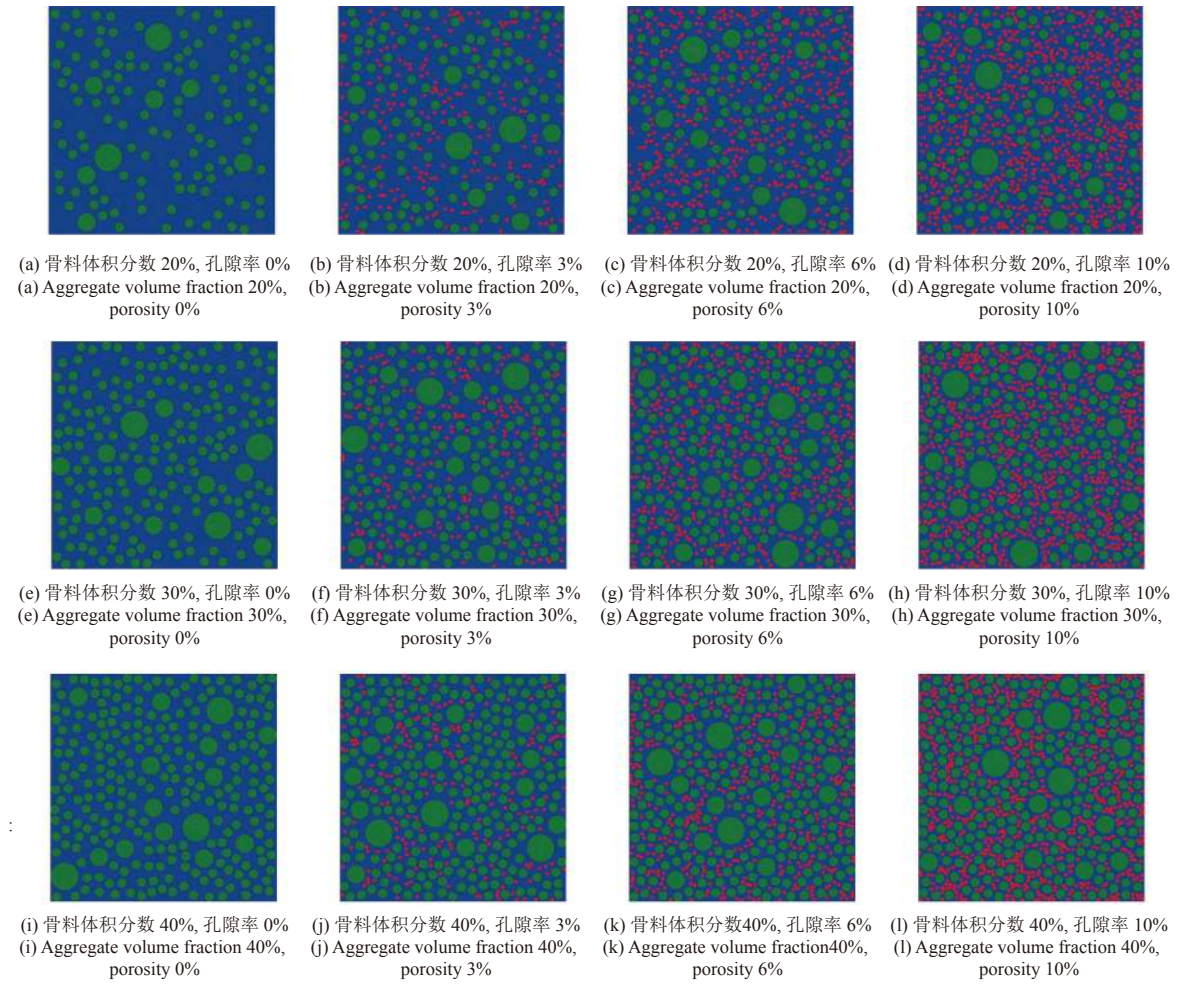


图 6 不同骨料体积分数和孔隙率下的部分数据集

Fig. 6 The partial image dataset with different aggregate volume fraction and porosity

一化处理, 最值归一化的公式如下

$$x_{\text{scale}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

其中, x_{scale} 为归一化后的图像灰度值; x 为归一化前的图像灰度值; $x_{\min}$ 为最小图像灰度值; $x_{\max}$ 为最大图像灰度值。

由数值模拟得到的应力-应变曲线中数据点的规模通常比较大, 直接将高维数组作为输出层会导致神经网络的学习难度增加。故本文将应力-应变曲线转化为 41 维的数组以降低神经网络的参数规模。原始“GoogLeNet”的输出层激活函数为“Sigmoid”函数, 考虑到本文的应力预测值的值域, 将“Sigmoid”函数改为“ReLU”函数。这种做法不但提升了计算效率, 且可以在一定程度上减少梯度消失的概率^[29]。最后, 将平均绝对误差 (MAE) 作为评价指标, 将原始“GoogLeNet”的损失函数改为自定义的均方根误差损失函数 (RMSE), 具体表达式如下

$$RMSE = \sqrt{\frac{\sum_{n=1}^n \sum_{i=1}^{41} (y_{ni} - \hat{y}_{ni})^2}{41n}} \quad (4)$$

$$MAE = \frac{\sum_{n=1}^n \sum_{i=1}^{41} |y_{ni} - \hat{y}_{ni}|}{41n} \quad (5)$$

其中, n 为训练样本个数; i 为应力-应变曲线中的数据点; y_{ni} 为第 n 个训练样本上的第 i 个数据点; $\hat{y}_{ni}$ 为第 n 个训练样本的第 i 个数据点的预测值。

4 神经网络训练

表 2 展示了训练过程所采用的计算机配置信息以及深度学习开发环境。本文在 PyCharm 中使用 Keras 开源框架搭建 AM-GoogLeNet 卷积神经网络, 包括 inception 模块、GoogLeNet 模块、attention 模块、训练集和测试集。

由于不同注意力模块的实际性能往往取决于具

表 2 试验环境的硬件和软件参数

Table 2 Hardware and software parameters of the experimental environment

Name	Parameters
central processing unit	Inter Core i7-11800 H CPU @ 2.3 GHz
memory	DDR4 memory 8 GB
graphics card	NIVIDA GeForce RTX3060
system	Windows 10
environment	Python 3.6 TensorFolw 2.8.0 Keras 2.8.0
compute unified device	NUMPY 1.22.2
architecture	CUDA 11.2

体问题的训练集和参数规模,故本文对三种不同的注意力模块分别进行了测试和对比.用于训练的数据包括训练集 1248 组,验证集 138 组和测试集 462 组.为了提高训练效率,将学习率自动下降机制加入神经网络,即初始学习率为 1×10^{-3} ,在训练过程中每 3 轮准确率不上升就将学习率减小为原来的 1/2,训练轮次设定为 50.图 7 展示了四种不同 GoogLeNet

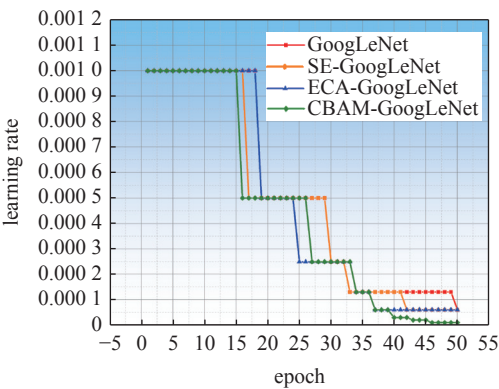


图 7 不同 GoogLeNet 学习率曲线

Fig. 7 The curves of the learning rate of different GoogLeNets

的训练过程学习率下降情况,可以看出 CBAM-GoogLeNet 在训练后期学习率仍处于不断稳定下降的趋势,符合学习率衰减的基本思想,即学习率随着训练的进行逐渐衰减,使得训练结果易于收敛到损失函数的最小值而不是在最小值附近震荡,从而保证了训练的稳定性 and 结果的精度.图 8(a) 和 图 8(b)

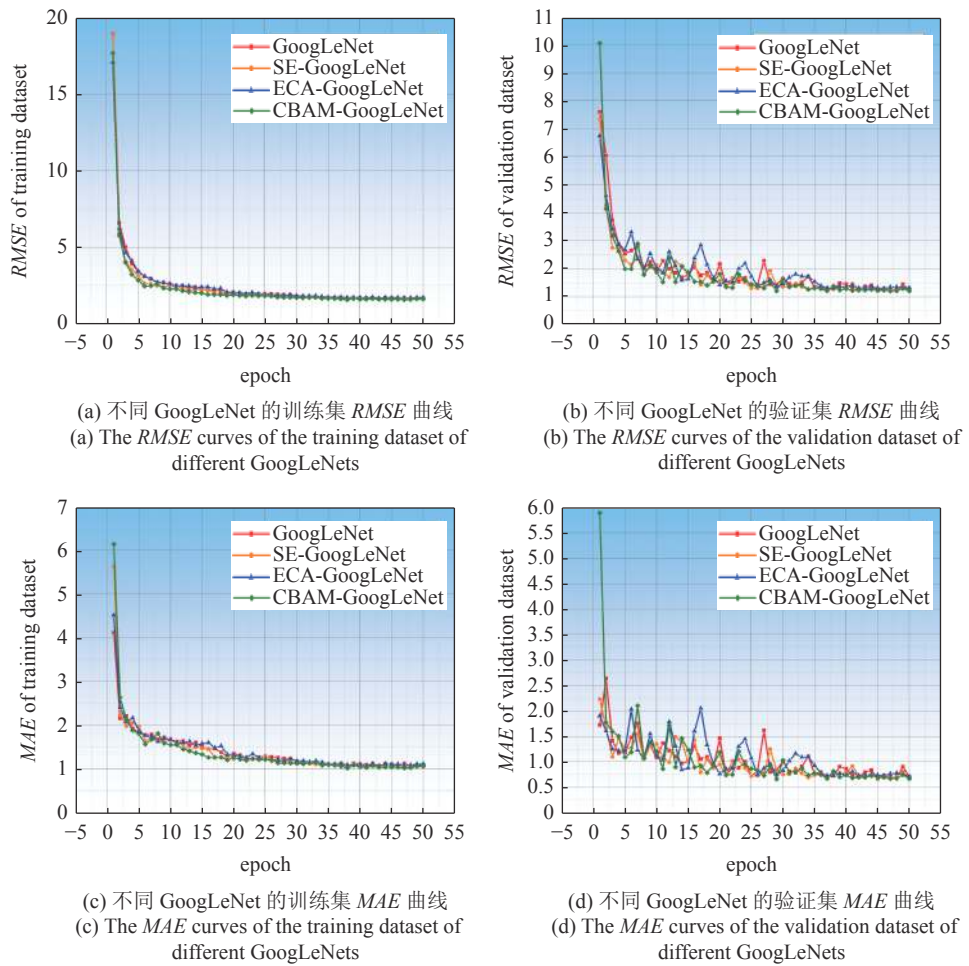


图 8 不同 GoogLeNet 的训练结果对比

Fig. 8 The comparison of the training results of different GoogLeNets

为不同 GoogLeNet 分别在训练集和验证集的 $RMSE$ 曲线, 可以看出, CBAM-GoogLeNet 在训练集曲线的拐点处下降得更快且在验证集曲线中展现了更小的波动幅度. 图 8(c) 和图 8(d) 为不同 GoogLeNet 分别在训练集和验证集的 MAE 曲线, CBAM-GoogLeNet 在训练集曲线的不稳定阶段同样下降得更快且在验证集中具有更稳定的训练效果.

表 3 展示了四种不同 GoogLeNet 训练过程的参数统计, 可以看出, SE-GoogLeNet 的验证集损失虽

然较低, 但是由于注意力模块的参数较多, 训练时间增加明显. ECA-GoogLeNet 由于共享参数较多, 虽然几乎没有提升神经网络的规模, 但是验证集损失反而提升. CBAM-GoogLeNet 同时具有通道注意力机制和空间注意力机制, 相比于只具有通道注意力机制的 SE-GoogLeNet 和 ECA-GoogLeNet 可以更好地捕捉骨料和孔隙的空间分布和对应力-应变曲线影响的权重, 在没有显著增加参数规模和训练时间的情况下, 得到了最小的验证集损失.

表 3 不同 GoogLeNet 的训练过程参数

Table 3 The training process parameters of different GoogLeNet

CNN	Parameters of AM	Total parameters	Model size/MB	Epoch	Time/s	Validation loss
GoogLeNet	—	6015 577	22.95	50	6254	1.245 4
SE-GoogLeNet	8352	6541 145	24.95	50	6466	1.207 0
ECA-GoogLeNet	5	6015 582	22.95	50	6321	1.325 0
CBAM-GoogLeNet	4194	6277 819	23.95	50	6280	1.205 3

图 9(a)~图 9(f) 展示了四种 GoogLeNet 在不同混凝土骨料体积分数和孔隙率下的应力-应变曲线预测结果对比. 结果表明, CBAM-GoogLeNet 在应力-应变曲线的线弹性阶段和峰值应力都有着较高的预测精度. 图 10(a) 和图 10(b) 分别为 CBAM-GoogLeNet 在测试集 1~231 组和 232~462 组的应力-应变曲线峰值应力预测结果. 采用平均相对误差 (RE) 定量评估峰值应力的预测误差, 具体计算公式如下

$$RE = \frac{\sum_{i=1}^n \frac{|\sigma_{\max} - \hat{\sigma}_{\max}|}{\sigma_{\max}}}{n} \quad (6)$$

其中, n 为测试集的样本数量; $\sigma_{\max}$ 为真实峰值应力; $\hat{\sigma}_{\max}$ 为预测峰值应力.

经过计算, CBAM-GoogLeNet 在 462 组测试集上的峰值应力预测平均相对误差为 3.01%. 综上所述, 本文最终选取 CBAM-GoogLeNet 的预测结果作

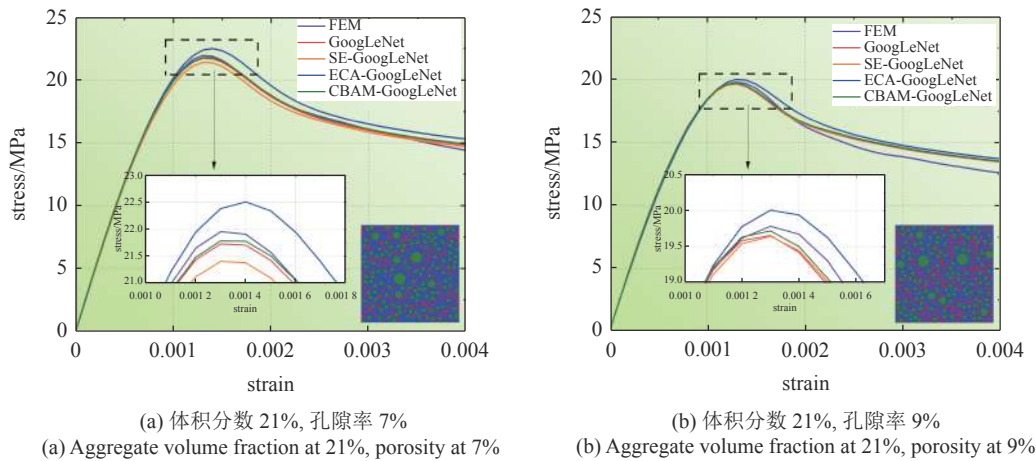


图 9 不同骨料体积分数和孔隙率下的 CBAM-GoogLeNet 预测结果

Fig. 9 The prediction results of CBAM-GoogLeNet with different aggregate volume fraction and porosity

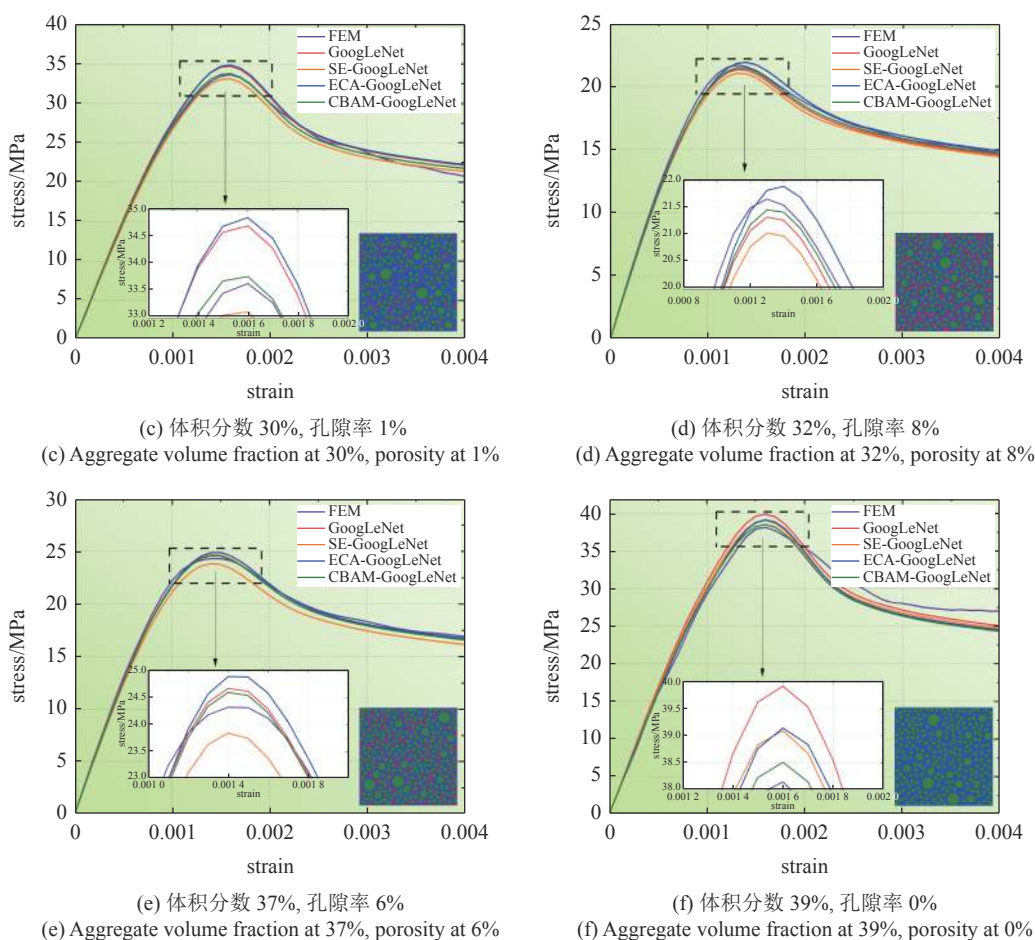


图 9 不同骨料体积分数和孔隙率下的 CBAM-GoogLeNet 预测结果 (续)

Fig. 9 The prediction results of CBAM-GoogLeNet with different aggregate volume fraction and porosity (continued)

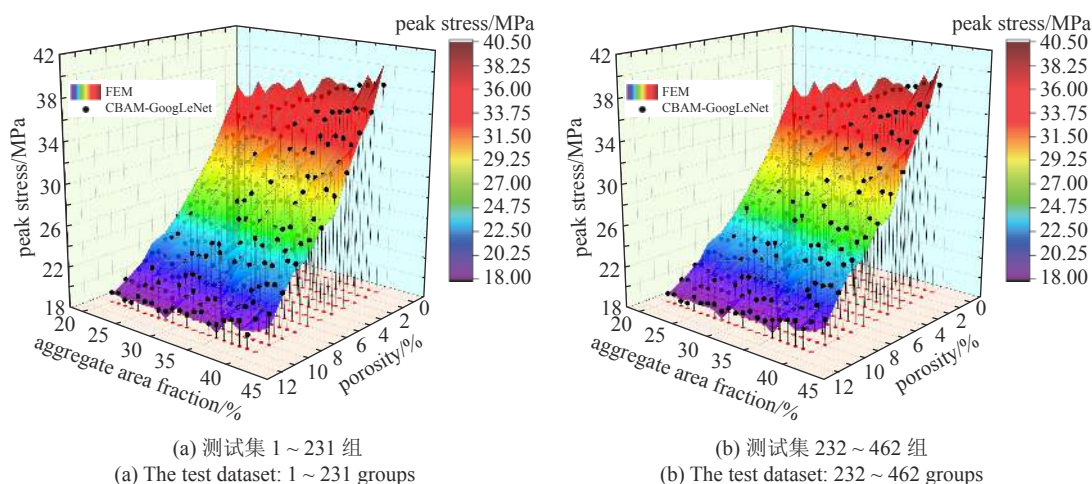


图 10 CBAM-GoogLeNet 测试集峰值应力预测结果

Fig. 10 The prediction results of the peak stress of the CBAM-GoogLeNet

为 BP 神经网络的训练集,同时提取预测应力-应变曲线中的重要参数峰值应力作为下一部分的主要研究对象。

5 物理参数的引入

在 AM-GoogLeNet 的训练过程中,通过进一步增大数据集等方法对提升预测精度的“收益”变得很

低. 因此本文采用两种将物理参数引入神经网络的方法进一步提升峰值应力的预测精度, 并对结果进行了分析和对比.

5.1 将物理参数引入 CBAM-GoogLeNet

首先将每个混凝土样本的骨料体积分数和孔隙率加入 CBAM-GoogLeNet 输入图像的特征层, 见图 11(a). 输入特征层共包括两个部分, 即几何特征层和物理参数层. 将原训练方案中几何特征层的高度由 224 压缩至 222, 同时加入两个物理参数层, 维度均为 $1 \times 224 \times 3$, 每层包括一个物理参数 (骨料体积分数、孔隙率) 和填充参数 0^[30]. 同时, 为了保证

输入特征层所有数据均为同一尺度 (即图像灰度值范围: 0 ~ 255), 以提高神经网络的收敛性^[31], 需要将物理参数 (例如骨料体积分数 20%, 孔隙率 3%) 放大 100 倍. 最后使用 CBAM-GoogLeNet 进行训练, 预测混凝土应力-应变曲线的峰值应力. 用于训练的数据包括训练集 1248 组, 验证集 138 组和测试集 462 组, 神经网络采用的参数与训练环境同第 4 节.

训练过程中训练集和验证集的 *RMSE* 和 *MAE* 曲线如图 11(b) 所示, 可以看出损失曲线展现了较好的收敛性, 其中验证集的 *RMSE* 在最优轮次下的训练值仅为 0.8978, 证明本文采用的将物理参数直接加入卷积神经网络输入层的方法具有可行性.

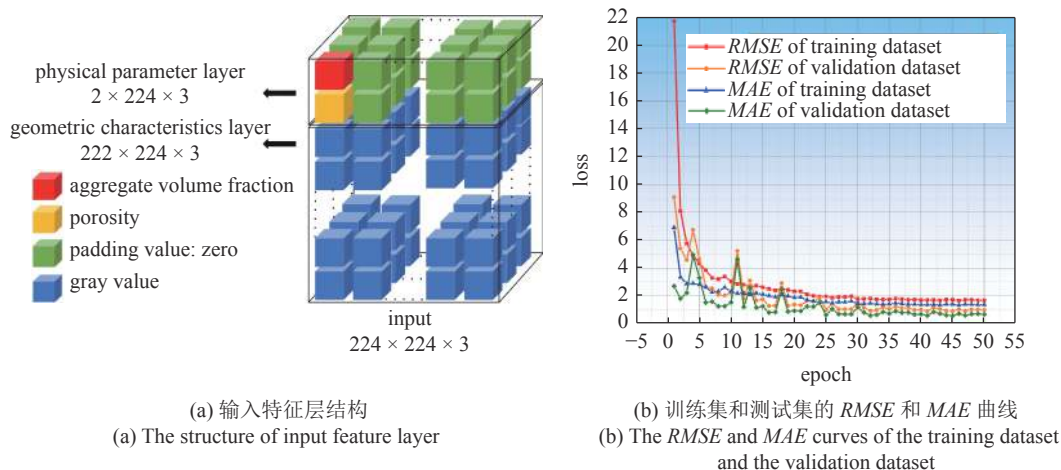


图 11 引入物理参数的 CBAM-GoogLeNet 训练过程

Fig. 11 The training process of the physics CBAM-GoogLeNet

5.2 BP 神经网络

将 CBAM-GoogLeNet 的峰值应力预测值以及对应的骨料体积分数和孔隙率作为输入数据, 数值模拟的峰值应力结果作为输出数据建立 BP 神经网络^[32-33]的训练集, 即采用二次训练的方式将 AM-GoogLeNet 和 BP 神经网络通过两个物理参数 (骨料体积分数、孔隙率) 结合成一个整体, 从而进一步提升预测精度.

本文使用的 BP 神经网络架构如图 12 所示, CBAM-GoogLeNet 的峰值应力预测值、对应的骨料体积分数和孔隙率作为 x_1 , x_2 和 x_3 输入, 隐藏层包括 6 个神经元, 鉴于数据集较小, 为了防止过拟合和具有更快的收敛速度, 采用“PReLU”激活函数, 输出为峰值应力. 采用“Adam”优化器, *MAE* 作为评价指标. 设置训练集 231 组, 测试集 231 组, 训练轮次设置为 50, “batch_size”设置为 2. 在训练末期, *MAE* 逐

渐趋于稳定, 在最后一轮训练的 *MAE* 为 0.647 53. 图 12 展示了 BP 神经网络测试集前 20 组的预测结果对比. 由此可见, 通过将物理参量引入神经网络有效提高了对混凝土二维细观模型力学性能的预测精度.

仍采用 *RE* 定量评估引入物理参数的 CBAM-GoogLeNet 方法和使用 BP 神经网络进行二次训练方法的预测误差. 经过计算, CBAM-GoogLeNet 在 231 组测试集的 *RE* 为 3.02%, 引入物理参数的 CBAM-GoogLeNet 的 *RE* 为 2.45%, BP 神经网络的 *RE* 为 2.41%. 两种方法在本文的研究问题中展现了相近的提升预测精度的能力. 引入物理参数的 CBAM-GoogLeNet 的训练方法仅需要卷积神经网络, 但是物理参量的引入使图像几何特征层的参数规模被压缩, 导致预测精度下降^[19,34]. 而本文采用的联合数据驱动方法在 AM-GoogLeNet 的输入部分无

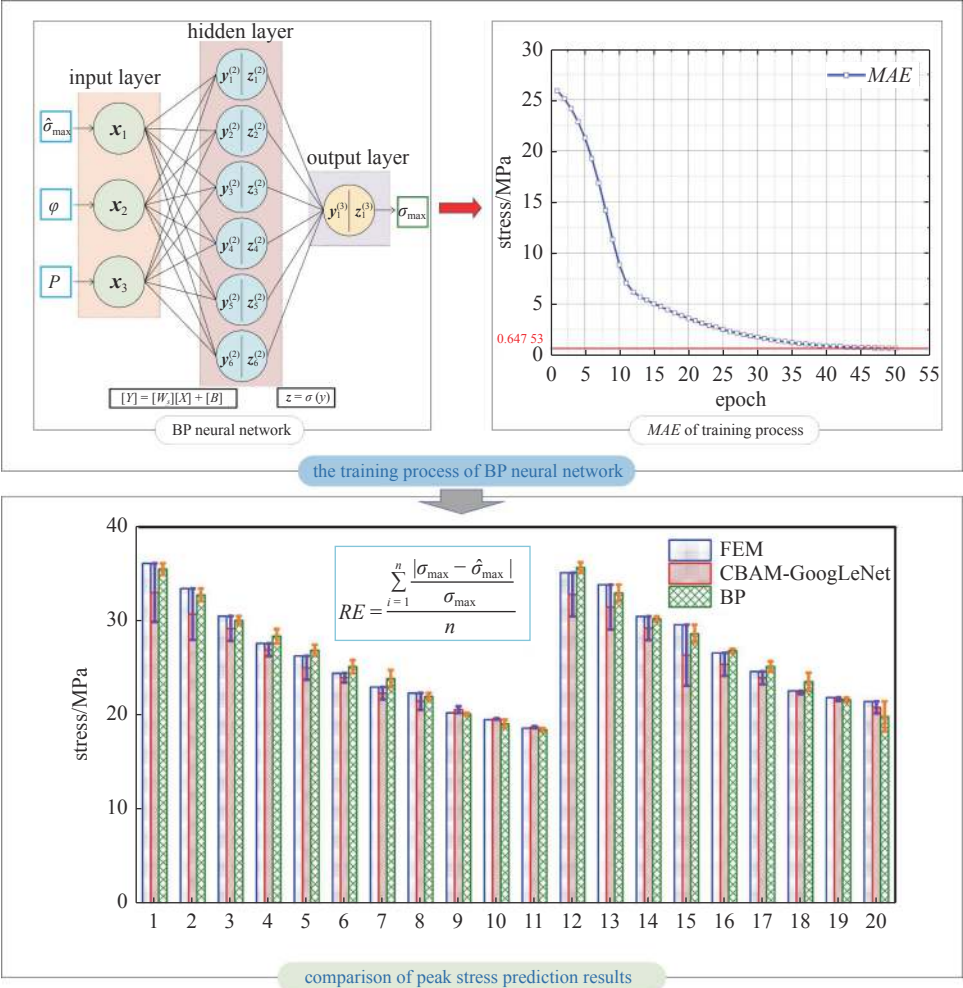


图 12 BP 神经网络架构、训练过程和结果评估

Fig. 12 The architecture, training process and results evaluation of the BP neural network

需加入物理参数,提高了数据集的制作效率,降低了神经网络的训练难度.同时采用BP神经网络进行二次训练便于提取神经网络的权重值以进行后续的参数分析.

Aleboyeh 等^[35]提出了利用神经网络的权重值

计算输入参量对输出变量影响大小的公式,即式(7).为了进一步定量分析骨料体积分数和孔隙率对峰值应力的影响,本文首先提取了BP神经网络最后一轮的训练权重和偏置值,如表4所示,然后利用式(7)进行计算.计算结果表明, $I_1 = 2.2831$, $I_2 = 3.31722$,

表 4 BP 神经网络的权重和偏置值

Table 4 The weights and biases of BP neural network

Neurons	W_{input}			Bias	W_{output}	Bias
	Peak stress	Volume fraction	Porosity			
1	-0.00172	-0.12354	0.59741	-2.15972	-2.69027	1.21105
2	-0.17380	1.08733	-0.53564	1.61095	2.00213	
3	-1.36082	0.33879	0.01926	1.44915	1.63320	
4	0.33198	-1.49304	0.67236	1.67449	1.40409	
5	0.03562	-0.19979	-1.14093	1.75658	1.35286	
6	1.12495	-0.56972	-1.04270	1.19552	1.38380	

$I_3 = 4.86572$, 故骨料体积分数和孔隙率对峰值应力影响的权重分别为 40.54% 和 59.46%

$$I_j = \frac{\sum_{m=1}^{N_h} \left[\left(\frac{|w_{jm}^{ih}|}{\sum_{k=1}^{N_i} |w_{km}^{ih}|} \right) |w_{mn}^{ho}| \right]}{\sum_{k=1}^{N_i} \left[\sum_{m=1}^{N_h} \left(\frac{|w_{km}^{ih}|}{\sum_{k=1}^{N_i} |w_{km}^{ih}|} \right) |w_{mn}^{ho}| \right]} \quad (7)$$

其中, I_j 是第 j 个输入变量对输出变量的相对重要性, N_i 和 N_h 分别是输入和隐藏神经元的个数, w_s 是连接权重, 上标“ i ”, “ h ”, “ o ”分别表示输入层、隐藏层和输出层, 下标“ k ”, “ m ”, “ n ”分别指输入、隐藏和输出神经元。

6 结论

本文以混凝土这种具有代表性的复合材料为研究对象, 通过“AM-GoogLeNet + BP”联合数据驱动的方法预测其压缩应力-应变曲线, 研究表明, 引入注意力机制后的“CBAM-GoogLeNet”和传统“GoogLeNet”预测方法的均方根误差分别为 1.21 和 1.25。CBAM-GoogLeNet 在提升预测精度的同时, 具有更好的泛化能力, 即对于不同的骨料体积分数和孔隙率均有较好的预测结果。将 CBAM-GoogLeNet 的峰值应力预测结果与其相应的骨料体积分数及孔隙率共同引入 BP 神经网络进行训练, 结果表明得到的峰值应力平均相对误差由 3.02% 降低为 2.41%。通过提取 BP 神经网络的权重参数可以得出: 在本文研究范围内, 骨料体积分数和孔隙率对峰值应力影响的权重分别为 40.54% 和 59.46%。本文所采用方法在提升预测效率的同时兼顾了物理意义层面的准确度, 提升了神经网络的物理可解释性, 为复合材料宏观跨尺度力学行为的进一步研究提供了思路。

参 考 文 献

- 1 Tian ZY, Yan Y, Li J, et al. Progressive damage and failure analysis of three-dimensional braided composites subjected to biaxial tension and compression. *Composite Structures*, 2018, 185: 496-507
- 2 Yen CF, Kaste B, Chen CCT, et al. Modeling and simulation of carbon composite ballistic and blast behavior. *Journal of Composite Materials*, 2020, 54(4): 485-499
- 3 金浏, 李健, 余文轩等. 混凝土动态双轴拉压破坏准则细观数值模拟研究. *力学学报*, 2022, 54(3): 800-809 (Jin Liu, Li Jian, Yu Wenxuan, et al. Mesoscopic numerical Simulation on dynamic biaxial tension compression failure criterion of concrete. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, 54(3): 800-809 (in Chinese))
- 4 Zhang J, Wang ZY, Yang HW, et al. 3D meso-scale modeling of reinforcement concrete with high volume fraction of randomly distributed aggregates. *Construction and Building Materials*, 2018, 164: 350-361
- 5 Guo L, Wu J, Li JH. Complexity at Mesoscales: A common challenge in developing artificial intelligence. *Engineering*, 2019, 5: 924-929
- 6 Li X, Liu ZL, Cui SQ, et al. Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. *Computer Methods in Applied Mechanics and Engineering*, 2019, 347: 735-753
- 7 瞿同明, 冯云田, 王孟琦等. 基于深度学习和细观力学的颗粒材料本构关系研究. *力学学报*, 2021, 53(9): 2404-2415 (Qu Tongming, Feng Yuntian, Wang Mengqi, et al. Constitutive relations of granular materials by integrating micromechanical knowledge with deep learning. *Chinese Journal of Theoretical and Applied Mechanics*, 2021, 53(9): 2404-2415 (in Chinese))
- 8 Yang C, Kim Y, Ryu S, et al. Prediction of composite microstructure stress-strain curves using convolutional neural networks. *Materials & Design*, 2020, 189: 108509
- 9 Seolhyun N. Analysis of gradient vanishing of RNNs and performance comparison. *Information*, 2021, 12(11): 442
- 10 梁雪慧, 程云译, 张瑞杰等. 基于卷积神经网络的桥梁裂缝识别和测量方法. *计算机应用*, 2020, 40(4): 1056-1061 (Liang Xuehui, Cheng Yunyi, Zhang Ruijie, et al. Bridge crack classification and measurement method based on deep convolutional neural network. *Journal of Computer Applications*, 2020, 40(4): 1056-1061 (in Chinese))
- 11 Chaudhari S, Polatkan G, Ramanath R, Mithal V. An attentive survey of attention models//IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2019: 1-33
- 12 Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention//IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2014: 1-12
- 13 王晓玲, 李克, 张宗亮等. 耦合 ALO-LSTM 和特征注意力机制的土石坝渗压预测模型. *水利学报*, 2022, 53(4): 403-412 (Wang Xiaoling, Li Ke, Zhang Zongliang, et al. Coupled ALO-LSTM and feature attention mechanism prediction model for seepage pressure of earth-rock dam. *Journal of Hydraulic Engineering*, 2022, 53(4): 403-412 (in Chinese))
- 14 Ye S, Li B, Li QY, et al. Deep neural network method for predicting the mechanical properties of composites. *Applied Physics Letters*, 2019, 115(16): 161901
- 15 赵地, 赵莉芝, 甘永进等. 基于支撑先验与深度图像先验的无预训练磁共振图像重建方法. *物理学报*, 2022, 71(5): 350-362 (Zhao Di, Zhao Lizhi, Gan Yongjin, et al. Undersampled magnetic resonance image reconstruction based on support prior and deep image prior without pre-training. *Acta Physica Sinica*, 2022, 71(5): 350-362 (in Chinese))
- 16 Hu ZL, Zhao Q, Wang J. The prediction model of cotton yarn intensity based on the CNN-BP neural network. *Wireless Personal Communications*, 2018, 102(2): 1905-1916
- 17 Bahdanau D, Cho K, Bengio Y. Neural Machine translation by jointly learning to align and translate//IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2014: 1-15
- 18 Hu J, Shen L, Samuel A, et al. Squeeze-and-excitation networks.

- IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 42(8): 1-13
- 19 Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 1-12
- 20 Woo S, Park J, Lee JY, et al. CBAM: Convolutional block attention module//European Conference on Computer Vision, 2018: 1-17
- 21 王溢琴, 董云云, 刘慧玲. 基于 GoogLeNet 和空间谱变换的高光谱图像超分辨率方法. *光学技术*, 2022, 48(1): 93-101 (Wang Yingqin, Dong Yunyun, Liu Huiling. Super-resolution method of hyperspectral image based on GoogLeNet and spatial spectrum transformation. *Optical Technique*, 2022, 48(1): 93-101 (in Chinese))
- 22 周杰, 赵婷婷, 陈青青等. 基于 GoogLeNet 的混凝土细观模型应力-应变曲线预测. *应用数学和力学*, 2022, 43(3): 290-299 (Zhou Jie, Zhao Tingting, Chen Qingqing, et al. Prediction of concrete meso-model stress-strain curves based on GoogLeNet. *Applied Mathematics and Mechanics*, 2022, 43(3): 290-299 (in Chinese))
- 23 狄少丞, 冯云田, 瞿同明等. 基于深度学习和细观力学的颗粒材料本构关系研究. *力学学报*, 2021, 53(10): 2712-2723 (Di Shaocheng, Feng Yuntian, Qu Tongming, et al. Data-driven stress-strain modeling for granular materials through deep reinforcement learning. *Chinese Journal of Theoretical and Applied Mechanics*, 2021, 53(10): 2712-2723 (in Chinese))
- 24 Efe GD, Zohar R, Sebastian HD, et al. Symmetries and phase diagrams with real-space mutual information neural estimation. *Physical Review E*, 2021, 104(6-1): 064106
- 25 胡丹丹, 张忠婷, 牛国臣. 融合 CBAM 注意力机制与可变形卷积的车道线检测. *北京航空航天大学学报*, 2022, doi: [10.13700/j.bh.1001-5965.2022.0601](https://doi.org/10.13700/j.bh.1001-5965.2022.0601) (Hu Dandan, Zhang Zhongting, Niu Guochen. Lane line detection incorporating CBAM attention mechanism and deformable convolution. *Journal of Beijing University of Aeronautics and Astronautics*, 2022, doi: [10.13700/j.bh.1001-5965.2022.0601](https://doi.org/10.13700/j.bh.1001-5965.2022.0601) (in Chinese))
- 26 Zhang YH, Chen QQ, Wang ZY, et al. 3D mesoscale fracture analysis of concrete under complex loading. *Engineering Fracture Mechanics*, 2019, 220: 106646
- 27 Li BB, Jiang JF, Xiong HB, et al. Improved concrete plastic-damage model for FRP-confined concrete based on true tri-axial experiment. *Composite Structures*, 2021, 269: 114051
- 28 陈青青, 张煜航, 张杰等. 含孔隙混凝土二维细观模型建模方法研究. *应用数学和力学*, 2020, 41(2): 182-194 (Chen Qingqing, Zhang Yuhang, Zhang Jie, et al. Study on a 2D mesoscopic modeling method for concrete with voids. *Applied Mathematics and Mechanics*, 2020, 41(2): 182-194 (in Chinese))
- 29 Krizhevsky A, Sutskever L, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84-90
- 30 姜克杰, 胡松, 韩强. 基于长短期记忆网络的 FRP 约束混凝土圆柱循环轴压应力应变预测模型. *工程力学*, 2022, 39: 1-15 (Jiang Kejie, Hu Song, Han Qiang. Cyclic axial compressive stress-strain prediction model for FRP-constrained concrete cylinder based on long short-term memory networks. *Engineering Mechanics*, 2022, 39: 1-15 (in Chinese))
- 31 Loffe S, Szegedy C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift//IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2015: 1-8
- 32 马闯, 杨晓龙, 陈含爽等. 基于平均场近似的 BP 算法求解随机块模型. *物理学报*, 2021, 70(22): 345-356 (Ma Chuang, Yang Xiaolong, Chen Hanshuang, et al. A mean-field approximation based BP algorithm for solving the stochastic block model. *Acta Physica Sinica*, 2021, 70(22): 345-356 (in Chinese))
- 33 汪恩良, 田雨, 刘兴超等. 基于 WOA-BP 神经网络的超低温冻土抗压强度预测模型研究. *力学学报*, 2022, 54(4): 1145-1153 (Wang Enliang, Tian Yu, Liu Xingchao, et al. Prediction model of compressive strength of ultra low temperature frozen soil based on WOA-BP neural network. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, 54(4): 1145-1153 (in Chinese))
- 34 陈苏, 丁毅, 孙浩等. 物理驱动深度学习波动数值模拟方法及应用. *力学学报*, 2023, 55(1): 272-282 (Chen Su, Ding Yi, Sun Hao, et al. Methods and applications of physical information deep learning in wave numerical simulation. *Chinese Journal of Theoretical and Applied Mechanics*, 2023, 55(1): 272-282 (in Chinese))
- 35 Aleboye A, Kasiri MB, Olya ME, et al. Prediction of azo dye decolorization by UV/H₂O₂ using artificial neural networks. *Dyes and Pigments*, 2008, 77(2): 288-294