

光学学报

基于深度学习的散斑图像大变形测量方法

萧红, 李成南, 冯明驰*

重庆邮电大学先进制造工程学院, 重庆 400065

摘要 本文提出了一种针对复杂大变形的散斑图像位移场的测量方法,该方法利用注意力机制和深度可分离卷积改进一个现有的卷积神经网络模型用于测量大变形位移场。为训练该模型,构建了符合真实情况的包含多种类型的散斑图像和复杂大变形位移场的数据集并提出了新的损失函数。该方法与传统数字图像相关方法以及最新的深度学习方法分别在自建数据集和公开数据集上进行了对比实验和结果分析。结果表明,所提方法在模型参数量最小的情况下取得了最高的平均精度,位移场计算速度也远超传统方法,能够满足大变形位移场的实际实时测量需求。

关键词 测量; 数字图像相关; 大变形位移场; 深度学习; 位移场测量

中图分类号 O348.1

文献标志码 A

DOI: 10.3788/AOS222084

1 引言

数字图像相关(DIC)法是一种基于可见光的非接触式全局位移应变测量方法,仅依靠相机即可完成对被测物体全局位移场的测量,相比于其他方法,其具有测量设备成本低、能够实现全局测量等优点,使得其在实验力学领域应用广泛。传统DIC算法通过划分子区和建立相关性准则来计算参考图像变化到目标图像的位移场。自DIC法在20世纪80年代被提出^[1-3]以来,众多学者对DIC算法的性能、精度、稳定性等进行了改进,拓展了其应用范围^[4-8]。目前最先进的DIC算法已经被应用于热变形测量^[9]、物体位姿测量^[10]、水下变形测量^[11]等领域。但是在实际应用中,实时性是一个关键要素,但是到目前为止,传统DIC方法无法满足实时测量要求。

Bruck等^[12]在1989年提出亚像素匹配方法,并采用正向增量牛顿拉夫逊(FA-NR)算法来计算位移场,提高了数字图像相关法的测量精度,极大地推动了DIC算法的发展。但是该方法计算效率较低,因此为了提高DIC的计算效率,Baker等^[13-14]提出了反向组合高斯牛顿迭代(IC-GN)算法来进行图像子区配准。IC-GN算法与FA-NR算法具有等效的测量精度,但是其计算速度是FA-NR算法的3~6倍,是目前主流的DIC亚像素匹配方法,采用该方法的开源软件有Ncorr^[15]、OpenCorr^[16]等。针对DIC算法适合并行计算的特点,Zhang等^[17]利用图形处理器(GPU)设备来加速OpenCorr的计算过程,每秒可以计算超过十万个感

兴趣像素点,但是该方法仍然无法达到实时运行的要求。

随着深度学习在计算机视觉领域中的飞速发展,深度学习方法逐渐开始被应用于DIC领域。得益于可以利用GPU设备进行高效计算,基于深度学习的散斑位移场测量方法可以更加容易地实现实时在线计算,在需要实时测量的应用中更具优势。刘小勇等^[18]和Min等^[19]开始采用深度学习来代替传统DIC法中的部分步骤,但是都存在计算精度和效率低的问题。随后,端到端的深度学习模型也逐渐被提出用于测量散斑位移场。Boukhtache等^[20]借鉴了基于深度学习的光流法,采用计算机模拟的散斑图像数据集重新训练了和微调了多个现有的深度学习光流法框架并提出了StrainNet,获得了与传统DIC方法相当的位移场测量精度。但是该方法使用的数据集只包含并不符合实际情况的亚像素级别的随机位移,导致该网络只能测量小于一个像素的位移场,不符合实际的测量需求。Yang等^[21]提出了一种基于深度学习的DIC方法DeepDIC,它包括了两个独立的卷积神经网络来分别进行端到端的位移场和应变场测量。DeepDIC的测量结果与商业DIC软件高度一致,但是该方法忽略了位移场和应变场之间的内在联系,增大了网络的计算量。DeepDIC使用的数据集只包含模拟散斑,缺少真实拍摄的散斑图像,数据集中的位移场绝大部分都是仅有几个像素的小位移,这会影响DeepDIC在真实应用场景下的泛化性能。黄举等^[22]提出了一种具有通道拼接结构的卷积神经网络来实现散斑图像位移场测

收稿日期: 2022-12-02; 修回日期: 2023-02-17; 录用日期: 2023-03-20; 网络首发日期: 2023-04-21

基金项目: 国家自然科学基金(51505054, 51605064)

通信作者: *fengmc@cqupt.edu.cn

量,网络模型使用了一个具有三种变形模式的模拟散斑数据集来进行训练。相比于传统方法,该方法具有更高的精度和更快的速度。但是该方法使用的数据集变形模式较为单一且缺乏大位移场样本,导致网络模型不适用存在复杂变形的较大位移场的情况。以上提到的端到端的深度学习模型运行速度远超传统方法,但是由于数据集不够全面,模型无法准确测量在实际应用中的复杂的大变形位移场。

综上所述,现阶段仍缺少计算速度快且能够测量符合实际情况的复杂大变形位移场的有效方法。为了解决以上问题,本文提出了一种基于UNet^[23]的带注意力机制^[24]和深度可分离卷积^[25]的散斑图像位移场测量方法。该方法具有以下优势:1)自建了符合实际情况的具有大变形位移场的散斑图像数据集;2)改进了UNet网络使得模型能够准确测量复杂的大位移场;3)建立了新的损失函数,加速了模型训练过程,提高了模型精度;4)位移场测量网络模型参数量小且计算速度快。实验结果表明,本文提出的方法不仅在自建数据集上的测量精度高于传统DIC方法,在公开的DIC挑战数据集^[26]上同样取得了较好的结果,并且在计算速度上远超传统方法。本研究的源代码、网络预训练权重开源地址为<https://github.com/donotbreeze/Large-deformation-measurement-method-of-speckle-image-based-on-deep-learning>,数据集开源地址为https://pan.baidu.com/s/1KzC9g_GIkvMnGFumDYGyBA?pwd=fd5x。

2 散斑图像大变形位移场测量方法

2.1 符合实际情况的具有大变形位移场的散斑图像数据集

数据集的质量、规模、多样性直接影响着网络模型

的准确性和泛化性能。在散斑位移场测量领域,散斑位移场图像数据集中的一个样本包括变形前的参考图像 I_{ref} ,变形后的目标图像 I_{tar} ,以及从 I_{ref} 变化到 I_{tar} 所产生的 x 和 y 方向的位移场 d_x 和 d_y 。由于现阶段缺乏能够准确测量全局位移场的设备,无法通过真实实验来获取位移场真值。此外,深度学习模型训练需要大量的样本,而通过真实实验来获取大量散斑图像效率极低且成本极高。因此,本文通过多种方式来获取大量散斑图像,并自定义了具有多种变形模式的大变形位移场,从而构建出可靠的具有多样性和大变形位移场的散斑图像数据集。

本文通过多种方式获得了1600张散斑图像用于制作数据集,每张图像尺寸为512 pixel×512 pixel,按照散斑图像的特点及其作用可以分为四大类,如图1所示。I类散斑图像共200张,此类图像是将众多DIC相关论文提到的散斑图像经过缩放、裁剪等处理后得到的。这些散斑图像大部分为真实实验拍摄得到且种类丰富,用于增加数据集的真实性。II类散斑图像共800张,此类图像是利用文献^[27]提出的计算机模拟方法产生的。该方法计算量大,但生成的散斑图像与真实散斑图像具有高度的相似性,用于代替真实散斑图像,降低数据集制作成本。III类散斑图像共200张,此类图像在白(黑)色的背景上生成大量随机的黑(白)色小椭圆散斑并添加高斯噪声。此类图像生成方法简单、计算量小,但与真实散斑存在一定差异,用于增强网络模型的泛化性能。IV类散斑图像共400张,生成方法与II类相同,但是参数不同,用于生成测试集样本。II类和IV类散斑图像生成参数如表1所示,其中 $R_{speckle}$ 表示散斑半径, $N_{speckle}$ 表示散斑数量, T 表示对比度, N_{image} 表示生成的散斑图像数量, $T_{speckle}$ 代

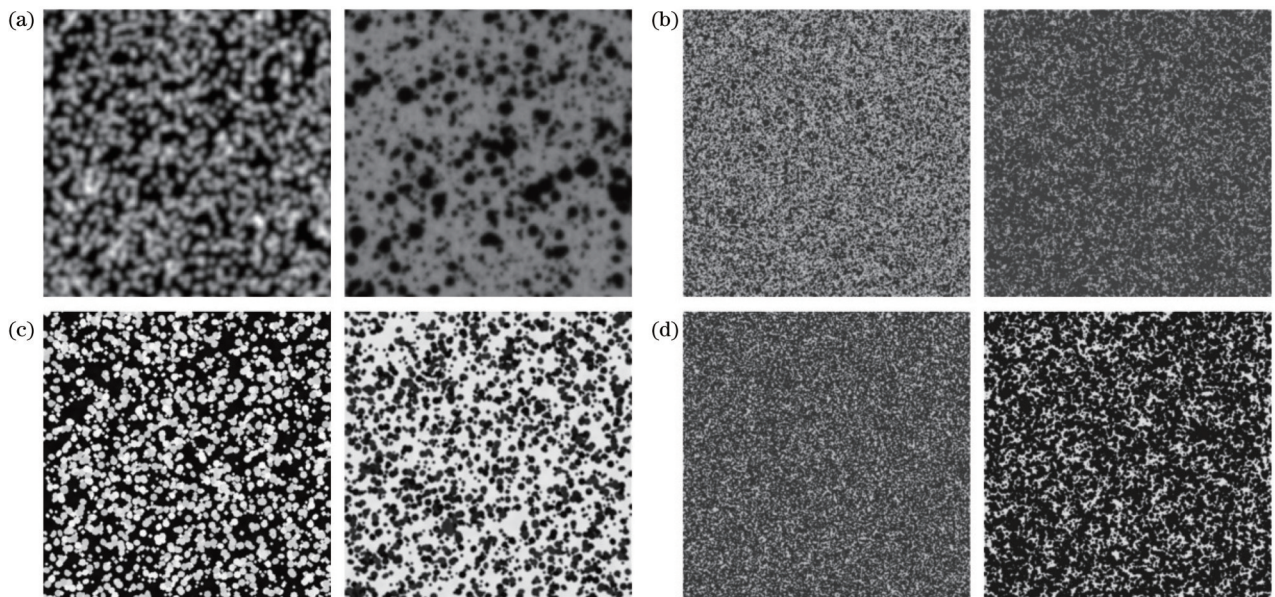


图1 四类典型散斑图像。(a) I类散斑图;(b) II类散斑图;(c) III类散斑图;(d) IV类散斑图

Fig. 1 Four kinds of typical speckle images. (a) Speckle image I; (b) speckle image II; (c) speckle image III; (d) speckle image IV

表散斑类型, E、P、U、L 分别代表指数分布散斑、均匀分布散斑、泊松分布散斑、正态分布散斑。

表 1 II、IV 类散斑图像生成参数
Table 1 Generation parameters of class II and IV speckle images

Class	$R_{\text{speckle}} / \text{pixel}$	N_{speckle}	Γ	T_{speckle}	N_{image}
II	0.5	150000	0.6	E	200
	0.7	150000	0.5	E	200
	1	10000	0.5	P	200
	3	6000	0.6	P	200
IV	0.7	150000	0.5	U	100
	1.5	50000	0.6	L	100
	1	50000	0.7	E	100
	2	20000	0.8	P	100

为了能够获得真实的位移场数据, 本文定义多种类型的随机大变形位移场 d_x 、 d_y 并将位移场施加在参考散斑图像 I_{ref} 上从而生成目标图像 I_{tar} 。生成目标图像的基本思路是假设其亮度不变, 对于特定的像素点 (i, j) 以及其 x 和 y 方向上的位移量 $u(i, j)$ 和 $v(i, j)$, 则变形后的像素点 (i', j') 满足:

$$i' = i + u(i, j), j' = j + v(i, j). \quad (1)$$

在真实实验中, 物体受到外力发生的复杂变形可以分解为多个简单变形, 因此本文自建的散斑位移场图像数据集采用包括平移、拉伸、压缩、旋转、高斯、剪切等多种基础变形组合而成的复合变形。具体实现方式如下式所示:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} k_x & \gamma_x \\ \gamma_y & k_y \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} + \begin{bmatrix} a_x \exp \left[-\frac{(i - i_0)^2}{c_{x0}^2} / 2 - \frac{(j - j_0)^2}{c_{y0}^2} / 2 \right] \\ a_y \exp \left[-\frac{(i - i_1)^2}{c_{x1}^2} / 2 - \frac{(j - j_1)^2}{c_{y1}^2} / 2 \right] \end{bmatrix}, \quad (2)$$

式中: 下标 x 和 y 分别代表水平和垂直方向; u 、 v 表示在坐标 (i, j) 处 x 和 y 方向的位移量; t 表示平移量; k 表示拉压变形程度; γ 表示剪切变形程度; θ 表示旋转变形角度; a 表示高斯变形幅度; (i_0, j_0) 、 (i_1, j_1) 表示高斯变形中心位置; c_{x0} 、 c_{y0} 、 c_{x1} 、 c_{y1} 表示高斯变形的标准差。

由于叠加后的复合位移场像素可能较大, 因此本文利用下式将最大位移限制在 16 个像素以内:

$$d = \frac{32 \times (d - d_{\min})}{d_{\max} - d_{\min}} - 16 \quad (d_{\min} < -16 \cup d_{\max} > 16), \quad (3)$$

式中: d 表示 x 或 y 方向的位移场; d_{\max} 与 d_{\min} 分别表示

最大和最小位移。最大位移超过 10 个像素的位移场可以被称为大变形位移场, 现有的公开可用于深度学习训练的数据集基本上不包含大变形位移场。本研究将最大位移限制为 16 个像素已经基本满足实际大变形测量需求, 同时方便制作数据集时对图像进行裁切以避免目标图像中出现黑边。在深度学习中, 利用大位移数据集训练的模型可以测量小位移场, 而利用小位移数据集训练的模型无法准确测量大位移场, 因此本数据集具有更强的通用性。

直接在参考图像上叠加位移场计算出的像素位置通常是非整像素位置的, 因此需要通过插值来计算出变形后图像的整数位置的像素值。最后对 $512 \text{ pixel} \times 512 \text{ pixel}$ 的图像四周裁剪 16 个像素, 最终得到 $480 \text{ pixel} \times 480 \text{ pixel}$ 的图像。利用上述方法, 本文制作了一个具有 4000 个样本的数据集, 其中训练集、验证集、测试集分别有 2400、1200、400 个样本。I、II、III 类散斑图像共 1200 张, 每张散斑图像分别施加三次不同的位移场用于生成三个不同的样本, 其中两个样本用于训练集, 一个样本用于验证集。IV 类散斑图像共 400 张, 用于生成 400 个样本作为测试集。

本文的位移场生成方法与文献[21]类似, 但是该文献用到的数据集的位移较小且数据集图像尺寸仅为 $128 \text{ pixel} \times 128 \text{ pixel}$ 。而本文的像素最大位移达到了 16 个像素, 且图像分辨率为 $480 \text{ pixel} \times 480 \text{ pixel}$, 这对后续本文提出的网络模型提出了更大的挑战, 部分数据集图像可视化结果如图 2 所示, 第一列和第二列分别表示参考图像和目标图像, 第三列和第四列分别表示 x 和 y 方向的位移场可视化结果。本文提出的散斑位移场图像数据集相比其他现有的数据集更加完善: 1) 散斑图像来源更加多样化, 包括了真实散斑图像和高质量模拟散斑图像; 2) 散斑图像尺寸更大, 具有更高的分辨率; 3) 散斑位移场绝对位移更大。

2.2 基于 UNet 改进的散斑位移场测量网络 DICNet

传统的散斑位移场测量方法是 DIC 法, 其本质是通过图像相关来匹配变化前后相对应的区域。卷积神经网络的本质也是相关运算, 因此本文选择采用卷积神经网络来搭建一个用于测量散斑图像位移场的深度学习模型。

本文以现有的卷积神经网络模型为基础, 对比分析了与本文任务类似的经典光流网络 FlowNetS、FlowNetC^[28] 和经典语义分割网络 UNet 在散斑位移场测量任务中的表现, 具体网络选型过程及实验见 3.1 节。结果表明, UNet 在大位移场散斑图像数据集上具有更好的表现, 因此本文在 UNet 的基础上加入混合注意力机制和深度可分离卷积来搭建用于大变形散斑位移场测量的网络 DICNet。

针对编码器部分的特征提取, 采用混合注意力机制来增加特征匹配和融合的效率, 从而为后续的解码

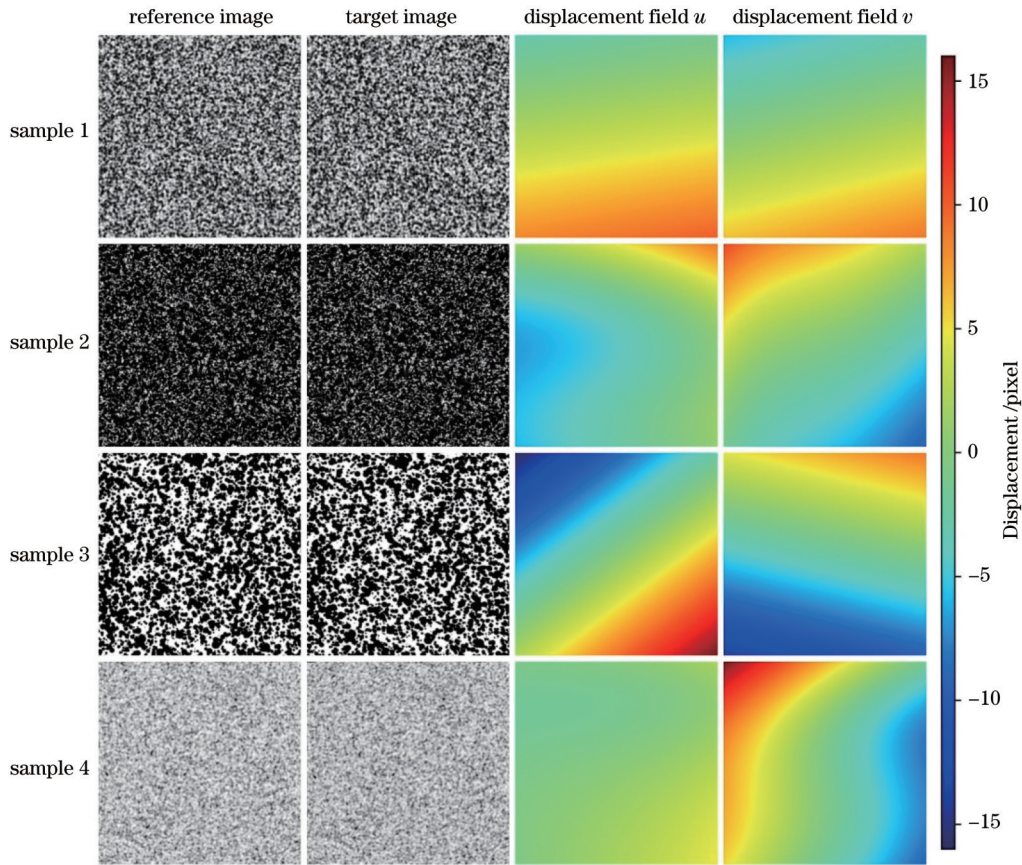


图 2 散斑位移场图像数据集部分样本

Fig. 2 Partial samples of speckle displacement field image dataset

网络提供更加准确的细化特征用于计算位移场。注意力机制可以分为通道注意力机制、空间注意力机制以及混合注意力机制。注意力机制利用平均池化层 (AvgPool)、最大池化层 (MaxPool) 和多层感知机 (MLP) 来实现特征细化, 通道注意力机制实现方式如图 3(a) 和式 (4) 所示, 空间注意力机制实现方式如图 3(b) 和式 (5) 所示, 图 3(c) 所示为混合注意力机制, 其同时具有通道注意力机制和空间注意力机制的优点, 能同时关注通道和空间信息。

$$M_c(F) =$$

$$\sigma \left\{ \text{MLP}[\text{AvgPool}(F)] + \text{MLP}[\text{MaxPool}(F)] \right\}, \quad (4)$$

$$M_s(F) = \sigma \left\{ f^{7 \times 7} \left[[\text{AvgPool}(F); \text{MaxPool}(F)] \right] \right\}, \quad (5)$$

式中: σ 表示 sigmoid 函数; $f^{7 \times 7}$ 表示卷积核尺寸为 7×7 的卷积层; F 为输入特征图; M_c 和 M_s 分别表示通道和空间注意力机制; $[\]$ 表示拼接特征图。

通道注意力机制可以分析不同通道间特征图的关系, 提取出与位移场计算相关度更高的特征图, 提高整体位移场测量精度。空间注意力机制可以分析特征图的空间分布情况, 增加网络对于局部和细节的关注程度, 提升变形复杂的局部位移场测量精度。在卷积神经网络中加入混合注意力模块, 能使网络同时关注重要的特征和空间信息, 进而提升网络的

特征表达能力, 同时提升位移场整体和局部的测量精度。

在编码器网络阶段, 采用深度可分离卷积代替部分卷积层, 且增加部分卷积层的卷积核尺寸, 增大感受野, 提升特征提取能力, 提高大位移场测量精度, 同时有效避免了因为引入混合注意力机制和增大卷积核尺寸导致的网络参数量增加的问题。

深度可分离卷积可以使用更少的计算量和参数完成特征提取和融合, 且计算过程中保证了空间信息的捕捉和跨通道信息的融合与普通卷积相同。设一个卷积层的输入特征图通道数为 C_{in} , 输出特征图通道数为 C_{out} , 卷积核尺寸为 $K \times K$ 。采用如图 4(a) 所示的普通卷积方法, 则参数量为 $C_{in} \times C_{out} \times K \times K$ 。采用如图 4(b) 所示的深度可分离卷积, 则可以将卷积过程分为逐通道卷积 (depthwise convolution) 和逐点卷积 (pointwise convolution), 其参数量为 $C_{in} \times K \times K + C_{out} \times C_{in}$ 。在不损失空间信息和通道信息的前提下, 深度可分离卷积模块的参数量仅为普通卷积模块的 $(C_{out} + K \times K) / (C_{out} \times K \times K)$ 。

通过以上改进, 本文搭建的用于测量具有大变形位移场散斑图像的网络模型 DICNet 结构如图 5 所示。图中深度可分离卷积模块被用于在特征提取阶段代替普通卷积, 混合注意力机制模块设置在特征提取后用

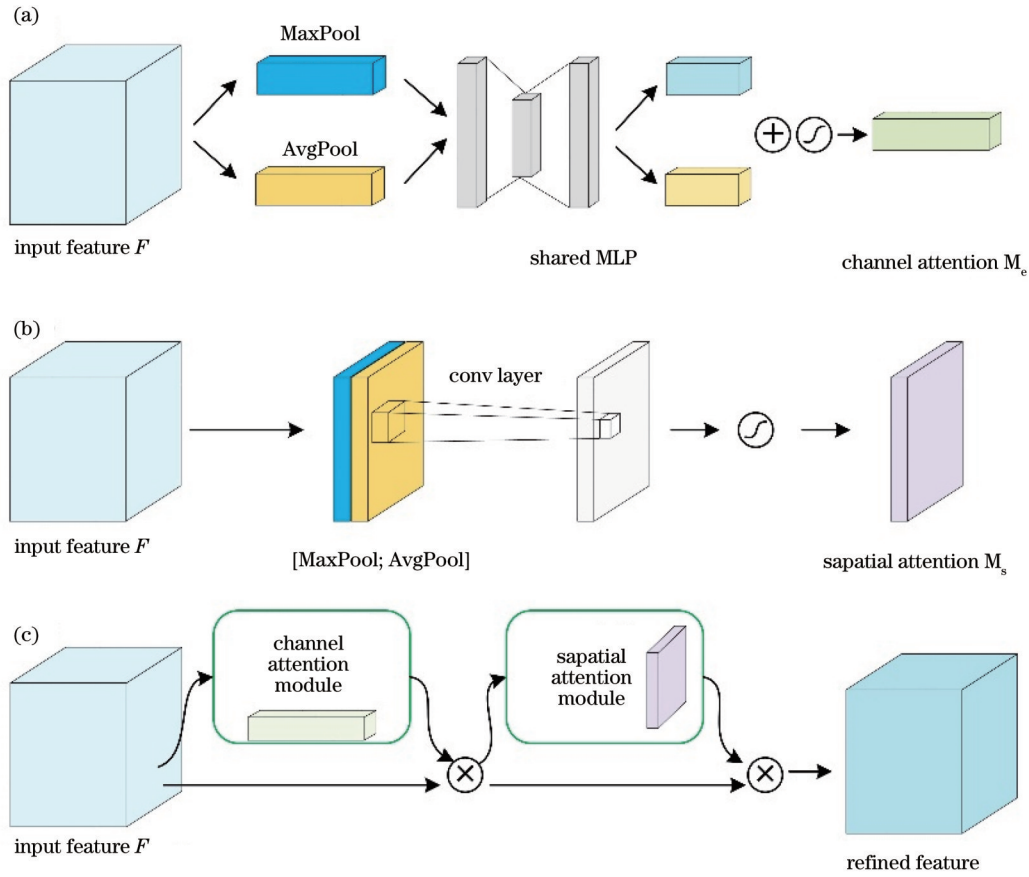


图 3 注意力机制模块。(a)通道注意力模块;(b)空间注意力模块;(c)混合注意力模块

Fig. 3 Convolutional block attention module. (a) Channel attention module; (b) sapatial attention module; (c) convolutional block attention module

于进一步细化特征,网络中的第一和第二个卷积层的卷积核尺寸增加到了 7×7 和 5×5 ,增大了特征提取的感受野,这对于大位移场测量起到了至关重要的作用。

在网络训练阶段,为了提高模型的收敛速度,增加模型的精度。本文结合全局形态损失函数 L_{shape} 和全局绝对损失函数 L_{error} 来提升网络的训练效果。

本文通过定义全局形态损失函数 L_{shape} 来计算预测的位移场与真实位移场之间的形态差异。具体的计算方式如下式所示:

$$L_{shape} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \left(\frac{\hat{d}_{ij} - \hat{d}_{min}}{\hat{d}_{max} - \hat{d}_{min}} - \frac{d_{ij} - d_{min}}{d_{max} - d_{min}} \right)^2, \quad (6)$$

式中: W 和 H 分别为图像的宽和高; d_{ij} 和 \hat{d}_{ij} 分别表示坐标为 (i, j) 位置的像素的真实和预测位移场大小; 下标 max 与 min 分别表示全局最大和最小的位移场。

该损失函数考虑全局位移场,从位移场构成形态的角度出发,衡量真实位移场与测量位移场之间的形态差异。全局形态损失能够将位移场归一化到相同尺度来计算误差,在训练前期,模型预测的位移场形态与真实位移场存在较大差异,因此全局形态损失能提供一个较大的损失来帮助网络更快地拟合位移场形态,从而在提升训练速度的同时避免局部过拟合的

情况。

全局绝对损失函数 L_{error} 用来计算预测位移场与真实位移场之间的绝对误差,它本质上是一个均方误差,如下式所示:

$$L_{error} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (\hat{d}_{ij} - d_{ij})^2. \quad (7)$$

全局绝对损失函数能够最直接准确地反映真实位移场与测量位移场之间的误差,是提高网络测量精度的关键损失函数。

因此,用于训练网络的总损失函数 L_{total} 如下式所示:

$$L_{total} = L_{shape} + L_{error}. \quad (8)$$

当不同的预测结果产生相近的绝对误差时,形态误差可以根据形态差异提供更具有指导性的误差值。当不同的预测结果形态相近时,绝对误差损失可以进一步提高测量精度。

3 实验论证

本文所有深度学习模型训练任务均是在 NVIDIA RTX 3090 显卡上完成的,但是考虑到实际应用中一般会配备成本更低的硬件,因此为了更加符合实际应用需求,本文除了模型训练外的其他计算任

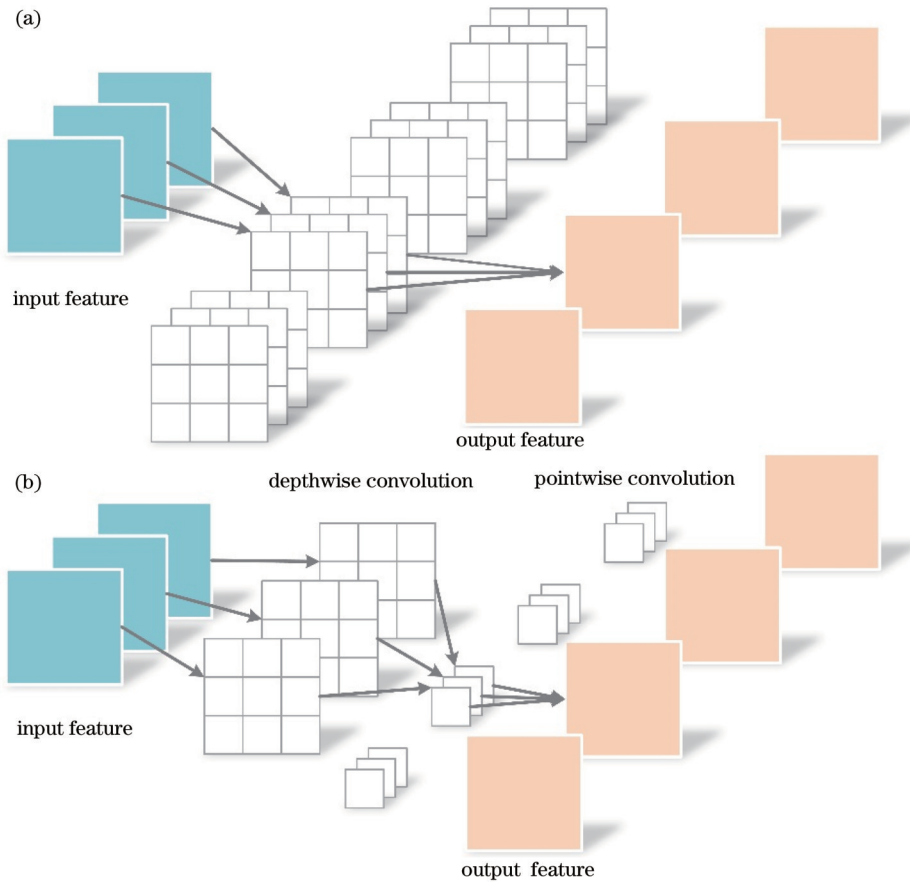


图 4 两种不同的卷积模块。(a)普通卷积;(b)深度可分离卷积

Fig. 4 Two different convolution modules. (a) General convolution; (b) depthwise separable convolution

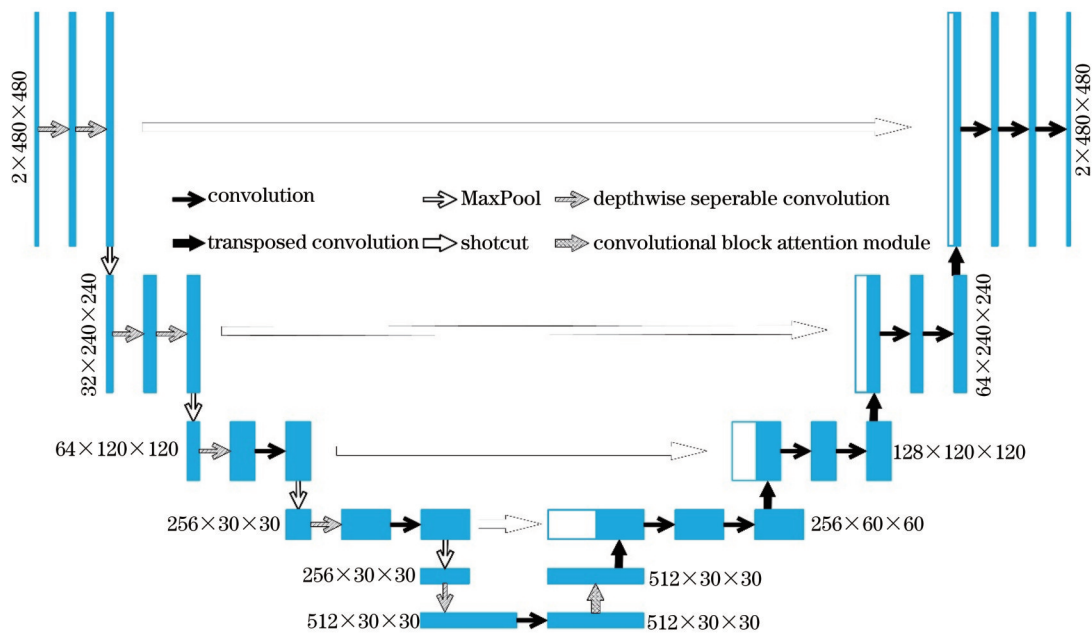


图 5 DICNet 网络结构图

Fig. 5 Structure diagram of DICNet

务的硬件配置均为:处理器为 i7-7700,内存为 16 GB,显卡为 NVIDIA GTX 1080Ti。传统 DIC 算法在中央处理器(CPU)上运行时均用到了多线程加速技术,深度学习方法均通过 GPU 加速执行。

3.1 网络选型实验

在网络选型初期,本文考虑了与散斑位移场测量具有相似性的光流和语义分割。光流用于估计两帧图像中的场景变化或物体运动,但其针对通用场景而非

散斑图像。语义分割虽然是像素分类任务,但是其本质是通过 Softmax 层对每一个像素的推理结果进行分类,这表明语义分割模型的预测精度能够精确每一个像素。因此,本文去掉 Softmax 层并针对性修改损失函数,即可将分类任务转化为亚像素级别的回归任务,这十分符合散斑位移场测量需求。因此,本实验对比分析了经典光流网络 FlowNetS、带相关模块的光流网络 FlowNetC,以及经典语义分割网络 UNet 在散斑位移场测量任务上的表现。

本实验将参考图像和目标图像组合后作为输入数据传入到网络模型,网络模型输出为 x 方向和 y 方向的位移场,使用均方误差(MSE)作为损失函数用于训练网络。由于现阶段的任务仅仅只是定性评估三个网络在散斑位移场测量任务中的表现,因此仅训练了 100

个迭代轮次来观察三个网络模型在训练前期的表现。三个网络模型的参数量 N_{params} 、训练耗时 T_{train} 、在训练集和验证集上的 MSE 损失 $L_{\text{MSE}}^{\text{tr}}$ 和 $L_{\text{MSE}}^{\text{val}}$ 、单样本平均推理时间 $T_{\text{inference}}$ 等信息如表 2 所示,训练过程中的损失曲线如图 6 所示。

表 2 FlowNetS、FlowNetC、Unet 的模型信息和训练结果
Table 2 Model information and training results of FlowNetS, FlowNetC, and UNet

Model	N_{params} / M	T_{train} / min	$L_{\text{MSE}}^{\text{tr}}$ / pixel ²	$L_{\text{MSE}}^{\text{val}}$ / pixel ²	$T_{\text{inference}}$ / ms
FlowNetC	39.175	147	0.358	0.477	8.369
FlowNetS	38.675	80	0.426	0.293	5.272
UNet	7.760	99	0.056	0.073	5.532

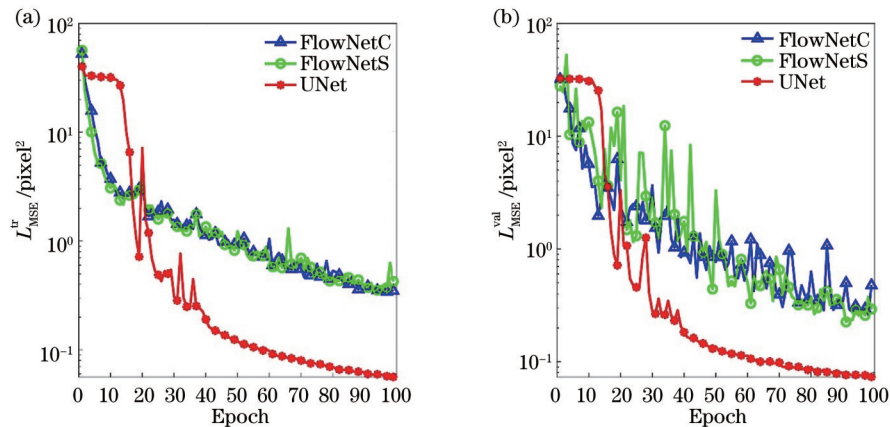


图 6 FlowNetS、FlowNetC、UNet 的训练损失下降曲线。(a) 训练集损失;(b) 验证集损失

Fig. 6 Loss decline curves of FlowNetS, FlowNetC, and UNet during training. (a) Loss of training set; (b) loss of validation set

由以上实验可得,在同样的训练周期下,UNet 在训练集和验证集上的精度都优于 FlowNetC 和 FlowNetS,同时 UNet 具有更小的参数量,训练时间和推理速度都与 FlowNetS 相当,但推理和训练所需要的显存占用更低。虽然本实验没有充分训练以上三个网络,但是从图 6 可以看出,UNet 相比其他网络收敛速度更快,且在验证集上的表现更加稳定,在位移场测量方面,UNet 更具有优势和潜力。因此,本文选用 UNet 作为基础网络,并根据散斑位移场的特性来改进网络以进一步提高网络的性能。此外,充分训练后的改进后的 FlowNet 与 UNet 将在 3.2 节的实验中进行对比分析。

3.2 自建数据集实验对比验证

目前提出的用于散斑位移场测量的深度神经网络有 DeepDIC、StrainNet,但是由于用于训练的数据集不同,模型适用范围也不一致,因此无法直接在某一数据集上对比不同模型的精度。例如,StrainNet 仅在位移场小于 1 的数据集上进行了训练,在本文自制的大位移场数据集上,原始 StrainNet 的测量结果是完全不可靠的,DeepDIC 也存在相同的问题。为了对比不同

的模型在大位移场测量上的表现,本文将 DeepDIC、StrainNet、UNet、DICNet 在自制数据集上训练 500 个迭代轮次至模型收敛,使这些模型在大位移场测量上也具有最佳的表现。训练的损失函数采用本文 2.2 节中提出的新的损失函数,初始学习率设置为 0.0005,批量大小设置为 12,每 20 个迭代轮次学习率衰减为原来的 0.7。需要注意的是,StrainNet 本质上就是 FlowNet 的改进版本,因此本实验没有重复加入 FlowNet 作为对比模型。

为了充分验证本文方法的有效性和鲁棒性,本文还加入了传统 DIC 方法 Ncorr、OpenCorr-CPU、OpenCorr-GPU 作为对照实验。其中 Ncorr 是一个基于 IC-GN 法的 DIC 方法,OpenCorr 是一个最新开源的面向研究人员的 DIC 算法程序,其中 OpenCorr-CPU 和 OpenCorr-GPU 分别对应的是采用 CPU 运算和利用 GPU 加速的版本。

本实验从测量误差、测量标准差、测量时间三个指标综合对比分析本文提出的方法与其他方法在大位移散斑位移场测量任务中的表现。测量误差采用计算测量结果与真值的均方根误差(RMSE)来衡量,其计算

方法为

$$L_{\text{RMSE}} = \sqrt{\frac{1}{2HW} \sum_{i=1}^H \sum_{j=1}^W [(\hat{u}_{ij} - u_{ij})^2 + (\hat{v}_{ij} - v_{ij})^2]}, \quad (9)$$

式中符号含义与式(6)一致, u 和 v 分别表示 x 和 y 方向的位移。不同方法和模型在训练集和验证集上的均方根误差 $L_{\text{RMSE}}^{\text{tr}}$ 和 $L_{\text{RMSE}}^{\text{val}}$ 以及单样本测量平均耗时 $T_{\text{inference}}$ 如表 3 所示, 在测试集上的结果如表 4 所示。

由表 3 可以看出, 在测量精度上, 深度学习整体精度优于传统方法, 其中 DICNet 的 RMSE 最小, 在训练集和验证集上的 RMSE 分别为 0.056 pixel 和 0.055 pixel, 优于传统的 DIC 算法。与最新的基于深度学习的 DIC 算法相比, DICNet 的表现也是最优的, 在训练集上, DICNet 的 RMSE 比 DeepDIC 低 67.82%, 比 StrainNet 低 70.05%。在验证集上, DICNet 的 RMSE 比 DeepDIC 低 68.39%, 比 StrainNet 低 69.95%。与 UNet 相比, DICNet 在训练集和验证集上的 RMSE 分别降低了 39.13% 和 39.56%, 说明本文提出的混合注意力机制和深度可分离卷积改进方法能够真实有效地提升大位移场测量精度。

通过分析标准差可以发现深度学习模型的稳定性

表 3 DICNet 在自建训练集和验证集上的表现

Table 3 Performance of DICNet on self-built training set and validation set

Algorithm and model	RMSE on training set $L_{\text{RMSE}}^{\text{tr}}$ / pixel				RMSE on validation set $L_{\text{RMSE}}^{\text{val}}$ / pixel				Mean time $T_{\text{inference}}$ / ms
	Mean	Maximum	Minimum	Standard deviation	Mean	Maximum	Minimum	Standard deviation	
Ncorr	0.512	1.309	0.056	0.241	0.519	1.245	0.046	0.239	8089.898
OpenCorr-CPU	0.307	2.897	0.004	0.193	0.317	1.619	0.036	0.196	16492.512
OpenCorr-GPU	0.270	2.877	0.003	0.171	0.275	1.276	0.013	0.163	2383.299
DeepDIC	0.174	0.325	0.097	0.019	0.174	0.363	0.089	0.019	15.800
StrainNet	0.187	0.784	0.055	0.071	0.183	0.746	0.054	0.066	3.727
UNet	0.092	0.259	0.049	0.026	0.091	0.241	0.047	0.026	4.015
DICNet	0.056	0.209	0.032	0.019	0.055	0.189	0.034	0.018	4.919

在计算速度方面, 深度学习模型的速度均远超传统算法。结果如表 3 所示, 在 CPU 上运行的传统算法计算量巨大, 在 480 pixel×480 pixel 的图像上, Ncorr 算法单样本平均耗时 8089.898 ms, OpenCorr-CPU 算法单样本平均耗时 16492.512 ms, 即使是通过 GPU 加速后的 OpenCorr-GPU 算法单样本平均耗时仍然达到了 2383.299 ms。而深度学习模型中运行最慢的 DeepDIC 单样本平均耗时仅 15.800 ms, 而 StrainNet、UNet、DICNet 耗时仅为 3~5 ms。由此可见, 深度学习模型能够充分利用 GPU 资源加速计算过程, 能够进行实时在线位移场测量。

DICNet 与其他方法在测试集上的表现如表 4 所示, 测试集中的数据从未被用于训练或验证, 且图像的生成参数与训练集和验证集不相同, 散斑图像存在着明显的差异。虽然 DICNet 在测试集上的表现相比训练集和验证集要差一些, 但是与其他算法和

优于传统算法, DICNet 在训练集上的标准差仅为 0.019, 比 Ncorr 算法低 92.12%, 比 OpenCorr 算法平均低 89.56%, 比 StrainNet 模型低 73.24%, 比 UNet 模型低 26.92%, 与 DeepDIC 相同。DICNet 在验证集上标准差为 0.018, 是以上提到的所有算法和模型中最小的。通过分析本文算法在数据集上的最大和最小误差也可以得出相似结论。在最大误差方面, 传统算法表现相比深度学习模型表现差很多, 传统算法的最大 RMSE 已经达到了 2.897 pixel, 相比之下, 深度学习模型中最大误差表现最差的模型 StrainNet 仅 0.784 pixel, 表现最好的 DICNet 仅 0.189 pixel。在最小误差方面, OpenCorr 算法具有最低的 RMSE, 在训练集上最低 RMSE 仅为 0.003 pixel, 在验证集上的最低 RMSE 仅为 0.013 pixel。但是需要特别说明的是, OpenCorr 算法在测量位移场时会产生少量“坏点”, 在本文自制数据集上, 产生“坏点”的概率约为 0.0122%, 在一张 480 pixel×480 pixel 的图像上, 约产生 28 个“坏点”。虽然“坏点”的数量较少, 但是“坏点”处的位移值却对结果有着极大的影响, 因此本文在计算 OpenCorr 的 RMSE 时剔除了这些“坏点”的值。

模型相比, DICNet 的表现仍然是最佳的。需要注意的是, DICNet 的参数量仅有 6.398 M, 是所有深度学习模型中最小的, 仅为 DeepDIC 的 12.81%, 同时 DICNet 在测试集上的平均 RMSE 为 0.158 pixel, 是所有方法中最优的。DICNet 模型的测量标准差也仅有 0.050 pixel, 仅次于 DeepDIC。图 7 给出了本文提到的几种方法在测试集上的实际位移场测量可视化结果。图中第一行和第二行表示 x 和 y 方向的位移场, 每一列表示不同算法和模型测量的位移场结果。从图中可以看出, 传统算法在图像边缘表现不稳定以及存在大量“坏点”。虽然测试集位移场较为复杂且都是不规则位移场, 深度学习整体表现一致, 都取得了较好的测量结果, 但在部分样本上如图 7(b)、7(c) 所示, DICNet 在图像边角处的表现更好。虽然从图 7 中无法特别直观地展现出各个模型的表

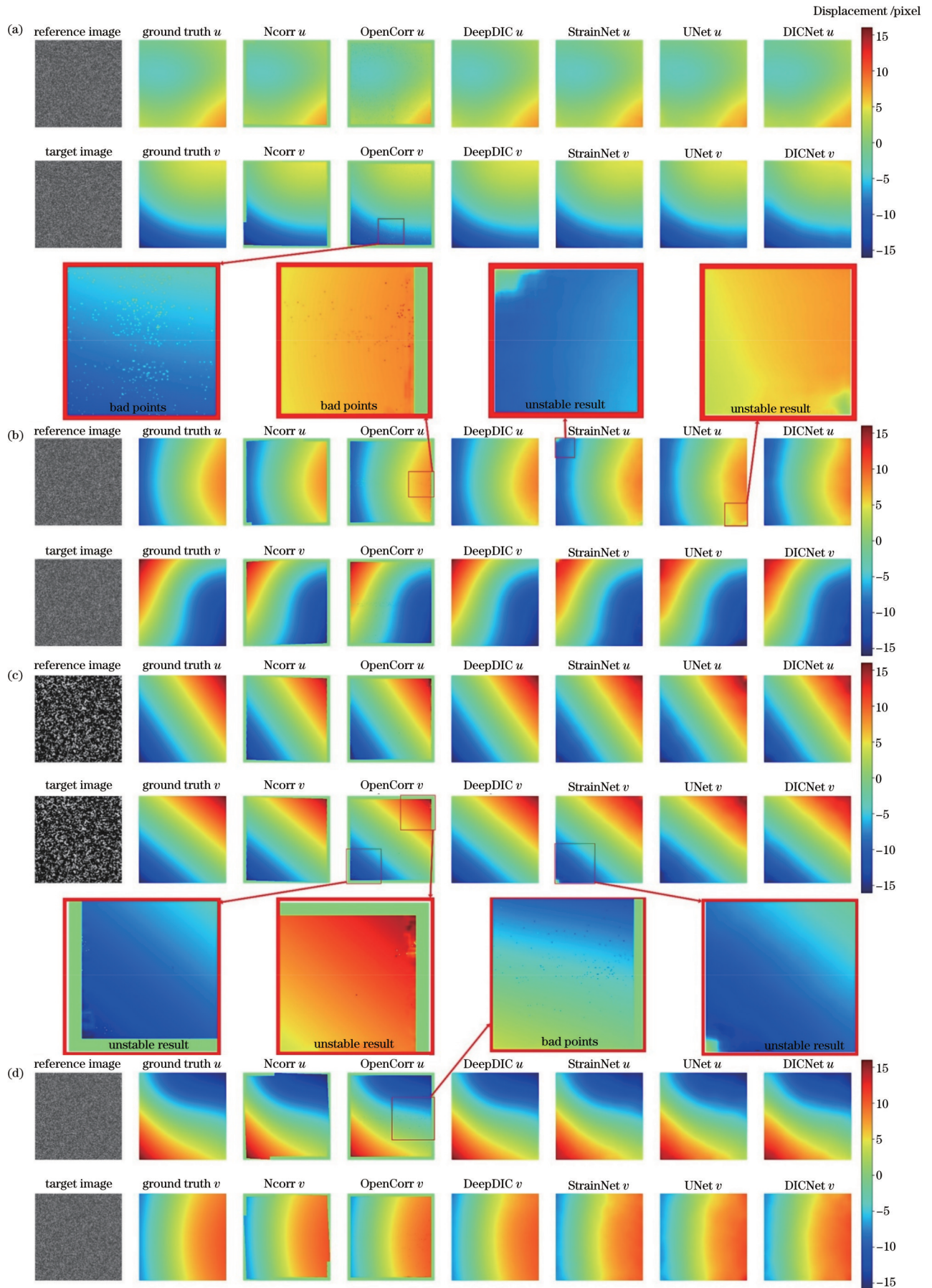


图 7 算法和模型在测试集上的表现。(a)样本 1;(b)样本 2;(c)样本 3;(d)样本 4

Fig. 7 Performance of algorithms and models on test set. (a) Sample 1; (b) sample 2; (c) sample 3; (d) sample 4

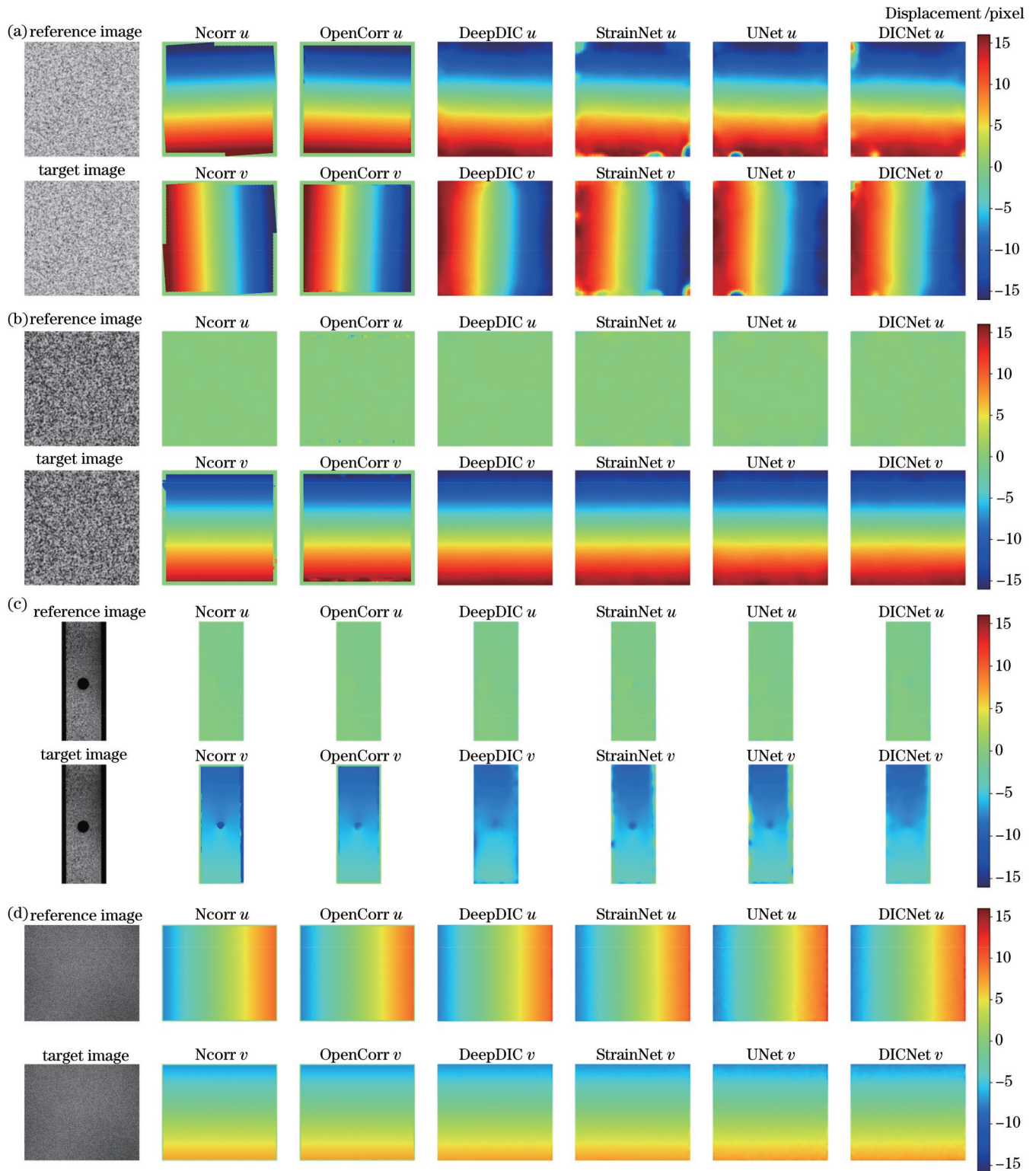


图 8 算法和模型在 DIC 挑战数据集上的表现。(a) 样本 8; (b) 样本 10; (c) 样本 12; (d) 样本 17

Fig. 8 Performance of algorithms and models on DIC challenge dataset. (a) Sample 8; (b) sample 10; (c) sample 12; (d) sample 17

现差异,但是结合表 4 可知 DICNet 具有更高的精度和稳定性。

3.3 公开数据集实验对比验证

为了进一步验证 DICNet 的泛化性能,本文将 DICNet 在公开的 DIC 挑战数据集上进行了实验,并与传统 DIC 算法 Ncorr、OpenCorr 以及深度学习模型

DeepDIC、StrainNet、UNet 作对比分析。由于 DIC 挑战数据集中的样本的位移场真值没有公开,因此本实验仅通过可视化位移场的方式对实验结果作出定性分析。本实验仅对比分析 DICNet 在样本 8~12、15、17 共 7 个样本上的表现,因为这些样本具有大位移场,与 DICNet 的设计初衷相符。

表 4 DICNet 在自建测试集上的表现
Table 4 Performance of DICNet on self-built test set

Algorithm and model	RMSE on test set L_{RMSE}^e /pixel				Number of model parameter N_{params} / M
	Mean	Maximum	Minimum	Standard deviation	
Ncorr	0.509	1.127	0.071	0.223	—
OpenCorr-CPU	0.165	0.596	0.029	0.091	—
OpenCorr-GPU	0.176	0.732	0.026	0.109	—
DeepDIC	0.183	0.250	0.140	0.019	49.970
StrainNet	0.206	0.557	0.067	0.072	38.663
UNet	0.174	0.413	0.070	0.050	7.760
DICNet	0.158	0.267	0.078	0.039	6.398

由于实验用到的多种方法在 DIC 挑战数据集上的表现基本一致,因此仅展示如图 8 所示的 4 个实验结果存在一定差异的样本。图中第一行和第二行分别表示 x 和 y 方向的位移场,每一列表示不同算法和模型测量的位移场结果。由图 8 可知,除了在样本 8 上,StrainNet、UNet、DICNet 在图像的边角处都出现了不稳定的表现,在其他所有样本上,深度学习方法与传统算法得出结果具有高度一致性。需要特别说明的是,样本 8 的参考图像是通过旋转变化的目标图像,因此变化前后的图像边角处像素会存在缺失导致计算失败,这属于正常误差。仔细观察可以发现,Ncorr 和 OpenCorr 算法在所有样本上的测量结果边界处都存在绿色的区域,这些区域的预测值为 0,这是由于传统算法需要选定一定大小的子区来完成匹配,所以导致边界区域无法正常计算。相比之下,深度学习模型在大多数样本上的边界处可以正确测量位移场。由该实验可知,虽然无法通过与真实位移场作对比来定量分析 DICNet 的鲁棒性,但是 DICNet 的测量结果与传统算法 Ncorr 与 OpenCorr 具有高度的一致性,表明 DICNet 在公开数据集上依然具有良好的泛化性能。

4 结 论

提出了一种基于深度学习的散斑位移场测量方法,在 UNet 的基础上引入了混合注意力机制和深度可分离卷积设计了一个参数量仅为 6.398 M 的卷积神经网络模型 DICNet。为了对网络模型进行有效训练,在数据集方面,本文自建了具有大位移场、散斑多样性、位移场多样性的散斑图像数据集,在损失函数方面,本文结合全局形态损失和全局绝对损失来加速模型收敛过程。

为了验证所提方法的有效性和鲁棒性,将 DICNet 与其他常用的传统 DIC 方法和深度学习模型在自建数据集以及公开数据集上进行了对比分析。结果表明,在自建数据集上,DICNet 在精度、稳定性方面都优于其他方法,且计算耗时远低于传统方法。在公开 DIC 挑战赛数据集上,DICNet 的测量结果与传统方法

表现出了高度的一致性,表明了 DICNet 在不同的数据集上都具有良好的鲁棒性。

参 考 文 献

- [1] Yamaguchi I. A laser-speckle strain gauge[J]. Journal of Physics E: Scientific Instruments, 1981, 14(11): 1270-1273.
- [2] Peters W H, Ranson W F. Digital imaging techniques in experimental stress analysis[J]. Optical Engineering, 1982, 21(3): 427-431.
- [3] Sutton M A, Wolters W J, Peters W H, et al. Determination of displacements using an improved digital correlation method[J]. Image and Vision Computing, 1983, 1(3): 133-139.
- [4] Pan B. Bias error reduction of digital image correlation using Gaussian pre-filtering[J]. Optics and Lasers in Engineering, 2013, 51(10): 1161-1167.
- [5] Su Z L, Lu L, He X Y, et al. Recursive-iterative digital image correlation based on salient features[J]. Optical Engineering, 2020, 59(3): 034111.
- [6] Mazzoleni P, Matta F, Zappa E, et al. Gaussian pre-filtering for uncertainty minimization in digital image correlation using numerically-designed speckle patterns[J]. Optics and Lasers in Engineering, 2015, 66: 19-33.
- [7] Jiang L B, Xie H M, Pan B. Speeding up digital image correlation computation using the integral image technique[J]. Optics and Lasers in Engineering, 2015, 65: 117-122.
- [8] 谭芳喜, 肖世德, 李晟尧, 等. 基于密集特征匹配的数字图像相关法[J]. 激光与光电子学进展, 2021, 58(16): 1612001.
- [9] Tan F X, Xiao S D, Li S Y, et al. Digital image correlation method based on dense feature matching[J]. Laser & Optoelectronics Progress, 2021, 58(16): 1612001.
- [9] 李桂华, 马万龙, 朱天天, 等. 基于数字图像相关方法的空心圆盘热变形测量[J]. 激光与光电子学进展, 2021, 58(14): 1412003.
- [9] Li G H, Ma W L, Zhu T T, et al. Thermal deformation measurement of hollow disk based on digital image correlation method[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1412003.
- [10] 王永红, 朱奕磊, 高启学, 等. 基于数字图像相关法的空间目标位姿测量[J]. 光学学报, 2022, 42(8): 0812001.
- [10] Wang Y H, Zhu Y L, Gao Q X, et al. Position and pose measurement of spatial object based on digital image correlation [J]. Acta Optica Sinica, 2022, 42(8): 0812001.
- [11] 潘济宇, 张水强, 苏志龙, 等. 基于数字图像相关的水下螺旋桨三维变形测量[J]. 光学学报, 2021, 41(12): 1212001.
- [11] Pan J Y, Zhang S Q, Su Z L, et al. Measuring three-dimensional deformation of underwater propellers based on digital image correlation[J]. Acta Optica Sinica, 2021, 41(12): 1212001.

- [12] Bruck H A, McNeill S R, Sutton M A, et al. Digital image correlation using Newton-Raphson method of partial differential correction[J]. *Experimental Mechanics*, 1989, 29(3): 261-267.
- [13] Baker S, Matthews I. Equivalence and efficiency of image alignment algorithms[C]//*Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR, December 8-14, 2001, Kauai, HI, USA*. New York: IEEE Press, 2003.
- [14] Baker S, Matthews I. Lucas-kanade 20 years on: a unifying framework[J]. *International Journal of Computer Vision*, 2004, 56(3): 221-255.
- [15] Blaber J, Adair B, Antoniou A. Ncorr: open-source 2D digital image correlation Matlab software[J]. *Experimental Mechanics*, 2015, 55(6): 1105-1122.
- [16] Jiang Z Y, Qian K M, Miao H, et al. Path-independent digital image correlation with high accuracy, speed and robustness[J]. *Optics and Lasers in Engineering*, 2015, 65: 93-102.
- [17] Zhang L Q, Wang T Y, Jiang Z Y, et al. High accuracy digital image correlation powered by GPU-based parallel computing[J]. *Optics and Lasers in Engineering*, 2015, 69: 7-12.
- [18] 刘小勇, 官岩, 李荣丽, 等. 基于 BP 神经网络的数字图像相关非迭代灰度梯度算法[J]. *机床与液压*, 2018, 46(1): 7-11.
- Liu X Y, Gong Y, Li R L, et al. Non-iterative gray-gradient algorithm based on BP artificial neural network in digital image correlation[J]. *Machine Tool & Hydraulics*, 2018, 46(1): 7-11.
- [19] Min H G, On H I, Kang D J, et al. Strain measurement during tensile testing using deep learning-based digital image correlation [J]. *Measurement Science and Technology*, 2020, 31(1): 015014.
- [20] Boukhtache S, Abdelouhab K, Berry F, et al. When deep learning meets digital image correlation[J]. *Optics and Lasers in Engineering*, 2021, 136: 106308.
- [21] Yang R, Li Y, Zeng D, et al. Deep DIC: deep learning-based digital image correlation for end-to-end displacement and strain measurement[J]. *Journal of Materials Processing Technology*, 2022, 302: 117474.
- [22] 黄举, 孙翠茹, 林祥龙. 基于卷积神经网络的散斑图像位移场测量方法[J]. *光学学报*, 2021, 41(20): 2012002.
- Huang J, Sun C R, Lin X L. Displacement field measurement of speckle images using convolutional neural network[J]. *Acta Optica Sinica*, 2021, 41(20): 2012002.
- [23] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [24] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 3-19.
- [25] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [26] Reu P L, Toussaint E, Jones E, et al. DIC challenge: developing images and guidelines for evaluating accuracy and resolution of 2D analyses[J]. *Experimental Mechanics*, 2018, 58 (7): 1067-1099.
- [27] Sur F, Blaysat B, Grédiac M. Rendering deformed speckle images with a Boolean model[J]. *Journal of Mathematical Imaging and Vision*, 2018, 60(5): 634-650.
- [28] Dosovitskiy A, Fischer P, Ilg E, et al. FlowNet: learning optical flow with convolutional networks[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 2758-2766.

Large Deformation Measurement Method of Speckle Images Based on Deep Learning

Xiao Hong, Li Chengnan, Feng Mingchi*

School of Advanced Manufacturing Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract

Objective As the demand for materials with excellent mechanical properties is increasing in scientific research and engineering, determining how to accurately measure the global displacement field of materials in mechanical experiments has become an important scientific research issue. Digital image correlation (DIC) algorithm is a non-contact optical method for measuring global speckle displacement fields based on visible light, which is widely used in experimental mechanics and engineering fields. It has the advantages of low measurement costs, high precision, high sensitivity, strong anti-interference ability, and global measurement. However, traditional DIC algorithm cannot meet the requirements of real-time measurement in practical applications, which greatly limits the development and promotion of this method. With the rapid development of deep learning in computer vision, deep learning methods gradually come into use in DIC algorithm. Thanks to the efficient calculation by general processing unit (GPU) devices, the deep learning-based method for measuring the speckle displacement field can more easily achieve real-time online calculation. Although the method is much faster than the traditional one, the model cannot accurately measure the complex large deformation displacement field in practical applications due to the incomplete dataset. Hence, this work aims to construct a more realistic and comprehensive speckle image dataset with a large deformation displacement field and propose a fast and high-precision deep learning model to measure the displacement field of speckle images with large deformation.

Methods A large number of different types of speckle images is obtained in various ways (Fig. 1) to construct a speckle image dataset with a large deformation displacement field in line with the actual situation. These speckle images are obtained from real experiments and computer simulations under different parameter combinations (Table 1). Then, a composite deformation composed of translation, stretching, compression, rotation, Gauss, shear, and other basic deformations is used to define the random displacement field. Finally, a speckle-image displacement-field dataset with a maximum displacement of 16 pixel and large deformation in line with the actual deformation is produced. In terms of the deep learning network model, a fast and high-precision network model DICNet (Fig. 5) for measuring the speckle images with a large-deformation displacement field is built upon the improvement on UNet. DICNet introduces a convolutional block attention module to increase the efficiency of feature extraction and fusion, uses depthwise separable convolution to replace some ordinary convolutional layers, and increases the convolution kernel size of some convolutional layers. It improves the displacement-field measurement accuracy and reduces the number of parameters of the network model. At the network training stage, a combination of the global shape loss function and global absolute loss function is proposed to improve the convergence speed and accuracy of the model.

Results and Discussions Network selection experiments are conducted to prove that UNet is a rational basic network model for measuring the large-deformation displacement (Table 2). It has higher measurement accuracy of the displacement field, a smaller number of parameters, and faster inference speed. The DICNet proposed in this work is compared with the traditional DIC algorithm and the latest deep learning methods on the self-built dataset, and the performance of these methods in the measurement task of the large-displacement displacement field is comprehensively compared in terms of three indicators, i. e. , the root-mean-square error (RMSE), the standard deviation, and mean time (Table 3). The results show that the measurement accuracy of the deep learning method is better than that of the traditional method. The RMSE of DICNet on the training set and the validation set is 0.056 pixel and 0.055 pixel, respectively, which is 67%–70% lower than that of other existing methods and about 39% lower than that of the original UNet network. On the test set, DICNet still has the smallest RMSE and the most stable performance (Table 4). The experiments of DICNet are also conducted on the public DIC challenge dataset (Fig. 8). The results show that the measurement results of the proposed method are highly consistent with those of the traditional algorithms, which indicates that the proposed method still has good generalization performance on the public dataset.

Conclusions This work proposes a displacement field measurement method for speckle images with complex large deformation. This method uses the convolutional block attention module and depthwise separable convolution to improve the UNet network for measuring large deformation displacement fields. To train the model, this work constructs a dataset containing multiple types of speckle images and complex large-deformation displacement fields in line with the real situation and proposes a new loss function. This method is compared with traditional DIC algorithm and the latest deep learning methods on the self-built dataset and public dataset separately. The results show that the measurement results of DICNet are highly consistent with those of other methods, and the method in this work achieves the highest average accuracy with the smallest number of model parameters. The measurement speed of the displacement field is far higher than those of traditional methods, which can meet the actual real-time measurement requirements of a large deformation displacement field. The source code and network pre-trained weights of this study are available at <https://github.com/donotbreeze/Large-deformation-measurement-method-of-speckle-image-based-on-deep-learning>. The dataset is available at https://pan.baidu.com/s/1KzC9g_GIkvMnGFumDYGyBA?pwd=fd5x.

Key words measurement; digital image correlation; large deformation displacement field; deep learning; displacement field measurement