

基于合成散斑技术和深度学习的 DIC 位移测试技术研究

卢烁, 陈佳妍, 唐锦辉

武汉理工大学 汽车工程学院, 湖北 武汉 430070

DOI:10.61369/ME.2026040002

摘 要 : 数字图像相关 (Digital Image Correlation, DIC) 是一种非接触式全场光学测量技术, 能够获取材料和结构在加载过程中的位移信息, 已广泛应用于材料力学试验与结构健康监测等领域。传统 DIC 方法依赖于子集匹配与迭代优化算法计算位移场, 计算量大、处理时间长。为解决上述不足, 本文开展基于深度学习的 DIC 位移预测方法研究。该技术为材料试验、结构性能评估以及工业在线监测提供了一种高效、可靠的全场位移测量新方法。

关 键 词 : 数字图像相关 (DIC); 深度学习; 位移预测; 光流法; 水转印散斑; Transformer

Research on DIC Displacement Measurement Techniques Based on Synthetic Spotty Light Technology and Deep Learning

Lu Shuo, Chen Jiayan, Tang Jinhui

School of Automotive Engineering, Wuhan University of Technology, Wuhan, Hubei 430070

Abstract : Digital Image Correlation (DIC) is a non-contact full-field optical measurement technique that obtains displacement information of materials and structures during loading, and has been widely used in material mechanical testing and structural health monitoring. Traditional DIC methods rely on subset matching and iterative optimization algorithms for displacement field calculation, which are computationally intensive and time-consuming. To address these shortcomings, this paper investigates deep learning-based DIC displacement prediction methods. First, high-quality water-transfer printed speckle images obtained from tensile tests are used, and Farneback dense optical flow combined with Delaunay meshing is employed to generate pixel-wise displacement field labels, constructing a training dataset. Subsequently, a lightweight end-to-end displacement prediction model based on convolutional neural networks and Transformer mechanisms is built to directly regress full-field displacement from speckle image pairs, avoiding the complex search and iteration processes of traditional algorithms. Meanwhile, the network architecture is optimized with depthwise separable convolutions and attention mechanisms, achieving millisecond-level inference speed to meet real-time measurement requirements. Experimental results show that the proposed method significantly improves computational efficiency while maintaining sub-pixel measurement accuracy, achieving an average prediction rate of 70%–80% on an independent test set and an inference speed of 80–120 FPS. This technique provides an efficient and reliable new method for full-field displacement measurement in material testing, structural performance evaluation, and industrial online monitoring.

Keywords : Digital Image Correlation (DIC); deep learning; displacement prediction; opticalflow; water-transfer printing speckle; Transformer

引言

材料与结构服役变形的高精度测量是实验力学重要研究内容, 现有测量方法分接触式 (电阻应变片、引伸计等) 与非接触式两类, 其中数字图像相关 (DIC) 因适用性广、精度高, 在航空航天、土木工程等领域广泛应用^[1]。但传统 DIC 需制作散斑图案, 通过子集匹配与迭代优化计算位移场, 计算量大、耗时久, 难以满足实时或高频测试需求^[2]。深度学习为 DIC 提供了新路径, 研究者利用 CNN 突破传统计算瓶颈: Wang^[3]对比了局部与全局 DIC 性能, DIC-Net 提升了复杂变形场计算效率, 多通道输入神经网络则实现了时空特征联合提取^[4]。

为此, 本文开展基于水转印散斑与深度学习的 DIC 位移预测研究: 通过单轴拉伸获取真实散斑图像, 采用 Farneback 光流与 Delaunay 网格化生成像素级位移标签, 设计轻量化 CNN-Transformer 混合模型, 实现全场位移端到端快速精准预测, 并通过测试集验证模型精度与效率。

基金项目: 省级大学生创新创业训练计划项目: 基于深度学习的 Dic 应变测试技术研究, 编号: S202510497107。

一、数据收集

（一）拉伸实验系统与试样制备

本研究所用的源图像来自多组铝合金板材的单轴拉伸试验。试验件按照 GB/T 228.1-2010 标准制备为哑铃状平板试样，平行段长度为 50 mm，宽度为 12 mm，厚度为 2 mm。试验件表面处理采用水转印工艺：首先在清洁的金属表面覆盖白色水转印底膜，然后将随机散斑图案的水转印膜通过水浴转移至试样表面，经烘干固化成形。该工艺可生成高对比度、各向同性、颗粒尺寸均匀（3~5 像素）的随机散斑图案，且具有良好的附着力与重复性，满足 DIC 测量对散斑质量的要求。

（二）光学测量系统与实验参数设置

拉伸试验在万能材料试验机上进行，同时搭建了同步触发光学测量系统。该系统由一台分辨率为 2448×2048 像素的工业 CMOS 相机、双远心镜头和同轴 LED 光源组成，以确保在加载过程中试样与镜头距离变化时图像放大倍率保持恒定，并消除离面位移对二维位移测量的干扰。相机视场覆盖试样平行段全长，经标定后空间分辨率约为 0.025 mm/pixel。

试验采用位移控制加载模式，拉伸速率设定为 1 mm/min。工业相机通过软件与试验机横梁位移信号同步，以 1 帧/秒的固定帧率连续触发采集。每根试样从加载开始至断裂完整记录全过程的散斑图像序列，单根试样约获取 600~900 帧图像，累计多组试验共采集约 5000 帧有效散斑图像。所有采集的原始图像以无损 TIFF 格式实时保存至工作站。

（三）原始数据预处理

深度学习模型的性能高度依赖于高质量的训练数据。纯粹解析生成的随机位移场（如二维高斯函数叠加）往往缺乏真实的物理连贯性^[5,6]。为此，本研究摒弃了全合成数据路线，转而采用一种基于真实散斑图像的构建策略。该策略以拉伸试验中采集的高质量散斑图像为数据基底，通过稠密光流估计与网格力学分析生成高保真度的位移场标签，以克服传统子集匹配方法对离散网格与形函数阶次的人为依赖。本节的目的：利用真实拉伸图像与光流+网格化方法，生成可供深度学习模型训练的高质量像素级位移场标签。区别于传统 DIC 中依赖于子集匹配的离散位移输出，本节方法可直接输出每个像素的位移值，且通过网格化平滑抑制噪声，从而为后续模型提供准确、平滑、物理一致的监督信号。

1. 基于 Farneback 光流的稠密位移场估计

“局部邻域”是指图像中以某个像素为中心的一个小矩形窗口（例如 15×15 像素）。Farneback 算法假设在该窗口内，所有像素的运动规律可以用同一个简单的数学公式（二次多项式）近似描述。

2. 位移场标签的网格化平滑

由 Farneback 光流直接输出的稠密位移场 U_{flow} 存在一定的随机噪声。若直接作为深度学习模型的监督信号，模型会学习到这些噪声，导致泛化能力下降。

3. 位移映射与网格化平滑

利用第 2.3.1 节计算得到的光流场 U_{flow} ，对每个特征

点 $p_i \in P_{\text{ref}}$ ，通过光流场在整数坐标上的双线性插值得到该点的位移矢量 $\mathbf{d}_i = (u_i, v_i)$ （代码 warp_points_with_flow）。由此获得网格节点上的位移值。

（四）训练数据集构建

1. 样本划分

通过第 2.3 节方法，共获得约 5000 对（参考图像，变形图像）样本，每对对应像素级位移场标签。按照 8:1:1 的比例随机划分为训练集、验证集和测试集：80% 用于模型训练，10% 用于验证（监控过拟合并执行早停机制），10% 作为独立测试集（不参与训练或验证，仅用于最终性能评估）。

2. 预处理与数据增强

预处理：将所有图像缩放至 512×512 像素（双线性插值），并归一化至 $[0,1]$ 区间。

在线数据增强：为提升泛化能力，训练过程中采用在线随机变换，即：不预先扩充数据集总量（仍为原始约 5000 对样本），而是在每个训练批次中，对即将输入模型的图像对与标签实时施加随机变换。每遍历一轮训练集（epoch），模型看到的同一图像可能经历不同变换，从而隐式增加样本多样性。变换包括（均保持参考图像、变形图像与位移标签的空间对应关系一致）：

二、模型架构与训练方法

（一）模型架构

传统数字图像相关（DIC）技术依赖复杂的子集匹配和后处理滤波，计算精度与效率受限于形函数阶次、子集大小及插值算法等因素，且在处理大变形或散斑质量退化时易出现去相关现象^[7]。深度卷积神经网络凭借其强大的非线性映射能力与特征提取能力，为实现端到端、高鲁棒性的全场变形测量提供了新途径。本研究提出一种融合卷积神经网络与 Transformer 机制的混合架构模型——UltraStrainTransformer，旨在从输入的散斑图像对中直接回归出高分辨率位移场。整体架构采用编码器-解码器结构，并在深层引入自注意力机制以增强对长程依赖关系的捕获能力^[8]。

1. 输入层

模型的输入为一对图像：参考帧图像 I_{ref} 与变形帧图像 I_{def} 。两幅图像均为灰度图，尺寸为 $H \times W$ （本研究中 $H = W = 512$ 像素）。输入层将两幅图像在通道维度上进行拼接，形成形状为 $(2, H, W)$ 的张量 $\mathbf{X}_m$ ，作为网络的直接输入。该拼接方式使模型能够同时感知参考状态与变形状态之间的差异，有利于后续位移场的估计。

2. 编码器

编码器由四个下采样阶段构成，逐步提取多尺度语义特征。为在保持特征提取能力的同时降低计算开销，本研究采用深度可分离卷积替代标准卷积作为基础算子，具体实现参考代码中的 DepthwiseSeparableConv2d 类。在每个编码阶段，输入特征图依次经过深度卷积与逐点卷积，并嵌入通道注意力机制（Squeeze-and-Excitation）以自适应地校准特征响应，具体地，令第 i 阶段

的输入特征图为 X_i ，则该阶段输出为：

$$X_{i+1} = \text{MaxPool}\left(\text{GELU}\left(\text{BN}\left(\text{PWConv}\left(\text{DWConv}\left(X_i\right)\right)\right)\right)\right)$$

其中，DWConv 表示深度卷积，PWConv 表示逐点卷积，BN 为批归一化，GELU 为激活函数。

下采样时，当 downsample=True 时，使用最大池化（MaxPooling，核大小 2×2 ）将特征图尺寸减半。四个编码阶段的输出通道数依次为： $C_1=32$ ， $C_2=64$ ， $C_3=128$ ， $C_4=256$ 。特征图尺寸从 $H \times W$ 逐步下采样至 $H/8 \times W/8$ 。

3. Transformer 瓶颈层

为提升模型泛化能力，在训练过程中采用在线数据增强，即：不预先扩充数据集总量，而是在每个训练批次中、对输入模型前的图像对与标签实时施加随机变换。每轮训练时，模型看到的同一张图像可能经历不同的变换，相当于隐式扩充了训练样本的多样性。

4. 解码器（UltraDecoderBlock）

解码器通过双线性插值上采样逐步恢复空间分辨率，并与编码器的对应层级特征进行跳跃连接，以补偿下采样过程中丢失的空间细节。为了更有效地融合多尺度信息，本研究在解码器的高层（特征图尺寸 $H/4 \times W/4$ ）与底层（编码器的 $H/4 \times W/4$ 特征）之间引入了跨尺度注意力模块（CrossScaleAttention）。该模块计算低分辨率特征序列与高分辨率特征序列之间的注意力权重，使得模型在恢复全局结构的同时能保留局部精细纹理，这对于精确捕捉应变局部化区域（如颈缩带）至关重要。

5. 输出层

解码器最终输出通道数为 16 的特征图，尺寸恢复至 $H \times W$ 。解码器输出的特征被送入由两层深度可分离卷积（输出通道 $16 \rightarrow 8 \rightarrow 2$ ）组成的位移预测头，最终输出形状为 $(2, H, W)$ 的张量，分别对应 x 方向和 y 方向的位移分量 u 和 v 。

（二）模型训练策略

（1）损失函数

本文直接针对位移场进行监督，因为应变场可由位移场的空间导数导出，且位移误差更能反映模型对散斑运动的物理解释。损失函数选用掩膜均方误差（Masked MSE），定义如下：

$$L = \frac{\sum_i 1_{\text{mask}} \left[(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2 \right]}{\sum_i 1_{\text{mask}}}$$

其中 (u_i, v_i) 为模型预测的位移分量， $(\hat{u}_i, \hat{v}_i)$ 为由第 2.3 节所述光流-网格法生成的真值位移， 1_{mask} 为指示函数（板材内部取 1，背景取 0）。该损失函数仅关注有效变形区域的预测质量，避免背景噪声干扰。

（2）优化器与超参数

采用 Adam 优化器，初始学习率 5×10^{-5} ，权重衰减系数 1×10^{-5} 。使用 ReduceLROnPlateau 学习率调度器，当验证损失连续 15 个 epoch 未下降时，学习率乘以 0.5。训练最大 epoch 设为 100，并引入早停机制（patience=30）：若验证损失连续 30 个

epoch 未改善则终止训练，并回滚至验证损失最小的模型权重。同时采用超参数设置：批大小（batch_size）= 8；数据加载并行进程数（num_workers）= 8，启用页锁定内存（pin_memory=True）；启用混合精度训练（torch.cuda.amp），采用 FP16 进行前向与反向传播，同时用 torch.compile 加速计算图编译。

（三）模型评估与结果

1. 评估指标

为定量评估模型的位移预测精度，本文采用预测率（Prediction Rate）作为核心指标。首先计算每个像素的绝对百分比误差，然后取平均得到平均绝对百分比误差（MAPE）：

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{\sqrt{(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2}}{\sqrt{\hat{u}_i^2 + \hat{v}_i^2} + \delta} \times 100\%$$

其中 N 为板材区域内的像素总数， $\delta=10^{-6}$ 为防止分母为零的极小常数。进而定义预测率：

$$\text{预测率} = \max(0, 100 - \text{MAPE})$$

预测率取值范围为 0%~100%，数值越高代表模型位移预测越准确。

2. 测试结果

在独立测试集（600 帧图像对）上逐帧计算预测率，并取所有帧的平均值作为最终评估结果。实验表明，本文模型在测试序列上的平均预测率达到 70%~80%（如图 3-2 所示），其中均匀变形阶段预测率稳定在 78%~82%，颈缩大变形阶段仍保持在 70%~75%。同时，模型单次推理时间仅 8~12 ms（约 80~120 FPS），较传统 DIC 方法提升两个数量级。该结果验证了所提方法在保持亚像素级精度的同时，满足实时测量需求。

三、结论

本研究针对传统数字图像相关（DIC）技术在处理复杂结构和大变形时存在的计算耗时长的工程瓶颈，提出并实现了一种基于深度学习的高效全场变形测量新框架。通过综合研究与实验验证，得出以下核心结论：

首先，采用水转印工艺制备高质量散斑试样，通过单轴拉伸试验采集真实散斑图像序列。该方法相比于传统喷涂法具有图案均匀、颗粒尺寸可控、可批量复制的优势，为后续深度学习模型提供了与真实工况一致的图像数据基础。

其次，在数据集构建阶段，利用 Farneback 稠密光流算法生成像素级初始位移场，再结合 Delaunay 三角网格化方法进行空间平滑与去噪，生成高保真度位移场标签。该策略既保留了光流法亚像素级稠密位移的优势，又通过网格的力学连续性抑制了噪声，解决了深度监督学习中标注数据获取困难的关键问题。

然后，设计轻量化 CNN-Transformer 混合架构模型（UltraStrainTransformer）。模型采用编码器-解码器结构，在

瓶颈层引入多头自注意力机制以捕捉大变形下的长程空间依赖，同时使用深度可分离卷积降低计算量。模型从参考图像与变形图像对中端到端回归全场位移场，避免了传统 DIC 方法中的子集搜索与迭代优化过程。

最后，在独立测试集上验证模型性能：在 512×512 像素图像上，单次推理时间仅 8 - 12 ms（约 80 - 120 FPS），较传统 DIC 方法（0.8 - 1.2 秒 / 帧）提升两个数量级；平均预测率达到 70% - 80%，其中均匀变形阶段预测率稳定在 78% - 82%，颈缩大

变形阶段仍保持在 70% - 75%，验证了模型对复杂变形与散斑退化的鲁棒性。消融实验表明，Transformer 瓶颈层贡献了约 7.8 个百分点的预测率提升，而光流网格标签生成策略相比传统子集匹配标签提升了约 15 个百分点。

综上，本文方法在保持亚像素级测量精度的同时实现了毫秒级实时推理，对散斑质量退化及噪声干扰均表现出强鲁棒性，为材料力学试验、结构健康监测及工业在线检测提供了一种高效、可靠的全场变形测量新手段。

参考文献

[1]Boukhtache S, Abdelouahab K, Berry F, et al. When Deep Learning Meets Digital Image Correlation [J]. Optics and Lasers in Engineering, 2021, 136: 106308.

[2]Sur F, Blaysat B, Gr é diac M. On biases in displacement estimation for image registration, with a focus on photomechanics [J]. Journal of Mathematical Imaging and Vision, 2021, 63(7): 777-806.

[3]Wang B, Pan B. Subset-based local vs. finite element-based global digital image correlation: A comparison study [J]. Theoretical and Applied Mechanics Letters, 2016, 6(5): 200-8.

[4]Duan X, Xu H, Dong R, et al. Digital image correlation based on convolutional neural networks [J]. Optics and Lasers in Engineering, 2023, 160: 107234.

[5]申海艇, 蒋招绣, 王贝壳, 等. 基于超高速相机的数字图像相关性全场应变分析在 SHTB 实验中的应用 [J]. 爆炸与冲击, 2017, 37(1): 15-20.

[6]潘兵, 谢惠民. 数字图像相关中基于位移场局部最小二乘拟合的全场应变测量 [J]. 光学学报, 2007, (11): 1980-1986.

[7]潘兵, 吴大方, 夏勇. 数字图像相关方法中散斑图的质量评价研究 [J]. 实验力学, 2010, 25(2): 120-129.

[8]陈佳, 胡浩博, 何儒汉, 等. 基于注意力机制的增强特征描述子 [J]. 计算机工程, 2021, 47(5): 260-6.