



Research  
High Performance Structures: Building Structures and Materials—Review

## 基于计算机视觉的民用基础设施的检查与监测研究进展

Billie F. Spencer Jr.<sup>a,\*</sup>, Vedhus Hoskere<sup>a,b</sup>, Yasutaka Narazaki<sup>a</sup>

<sup>a</sup> Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801-2352, USA

<sup>b</sup> Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801-2302, USA

### ARTICLE INFO

#### Article history:

Received 3 August 2018

Revised 14 October 2018

Accepted 15 November 2018

Available online 7 March 2019

#### 关键词

结构检查和监测

人工智能

计算机视觉

机器学习

光流

### 摘要

计算机视觉技术与远程摄像机和无人机(UAVs)的采集相结合,为民用基础设施状况评估提供了前景良好的非接触式解决方案。这种系统的最终目标是自动且稳健地将图像或视频数据转换为可操作的信息。本文概述了将计算机视觉技术应用于民用基础设施状态评估的最新进展。特别介绍了计算机视觉、机器学习和结构工程领域的相关研究。评估工作分为两类:检查应用和监测应用。检查应用包括识别环境,如结构构件,表征局部和全部的可见损坏,以及检测参考图像的变化。监测应用包括应变和位移的静态测量,以及模态分析的位移动态测量。最后,文章指出了为实现基于自动化视觉的民用基础设施和监测目标而持续存在的一些关键挑战,以及为解决这些挑战而正在进行的工作。

© 2019 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. 引言

许多为当今社会服务的重要基础设施,包括桥梁、水坝、高速公路、生命线系统和建筑物,都是几十年前建成的,远远超过了其自身的设计寿命。例如,美国土木工程师协会发布的《2017年基础设施报告》显示,美国有超过56 000座桥梁存在结构缺陷,需要花费1230亿美元的巨额资金来修复[1]。修复工作的经济意义需要通过仔细了解基础设施的现状来确定系统的优先次序。

民用基础设施状况评估是利用检查和(或)监测过程所获得的信息进行的。评估民用基础设施状况的传统技术通常包括目视检查,是由经过培训的人类检查员结合相关决策标准(如ATC-20[2]、国家桥梁检查标准[3])

进行的。然而,这种检查耗时、费力、成本高,又有一定危险性(图1)[4]。监测可用于通过测量物理量[如加速度、应变和(或)位移]来定量了解结构的现状;这些方法可以实时且连续地观察结构完整性,目的是增强安全性和可靠性,并降低维护和检查成本[5–8]。虽然这些方法可以产生可靠的数据,但它们通常具有有限的空间分辨率或者需要安装密集的传感器阵列。另一个问题是,一旦安装密集的传感器阵列,传感器的访问将受到限制,这使得常规系统维护具有挑战性。如果仅需要偶尔监测,则接触式传感器的安装是困难且耗时的。为了解决其中的一些问题,必须开发并测试改良的检查和监测方法,减少人为干预,降低成本和提高空间分辨率,以推进和实现自动化民用基础设施状况评估带来的全部益处。

\* Corresponding author.

E-mail address: [bfs@illinois.edu](mailto:bfs@illinois.edu)



图1. 美国陆军工程兵团的检查员从闸门垂降，检查表面是否有损坏。

计算机视觉技术已被公认为土木工程领域改进检查和监测方法的关键组成部分。图像和视频是计算机视觉技术分析数据的两种主要模式。图像捕捉的视觉信息与人类检查员获取的信息相似。由于这一相似性，可预知计算机实施的结构检查类似于人类检查员进行的目视检查。此外，图像可以以非接触方式对来自整个视野的信息进行编码，这有可能解决使用接触式传感器监测所面临的难题。视频是一组图像序列，从由多个视图收集图像时环境的同化，到使用高采样率时结构的动态响应，额外的时间维度为检查和监测应用提供了重要信息。土木工程界的大量研究都集中于开发和调整计算机视觉技术以完成检查和监测任务。此外，这种基于视觉的方法与摄像机和无人机（UAVs）配合使用，为民用基础设施状况评估提供了快速、自动化的检查和监测潜力。

本文对近年来基于视觉的民用基础设施状况评估的研究进行了综述。为了将本文所述的研究以适当的技术观点展开，第2节首先讲述了计算机视觉研究简史。第3节详细回顾了近年来关于计算机视觉技术在民用基础设施评估的检查应用中的一些成就。第4节重点介绍了监测应用。第5节概述了实现自动化结构检查和监测所面临的挑战。第6节讨论了作者为实现自动化检查目标正在进行的工作。第7节阐述了本文的结论。

## 2. 计算机视觉研究简史

计算机视觉是一个跨学科的科学领域，涉及从图像数据中自动提取有用的信息，以便定性或定量地理解或表示潜在的物理世界。计算机视觉方法可以使人类视觉皮层的任务自动化。最早开始尝试应用计算机视觉方法是在20世纪60年代，并试图利用边缘和原始形状（如

盒子）提取物体的形状信息[9]。随着图像模式不同表示方式的发展，计算机视觉方法开始考虑更为复杂的感知问题。最受关注的是光学字符识别（OCR），因为美国邮政总局[10]、车牌识别[11]等部门为了提高自动化程度，需要识别任何字体的字符和数字。人脸识别也是一个非常活跃的研究领域，通过使用手工制作或学习过滤器，在特征空间中对输入图像进行评估，以检测代表人脸的图案[12,13]。近年来，由于对监测和交通监测需求的增加，其他目标检测问题，如行人检测和车辆检测已开始出现显著改善（如参考文献[14]）。计算机视觉技术也被应用于体育广播，如球的跟踪和虚拟回放[15]。

计算机视觉技术的最新进展主要通过使用人工神经网络（ANNs）和卷积神经网络（CNNs）的端到端学习来推动。在ANNs和CNNs中，复杂的数据输入-输出关系由一个参数化的非线性函数来近似，该函数使用节点单元定义[16]。每个ANN节点的输出由下式计算：

$$y_n = \sigma_n(\mathbf{w}_n^T \mathbf{x}_n + \mathbf{b}_n) \quad (1)$$

式中， $\mathbf{x}_n$ 为节点 $n$ 的输入向量； $y_n$ 是该节点的输出标量； $\mathbf{w}_n$ 和 $\mathbf{b}_n$ 分别为权向量和偏置参数； $\sigma_n$ 是一个非线性激活函数，如S型函数和整流器（整流线性单元或ReLU[17]）。类似地，CNNs每个节点都应用卷积，然后得到一个非线性激活函数：

$$y_n = \sigma_n(W_n * \mathbf{x}_n + \mathbf{b}_n) \quad (2)$$

式中， $*$ 表示卷积； $W_n$ 为卷积核。CNNs的最后一层通常是全连接层（FCL），它与输出有紧密的连接，类似于ANN的层。CNN对于图像和视频数据尤其有效，因为使用CNNs进行识别，在参数数量有限的情况下，对翻译是稳健的。通过增加相互连接的节点数量，可以实现输入-输出关系的任意复杂参数化[如每层有多个隐含层和（或）多个节点的多层感知器，深度卷积神经网络（DCNNs）等]。使用一些输入和输出数据（训练数据）对ANNs/CNNs的参数进行优化（如参考文献[18,19]）。

这些算法在构建高度复杂的视觉问题的感知系统方面取得了显著成功。CNNs对修订后的美国国家标准与技术研究院（NIST）手写数字分类问题的准确率达到99.5%以上[图2（a）][20]。此外，在1000类ImageNet分类问题上，最先进的CNN体系结构的前五大错误率还不到5%（实际分类不标记前5个分类得分的数据比率）[21][图2（b）][22]。

CNNs的作用并不限于图像分类（如推断每个图像的单个标签）。DCNN使用了多个非线性滤波器并计算滤波器响应映射（如图3所示，称为“特征映射”）。DCNN不需要同时使用所有滤波器响应来获得每个图像类别（图3中的上部流程），而是可以单独使用映射中每个位置的滤波器响应来提取关于对象类别及其位置的信息。通过使用特征映射，语义分割算法可为图像的每个像素赋予适当的标签[23–26]。目标检测算法通常通过绘制研究对象的边界框，使用特征映射来检测和定位感兴趣的对象[27–31]。实例分割算法[32–34]进一步处理特征映射以区分对象的每个实例（如为图像中的每个人赋予单独的标签，而不是赋予所有人相同的标签）。在处理视频数据时，为便于分割，还可以利用附加的时间信息来进行时空分析[35–37]。

监督学习技术的弱点是需要高质量标记数据（如已经识别了对象的图像）用于训练目的。尽管已经创建了许多软件应用程序来帮助简化标记过程（如参考文献[38,39]），但手动标记仍然是一项非常繁琐的工作。所以，针对目标检测和定位任务，本文提出了一种弱监督训练方法，该方法不需要对图像进行像素级或目标级标记[40,41]；同时，CNN被训练成智能图像标签以获得对

象类别和图像中的近似位置。

无监督学习技术通过识别观测数据中的潜在概率结构，进一步减少了对标记数据的需求。例如，聚类算法（如 $k$ -均值算法[11]）假设数据（如图像块）由多个数据源（如不同的材质类型）生成，并基于最大似然法（ML）将每个数据样本分配给其中一个数据源。例如，DeGol等[42]使用 $k$ -均值算法对成像表面进行材料识别。通过将参数化概率模型拟合到观测数据[如高斯混合模型（GMM）[16]，玻尔兹曼机[43,44]]，可以提取更为复杂的概率结构。在图像处理背景下，研究人员对基于CNN的无监督学习体系结构进行了积极的研究，如自动编码器[45–47]和生成式对抗网络（GANs）[48–50]。这些方法可以自动学习输入图像的压缩表示和（或）从压缩图像恢复/生成的过程，而不需要手动标记数据。参考文献[51]对不同的有监督和无监督学习算法进行了全面而简明的回顾。

另一组算法，即光流法，促进了计算机视觉和人工智能（AI）在许多应用领域的重大进步。光流通过两个图像帧之间的像素对应来估计运动场。光流法的算法主要有四类：①微分法，②区域匹配法，③能量法和④基于相位的方法。具体细节和参数见参考文献[52]。从视频压缩[53]到视频分割[54]、运动放大[55]以及基于视觉的UAV导航[56]，光流在视频数据处理中有着广泛的应用。

随着这些技术的进步，计算机视觉技术已经被用于实现各种各样的前沿应用。例如，通过使用计算机视觉技术（图4）[57–59]，自动驾驶汽车可以识别和应对驾驶过程中可能遇到的潜在风险。准确的人脸识别算法增强了社交媒体[60]的能力，也被用于监控应用（如机场执法[61]）。其他成功的应用包括城市自动测绘[62]和医学成像增强[63]。计算机视觉技术在许多领域的显著改进和成功应用，为学者们开发土木工程问题的计算机视

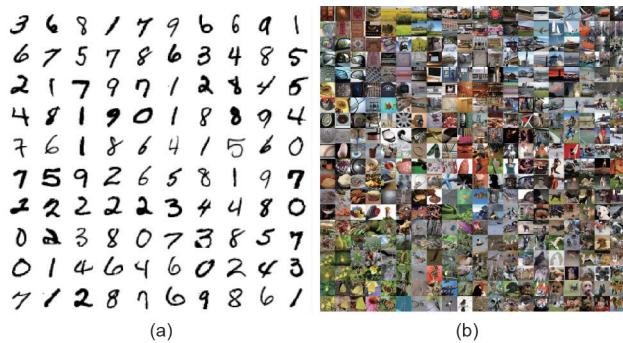


图2. 热门图像分类数据集。(a) MNIST数据集中的示例图像[20]；(b) ImageNet示例图像，通过 $t$ 分布随机邻域嵌入（ $t$ -SNE）进行可视化[22]。

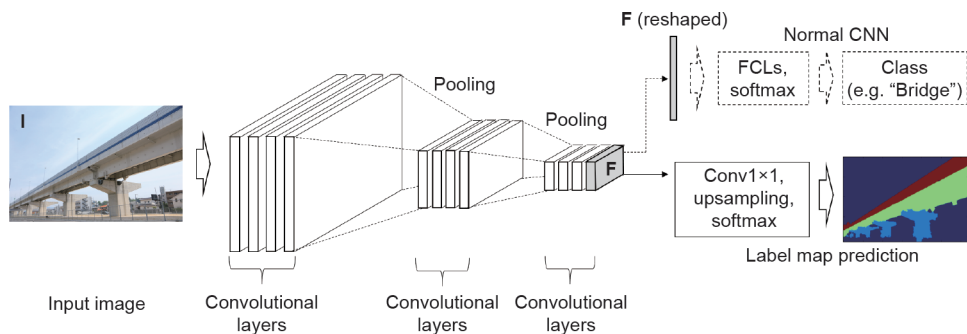


图3. 全卷积神经网络（FCNs）[23]。

觉解决方案提供了越来越多的动力。事实上，使用计算机视觉是改善民用基础设施监测和检查的一个常规步骤。以下各节将以这段简史为背景，介绍为适应和进一步发展用于检查和监测民用基础设施的计算机视觉技术而进行的研究工作。

### 3. 检测应用

研究人员时常在构想一个自动化的检测框架，它由两个主要步骤组成：①利用UAV远程获取自动化数据；②利用计算机视觉技术进行数据检查和检查。智能UAV已不再遥不可及，而且过去几年UAV工业的快速增长已使UAV成为数据采集的可行选择。事实上，美国的一些联邦和州机构以及其他研究机构正在部署UAV（如明尼苏达州交通部[64,65]、佛罗里达州交通部[66]、佛罗里达州大学[67]、密歇根州交通部[68]和南达科他州立大学[69]）。这些工作主要集中于拍照和录视频，工程师用它们进行现场评估或者随后的虚拟测试。然而，要将图像或视频数据自动而稳定地转换为可执行信息还有一定难度。为了实现这一目标，以下第一部分简述了关于损伤探测的文献，第二部分回顾了结构组件的识别，第三部分简述了结合这两个方面的一次演示：损伤探测和结构水平上的相容性。

#### 3.1. 损伤探测

自动化探伤在任何自动或半自动检测系统中都是至关重要的一个环节。当用结构体表面损伤部分和未损伤部分的像素比来表示时，结构体影像中的缺陷就很难被发现。因此，对可见缺陷进行高精度和记忆性探测是一项十分困难的工作。由于存在类似损伤的特征，这个问题更加复杂（如类似凹槽的深色边缘可能会被误认作裂



图4. Waymo的自动驾驶汽车系统[58,59]。†

缝)。如下所述，目前人们已经做了大量研究，致力于开发可靠地识别不同视觉缺陷的方法和技术，包括混凝土裂缝、剥落与脱层，疲劳裂纹，钢筋锈蚀，沥青裂缝。以下讨论了三种不同的探伤途径：①启发式特征提取法；②基于深度学习的损伤探测；③变化检测。

#### 3.1.1. 启发式特征提取法

研究人员为使用图像数据探伤制定了不同的启发式方法。原则上，这些方法在运作时会针对特定的损伤类型给手工过滤器的输出加上一个阈值或者机器学习分类器。本章节描述了几种关键的损伤类型，并针对这些损伤已开发出相应的启发式特征提取法。

(1) 混凝土裂缝。基于视觉的损伤探测，其早期工作大多集中于基于启发式滤波器的混凝土裂缝的识别（如参考文献[70–80]）。边缘探测滤波器是第一种用于损伤探测的启发式方法（如参考文献[70]）。关于该方法的早期调查可以在参考文献[71]中找到。Jahanshahi和Masri[72]根据形态特征，结合分类器（神经网络和支持向量机）确定了不同厚度层中的裂缝。本研究的结果如图5[72,81]所示，其中第一列显示本研究中使用的原始图像，后续各列显示了应用Bottom Hat方法、Canny方法和参考文献[72]中的算法得出的结果。本文还提出了一种通过识别裂纹中心线和计算裂纹边缘距离来量化裂纹厚度的方法。Nishikawa等[74]提出了用于裂纹检测和性能评估的多序列图像滤波。其他研究人员也开发了评估混凝土裂缝特性的方法。Liu等[79]提出了一种通过图像自适应处理自动进行裂缝评估的方法，该方法利用中值滤波器将裂缝的骨架和边缘分离。在参考文献[81,80]中，深度与三维（3D）信息也被用来进行定量损伤评估。Erkal和Hajjar[82]开发并评估了聚类处理技术，利用基于表面法线的损伤检测对彩色激光扫描数据中的裂纹、腐蚀、破裂和剥落等缺陷进行自动分类。在本文讨论的许多方法中，二值化是管道裂纹检测中常用的一种方法。Kim等[83]比较了几种不同的二值化方法。这些方法已被应用在多种民用基础设施中，包括桥梁（如参考文献[84,85]）、隧道衬砌（如参考文献[76]）和震后建筑评估（如参考文献[86]）。

(2) 混凝土剥落。本文还提出了识别混凝土中其他缺陷的方法，如剥落法。Adhikari等[87]采用了一种与桥梁状况指数相结合的新型正交变换法来量化退化过

† 所有图像版权与 Waymo 公司图片的使用均符合“《美国法典》第 17 条第 107 款对专有权利的限制：合理使用。”

程，随后再映射到状况等级中。作者对其数据集进行的剥落探测能够达到85%的合理精确度，但是无法处理裂缝和剥落同时存在的情况。Paal等[88]采用分割、模板匹配和形态预处理相结合的方法进行层裂检测和混凝土柱评估。

(3) 钢材的疲劳裂纹。疲劳裂纹是钢桥面板的一个重大问题，因为它们会显著缩短钢结构的寿命。然而，关于民用基础设施钢材疲劳裂纹探测的研究却相当有限。Yeum和Dyke[89]在一根钢梁上手动制造了一些损伤去模拟疲劳裂纹（图6）。然后，他们采用目标检测和过滤技术相结合的区域定位方式确定了疲劳裂纹状缺陷。他们做了一个有意思且有用的假设：疲劳裂纹通常围绕着螺栓孔延展；然而这一假设对其他焊接而成的钢结构主要部件来说可能并不合理，如人字门等导航基础设施[90]。Jahanshahi等[91]提出了一种区域成长法，用于分割核反应堆内部零件的细小裂纹。

(4) 钢铁腐蚀。研究人员现在使用纹理、光谱和色彩信息来识别腐蚀。Ghanta等[92]提出利用小波特征和主成分分析对图像中的腐蚀进行百分比估计。Jahanshahi和Masri[93]对基于小波的腐蚀算法性能进行了参数化评估。有研究者提出并评估了使用纹理和色彩信息的方法（如参考文献[94,95]）。也有人提议用机械化和基于智能手机的维护系统的自动算法来实施图像化腐蚀检测（如参考文献[96,97]）。在参考文献[98]中有一项关于利用计算机视觉进行腐蚀检测方法的调查。

(5) 沥青缺陷。使用启发式特征提取技术检测和评估沥青路面裂缝和缺陷的方法很多[99-105]。Hu和

Zhao[101]采用了一种局部二元模式（LBP）算法来识别路面裂缝。Salman等[100]提议使用Gabor滤波器。Koch和Brilakis[99]使用直方图阈值法来自动检测路面上的坑槽。除了RGB数据外（RGB指三种颜色通道，分别代表红、绿、蓝光波长），深度数据也被用于道路状况评定。例如，Chen等[106]指出，他们使用了一种廉价的

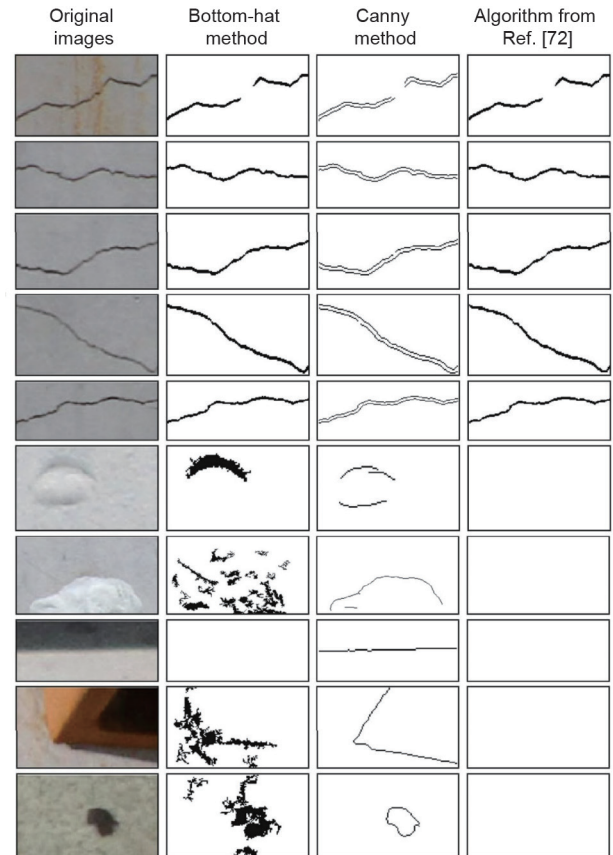


图5. 参考文献[81]中实施的不同探伤方法的比较。

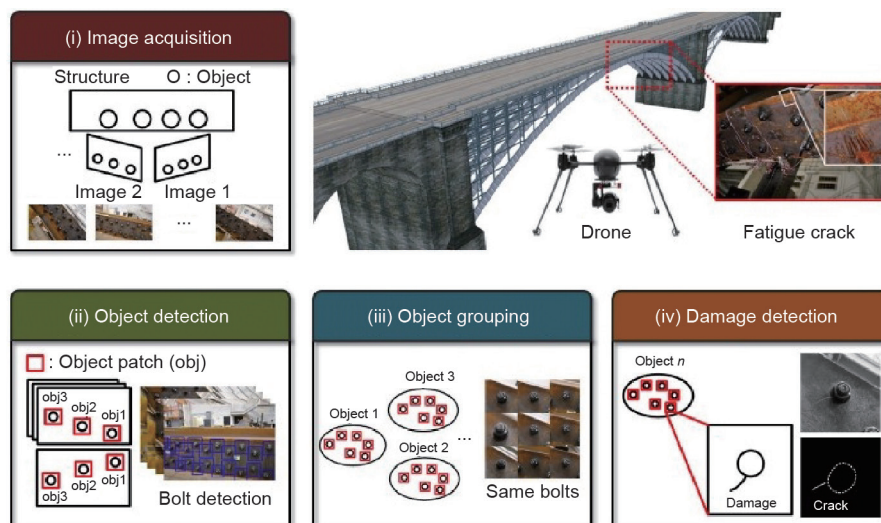


图6. 参考文献[89]中用于桥梁检测的基于视觉的自动化裂缝探测。

RGB-D传感器 (Microsoft Kinect) 来探测、量化、定位路面缺陷。有关沥青缺陷检测方法的详细回顾, 请参见参考文献[107]。

为了进一步研究这些缺陷的识别方法, Koch等[108]对2015年之前开发的计算机视觉缺陷检测技术进行了全面回顾, 并根据其应用的结构进行了分类。

### 3.1.2. 基于深度学习的损伤探测

迄今为止, 我们所讨论的研究和技术可分为两类: 利用机器学习技术, 或者依赖于启发式特征和分类器的组合。然而实际上, 此类技术在自动化结构检测环境中的应用还很有限, 因为这些技术并未采用缺陷所在区域附近的信息, 如材料性质或结构构件。这些基于启发式过滤的技术需要根据监测目标结构的外观来进行手动或半自动调整。真实情况变化多端, 要想手工制定一套适用于一般情形的通用算法非常困难。最近, 计算机视觉深度学习[21,51]在一些领域取得了成功, 如常规的图像分类[109]、自动传输系统[57]和医学成像[63], 推动了其在民用基础设施检查和监测中的应用。深度学习极大地扩展了基于视觉的传统的损伤检测能力和稳健性, 用于从裂纹、剥落到腐蚀等多种视觉缺陷的检测。目前人们已经研究了几种不同的探测方法, 包括①图像分类法, ②目标检测或区域提议法, 以及③语义分割法。以下是对这些应用的讨论。

(1) 图像分类法。CNNs可用于对钢板层[110]、沥青路面[111]、混凝土表面[112]裂缝的探测, 并且在所有情况下都非常精确。Kim等[113]提出了一种分类框架, 利用CNN和加速稳健特征 (SURF) 识别类裂纹模板中的裂缝, 并且利用图像二值化确定像素位置。像Alexnet这样的结构已经针对裂纹检测进行了微调[114,115], 而GoogleNet也同样针对剥落进行了微调[116]。Atha和Jahanshahi[117]针对腐蚀探测分析评估了几种深度学习技术, Chen和Jahanshahi[118]建议利用朴素贝叶斯数据与CNN相结合来进行裂纹探测。为了简化检测过程, Yeum[119]利用CNNs提取了高速公路桁架结构的重要部位。

Xu等[110,120]利用深度学习神经网络系统地研究了长跨桥钢面疲劳裂纹检测, 包括一台受限的玻尔兹曼机和融合CNN。在内场测试的复杂背景下, 这种新型的融合CNN能够精确识别多种尺度下的微小裂缝。Maguire等[121]整理了一套用于机器学习应用的混凝土

裂缝图像数据集, 其包含56 000张图像, 分为有裂缝和无裂缝两类。

Bao等[122]建议将DCNNs作为异常检测器, 帮助检查员从记录加速度数据的桥梁结构健康监测 (SHM) 系统中过滤异常数据。Dang等[123]利用UAV采集桥梁的特写照片, 然后将CNNs应用到图像块中, 自动检测结构损伤。

(2) 目标检测法。目标检测法最近已被用于损伤探测[112,119]。目标检测法是在损伤区域附近划分出一个边界框, 而不是对整张图片分类。Yeum等[124]使用具有CNN特征的区域 (R-CNNs) 在灾后场景中进行了层裂缝检测, 但结果 (59.39%的真阳性) 仍有提升空间。到目前为止所论述的方法仅适用于单个DT。与之相反, 深度学习法可以在极多类型的图像中学到可识别特征的一般表示。例如, DCNNs已成功解决了超过1000种类型的分类问题[21]。目前, 针对多种DTs的检测技术研究很有限。Cha等[125]研究了Faster R-CNN法, 该方法是由Ren等[30]提出的一种基于区域的方法, 用来识别包括混凝土裂缝和不同等级腐蚀和分层在内的多种损伤类型。

(3) 语义分割法。基于目标探测的方法不能准确地将其分离出的损伤轮廓描绘出来, 因为它们仅仅是为了与相关区域周围的矩形相适应。另一种分离图片中相关区域的方法称为语义分割法。更准确地说, 语义分割是将图像中的每一像素划分为不固定数量的类。其结果是在一张分割过的图像中, 每个部分被划分为一个特定类别。因此, 在进行损伤检测时, 语义分割法可以描绘出损伤的精确位置和形状。

Zhang等[126]提出了CrackNet, 它是一种有效针对路面裂缝的语义分割体系。对象实例分割技术Mask R-CNN[32]最近也被应用于裂缝、剥落、钢筋外露和风化的探测。尽管Mask R-CNN提供了像素级损伤描述, 它也只能分割“目标”所在区域的部分图像, 而不是对整个图像进行语义分割。

Hoskere等[127,128]对两种用于多种DTs常规定位和分类的方法进行了评估: ①多尺度像素DCNN, ②全卷积神经网络 (FCN)。如图7所示[127], 研究人员考虑了六种不同类型的DTs: 混凝土裂缝、混凝土剥落、钢筋外露、钢铁锈蚀、钢铁断裂与疲劳裂纹、沥青裂缝。参考文献[127]提出了一种新型的网络配置和数据集。数据集由各种结构的图像组成, 包括桥梁、建筑物、路

面、堤坝和实验室标本。该技术由两种网络的平行配置组合而成——DP网和DT网——能提高损伤探测的效率。该数据集中损伤规模的多样性证明了这一技术具有标度不变性。

### 3.1.3. 变化探测

当某一结构体必须接受定期检查时，首先应建立一条代表这一结构体的基线。在随后检测中，可将获取到的数据与该基线进行比对。在与基线的比对中，对结构体的任何新的视觉损伤都将显示为一种变化。识别并定位这些变化，将有助于减少处理从UAV检测中获取数据的工作量。由于任何损伤必定会显示为一种变化，因此，在实施损伤探测前，采用变化探测法可降低检测误差的数量。因为在这两个状态中都可能存在类似损伤的纹理。变化探测技术在计算机视觉中已有研究，从环境

监测到视频监控都有其应用。本小节中，我们研究了两种主要的变化探测方法：①点云变化探测，②图像变化探测。

(1) 点云变化探测。运动恢复结构 (SFM) 和多视角立体视觉 (MVS) [129] 是基于视觉的技术，它们可使结构体产生点云。在实施变化探测法之前，必须先建立一条点云基线。正如参考文献[130,131]所描述的，即使在桁架桥或堤坝这类复杂的民用基础设施中，这些点云的精确度也很高。后续的扫描将会被注册到云基线，校准将由迭代最近点 (ICP) 算法[132]代为执行。ICP算法已在MeshLab[133]和CloudCompare[134]等开源软件中得到了应用。校准之后，就可以开始执行变化探测的各个程序了。这些技术同时适用于激光扫描点云和从摄影测量中产生的点云。早期研究将云与云之间 (C2C) 的豪斯多夫距离作为在3D空间中识别变化的度量标准

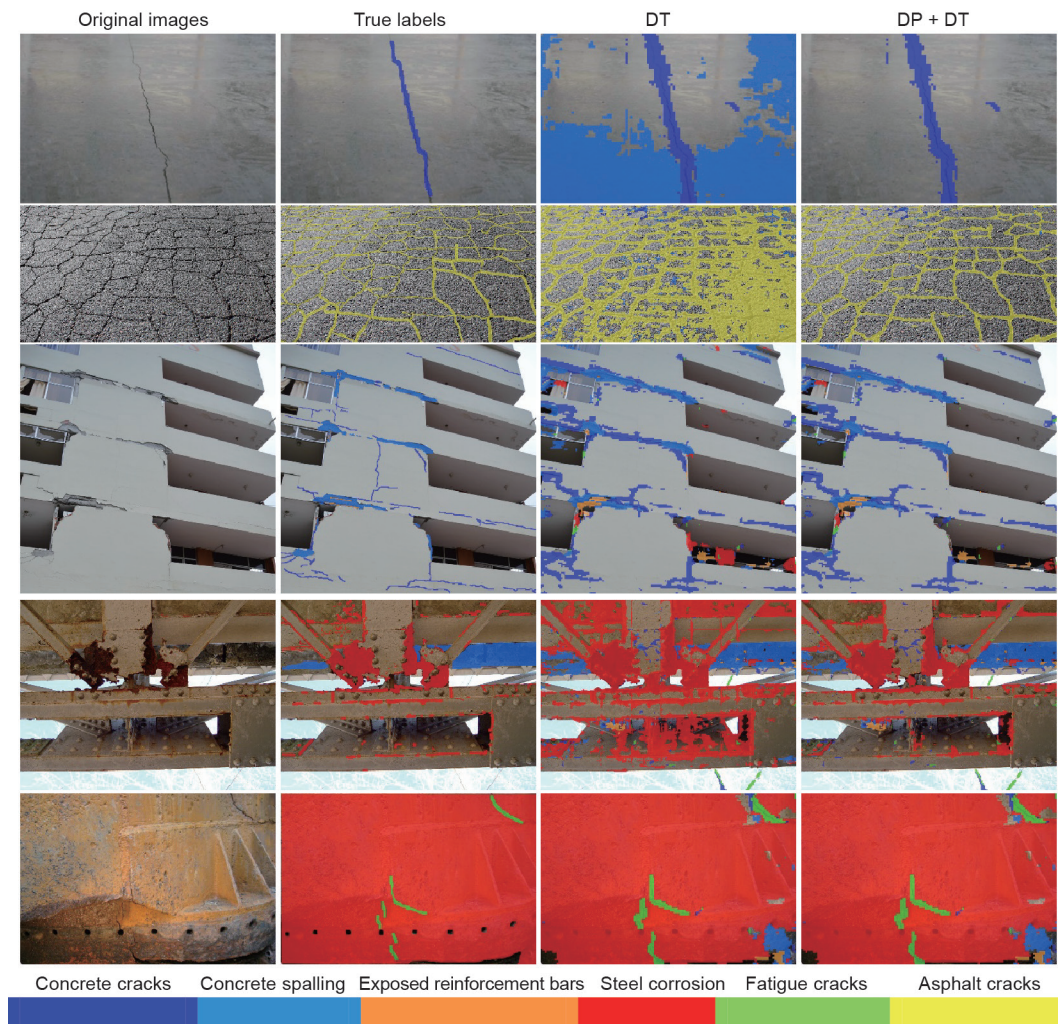


图7. 对多种结构性DTs进行基于深度学习的语义分割。

[135]。其他技术包括数字高程差分模型（DoD）、云-网格（C2M）法、多尺度模型到模型云的比较（M3C2）法。参考文献[136]中有关于这些技术的概述。

结合UAV数据采集，可将变化探测法用于民用基础设施。例如，Morgenthal和Hallerman[137]使用正射投影（aligned orthomosaics）对挡土墙内的变化进行人工识别，用于平面内变化；使用CloudCompare程序包进行C2C比较，用于平面外变化。Khaloo和Lattanzi[139]利用在不同色彩空间中的像素色调值来辅助探测一座重力坝的重要变化。Jafari等[140]提出了一种测量变形的新方法，即使用直接的逐点距离协同统计抽样将数据完整性最大化。点云变化探测的另一个有趣的应用是有限元模型更新。通过对实验室结构组件的两个点云的对比分析，Ghahremani等[141]用基于视觉的方法自动定位、识别并量化损伤；之后这些信息将用于更新这一组件的有限元模型。当点深度足以被识别的时候，就可以用点云变化探测法。在寻找不会引发足够的几何变化的可视变化时，可以利用图像变化探测法。

（2）基于图像的变化探测。在计算机视觉中，关于图像的变化探测是一个研究热点，这是由于其应用范围

十分宽泛[142]。遥感卫星图像是基于图像的变化探测最普遍的应用案例之一，应用范围从土地覆盖和土地利用检测延伸到损伤评估和灾害监测。参考文献[143]对高分辨率卫星图像的变化探测进行了深入评论。在进行变化探测之前，图像要先经过预处理，以便排除大气和辐射度等环境变量的影响，之后才会进行图像配准。与损伤检测类似，基于启发式和深度学习的技术以及基于点云和对象检测的技术都是可用的。参考文献[144]给出了关于这些方法的概述。

虽然遥感卫星图像可以让人们对城市规模的损伤有所了解，但对于私人建筑来说，这种图像的分辨率和视角却阻碍了有用信息的提取。对于来自UAV或地面车辆调查的图像，变化探测可作为损伤探测的先导，以帮助定位可能代表损伤的候选像素或区域。为此，Sakurada等[145]提出了一种方法，即从不同时间点拍下的多视角图像中，利用概率估计的场景深度探测户外场景的3D变化。CNNs也被用来识别城市场景的变化[146]。Stent等[147]提出利用CNN来识别隧道衬砌中的变化，然后再用集群的方法根据重要性将这些变化分级。图8展示的是Stent等[147]方案的原理图。

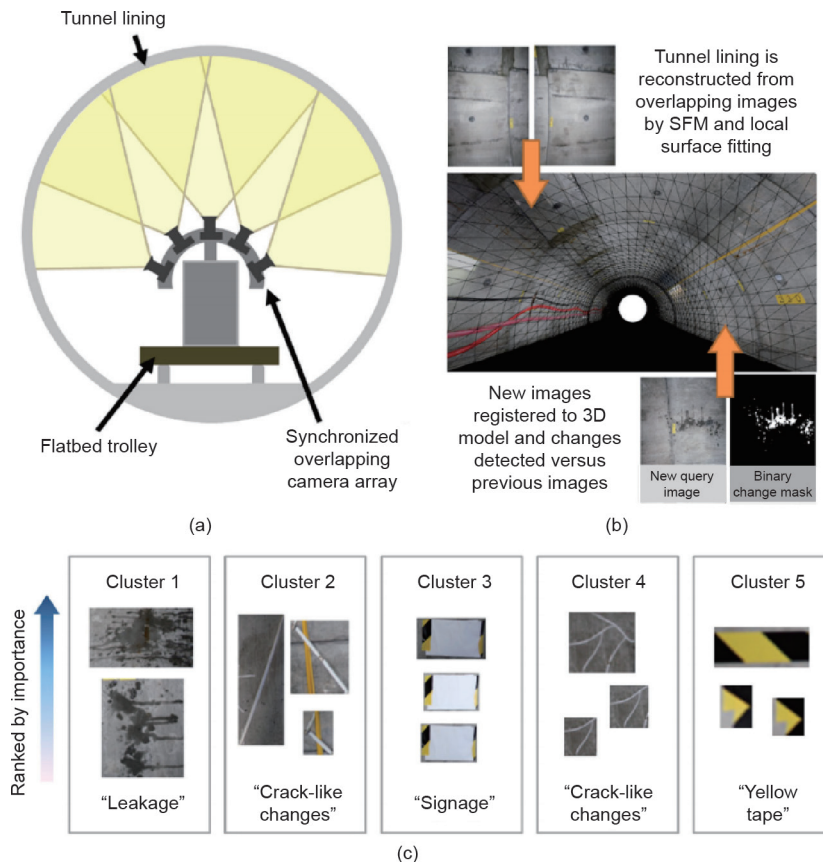


图8. 参考文献[147]提到的系统图解。(a) 数据捕获硬件；(b) 通过在重建参考模型中定位而探测到的变化；(c) 样本输出，其中检测到的变化根据外观进行了集群处理。



### 3.2. 建筑物结构构件识别

建筑物结构构件识别是对建筑物典型构件进行检测、定位和分类的过程，也是实现基础设施自动化检测的关键步骤。建筑物结构构件信息可为原始图片和3D点云数据添加图像和数据语义，这样的图像和数据语义能够帮助人们了解建筑物当前的状态，并能在现场环境中使容易出错的数据保持一致[148–150]。例如，通过对点云数据设定“柱子”标签，一个点集合可以被识别成单个的结构构件（竣工模型）。在检测施工进度的环境下，竣工模型的柱子可以与在设计阶段开发的3D模型的柱子(计划模型)相对应，从而可以为评估柱子的当前状态做参考。在评估过程中，可以忽略没有被标记“柱子”的点，因为这些点被认为来自不相关的事物或者错误的信息。从这个意义上来说，结构组件的信息是已竣工模型的基本属性之一，用于以有效和一致的方式表示结构的当前状态。

建筑物结构构件识别也为土建结构视觉损伤的自动化评估提供了强有力的支持信息。与竣工模型类似，通过删与建筑物结构互相关联的构件之外的对象上的类似损伤模式(如在树中检测到的裂纹属检测误差)，利用建筑物结构构件的信息可提升自动化损伤检测方法的一致性。此外，为了在大多数现行的结构检测准则中得出安全等级，需要对损伤和出现损伤的结构构件进行联合评估，所以结构构件信息对整个建筑物结构构件的安全评价很有必要（ATC-20[2]、国家桥梁检测标准[3]）。

在实现完全自主检测的过程中，结构构件识别有望成为机器人平台(如UAV)自主导航和数据采集算法的一个组成部分。根据机载摄像机识别的结构构件的类型和位置，自主机器人有望能够规划合适的导航路径和数据采集行为。虽然目前还没有实现结构构件检测的完全自动化，但是在农业领域已有基于视觉的周围环境识别的自动机器人的例子（如TerraSentia机器人[151]）。

#### 3.2.1. 利用图像数据的启发式结构构件识别

在早期的研究中，人们使用手工制作图像滤波器和图像启发式方法从图像中提取结构构件。例如，利用线段组识别图像中的钢筋混凝土（RC）柱子（图9）[152,153]。为了将柱子和其他不相关的线段组进行区分，该方法采用了一个阈值来选择具有预定长宽比范围的近似平行组。该研究的作者用此方法检测了20幅以柱子为主要拍摄对象的图像，从51个柱子中检测出38个，其中7个为检测误差。这种方法虽然简单，但却严重依

赖阈值，并且往往无法找到部分闭塞或相对较远的柱子。此外，在此方法中并没有对场景做进一步的了解，任何满足阈值的线段都会被识别为柱子。所以为了改进结果并减少检测误差，高层场景需要以不同的比例进行合并。

#### 3.2.2. 利用3D点云数据的结构构件识别

结构构件识别的另一个重要方案是利用可用的密集3D点云数据识别构件。针对使用密集3D点云数据进行结构构件识别，可以采用不同的分割和分类方法来执行。Xiong等[154]研究了一种自动化方法，可以将密集3D点云数据从空间转换为语义丰富的3D模型，该模型由平面墙壁、地板、天花板和矩形开口表示[该过程称为扫描-建筑信息模型（BIM）]。Perez等[155]采用高维特征（语义特征为193维，几何特征为553维）对室内空间进行结构和非结构构件识别。该方法利用提取特征所携带的丰富信息和使用条件随机场执行的后处理，能够准确地标注平面和复杂非平面表面，如图10所示[155]。Armeni等[156]提出了一种针对密集3D点云数据进行过滤、分割和分类的方法，并通过将整个建筑解析为平面构件来演示该方法。

Golarvar-Fard等[157]对基于图像的点云与激光扫描自动化性能检测技术进行了详细比较，包括3D重建、形状建模、生成可视化的准确性和可用性。通过比较发现，虽然基于图像的技术并不准确，但它们



图9. 钢筋混凝土柱子的识别结果[153]。

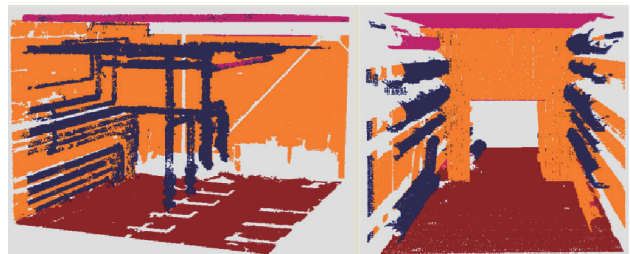


图10. Perez-perez等[155]利用密集3D点云数据进行的室内语义分割。

为可视化及丰富的语义信息的提取提供了巨大便利。Golparvar-Fard等[158]提出了一种自动监测3D建筑元素变化的方法。该方法将无序的照片集合与使用SFM的建筑信息建模相融合，然后对基于体素的场景进行量化。最近，Lu等[159]提出了一种方法，即通过自顶向下的方式从钢筋混凝土桥梁的点云中准确地检测出桥梁的四种构件类型。

本节讨论的3D方法的有效性取决于解决当前问题的可用数据。与图像数据相比，密集3D点云数据以其额外的维度携带更丰富的信息，能够识别形状复杂的结构构件和（或）识别定位精度要求较高的任务。另一方面，为了获得准确且密集的3D点云数据，需要对被检查结构的每个部分以足够的分辨率和重叠方式进行拍摄，这就需要增加数据收集的工作量。此外，离线后处理也是非常必要的，这对应用3D方法进行实时处理任务提出了挑战。对于这种情况，利用图像数据进行基于深度学习的结构构件识别是另一种执行结构构件识别任务的可行的方法。下一节将对此进行讨论。

### 3.2.3. 利用图像数据的基于深度学习的结构构件识别

近年来，基于机器学习的结构构件识别方法得到了广泛研究。图像分类是CNNs的主要应用之一，其中单个代表性标签是从输入图像中预估出来的。Yeum等[160]利用CNNs对某公路标志桁架结构的焊接接头候选图像块进行了分类，从而准确地识别出兴趣区域。Gao和Mosalam[161]使用CNNs把输入图像分为合适的结构构件和损伤的结构构件两类。然后，作者根据最后卷积层的输出结果推断出目标对象的粗略位置（弱监督学习；如图11[161]所示，用于结构构件识别结果）。目标检测算法也可用于结构构件识别。Liang[162]采用Faster

R-CNN算法，通过自动绘制桥梁组件周围的边界框对其进行检测和定位。

语义分割是解决结构构件识别问题的另一种可行途径[163–165]。语义分割算法不需要绘制边界框，也不需要每幅图像的标签来推断目标的大概位置，而需要输出与输入图像分辨率相同的标签映射。这对于精确检测、定位和分类复杂形状的结构组件尤其有效。为了得到与高层场景结构一致的高分辨率桥梁构件识别结果，Narazaki等[164]研究了三种不同配置的FCNs：①原始配置，即直接从输入图像预估标签映射；②平行配置，即根据高层场景类和平行运行的桥梁构件类的语义分割结构预估标签映射[图12（a）]；③序列配置，即根据场景分割结果和输入图像预估标签映射[图12（b）]。桥梁构件识别结果如图13所示。除了第三张和第七张图像（见图13中的Input image），所有的配置都能够识别结构构件，包括远距离的柱子或被部分遮挡的柱子。在非桥梁图像中可观察到显著的差异（图13[164]中最后两幅图像）。对于原始配置和平行配置，在建筑物和路面

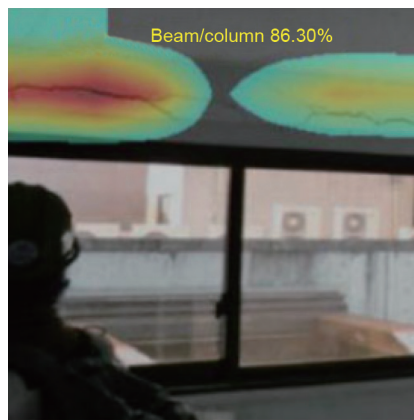


图11. 弱监督学习的结构构件识别结果[161]。

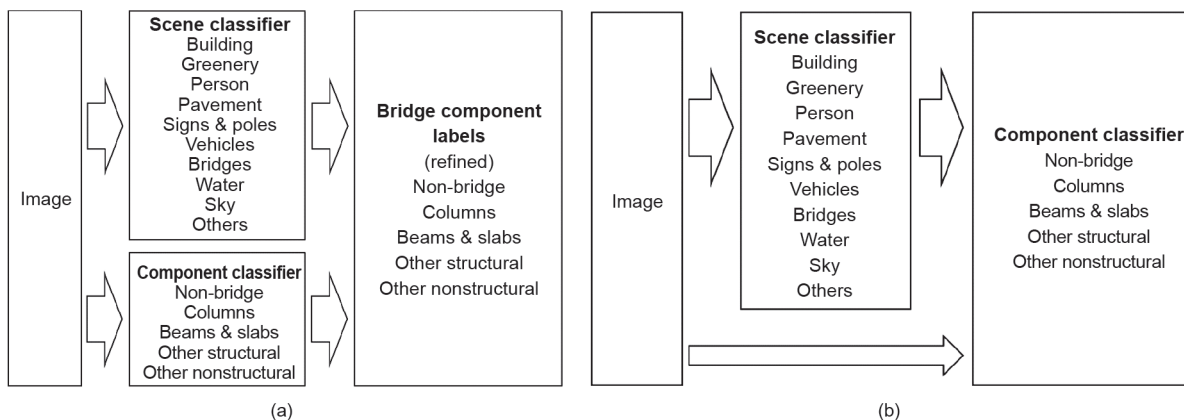


图12. 网络配置增强场景级一致性[164]。

像素中发现了检测误差。相反，在顺序配置的FCNs没有发现错误。（表1给出了非桥梁图像检测结果中的误差检测率）。因此，顺序配置能够有效地将高层场景一致性应用到桥梁构件识别中，以便提高复杂场景图像识别的鲁棒性。

### 3.3. 结构级一致性的损伤检测

结构构件和其损伤状态信息的结合对进行自动化评估至关重要。German等[166]提出了一种地震后快速评估建筑物的自动化框架。在这个框架中，对受损的建筑物内部情况进行视频拍摄，并在每一帧中搜索是否存在

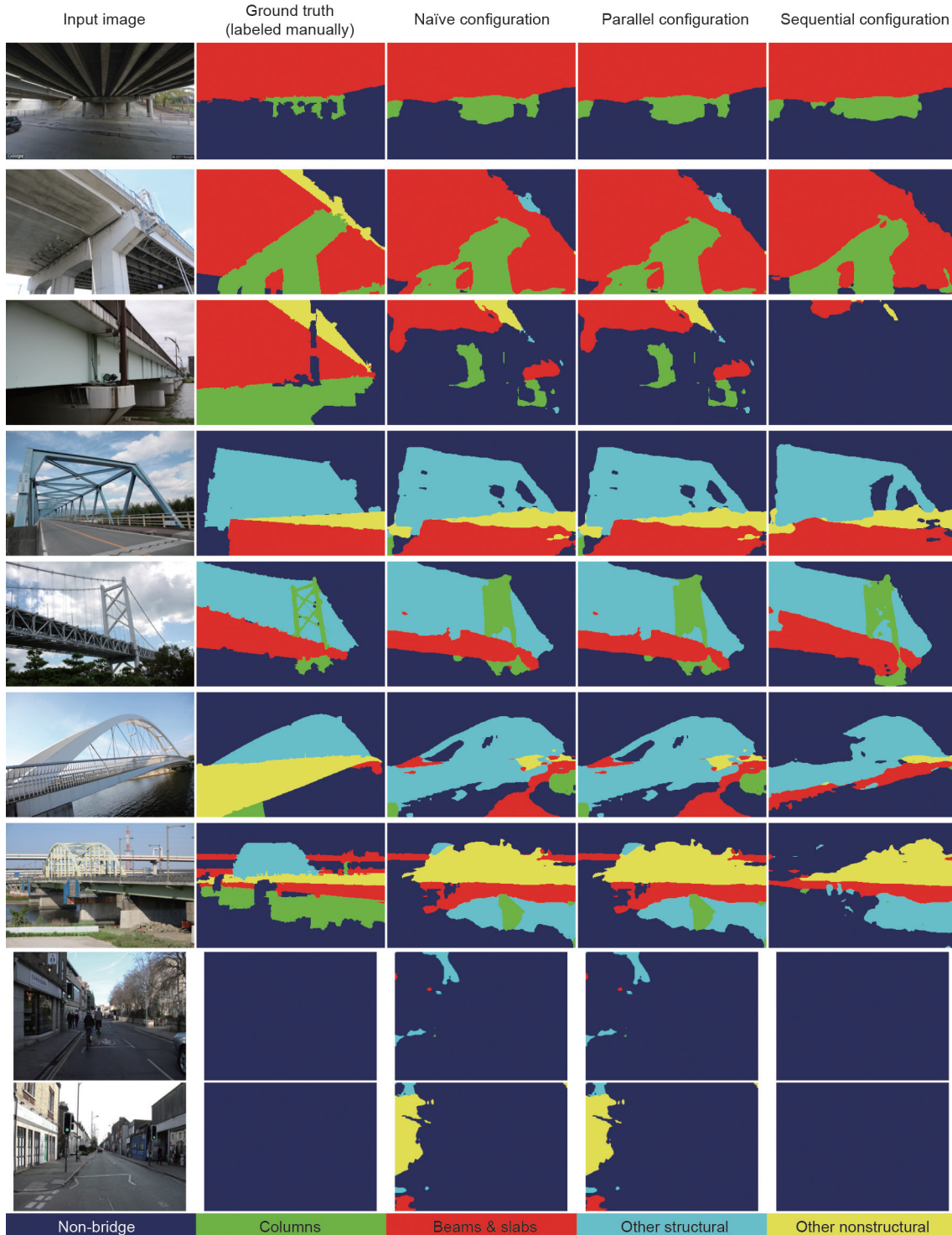


图13. 桥构件识别结果示例[164]。

柱子[152]，然后为每一个柱子分配一个损伤指数。损伤指数[88]是采用参考文献[73,167]提出的方法，根据裂缝、剥落和外露钢筋的位置和严重程度，将柱子的破坏模式分为剪切或弯曲破坏来估算的[73,167]。然后，人工记录建筑物的结构布局，用于查询易损性数据。该数据提供了物体易处于某种损坏状态概率的信息。

Anil等[168]确定了一些信息需求，以适当地表示地震后结构墙体的视觉损伤信息，并根据17个不同损伤敏感性的损伤参数将其分为5类。这些信息用来描述参考文献[169]中以BIM为基础的方法，帮助工程分析自动引入一些启发式方法，以结合强度分析和视觉损伤评估信息。Wei和Kasireddy[170]详细回顾了建筑和基础设施管理3D成像技术的现状及其面临的持续和紧急的挑战。

Hoskere等[128]利用FCNs对损伤进行划分，并构建构件的图像信息，用于生成类似于检测的语义信息。这个过程使用了三种不同的网络：一种用于场景和建筑物（SB）信息，一种用于识别DP，另一种用于识别DT。SB网络的平均准确率为88.8%，DP和DT联合网络的平均准确率为91.1%。这种方法能够成功地识别出损伤的位置和类型，也能识别出一些关于SB存在的场景。与以前的实验相比，这种方法适用于更普遍的环境。如图14所示的多幅图像定性结果，其中最右栏显示的是对准确检测和误报及漏报的评价。

### 3.4. 小结

损伤探测、变化探测和结构构件识别是实现建筑结构自动化检查的关键步骤。虽然建筑结构检查为评估基础设施状况提供了有价值的指标，但往往还需要对建筑结构响应进行更多的定量测量。为了实现建筑结构状态评估，还需要用基于视觉的技术对位移和应变等物理量进行测量。本文下一节将会介绍使用视觉技术的民用基础设施的监测应用。

## 4. 监测应用

监测的目的是通过测量加速度、应变和（或）位移

等物理量，定量了解民用基础设施的当前状态。监测工作通常使用有线或无线接触式传感器来完成[171–178]，尽管许多应用程序都可以使用接触式传感器有效地收集数据，但这些传感器的安装成本往往很高，维护起来也很困难。基于视觉的技术为非接触式方法提供了优势，克服了使用接触式传感器带来的一些问题。如第2节所述，能够执行测量任务的关键计算机视觉算法是光流算法，它能估算两个图像帧之间每个像素的平移运动[179]。光流算法是一种通用计算机视觉技术，它通过优化目标函数，如误差平方和（SSD）、归一化互相关（NCC）标准、全局代价函数[179]或局部和全局综合函数[180,181]，将参考图像中的像素与不同视角下同一场景的另一个图像的对应像素相关联。参考文献[182]对不同代价函数和优化算法的方法进行了比较。本节其余部分讨论了基于视觉的民用基础设施监测技术研究。本节主要分为两小节：静态应用和动态应用。

### 4.1. 静态应用

基于视觉技术的民用基础设施静态位移和应变的测量通常采用数字图像相关（DIC）技术进行。根据Sutton等[183]的研究，DIC是指“一种非接触式方法，它能获取物体的图像，以数字形式存储图像，并进行图像分析以提取全部形状、变形和（或）运动测量值。”（p.1）除了估算图像平面内的位移场，DIC算法还包括计算二维（2D）平面内应变场（2D DIC）、平面外位移和应变场（3D DIC）、体积测量（VDIC）的不同后处理步骤。目前已有高度可靠的商业DIC解决方案（如VIC-2D™[184]和GOM Correlate[185]）。有关一般DIC应用的详细介绍，请参阅参考文献[186,183]。

DIC方法已被应用于土木工程中位移和应变的测量。Hoult等[187]在单轴载荷下使用了钢样品，将结果与应变计测量结果进行比较，评估了2D DIC技术的性能（图15）。然后，研究人员提出了一种补偿平面外变形影响的方法。研究人员还使用钢和钢筋混凝土梁试样测试了2D DIC技术的性能，通过应变计得到了应变的理论值和应变测量数据[188]。在参考文献[189]中，以

表1 九类场景的误差检测率

Configuration	Scene classes								
	Building	Greenery	Person	Pavement	Signs & poles	Vehicles	Water	Sky	Others
Naïve (%)	18.60	5.80	3.70	3.10	6.80	4.60	9.90	0.90	13.20
Parallel (%)	16.90	5.20	3.10	2.70	5.80	3.90	8.90	0.80	11.80
Sequential (%)	0.48	0.08	0.03	0.05	0.21	0.18	0.81	0.03	0.50

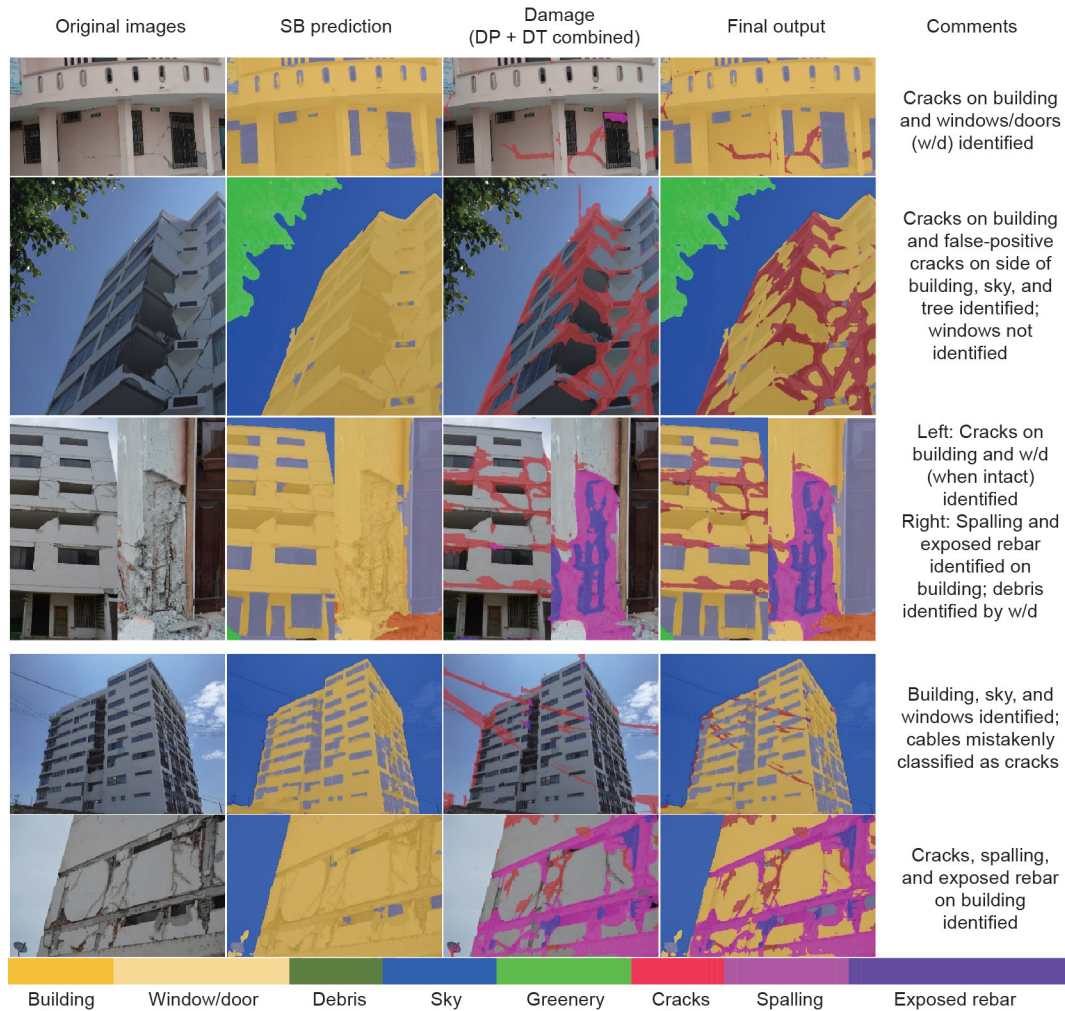


图14. 参考文献[128]中的定性结果。

3D DIC系统为参考，测量了实验室试件的静态位移。这些试验获得了位移的亚像素精度，而且应变估算值与应变计测量值和理论值一致。

DIC方法也被应用于民用建筑结构的位移和应变的现场测量。McCormick和Lord[190]采用2D DIC技术测量了静载4辆32 t卡车的高速公路桥面的垂直位移。Yoneyama等[191]使用2D DIC技术估算了负载一辆20 t卡车的桥梁的挠度。作者利用位移传感器的数据评估了有和没有人工模式的挠度测量的准确性。Yoneyama和Ueda[192]采用2D DIC技术测量了工作荷载下的桥梁挠度。Helfrick等[193]采用3D DIC技术进行了全场振动测量。Reagan[194]使用携带立体摄像机的UAV，将3D DIC技术应用于桥梁变形的长期监测。

DIC方法在土木工程领域的另一个具有前景的应用是裂缝映射，在此方法中，3D DIC被用于提取具有大应变特征的裂纹区域。Mahal等[195]成功地提取了RC

试样上的裂缝，Ghorbani等[196]将这种裂缝映射方法推广到了循环荷载作用下的全尺寸砌石墙试样中（图16）。所得到的裂缝图不仅对分析实验室测试结果有一定的参考价值，而且对增加结构检测的信息量也很有用。

#### 4.2. 动态应用

系统识别和模态分析是SHM的有力工具，能够为相关结构系统的动态特性提供有价值的信息。系统识别和其他与SHM相关的任务通常使用有线或无线加速度计来完成，因为这些传感器可靠并且安装方便[171–178]。与传统方法相比，基于视觉的技术提供了非接触式方法的优势。随着市场上低成本相机的普及和计算能力的提高，基于视频的方法已经成为结构位移测量的一种简便方法。目前，有几种算法可以实现位移提取，这些方法的原理是通过模板匹配，或者通过追踪恒定相位或与时间有关的强度轮廓来实现的[55,197,198]。光流方法已

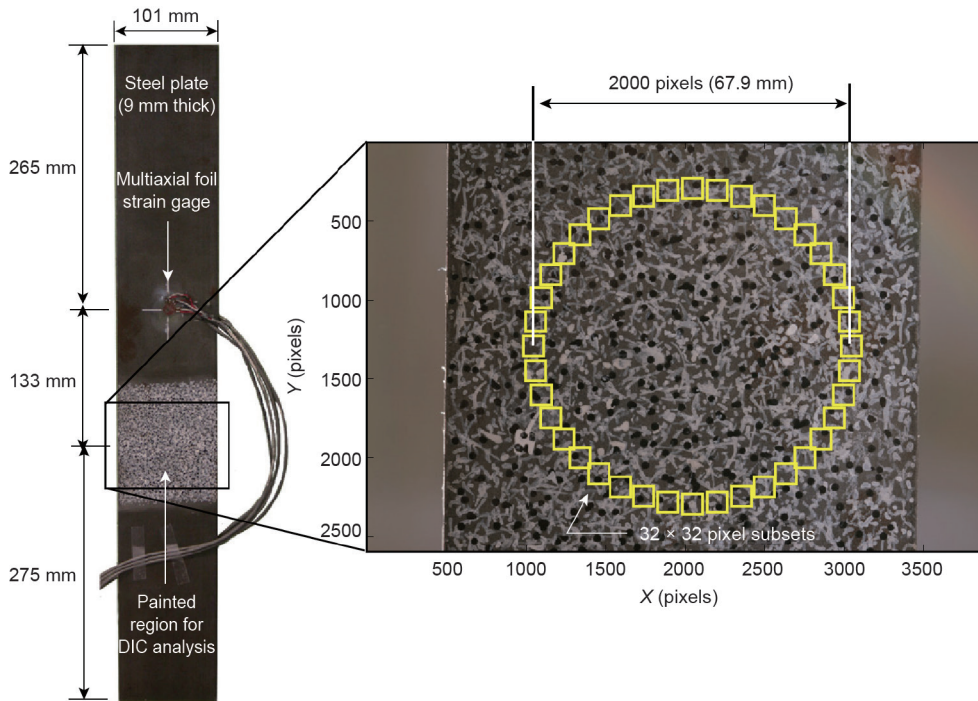


图15. Hoult等[187]进行单轴测试所使用的钢板试样。

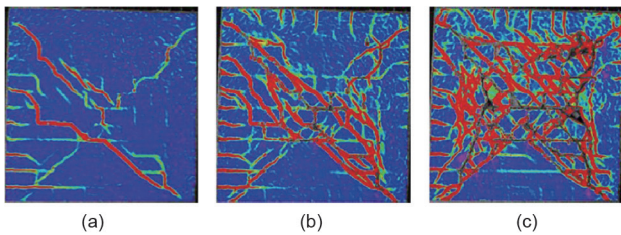


图16. 使用3D DIC技术制作的裂缝图[196]。(a) 第一次裂缝；(b) 最大载荷；(c) 极限状态。红色部分相当于 $+3000 \mu\text{m}\cdot\text{m}^{-1}$

被用于测量几个应用的动态和伪静态响应，包括系统识别、模态分析、模型更新和基于阈值的适用性变化的直接指示。为了获得更多的信息，Ye等[199]、Feng D和Feng M Q[200]对使用计算机视觉技术的动态监测应用进行了综述。这些研究通过实验室试验和实地验证提出和（或）评估了位移测量的不同算法，本节对这些研究进行了讨论。

#### 4.2.1. 实验室试验

早期用于动态监测的光流算法侧重于固有频率估算[201]和位移测量[202,203]。通常采用标记对关注点进行精细和精确的检测和追踪。Min等[204]设计了高对比度标记，它能辅助智能手机设备和长焦镜头测量位移，并在实验室测试中取得了很好的结果（图17）。

Dong等[205]提出了一个多点同步测量结构动态位移的方法。Celik等[206]评估了几种不同的基于视觉的

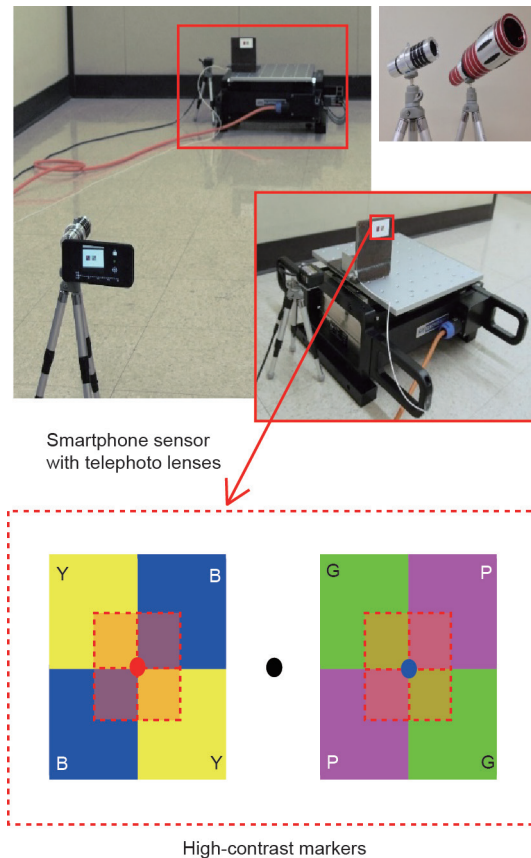


图17. Min等[204]提出了一种基于智能手机的位移测量系统，包括长焦镜头和高对比度标记。B: 蓝色；G: 绿色；P: 粉色；Y: 黄色。

技术，用以测量结构上的人体负荷。Lee等[207]提出了一种位移测量方法，该方法是为实地测试量身定制的，

且在强光下具有较强的鲁棒性。Park等[208]证明了基于视觉的图像与加速度数据融合在扩展动态范围和降低信号噪声方面的功效。视觉算法的应用目前已经扩展到了实验结构的系统识别中。Shumacher和Shariati[209]提出了虚拟视觉传感器的概念，利用虚拟视觉传感器可进行结构的模态分析。Yoon等[210]利用一个Kanade-Lucas-Tomasi (KLT) 追踪器识别了实验室规模的六层建筑模型(图18)。Ye等[211]在一个小尺寸模型的拱桥上进行了多点位移测量，并利用线性可变差动变压器(LVDTs)对测量结果进行了验证。Abdelbarr等[212]使用廉价的RGB-D传感器测量了3D动态位移。Ye等[213]在振动台上进行了一个研究，确定了影响基于视觉测量的系统性能的因素。Feng D和Feng M Q[214]利用上采样的互相关性实现了一种模板追踪法，获取了振动结构上的多点位移。研究人员还利用UAV捕获的视觉数据对实验结构进行了系统识别[215]。这些作者还提出了一种使用UAV在背景中结合静止坐标来测量动态位移的方法。

Wadhwa等[55]提出了一种运动放大技术，该技术通过具有微小变形的带通视频来提取和放大特定频率的运动。这个过程包括以多种比例分解视频，对每一比例视频应用一个滤波器，然后重新组合过滤后的空间频段。随后，在运动放大技术的启发下，研究人员发表了许多关于使用基于视觉的方法对结构进行全场模态分析的论文。Chen等[216]成功地将运动放大技术应用于实验结构工作振型的可视化(图19)。Cha等[217]使用基于相位的方法和无迹卡尔曼滤波器通过噪声位移测量法进行系统识别。Yang等[218]采用该方法将多尺度小波滤波器与复杂可控滤波器进行盲源分离，以获得实验样品的

全场模态。Yang等[219]提出了一种利用高空间分辨率模态模型和视频操作对结构响应进行高保真、逼真模拟的方法。

#### 4.2.2. 实地验证

在过去几年里，实验室基于视觉的振动测量技术的成功已经带来了许多实际应用。最常见的应用是测量全尺寸桥梁结构的位移[220–223]，包括测量桥面板、桁架和机库电缆等不同构件的位移。基于相位的方法也被用来估算天线塔的位移和频率，从而获得桁架桥结构的部分振型[224]。

一些研究人员利用视觉传感器测量了一种结构的多点位移。Yoon[225]使用摄像系统测量了列车通过时铁路桥梁的位移。如图20[225]所示，实测位移与以列车荷载为输入的有限元模型的预测值非常接近，差异主要是因为列车的速度不恒定。Mas等[226]通过对高速视频序列的分析，开发了一种同步多点测量振动频率的方法，并在一座钢制人行天桥上验证了他们的算法。

Chen等[227]研究了载荷估算的一个有趣应用，他们自动检测通过桥梁的车辆类型，并将这些车辆类型与从桥梁某一横截面的动态称重系统中得到的信息相结合，使用计算机视觉技术识别了车辆荷载在通过桥梁时在空间和时间上的分布。与以往只能在桥梁某一横截面测量车辆荷载的方法相比，该系统能够精确地测定整座桥梁的荷载。

利用计算机视觉技术进行结构的系统识别在一定程度上存在局限性；在单个视频帧内测量大型结构上的所有点，通常会导致像素分辨率不足，不能获取精确的结

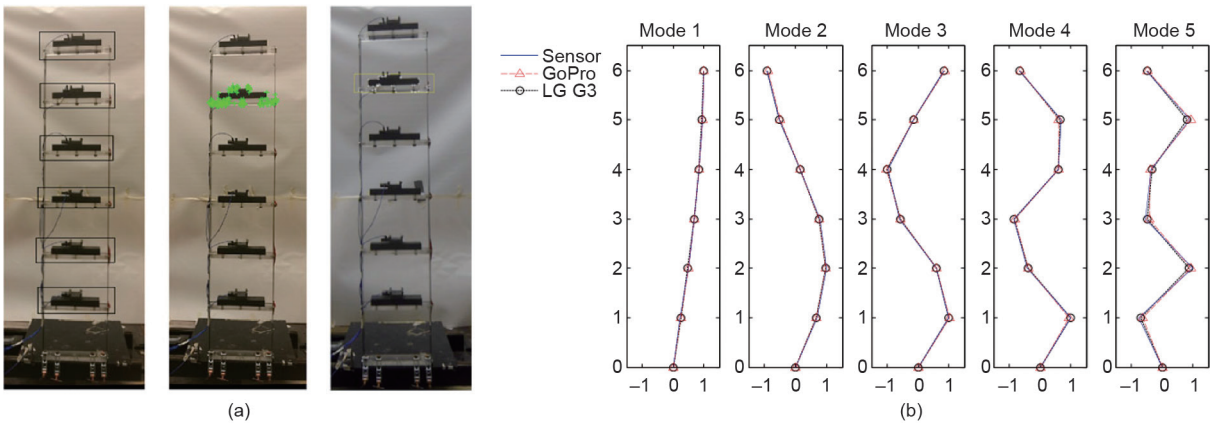


图18. 使用消费者级别相机进行基于视觉结构的无目标系统识别方法[210]。(a) 目标追踪截图；(b) 从不同传感器中提取的模态形状。GoPro和LG G3是测试中所使用的相机。

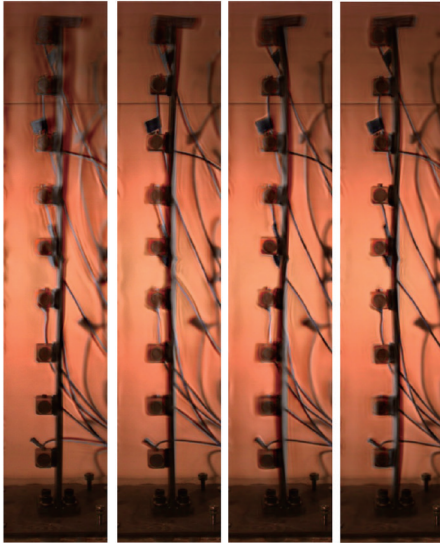
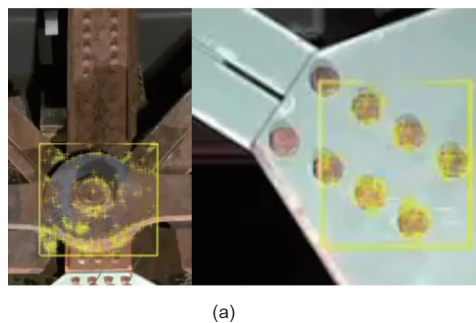


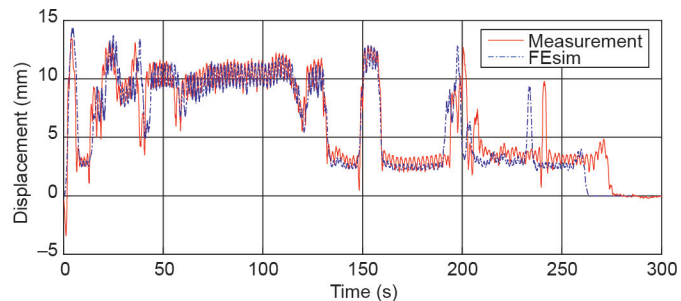
图19. 参考文献[216]中悬臂梁振动的运动放大视频截图。

构位移。此外，在城市环境中，要找到一个好的位置来放置相机是很困难的，由于使用变焦镜头进行远距离监测，导致出现了包含透视畸变和大气畸变的视频数据[216]。最后，当使用远程相机时，只能监测从所选相机位置容易看到的结构上的点。从视频数据中分离模态信息，通常需要选择手动生成掩模或关注区域，这使得整个过程十分繁杂。

Xu等[222]提出了一种低成本、非接触式的基于视觉的多点位移测量系统，该系统基于消费级摄像机进行视频采集，并且他们使用这个系统获得了人行天桥的振型。Hoskere等[228]研发了一种分而治之的方法，利用UAV获取全尺寸基础设施的振型。这个方法直接解决了与使用基于视觉的方法对全尺寸基础设施进行模态分析相关的许多困难。在实验室环境下，利用振动台上六层剪力模型对该方法进行了初步评价。随后，对一座全尺寸人行吊桥进行了现场测试，以获得其固有频率和振型(图21)[151]。



(a)



(b)

图20. 利用计算机视觉技术测量铁路桥梁位移[225]。(a) 铁路构件的光学追踪图像；(b) 基于视觉的位移测量与FE模拟估测的比较。FEsim: 有限元模拟。

## 5. 基于视觉的民用基础设施自动化检查和监测所面临的挑战

尽管近年来研究界取得了重大进展，但在使用基于视觉的技术完全实现自动化SHM之前，必须克服许多技术障碍。其中主要的困难在于将基于视觉的方法所提取的特征和信号转换成更具可操作性的信息，从而有助于更高层次的决策。

### 5.1. 自动化结构检查需要对损坏情况和背景进行全面了解

执行视觉检查的人类具有非凡的感知能力，这是视觉和深度学习算法难以复制的。训练有素的检查员能够识别出对结构整体健康有重要意义的区域（如关键的结构构件、结构上明显的损坏等）。当结构受损时，根据损坏的形状、大小和位置，以及损坏部件的类型和重要性，训练有素的检查员可以推断出损坏结构的重要性。检查员能够理解多种损坏存在的影响。因此，虽然目视检查已经取得了重大进展，但仍然需要更高精度的损伤检测和构件识别。此外，关于解释已识别损伤的结构意义、将局部信息与全局信息同化以进行结构级评估这类研究，几乎鲜有文献提及。解决这些问题对于实现基于视觉的全自动检查至关重要。

### 5.2. 深层网络的普遍性取决于数据的普遍性

从推断数据中提取的特征如果与训练数据存在显著差异，那么训练后的DCNN模型往往会表现不佳。因此，经过训练的深层模型的质量直接取决于基础数据集。DCNN模型的感知能力对诸如凹槽或关节等类似损伤的特征还不具有鲁棒性，因此在推断期间无法区分这些纹理。为提高DCNN对自动检查的感知能力，就必须克服用于检测结构损坏的数据集的有限性。



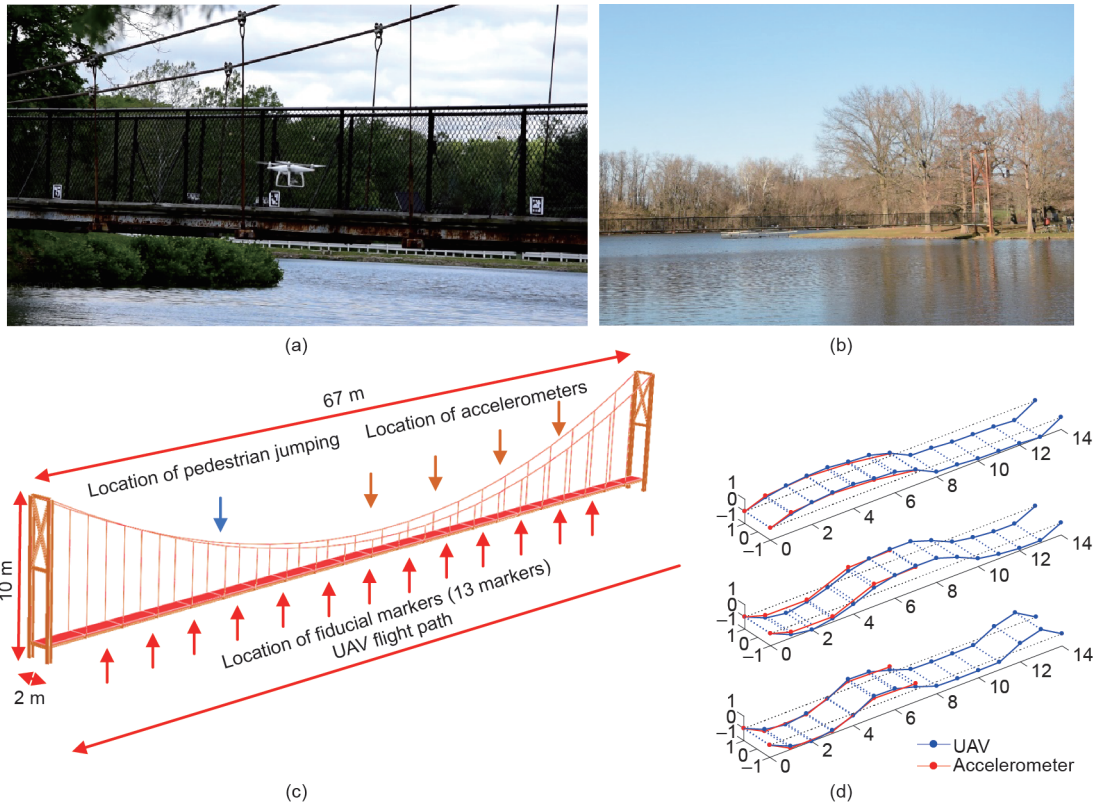


图21. (a) Phantom 4拍摄的一座振动桥梁的视频图像（每秒30帧，像素3840×2160）；(b) 美国伊利诺伊州Mahomet市的伍兹湖人行天桥；(c) 桥梁有限元模型；(d) 提取的模态形状[151]。

### 5.3. 检查的人类感知需要理解顺序视图

单个图像并不总能为损伤检测和构件识别提供足够的信息。例如，当图像是构件的特写视图时，损伤识别最有可能成功；但是，对这类图像进行构件识别就非常困难。在极端情况下，检查员可能非常靠近构件，以至于无法区分混凝土柱与混凝土梁或混凝土墙。在进行人工检查时，可通过先检查整个结构，然后靠近结构构件，并且同时牢记目标结构构件，就能很容易地解决这个问题。为了复制这种功能，必须将观看顺序（如使用视频数据）合并到检查过程中，并且必须基于当前帧以及以前的帧来执行识别任务。

### 5.4. 位移通常很小并且难以捕捉

对于监测应用，最近的工作成功地证明了用基于视觉的方法来测量模态信息以及实验室和现场结构的位移和应变是可行的。另一方面，对现场民用基础设施的精确位移和应变的测量很少是直接测得的。现场试验中预期的位移和应变范围通常小于实验室试验中的位移和应变范围，因为现场的目标结构会对操作荷载做出反应。在现场环境中，重要结构构件的可访问性通常是有限

的。在这种情况下，无法达到高质量测量的最佳摄像机位置，也无法放置引导位移测量的标记。对于静态应用，一般通过人为添加表面纹理（如斑点图案）在DIC方法中进行图像匹配[183]，然而这对于可访问性有限的结构而言也很困难。为了在这种操作情况下应用基于视觉的静态位移/应变测量，在硬件和软件方面都需要进行更深的研究和开发工作。

### 5.5. 照明和环境影响

基于视觉的方法非常容易受到与能见度相关的环境变化的影响，如下雨和起雾。尽管上述问题难以规避，但其他环境因素，如光线、阴影和大气干扰的变化是可以归一化的，尽管需要做更多的工作来提高鲁棒性。

### 5.6. 大数据需要大数据管理

基于视觉的连续和自动化监测的实现对生成的大量数据提出了挑战，这些数据在长期应用中很难存储和处理。为了减少存储的数据量，自动实时信号提取是必要的。处理并加工通过视频带通技术获得的全场模态信息的方法也是一个有待研究的领域。

## 6. 仍在进行的自动化检查工作

为了实现自动化检查目标，基于视觉的感知仍然是一个需要大量关注的开放性研究问题。本节讨论了伊利诺伊大学正在进行的旨在解决以下挑战的工作，在第5节对这些挑战已进行了概述：①结合背景以生成状态感知模型；②使用基于物理的逼真图形模型生成合成标记数据，以满足对更一般数据的需求；③利用视频序列对结构成分进行人类识别。

### 6.1. 结合背景以生成状态感知模型

如第5.1节所述，了解损害发生的背景是进行自动化和高级别评估以提供详细检查判断的关键。为了解决这个问题，Hoskere等[128]提出了一种新的程序，其中有关结构类型、各种构件以及每个构件的状态信息被合并为一个单独的模型，即状态感知模型。此类模型可被视为类似于建筑和设计行业中使用的竣工模型，但在此处被用于检查和维护。状态感知模型是自动生成的注释，可显示结构上存在的视觉缺陷。根据所考虑的特定检查应用，所需的状态感知模型的保真度也各不

相同。与直接使用图像相比，构建状态感知模型的主要优点是结构的背景和损坏的规模是很容易被识别的。此外，全局3D几何信息对评估过程也有帮助。该模型作为一个方便的实体，可以快速、自动地记录结构上可见的缺陷。

Hoskere等[128]提出了用于生成灾后快速自动检查的状态感知模型框架，如图22所示。3D网格模型由UAV结构测量中的多视图立体生成。然后在同一组图像上进行基于深度学习的条件推断，从而对损伤和建筑环境进行语义分割。使用UV映射将生成的标签投影到网格上（将2D图像投影到3D模型的3D建模过程），生成一个在每个单元上叠加了平均损伤和背景标签的状态感知模型。图23显示了在2017年9月墨西哥中部地震期间，使用此程序为受损的建筑物开发的状态感知模型。

### 6.2. 使用基于物理的逼真图形模型生成合成标记数据

如第5.2节所述，对于针对自动化检查的深度学习技术，由于缺乏大量标记数据，使其难以在各种结构和环境条件下推广培训模型。每一种土木工程结构都是独一无二的，这使得损伤识别更具挑战性。例如，涂在建

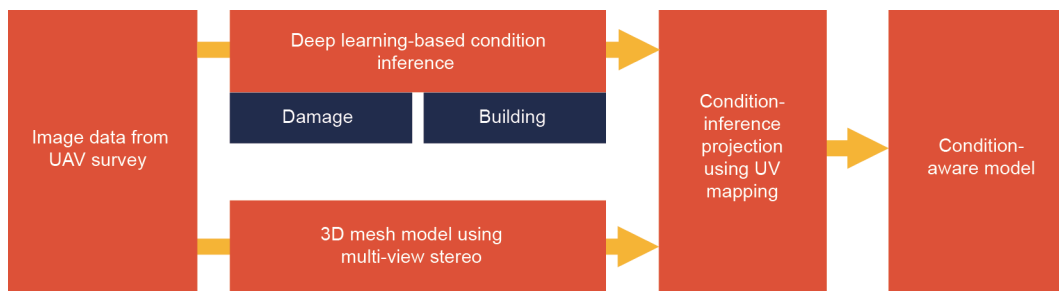


图22. 生成灾后快速检查的状态感知模型框架。

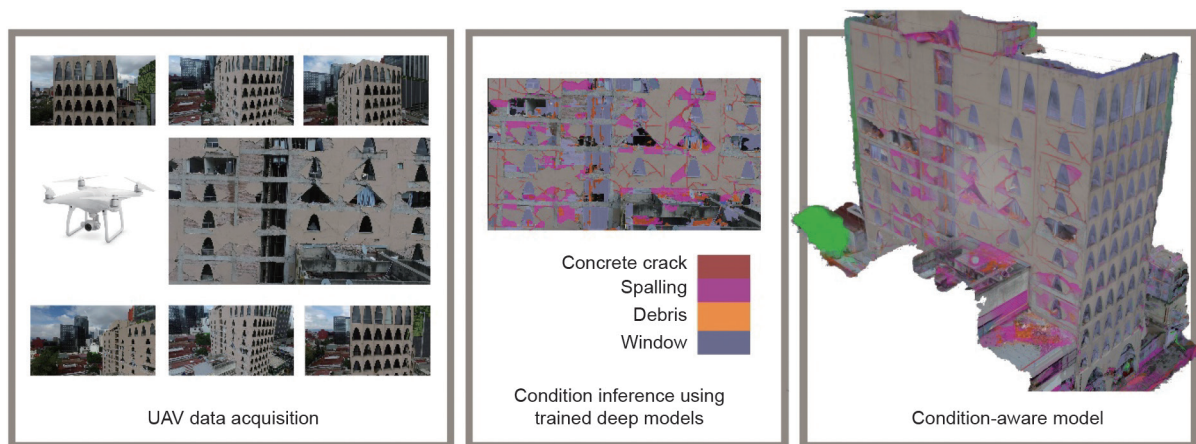


图23. 2017年9月墨西哥中部地震中受损建筑的状态感知模型[128]。

建筑物上的各种颜色（这一参数肯定会对损伤检测的结果产生影响，特别是对于腐蚀）；因此，在不考虑这些问题的情况下，开发用于损伤检测的通用算法是困难的。然而，更严重的问题是，因为受损结构并不常见，所以从受损结构中获取高质量数据也相对困难。

在过去十年中，计算机图形学领域取得的重大进展使人们能够创建出逼真的图像和视频。这里的合成数据指的是从图形模型中生成的数据，而不是来自现实世界中的相机。近年来，合成数据已被应用于计算机视觉领域，用于训练深层神经网络对城市场景进行语义分割，而且基于合成数据的模型在实际数据上显示出良好的性能[229]。使用合成数据有很多好处。有两种类型的平台可用于生成合成数据：①使用光栅化以低计算成本来渲染图形图像的实时游戏引擎，但它缺少准确性和真实性；②使用基于物理的光线跟踪引擎精确模拟光和材料的渲染器，以高计算成本来生成逼真的图形。合成数据的生成有助于解决数据标记问题，因为任何基于算法生成的图形模型的数据都将在像素级和图像级自动标记。图形模型还可以为视觉算法提供具有重复条件的测试平台。测试平台可以模拟不同的环境条件（如照明），并且可以使用不同的相机参数和UAV数据采集的飞行路径来研究算法。在这些虚拟测试平台中，有效的算法将更有可能在真实数据集上奏效。

3D建模、模拟和渲染工具（如Blender[230]）可以更好地模拟现实环境的影响。结合有限元模型的变形网格，这些工具可用于创建受损结构的图形模型。了解结构的损坏情况需要背景感知。例如，同一种结构不同位置的相同裂纹可能对结构的整体健康产生不同的影响。同样，桥梁中的裂缝必须与建筑物墙壁中的裂缝区别对待。Hoskere等[231]提出了一种新的框架（图24），即使用基于物理的结构模型来创建具有代表性的受损结

构的合成图形图像。该框架主要有五个步骤：①使用参数化的有限元模型对各种形状、尺寸和材料的代表性结构进行结构建模；②利用非线性有限元分析方法对生成模型的结构热点进行识别；③应用材料图形特性对生成模型进行真实绘制；④使用有限元模型中的热点生成程序损伤；⑤利用生成的合成数据训练用于评估的深度学习模型。

基于物理的图形模型可用于生成各种各样的损伤场景。由于生成的数据类似于真实数据，因此可以确定用于识别重要损伤和结构特征的深度学习方法的局限性。这些模型在多个环境层面都提供了高质量的标记数据，包括：①整体结构属性，如楼层和间隔的数量以及结构系统；②结构和非结构构件以及关键区域；③不同类型的局部和全局损坏，如裂缝、剥落、薄弱层和柱的屈曲，以及其他如坠落部分等危险。这种更高级别的环境信息有望提供更可靠的自动检查，而对局部损坏图像进行训练的方法却难以达到该效果。通过使用基于物理的客观模型作为训练数据的基础，而不是使用主观的手工标记数据，可以大大减少现场检查人员的固有主观性。

对利用合成数据进行基于视觉的检测应用的研究一直以来都有限。Hoskere等[232]创建了一个基于物理的人字门图形模型，并训练深度语义分割，以识别合成环境中门的重要变化。网络训练数据是使用基于物理的图形模型生成的，包括裂缝和腐蚀等缺陷，同时适应照明的变化（图25）。

目前正在进行一项研究，目的是使以合成数据为训练对象的成功深度学习模型同样适用于实际数据。

### 6.3. 利用视频序列对结构成分进行人类识别

人类检查员首先对整个结构进行调查，然后进一步对受损的结构构件进行详细评估。在进行详细检查时，

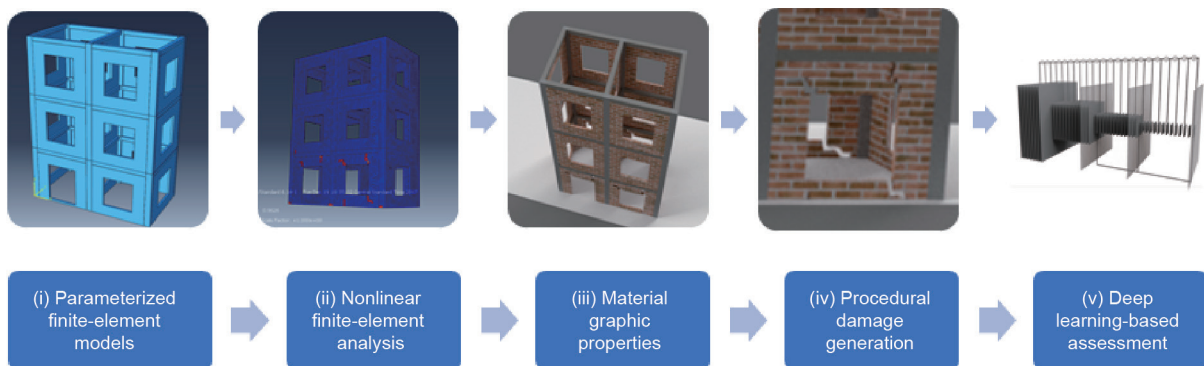


图24. 基于物理的图形生成框架，用于使用深度学习进行自动评估[231]。

他们会记住受损部件是如何适应整体结构环境的；这对于了解损坏对结构安全的相关性至关重要。然而，正如第5.3节所讨论的，用于损伤检测的计算机视觉策略通常基于逐帧操作，即独立使用单个图像；特写图像不包含有关全局结构背景的必要信息。对于检查员来说，观看历史（如视频中的相关图像序列）可提供这种背景信息。本节讨论了嵌入到视频序列中的观看历史，以便在整个检查过程中实现更精确的结构构件识别。

Narazaki等[233]利用视频数据将循环神经网络（RNNs）应用于桥梁构件的识别中，其中包括整体结构视图和结构构件表面的特写细节。研究中使用的网络架构如图26所示。首先，应用一个基于深度单一图像的FCN来提取标签预测图。接下来，在最低分辨率预测层之后添加三个较小的RNN层。最后，将RNN层和其他具有高分辨率的跳跃层的输出相结合，生成最终的预估标签映射。RNN单元仅在最低分辨率预测层之后插入，因为研究中的RNN单元是被用来记忆视频聚焦位置的，

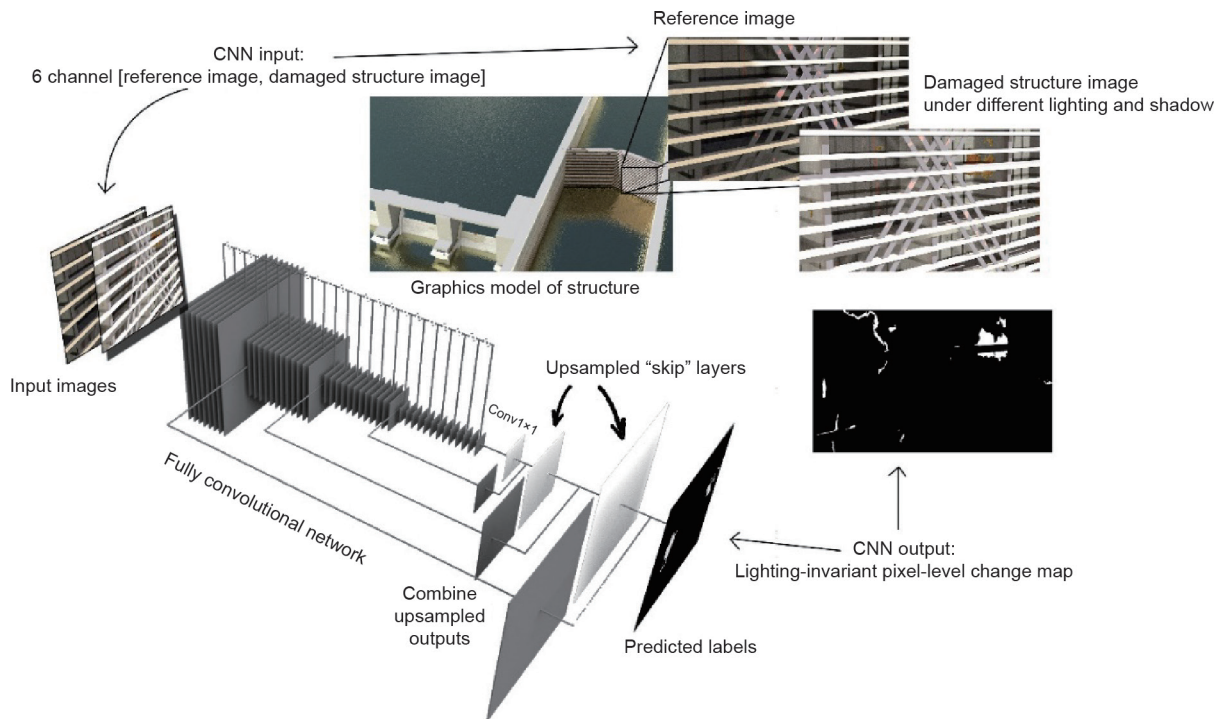


图25. 基于深度学习的变化探测[232]。

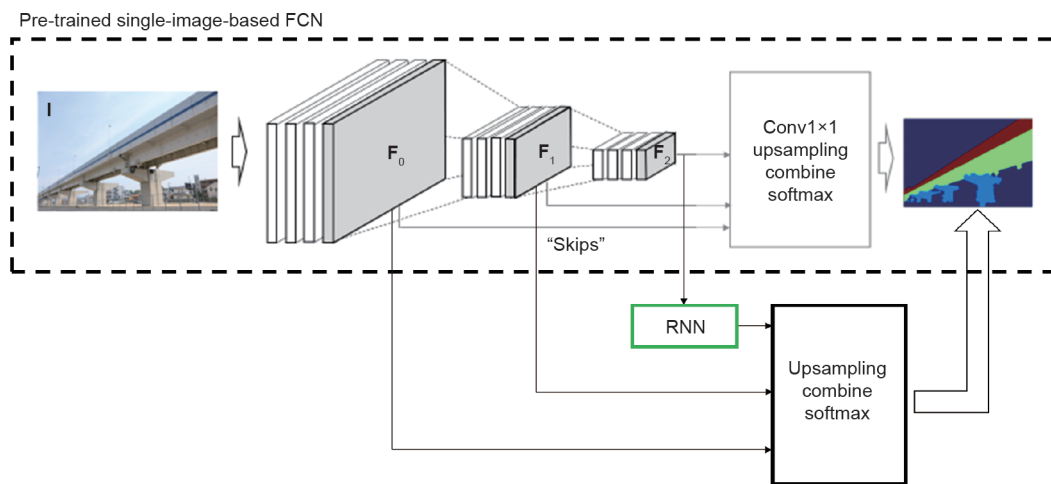


图26. 参考文献[186]中使用的网络架构图。

而不是改善预估映射的细节水平的。

在该研究中测试了两种类型的RNN单元[186]：简单的RNN单元和卷积长短期记忆（ConvLSTM）单元[234]。在简单的RNN单元中，前一时段的输出增大了现在时段的输入，并采用了ReLU激活函数的卷积。另外，ConvLSTM单元被插入到架构的RNN中，以此有效地模拟长期模式。

对RNN进行视频处理训练和测试的主要挑战之一是采集视频数据及其地面实况标签。手动标记每一帧的视频数据是很不切实际的。在第6.2节讨论了合成数据的优点之后，有研究[186]通过使用Unity3D游戏引擎[235]的实时渲染功能解决了这一问题。通过对UAV在混凝土高架桥上的航行进行模拟，我们创建了视频数据集。用于创建数据集的步骤与用于创建SYNTHIA数据集的步骤相似[229]。然而，这个数据集在3D空间中随着航向、倾斜程度和飞行高度的变化随意航行。视频的分辨率设置为 $240 \times 320$ ，自动生成37 081个训练图像和2000个测试图像以及相应的地面实况标签。视频的示例帧如图27所示。此外，还检索了深度图，尽管该数据并未用于研究。

图28中的示例结果显示，当FCN不能正确识别桥梁构件时，重复单元仍然有效。这些结果表明，即使全局结构的视觉线索暂时不可用，ConvLSTM单元与预先训练的FCN相结合还是一种有效的桥梁构件自动识别方法。基于单个图像的FCN的总像素精度为65.0%。相比之下，简单RNN和ConvLSTM单元的总像素精度分别为74.9%和80.5%。数据集、训练和测试的其他细节见参考文献[233]。

目前，这项研究正被用于开发地震后交通基础设施的快速检测策略。

## 7. 结论

本文概述了基于计算机视觉的民用基础设施检查和监测的最新进展。目前，人工目视检测是评估土木基础设施状况的主要手段。针对民用基础设施检查和监测的计算机视觉技术是一种自然的进步，人们可以轻易地用它来帮助并最终代替人工目视检测，同时，它还提供了新的优势和机遇。然而，图像数据的使用可能是一把双刃剑；尽管每幅图像中都存在着丰富的空间、纹理和背景信息，但从这些图像中提取可操作信息的过程是有难度的。从深度学习到光流，研究界已经成功地证明了视觉算法的可行性。本文讨论的检查应用分为以下三类：

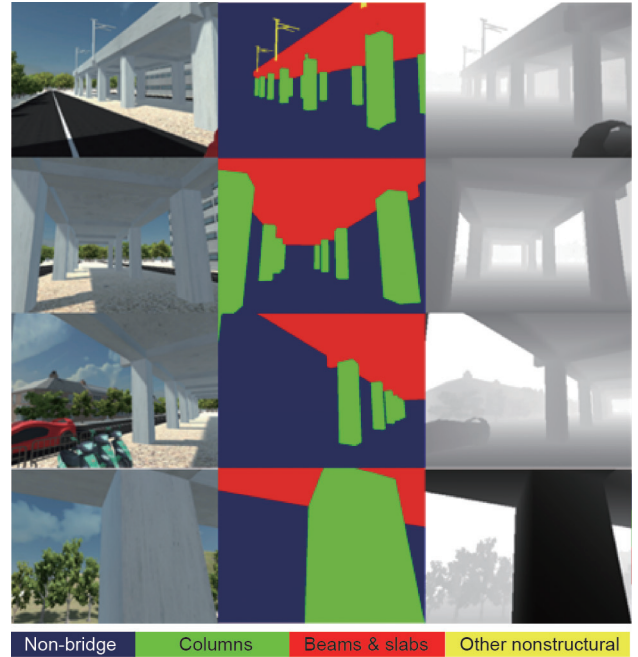


图27. 具有地面实况标签和地面实况深度图的新视频数据集的示例帧。

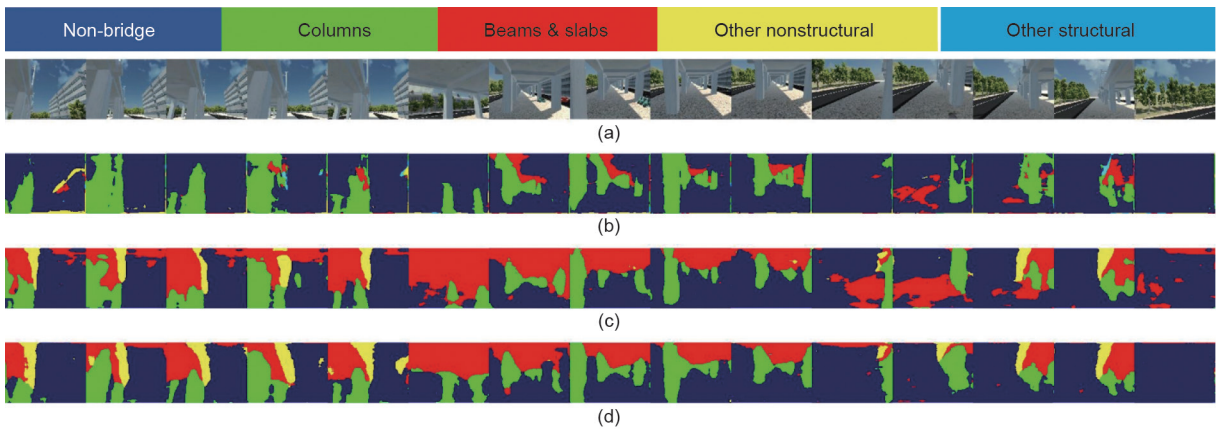


图28. 示例结果。(a) 输入图像；(b) FCN；(c) FCN-简单RNN；(d) FCN-ConvLSTM[233]。

表征局部和全局可见损伤、检测参考图像的变化以及结构构件识别。自动化检测的最新进展源于以数据驱动检测替代启发式方法，在这种方法中，通过对大量数据集进行训练来建立深层模型。监控应用程序包括静态和动态两种。全面实地测量技术的应用以及实验室技术在全面基础设施中的推广为该方法进一步的发展提供了动力。

本文还介绍了研究界在实现基于视觉的自动检查和监测方面面临的关键挑战。这些挑战主要在于将基于视觉方法提取的特征和信号转换为可操作的数据，从而在更高的层次上帮助决策。

最后，本文提出了正在进行的旨在实现自动化检查的三个研究领域：状态感知模型的生成、通过图形模型生成合成数据以及从视频中提取数据的方法。本文所述的基于计算机视觉的民用基础设施检查和监测的研究正在快速发展，这将为最终实现自动化的民用基础设施检查和检测获得更高的时间效率和成本效益，同时，预示了基础设施维护和管理方式的革命即将来临，最终将使世界各地的城市变得更安全、更有弹性。

## Acknowledgements

This research was supported in part by funding from the US Army Corps of Engineers under a project entitled “Cybermodeling: A Digital Surrogate Approach for Optimal Risk-Based Operations and Infrastructure” (W912HZ-17-2-0024).

## Compliance with ethics guidelines

Billie F. Spencer Jr., Vedhus Hoskere, and Yasutaka Narazaki declare that they have no conflict of interest or financial conflicts to disclose.

## References

- [1] ASCE's 2017 infrastructure report card: bridges [Internet]. Reston: American Society of Civil Engineers; 2017 [cited 2017 Oct 11]. Available from: <https://www.infrastructurereportcard.org/cat-item/bridges/>.
- [2] R. P. Gallagher Associates, Inc. ATC-20-1 Field manual: postearthquake safety evaluation of buildings (second edition). Redwood City: Applied Technology Council; 2005.
- [3] Federal Highway Administration, Department of Transportation. National bridge inspection standards. Fed Regist [Internet] 2004 Dec [cited 2018 Jul 30];69(239):74419–39. Available from: <https://www.govinfo.gov/content/pkg/FR-2004-12-14/pdf/04-27355.pdf>.
- [4] Corps of Engineers inspection team goes to new heights. Washington, DC: US Army Corps of Engineers; 2012.
- [5] Ou J, Li H. Structural health monitoring in mainland China: review and future trends. *Struct Health Monit* 2010;9(3):219–31.
- [6] Carden EP, Fanning P. Vibration based condition monitoring: a review. *Struct Health Monit* 2004;3(4):355–77.
- [7] Brownjohn JMW. Structural health monitoring of civil infrastructure. *Philos Trans A Math Phys Eng Sci* 1851;2007(365):589–622.
- [8] Li HN, Ren L, Jia ZG, Yi TH, Li DS. State-of-the-art in structural health monitoring of large and complex civil infrastructures. *J Civ Struct Heal Monit* 2016;6(1):3–16.
- [9] Roberts LG. Machine perception of three-dimensional solids [dissertation]. Cambridge: Massachusetts Institute of Technology; 1963.
- [10] Postal mechanization and early automation [Internet]. Washington, DC: United States Postal Service; 1956 [cited 2018 Jun 22]. Available from: [https://about.usps.com/publications/pub100/pub100\\_042.htm](https://about.usps.com/publications/pub100/pub100_042.htm).
- [11] Szeliski R. Computer vision: algorithms and applications. London: Springer; 2011.
- [12] Viola P, Jones MJ. Robust real-time face detection. *Int J Comput Vis* 2004;57(2):137–54.
- [13] Turk MA, Pentland AP. Face recognition using eigenfaces. In: Proceedings of 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 1991 Jun 3–6; Maui, HI, USA. Piscataway: IEEE; 1991. p. 586–91.
- [14] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2005 Jun 20–25; San Diego, CA, USA. Piscataway: IEEE; 2005. p. 886–93.
- [15] Pingali G, Opalach A, Jean Y. Ball tracking and virtual replays for innovative tennis broadcasts. In: Proceedings of the 15th International Conference on Pattern Recognition; 2000 Sep 3–7; Barcelona, Spain. Piscataway: IEEE; 2000. p. 152–6.
- [16] Bishop CM. Pattern recognition and machine learning. New York City: Springer-Verlag; 2006.
- [17] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Gordon G, Dunson D, Dudík M, editors. Proceedings of the 14th International Conference on Artificial Intelligence and Statistics; 2011 Apr 11–13; Fort Lauderdale, FL, USA; 2011. p. 315–23.
- [18] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by backpropagating errors. *Nature* 1986;323(6088):533–6.
- [19] Kingma DP, Ba J. Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations; 2015 May 7–9; San Diego, CA, USA; 2015. p. 1–15.
- [20] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE* 1998;86(11):2278–324.
- [21] Wu S, Zhong S, Liu Y. Deep residual learning for image steganalysis. *Multimed Tools Appl* 2018;77(9):10437–53.
- [22] Karpathy A. t-SNE visualization of CNN codes [Internet]. [cited 2018 Jul 27]. Available from: <https://cs.stanford.edu/people/karpathy/cnembed/>.
- [23] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition; 2015 Jun 7–12; Boston, MA, USA. Piscataway: IEEE; 2015. p. 3431–40.
- [24] Farabet C, Couprie C, Najman L, LeCun Y. Learning hierarchical features for scene labeling. *IEEE Trans Pattern Anal Mach Intell* 2013;35(8):1915–29.
- [25] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017;39(12):2481–95.
- [26] Couprie C, Farabet C, Najman L, LeCun Y. Indoor semantic segmentation using depth information. arXiv:1301.3572. 2013.
- [27] He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. Computer vision—ECCV 2014. Cham: Springer; 2014. p. 346–61.
- [28] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition; 2014 Jun 23–28; Columbus, OH, USA. Piscataway: IEEE; 2014. p. 580–7.
- [29] Girshick R. Fast R-CNN. In: Proceedings of 2015 IEEE International Conference on Computer Vision; 2015 Dec 7–13; Santiago, Chile. Piscataway: IEEE; 2015. p. 1440–8.
- [30] Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 2017;39(6):1137–49.
- [31] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, realtime object detection. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway: IEEE; 2016. p. 779–88.
- [32] He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: Proceedings of 2017 IEEE International Conference on Computer Vision; 2017 Oct 22–29; Venice, Italy. Piscataway: IEEE; 2017. p. 2980–8.
- [33] De Brabandere B, Neven D, Van Gool L. Semantic instance segmentation with a discriminative loss function. arXiv:1708.02551. 2017.
- [34] Zhang Z, Schwing AG, Fidler S, Urtasun R. Monocular object instance segmentation and depth ordering with CNNs. arXiv:1505.03159. 2015.
- [35] Werbos PJ. Backpropagation through time: what it does and how to do it. *Proc IEEE* 1990;78(10):1550–60.
- [36] Perazzi F, Khoreva A, Benenson R, Schiele B, Sorkine-Hornung A. Learning video object segmentation from static images. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition; 2017 Jul 21–26;

- Honolulu, HI, USA. Piscataway: IEEE; 2017. p. 2663–72.
- [37] Nilsson D, Sminchisescu C. Semantic video segmentation by gated recurrent flow propagation. arXiv:1612.08871. 2016.
- [38] Labelme.csail.mit.edu [Internet]. Cambridge: MIT Computer Science and Artificial Intelligence Laboratory; [cited 2018 Aug 1]. Available from: <http://labelme.csail.mit.edu/Release3.0/>.
- [39] Hess MR, Petrovic V, Kuester F. Interactive classification of construction materials: feedback driven framework for annotation and analysis of 3D point clouds. In: Hayes J, Ouimet C, Santana Quintero M, Fai S, Smith L, editors. Proceedings of ICOMOS/ISPRS International Scientific Committee on Heritage Documentation (volume XLII-2/W5); 2017 Aug 28–Sep 1; Ottawa, ON, Canada; 2017. p. 343–7.
- [40] Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway: IEEE; 2016. p. 2921–9.
- [41] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proceedings of 2017 IEEE International Conference on Computer Vision; 2017 Oct 22–29; Venice, Italy. Piscataway: IEEE; 2017. p. 618–26.
- [42] DeGol J, Golparvar-Fard M, Hoiem D. Geometry-informed material recognition. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway: IEEE; 2016. p. 1554–62.
- [43] Hinton GE. Training products of experts by minimizing contrastive divergence. *Neural Comput* 2002;14(8):1771–800.
- [44] Salakhutdinov R, Larochelle H. Efficient learning of deep Boltzmann machines. In: Teh YW, Titterton M, editors. Proceedings of the 13th International Conference on Artificial Intelligence and Statistics; 2010 May 13–15; Sardinia, Italy; 2010. p. 693–700.
- [45] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science* 2006;313(5786):504–7.
- [46] Pathak D, Krähenbühl P, Donahue J, Darrell T, Efros AA. Context encoders: feature learning by inpainting. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway: IEEE; 2016. p. 2536–44.
- [47] Zhang R, Isola P, Efros AA. Split-brain autoencoders: unsupervised learning by cross-channel prediction. In: Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition; 2017 Jul 21–26; Honolulu, HI, USA. Piscataway: IEEE; 2017. p. 645–54.
- [48] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434. 2015.
- [49] Denton E, Chintala S, Szlam A, Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems: volume 1; 2015 Dec 7–12; Montreal, QC, Canada. Cambridge: MIT Press; 2015. p. 1486–94.
- [50] Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved techniques for training GANs. In: Proceedings of the 30th International Conference on Neural Information Processing Systems; 2016 Dec 5–10; Barcelona, Spain. New York City: Curran Associates Inc.; 2016. p. 2234–42.
- [51] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–44.
- [52] Barron JL, Fleet DJ, Beauchemin SS. Performance of optical flow techniques. *Int J Comput Vis* 1994;12(1):43–77.
- [53] Richardson IEG. H.264 and MPEG-4 video compression: video coding for next-generation multimedia. Chichester: John Wiley & Sons, Ltd.; 2003.
- [54] Grundmann M, Kwatra V, Han M, Essa I. Efficient hierarchical graph-based video segmentation. In: Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2010 Jun 13–18; San Francisco, CA, USA. Piscataway: IEEE; 2010. p. 2141–8.
- [55] Wadhwa N, Rubinstein M, Durand F, Freeman WT. Phase-based video motion processing. *ACM Trans Graph* 2013;32(4):80:1–9.
- [56] Engel J, Sturm J, Cremers D. Camera-based navigation of a low-cost quadcopter. In: Proceedings of 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2012 Oct 7–12; Vilamoura, Portugal. Piscataway: IEEE; 2012. p. 2815–21.
- [57] Bojarski M, Del Testa D, Dworakowski D, Firner B, Flepp B, Goyal P, et al. End to end learning for self-driving cars. arXiv:1604.07316v1. 2016.
- [58] Goodfellow IJ, Bulatov Y, Ibarz J, Arnoud S, Shet V. Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv:1312.6082v4. 2014.
- [59] Templeton B, inventor; Google LLC, Waymo LLC, assignees. Methods and systems for transportation to destinations by a self-driving vehicle. United States Patent US 9665101B1. 2017 May 30.
- [60] Taigman Y, Yang M, Ranzato MA, Wolf L. DeepFace: closing the gap to human level performance in face verification. In: Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition; 2014 Jun 23–28; Columbus, OH, USA. Piscataway: IEEE; 2014. p. 1701–8.
- [61] Airport's facial recognition technology catches 2 people attempting to enter US illegally [Internet]. New York City: FOX News Network, LLC.; 2018 Sep 12 [cited 2018 Sep 25]. Available from: <http://www.foxnews.com/travel/2018/09/12/airports-facial-recognition-technology-catches-2-people-attemptingto-enter-us-illegally.html>.
- [62] Wojna Z, Gorban AN, Lee DS, Murphy K, Yu Q, Li Y, et al. Attention-based extraction of structured information from street view imagery. In: Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition: volume 1; 2017 Nov 9–15; Kyoto, Japan. Piscataway: IEEE; 2018. p. 844–50.
- [63] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60–88.
- [64] Zink J, Lovelace B. Unmanned aerial vehicle bridge inspection demonstration project: final report. St. Paul: Minnesota Department of Transportation; 2015 Jul. Report No.: MN/RC 2015–40.
- [65] Wells J, Lovelace B. Unmanned aircraft system bridge inspection demonstration project phase II: final report. St. Paul: Minnesota Department of Transportation; 2017 Jun. Report No.: MN/RC 2017–18.
- [66] Otero LD, Gagliardo N, Dalli D, Huang WH, Cosentino P. Proof of concept for using unmanned aerial vehicles for high mast pole and bridge inspections: final report. Tallahassee: Florida Department of Transportation; 2015 Jun. Contract No.: BDV28 TWO 977-02.
- [67] Tomiczek AP, Bridge JA, Ifju PG, Whitley TJ, Tripp CS, Ortega AE, et al. Small unmanned aerial vehicle (sUAV) inspections in GPS denied area beneath bridges. In: Soules JG, editor. Structures Congress 2018: bridges, transportation structures, and nonbuilding structures; 2018 Apr 19–21; Fort Worth, TX, USA. Reston: American Society of Civil Engineers; 2018. p. 205–16.
- [68] Brooks C, Dobson RJ, Banach DM, Dean D, Oommen T, Wolf RE, et al. Evaluating the use of unmanned aerial vehicles for transportation purposes: final report. Lansing: Michigan Department of Transportation; 2015 May. Report No.: RC-1616, MTRI-MDOTUAV-FR-2014. Contract No.: 2013-067.
- [69] Duque L, Seo J, Wacker J. Synthesis of unmanned aerial vehicle applications for infrastructures. *J Perform Constr Facil* 2018;32(4):04018046.
- [70] Abdel-Qader I, Abudayyeh O, Kelly ME. Analysis of edge-detection techniques for crack identification in bridges. *J Comput Civ Eng* 2003;17(4):255–63.
- [71] Jahanshahi MR, Kelly JS, Masri SF, Sukhatme GS. A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. *Struct Infrastruct Eng* 2009;5(6):455–86.
- [72] Jahanshahi MR, Masri SF. A new methodology for non-contact accurate crack width measurement through photogrammetry for automated structural safety evaluation. *Smart Mater Struct* 2013;22(3):035019.
- [73] Zhu Z, German S, Brilakis I. Visual retrieval of concrete crack properties for automated post-earthquake structural safety evaluation. *Autom Construct* 2011;20(7):874–83.
- [74] Nishikawa T, Yoshida J, Sugiyama T, Fujino Y. Concrete crack detection by multiple sequential image filtering. *Comput Civ Infrastruct Eng* 2012;27(1):29–47.
- [75] Yamaguchi T, Hashimoto S. Fast crack detection method for large-size concrete surface images using percolation-based image processing. *Mach Vis Appl* 2010;21(5):797–809.
- [76] Zhang W, Zhang Z, Qi D, Liu Y. Automatic crack detection and classification method for subway tunnel safety monitoring. *Sensors (Basel)* 2014;14(10):19307–28.
- [77] Olsen MJ, Chen Z, Hutchinson T, Kuester F. Optical techniques for multiscale damage assessment. *Geomat Nat Haz Risk* 2013;4(1):49–70.
- [78] Lee BY, Kim YY, Yi ST, Kim JK. Automated image processing technique for detecting and analysing concrete surface cracks. *Struct Infrastruct Eng* 2013;9(6):567–77.
- [79] Liu YF, Cho SJ, Spencer BF Jr, Fan JS. Automated assessment of cracks on concrete surfaces using adaptive digital image processing. *Smart Struct Syst* 2014;14(4):719–41.
- [80] Liu YF, Cho SJ, Spencer BF Jr, Fan JS. Concrete crack assessment using digital image processing and 3D scene reconstruction. *J Comput Civ Eng* 2016;30(1):04014124.
- [81] Jahanshahi MR, Masri SF, Padgett CW, Sukhatme GS. An innovative methodology for detection and quantification of cracks through incorporation of depth perception. *Mach Vis Appl* 2013;24(2):227–41.
- [82] Erkal BG, Hajjar JF. Laser-based surface damage detection and quantification using predicted surface properties. *Autom Construct* 2017;83:285–302.
- [83] Kim H, Ahn E, Cho S, Shin M, Sim SH. Comparative analysis of image binarization methods for crack identification in concrete structures. *Cement Concr Res* 2017;99:53–61.
- [84] Prasanna P, Dana KJ, Gucunski N, Basily BB, La HM, Lim RS, et al. Automated crack detection on concrete bridges. *IEEE Trans Autom Sci Eng* 2016;13(2):591–9.
- [85] Adhikari RS, Moselhi O, Bagchi A. Image-based retrieval of concrete crack properties for bridge inspection. *Autom Construct* 2014;39:180–94.
- [86] Torok MM, Golparvar-Fard M, Kochersberger KB. Image-based automated 3D crack detection for post-disaster building assessment. *J Comput Civ Eng* 2014;28(5):A4014004.
- [87] Adhikari RS, Moselhi O, Bagchi A. A study of image-based element condition index for bridge inspection. In: Proceedings of the 30th International Symposium on Automation and Robotics in Construction and Mining: building the future in automation and robotics (volume 2); 2013 Aug 11–15; Montreal, QC, Canada. International Association for Automation and Robotics in Construction; 2013. p. 868–79.
- [88] Paal SG, Jeon JS, Brilakis I, DesRoches R. Automated damage index estimation of reinforced concrete columns for post-earthquake evaluations. *J Struct Eng*

- 2015;141(9):04014228.
- [89] Yeum CM, Dyke SJ. Vision-based automated crack detection for bridge inspection. *Comput Civ Infrastruct Eng* 2015;30(10):759–70.
- [90] Greimann LF, Stecker JH, Kao AM, Rens KL. Inspection and rating of miter lock gates. *J Perform Constr Facil* 1991;5(4):226–38.
- [91] Jahanshahi MR, Chen FC, Joffe C, Masri SF. Vision-based quantitative assessment of microcracks on reactor internal components of nuclear power plants. *Struct Infrastruct Eng* 2017;13(8):1013–26.
- [92] Ghanta S, Karp T, Lee S. Wavelet domain detection of rust in steel bridge images. In: *Proceedings of 2011 IEEE International Conference on Acoustics, Speech and Signal Processing*; 2011 May 22–27; Prague, Czech Republic. Piscataway: IEEE; 2011. p. 1033–6.
- [93] Jahanshahi MR, Masri SF. Parametric performance evaluation of waveletbased corrosion detection algorithms for condition assessment of civil infrastructure systems. *J Comput Civ Eng* 2013;27(4):345–57.
- [94] Medeiros FNS, Ramalho GLB, Bento MP, Medeiros LCL. On the evaluation of texture and color features for nondestructive corrosion detection. *EURASIP J Adv Signal Process* 2010;2010:817473.
- [95] Shen HK, Chen PH, Chang LM. Automated steel bridge coating rust defect recognition method based on color and texture feature. *Autom Construct* 2013;31:338–56.
- [96] Son H, Hwang N, Kim C, Kim C. Rapid and automated determination of rusted surface areas of a steel bridge for robotic maintenance systems. *Autom Construct* 2014;42:13–24.
- [97] Igoe D, Parisi AV. Characterization of the corrosion of iron using a smartphone camera. *Instrum Sci Technol* 2016;44(2):139–47.
- [98] Ahuja SK, Shukla MK. A survey of computer vision based corrosion detection approaches. In: Satapathy S, Joshi A, editors. *Information and communication technology for intelligent systems (ICTIS 2017)*—volume 2. Cham: Springer; 2018. p. 55–63.
- [99] Koch C, Brilakis I. Pothole detection in asphalt pavement images. *Adv Eng Inform* 2011;25(3):507–15.
- [100] Salman M, Mathavan S, Kamal K, Rahman M. Pavement crack detection using the Gabor filter. In: *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems*; 2013 Oct 6–9; The Hague, the Netherlands. Piscataway: IEEE; 2013. p. 2039–44.
- [101] Hu Y, Zhao C. A novel LBP based methods for pavement crack detection. *J Pattern Recognit Res* 2010;5(1):140–7.
- [102] Zou Q, Cao Y, Li Q, Mao Q, Wang S. CrackTree: automatic crack detection from pavement images. *Pattern Recognit Lett* 2012;33(3):227–38.
- [103] Kirschke KR, Velinsky SA. Histogram-based approach for automated pavement-crack sensing. *J Transp Eng* 1992;118(5):700–10.
- [104] Li L, Sun L, Tan S, Ning G. An efficient way in image preprocessing for pavement crack images. In: *Proceedings of the 12th COTA International Conference of Transportation Professionals*; 2012 Aug 3–6; Beijing, China. Reston: American Society of Civil Engineers; 2012. p. 3095–103.
- [105] Adu-Gyamfi YO, Okine NOA, Garateguy G, Carrillo R, Arce GR. Multiresolution information mining for pavement crack image analysis. *J Comput Civ Eng* 2012;26(6):741–9.
- [106] Chen YL, Jahanshahi MR, Manjunatha P, Gan WP, Abdelbarr M, Masri SF, et al. Inexpensive multimodal sensor fusion system for autonomous data acquisition of road surface conditions. *IEEE Sens J* 2016;16(21):7731–43.
- [107] Zakeri H, Nejad FM, Fahimifar A. Image based techniques for crack detection, classification and quantification in asphalt pavement: a review. *Arch Comput Methods Eng* 2017;24(4):935–77.
- [108] Koch C, Georgieva K, Kasireddy V, Akinci B, Fieguth P. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv Eng Inform* 2015;29(2):196–210. *Corrigendum in: Adv Eng Inform* 2016;30(2):208–10.
- [109] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *Proceedings of International Conference on Learning Representations 2015*; 2015 May 7–9; San Diego, CA, USA; 2015. p. 1–14.
- [110] Xu Y, Bao Y, Chen J, Zuo W, Li H. Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images. *Struct Health Monit*. Epub 2018 Apr 2.
- [111] Zhang L, Yang F, Zhang YD, Zhu YJ. Road crack detection using deep convolutional neural network. In: *Proceedings of 2016 IEEE International Conference on Image Processing*; 2016 Sep 25–28; Phoenix, AZ, USA. Piscataway: IEEE; 2016. p. 3708–12.
- [112] Cha YJ, Choi W, Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks. *Comput Civ Infrastruct Eng* 2017;32(5):361–78.
- [113] Kim H, Ahn E, Shin M, Sim SH. Crack and noncrack classification from concrete surface images using machine learning. *Struct Health Monit*. Epub 2018 Apr 23.
- [114] Kim B, Cho S. Automated vision-based detection of cracks on concrete surfaces using a deep learning technique. *Sensors (Basel)* 2018;18(10):3452.
- [115] Kim B, Cho S. Automated and practical vision-based crack detection of concrete structures using deep learning. *Comput Civ Infrastruct Eng*. In press.
- [116] Lee Y, Kim B, Cho S. Image-based spalling detection of concrete structures incorporating deep learning. In: *Proceedings of the 7th World Conference on Structural Control and Monitoring*; 2018 Jul 22–25; Qingdao, China; 2018.
- [117] Atha DJ, Jahanshahi MR. Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Struct Health Monit* 2018;17(5):1110–28.
- [118] Chen FC, Jahanshahi MR. NB-CNN: deep learning-based crack detection using convolutional neural network and naïve Bayes data fusion. *IEEE Trans Ind Electron* 2018;65(5):4392–400.
- [119] Yeum CM. Computer vision-based structural assessment exploiting large volumes of images [dissertation]. West Lafayette: Purdue University; 2016.
- [120] Xu Y, Li S, Zhang D, Jin Y, Zhang F, Li N, et al. Identification framework for cracks on a steel structure surface by a restricted Boltzmann machines algorithm based on consumer-grade camera images. *Struct Contr Health Monit* 2018;25(2):e2075.
- [121] Maguire M, Dorafshan S, Thomas RJ. SDNET2018: a concrete crack image dataset for machine learning applications [dataset]. Logan: Utah State University; 2018.
- [122] Bao Y, Tang Z, Li H, Zhang Y. Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Struct Heal Monit* 2019;18(2):401–21.
- [123] Dang J, Shrestha A, Haruta D, Tabata Y, Chun P, Okubo K. Site verification tests for UAV bridge inspection and damage detection based on deep learning. In: *Proceedings of the 7th World Conference on Structural Control and Monitoring*; 2018 Jul 22–25; Qingdao, China; 2018.
- [124] Yeum CM, Dyke SJ, Ramirez J. Visual data classification in post-event building reconnaissance. *Eng Struct* 2018;155:16–24.
- [125] Cha YJ, Choi W, Suh G, Mahmoudkhani S, Büyüköztürk O. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Comput Civ Infrastruct Eng* 2018;33(9):731–47.
- [126] Zhang A, Wang KCP, Li B, Yang E, Dai X, Peng Y, et al. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. *Comput Civ Infrastruct Eng* 2017;32(10):805–19.
- [127] Hoskere V, Narazaki Y, Hoang TA, Spencer BF Jr. Vision-based structural inspection using multiscale deep convolutional neural networks. In: *Proceedings of the 3rd Huixian International Forum on Earthquake Engineering for Young Researchers*; 2017 Aug 11–12; Urbana, IL, USA; 2017.
- [128] Hoskere V, Narazaki Y, Hoang TA, Spencer BF Jr. Towards automated postearthquake inspections with deep learning-based condition-aware models. In: *Proceedings of the 7th World Conference on Structural Control and Monitoring*; 2018 Jul 22–25; Qingdao, China; 2018.
- [129] Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition: volume 1*; 2006 Jun 17–22; New York City, NY, USA. Piscataway: IEEE; 2006. p. 519–28.
- [130] Khaloo A, Lattanzi D, Cunningham K, Dell'Andrea R, Riley M. Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modelling. *Struct Infrastruct Eng* 2018;14(1):124–36.
- [131] Khaloo A, Lattanzi D, Jachimowicz A, Devaney C. Utilizing UAV and 3D computer vision for visual inspection of a large gravity dam. *Front Built Environ* 2018;4:31.
- [132] Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* 1992;14(2):239–56.
- [133] Meshlab.net [Internet]. [cited 2018 Jun 20]. Available from: <http://www.meshlab.net/>.
- [134] CloudCompare: 3D point cloud and mesh processing software open source project [Internet]. [cited 2018 Jun 20]. Available from: <http://www.danielgm.net/cc/>.
- [135] Girardeau-Montaut D, Roux M, Marc R, Thibault G. Change detection on points cloud data acquired with a ground laser scanner. In: *Proceedings of the ISPRS Workshop—Laser Scanning 2005*; 2005 Sep 12–14; Enschede, the Netherlands; 2005. p. 30–5.
- [136] Lague D, Brodu N, Leroux J. Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z). *ISPRS J Photogramm Remote Sens* 2013;82:10–26.
- [137] Morgenthal G, Hallermann N. Quality assessment of unmanned aerial vehicle (UAV) based visual inspection of structures. *Adv Struct Eng* 2014;17(3):289–302.
- [138] Hallermann N, Morgenthal G, Rodehorst V. Unmanned aerial systems (UAS)— case studies of vision based monitoring of ageing structures. In: *Proceedings of the International Symposium Non-Destructive Testing in Civil Engineering*; 2015 Sep 15–17; Berlin, Germany; 2015.
- [139] Khaloo A, Lattanzi D. Automatic detection of structural deficiencies using 4D hue-assisted analysis of color point clouds. In: Pakzad S, editor. *Dynamics of civil structures, volume 2*. Cham: Springer; 2019. p. 197–205.
- [140] Jafari B, Khaloo A, Lattanzi D. Deformation tracking in 3D point clouds via statistical sampling of direct cloud-to-cloud distances. *J Nondestruct Eval* 2017;36:65.
- [141] Ghahremani K, Khaloo A, Mohamadi S, Lattanzi D. Damage detection and finite-element model updating of structural components through point cloud analysis. *J Aerosp Eng* 2018;31(5):04018068.
- [142] Radke RJ, Andra S, Al-Kofahi O, Roysam B. Image change detection algorithms: a systematic survey. *IEEE Trans Image Process* 2005;14(3):294–307.
- [143] Hussain M, Chen D, Cheng A, Wei H, Stanley D. Change detection from remotely sensed images: from pixel-based to object-based approaches. *ISPRS J Photogramm Remote Sens* 2013;80:91–106.
- [144] Tewkesbury AP, Comber AJ, Tate NJ, Lamb A, Fisher PF. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sens*



- Environ 2015;160:1–14.
- [145] Sakurada K, Okatani T, Deguchi K. Detecting changes in 3D structure of a scene from multi-view images captured by a vehicle-mounted camera. In: Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition; 2013 Jun 23–28; Portland, OR, USA. Piscataway: IEEE; 2013. p. 137–44.
- [146] Alcantarilla PF, Stent S, Ros G, Arroyo R, Gherardi R. Street-view change detection with deconvolutional networks. In: Proceedings of 2016 Robotics: Science and Systems Conference; 2016 Jun 18–22; Ann Arbor, MI, USA; 2016.
- [147] Stent S, Gherardi R, Stenger B, Soga K, Cipolla R. Visual change detection on tunnel linings. *Mach Vis Appl* 2016;27(3):319–30.
- [148] Kiziltas S, Akinci B, Ergen E, Tang P, Gordon C. Technological assessment and process implications of field data capture technologies for construction and facility/infrastructure management. *Electron J Inf Technol Constr* 2008;13:134–54.
- [149] Golparvar-Fard M, Peña-Mora F, Savarese S. D4AR—a 4-dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication. *J Inf Technol Constr* 2009;14:129–53.
- [150] Huber D. ARIA: the Aerial Robotic Infrastructure Analyst. SPIE Newsroom [Internet]. 2014 Jun 9 [cited 2019 Feb 24]. Available from: <http://spie.org/newsroom/5472-aria-the-aerial-robotic-infrastructure-analyst?SSO=1>.
- [151] Earthsense.co [Internet]. Champaign: EarthSense, Inc.; [cited 2019 Feb 25]. Available from: <https://www.earthsense.co>.
- [152] Zhu Z, Brilakis I. Concrete column recognition in images and videos. *J Comput Civ Eng* 2010;24(6):478–87.
- [153] Koch C, Paal SG, Rashidi A, Zhu Z, König M, Brilakis I. Achievements and challenges in machine vision-based inspection of large concrete structures. *Adv Struct Eng* 2014;17(3):303–18.
- [154] Xiong X, Adan A, Akinci B, Huber D. Automatic creation of semantically rich 3D building models from laser scanner data. *Autom Construct* 2013;31:325–37.
- [155] Perez-Perez Y, Golparvar-Fard M, El-Rayes K. Semantic and geometric labeling for enhanced 3D point cloud segmentation. In: Perdomo-Rivera JL, González-Quevedo A, del Puerto CL, Maldonado-Fortunet F, Molina-Bas OI, editors. Construction Research Congress 2016: old and new; 2016 May 31–June 2; San Juan, Puerto Rico. Reston: American Society of Civil Engineers; 2016. p. 2542–52.
- [156] Armeni I, Sener O, Zamir AR, Jiang H, Brilakis I, Fischer M, et al. 3D semantic parsing of large-scale indoor spaces. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway: IEEE; 2016. p. 1534–43.
- [157] Golparvar-Fard M, Bohn J, Teizer J, Savarese S, Peña-Mora F. Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Autom Construct* 2011;20(8):1143–55.
- [158] Golparvar-Fard M, Peña-Mora F, Savarese S. Monitoring changes of 3D building elements from unordered photo collections. In: Proceedings of 2011 IEEE International Conference on Computer Vision Workshops; 2011 Nov 6–13; Barcelona, Spain. Piscataway: IEEE; 2011. p. 249–56.
- [159] Lu R, Brilakis I, Middleton CR. Detection of structural components in point clouds of existing RC bridges. *Comput Civ Infrastruct Eng* 2019;34 (3):191–212.
- [160] Yeum CM, Choi J, Dyke SJ. Automated region-of-interest localization and classification for vision-based visual assessment of civil infrastructure. *Struct Health Monit*. Epub 2018 Mar 27.
- [161] Gao Y, Mosalam KM. Deep transfer learning for image-based structural damage recognition. *Comput Civ Infrastruct Eng* 2018;33(9):748–68.
- [162] Liang X. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization. *Comput Civ Infrastruct Eng*. Epub 2018 December 5.
- [163] Narazaki Y, Hoskere V, Hoang TA, Spencer BF Jr. Vision-based automated bridge component recognition integrated with high-level scene understanding. In: Proceedings of the 13th International Workshop on Advanced Smart Materials and Smart Structures Technology; 2017 Jul 22–23; Tokyo, Japan; 2017.
- [164] Narazaki Y, Hoskere V, Hoang TA, Spencer BF Jr. Vision-based automated bridge component recognition with high-level scene consistency. *Comput Civ Infrastruct Eng*. In Press.
- [165] Narazaki Y, Hoskere V, Hoang TA, Spencer BF. Automated vision-based bridge component extraction using multiscale convolutional neural networks. In: Proceedings of the 3rd Huixian International Forum on Earthquake Engineering for Young Researchers; 2017 Aug 11–12; Urbana, IL, USA; 2017.
- [166] German S, Jeon JS, Zhu Z, Bearman C, Brilakis I, DesRoches R, et al. Machine vision-enhanced postearthquake inspection. *J Comput Civ Eng* 2013;27(6):622–34.
- [167] German S, Brilakis I, Desroches R. Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for postearthquake safety assessments. *Adv Eng Informatics* 2012;26(4):846–58.
- [168] Anil EB, Akinci B, Garrett JH, Kurc O. Information requirements for earthquake damage assessment of structural walls. *Adv Eng Inform* 2016;30(1):54–64.
- [169] Anil EB, Akinci B, Kurc O, Garrett JH. Building-information-modeling-based earthquake damage assessment for reinforced concrete walls. *J Comput Civ Eng* 2016;30(4):04015076.
- [170] Wei Y, Kasireddy V, Akinci B, Smith I, Domer B. 3D imaging in construction and infrastructure management: technological assessment and future research directions. In: Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition; 2013 Jun 23–28; Portland, OR, USA. Piscataway: IEEE; 2013. p. 137–44.
- [171] Farrar CR, James III GH. System identification from ambient vibration measurements on a bridge. *J Sound Vibrat* 1997;205(1):1–18.
- [172] Lynch JP, Sundararajan A, Law KH, Kiremidjian AS, Carryer E, Sohn H. Field validation of a wireless structural monitoring system on the Alamosa Canyon Bridge. In: Liu SC, editor. Proceedings volume 5057: Smart Structures and Materials 2003: smart systems and nondestructive evaluation for civil infrastructures; 2003 Mar 2–6; San Diego, CA, USA; 2003.
- [173] Jin G, Sain MK, Spencer Jr BF. Frequency domain system identification for controlled civil engineering structures. *IEEE Trans Contr Syst Technol* 2005;13(6):1055–62.
- [174] Jang S, Jo H, Cho S, Mechtov K, Rice JA, Sim SH, et al. Structural health monitoring of a cable-stayed bridge using smart sensor technology: deployment and evaluation. *Smart Struct Syst* 2010;6(5–6):439–59.
- [175] Rice JA, Mechtov K, Sim SH, Nagayama T, Jang S, Kim R, et al. Flexible smart sensor framework for autonomous structural health monitoring. *Smart Struct Syst* 2010;6(5–6):423–38.
- [176] Moreu F, Kim R, Lei S, Diaz-Fañás G, Dai F, Cho S, et al. Campaign monitoring railroad bridges using wireless smart sensors networks. In: Proceedings of 2014 Engineering Mechanics Institute Conference [Internet]; 2014 Aug 5–8; Hamilton, OT, Canada, 2014 [cited 2017 Oct 11]. Available from: <http://emi2014.mcmaster.ca/index.php/emi2014/EMI2014/paper/viewPaper/139>.
- [177] Cho S, Giles RK, Spencer Jr BF. System identification of a historic swing truss bridge using a wireless sensor network employing orientation correction. *Struct Contr Health Monit* 2015;22(2):255–72.
- [178] Sim SH, Spencer Jr BF, Zhang M, Xie H. Automated decentralized modal analysis using smart sensors. *Struct Contr Health Monit* 2010;17(8):872–94.
- [179] Horn BKP, Schunck BG. Determining optical flow. *Artif Intell* 1981;17(1–3):185–203.
- [180] Bruhn A, Weickert J, Schnörr C. Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. *Int J Comput Vis* 2005;61(3):211–31.
- [181] Liu C. Beyond pixels: exploring new representations and applications for motion analysis [dissertation]. Cambridge: Massachusetts Institute of Technology; 2009.
- [182] Baker S, Scharstein D, Lewis JP, Roth S, Black MJ, Szeliski R. A database and evaluation methodology for optical flow. *Int J Comput Vis* 2011;92(1):1–31.
- [183] Sutton MA, Ortu JJ, Schreier HW. Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications. New York City: Springer Science & Business Media; 2009.
- [184] The VIC-2DTM system [Internet]. Irmo: Correlated Solutions, Inc.; [cited 2019 Feb 24]. Available from: <https://www.correlatedsolutions.com/vic-2d/>.
- [185] GOM Correlate: evaluation software for 3D testing [Internet]. Braunschweig: GOM GmbH; [cited 2019 Feb 24]. Available from: <https://www.gom.com/3dsoftware/gom-correlate.html>.
- [186] Pan B, Qian K, Xie H, Asundi A. Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. *Meas Sci Technol* 2009;20(6):062001.
- [187] Hoult NA, Take WA, Lee C, Dutton M. Experimental accuracy of two dimensional strain measurements using digital image correlation. *Eng Struct* 2013;46:718–26.
- [188] Dutton M, Take WA, Hoult NA. Curvature monitoring of beams using digital image correlation. *J Bridge Eng* 2014;19(3):05013001.
- [189] Ellenberg A, Branco L, Krick A, Bartoli I, Kontsos A. Use of unmanned aerial vehicle for quantitative infrastructure evaluation. *J Infrastruct Syst* 2015;21(3):04014054.
- [190] McCormick N, Lord J. Digital image correlation for structural measurements. *Proc Inst Civ Eng-Civ Eng* 2012;165(4):185–90.
- [191] Yoneyama S, Kitagawa A, Iwata S, Tani K, Kikuta H. Bridge deflection measurement using digital image correlation. *Exp Tech* 2007;31(1):34–40.
- [192] Yoneyama S, Ueda H. Bridge deflection measurement using digital image correlation with camera movement correction. *Mater Trans* 2012;53(2):285–90.
- [193] Helfrick MN, Niezrecki C, Avitabile P, Schmidt T. 3D digital image correlation methods for full-field vibration measurement. *Mech Syst Signal Process* 2011;25(3):917–27.
- [194] Reagan D. Unmanned aerial vehicle measurement using three-dimensional digital image correlation to perform bridge structural health monitoring [dissertation]. Lowell: University of Massachusetts Lowell; 2017.
- [195] Mahal M, Blanksvård T, Täljsten B, Sas G. Using digital image correlation to evaluate fatigue behavior of strengthened reinforced concrete beams. *Eng Struct* 2015;105:277–88.
- [196] Ghorbani R, Matta F, Sutton MA. Full-field deformation measurement and crack mapping on confined masonry walls using digital image correlation. *Exp Mech* 2015;55(1):227–43.
- [197] Tomasi C, Kanade T. Detection and tracking of point features. Technical report. Pittsburgh: Carnegie Mellon University; 1991 Apr. Report No.: CMUCS-91-132.
- [198] Lewis JP. Fast normalized cross-correlation. *Vis. Interface* 1995;10:120–3.

- [199] Ye XW, Dong CZ, Liu T. A review of machine vision-based structural health monitoring: methodologies and applications. *J Sens* 2016;2016:7103039.
- [200] Feng D, Feng MQ. Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection—a review. *Eng Struct* 2018;156:105–17.
- [201] Nogueira FMA, Barbosa FS, Barra LPS. Evaluation of structural natural frequencies using image processing. In: *Proceedings of 2005 International Conference on Experimental Vibration Analysis for Civil Engineering Structures*; 2005 Oct 26–28; Bordeaux, France; 2005.
- [202] Chang CC, Xiao XH. An integrated visual-inertial technique for structural displacement and velocity measurement. *Smart Struct Syst* 2010;6(9):1025–39.
- [203] Fukuda Y, Feng MQ, Narita Y, Kaneko S, Tanaka T. Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm. *IEEE Sens J* 2013;13(12):4725–32.
- [204] Min JH, Gelo NJ, Jo H. Real-time image processing for non-contact monitoring of dynamic displacements using smartphone technologies. In: Lynch JP, editor. *Proceedings volume 9803: sensors and smart structures technologies for civil, mechanical, and aerospace systems 2016*; 2016 Mar 20–24; Las Vegas, NV, USA; 2016. p. 98031B.
- [205] Dong CZ, Ye XW, Jin T. Identification of structural dynamic characteristics based on machine vision technology. *Measurement* 2018;126:405–16.
- [206] Celik O, Dong CZ, Catbas FN. Measurement of human loads using computer vision. In: Pakzad S, editor. *Dynamics of civil structures, volume 2: proceedings of the 36th IMAC, a conference and exposition on structural dynamics 2018*; 2018 Feb 12–15; Orlando, FL, USA. Cham: Springer; 2019. p. 191–5.
- [207] Lee J, Lee KC, Cho S, Sim SH. Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges. *Sensors (Basel)* 2017;17(10):2317.
- [208] Park JW, Moon DS, Yoon H, Gomez F, Spencer Jr BF, Kim JR. Visual-inertial displacement sensing using data fusion of vision-based displacement with acceleration. *Struct Contr Health Monit* 2018;25(3):e2122.
- [209] Schumacher T, Shariati A. Monitoring of structures and mechanical systems using virtual visual sensors for video analysis: fundamental concept and proof of feasibility. *Sensors (Basel)* 2013;13(12):16551–64.
- [210] Yoon H, Elanwar H, Choi H, Golparvar-Fard M, Spencer Jr BF. Target-free approach for vision-based structural system identification using consumergrade cameras. *Struct Contr Health Monit* 2016;23(12):1405–16.
- [211] Ye XW, Yi TH, Dong CZ, Liu T, Bai H. Multi-point displacement monitoring of bridges using a vision-based approach. *Wind Struct* 2015;20(2):315–26.
- [212] Abdelbarr M, Chen YL, Jahanshahi MR, Masri SF, Shen WM, Qidwai UA. 3D dynamic displacement-field measurement for structural health monitoring using inexpensive RGB-D based sensor. *Smart Mater Struct* 2017;26(12):125016.
- [213] Ye XW, Yi TH, Dong CZ, Liu T. Vision-based structural displacement measurement: system performance evaluation and influence factor analysis. *Measurement* 2016;88:372–84.
- [214] Feng D, Feng MQ. Vision-based multipoint displacement measurement for structural health monitoring. *Struct Contr Health Monit* 2016;23(5):876–90.
- [215] Yoon H, Hoskere V, Park JW, Spencer BF Jr. Cross-correlation-based structural system identification using unmanned aerial vehicles. *Sensors (Basel)* 2017;17(9):2075.
- [216] Chen JG, Wadhwa N, Cha YJ, Durand F, Freeman WT, Buyukozturk O. Modal identification of simple structures with high-speed video using motion magnification. *J Sound Vibrat* 2015;345:58–71.
- [217] Cha YJ, Chen JG, Büyüköztürk O. Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters. *Eng Struct* 2017;132:300–13.
- [218] Yang Y, Dorn C, Mancini T, Talken Z, Kenyon G, Farrar C, et al. Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. *Mech Syst Signal Process* 2017;85:567–90.
- [219] Yang Y, Dorn C, Mancini T, Talken Z, Kenyon G, Farrar C, et al. Spatiotemporal video-domain high-fidelity simulation and realistic visualization of full-field dynamic responses of structures by a combination of high-spatial-resolution modal model and video motion manipulations. *Struct Contr Health Monit* 2018;25(8):e2193.
- [220] Zaurin R, Khuc T, Catbas FN. Hybrid sensor-camera monitoring for damage detection: case study of a real bridge. *J Bridge Eng* 2016;21(6):05016002.
- [221] Pan B, Tian L, Song X. Real-time, non-contact and targetless measurement of vertical deflection of bridges using off-axis digital image correlation. *NDT E Int* 2016;79:73–80.
- [222] Xu Y, Brownjohn J, Kong D. A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge. *Struct Contr Health Monit* 2018;25(5):e2155.
- [223] Kim SW, Kim NS. Dynamic characteristics of suspension bridge hanger cables using digital image processing. *NDT E Int* 2013;59:25–33.
- [224] Chen JG. *Video camera-based vibration measurement of infrastructure [dissertation]*. Cambridge: Massachusetts Institute of Technology; 2016.
- [225] Yoon H. *Enabling smart city resilience : post-disaster response and structural health monitoring [dissertation]*. Urbana: University of Illinois at Urbana-Champaign; 2016.
- [226] Mas D, Ferrer B, Acevedo P, Espinosa J. Methods and algorithms for videobased multi-point frequency measuring and mapping. *Measurement* 2016;85:164–74.
- [227] Chen Z, Li H, Bao Y, Li N, Jin Y. Identification of spatio-temporal distribution of vehicle loads on long-span bridges using computer vision technology. *Struct Contr Health Monit* 2016;23(3):517–34.
- [228] Hoskere V, Park JW, Yoon H, Spencer BF Jr. Vision-based modal survey of civil infrastructure using unmanned aerial vehicles. *J Struct Eng*. In press.
- [229] Ros G, Sellart L, Materzynska J, Vazquez D, Lopez AM. The SYNTHIA dataset: a large collection of synthetic images for semantic segmentation of urban scenes. In: *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*; 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway: IEEE; 2016. p. 3234–43.
- [230] Blender.org [Internet]. Amsterdam: Stichting Blender Foundation; [cited 2018 Aug 1]. Available from: <https://www.blender.org/>.
- [231] Hoskere V, Narazaki Y, Spencer BF Jr, Smith MD. Deep learning-based damage detection of miter gates using synthetic imagery from computer graphics. In: *Proceedings of the 12th International Workshop on Structural Health Monitoring*; 2019 Sep 10–12; Stanford, CA, USA; 2019. In press.
- [232] Hoskere V, Narazaki Y, Spencer BF Jr. Learning to detect important visual changes for structural inspections using physics-based graphics models. In: *9th International Conference on Structural Health Monitoring of Intelligent Infrastructure*; 2019 Aug 4–7; Louis, MO, USA; 2019. In press.
- [233] Narazaki Y, Hoskere V, Hoang TA, Spencer BF Jr. Automated bridge component recognition using video data. In: *Proceedings of the 7th World Conference on Structural Control and Monitoring*; 2018 Jul 22–25; Qingdao, China; 2018.
- [234] Shi X, Chen Z, Wang H, Yeung DY, Wong W, Woo W. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems*: volume 1; 2015 Dec 7–12; Montreal, QC, Canada. Cambridge: MIT Press; 2015. p. 802–10.
- [235] Unity3d.com [Internet]. San Francisco: Unity Technologies; c2019 [cited 2019 Feb 24]. Available from: <https://unity3d.com/>.