

DOI:10.3969/j.issn.1001-0505.2025.05.012

多尺度单目深度信息辅助的铁路桥梁 视觉识别方法

杨涵¹, 徐庆凯², 章金勇², 蒋友², 于泓川¹, 舒江鹏^{1,3}, 徐声亮⁴

(1. 浙江大学建筑工程学院, 杭州 310058; 2. 中铁十二局集团城市发展建设有限公司, 苏州 215163; 3. 浙江大学
长三角智慧绿洲创新中心, 嘉兴 314100; 4. 宁波市政工程建设集团股份有限公司, 宁波 315012)

摘要: 为提高利用深度学习模型进行基础设施检测的计算机视觉任务的准确度和效率,提出了一种多尺度单目深度信息辅助的铁路桥梁视觉识别方法。建立了多任务神经网络,在单目深度图的辅助下,同时完成构件识别和病害分割2项任务。为充分利用深度图与2项任务在不同尺度上的相关性,在多任务神经网络结构上提出了多尺度特征交互机制。建立了多尺度多模态蒸馏模块、跨尺度特征传递模块和特征融合模块,实现了多尺度特征共享和交互,并使用大规模图片数据集对该方法进行验证。结果表明,构件识别和病害分割的平均F1分数分别为93.48%和85.93%,较无深度信息辅助的单任务网络分别提高1.87%和4.41%。与独立训练的特定于任务的网络相比,多任务网络的浮点运算数和预测时长分别降低17.89%和30.48%。由此表明,利用所提方法可以提升铁路桥梁视觉识别的准确度和效率。

关键词: 铁路桥梁; 构件识别; 病害分割; 多任务神经网络; 多尺度; 单目深度图

中图分类号: U24 **文献标志码:** A **文章编号:** 1001-0505(2025)05-1319-09

Multi-scale monocular depth information-assisted railway bridge vision inspection method

YANG Han¹, XU Qingkai², ZHANG Jinyong², JIANG You², YU Hongchuan¹,
SHU Jiangpeng^{1,3}, XU Shengliang⁴

(1. College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China; 2. China Railway 12th Bureau
Group Urban Development and Construction Co., Ltd., Suzhou 215163, China; 3. Innovation Center of Yangtze River Delta, Zhe-
jiang University, Jiaxing 314100, China; 4. Ningbo Municipal Engineering Group Co., Ltd., Ningbo 315012, China)

Abstract: To improve the accuracy and efficiency of computer vision tasks for infrastructure inspection by deep learning models, a multi-scale monocular depth information-assisted railway bridge vision inspection method was proposed. A multi-task neural network was developed. With the assist of the monocular depth map, two tasks including component recognition and damage segmentation were fulfilled simultaneously. To make full use of the correlation at different scales of the depth map and the two tasks, a multi-scale feature interaction strategy was proposed based on the multi-task neural network. The multi-scale multi-modal distillation module, cross-scale feature propagation module, and feature aggregation module were developed to share features and achieve multi-scale feature interaction. A large-scale image dataset was used to validate the proposed method. The results show that the mean F1-scores of component recognition and damage segmentation are 93.48% and 85.93%, respectively, which are 1.87% and 4.41% higher than those of the single-task network without depth information. Compared with task-specific networks trained separately, floating-point op-

收稿日期: 2025-02-28; 网络出版日期: 2025-05-19.

作者简介: 杨涵(1996—),男,博士生;舒江鹏(通信作者),男,博士,研究员,博士生导师, jpesu@zju.edu.cn.

基金项目: 国家自然科学基金资助项目(W2412092);宁波市重点研发计划国际科技合作资助项目(2024H013);浙江省重点研发计划资助项目(2024C01132).

引用格式: 杨涵,徐庆凯,章金勇,等. 多尺度单目深度信息辅助的铁路桥梁视觉识别方法[J]. 东南大学学报(自然科学版),2025,55(5): 1319-1327. DOI:10.3969/j.issn.1001-0505.2025.05.012.

erations and inference duration of the multi-task network are reduced by 17.89% and 30.48%, respectively. Therefore, the proposed method can improve the accuracy and efficiency of railway bridge vision inspection.

Key words: railway bridge; component recognition; damage segmentation; multi-task neural network; multi-scale; monocular depth map

铁路基础设施检测的传统方法包括人工目视检查和接触式传感测量,前者消耗人力物力且效率较低,会造成运营干扰甚至中断;后者依赖于密集传感器矩阵,安装与维护成本较高。近年来,无人机等设备的广泛应用使图像采集愈发便利,结构检测的计算机视觉技术成为研究热点。然而,采集到的海量图像质量良莠不齐,背景复杂,仍需人工整理识别。深度学习模型有助于实现图像信息提取的自动化与智能化,提高效率并降低成本,减少对铁路运营的影响。铁路桥梁的计算机视觉识别主要从构件和局部病害2个方面进行。构件识别是铁路桥梁健康评估的基础,综合考虑构件类型与局部病害信息可评估结构损伤程度,深度学习模型可从图片中识别并分割出桥梁的各类构件^[1-4]。病害检测是结构健康状态评估最重要的步骤之一,当前多种深度学习模型已实现了结构病害的检测、识别与分割^[5-6]。裂缝是结构退化的重要信号,既有研究首先聚焦于裂缝的识别^[7-8],但图像分类及目标检测模型会忽略裂缝形态细节,难以进行定量研究^[9-10];使用语义分割或实例分割模型对图像中的裂缝进行像素级识别是一种更加精细化的方法^[11-13]。此外,深度学习模型还可用于检测识别多类型病害^[14-16]。

用于基础设施视觉识别的深度学习模型的准确度仍需进一步提升。特别是对于图像中通常表现为小目标的局部病害,提升其识别精度是深度学习模型研究中的难点之一。多数深度学习模型是特定于任务的,即独立训练一个模型只能处理单一任务,而在工程实践中,结构健康评估需同时考虑构件类型和局部病害等多类型多模态的信息。利用单任务神经网络处理其他任务时,需重新部署训练模型,浪费计算资源,且检测效率较低。因此,将特定于任务的网络用于铁路桥梁视觉识别时难以满足工程实践中的检测需求。此外,结构视觉识别任务间通常具有一定的相关性,如某类型构件为结构薄弱环节,出现病害的概率较大。忽略此类关联重新训练模型将浪费计算资源,利用任务间相关性则可提高识别精度和效率。因此,建立可同时完成构件和病害识别的多任务深度学习模型,并利用2项任务的相关性来提高识别精度和效率,具有重要的研究价值与实践意义。当前,关于结构视觉识别

的多任务学习模型的相关研究较少。文献[17]利用多任务神经网络同时完成了损伤和材料的检测;文献[18-19]建立了构件和损伤识别的多任务网络。多数多任务深度学习网络都遵循同一个模式,即任务在指定尺度或特定感受野进行特征融合。但在处理视觉任务时,通常会利用任务在不同尺度的相关性来提高判断的准确性^[20]。

单目深度图是一种常用于计算机视觉任务的数据,其中各像素均被标记一个深度值,用以表示此像素对应的物理位置与图像采集点间的相对距离。铁路桥梁的深度图既可提供结构的几何属性,也可评估结构的完整性。深度图与结构视觉识别任务之间具有相关性,例如构件类型的变化会导致相对距离突变。目前,关于单目深度图的结构视觉识别方面的工作鲜见述及。

鉴于此,本文提出了一种多尺度单目深度信息辅助的铁路桥梁视觉识别方法。建立了多任务深度神经网络,在单目深度图的辅助下,同时完成构件识别和病害分割2项任务。为充分利用深度图与2项任务在不同尺度上的相关性,在多任务神经网络结构上提出了多尺度特征交互机制。深度图的辅助与多尺度特征交互机制可提高2项任务的准确度;同时,多任务神经网络可共享任务特征,避免重复运算,提高识别效率,节省计算资源。所提方法可为后续健康状态评估提供参考依据。

1 多任务深度神经网络

1.1 多任务深度神经网络结构

目前,较成熟的用于计算机视觉的多任务深度神经网络结构主要分为以下2类:以编码器为中心和以解码器为中心^[21]。以编码器为中心的网络结构在编码阶段共享任务特征,即在特征传入特定于任务的预测头前完成特征共享。任务的预测结果只并行或顺序生成一次,后续不再优化;典型网络包括多任务注意力网络(multi-task attention networks, MTAN)、完全自适应特征共享网络(fully-adaptive feature sharing, FAFS)等。以解码器为中心的网络结构在解码阶段交换共享特征,网络的初始预测以一次性或递归的方式来改进每个任务的最终输出;典型网络包括 PAD-Net、

PAP-Net、联合任务递归学习网络 (joint task-recursive learning) 等。两者相比,以编码器为中心的网路结构在一个周期中直接预测所有任务输出,无法捕捉任务间的相关性,而这些关联和差异对提高每个任务的预测准确度是非常重要的。因此,本文选择以解码器为中心的网路结构来进行铁路桥梁视觉识别的多任务学习。多任务学习具有以下优势:①由于各任务共用网络的共享层,网络训练和预测时的内存占用较少;②共享层中各任务的特征无需重复计算,效率较高,可节省计算资源;③具有相关性的任务共享特征,各任务的准确度较高,模型具有良好的泛化能力和鲁棒性^[21]。

1.2 多任务多尺度特征交互

以解码器为中心的网路结构遵循同一个模式,即在某特定尺度上进行任务间的特征交互和多模态蒸馏。这种模式默认遵循了一个严格的假设,即所有相关任务都只能使用特定感受野,通过单一过滤器进行特征提取,过滤器尺寸对任务相关性没有影响。然而,这种假设有时无法成立,甚至不符合直观认知。如图1所示,在构件识别的场景中,相较于尺度较大的全局图片块,深度图中尺度较小的局部图片块为构件分割提供的信息较少。感受野扩大时,可以更清晰地呈现出深度图中构件的几何

特征,其中暗含了构件分割的语义信息,为构件类型识别分割提供有益的帮助。另一方面,构件语义分割时,深度图中较小尺度的图片块可以改善边缘的对齐精度。由此可知,不同尺度下任务间的相关性是不同的。因此,需要利用深度图的多种尺度来提高铁路桥梁视觉识别2项任务的预测准确性。本文提出了一种多尺度特征交互机制,用以改进多任务神经网络结构,提高其学习、泛化能力和鲁棒性,形成多任务多尺度特征交互网络。针对铁路桥梁视觉识别,建立了如下3个模块来实现多尺度特征交互^[20]:

(1)多尺度多模态蒸馏模块。如图2所示,在每个尺度上,模型通过骨干网络从输入的图像中提取特征,并进行初始预测,从而在各尺度上获得一组特定于任务的表征。随后,对于每个任务,利用空间注意力机制从其他任务中提取信息来细化特

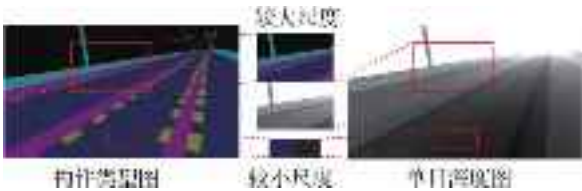


图1 尺度对任务相关性的影响
Fig. 1 Influence of scale on task correlation

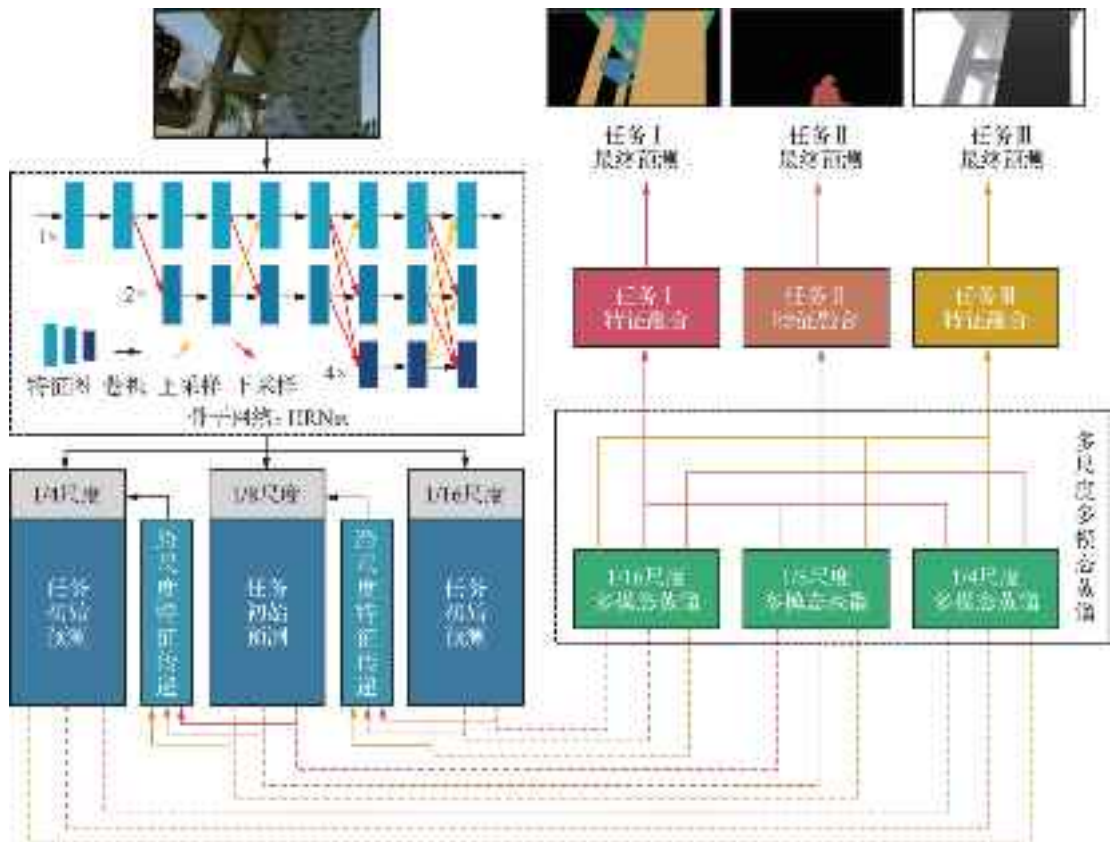


图2 多任务多尺度特征交互网络
Fig. 2 Multi-task multi-scale feature interaction network

征表达。作为门函数,空间注意力机制可筛选并增强来自其他任务的有效语义特征。这一过程被称为多模态蒸馏。在每个尺度上重复进行此操作,便可实现多尺度多模态蒸馏。

(2)跨尺度特征传递模块。通过一个特征协调单元,将较小分辨率尺度的各任务的特征组合成共享特征,用于改进小尺度特征。改进后的任务特征传递至下一尺度对应任务的预测头。跨尺度特征传递模块中的第1步为特征协调操作。将 N 个任务的尺寸为 $C \times H \times W$ 的特征输入该模块并进行连接,得到尺寸为 $NC \times H \times W$ 的特征,使用非线性函数处理

该特征。利用Softmax函数沿任务维度生成注意力掩码,并对特征进行进一步处理,从而将通道数量由 NC 减少至 C 。特征协调操作的最终输出为包含来自所有任务的共享特征。跨尺度特征传递模块中的第2步为特征改进。通过一个卷积操作 F_w ,将尺寸为 $C' \times H' \times W'$ 的特征转换成尺寸为 $C \times H \times W$ 的特征。特征改进利用特定于任务的低计算量剪枝通道门控(channel gating)方法,从共享表征中选择对结果贡献较大的部分来实现。本文网络中,通道门控方法通过挤压-激发(squeeze-and-excitation)块来实现。跨尺度特征传递模块如图3所示。

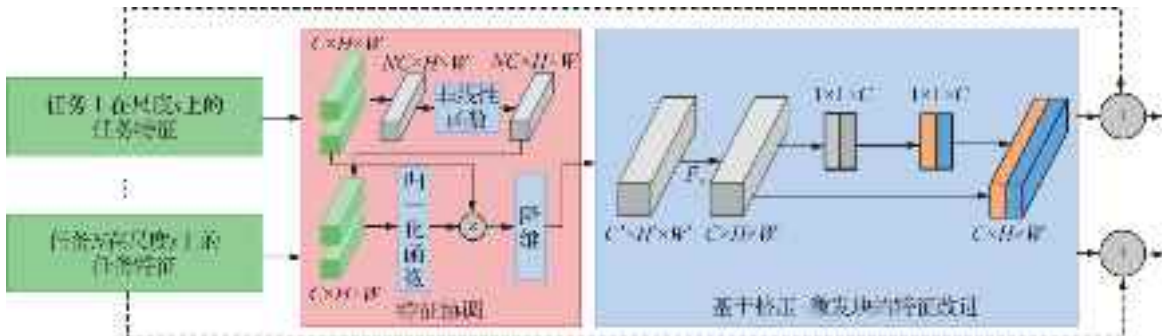


图3 跨尺度特征传递模块
Fig. 3 Cross-scale feature propagation module

(3)特征融合模块。在多尺度多模态蒸馏后,每个尺度的蒸馏特征均上采样至最大尺度并进行连接,得到每个任务的最终表征。通过特定于任务的预测头解码最终表征,从而获得每个任务的最终预测。

HRNet可以保持图像的较高分辨率,不会忽略空间语义特征。因此,将HRNet作为骨干网络,提取特定于任务的特征。

本文提出的多任务多尺度特征交互网络的整体结构如下:首先,骨干网络HRNet在各尺度提取特定于任务的特征,并在每个尺度上进行任务的初始预测;其次,在每个尺度上蒸馏任务特征,捕捉任务间的相关性;然后,连接所有尺度中蒸馏出的特征,进行特征融合,得到各任务的最终预测结果。为进一步提高性能,还建立了跨尺度特征传递模块,将经过蒸馏的特征从较小尺度传递至较大尺度。

1.3 任务亲和度

为使构件识别和病害分割2项任务能够充分利用与单目深度图的相关性和其提供的语义信息,需确定执行特征交互的尺度。为此,需要量化任务共享局部结构的程度和尺度间的关系,即任务相关度。对于同一场景,在视觉识别的常用尺度上计算构件类型标注图、病害标注图、单目深度图三者的相关度。

对于构件识别和病害分割2项任务,当一个像

素对属于同一类型构件或同一种类病害时,认为两者存在关联。对于单目深度图,当一个像素对的深度值差值在阈值以下时,认为两者存在关联。若一个像素对在同一场景的3种图中均存在或不存在关联,则认为该像素对是相关的。相关像素对占所有像素对的比例即为任务相关度^[22]。在各尺度上计算任务相关度,尺度的改变依靠核膨胀实现,各尺度的任务相关度见图4。图中, m 为尺度因子,且尺度 $s = (1/2)^m$ 。由图可知,2项任务与单目深度图在各尺度上的相关度均较高,说明利用多尺度的单目深度图辅助构件识别和病害分割进行多任务学习并完成铁路桥梁视觉识别是可行性的。此外,任

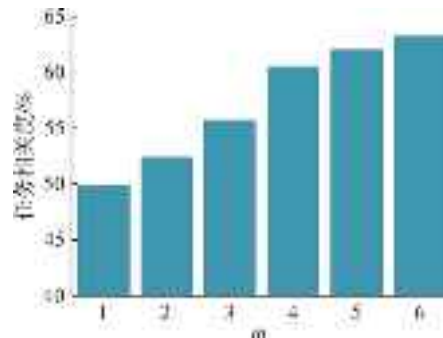


图4 2项任务与单目深度图在不同尺度的任务相关度
Fig. 4 Task correlation of 2 tasks and monocular depth map at different scales

务相关度高度依赖于尺度,从而证明了进行多尺度特征交互的必要性。尺度变大时,任务亲和度下降。大尺度的图像块会造成计算量剧增,过小的尺度则会丢失某些全局的语义信息。因此,选取1/4、1/8、1/16三个尺度进行多任务多尺度特征交互。

2 数据集

使用大规模图像数据集——Tokaido 数据集^[23],验证多任务多尺度特征交互神经网络的有效性。该数据集中,场景和目标结构均由合成环境随机生成。在合成环境中,模拟无人机围绕铁路桥梁飞行,并进行图像采集。该数据集中的图像具有200个不同的合成环境,包含2 000座铁路桥梁。每组数据包含4张图像,除原场景图外,图中每个像素均被分别标注构件类型、病害类型和深度值,从而形成构件类型标注图、病害标注图和单目深度图。对于构件识别任务,数据集中包含7种常见于铁路桥梁的构件,即背景、板、梁、柱、非结构构件、轨道和轨枕。对于病害分割任务,数据中包含的病害种类为混凝土损伤和钢筋暴露。单目深度图中,深度值范围为[0.5,30] m,步长为0.5 m。

数据集共包含1 900组图像,将其随机分成训练集和测试集,其中训练集占75%,测试集占25%,即训练集中包含1 400组图像,测试集中包含500组图像。所有原图及标注图均调整至640×360像素。数据集中具有代表性的原图及其对应的构件分割和深度标注图见图5。

3 模型训练与实验分析

所有实验均在一台配备有Intel Xeon E5-2678 v3 @ 2.50 GHz、64 GB RAM CPU 和 NVIDIA RTX2080TI、11 GB RAM GPU 的服务器上进行。

3.1 网络配置及超参数设置

选用交叉熵损失作为构件识别和病害分割任

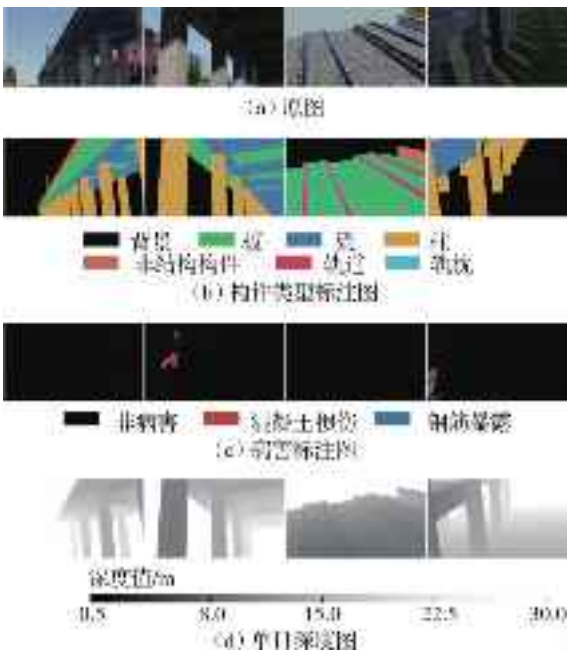


图5 数据集示例
Fig. 5 Examples of dataset

务的损失函数,将L1损失作为相对位置判断任务的损失函数。经实验,在参数空间中得到较合理的训练超参数和选项:批处理大小(batch size)为4,优化器为ADAM,训练轮次(epoch)为250,初始学习率为0.001,学习率衰减周期为50训练轮次,学习率衰减系数为0.5,数据集置乱频率为1训练轮次。为方便下文讨论,将该多任务多尺度特征交互网络命名为Net-M。

为评价网络性能,对于构件识别和病害分割任务,采用精确率、召回率、交并比、F1分数作为评价指标;对于单目深度图,采用均方根误差(RMSE)作为评价指标。

3.2 网络训练、测试结果及分析

多任务网络Net-M中,构件识别任务的评价指标见表1。由表可知,多任务网络的精确率、召回率、交并比、F1分数均较高。对于板、梁、柱3个重

表1 构件识别任务的训练及测试评价指标
Table 1 Training and testing evaluation metrics of component recognition task %

构件类型	训练集				测试集			
	精确率	召回率	交并比	F1分数	精确率	召回率	交并比	F1分数
背景	99.53	99.66	99.20	99.59	99.19	99.31	98.52	99.25
板	98.28	98.17	96.51	98.22	95.58	96.37	92.23	95.97
梁	97.59	97.66	95.36	97.62	94.14	92.44	87.41	93.28
柱	98.85	98.93	97.81	98.89	97.76	97.85	95.70	97.80
非结构构件	97.78	96.16	94.10	96.96	95.58	94.28	90.34	94.93
轨道	95.64	96.80	92.71	96.22	92.91	94.18	87.87	93.54
轨枕	86.80	83.73	74.27	85.24	80.75	78.46	66.10	79.59
平均值	96.35	95.87	92.85	96.11	93.70	93.27	88.31	93.48

要的结构构件,训练集的 4 个评价指标均高于 95%,测试集的 4 个评价指标均高于 87%。对于构件识别,轨枕构件属于较小目标,存在类别不平衡问题,其训练集和测试集的 F1 分数分别达到 85.24% 和 79.59%。训练集中 7 种构件的平均 F1 分数高达 96.11%,测试集的平均 F1 分数及交并比也分别达到 93.48% 和 88.31%。以上数据均表明多任务网络具有良好的学习和泛化能力。

构件识别任务的代表性结果见图 6。由图可知,网络对于构件的预测结果与标注图接近,说明网络输出的预测图在工程实践中是可接受的。

病害分割任务的评价指标见表 2。由表可知,多任务网络 Net-M 对于混凝土损伤和钢筋暴露 2 种病害分割的 F1 分数均高于 60%,训练集的平均 F1 分数和交并比分别为 86.14% 和 75.54%,测试集的平均 F1 分数和交并比分别为 85.93% 和 75.22%,测试集中混凝土损伤和钢筋暴露 2 种病害的交并比均大于 50%。这些评价指标都是阳性

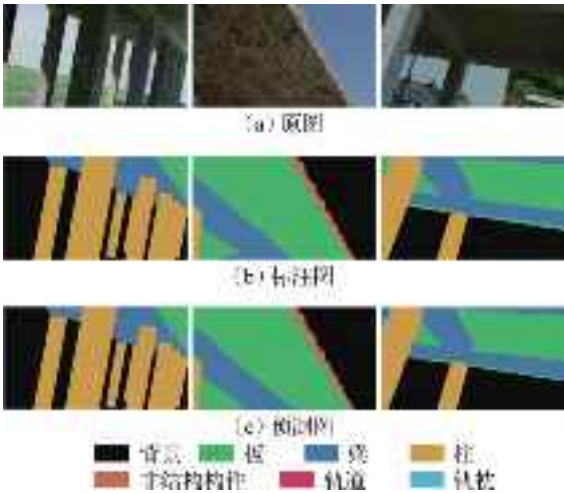


图 6 构件识别任务的代表性结果
Fig. 6 Typical results of component recognition task

样本敏感的,但在病害分割任务里,图像中大部分像素都属于非病害的阴性样本,2 类病害均为小目标,故与构件识别任务相比,这些评价指标的数值均不高。

表 2 病害分割任务的训练及测试评价指标
Table 2 Training and testing evaluation metrics of damage segmentation task %

病害类型	训练集				测试集			
	精确率	召回率	交并比	F1 分数	精确率	召回率	交并比	F1 分数
非病害	99.80	99.38	99.18	99.59	99.79	99.36	99.15	99.57
混凝土损伤	86.21	87.35	76.65	86.78	85.85	87.21	76.25	86.52
钢筋暴露	51.84	96.14	50.78	67.36	51.33	96.05	50.27	66.91
平均值	79.28	94.29	75.54	86.14	78.99	94.21	75.22	85.93

病害分割任务的代表性结果见图 7。由图可知,虽然评价指标数值不高,但可视化后的预测图与标注图相差不大,在工程实践中是可以接受的。

作为辅助任务,多任务网络 Net-M 同样完成了单目深度图中的深度值判断。为综合判断本文提

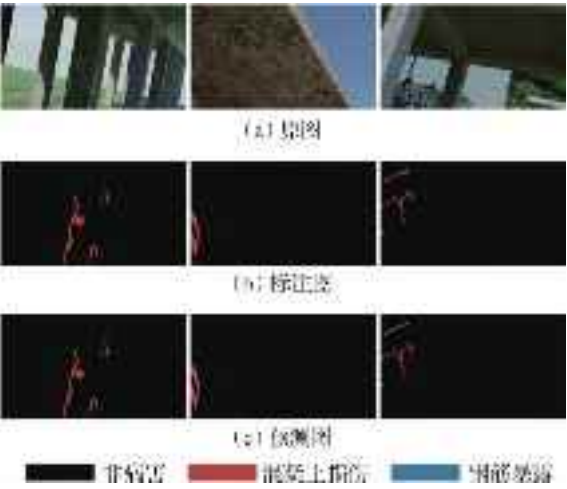


图 7 病害分割任务的代表性结果
Fig. 7 Typical results of damage segmentation task

出的多任务多尺度特征交互网络的优势,量化了深度值判断任务的评价指标,并对预测结果进行了可视化处理。数据集中深度最大值为 30 m,而多任务网络 Net-M 训练集和测试集的均方根误差分别为 0.611 和 1.531 m,在工程实践中是可接受的。多任务网络 Net-M 对于深度值判断的代表性结果见图 8。图 8(d)给出了 Net-M 输出的预测结果与真实深度值的绝对误差,通过色阶映射,直观展示出多任务深度学习方法在深度值判断这一辅助任务上的出色表现。

4 讨论

4.1 与单任务模型对比

为充分说明本文提出的多任务多尺度特征交互网络的优势,通过训练普通单任务神经网络,完成构件识别、病害分割和单目深度图中深度值判断 3 个任务,并将对应的 3 个网络分别命名为 Net-R、Net-S 和 Net-D。单任务网络的基本结构依旧选用 HRNet。

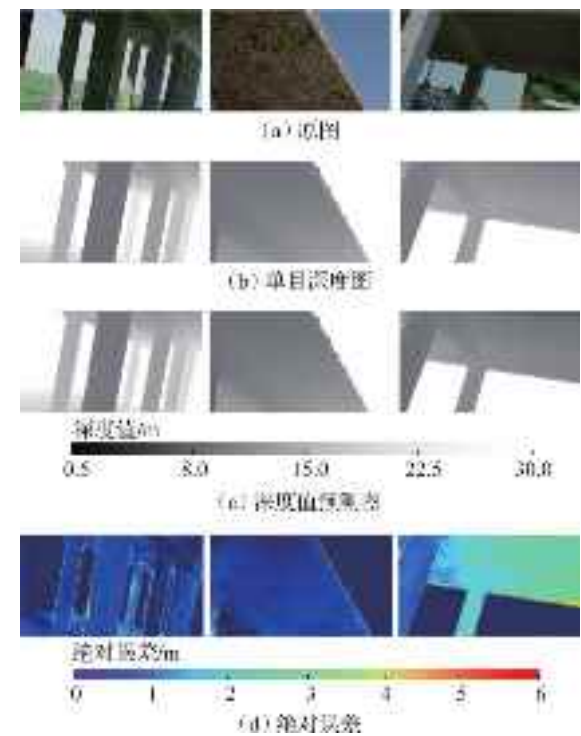


图8 单目深度图中深度值判断的代表性结果
Fig.8 Typical results of depth value determination in monocular depth map

用于构件识别的单任务网络Net-R的评价指标见表3。由表可知,多任务网络训练集和测试集的平均F1分数相较于单任务网络分别提高1.93%和1.87%。多任务网络训练集和测试集中每种构件的F1分数均高于相应单任务网络;数值差距最大的构件为训练集中的轨枕,多任务网络较单任务网络提高3.97%。由此可知,本文提出的多任务网络对于构件识别任务的表现优于传统单任务神经网络。

表3 单任务网络Net-R对于构件识别的F1分数 Table 3 F1-scores of component recognition of single-task network Net-R									
数据集	背景	板	梁	柱	非结构构件	轨道	轨枕	平均值	
训练集	98.55	96.53	95.57	97.67	95.43	94.26	81.27	94.18	
测试集	98.24	94.20	91.21	96.52	93.23	91.62	76.28	91.61	

用于病害分割的单任务网络Net-S的评价指标见表4。由表可知,多任务网络训练集和测试集的平均F1分数相较于单任务网络分别提高3.54%和4.41%。多任务网络中混凝土损伤病害的训练集和测试集F1分数较单任务网络分别提高2.37%和3.38%,而对于钢筋暴露病害,两者分别提高3.50%和4.90%,充分说明了多任务多尺度特征交互网络的优势。

对用于辅助任务深度值判断的单任务网络

表4 单任务网络Net-S对于病害分割的F1分数
Table 4 F1-scores of damage segmentation of single-task network Net-S

数据集	非病害	混凝土损伤	钢筋暴露	平均值
训练集	99.49	84.43	63.86	82.60
测试集	99.43	83.14	62.01	81.52

Net-D的评价指标进行量化处理,训练集和测试集的均方根误差分别为1.022和2.013 m。多任务网络训练集和测试集的均方根误差较单任务网络分别降低40.2%和23.9%。虽然单目深度图中的深度值判断为辅助任务,但多任务多尺度特征交互网络仍可获得较低的均方根误差,从而证明其优秀的学习、泛化能力以及所提方法的有效性。

4.2 模型效率

为进一步说明本文方法的优势,利用训练时间、单次预测时长、模型参数量和浮点运算总数4个指标,量化模型占用的计算资源及效率。训练时间和预测时长较短,表示模型效率较高;模型参数量和浮点运算数较少,表示所用的计算资源较少。多任务模型Net-M和3个单任务模型Net-R、Net-S、Net-D的量化和对比结果见表5。以3个单任务模型的总和为基准,进行对比分析。由表可知,相较于3个模型的总和,所提多任务多尺度特征交互网络的训练时间和单次预测时长分别降低18.64%和30.48%,模型参数量和浮点运算总数分别减少12.85%和17.89%。这是因为多任务网络中的特征无需重复运算,从而节省了计算资源,提高了模型效率。

表5 模型效率分析
Table 5 Model efficiency analysis

模型	训练时间/s	单次预测时长/ms	模型参数量/10 ⁶	浮点运算总数/10 ⁹
Net-R、Net-S和Net-D	16 728.59	105	17.9	12.3
Net-M	13 610.22	73	15.6	10.1

5 结论

(1) 提出的多任务多尺度特征交互网络可以同时完成铁路桥梁构件识别和病害分割2项任务。多尺度特征交互机制有效利用了2项任务与单目深度图在不同尺度的相关性,且任务间共享特征。多任务网络对于构件识别和病害分割的测试集平均F1分数较单任务网络分别提高1.87%和4.41%。本文方法有效提高了铁路桥梁视觉识别任务的准确度。

(2) 提出的多任务网络结构可共享特征,避免重复运算,从而提高了铁路桥梁视觉识别的效率,

节省了计算资源。与独立训练的特定于任务的网络相比,多任务多尺度特征交互网络的单次预测时长降低30.48%,浮点运算总数减少17.89%。

(3) 作为辅助任务,多任务模型完成了单目深度图中的深度值判断,准确度较高。多任务模型的测试集均方根误差为1.531 m,相较于单任务网络降低23.9%,再次证明了所提网络在识别准确度方面的优势。

参考文献(References)

- [1] 杜文康,王浩,雷冬,等. 基于机器视觉和边缘重构的拉索振动响应识别[J/OL]. 东南大学学报(自然科学版), 2025[2025-04-16]. <http://kns.cnki.net/kcms/detail/32.1178.N.20250226.1007.002.html>.
DU W K, WANG H, LEI D, et al. Vibration response identification of cable based on machine vision and edge reconstruction[J/OL]. Journal of Southeast University (Natural Science Edition), 2025[2025-04-16]. <http://kns.cnki.net/kcms/detail/32.1178.N.20250226.1007.002.html>. (in Chinese)
- [2] JIA J W, NI Y H, MAO J X, et al. Geometric parameter identification of bridge precast box girder sections based on deep learning and computer vision[J]. Journal of Southeast University (English Edition), 2025, 41(3): 278-285.
- [3] NARAZAKI Y, HOSKERE V, CHOWDHARY G, et al. Vision-based navigation planning for autonomous post-earthquake inspection of reinforced concrete railway viaducts using unmanned aerial vehicles[J]. Automation in Construction, 2022, 137: 104214.
- [4] KARIM M M, QIN R W, CHEN G D, et al. A semi-supervised self-training method to develop assistive intelligence for segmenting multiclass bridge elements from inspection videos[J]. Structural Health Monitoring, 2022, 21(3): 835-852.
- [5] SPENCER B F, HOSKERE V, NARAZAKI Y. Advances in computer vision-based civil infrastructure inspection and monitoring[J]. Engineering, 2019, 5(2): 199-222.
- [6] 刘宇飞,冯楚乔,陈伟乐,等. 基于机器视觉法的桥梁表观病害检测研究综述[J]. 中国公路学报, 2024, 37(2): 1-15.
LIU Y F, FENG C Q, CHEN W L, et al. Review of bridge apparent defect inspection based on machine vision[J]. China Journal of Highway and Transport, 2024, 37(2): 1-15. (in Chinese)
- [7] 刘宇飞,樊健生,聂建国,等. 结构表面裂缝数字图像法识别研究综述与前景展望[J]. 土木工程学报, 2021, 54(6): 79-98.
LIU Y F, FAN J S, NIE J G, et al. Review and prospect of digital-image-based crack detection of structure surface[J]. China Civil Engineering Journal, 2021, 54(6): 79-98. (in Chinese)
- [8] 邓露,褚鸿鹄,龙砺芝,等. 基于深度学习的土木基础设施裂缝检测综述[J]. 中国公路学报, 2023, 36(2): 1-21.
DENG L, CHU H H, LONG L Z, et al. Review of deep learning-based crack detection for civil infrastructures[J]. China Journal of Highway and Transport, 2023, 36(2): 1-21. (in Chinese)
- [9] 崔弥达,王超,陈金桥,等. 基于ROS及YOLOv3的混凝土桥梁裂缝实时检测系统[J]. 东南大学学报(自然科学版), 2023, 53(1): 61-66.
CUI M D, WANG C, CHEN J Q, et al. Real-time concrete bridge cracks detection system based on ROS and YOLOv3[J]. Journal of Southeast University (Natural Science Edition), 2023, 53(1): 61-66. (in Chinese)
- [10] CHA Y J, CHOI W, BÜYÜKÖZTÜRK O. Deep learning-based crack damage detection using convolutional neural networks[J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(5): 361-378.
- [11] 倪有豪,陆欢,季超,等. 基于语义分割的桥梁锈蚀病害识别对比分析[J]. 东南大学学报(自然科学版), 2023, 53(2): 201-209.
NI Y H, LU H, JI C, et al. Comparative analysis on bridge corrosion damage detection based on semantic segmentation[J]. Journal of Southeast University (Natural Science Edition), 2023, 53(2): 201-209. (in Chinese)
- [12] 乔朋,梁志强,段长江,等. 基于改进Faster R-CNN与U-Net算法的桥梁病害识别与量化方法[J]. 东南大学学报(自然科学版), 2024, 54(3): 627-638.
QIAO P, LIANG Z Q, DUAN C J, et al. Bridge defects detection and quantifying method based on modified Faster R-CNN and U-Net[J]. Journal of Southeast University (Natural Science Edition), 2024, 54(3): 627-638. (in Chinese)
- [13] 丁威,夏哲,舒江鹏,等. 基于负压吸附爬壁机器人和Transformer的混凝土桥塔裂缝识别检测[J]. 中国公路学报, 2024, 37(2): 53-64.
DING W, XIA Z, SHU J P, et al. Recognition and detection of concrete bridge tower cracks using a negative pressure adhesion wall-climbing robot and transformer[J]. China Journal of Highway and Transport, 2024, 37(2): 53-64. (in Chinese)
- [14] LI S, ZHAO X F, ZHOU G Y. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network[J]. Computer-Aided Civil and Infrastructure Engineering, 2019, 34(7): 616-634.
- [15] 陈江,原野,郎洪,等. 基于多分支深度学习的沥青路面多病害检测方法[J]. 东南大学学报(自然科学版), 2023, 53(1): 123-129.
CHEN J, YUAN Y, LANG H, et al. Multi-distress detection method for asphalt pavements based on

<http://journal.seu.edu.cn>

- multi-branch deep learning [J]. Journal of Southeast University (Natural Science Edition), 2023, 53(1): 123-129. (in Chinese)
- [16] WANG S, ZHANG L J, YIN G J. Defect identification method for steel surfaces based on improved YOLOv5 [J]. Journal of Southeast University (English Edition), 2024, 40(1): 49-57.
- [17] HOSKERE V, NARAZAKI Y, HOANG T A, et al. MaDnet: Multi-task semantic segmentation of multiple types of structural materials and damage in images of civil infrastructure [J]. Journal of Civil Structural Health Monitoring, 2020, 10(5): 757-773.
- [18] YE X W, MA S Y, LIU Z X, et al. Post-earthquake damage recognition and condition assessment of bridges using UAV integrated with deep learning approach [J]. Structural Control and Health Monitoring, 2022, 29(12): e3128.
- [19] XU Y, QIAO W D, ZHAO J, et al. Vision-based multi-level synthetical evaluation of seismic damage for RC structural components: A multi-task learning approach [J]. Earthquake Engineering and Engineering Vibration, 2023, 22(1): 69-85.
- [20] VANDENHENDE S, GEORGOULIS S, VAN GOOL L. MTI-Net: Multi-scale task interaction networks for multi-task learning [C]// European Conference on Computer Vision. Glasgow, UK, 2020: 527-543.
- [21] VANDENHENDE S, GEORGOULIS S, VAN GANSBEKE W, et al. Multi-task learning for dense prediction tasks: A survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(7): 3614-3633.
- [22] ZHANG Z Y, CUI Z, XU C Y, et al. Pattern-affinitive propagation across depth, surface normal and semantic segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: 2019: 4101-4110.
- [23] NARAZAKI Y, HOSKERE V, YOSHIDA K, et al. Synthetic environments for vision-based structural condition assessment of Japanese high-speed railway viaducts [J]. Mechanical Systems and Signal Processing, 2021, 160: 107850.