

基于图像语义分割的半监督裂纹检测方法

刘培^{1,2}, 黄雅平^{1,2*}

¹ 北京交通大学计算机与信息技术学院交通数据分析与挖掘北京市重点实验室, 北京 100044;

² 北京交通大学计算机与信息技术学院, 北京 100044

摘要 基于深度学习的裂纹检测方法严重依赖大量的像素级标注信息, 为此提出一种基于半监督学习的裂纹检测方法。该方法将多尺度模块引入到裂纹检测的网络模型中, 仅利用小部分的像素级标注数据进行全监督训练。对于无标签数据, 融合多种显著性区域检测方法生成伪标签, 可以减少对像素级标注信息的依赖。在裂纹数据集上对改进网络进行实验验证, 并与常用语义分割网络和弱监督实验基准从主观评价、精度、召回率和 F1-score 的角度进行比较。实验结果表明, 改进网络可以有效提升裂纹的识别准确率, 提出的半监督训练策略在仅需 6.25% 像素级标注信息的情况下, 能够取得与全监督方法相当的识别精度和召回率。

关键词 机器视觉; 裂纹检测; 语义分割; 多尺度模块; 半监督

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.0615004

Semi-Supervised Crack-Detection Method Based on Image-Semantic Segmentation

Liu Pei^{1,2}, Huang Yaping^{1,2*}

¹ Beijing Key Lab of Traffic Data Analysis and Mining, School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China;

² School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

Abstract The crack-detection method based on deep learning relies heavily on a large amount of pixel-level annotation information. Thus, a crack-detection method based on semi-supervised learning is proposed. The proposed method introduces multiscale modules into the network model of crack-detection. It uses only a small part of pixel-level annotation data for fully-supervised training. For the unlabeled data, the fusion of multiple saliency area detection methods to generate pseudolabels can reduce pixel-level reliance on the labeled information. The improved network is experimentally verified on the crack dataset. It is compared with the commonly used semantic segmentation network and weakly supervised experimental benchmarks from the perspective of subjective evaluation, accuracy, recall rate, and F1-score. The experimental results show that the improved network can effectively enhance crack recognition accuracy. The proposed semi-supervised training method can achieve recognition accuracy and recall rate equivalent to the fully-supervised method when only 6.25% pixel-level label information is required.

Key words machine vision; crack-detection; semantic segmentation; multi-scale module; semi-supervision

OCIS codes 150.0155; 150.1135; 150.3040

1 引言

裂纹是一种十分常见且严重影响建筑物使用寿命

的缺陷, 及时准确地检测裂纹对建筑物及道路桥梁等维护工作至关重要。随着计算机视觉技术的发展, 自动化检测的方法逐渐代替了人工巡检的方式。

收稿日期: 2020-07-10; 修回日期: 2020-08-07; 录用日期: 2020-09-03

* E-mail: yphuang@bjtu.edu.cn

在裂纹检测的任务中,传统方法主要包含阈值分割^[1-2]、边缘检测^[3-4]、最小路径搜索^[5]和小波变换^[6]等。基于手工特征的聚类^[7]、随机森林^[8]和支持向量机^[9]等机器学习方法在裂纹检测的任务中取得了较大的进步,但这些方法仍然存在一定的局限性,如忽略了裂纹宽度。随着深度学习的发展,基于语义分割^[10]、像素块分类^[11]或显著性检测^[12]等方法陆续被提出。随着全卷积神经网络^[13]的提出,图像语义分割在识别准确率上获得了十分显著的提升,受到了科研人员的广泛关注。

目前,基于图像语义分割的裂纹检测方法的检测性能得到了很大的提升,但是仍然存在着许多不足。如基于图像语义分割的裂纹检测方法,通常直接使用已有的通用语义分割网络在表面裂纹的数据集上进行训练,但未对裂纹检测任务中表面裂纹的特征进行针对性的设计,仅有少量工作对裂纹检测方法进行改进,如 DeepCrack^[14],但还是未充分利用裂纹不同尺度的上下文信息,所以在裂纹的识别准确率方面有进一步提升的空间;基于深度学习的裂纹检测方法大多为全监督学习方式,需要在大规模的像素级标注的数据集上进行训练,然而像素级标注的这一过程十分费时费力。因此,如何减少对标注信息的依赖是裂纹检测方法在实际应用中需要解决的一个关键问题。目前,基于对抗生成的半监督语义分割方法使用小部分手工标注的数据集进行训练,对未标注的图像生成伪标签并参与网络训练,该方法的训练过程复杂,不易收敛,因此在裂纹检测

任务中需要研究更有效的半监督学习方法以克服对像素级标签的依赖。

鉴于此,本文提出一种简单有效的半监督训练策略,其与采用对抗生成思想使用部分手工标注的数据集进行训练生成伪标签的方法不同,利用简单易得的多种无监督算法生成的预测结果融合成伪标签,以此简化伪标签生成的训练过程,解决对抗网络的训练困难以及难以收敛的问题。在训练过程中引入部分手工标注的数据集,给出正确预测的准确定义,从而提高网络的预测和判别能力。同时,在 DeepCrack 网络的基础上,提出基于多尺度的裂纹检测模块(MSCM),使用该模块对改进后的网络进行半监督训练。

2 半监督裂纹检测方法

深度学习中的全监督学习方式严重依赖数据集所提供的标注信息,所以如何减少对标注信息的依赖是深度学习在实际使用中所面临的挑战之一。为了简化半监督学习方式生成伪标签的训练过程,提出融合多种显著性区域检测方法的半监督训练框架,半监督训练框架如图 1 所示,其中 GC 为 Global Contrast, SF 为 Saliency Filters, wCtr 为 Weighted Contrast。首先对无标签数据采用多种无监督方法生成预测结果,然后对其进行融合生成伪标签后参与网络训练,最后在裂纹检测网络 DeepCrack 中引入 MSCM,使用该模块对改进后的网络进行半监督训练。

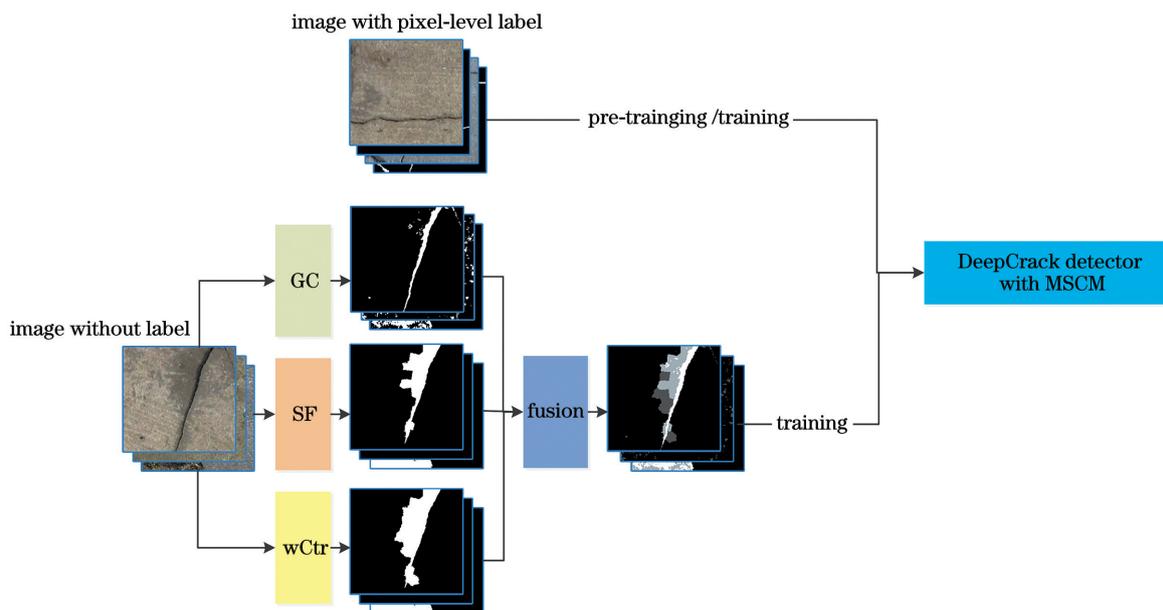


图 1 半监督训练框架

Fig. 1 Semi-supervised training framework

2.1 DeepCrack 网络框架

实验选取 DeepCrack 作为基础网络,其由编码网络和解码网络组成,同时包含深监督机制^[15]。DeepCrack 是一种全卷积网络,输入一张任意尺寸的 RGB(Red, Green, Blue)三通道图像,则输出一个单通道的预测结果,其中预测结果经过激活函数计算后在对应的像素位置处输出该像素属于裂纹的概率。与传统的卷积神经网络不同,全卷积网络结构中利用卷积层替代了原本的全连接层,以此保留图像特征的空间信息来达到定位的目的。在 DeepCrack 网络中,编码网络为去掉全连接层的 VGG 16(Visual Geometry Group 16),使用该网络对图像特征进行提取,其包含 13 个卷积核大小为 3×3 的卷积层(步长为 1)和 5 个最大池化层。在网络级联的卷积层与池化层中,图像特征的尺寸逐渐缩小,感受野逐渐扩大,特征抽象的程度不断提升。解码网络的结构与编码网络相对称,不同的则是编码网络中的池化层改为解码网络中的上采样层。根据下采样阶段保存的最大池化元素位置来索引信息并对其进行非线性上采样,从而保证最大元素恢复到原有位置,而其他位置用零来填充。上采样后得到的稀疏特征经过卷积之后得到稠密特征,可以避免上采样阶段的学习过程,使得网络参数减少。

采用最大池化索引的方式对目标进行上采样,使得上采样阶段尺寸较小的目标被忽略,但 DeepCrack 网络引入了深监督机制,其可以充分利用编码与解码网络中的各阶段特征并对其进行融合监督。首先融合每一个对应阶段编码网络中池化层前卷积层输出的特征,以及解码网络中卷积层输出的特征,经过跳层结构对特征进行融合和上采样,获得一个与输入图像尺寸相同的单通道输出,将各个阶段的融合结果通过跳层结构进一步融合,得到的单通道输出作为网络最终的输出结果。跳层结构由一个卷积核为 1×1 的卷积层和一个上采样层组成,使用卷积核为 1×1 的卷积层对融合特征进行降维以得到单通道输出,上采样层将得到的单通道输出上采样到输入图像的尺寸。

深监督机制中,由于裂纹检测中只包含两个类别,所以各阶段的融合输出采用的损失函数均为二元交叉熵(BCE)函数,表达式为

$$L_{\text{BCE}}(F_i; W) = \sum_{i=1}^I \left\{ \omega_1 \cdot y \text{lb} [V(F_i; W)] + \omega_2 \cdot (1 - y) \text{lb} [1 - V(F_i; W)] \right\}, \quad (1)$$

式中: I 表示数据集中图像总数; W 表示网络参数; F_i 表示第 i 阶段的特征融合输出; ω_1 和 ω_2 表示各类别所赋予的权重; y 表示真实值; V 表示预测值。

结合深监督机制中各阶段融合输出的损失以及网络最终输出结果的损失,则网络总的损失函数可以表示为

$$L(W) = \sum_{i=1}^I \left[\sum_{k=1}^K L_{\text{BCE}}(F_i^{(k)}; W) + L_{\text{BCE}}(F_i^{\text{fusion}}; W) \right], \quad (2)$$

式中: K 表示网络中的融合阶段总数; F_i^{fusion} 表示各阶段特征融合输出的最终融合结果,同样为网络的最终输出。

2.2 MSCM

在不同尺度的感受野下,不同尺寸的裂纹,响应程度不同,考虑到 DeepCrack 网络中对裂纹的上下文信息利用不够充分,仅有几个不同尺寸且级联递增的感受野。为了进一步利用不同尺度的上下文信息,提高不同尺寸与形状的表面裂纹的识别准确率,提出基于多尺度的 MSCM,使用该模块提高裂纹检测的性能。随着图像特征在级联的卷积层与池化层中的传递,图像特征尺寸逐渐减小,感受野逐渐增大,抽象程度逐渐增高,这有利于分类效果的提升,但是同时也丢失了一定的空间信息,不利于目标的精确定位。为了解决这一矛盾,利用膨胀卷积在保证图像特征尺寸不会减小的前提下获得更多的上下文信息。膨胀卷积最初在一维信号中使用,表达式为

$$x_1(j) = \sum_{s=1}^S x(j + r \cdot s) \cdot \omega_k(s), \quad (3)$$

式中: j 表示一维信号上第 j 个被卷积的窗口; x 表示输入的一维信号; x_1 表示输出; ω_k 表示尺寸为 S 的卷积核; r 表示膨胀间距。

膨胀卷积在保持卷积核原有参数数目的前提下,通过在原有卷积核的间隙插入零的方式来扩大卷积核的尺寸以获得更大的感受野,很好地解决了图像特征减小与感受野增大的这一矛盾。在改进网络中引入的 MSCM,则是通过引入不同膨胀率的膨胀卷积层来获得不同尺度的上下文信息,通过引入全局平均池化层来获得图像级特征。MSCM 结构如图 2 所示。MSCM 包含 4 个具有不同膨胀率的膨胀卷积层和一个全局平均池化层与上采样层。当膨胀卷积层的膨胀率为 1 时,则其等同于普通的卷积层。通过多尺度模块来保证图像特征尺寸在不受影响的前提下,获得了更多尺度的上下文信息,使得

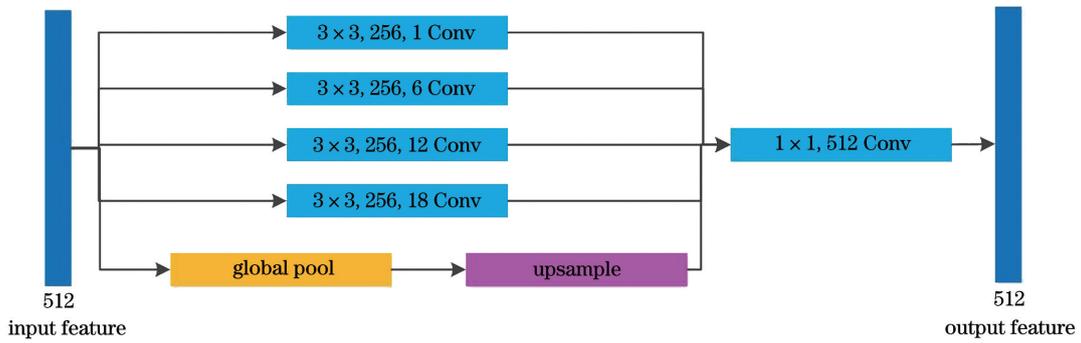


图 2 MSCM 结构

Fig. 2 MSCM structure

分类效果得到进一步提升,更好地识别不同尺寸与 形状的 表面裂纹。改进后的网络框架如图 3 所示。

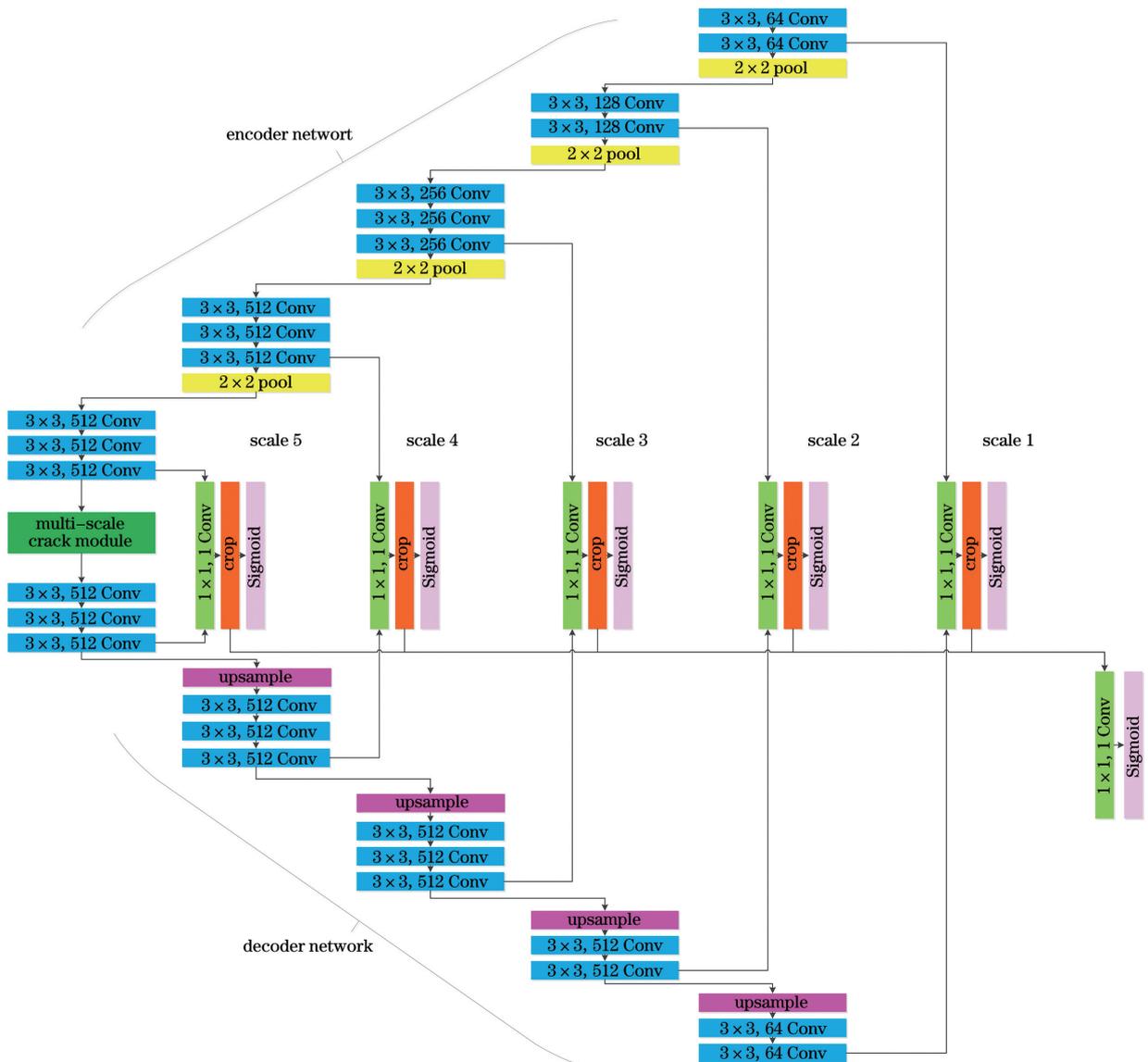


图 3 改进后的网络框架

Fig. 3 Improved network framework

2.3 半监督训练策略

目前,基于图像语义分割的裂纹检测方法通常

采用全监督训练的方式,在数据集中提供原始图像的同时还需要提供像素级标注的标签信息,但手工

标注的这一过程十分费时费力,代价高昂。为了缓解对标签数据的依赖,在语义分割任务中很多研究人员关注半监督语义分割任务。根据对抗生成的思想利用对抗生成网络在小规模的数据集上进行训练后,对未标注图像进行预测,预测结果作为伪标签并参与网络的训练。考虑到以上方法存在训练不易收敛和训练过程慢等缺点,提出一种简单有效的半监督训练策略。该策略将由简单易得的无监督方法生成的预测结果经过融合操作后作为伪标签,同时使少量的手工标签参与训练,在保证识别准确率的前提下极大地减少手工标注的工作。这一策略的核心思想是根据卷积神经网络强大的学习能力,利用部分的手工标注数据集给出正确预测的准确定义,并将其与大量融合后的伪标签进行训练,从而提升网络的抗噪声能力,进而获得与全监督方法相当的检测效果。

综上,实验提出基于多种无监督显著性区域检测方法融合生成伪标签并参与训练的半监督训练策略。该训练策略选取了 SF^[16]、wCtr^[17]和 GC^[18]三种无监督显著性区域检测方法,并作为生成伪标签的基础方法,用来对无标签数据进行显著性区域检测,检测结果作为该方法对无标签图像中裂纹的检测结果。获得无监督检测结果后,对其进行二值化处理并将其划分为裂纹与背景以提高召回率。二值化过程可表示为

$$B = \begin{cases} 1, & \hat{B} > \sigma \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

式中: σ 表示二值化处理所设定的阈值; \hat{B} 表示无监督方法中样本的预测值; B 表示二值化后取值;0表示背景;1表示裂纹。

得到预测结果后,根据融合策略对其进行融合,将融合所得的结果作为无标签数据的伪标签。在裂纹检测的问题中只包含背景和裂纹两个类别,所以融合后所获得的伪标签也可以代表所属裂纹类别的置信度,融合结果越大表示该样本越可能是裂纹。融合策略可表示为

$$C_{\text{fusion}} = \sum_{l=1}^L a_l \cdot b_l, \quad (5)$$

式中: L 表示无监督方法的个数; b_l 表示第 l 种无监督方法经二值化后的预测结果; a_l 表示第 l 种无监督方法的融合系数; C_{fusion} 表示融合后的伪标签。

提出的半监督训练策略选取改进后的网络作为基础网络并对其进行训练。网络训练过程中,如果未能正确预测并给出准确定义,尤其对于困难样本

来说,卷积神经网络很难在训练过程中获得较好的预测判别结果,所以训练过程中利用少量的全监督标记样本数据集进行预训练,接着迁移到包含少量的全监督数据集和大量的无监督数据集上进行训练,从而提升网络对困难样本的预测判别能力和网络的抗噪声能力。

3 实验与分析

3.1 实验设置与数据集

实验环境为 64 位 Ubuntu 操作系统, GeForce RTX 2080 Ti GPU;软件方面采用 PyTorch 深度学习框架,编程语言为 Python3.6。以改进后的卷积神经网络作为基本网络,在自建的数据集上构建表面裂纹检测模型,采用 Adam 算法进行训练调优。训练过程中,为了增强手工标注标签的作用以及提升网络的预测能力,在有全监督数据集参与训练的实验中,先将网络在全监督数据集中进行预训练,接着将其迁移到包含少量的全监督数据集和大量的无监督数据集中进行训练,从而提升网络在大规模数据集中的预测能力和抗噪声能力。检测过程中,利用划分出的测试集来测试网络模型的裂纹检测能力。

根据调研可知,目前的道路裂纹数据集中含有大量的标注噪声,并不能真实评测算法的效率,所以实验收集整理了道路裂纹数据集(RCD),使用该裂纹数据集对目前公开的数据集进行整理和修正。考虑到背景纹理对裂纹检测任务的影响,所以收集的 RCD 中包含多种纹理背景的表面裂纹图像,同时数据集中还包含横向、纵向和网状等多种形状与尺寸的表面裂纹,可以保证自建数据集中裂纹的多样性以及裂纹识别的难度。自建表面数据集中共包含 350 张图像,图像的分辨率为 296 pixel × 306 pixel。在数据增强阶段,图像尺寸调整为 224 pixel × 224 pixel 并输入网络中。数据集划分为训练集、验证集和测试集,分别有 210 张、70 张和 70 张图像。由于数据集的样本数较少,所以训练前的训练集需经过数据增强处理,包含旋转、翻转和剪切等手段,最终无标签训练集包含 9000 张表面裂纹图像。因为裂纹像素在图像中占比较小,所以存在类别不均衡的问题,实验通过在二元交叉熵损失函数中赋予不同类别不同的权重来均衡在损失函数中具有不同比例的两个类别。

实验所采用的无监督方法分别为 SF^[16]、wCtr^[17]和 GC^[18],这些方法在众多的无监督方法中

检测效果较好,生成的预测结果经过二值化处理 and 融合后作为伪标签并参与网络训练,图 4 为对应的二值化处理后的效果。实验所采取的定量评价指标为图像语义分割问题中常用的精度、召回率和 F1-score 等,其中 F1-score 综合衡量了裂纹的检测效果,表达式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (6)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (7)$$

$$x_{F1-score} = 2 \cdot \frac{P \cdot R}{P + R}, \quad (8)$$

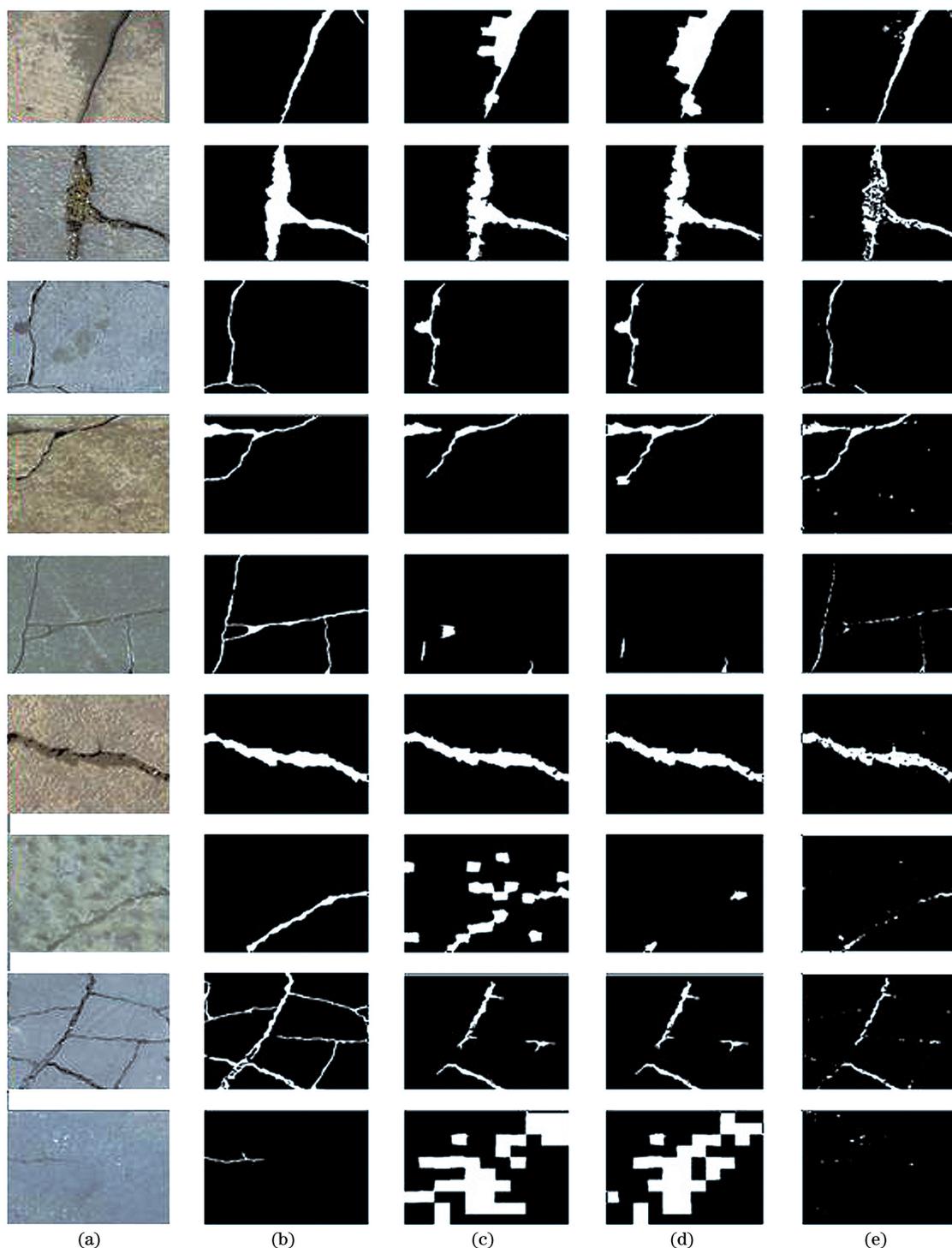


图 4 不同方法得到的伪标签。(a)原图;(b)真值图;(c)SF 方法;(d)wCtr 方法;(e)GC 方法
Fig. 4 Pseudo tags obtained by different methods. (a) Original images; (b) ground-truth; (c) SF method;
(d) wCtr method; (e) GC method

式中： N_{TP} 表示正样本被正确识别为正样本的数量； N_{FP} 表示负样本被识别为正样本的数量； N_{FN} 为正样本被识别为负样本的数量。综合手工标注的误差以及表面裂纹自身的特点，认定裂纹预测结果在距离正样本的 2 pixel 以内同样被认定为正样本。

3.2 实验分析

为了验证所提出的半监督训练策略的有效性，将半监督训练策略与常用语义分割网络 SegNet^[19] 和 DeepCrack 的全监督策略在 RCD 上进行比较，手工标签数据集的图像数目与弱标签数据集比例为

1/15，定量分析结果如表 1 所示。对不同类型和纹理背景的裂纹检测效果如图 5 所示。从表 1 可以看到，所提出的半监督训练策略在仅需 6.25% 像素级标注信息的情况下，能够取得与常用语义分割网络

表 1 不同网络的裂纹检测结果

Table 1 Crack-detection results of different networks

Network	Precision / %	Recall / %	F1-score
SegNet ^[19]	85.23	89.77	86.74
DeepCrack ^[14]	81.71	92.88	86.26
Proposed network	80.29	93.56	85.66

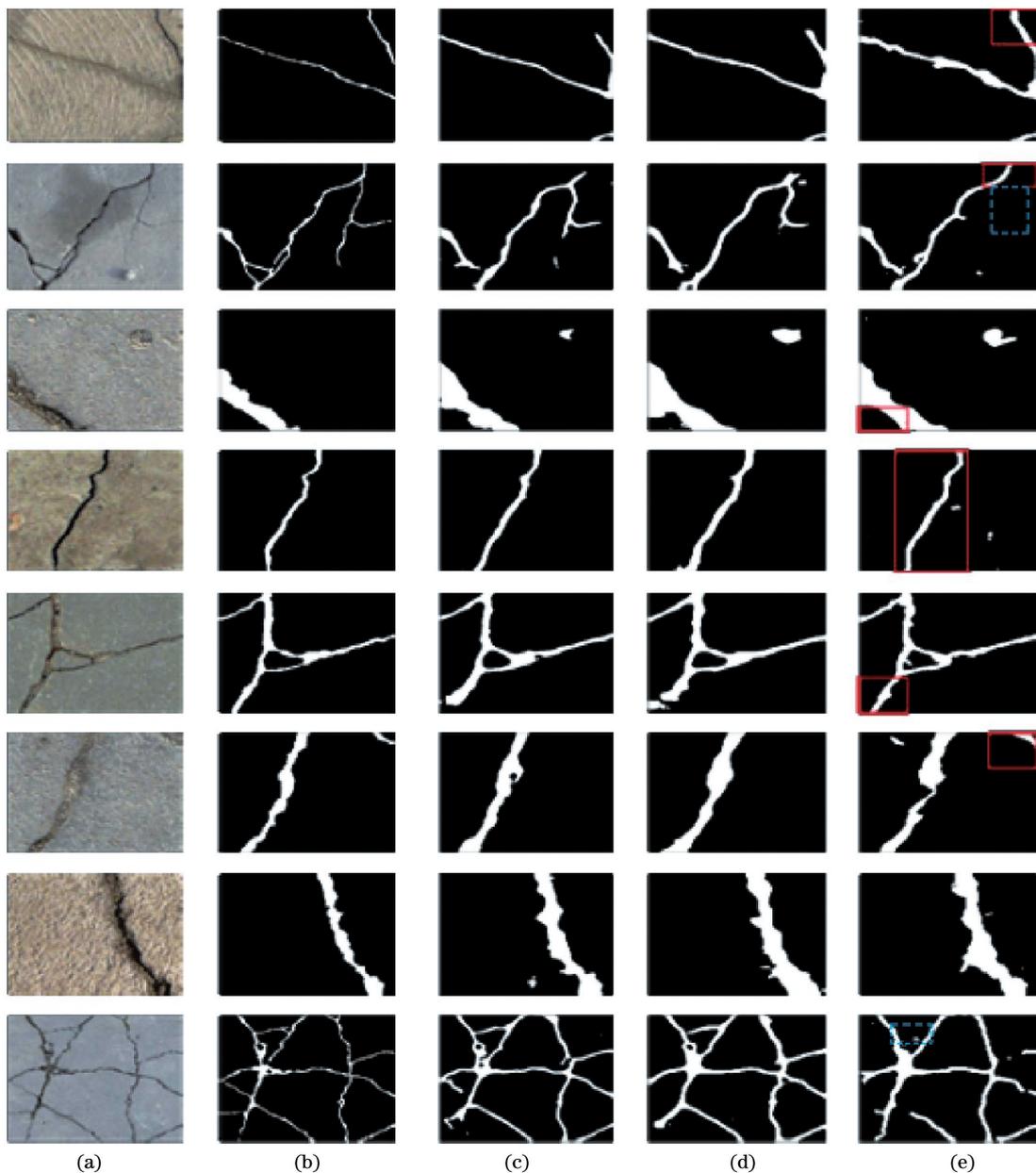


图 5 不同网络下的裂纹检测效果。(a)原图；(b)真值图；(c)SegNet 网络；(d)DeepCrack 网络；(e)所提网络

Fig. 5 Crack-detection effect under different networks. (a) Original images; (b) ground-truth; (c) SegNet network; (d) DeepCrack network; (e) proposed network

SegNet 和 DeepCrack 全监督策略相当的识别精度、召回率和检测效果,所提出的半监督训练策略可以有效减少裂纹检测任务中的数据标注。从图 5 可以看到,所提出的半监督训练策略的检测结果在整体上稍优于 SegNet 和 DeepCrack 的全监督策略,在实线框中表现出明显优势,说明半监督训练策略可以较为准确地识别全监督训练策略错误遗漏的部分,而虚线框为所提出的半监督训练策略错误遗漏的部分,结合图 4 可以看到,伪标签中对细小裂纹处的标注结果不准确,从而导致在所提出的半监督训练策略下,细小裂纹处的检测结果可能出现偏差。

为了进一步研究所提出的半监督训练策略在不同比例的标签数据和融合策略对实验结果的影响,以及多尺度模块在网络改进中的作用,进行如下扩展实验。

1) 不同比例的标签数据的性能

为了研究所提出的半监督训练策略中不同比例的手工标注数据集对检测效果的影响,构建了手工标注数据集与伪标签数据集在不同比例条件下的对比实验。表 2 为三种方法的定量分析结果。表 3 为手工标注数据集与伪标签数据集在不同比例条件下的定量分析结果。表 4 为在不同比例的手工标注数据集上预训练后的定量分析结果。从表 2~4 可以

表 2 三种方法的定量分析结果

Table 2 Quantitative analysis results of three methods

Method	Precision / %	Recall / %	F1-score
SF ^[16]	76.87	72.37	69.22
wCtr ^[17]	74.33	71.30	67.34
GC ^[18]	86.57	66.93	70.00

表 3 手工标注数据集与伪标签数据集在不同比例条件下的定量分析结果

Table 3 Quantitative analysis results of manual labeled dataset and pseudo labeled dataset in different proportions

Proportion	Precision / %	Recall / %	F1-score
0	72.93	85.21	74.52
1/150	72.74	88.62	76.59
1/65	77.48	88.50	80.47
1/30	77.68	92.67	83.39
1/15	80.29	93.56	85.66
1/6	85.99	92.17	88.42
1	89.34	97.48	92.31

表 4 不同比例的手工标注数据集上预训练后的定量分析结果

Table 4 Quantitative analysis results after pre-training on different proportions of manually annotated datasets

Proportion	Precision / %	Recall / %	F1-score
1/150	14.26	50.35	21.46
1/65	25.85	60.38	35.04
1/30	49.83	63.80	54.81
1/15	69.48	83.51	75.21
1/6	82.16	91.59	86.22

看到,伪标签数据集参与训练对网络检测效果有极大的提升效果,精度和召回率提升效果明显;当手工标注数据集与伪标签数据集的比例达到 1/15 及以上时,虽然预训练后的测试结果已经较高,但在伪标签数据集上的训练可以进一步提升网络的检测效果和泛化能力,同时因为伪标签标注不够准确且包含一定的噪音,所以也提升了网络的抗噪声能力和鲁棒性;所提出的半监督训练策略中,不同比例的手工标注数据集对网络的识别检测效果的提升均有一定的效果,F1-score 均高于三种无监督方法,且随着所占比例的提升,提升效果增强,精度和召回率提升明显。图 6 为手工标注数据集与伪标签数据集在不同比例情况下的半监督裂纹检测效果。从图 6 可以看到,随着手工标注标签所占比例的增大,所提出的半监督裂纹检测效果不断提升,背景与裂纹的判别和裂纹边界的定位更准确。

2) 三种无监督方法的标签融合策略

为了研究对比多种无监督显著性区域检测方法融合策略中融合比例的影响,构建了在不同融合比例下的对比实验。表 5 为 SF、wCtr 和 GC 方法在不同融合系数下的裂纹检测结果,其中 union 表示三种方法的预测结果中一种方法预测该像素为裂纹,则伪标签中将该像素标注为裂纹。结合表 2 和表 5 可以看到,网络检测的召回率较高而精度相对较低,当融合过程中给予精度较高而召回率较低的无监督检测结果更大的融合权重时,对网络检测效果的提升较为明显,但为了减少对单一无监督方法的依赖,尤其是当选用的无监督方法较多时,综合多种无监督方法的预测结果,所以实验默认选用 1 : 1 : 1 的融合比例。

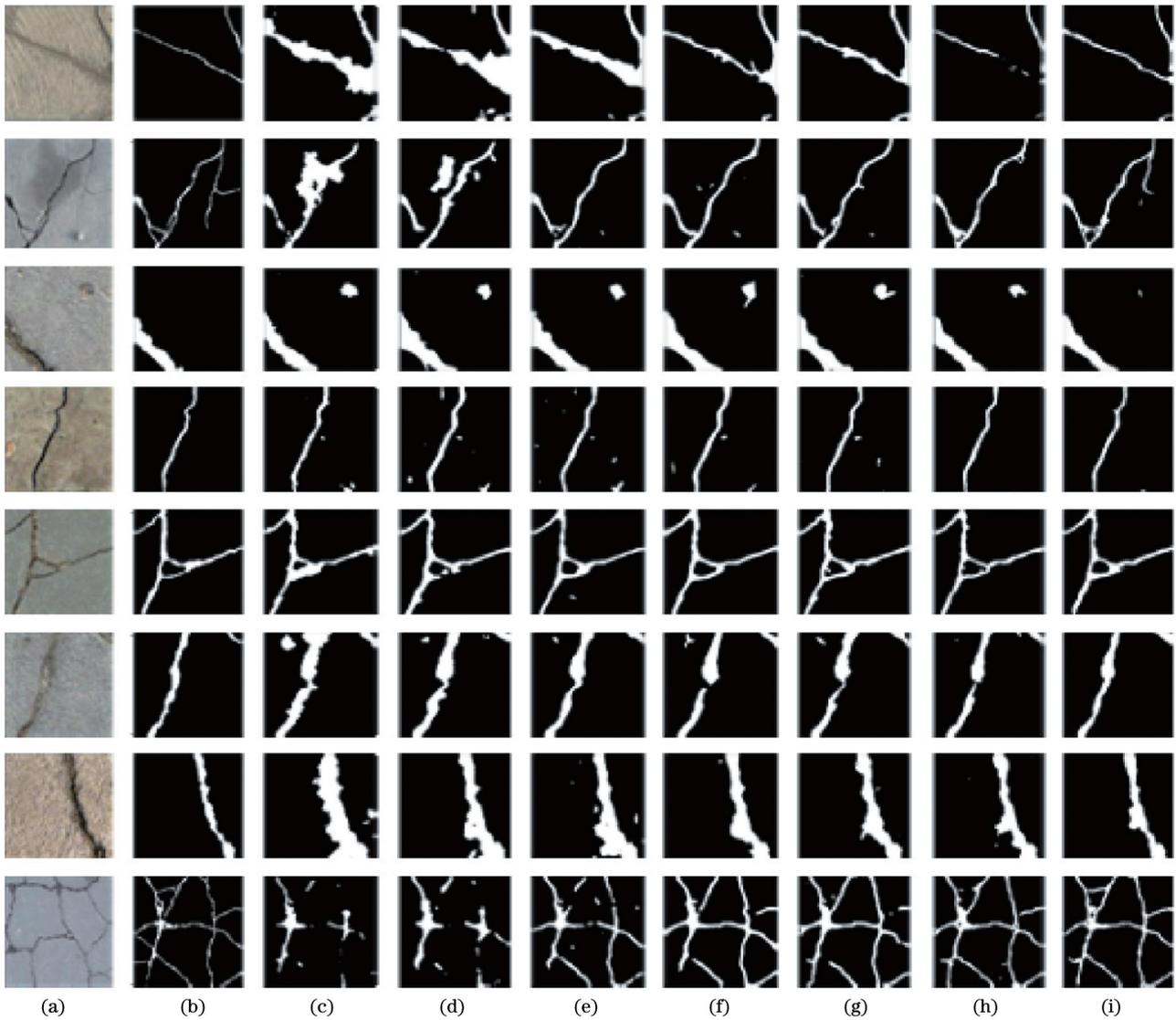


图 6 手工标签与伪标签的不同比例参与训练后的裂纹检测效果。(a)原图;(b)真值图;(c)0;(d)1/150;(e)1/65;(f)1/30;(g)1/15;(h)1/6;(i)1

Fig. 6 Crack detection effect after training with different proportion of manual tag and pseudo tag. (a) Original images; (b) ground-truth; (c) 0; (d) 1/150; (e) 1/65; (f) 1/30; (g) 1/15; (h) 1/6; (i) 1

表 5 SF, wCtr 和 GC 方法在不同融合系数下的裂纹检测结果

Table 5 Crack detection results of SF, wCtr and GC methods under different fusion coefficients

Fusion coefficient	Precision /%	Recall /%	F1-score
1 : 0 : 0	81.78	89.81	84.81
0 : 1 : 0	81.07	88.14	83.68
0 : 0 : 1	85.32	95.03	89.34
1 : 1 : 1	80.29	93.56	85.66
1 : 1 : 2	85.93	91.33	87.69
Union	80.39	91.67	84.77

3) 多尺度模块的影响

为了研究 MSCM 对裂纹检测网络的提升效果,构建了全监督条件下常用语义分割网络 SegNet 和 DeepCrack, 以及引入 MSCM 后改进网络的对比实验。表 6 为在自建表面裂纹数据集

表 6 不同网络在全监督方式下的裂纹检测结果
Table 6 Crack detection results of different networks under full supervision

Network	Precision /%	Recall /%	F1-score
SegNet ^[19]	85.23	89.77	86.74
DeepCrack ^[14]	81.71	92.88	86.26
Proposed network	89.34	96.48	93.31

RCD上,常用语义分割网络 SegNet 和 DeepCrack 以及改进后的网络在全监督方式下的定量分析结果。从表 6 可以看到,所提出的改进网络在精度、召回率和 F1-score 上有显著的提升,相比于实验基准网络 DeepCrack 分别提升 7.63,3.60,7.05 个百分点,检测效果优于 SegNet 与 DeepCrack。图 7 为不同类型裂纹的检测效果。从图 7 可以看到,

改进后的网络在细小裂纹和特大尺寸裂纹的识别准确率上均有较大的提升,实线框中的部分提升效果明显,裂纹边缘定位更精准,对困难样本的判别更准确。实验定性与定量结果均证明引入的 MSCM 可以有效利用不同尺度的上下文信息,使得对不同尺度和形状裂纹的检测效果有一定的提升。

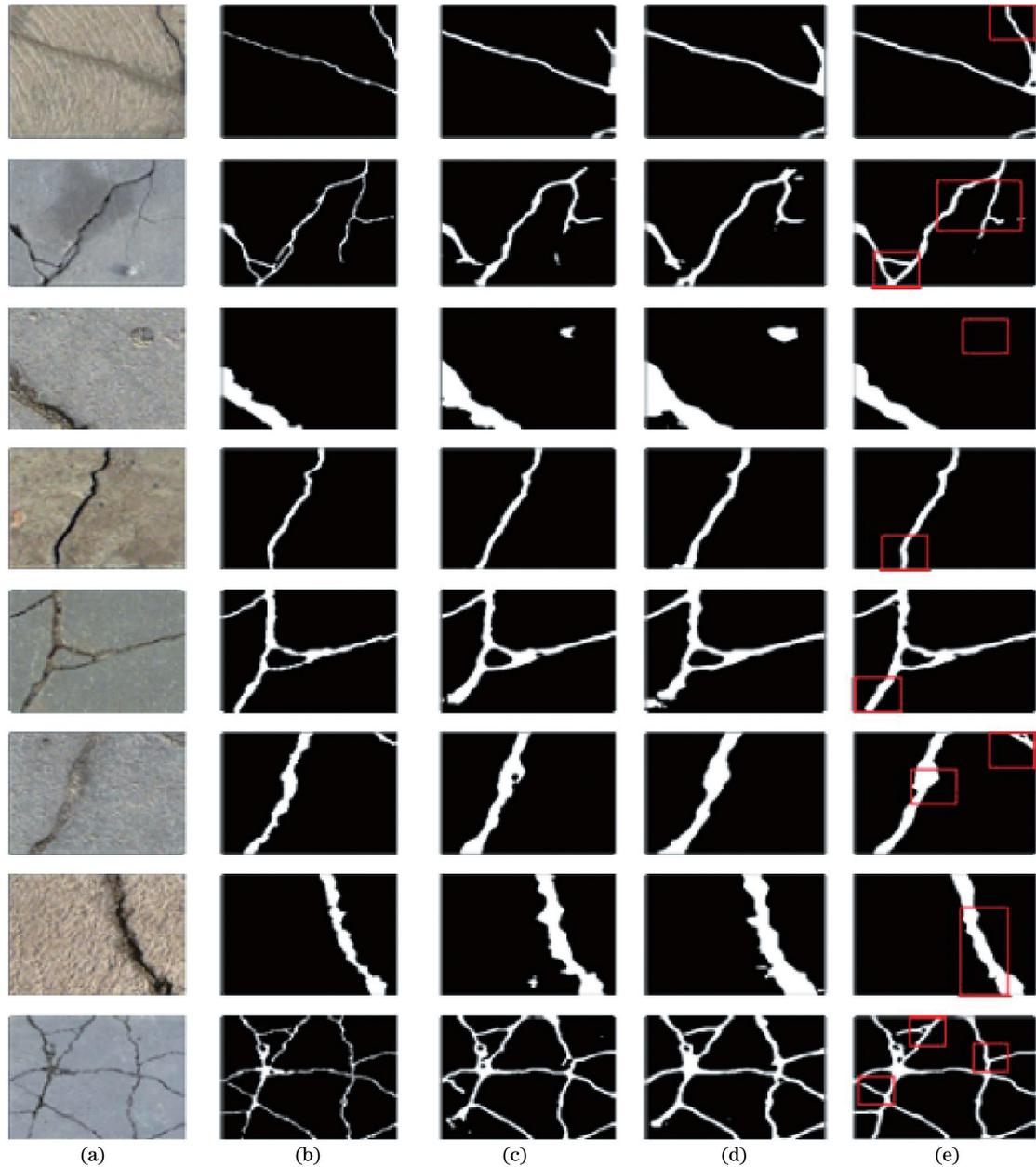


图 7 不同网络在全监督方式下不同裂纹的检测效果。(a)原图;(b)真值图;(c)SegNet 网络;(d)DeepCrack 网络;(e)所提网络

Fig. 7 Detection effect of different cracks in different networks under full supervision. (a) Original images; (b) ground-truth; (c) SegNet network; (d) DeepCrack network; (e) proposed network

综上所述,所提出的半监督训练策略在只有少量手工标注标签参与训练的条件下,可以取得较好

的裂纹检测效果,其中利用无监督融合策略生成的伪标签数据集不仅提升了网络的检测效果,还可以

提高网络的抗噪声能力。当参与训练手工标注数据集达到一定的规模时,半监督策略训练所获得的网络模型的检测能力能够达到部分全监督网络的检测效果。裂纹检测网络中引入的 MSCM 可以有效利用了不同尺度的上下文信息,使得网络的检测效果有一定程度的提升。

4 结 论

针对表面裂纹的检测问题,提出一种简单有效的半监督训练策略,采用无监督方法生成的预测融合结果作为伪标签并参与卷积神经网络的训练,在保证裂纹检测效果的前提下极大地减少手工标注的工作。在基础网络中,根据膨胀卷积的特点引入 MSCM,可以提高网络对不同形状和尺寸裂纹的检测能力。在现有利用图像语义分割解决表面裂纹检测的工作中,全监督方式下改进后网络的检测精度、召回率和 F1-score 分别提升了 7.63, 3.60, 7.02 个百分点,半监督方式下只利用了极少部分的手工标签,便可以取得与全监督方法相当的识别精度和召回率。未来的工作将围绕进一步降低手工标注数据集所占的比例,以及提高网络在伪标签数据集中训练后的识别效果。

参 考 文 献

- [1] Li Q Q, Liu X L. Novel approach to pavement image segmentation based on neighboring difference histogram method[C]//2008 Congress on Image and Signal Processing, May 27-30, 2008, Sanya, Hainan, China. New York: IEEE Press, 2008: 792-796.
- [2] Li Q Q, Zou Q, Zhang D Q, et al. FoSA: F* seed-growing approach for crack-line detection from pavement images[J]. Image and Vision Computing, 2011, 29(12): 861-872.
- [3] Arbeláez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
- [4] Zhao H L, Qin G F, Wang X J. Improvement of canny algorithm based on pavement edge detection [C]//2010 3rd International Congress on Image and Signal Processing, October 16-18, 2010, Yantai, China. New York: IEEE Press, 2010: 964-967.
- [5] Amhaz R, Chambon S, Idier J, et al. Automatic crack detection on two-dimensional pavement images: an algorithm based on minimal path selection [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10): 2718-2729.
- [6] Subirats P, Dumoulin J, Legeay V, et al. Automation of pavement surface crack detection using the continuous wavelet transform [C] // 2006 International Conference on Image Processing, October 8-11, 2006, Atlanta, GA, USA. New York: IEEE Press, 2006: 3037-3040.
- [7] Oliveira H, Correia P L. Automatic road crack detection and characterization[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(1): 155-168.
- [8] Shi Y, Cui L M, Qi Z Q, et al. Automatic road crack detection using random structured forests[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(12): 3434-3445.
- [9] Qu Z, Bai L, An S Q, et al. Lining seam elimination algorithm and surface crack detection in concrete tunnel lining [J]. Journal of Electronic Imaging, 2016, 25(6): 063004.
- [10] Yang X C, Li H, Yu Y T, et al. Automatic pixel-level crack detection and measurement using fully convolutional network[J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(12): 1090-1109.
- [11] Zhang L, Yang F, Zhang Y, et al. Road crack detection using deep convolutional neural network [C]//2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 3708-3712.
- [12] Xu W, Tang Z M, Zhou J, et al. Pavement crack detection based on saliency and statistical features [C]//2013 IEEE International Conference on Image Processing, September 15-18, 2013, Melbourne, VIC, Australia. New York: IEEE Press, 2013: 4093-4097.
- [13] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [14] Zou Q, Zhang Z, Li Q Q, et al. DeepCrack: learning hierarchical convolutional features for crack detection [J]. IEEE Transactions on Image Processing, 2019, 28(3): 1498-1512.
- [15] Lee C, Xie S, Gallagher P W, et al. Deeply-supervised nets [EB/OL]. (2014-09-25) [2020-07-10]. <https://arxiv.org/abs/1409.5185v2>.
- [16] Perazzi F, Krähenbühl P, Pritch Y, et al. Saliency filters: contrast based filtering for salient region detection [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press,

- 2012: 733-740.
- [17] Zhu W J, Liang S, Wei Y C, et al. Saliency optimization from robust background detection[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 2814-2821.
- [18] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection [C] // CVPR 2011, June 20-25, 2011, Providence, RI, USA. New York: IEEE Press, 2011: 409-416.
- [19] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (12): 2481-2495.