光电技术及应用

DOI：10.16818/j.issn1001－5868.2024010304

基于图像处理的路面裂缝特征提取方法

代少升[1]，毛兴华[1]，余自安[2*]

（1. 重庆邮电大学 通信与信息工程学院，重庆 400065；2. 昆明云内动力股份有限公司，昆明 650200）

摘　要： 裂缝是一种常见的路面损坏形式，及时、准确地检测和评估路面裂缝的状态，对于路况监测和养护决策具有重要的意义。文章提出了一种基于图像处理的路面裂缝特征提取方法，首先，对分割得到的路面二值图像进行连通域分析，利用裂缝区域和干扰区域之间的特征差异，筛选出真正代表裂缝的连通域；然后，采用改进的快速并行细化算法提取出裂缝骨架，以端点所在分支的长度为标准来去除骨架中的毛刺，针对骨架断裂和向内缩短的问题，通过端点方向以及端点间距离来进行骨架的连接和生长。实验结果表明，该方法能够有效去除分割结果中的干扰区域，提取出清晰完整且无毛刺的裂缝单一像素骨架，较好地反映了裂缝的主要结构和形态特征。

关键词： 路面裂缝；图像处理；连通域分析；骨架提取

中图分类号： TP391.41　**文章编号：** 1001－5868(2024)03－0508－07

Pavement Crack Feature Extraction Based on Image Processing

DAI Shaosheng[1], MAO Xinghua[1], YU Zi'an[2]

(1. School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, CHN; 2. Kunming Yunnei Power Co., Ltd., Kunming 650200, CHN)

Abstract: Cracks are a common type of pavement damage. Thus, the accurate and timely detection and evaluation of the status of pavement cracks are crucial for road condition monitoring and maintenance decision-making. This study proposes a pavement crack feature extraction method based on image processing. First, we performed a connected-domain analysis on the segmented pavement binary image, whereby the connected domains representing cracks were selected based on the feature differences between the crack and interference regions. Subsequently, we applied an improved fast parallel-thinning algorithm to extract the crack skeleton, and the branch length of the endpoint was used as the standard to remove burrs in the skeleton. To address the problems of skeletal fracture and inward shortening, we used the direction of the endpoints and the distance between them to connect and grow the skeleton. The experimental results show that this method can effectively remove interfering regions from the segmentation results, thereby extracting a clear and complete single-pixel skeleton of cracks without burrs to better reflect the main structural and morphological characteristics of the crack.

Key words: pavement cracks; image processing; connected-domain analysis; skeleton extraction

0　引言

随着城市化进程的加速，道路交通网络的规模和复杂性不断增加，道路的质量和安全性对于确保交通流畅、减少交通事故、提高公众生活质量等方面起着至关重要的作用。然而，由于各种自然和人为因素（如气候变化、交通负荷、施工质量等）的影响，道路表面经常会出现裂缝、坑洞、沉降等病害[1]。路

收稿日期：2024－01－03.

* **通信作者：**余自安

面裂缝是一种常见的路面损坏形式，它不仅影响道路的平整性和美观性，还会降低道路的使用寿命和安全性，给交通运输和道路养护带来巨大的困难和成本。这些裂缝如果不及时检测和修复，可能会导致道路结构的进一步恶化，甚至可能引发交通事故。因此，及时、准确地检测和评估路面裂缝的状态，对于路况监测和养护决策具有重要的意义[2]。

传统的道路裂缝检测方法主要依赖于人工视觉检查，这种方法不仅效率低，而且结果受到检查员经验和主观判断的影响，无法保证准确性和一致性，难以满足大规模、高精度的路面裂缝检测需求。近年来，随着计算机视觉和深度学习技术的发展，自动道路裂缝检测技术[3]已经取得了显著的进步，这些技术可以自动从大量的道路图像中检测和识别裂缝，大大提高了检测的效率和准确性。

现有研究主要集中在裂缝的检测和识别上，为了获取裂缝的形态和结构等特征，本文首先使用深度学习的语义分割算法对路面图像进行像素级别的分类，将裂缝从背景中分离出来；然后，利用形态学运算对分割后的二值图像进行处理，以提高分割结果的质量；接下来，对形态学处理后的分割图像进行连通域分析，获取每个连通域的特征信息，并利用这些特征将干扰区域滤除；最后，通过改进的图像细化算法提取裂缝骨架，使用端点特征对裂缝骨架进行优化，去除骨架中的短小毛刺并完善裂缝骨架结构。

1　裂缝二值图像获取

1.1　路面裂缝分割

图像分割是一种将图像划分为不同区域的技术，它可以根据像素的灰度、纹理、边缘等特征，将图像中具有相似特征的像素归为一类，从而实现对图像中感兴趣的目标的提取。基于传统图像处理技术的方法[4]通常需要人为设定一些参数或阈值，且对图像的质量和噪声敏感，难以适应路面图像的复杂性和多变性。基于深度学习的方法[5]可以自动地学习图像的特征，提高分割的准确性和鲁棒性。本文采用了基于深度学习的语义分割方法，利用深度神经网络对路面图像进行像素级的分类，从而将裂缝从背景中分离出来。

1.2　形态学处理

数学形态学是一种基于集合论和拓扑学的图像分析方法。膨胀和腐蚀是形态学中的基本操作，其数学表示如下：

$$(I \oplus S)(x,y)=\max_{(i,j)\in S} I(x\pm i,y\pm j) \quad (1)$$

$$(I \ominus S)(x,y)=\min_{(i,j)\in S} I(x\pm i,y\pm j) \quad (2)$$

式中，I 表示输入图像，S 表示结构元素，$\oplus$表示膨胀操作，即用给定结构元素的图像局部最大值替换目标像素，从而扩大了对象边界；$\ominus$表示腐蚀操作，即用图像局部最小值替换目标边界像素，从而缩小了对象边界。

分割后得到的二值图像中通常会包含一些小的噪声和孔洞且目标边缘呈锯齿状。首先使用闭运算进行处理，它是对图像先膨胀后腐蚀的操作：

$$I \cdot S=(I \oplus S) \ominus S \quad (3)$$

之后再应用开运算处理，它是对图像先腐蚀后膨胀的操作：

$$I \circ S=(I \ominus S) \oplus S \quad (4)$$

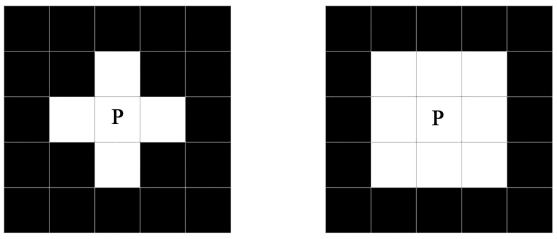
本文采用 5×5 的矩形结构元素进行运算，通过形态学处理，可以消除分割结果中的噪点，填补裂缝中的小孔洞，并平滑裂缝边缘。

2　裂缝特征提取方法

2.1　连通域分析

路面图像分割结果中可能呈现出多条裂缝，为了区分不同裂缝个体以及获取每条裂缝的特征信息，对分割后的二值图像进行连通域分析。

在图像中，所有像素值相同且位置也相邻的像素点形成了一个连通域，每个连通域对应于图像中的一个对象。像素点之间常见的邻接关系有两种，如图 1 所示，在四邻接的情况下，一个像素点仅与其上下左右四个像素点相邻；在八邻接的情况下，一个像素点与其周围的八个像素点相邻，不同的邻接关系会使得连通域的形状和大小产生一定的差异。



(a) 四连通　　(b) 八连通

图 1　像素邻接关系

连通域标记为每个连通域分配一个唯一的标识符，以使它们能够被区分，本文采用八连通关系进行标记，具体步骤如下：

(1)创建一个与待标记图像相同大小的标签图，其中所有像素都被初始化为零，同时创建一个空栈，

用于存储像素点的坐标。

(2)扫描分割图像,寻找一个未被标记的前景像素点。

(3)将该像素点作为起始种子并标记,然后将其相邻的未标记且未在栈中的像素点都压入栈中。

(4)弹出栈顶像素并标记,然后将其相邻的未标记且未在栈中的像素点都压入栈中;

(5)重复步骤(4),直到栈为空,即标记出了图像中的一个连通域。

(6)改变标签值以区分不同的连通区域,重复步骤(2)到(5),直到所有像素点都被标记。

通过上述步骤,可以将图像中的所有前景区域都标记出来,为每个连通域赋予了一个唯一的标签。

在标记了分割结果图像中的所有连通域之后,利用裂缝连通域和干扰连通域之间的特征差异[6]来对它们进行区分:

(1)灰度特征:裂缝像素点通常具有较低的灰度值,与周围区域相比,它们显得更暗。在原始路面灰度图像中,裂缝连通域对应的区域内的平均灰度值往往低于整个图像的平均灰度值。

(2)面积特征:裂缝区域通常比干扰区域有更大的面积,干扰区域往往是一些随机的小块组成,而裂缝则可能贯穿整个图片。通过统计连通域中的像素点的个数来代表其面积 S,剔除掉那些面积较小的区域。

(3)形状特征:裂缝的形状通常是蜿蜒曲折的,而非完全笔直的和规整的,其矩形度 R 和圆形度 C 比较小。矩形度为连通域的面积 S 与其最小外接矩形的面积 S_{minRect} 之比,圆形度为连通域的面积 S 与其最小外接圆的面积 S_{minCir} 之比,这两个指标越接近 1,表示连通域的形状越接近矩形或圆形,反之则越不规则。

$$R=\frac{S}{S_{\text{minRect}}} \tag{5}$$

$$C=\frac{S}{S_{\text{minCir}}} \tag{6}$$

通过对这些特征的综合分析,能够有效区分裂缝连通域和干扰连通域。

2.2 裂缝骨架提取

骨架是表示图像中对象的主要结构特征的一系列连接的线条,它反映了对象的整体形态和走向。在成功标记出路面图像中的裂缝目标后,通过裂缝骨架的提取,可以在保留裂缝主要拓扑结构的同时减少图像中的冗余信息,使得裂缝的关键特征更容易被捕捉和分析。在常用的骨架提取算法中,Zhang-Suen 算法无法保证骨架为单一像素[7],中轴变换算法会产生较多的毛刺[8]。

图像细化是一种通过迭代删除物体边界上的像素同时保持连通性,直到只剩下单一像素宽度骨架的方法。本文采用改进的快速并行细化算法[9]来提取裂缝骨架,该算法利用目标像素点的八邻域像素点之间的关系来决定该点是否需要被保留。中心像素点 P_0 及其八邻域点符号表示,如图 2 所示:

P_8	P_1	P_2
P_7	P_0	P_3
P_6	P_5	P_4

图 2 八邻域符号表示

如果一个裂缝像素点满足以下所有条件,则表明该像素点为外层的边界像素点,对其进行标记:

(1)$C(P_0)=1$;

(2)$2\leqslant N(P_0)\leqslant 3$;

(3)对于奇数次迭代:$(P_1 \vee P_2 \vee \widetilde{P}_4)\wedge P_3=0$;

(4)对于偶数次迭代:$(P_5 \vee P_6 \vee \widetilde{P}_8)\wedge P_7=0$;其中,符号 $\vee$ 表示逻辑或,符号 $\wedge$ 表示逻辑与,符号 $\sim$ 表示逻辑非,$C(P_0)$ 表示在 P_0 的八邻域中,出现邻边点(P_1,P_3,P_5,P_7)的值为 0 且延顺时针方向的两个相邻点至少有一个值为 1 的次数,计算公式如下:

$$\begin{aligned}C(P_0)=&\widetilde{P}_1\wedge(P_2\vee P_3)+\widetilde{P}_3\wedge(P_4\vee P_5)+\\&\widetilde{P}_5\wedge(P_6\vee P_7)+\widetilde{P}_7\wedge(P_8\vee P_1)\end{aligned} \tag{7}$$

$N(P_0)$ 表示两组相邻像素中,每一组包含像素 1 的对数的较小值,使用以下公式计算:

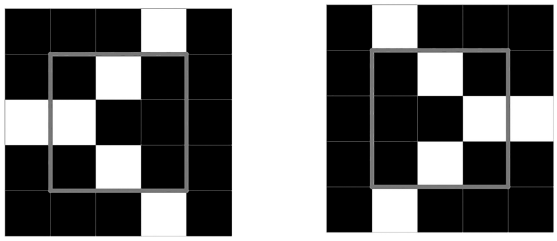
$$N(P_0)=\min(N_1(P_0),N_2(P_0)) \tag{8}$$

$$\begin{aligned}N_1(P_0)=&(P_1\vee P_2)+(P_3\vee P_4)+\\&(P_5\vee P_6)+(P_7\vee P_8)\end{aligned} \tag{9}$$

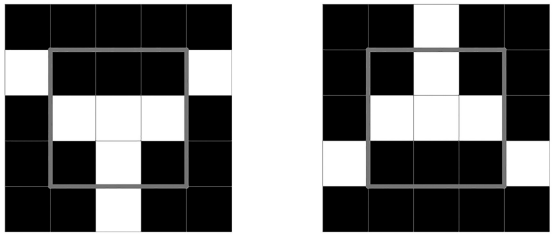
$$\begin{aligned}N_2(P_0)=&(P_1\vee P_8)+(P_7\vee P_6)+\\&(P_5\vee P_4)+(P_3\vee P_2)\end{aligned} \tag{10}$$

完成一次迭代后,将所有标记点删除,即置 0 转为背景点,当没有任何裂缝像素点满足条件时,迭代结束,图像中保留的点构成了裂缝骨架。

提取出的骨架在左 T 形和右 T 形结构处可以保持单一像素宽度,但在上下 T 形结构处无法保证为单一像素,如图 3 所示。



(a) 左 T 形　(b) 右 T 形



(c) 下 T 形　(d) 上 T 形

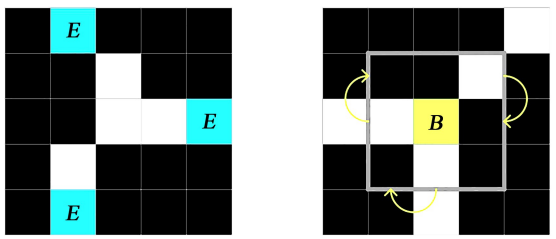
图 3　各类 T 形结构

对上下 T 形结构进一步精简，遍历细化后的图像，将八邻域点满足式(11)的像素点删除。

$$(P_2 \vee \widetilde{P}_3 \vee P_4)+(P_6 \vee \widetilde{P}_7 \vee P_8)+P_1+P_5=1 \tag{11}$$

由于裂缝形态的不规则性以及算法的局限性，经细化后得到的裂缝骨架仍然存在一些短小的毛刺，这不仅降低了骨架的准确性，也影响了对裂缝结构的分析，需要将其去除。

首先遍历细化后得到的骨架，找到所有的骨架端点和分支点，端点及分支点的示例如图 4 所示。其中，端点需满足：其八邻域内只有一个点为骨架点；仅利用八邻域内骨架点的数量[10]无法准确定位分支点，本文定义分支点需满足：其八邻域内沿顺时针方向由骨架点变为非骨架点的次数大于二。



(a) 端点　(b) 分支点

图 4　骨架关键点示例

选择一个端点作为起点，迭代地寻找相邻的骨架点，直到遇到另一个端点或者分支点时，将这些点构成的曲线作为该端点所在分支。统计该分支的像素点数量 N，如果 $N<50$，则判定该分支为毛刺，删除其所有像素点。重复以上步骤，直到所有端点都已被选择。

去除骨架毛刺后，通过骨架端点之间的特征对断裂的骨架进行连接。对于每个端点(x_0,y_0)，取与其连通的 n 个连续骨架点(x_i,y_i)构建点集 $P=\{(x_i,y_i)\}_{i=0}^{n}$，计算其协方差矩阵：

$$\boldsymbol{M}=\begin{bmatrix} d_{xx} & d_{xy} \\ d_{xy} & d_{yy} \end{bmatrix} \tag{12}$$

协方差矩阵 $\boldsymbol{M}$ 的元素由下列公式给出：

$$d_{xx}=\frac{\sum_{i=0}^{n}(x_i-\bar{x})^2}{n} \tag{13}$$

$$d_{xy}=\frac{\sum_{i=0}^{n}(x_i-\bar{x})\times(y_i-\bar{y})}{n} \tag{14}$$

$$d_{yy}=\frac{\sum_{i=0}^{n}(y_i-\bar{y})^2}{n} \tag{15}$$

式中，$\bar{x}$ 和 $\bar{y}$ 分别是点集 P 在 x 轴和 y 轴上的坐标均值。

得到协方差矩阵后，点集 P 构成的区域的主轴与 x 轴正半轴的夹角 α 可使用式(16)进行计算，α 的取值为[0,180)，单位为度。

$$\alpha=\frac{1}{2}\times\arctan\left(\frac{2d_{xy}}{d_{xx}-d_{yy}}\right) \tag{16}$$

由于一条直线可以对应两个相反的方向，根据端点和与之相邻的 n 个点坐标之间的关系，来确定端点的生长方向。

令 $d_x=\sum_{i=1}^{n}(x_0-x_i)$，$d_y=\sum_{i=1}^{n}(y_0-y_i)$，它们分别为端点沿 x 轴和 y 轴的变化方向。当 $d_y<0$，或者 $d_y=0$，$d_x<0$ 时，修正端点的生长方向为 $\theta=180+\alpha$；对于其他情况，$\theta=\alpha$。端点方向示意如图 5。

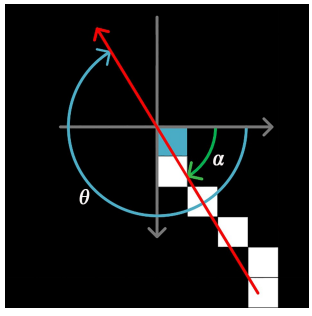


图 5　端点方向示意图

计算完所有骨架端点的生长方向角后，选择两个端点 E_1，E_2，以 E_1 为起点，E_2 为终点的连线向量的方向角为 $\beta_{1,2}=\arctan\left(\frac{y_2-y_1}{x_2-x_1}\right)$，与之相反的向量方向角为 $\beta_{2,1}=180+\beta_{1,2}$，它们之间的距离 $D=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$。

当两个端点之间的距离 $D<50$，且它们之间连线向量的方向角与两个端点的方向角之间的角度差满足：$|\beta_{1,2}-\theta_1|<45$ 且 $|\beta_{2,1}-\theta_2|<45$，则将两个端点用直线连接起来。

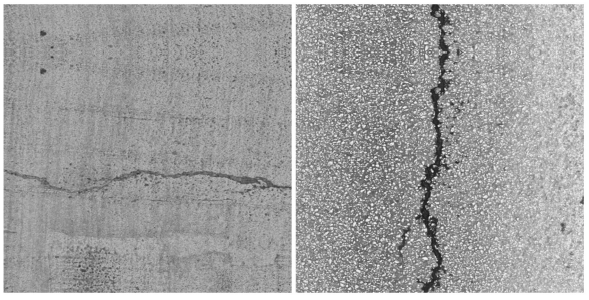
由于细化算法为迭代的侵蚀物体边缘，在端点处会出现向内缩短的问题，在细化过程中对裂缝末端进行标记来避免末端缩短的方法[11]需要预先假定裂缝形态，适用性较低。本文利用端点的生长方向来对骨架末端进行修复，以端点 $E_0(x_0,y_0)$ 为起点，沿着端点方向 θ 生长，直到生长点为非裂缝像素点，第 k 个生长点的坐标 (x_k,y_k) 为：

$$\begin{cases} x_k = x_0 + k \times \cos(\theta) \\ y_k = y_0 + k \times \sin(\theta) \end{cases} \tag{17}$$

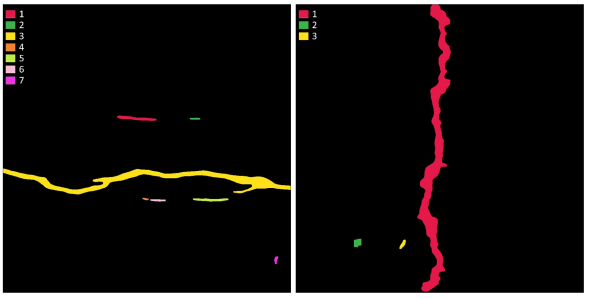
3 实验结果及分析

实验所用计算机配置如下：处理器为 Intel(R) Core(TM) i5-9500 CPU@3.00 GHz，内存为 16 GB，操作系统为 64 位 Windows10，开发环境为 PyCharm 2023.2，编程语言为 Python 3.7，用 OpenCV 和 NumPy 等库进行图像处理和数学计算。

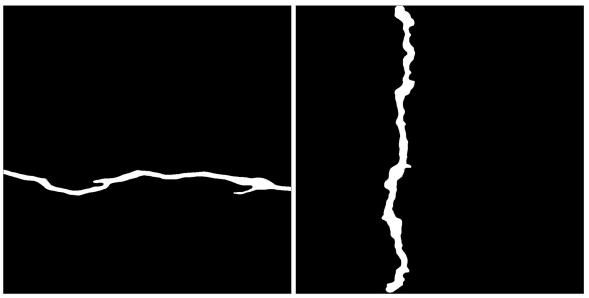
为了验证本文算法的有效性，以采集到的大小为 3 840×1 000 的实际沥青路面图像作为实验测试数据。首先通过神经网络对路面图像进行分割，对分割结果形态学处理后再进行连通域标记，使用不同颜色来代表不同连通域，如图 6 所示。



(a) 路面原始图像



(b) 分割结果连通域



(c) 保留的裂缝区域

图 6 连通域分析结果

表 1 为图 6(b)中不同连通域对应的特征，本文利用连通域的灰度、像素和形状特征来筛选出真正的裂缝目标，将满足以下任意条件的连通域认为是干扰区域并将其去除。

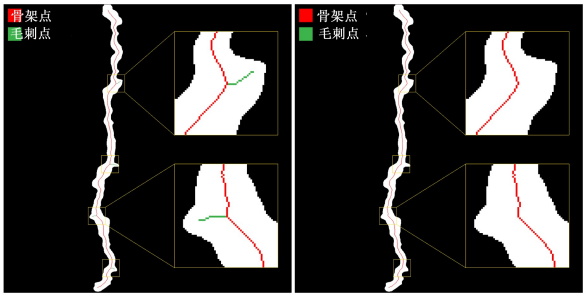
表 1 连通域特征

连通域	平均灰度	像素	圆形度	矩形度
1	137.27	1 107.50	0.08	0.74
2	127.73	175.00	0.18	0.83
3	116.88	19 543.00	0.02	0.22
4	147.10	101.00	0.29	0.72
5	136.51	842.00	0.07	0.69
6	128.64	283.00	0.13	0.78
7	125.07	208.00	0.37	0.67

条件包括：图像平均灰度值 G_1 与连通域平均灰度值 G_2 的差 $G_1-G_2<20$；像素 $S<300$；矩形度 $R>0.7$；圆形度 $C>0.7$。

去除干扰区域后的结果如图 6(c)所示，可以看出，分割结果中的干扰区域被有效剔除，只保留了裂缝区域。

使用改进的快速并行细化算法提取出的裂缝骨架如图 7(a)所示，经细化提取出的裂缝骨架存在一些短小的毛刺，通过本文的毛刺去除方法，成功地去除了这些毛刺，如图 7(b)所示。



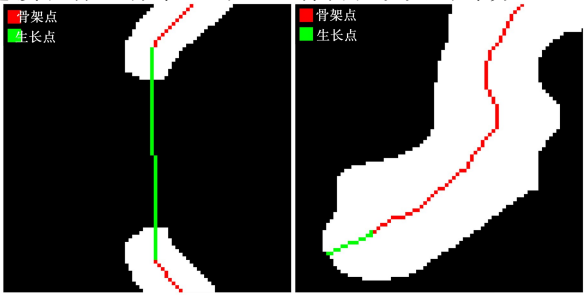
(a) 骨架提取 (b) 毛刺去除

图 7 骨架提取及毛刺去除结果

从图 7 中可以看出，本文的方法能够有效地提

取出裂缝的骨架，保留了裂缝的主要拓扑结构，同时能够去除骨架上的毛刺，使得骨架更加平滑和规整。

最后，本文使用端点连接及生长的方法来修复骨架上的断裂和缺失，利用端点的方向和距离来判断是否需要进行连接，同时沿着端点的方向进行生长，以补充骨架的细节，图 8 展示了骨架修复的结果。通过本文的端点连接及生长方法，能够有效地连接和补全骨架上的一些断裂和缺失的部分。



（a）端点连接　　（b）端点生长

图 8　骨架修复结果

为了定量评价本文提取出的骨架的效果，本文使用了以下两个指标：

精确率（P）：表示被算法正确标记为骨架的像素点数量 N_T 在所有被标记为骨架的像素点数量 N_P 中的比例，其数学表达式为：

$$P=\frac{N_T}{N_P} \tag{18}$$

精确率的值范围在 0 到 1 之间，越接近 1 表示算法在提取骨架方面的准确性越高。

召回率（R）：表示被算法正确标记为骨架的像素点数量 N_T 在所有实际骨架像素点数量 N_R 中的比例，其可表示为：

$$R=\frac{N_T}{N_R} \tag{19}$$

召回率的值范围同样在 0 到 1 之间，越接近 1 表示算法对实际骨架的捕捉能力越强。

表 2 展示了不同方法的指标比较结果，可以看出，本文方法相较于另外两种方法，精确率和准确率都更高，能够获得质量更好的骨架。

表 2　不同方法的比较

方法	P	R
Zhang-Suen 算法	0.872 3	0.918 9
中轴变换算法	0.821 3	0.926 1
本文方法	0.929 7	0.932 9

4　结论

本文提出了一种基于图像处理的路面裂缝特征提取方法。在使用深度神经网络对路面图像进行分割后，结合形态学处理和连通域分析，通过综合考虑裂缝的多个特征，包括灰度、面积和形状等方面的差异，有效地去除了干扰区域，只保留了真正代表裂缝的连通域，采用改进的快速并行细化算法提取出平滑、规整无毛刺的单一像素裂缝骨架，并通过端点方向和距离进行骨架的连接和生长，修复了骨架上的断裂和缺失，增强了骨架的连续性和完整性，有效地获得了裂缝的主要结构和形态特征。

参考文献：

[1] 林博煌，石立万，王悦婵，等. 沥青路面裂缝的数字图像测量与细观结构特征研究[J]. 公路，2022，67(7)：33-40.
Lin Bohuang，Shi Liwan，Wang Yuechan，et al. Research on measurement and meso-structural characteristics of asphalt pavement cracks based on digital image processing [J]. Highway，2022，67(7)：33-40.

[2] Kheradmandi N，Mehranfar V. A critical review and comparative study on image segmentation-based techniques for pavement crack detection[J]. Constr. Build. Mater.，2022，321：126162.

[3] 徐志刚，车艳丽，李金龙，等. 路面破损图像自动处理技术研究进展[J]. 交通运输工程学报，2019，19(1)：172-190.
Xu Zhigang，Che Yanli，Li Jinlong，et al. Research progress on automatic image processing technology for pavement distress [J]. J. of Traffic and Transportation Engineering，2019，19(1)：172-190.

[4] Vivekananthan V，Vignesh R，Vasanthaseelan S，et al. Concrete bridge crack detection by image processing technique by using the improved OTSU method [J]. Mater. Today Proc.，2023，74(4)：1002-1007.

[5] Pham M V，Ha Y S，Kim Y T. Automatic detection and measurement of ground crack propagation using deep learning networks and an image processing technique [J]. Measurement，2023，215：112832.

[6] 胡文魁，邓　晖，付志旭，等. 基于全卷积神经网络的桥梁裂缝分割和测量方法[J]. 工业建筑，2022，52(4)：192-201，218.
Hu Wenkui，Deng Hui，Fu Zhixu，et al. Bridge crack segmentation and measurement method based on full convolutional neural network [J]. Industrial Construction，2022，52(4)：192-201，218.

[7] Li Z，Yin C，Zhang X. Crack segmentation extraction and parameter calculation of asphalt pavement based on image processing[J]. Sensors，2023，23(22)：9161.

[8] Deng J，Lu Y，Lee V C S. A hybrid lightweight encoder-decoder network for automatic bridge crack assessment with real-world interference [J]. Measurement，2023，216：112892.

[9] Guo Z，Hall R W. Parallel thinning with two-subiteration

algorithms[J]. Commun. ACM, 1989, 32(3): 359-373.
[10] Bae H, An Y K. Computer vision-based statistical crack quantification for concrete structures[J]. Measurement, 2023, 211: 112632.
[11] Tang Y, Huang Z, Chen Z, et al. Novel visual crack width measurement based on backbone double-scale features for improved detection automation[J]. Eng. Struct., 2023, 274: 115158.

作者简介：

代少升(1974—)，男，河南省潢川县人，博士，教授，主要从事红外成像系统及 SOPC 嵌入式系统设计与开发；

余自安(1974—)，男，河南省光山县人，高级工程师，主要从事机械设计及理论研究。

(上接第 507 页)

[18] Yang Y, Gao X, Wang Y, et al. VAMYOLOX: An accurate and efficient object detection algorithm based on visual attention mechanism for UAV optical sensors[J]. IEEE Sens. J., 2022, 23(11): 11139-11155.
[19] Ge Z, Liu S, Wang F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. 2021: arXiv: 2107.08430. https://arxiv.org/abs/2107.08430.
[20] Wang G, Chen Y, An P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios[J]. Sensors, 2023, 23(16): 7190.
[21] Pandey V, Anand K, Kalra A, et al. Enhancing object detection in aerial images[J]. Math. Biosci. Eng., 2022, 19(8): 7920-7932.
[22] Tong Z, Chen Y, Xu Z, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[EB/OL]. 2023: arXiv: 2301.10051. https://arxiv.org/abs/2301.10051.
[23] Li X, Wang W, Wu L, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection[J]. Adv. in Neural Inf. Process. Syst., 2020, 33: 21002-21012.
[24] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]// Proc. of the AAAI Conf. on Artificial Intelligence, 2020: 12993-13000.
[25] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation[C]// Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2018: 8759-8768.
[26] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]// Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2020: 10781-10790.
[27] Ding X, Zhang X, Han J, et al. Diverse branch block: Building a convolution as an inception-like unit[C]// Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2021: 10886-10895.
[28] Yang J, Li C, Dai X, et al. Focal modulation networks[J]. Adv. in Neural Inf. Process. Syst., 2022, 35: 4203-4217.

作者简介：

贾　亮(1971—)，男，辽宁省大石桥市人，副教授，硕士生导师，主要从事信息获取与处理方面的研究；

林铭文(2000—)，男，福建省福州市人，硕士研究生，主要从事图像处理与目标检测方向的研究。