

# SU-Net：一种基于注意力机制的道路裂缝检测模型

郭晓彤<sup>1</sup> 王跃宝<sup>1</sup> 鲁王泽<sup>1</sup>  
GUO Xiaotong WANG Yuebao LU Wangze

## 摘要

道路裂缝的早期检测对于解决安全问题和减少维修成本非常重要。然而，由于裂缝的复杂性和环境因素的干扰，传统图像处理方法在裂缝检测任务中的表现面临诸多挑战。为解决上述问题，文章提出了一种基于 SE（squeeze-and-excitation）注意力机制和膨胀卷积的道路裂缝检测方法。通过改进 U-Net 模型，构建了一种新的编码解码结构模型 SU-Net。该模型在编码阶段引入 SE 注意力机制，通过自适应调整通道权重，增强了对裂缝细节的捕捉能力。此外，为在不增加计算复杂度和参数规模的前提下扩展模型的感受野，引入了膨胀卷积，通过对膨胀率的细致调整，使模型性能达到最优。实验结果表明，与传统的 U-Net 模型相比，SU-Net 模型在道路裂缝检测任务中取得了显著的性能提升。通过定量评估，验证了改进网络在提高裂缝检测精度和鲁棒性方面的有效性,为今后裂缝检测领域的研究和应用提供了重要参考。

## 关键词

道路裂缝；注意力机制；深度学习；改进 U-Net 模型；图像分割

doi: 10.3969/j.issn.1672-9528.2025.07.011

## 0 引言

随着交通领域的大力发展，我国境内建成了巨量道路。道路作为重要基础设施，并非一次性消费品，其使用寿命普遍较长，短则数十年，长则可达上百年。在长期使用过程中，各类道路路面不可避免地会出现不同程度的破损，其中裂缝问题尤为常见。裂缝不仅会影响路面的美观，还会对行车安全构成威胁，甚至会导致严重事故。所以裂缝的早期检测能有效防止道路进一步损坏，降低维修成本，并减少交通事故的发生<sup>[1]</sup>。

近年来，计算机视觉技术被广泛应用于自动裂缝检测任务，其中基于图像处理和机器学习的方法取得了显著进展<sup>[2]</sup>。在道路裂缝检测领域，已有大量研究利用传统图像处理技术和机器学习方法进行裂缝识别。例如，早期的裂缝检测方法多采用经典的图像处理技术，如阈值分割、边缘检测和形态学操作<sup>[3]</sup>。这些方法虽然在理想环境下能够有效检测裂缝，但是对于光照变化、噪声干扰、复杂背景纹理等复杂场景效果较差。为解决传统图像处理技术的问题，研究者们提出了基于机器学习的方法，例如支持向量机（SVM）和随机森林等<sup>[4]</sup>。然而，这些方法依赖人工设计特征，难以应对裂缝的多样性。

随着卷积神经网络的兴起，深度学习方法逐渐成为主流，为裂缝检测提供了更强的特征学习能力。基于 CNN

的端到端模型可以自动提取裂缝的高层次语义信息，提高检测精度和鲁棒性<sup>[5]</sup>。其中，U-Net 因为其编码 - 解码结构的特点，已被广泛地应用在医学图像分割，并被应用于裂缝检测任务<sup>[6]</sup>。但是传统 U-Net 在裂缝检测中的表现仍存在局限性，如对裂缝的细节特征提取不足、对背景噪声较为敏感等<sup>[7]</sup>。一是裂缝的形态多样性和环境干扰使得模型在复杂场景下的检测精度有限；二是传统 U-Net 模型在编码阶段对全局信息的捕捉能力较强，但在细节信息的提取上表现不足，导致对小尺度裂缝的检测效果不理想；三是现有方法在模型复杂性和计算效率之间的平衡也有待进一步优化。所以如何设计一种既能有效捕捉裂缝细节信息，又能在复杂环境下保持高鲁棒性的检测模型，成为当前研究的重点。

针对上述问题，本文提出了一种基于 SE（squeeze-and-excitation）注意力机制和膨胀卷积的道路裂缝检测方法，通过改进 U-Net 模型，构建了新型编码 - 解码结构模型 SU-Net。注意力机制在编码阶段通过全局池化操作获取特征图的全局信息，再利用全连接层学习通道间的依赖关系，从而自适应地对每个通道的特征进行加权。与传统方法不同，这种方式能够更加精准地聚焦于裂缝的细节特征，增强模型对细微裂缝和复杂背景下裂缝的识别能力。同时，模型还引入膨胀卷积，从而使卷积核能够覆盖到更广阔的图像区域，捕捉到更丰富的上下文信息，进而提取到更为细致和全面的裂缝特征。

1. 防灾科技学院 河北廊坊 065201

## 1 方法

### 1.1 模型结构

本文提出的道路裂缝检测模型 SU-Net 基于经典的 U-Net 架构，并融入注意力机制和膨胀卷积，改进模型的结构示意图如图 1 所示。模型采用编码-解码结构，使用卷积和池化操作提取特征，借助上采样和跳跃连接实现图像分割，以增强对裂缝细节信息的捕捉能力。

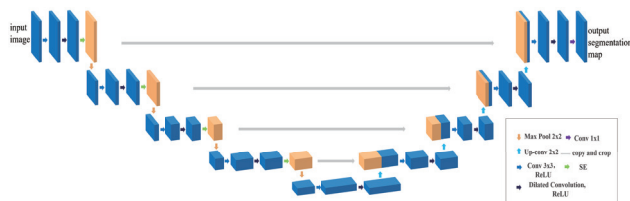


图 1 SU-Net 结构示意图

与 U-Net 模型不同的是，编码阶段中每层第一个卷积的膨胀率设置为 1，卷积后的图像不会发生改变，每层第二个卷积的膨胀率设置为 2，卷积后的图像尺寸不改变，获得了更大的感受野，并结合 ReLU 激活函数。两次卷积后的特征图添加 SE 注意力机制，使网络关注重要特征。然后通过最大池化把图像尺寸减小为原始的一半，同时要通道数加倍。编码器重复上述操作，逐层卷积和下采样，直到特征图通道数到 1 024，编码阶段结束。

解码器通过  $2 \times 2$  转置卷积上采样，恢复特征图分辨率，同时，利用跳跃连接将编码器中的低层次特征与解码器的高层次特征融合，在恢复图像分辨率的同时保留细节信息，提升裂缝检测精度。最后，将 64 通道的特征图通过  $1 \times 1$  卷积，将通道数调整为所需的分类个数，即输出二值分割图，阈值处理后输出裂缝检测结果。

### 1.2 注意力机制

SE 注意力机制是一种通道注意力机制<sup>[8]</sup>，旨在通过自适应地调整通道权重，使网络能够更好地关注重要特征，忽略非重要的信息，其工作流程主要包括 Squeeze（压缩）、Excitation（激励）和 Scale（重新加权）3 个步骤，如图 2 所示。

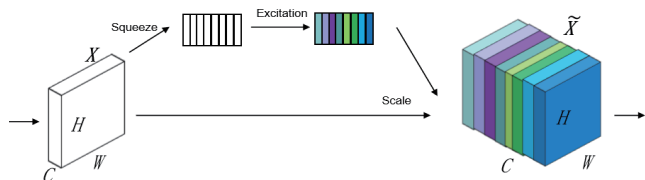


图 2 SE 模块

本文选择了 CBAM、ECA 和 SE 注意力进行实验，得出了 SE 注意力机制在裂缝检测中的效果最好。在每个卷积块后引入 SE 注意力机制，通过全局池化操作，将特征图在空

间维度压缩，计算每个通道的均值，得到  $1 \times 1 \times C$  的向量（ $C$  为通道数），获取全局信息，并通过两个全连接层，第一个全连接层将向量维度降低，激活函数采用 ReLU；第二个全连接层再将维度升回原通道数，激活函数为 Sigmoid，学习通道间的依赖关系，生成通道注意力权重。该权重与特征图通道相乘，增强裂缝特征并抑制背景噪声，提高模型鲁棒性。

引入 SE 注意力机制主要基于以下原因：

一是增强特征选择能力，通过显式建模通道依赖关系，自适应调整通道权重，突出裂缝特征，抑制背景干扰，提升模型在复杂环境下的鲁棒性。

二是捕捉全局上下文信息，弥补传统卷积操作只能捕捉局部感受野信息的不足，帮助模型更好地理解裂缝图像全局结构。

三是提升模型对细节的敏感性，针对道路裂缝细长、不规则且与背景对比度低的特点，动态调整通道权重，增强对裂缝细节的捕捉能力。

### 1.3 膨胀卷积

膨胀卷积（dilated convolution），又称空洞卷积或扩张卷积，是一种特殊的卷积操作。该操作通过在标准卷积核元素间引入固定间隔的方式，在不增加计算复杂度和参数规模的前提下，能够扩大模型的感受野。这种独特的设计使卷积核能够覆盖更广的图像区域，从而捕获更丰富的上下文信息，实现更精细、更全面的裂缝特征提取。本文采用的数据集具有尺寸较小的特点，所以在膨胀率的设置上进行了优化。为了确定最佳的膨胀率组合，令膨胀率为 1 或 2 或 3，进行了实验。实验结果表明，当网络结构中每层第一个卷积 dilation 为 1，第二个卷积 dilation 为 2 时，模型展现出最优的表现，所以模型结构选用此组合进行后续的卷积操作。

## 2 实验

### 2.1 数据集

在本实验中，为了构建一个丰富多样的数据集，整合了 CFD、Crack500 和 SUT-Crack 三个公开数据集，最终形成了 Crack 数据集，共包括 2 620 张图像。这些图像来源广泛，在分辨率、裂缝场景以及背景材质等方面存在显著差异，为模型训练提供了多元且丰富的数据样本，极大地提升了模型在复杂现实场景下的适应性和泛化能力。

CFD（crack forest dataset），包含 118 张高分辨率图像，每张图像分辨率为  $480 \text{ px} \times 320 \text{ px}$ 。该数据集涵盖了各类复杂的裂缝场景，背景涉及道路、墙壁等多种材质，并且提供了精确到像素级别的裂缝标注。Crack500 数据集包含 500 张高分辨率图像，主要用于混凝土、沥青等材料表面的裂缝检测<sup>[9]</sup>。该数据集的显著优势在于其裂缝类型极为丰富，包括

细小裂缝、宽大裂缝、直线裂缝和曲线裂缝等，同时背景复杂度各不相同，从相对单一的背景到包含各类纹理、光照变化及其他干扰因素的复杂背景都有涉及。SUT-Crack 数据集专注于建筑和工程结构表面的裂缝检测，包含数百张高质量图像，图像背景涵盖混凝土、沥青等多种材质表面<sup>[10]</sup>。其丰富的裂缝场景和优质的图像质量，使其非常适用于精确检测和分割任务。

在训练前，使用 OpenCV 清洗数据，剔除损坏图像，修正错误标注。对 RGB 通道进行颜色归一化。为了增强模型泛化能力，实施了几何变换（随机旋转、翻转）和颜色变换（亮度和对比度调节，添加高斯噪声）。

## 2.2 实验设计

在本研究中，设计了一系列实验来评估改进的 SU-Net 模型在裂缝图像分割任务中的性能。以下是实验设计的详细描述。

### （1）实验设置

为深入探索模型性能，开展了对比实验。为保障实验的公平性与可比性，均在相同的 Crack 数据集上进行训练与测试，且使用一致的超参数设置，包括初始学习率、训练轮数、批量大小等，以确保唯一变量为模型结构差异。将数据集按照 8:1:1 分为训练集、验证集和测试集，训练集用于模型的训练与参数优化，验证集用于调整超参数和防止过拟合，测试集用于评估模型的泛化能力和最终性能。

本实验的操作系统为 Ubuntu20.04，64 位，CPU 为 Intel(R) Xeon(R) Gold 6230R CPU @ 2.10 GHz，内存为 512 GB，深度学习框架使用的是 PyTorch1.8，cuda 版本为 11.4，python 环境为 3.7，GPU 为 RTX3090。本实验属于二分类问题。实验使用带有权重衰减率（设置为 0.000 1）和动量的 Adam 来优化模型，动量设置为 0.9，初始学习率设置为 0.000 5。模型的训练轮数设置为 100，批次大小设置为 4，在每轮后进行一次验证，并使用验证集上的最佳权重模型作为在测试集上进行性能评估的最终模型。

### （2）评估指标

为全面、客观地评估模型在裂缝检测任务中的性能，采用以下几种常用的评估指标。

精确率是指在所有被模型预测为正的样本中，实际为正的样本所占的概率。衡量了模型预测为正的结果中有多少是真正正确的。精确率越高，说明模型在预测正样本时的准确性越高。其计算公式为：

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

查全率是指在实际为正的样本中，被模型预测为正样本的概率。反映了模型能够正确识别出多少真正的正样本。

查全率越高，说明模型对正样本的捕捉能力越强。其计算公式为：

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

均像素精度是像素精度（PA）的一种简单提升。计算方式是先计算每个类内被正确分类像素数的比例，后求所有类的平均值。该指标从像素层面衡量了模型分类的准确性，能够反映模型对不同类别像素的分类能力。其计算公式为：

$$\text{MPA} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (3)$$

均交并比是语义分割任务中的标准度量指标。通过计算两个集合（真实值和预测值）的交集和并集之比来评估模型的性能。在语义分割问题中，这个比例可以变形为真正例数比上真正例、假负例、假正例（并集）之和。该指标会在每个类上分别计算交并比（IoU），然后求平均值。均交并比越高，说明模型预测结果与真实标签的重叠程度越高，分割效果越好。其计算公式为：

$$\text{MIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (4)$$

通过以上实验设计，能够全面评估 SU-Net 模型在道路裂缝图像分割任务中的性能，为道路裂缝检测提供更为精准且可靠的技术支撑。

## 3 实验结果与分析

为了全面评估 SU-Net 模型在裂缝图像分割任务中的性能，设置了消融实验和对比实验，并结合可视化结果直观展现模型的分割效果。

### 3.1 消融实验

为了验证 SE 注意力机制和膨胀卷积对原始 U-Net 模型带来的性能提升，在统一数据集和损失函数的基础上，设置了消融实验，实验结果如表 1 所示。由表中数据可以看出，U-Net 编码端加入 SE 注意力机制后，精确率、MIoU、MPA、查全率均有所提升，分别提升了 0.74%、3.32%、1.02%、0.91%。在此基础上，引入了膨胀卷积，分别提升了 1.23%、3.57%、3.74%、4.43%。说明了这些改进是有效的，提升了 U-Net 对裂缝的检测性能。

表 1 消融实验结果

| 单位：%           |           |       |       |        |
|----------------|-----------|-------|-------|--------|
| 模型             | Precision | MIoU  | MPA   | Recall |
| U-Net          | 88.36     | 75.13 | 84.10 | 83.07  |
| UNet+SE        | 89.10     | 78.45 | 85.12 | 83.98  |
| U-Net+SE+ 膨胀卷积 | 90.33     | 82.02 | 88.86 | 88.41  |

3.2 对比实验

为进一步验证 SU-Net 有效性，在同一数据集 Crack 和评价指标下设计对比试验，将其与 VGG16-UNet 和 Deeplabv3+ 进行对比。对比实验结果如表 2 所示。

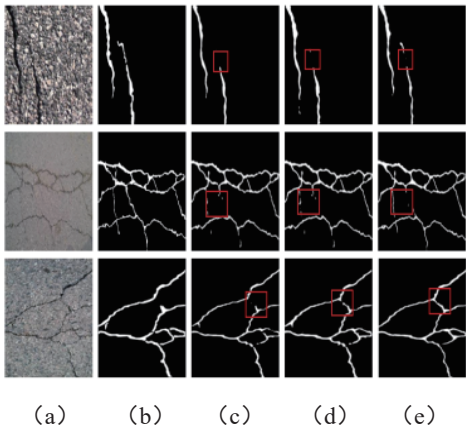
表 2 对比实验结果

| 模型         | 单位：%      |       |       |        |
|------------|-----------|-------|-------|--------|
|            | Precision | MIoU  | MPA   | Recall |
| VGG16-UNet | 89.32     | 80.32 | 85.23 | 84.95  |
| Deeplabv3+ | 89.54     | 81.84 | 88.73 | 88.20  |
| SU-Net     | 90.33     | 82.02 | 88.86 | 88.42  |

由表 2 中数据中可以看出，SU-Net 模型在 Precision、MIoU、MPA 和 Recall 四个分割评价指标上均优于其他两个模型。具体而言，在 Precision 指标上，SU-Net 比 VGG16-UNet 高 1.01%、比 Deeplabv3+ 高 0.79%，说明 SU-Net 在正确预测裂缝正类像素的能力强。在 MIoU 指标上，SU-Net 同样领先，比其他两个模型分别高 1.7%、0.18%，表明 SU-Net 在整体分割精度上有较好的表现，能够更准确地提取目标区域。在 MPA 指标上，SU-Net 分别优于其他两个模型 3.63%、0.13%，显示出其在像素级准确性方面的优势。在 Recall 值方面，SU-Net 略优于 Deeplabv3+0.22%，比 VGG16-UNet 高 3.47%，表明其在召回率上表现出色，能够更好地捕捉到真实裂缝目标。

3.3 可视化分析

在道路裂缝图像分割任务中，可视化分析能够直观展现不同模型的分割效果。本部分对 3 组图片进行分析，在极窄裂缝、多分支裂缝和复杂裂缝的场景下，对比以上 3 种模型的表现。分割预测结果如图 3 所示。



注：（a）原图；（b）标签图；（c）VGG16-UNet 分割预测结果图；（d）Deeplabv3+ 分割预测结果图；（e）SU-Net 分割预测结果图

图 3 分割预测结果

在裂缝形态还原上，VGG16-UNet 虽能捕捉主要裂缝走向，但在细微分支和复杂连接结构的还原上存在不足，部分细分支遗漏，连接点位置偏差，裂缝形态呈现松散感。

Deeplabv3+ 有所进步，主要分支呈现较完整，但在细节处仍有欠缺，细微分支和宽度变化的细腻度不够。SU-Net 在这方面表现卓越，无论是主要还是细小分支，均能完整精准呈现，连接点精准，宽度变化细腻，高度还原裂缝复杂结构。在裂缝边缘精细程度上，VGG16-UNet 的裂缝边缘粗糙，锯齿状明显，尤其是分支边缘，像素分类不够精细。Deeplabv3+ 锯齿现象减少，但边缘仍有细微不平整和模糊。SU-Net 的边缘平滑连续，无锯齿或不平整问题，主裂缝与分支裂缝边缘处理精细，体现出强大的边缘像素精准分类能力。在背景噪声控制上，VGG16-UNet 受背景干扰严重，路面纹理和颜色差异异常被误判为裂缝，分割结果噪声多。Deeplabv3+ 噪声控制能力提升，误分割噪声减少，但在复杂背景区域仍有误判。SU-Net 背景噪声控制效果最佳，背景干净，能有效区分裂缝与背景，抑制背景干扰，获取纯净准确的分割结果。

综上所述，通过对 3 组图片的可视化分析，SU-Net 在裂缝形态还原、边缘精细程度和背景噪声控制方面均优于 VGG16-UNet 和 Deeplabv3+，展现出在裂缝图像分割中的卓越性能，为实际裂缝检测应用提供了更可靠的技术支持。

4 结语

本研究通过引入注意力机制和膨胀卷积来优化网络结构，提出了 SU-Net 模型用于道路裂缝图像分割。实验结果表明，SU-Net 在精确率、查全率等指标上优于其他两个模型，在细微裂缝和复杂背景处理上表现出色。通过可视化对比图，进一步验证了模型的优越性。模型在背景干扰繁杂、裂缝特征微弱时，分割精度和检测能力会受限，但在常见裂缝检测场景中，SU-Net 能精准分割裂缝，为基础设施安全评估提供有力支撑，应用前景广阔。未来，研究将聚焦于进一步优化模型结构，提高模型在极端复杂场景下的性能，推动裂缝检测技术的不断发展。

参考文献：

[1] MANDAL V, UONG L, ADU-GYAMFI Y. Automated road crack detection using deep convolutional neural networks[C]//2018 IEEE International Conference on Big Data (Big Data). Piscataway:IEEE, 2018:5212-5215.

[2] MUNAWAR H S, HAMMAD A W A, HADDAD A, et al. Image-based crack detection methods: a review[J]. Infrastructures, 2021, 6(8): 115.

- [3] CANNY J. A computational approach to edge detection[J]. IEEE transactions on pattern analysis and machine intelligence, 1986,8(6): 679-698.
- [4] FUJITA Y, SHIMADA K, ICHIHARA M, et al. A method based on machine learning using hand-crafted features for crack detection from asphalt pavement surface images[C]// Proceedings of SPIE - The International Society for Optical Engineering. USA: SPIE, 2017: 117-124.
- [5] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[EB/OL].(2015-05-18)[2025-01-26].<https://doi.org/10.48550/arXiv.1505.04597>.
- [6] HAMISHEBAHAR Y, GUAN H, STEPHEN S, et al. A comprehensive review of deep learning-based crack detection approaches[J]. Applied sciences, 2022, 12(3): 1374.
- [7] GUO F, LIU J, LÜ C S, et al. A novel transformer-based network with attention mechanism for automatic pavement crack detection[J]. Construction and building materials, 2023, 391: 131852.
- [8] HU J, SHEN L, SUN G.Squeeze-and-excitation networks[C/OL]//2018 IEEE/CVF Conference on Computer Vision and

Pattern Recognition.Piscataway:IEEE,2018[2025-02-19].  
<https://ieeexplore.ieee.org/document/8578843>.DOI:10.1109/CVPR.2018.00745.

- [9] LIU Z H. Road crack detection system using image segmentation algorithm[C]//PCCNT '23: Proceedings of the 2023 International Conference on Power, Communication, Computing and Networking Technologies.NewYork: ACM, 2023: 1-6.
- [10]SABOURI M, SEPIDBAR A. SUT-Crack: a comprehensive dataset for pavement crack detection across all methods[J]. Data in brief, 2023, 51:109642.

#### 【作者简介】

郭晓彤（2001—），女，河北邯郸人，硕士研究生，研究方向：灾害信息处理技术。

王跃宝（1999—），女，河北石家庄人，硕士研究生，研究方向：灾害信息处理技术。

鲁王泽（2001—），男，安徽铜陵人，硕士研究生，研究方向：灾害信息处理技术。

（收稿日期：2025-03-13 修回日期：2025-07-07）

（上接第 47 页）

#### 4 结论

本文提出了一种基于双向 Dijkstra 算法的无人机路径规划，通过对路径进行正向与反向的同步搜索，当两个方向的搜索相遇时，合并获得最短路径。后对提出的双向 Dijkstra 算法进行时间、空间的复杂度分析，以及对算法的实际运行时间进行了分析对比。通过分析可知，双向 Dijkstra 算法在实际运行中是高效的，特别是在稀疏图 ( $E \ll V^2$ ) 和大规模图上。最后，通过一个算例验证了该算法的可行性和高效性。

#### 参考文献:

- [1] 王其, 王磊, 倪世松, 等. 基于改进 A\* 算法的应急救援无人机路径规划 [J]. 计算机仿真, 2024, 41(6):84-88.
- [2] 唐嘉宁, 彭志祥, 李孟霜, 等. 基于改进 A\* 算法的无人机路径规划研究 [J]. 电子测量技术, 2023, 46(8):99-104.
- [3] 高九州, 张焯. 基于改进 A\* 算法的无人机三维空间避障路径规划 [J]. 计算机测量与控制, 2023, 31(12):203-209.
- [4] 罗超. 基于改进 A\* 算法的无人机路径规划算法及应用 [D]. 长沙: 中南大学, 2023.

- [5] 程谦, 高嵩, 曹凯, 等. 基于 PRM 优化算法的移动机器人路径规划 [J]. 计算机应用与软件, 2020, 37(12):254-259.
- [6] 周金亮, 胡涛. 基于改进遗传算法的无人机路径规划研究 [J]. 机电技术, 2024(2):13-16.
- [7] 黄书召, 田军委, 乔路, 等. 基于改进遗传算法的无人机路径规划 [J]. 计算机应用, 2021, 41(2):390-397.
- [8] 杨晨, 肖博文, 刘郡怡, 等. 基于改进遗传算法的无人机路径规划方法:202311512310[P].2024-12-17.
- [9] 郭一凡. 基于深度强化学习的四旋翼无人机路径规划研究 [D]. 西安: 西安工业大学, 2024.
- [10] 骆文冠, 于小兵. 基于强化学习布谷鸟搜索算法的应急无人机路径规划 [J]. 灾害学, 2023, 38(2):206-212.
- [11] 王兴旺, 张清杨, 姜守勇, 等. 基于改进鲸鱼优化算法的动态无人机路径规划 [J]. 计算机应用, 2025, 45(3):928-936.

#### 【作者简介】

卢婷（1993—），女，河南周口人，硕士，助教，研究方向：最优化理论算法及应用，email: luting19@yeah.net。

（收稿日期：2025-03-06 修回日期：2025-07-02）