

doi: 10.6046/zrzyg.2023117

引用格式: 李世琦, 姚国清. 基于 CNN 与 SETR 的特征融合滑坡体检测[J]. 自然资源遥感, 2024, 36(4): 158-164. (Li S Q, Yao G Q. A landslide detection method using CNN- and SETR-based feature fusion[J]. Remote Sensing for Natural Resources, 2024, 36(4): 158-164.)

基于 CNN 与 SETR 的特征融合滑坡体检测

李世琦, 姚国清

(中国地质大学(北京)信息工程学院, 北京 100083)

摘要: 准确、及时地检测出滑坡体对减少山体滑坡自然灾害对人类生命和财产造成的威胁与损失具有重要意义。论文提出了一种基于卷积神经网络(convolutional neural network, CNN)与 Set Transformer(SETR)的特征融合滑坡体检测方法。基于 CNN 的网络模型选择了全卷积网络(fully convolutional network, FCN)、U-Net 和 Deeplabv3+, 基于 Transformer 的模型选择了 SETR。首先对 CNN 网络模型在滑坡检测中的效果进行评价, 然后在 CNN 网络模型的编码器部分引入 SETR, 并将 SETR 的输出融合到 CNN 的解码器结构中作为模型的整体输出。基于 LandSlide4Sense 数据集的实验结果表明, 典型 CNN 融合 SETR 后有效改善了模型的检测效果, FCN, U-Net, Deeplabv3+ 模型在融合 SETR 后 F1 分数分别从 0.672 6, 0.727 3, 0.687 3 提高到 0.686 9, 0.743 0, 0.705 5。因为滑坡与地形密切相关, 以效果最好的 U-Net 模型为基准, 在模型输入中引入数字高程模型之后 F1 分数从 0.732 5 提高到 0.750 3。

关键词: CNN; SETR; 滑坡体检测; 语义分割

中图法分类号: TP 79 **文献标志码:** A

文章编号: 2097-034X(2024)04-0158-07

0 引言

山体滑坡是地球表面自然过程的一部分。近年来, 受全球极端气候、地质活动、人类生产活动等影响, 滑坡地质灾害频频发生, 对人类的生命财产安全造成极大危害。随着硬件设备性能的提升, 深度学习技术得到快速发展, 广泛应用了深度学习的计算机视觉领域为滑坡体检测提供了新的研究方向。考虑到滑坡地质灾害危害性大、突发性强等特点, 模型对于滑坡体的检测应该满足漏检率低、定位准确、边界判定明确等要求, 从而最大限度地减小滑坡对人类带来的影响。

目前用于地学领域的算法和模型大多来自通用图像处理和计算机视觉领域, 融合地学特性尤其是滑坡等地质灾害的专用人工智能算法或专用深度学习框架仍有待研究, 地质灾害及其特征的自动化、智能化精准识别及提取方法亟待提高。传统的滑坡体检测方法以人为调查为主, 需耗费大量的人力与时间, 效率低且具有危险性^[1]。基于变化检测的滑坡检测方法适用于发生滑坡区域前后变化较为突出的场景, 方法原理相对简单, 但适合变化检测的数据难以获取^[2]。面向对象方法是基于传统的图像处理

技术, 算法相对简单, 但需手动提取滑坡体特征^[3]。

随着计算机技术的快速发展, 许多科研人员将机器学习方法应用于滑坡体检测问题上, 如: 张越等^[4]、余淙蔚等^[5]、许高程等^[6]分别采用决策树方法、逻辑回归方法和支持向量机方法应用于滑坡体检测, 但机器学习方法模型调参对模型评价结果影响较大, 对复杂函数表达能力有限, 泛化能力一般^[7]。随着深度学习技术的崛起, 许多科研人员考虑将滑坡体检测与深度学习技术融合^[8-10]。同时滑坡通常发生在陡峭的斜坡或者地形起伏较大的区域, 数字高程模型(digital elevation model, DEM)可以提供地表的高程信息, 通过对 DEM 进行分析, 可以确定地形的起伏和变化。如杨昭颖等^[11]将滑坡体检测作为目标检测任务进行研究, 将遥感影像与 DEM 特征信息融合, 验证了 DEM 数据在基于目标检测的滑坡检测方法中的重要性, 但只能将滑坡范围限定在一定区域, 无法区分滑坡体的边界, 当发生滑坡时无法第一时间紧急救援。实现像素级分类的语义分割技术可以有效改善无法定位目标边界的问题, 并且已被验证其可靠性^[12-14]。白石等^[15]以 U-Net 模型为基准, 引入密集上采样和非对称卷积来自动提取滑坡以避免空间信息损失, 并将 DEM 与高分辨率遥感影像融合进行特征提取, 进一步提高了

收稿日期: 2023-04-26; 修订日期: 2024-05-15

第一作者: 李世琦(1998-), 男, 硕士研究生, 主要从事计算机图像智能分析研究。Email: 895229712@qq.com。

通信作者: 姚国清(1964-), 男, 硕士, 教授, 主要从事遥感图像处理与信息提取研究。Email: gqyao@cugb.edu.cn。

模型对滑坡检测的效果。Transformer 模型最初是为自然语言处理任务设计的,但其自注意力机制和并行化计算的优势使其在计算机视觉领域也得到了广泛应用。2020 年 Vision Transformer (ViT)^[16] 将 Transformer 架构首次应用到计算机视觉领域。Set Transformer (SETR) 模型是一种基于 Transformer 架构的端到端的图像分割模型,利用全局自注意力机制和局部自注意力机制来处理图像语义信息。随后一些学者采用了 Transformer 结构模型分别用于目标检测和语义分割^[17-19],注意力机制在计算机视觉领域得到广泛应用。张思远等^[20]将注意力机制应用到滑坡检测,但 Transformer 模型对滑坡边界的判定比较模糊,且需大量数据作为实验支撑。

本文研究工作从以下 2 个方面展开:第一,改进网络模型,提高滑坡检测精度,所以本文将基于卷积神经网络 (convloutional neural network, CNN) 的全卷积网络 (fully convolutional network, FCN)、U-Net 和 Deeplabv3+ 网络模型分别与 SETR 模型进行融合,构建一个新的滑坡体检测模型,融合特征信息,从多尺度提取特征,降低漏检率,明确滑坡体边界,以期准确、及时地检测滑坡体;第二,引入 DEM 相

关数据作为滑坡体分割的辅助信息,进一步提高分割精度。

1 数据源与数据预处理

1.1 数据源

研究所需数据源来自于 Landslide4Sense 数据集,通过 Landslide4sense 官网下载 (<https://www.iairai.ac.at/landslide4sense/challenge/>),数据由 2 个部分组成,分别是 Sentinel-2 的 12 波段遥感影像 (除 water vapour (水蒸气) 波段),同时还包括 DEM 和坡度 2 种数据,组成一个 14 通道的数据集,并将所有波段都调整为每像素约 10 m 的分辨率,每张图像的大小为 128 像素×128 像素,并按像素进行标记。

1.2 数据预处理

官方一共提供了 3 799 张滑坡遥感影像。按 6:3:1 的比例划分为训练集、验证集和测试集。同时为了减少样本数量不平衡对模型性能带来的影像,采用数据增强的方法,增加正样本的权重,对遥感图像进行随机裁剪、翻转的操作,丰富数据多样性,从而提高模型的泛化能力。数据示例见图 1。

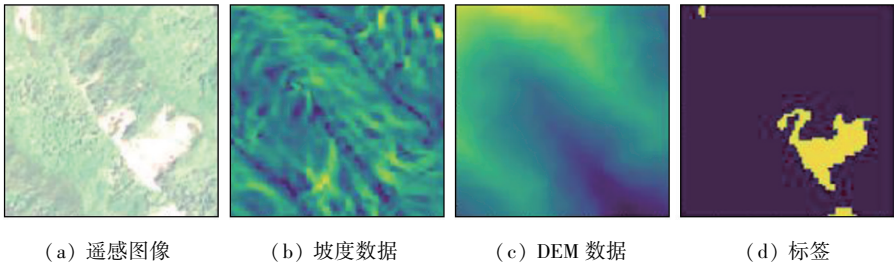


图 1 实验数据示意图

Fig.1 Schematic diagram of data

2 研究方法

为了精确且快速地提取滑坡体,本文构建了 CNN 与 SETR 融合的新模型。同时滑坡的发生通常

与地形因素密切相关,本文对设计的模型精度评价后,基于滑坡体检测性能更优的模型,引入 DEM 数据进一步提升模型检测效果。最终通过准确率、召回率和 F1 分数作为评价指标,并最终对模型输出结果进行评估。实验流程图如图 2 所示。

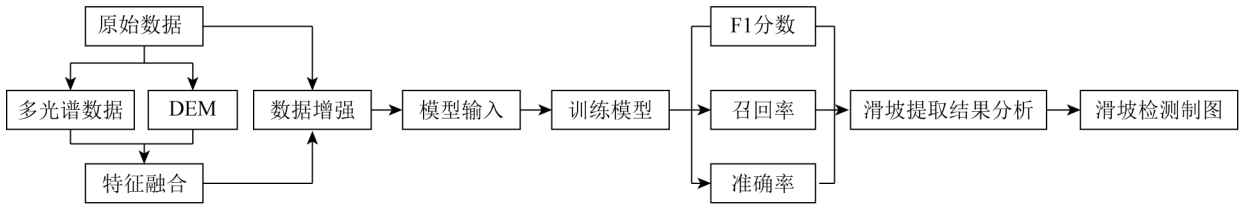


图 2 实验流程图

Fig.2 Experimental flow chart

2.1 经典语义分割方法

近几年,计算机视觉领域中语义分割技术已经

得到成熟运用,如 FCN, U-Net 和 Deeplabv3+ 等。FCN 是一种全卷积神经网络,由编码器和解码

器组成,采用逐点相加的方式,将编码器中不同层的特征图与解码器中相应层的特征图进行连接,用于端到端的像素级语义分割任务。

U-Net 模型具有“U”型结构和跳跃连接,用于高效地处理图像分割任务,它同样采用了编码器和解码器结构,但跳跃连接的方式是将特征在通道维度进行拼接,从而形成更加丰富的特征信息,并保证了各个尺度上特征的权重。

在 Deeplabv3+模型中利用空洞卷积和多尺度特征融合技术,通过将多个不同空洞率的空洞卷积的特征图融合在一起,从而在不同尺度上进行特征提取,来实现精细的图像语义分割。

FCN,U-Net 和 Deeplabv3+都是基于 CNN 结构的语义分割模型,可以有效地捕捉图像中的局部特征,对滑坡边界的判定更加精确。但由于受到卷积核感受野的范围较小,CNN 网络无法在全局信息上建立特征依赖。同时太深的网络结构可能会导致模型在训练过程中的梯度消失。

2.2 融合 CNN 与 SETR 的语义分割方法

SETR 是基于 Transformer 的语义分割模型。通过自注意力机制在特征图上建立全局依赖关系,有助于提高模型对特征图中复杂特征的学习能力。SETR 结构如图 3 所示。

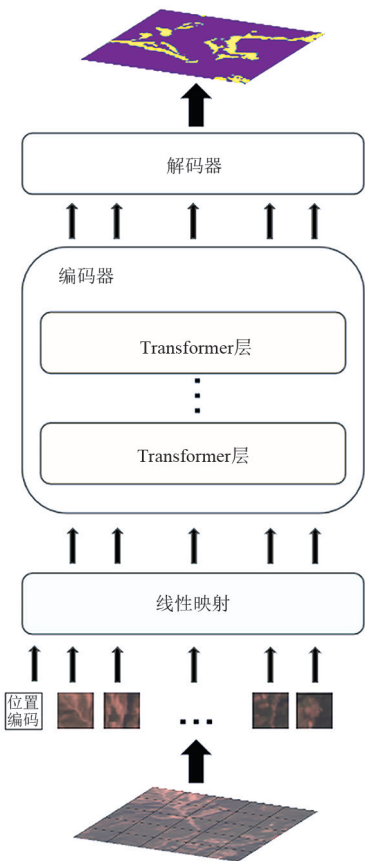


图 3 SETR 结构示意图
Fig.3 SETR structure diagram

将 CNN 与 Transformer 结合,可以在有效提取图像局部特征的同时,为图像中的所有特征建立全局依赖关系。从而提高模型对局部信息中细节的表达能力和对滑坡的识别能力。

本文利用 CNN 对边界判定更加精确的优点,同时结合 Transformer 在处理图像时捕获全局上下文关系的能力,创造出一个具有对滑坡体特征学习深刻,且对滑坡体边界判定精确的新模型。具体做法是:在 CNN 模型的编码器架构部分融合 SETR 模型,并将最终的融合结果在解码器架构部分进行输出;最终得到漏检率更低,且对边界的定位更加准确的模型。由于 Transformer 模型的复杂性和参数量较大,相对于传统的 CNN 模型,其训练时间可能有所加长。

U-Net 与 SETR,输入 14 通道的特征图,为了使模型提取到更多特征信息,首先将输入数据调整到 64 通道。将 U-Net 模型编码器中各个尺寸的特征图经过 SETR 模型建立特征的全局依赖关系,并将原始特征图和经过 SETR 融合的特征图以跳跃链接的方式与解码器中所对应的特征图融合。U-Net 与 SETR 融合过程如图 4 所示。

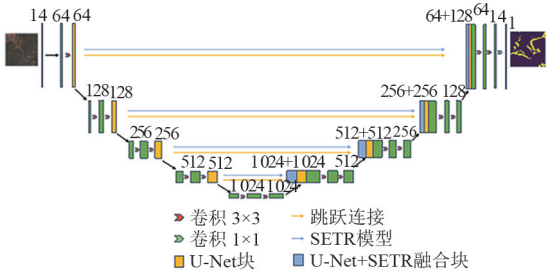


图 4 U-Net+SETR 示意图
Fig.4 Schematic diagram of U-Net+SETR

在本次实验中 Deeplabv3+的骨干神经网络采用 Xception^[21]网络模型。将原始数据分别输入到 Deeplabv3+模型的 Encoder 架构和 SETR 中,随后将经过 Xception 骨干神经网络提取的特征图输入到 SETR 建立全局依赖关系。将经过 Xception 学习的原始特征图、经过空洞空间金字塔池化 (atrous spatial pyramid pooling, ASPP) 模块处理的特征图以及经过 SETR 模型处理的特征图进行融合,经过上采样后再与最原始的经过 SETR 模型处理的特征图融合,共同做为模型的输出。其融合过程如图 5 所示。

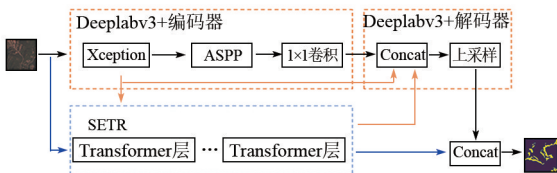


图 5 Deeplabv3+SETR 示意图
Fig.5 Deeplabv3+SETR schematic

FCN 与 SETR 融合是将 FCN 下采样过程中各个尺寸的特征图,输入到 SETR 模型中。并在 FCN

上采样中各个尺寸的特征图与相对应的 SETR 输出的特征图融合。融合的过程如图 6 所示。

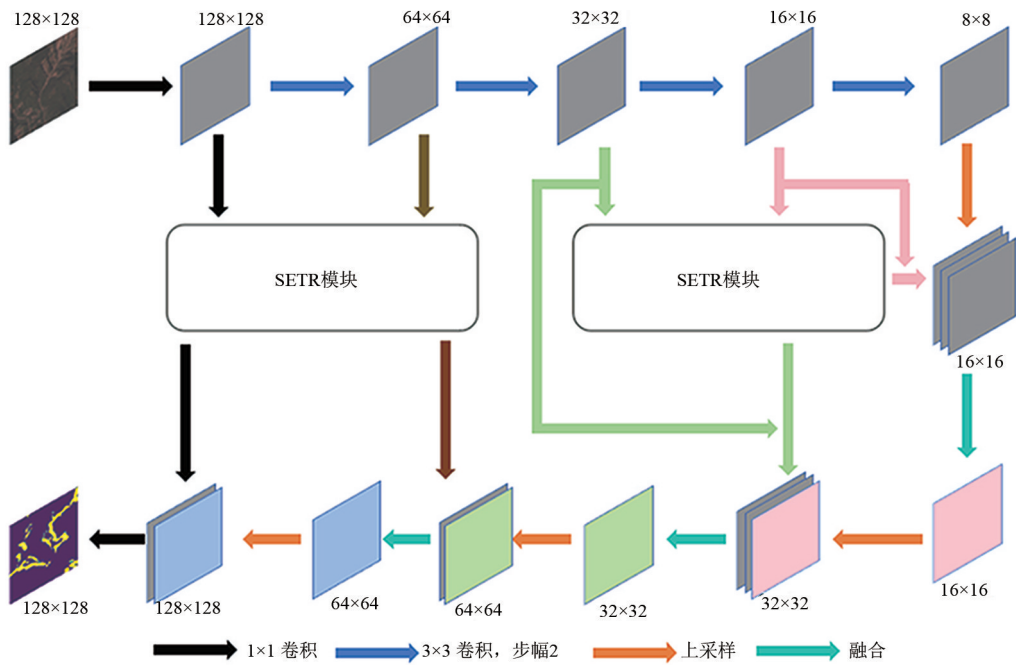


图 6 FCN+SETR 示意图
Fig.6 Schematic diagram of FCN+SETR

2.3 DEM 数据提取

滑坡的发生往往与地形信息密切相关,而 DEM 数据包含了各种地貌因子。在滑坡体检测中,融入 DEM 数据并加大 DEM 权重可以有效提高模型对滑坡检测的效果。本文将增大了 DEM 权重的数据输入到 U-Net 模型与融合了 SETR 的 U-Net 模型中。滑坡坡度数据和 DEM 提取流程图如图 7 所示。

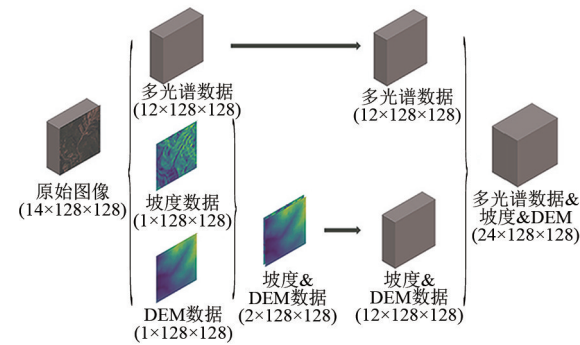


图 7 滑坡坡度数据和 DEM 提取流程
Fig.7 Flow chart of landslide slope data and DEM extraction

2.4 模型评价指标

采用 F1 分数来评价模型精度和性能。F1 分数是统计学中用来衡量二分类模型精度的一种指标,它同时兼顾了分类模型的精确率和召回率。精确率、召回率以及 F1 分数表达式分别为:

$$m_{\text{precision}} = \frac{TP}{TP + FP} \tag{1}$$

$$m_{\text{recall}} = \frac{TP}{TP + FN} \tag{2}$$

$$m_{F1} = \frac{2m_{\text{precision}} \cdot m_{\text{recall}}}{m_{\text{precision}} + m_{\text{recall}}} \tag{3}$$

式中: $m_{\text{precision}}$ 为模型的准确率; m_{recall} 为模型的召回率; m_{F1} 为模型的 F1 分数; TP , FP 和 FN 分别为检测结果真阳性、假阳性和假阴性的个数。

3 结果与分析

3.1 检测结果定性分析

CNN 融合 SETR 实验结果对比如表 1 所示。在当前数据集训练的模型下,U-Net 网络模型取得最好的检测效果。同时,从实验对比结果中可以看到,CNN 网络在融合了 SETR 模型后,对滑坡体边界的判定更加精确,大幅降低了滑坡体的漏检率,模型对滑坡体检测的视觉效果得到提高。以 U-Net 模型为基准,在加大 DEM 数据的权重后,模型对滑坡体特征的提取能力得到提高,漏检率进一步降低,对滑坡体边界的划分更加明确。加大 DEM 权重实验结果对比如表 2 所示。

表 1 实验结果对比图
Tab.1 Comparison of experimental results

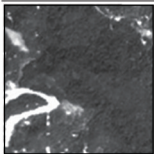
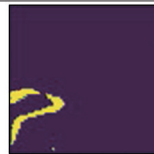
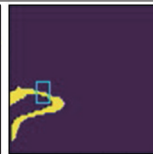
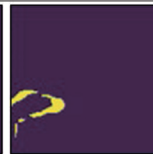
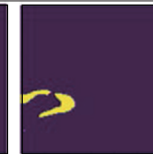
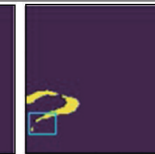
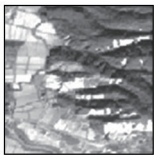
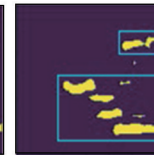
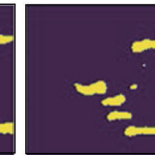
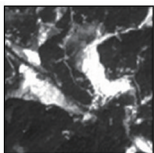
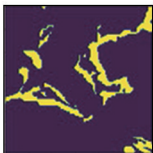
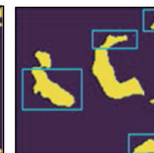
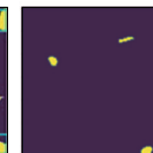
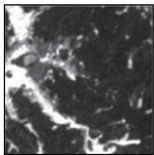
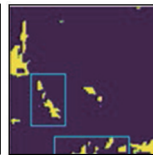
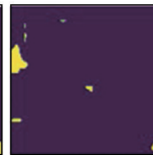
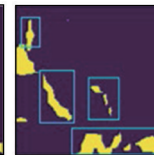
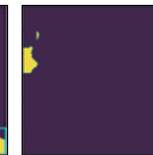
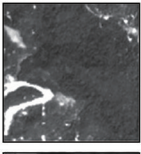
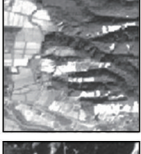
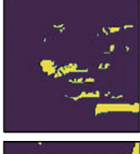
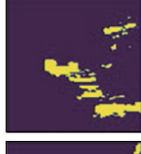
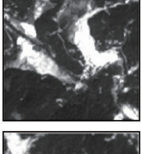
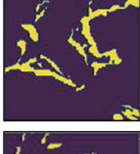
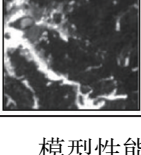
遥感影像	真实标签	U-Net	U-Net +SETR	Deeplabv3+	Deeplabv3+ +SETR	FCN	FCN+SETR
							
							
							
							

表 2 加大 DEM 权重实验结果对比图
Tab.2 Comparison of experimental results with
the higher weight of DEM

遥感影像	真实标签	U-Net+DEM	U-Net+DEM +SETR
			
			
			
			

模型性能与模型结构本身以及数据集有很大关系。在现实情况下正样本数量占比小,为了使模型达到理想效果,首先需要均衡正负样本数量,增加正样本的权重,以确保正负样本平衡。除了模型与数据本身之外,学习率以及模型的其他参数也可能会成为影响模型性能的关键。

为了尽快达到理想效果,首先采用较大的学习率,当模型性能遭遇瓶颈时,使用预训练方法改小模

型学习率。

3.2 检测结果定量分析

表 3 为 CNN 模型融合 SETR 前后的模型性能对比。FCN 融合 SETR 后召回率从 68.83%降低至 65.13%,准确率从 65.75%提高到 72.65%,F1 分数从 0.672 6 提高到 0.686 9; U-Net 融合 SETR 后召回率从 68.47%提升到 74.26%,准确率略有降低,从 77.56%降低至 74.34%,F1 分数从 0.727 3 提高到 0.743 0; Deeplabv3 + 融合 SETR 后,召回率从 61.58%提高到 66.13%,准确率从 77.77%降低至 75.60%,F1 分数从 0.687 3 提高到 0.705 5。融合后的模型虽准确率略有降低,但总体上召回率和 F1 分数得到大幅提升,召回率越高代表漏检率越低,更符合实际应用需求。

表 3 CNN 模型融合 SETR 性能对比
Tab.3 Comparison of SETR performance of
each model fusion

模型指标	U-Net	U-Net+ SETR	Deep- labv3+	Deeplabv3+ +SETR	FCN	FCN+ SETR
召回率/%	68.47	74.26	61.58	66.13	68.83	65.13
准确率/%	77.56	74.34	77.77	75.60	65.75	72.65
F1 分数	0.727 3	0.743 0	0.687 3	0.705 5	0.672 6	0.686 9
F1 提升/%	2.16		2.65		2.13	

以 U-Net 模型为基准,在加大 DEM 数据权重后,召回率为 69.15%,准确率为 77.88%,F1 分数为 0.732 5; 融合 SETR 后,召回率为 75.34%,准确率为 74.73%,F1 分数为 0.750 3。U-Net 模型加大 DEM

权重性能对比如表 4 所示。

表 4 U-Net 模型加大 DEM 权重性能对比
Tab.4 U-Net model increases DEM weight performance comparison

模型指标	U-Net	U-Net+ SETR	U-Net+ DEM	U-Net+ DEM+SETR
召回率/%	68.47	74.26	69.15	75.34
准确率/%	77.56	74.34	77.88	74.73
F1 分数	0.727 3	0.743 0	0.732 5	0.750 3
F1 提升/%	2.16		2.43	

以上结果表明,将 CNN 网络模型与 SETR 模型融合后,F1 分数均有所提高。其中 U-Net 与 Deep-labv3+在小幅损失准确率的前提下,提高了召回率,降低了模型漏检率;U-Net 模型在加大 DEM 数据权重后,模型的召回率、准确率、F1 分数均得到了提升。

3.3 多结构语义分割模型分析

CNN 是一种分层的数据表示方式,高层的特征表示依赖于底层的特征表示,由浅入深逐步抽象地提取更具备高级语义信息的特征。将每一层学习到的特征连接到一起,从而最终形成强大的特征表现能力。此外,CNN 通过共享卷积核来提取特征,这样一方面可以大大降低参数量来避免更多冗余的计算从而提高网络模型计算的效率,另一方面又结合卷积和池化使网络具备一定的平移不变性和平移等变性。拥有平移等变性,对于分割任务至关重要。

尽管 CNN 拥有诸多优势,但是通过不断的堆叠更深的卷积层,依然会造成模型臃肿以及计算量骤升等问题,这违背了深度学习的初衷。Transformer 的优势在于利用注意力的方式来捕获全局的上下文信息从而对目标建立起远距离的依赖,从而提取出更强有力的特征。然而 Transformer 模型的训练需要大量数据集才得以呈现出理想的效果。因此在缺乏足够多的数据集的情况下,将 CNN 与 Transformer 结合或许可以达到更加理想的效果,实验的最终结果也证实了这一点。

4 结论

本文主要讨论了 CNN 模型在融合基于 Trans-former 结构的 SETR 模型前后性能的对比。证明 CNN 模型在融合 SETR 模型后整体 F1 分数得到提升。虽然准确率略有降低,但召回率得到更大提升,从而降低了模型的漏检率,且对滑坡体边界的处理得到改善。

实验结果表明,CNN 模型在融合了 SETR 后,不论是视觉效果还是模型评价指标,对滑坡检测的精

度都得到提升。FCN 模型在融合 SETR 后,F1 分数从 0.672 6 提高到 0.686 9;U-Net 模型在融合 SETR 后,F1 分数从 0.727 3 提高到 0.743 0;Deeplabv3+模型在融合 SETR 后,F1 分数从 0.687 3 提高到 0.705 5;以性能最好的 U-Net 为基准,在加大了 DEM 权重后,U-Net 模型 F1 分数为 0.732 5,融合 SETR 后,F1 分数提高到 0.750 3。

由于滑坡数据集的样本数量少,且滑坡形成的影响因素多。提高数据集样本数量、构建包含丰富特征信息的数据集,可以进一步提高滑坡检测精度。

参考文献

[1] 王 宇.基于深度学习的遥感图像滑坡信息提取算法研究[D].哈尔滨:黑龙江大学,2020.
Wang Y. Research on landslide information extraction algorithm from remote sensing images based on deep learning[D].Harbin: Helongjiang University, 2020.

[2] 张帅娟.变化检测和面向对象结合的高分辨率遥感影像滑坡体提取方法研究[D].成都:西南交通大学,2017.
Zhang S J.Study on landslide extraction method from high-resolution remote sensing images by combining change detection with object-oriented method[D].Chengdu:Southwest Jiaotong University, 2017.

[3] 张 雨,明冬萍,赵文祎,等.基于高分光学卫星影像的泸定地震型滑坡提取与分析[J].自然资源遥感,2023,35(1): 161-170. doi: 10.6046/ zrzyyg.2022434.
Zhang Y,Ming D P,Zhao W Y,et al.The extraction and analysis of Luding earthquake-induced landslide based on high-resolution optical satellite images[J].Remote Sensing for Natural Resources, 2023,35(1): 161-170. doi: 10.6046/ zrzyyg.2022434.

[4] 张 越,宋炜炜.基于 BP 神经网络和决策树的昆明市东川区滑坡空间易发性评价[J].国土与自然资源研究,2023(2): 67-70.
Zhang Y,Song W W.Spatial susceptibility evaluation of landslide in Dongchuan District of Kunming based on BP neural network and decision tree[J].Territory & Natural Resources Study,2023(2): 67-70.

[5] 余淙蔚,柳 侃,殷 杰,等.一种适用于逻辑回归模型评价浅层滑坡易发性的网格尺度划分方法——以 2019 年福建省三明市群发浅层滑坡为例[J].山地学报,2022,40(1): 106-119.
Yu C W,Liu K,Yin J,et al.A grid-scale division method applicable to logistic regression models for evaluating the susceptibility of shallow landslides: Taking the 2019 cluster of shallow landslides in Sanming,Fujian as example[J].Mountain Research, 2022, 40(1): 106-119.

[6] 许高程,张文君,王卫红.支持向量机技术在遥感影像滑坡体提取中的应用[J].安徽农业科学,2009,37(6): 2781-2782,2822.
Xu G C,Zhang W J,Wang W H.Application of supporting vector machine technology in extraction of remote sensing image of landslide[J].Journal of Anhui Agricultural Sciences, 2009, 37(6): 2781-2782,2822.

[7] 牛朝阳,高欧阳,刘 伟,等.光学遥感图像滑坡检测研究进展[J].航天返回与遥感,2023,44(3): 133-144.

Niu C Y, Gao O Y, Liu W, et al. Research progress of landslide detection in optical remote sensing images[J]. Spacecraft Recovery & Remote Sensing, 2023, 44(3): 133–144.

[8] Yu B, Chen F, Xu C. Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015[J]. Computers & Geosciences, 2020, 135: 104388.

[9] Ji S, Yu D, Shen C, et al. Landslide detection from an open satellite imagery and digital elevation model dataset using attention boosted convolutional neural networks[J]. Landslides, 2020, 17(6): 1337–1352.

[10] Liu P, Wei Y, Wang Q, et al. Research on post-earthquake landslide extraction algorithm based on improved U-net model[J]. Remote Sensing, 2020, 12(5): 894.

[11] 杨昭颖, 韩灵怡, 郑向向, 等. 基于卷积神经网络的遥感影像及 DEM 滑坡识别——以黄土滑坡为例[J]. 自然资源遥感, 2022, 34(2): 224–230. doi: 10.6046/zrzyyg.2021204.

Yang Z Y, Han L Y, Zheng X X, et al. Landslide identification using remote sensing images and DEM based on convolutional neural network: A case study of loess landslide[J]. Remote Sensing for Natural Resources, 2022, 34(2): 224–230. doi: 10.6046/zrzyyg.2021204.

[12] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE, 2017: 640–651.

[13] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al, eds. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234–241.

[14] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//European Conference on Computer Vision. Cham: Springer, 2018: 833–851.

[15] 白石, 唐攀攀, 苗朝, 等. 基于高分辨率遥感影像和改进 U-Net 模型的滑坡提取——以汶川地区为例[J]. 自然资源遥感, 2024, 36(3): 96–107. doi: 10.6046/zrzyyg.2023132.

Bai S, Tang P P, Miao Z, et al. Information extraction of landslide based on high resolution remote sensing images and an improved U-Net model: A case study of Wenchuan, Sichuan[J]. Remote Sensing for Natural Resources, 2024, 36(3): 96–107. doi: 10.6046/zrzyyg.2023132.

[16] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J/OL]. arXiv, 2020. https://arxiv.org/abs/2010.11929v1.

[17] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//Vedaldi A, Bischof H, Brox T, et al. European Conference on Computer Vision. Cham: Springer, 2020: 213–229.

[18] Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[J/OL]. arXiv, 2021. https://arxiv.org/abs/2103.14030v2.

[19] Xie E Z, Wang W H, Yu Z D, et al. SegFormer: Simple and efficient design for semantic segmentation with Transformers[J/OL]. arXiv, 2021. https://arxiv.org/abs/2105.15203.

[20] 张思远, 谭晓玲, 朱木雷, 等. 基于改进 Swin Transformer 的滑坡分割算法[J]. 国外电子测量技术, 2023, 42(11): 49–56.

Zhang S Y, Tan X L, Zhu M L, et al. Landslide segmentation algorithm based on improved Swin Transformer[J]. Foreign Electronic Measurement Technology, 2023, 42(11): 49–56.

[21] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1800–1807.

A landslide detection method using CNN- and SETR-based feature fusion

LI Shiqi, YAO Guoqing

(School of Information Engineering, China University of Geosciences(Beijing), Beijing 100083, China)

Abstract: The accurate and timely detection of landslides is of great significance for reducing the threats to human life and properties, along with relevant losses, caused by landslides. This study proposed a landslide detection method using feature fusion based on convolutional neural networks (CNNs) and Segmentation Transformer (SETR). The CNN-based models utilized a fully convolutional network (FCN), U-Net, and Deeplabv3+, while the Transformer-based models used SETR. First, the landslide detection effects of the CNN-based models were evaluated. Then, SETR was introduced into the encoders of the CNN-based models, and the output of SETR was fused into the CNN decoder structure as the final output of the models. The experiments using the LandSlide4Sense dataset indicate that the fusion of typical CNNs with SETR can effectively improve the landslide detection effects. After SETR fusion, the FCN, U-Net, and Deeplabv3+ models exhibited higher F1-scores, which increased from 0.672 6, 0.727 3, and 0.687 3 to 0.686 9, 0.743 0, and 0.705 5, respectively. Given the close relationship between landslides and terrain, a digital elevation model (DEM) was incorporated into the U-Net model, which outperformed other models. As a result, the F1-score of the model increased from 0.732 5 to 0.750 3.

Keywords: CNN; SETR; landslide detection; semantic segmentation

(责任编辑: 张 仙)