

基于深度残差全卷积网络的 Landsat 8 遥感影像云检测方法

张家强^{1,2,3}, 李潇雁^{1,2,3}, 李丽圆^{1,2,3}, 孙鹏程^{2,4}, 苏晓峰^{1,2*}, 胡亭亮^{1,2}, 陈凡胜^{1,2}

¹中国科学院智能红外感知重点实验室, 上海 200083;

²中国科学院上海技术物理研究所, 上海 200083;

³中国科学院大学, 北京 100049;

⁴上海大学上海先进通信与数据科学研究所专业光纤与光接入网络重点实验室,
专业光纤与先进通信国际联合研究实验室, 上海 200444

摘要 为了实现量化应用目标,高精度的云层检测已成为遥感数据预处理的关键步骤之一。然而,传统的云检测方法存在特征复杂、算法步骤多、鲁棒性差,且无法将高级特征和低级特征相结合的缺陷,检测效果一般。针对以上问题,提出了一种基于深度残差全卷积网络的高精度云检测方法,能够实现对遥感影像云层目标像素级别的分割。首先,编码器通过残差模块的不断降采样提取图像深层特征;然后,应用双线性插值进行上采样,结合多层次编码后的图像特征完成解码;最后,将解码后的特征图与输入图像融合后再次进行卷积,实现端到端的云检测。实验结果表明,对于 Landsat 8 云检测数据集,所提方法的像素精度达到 93.33%,比原版 U-Net 提高了 2.29%,比传统 Otsu 方法提高了 7.78%。该方法可以为云层目标智能化检测研究提供有益参考。

关键词 遥感; 云检测; 深度学习; 语义分割; 全卷积网络; 残差网络

中图分类号 P237

文献标志码 A

doi: 10.3788/LOP57.102801

Landsat 8 Remote Sensing Image Based on Deep Residual Fully Convolutional Network

Zhang Jiaqiang^{1,2,3}, Li Xiaoyan^{1,2,3}, Li Liyuan^{1,2,3}, Sun Pengcheng^{2,4},
Su Xiaofeng^{1,2*}, Hu Tingliang^{1,2}, Chen Fansheng^{1,2}

¹Key Laboratory of Intelligent Infrared Perception, Chinese Academy of Sciences, Shanghai 200083, China;

²Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

³University of Chinese Academy of Sciences, Beijing 100049, China;

⁴Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute of Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China

Abstract In order to achieve the goal of quantitative application, high-precision cloud detection has become one of the key steps in remote sensing data preprocessing. However, traditional cloud detection methods have the disadvantages of complex features, multiple algorithm steps, poor robustness, inability to combine high-level features with low-level features, and ordinary detection performance. In view of the above problems, this paper proposes a high-precision cloud detection method based on deep residual fully convolutional network, which can achieve the target pixel level segmentation of cloud layer in remote sensing images. First, the encoder extracts the deep features of the image through continuous down-sampling of the residual module. Then, the bilinear interpolation is used for sampling, and the decoding is completed by combining the image features after multilevel

收稿日期: 2020-02-17; 修回日期: 2020-02-21; 录用日期: 2020-02-25

基金项目: 中组部“万人计划”青年拔尖人才、国家自然科学基金面上项目(61975222)

* E-mail: fishsu@mail.sitp.ac.cn

coding. Finally, the decoded feature map is fused with the input image and convolution is performed again to achieve end-to-end cloud detection. Experimental results show that, in terms of the Landsat 8 cloud detection data set, the pixel accuracy of the proposed method reaches 93.33%, which is 2.29% higher than that of the original U-Net, and 7.78% higher than that of the traditional Otsu method. This method can provide useful reference for research on intelligent detection of cloud targets.

Key words remote sensing; cloud detection; deep learning; semantic segmentation; fully convolutional network; residual network

OCIS codes 280.4788; 100.3008; 100.2980; 100.2960

1 引言

在光学遥感中,云作为一种广泛存在的物体是影响地物分析及遥感影像特征提取精度的重要因素。在 Landsat 5 和 Landsat 7 的图像中云覆盖率可达 40%^[1],较高的云覆盖率会显著降低遥感数据的准确性和应用价值,因此提高云检测精度对遥感影像的应用具有重要的现实意义。

目前的云检测方法可分为基于规则和基于机器学习两类。光谱分析法是最早且应用最广泛的基于规则的云检测方法,其主要特点是利用云在可见、短波红外、热红外等不同波段的辐射特性、反射特性进行云检测^[2]。刘心燕等^[3]基于云与地表的光谱特征,采用光谱差异分析技术完成了高分四号卫星数据的云检测,正确率可达 82% 以上。光谱分析法虽能够在特定传感器上取得相对不错的云检测效果,但这类算法只适用于某一特定数据集,因而普适性较差^[4]。为了弥补光谱法的缺陷,机器学习算法被引入云检测研究,用于图像特征的分类^[5-6]。Li 等^[7]应用支持向量机对反射率和灰度共生矩阵分类,对小型数据集的云检测精度超过 90%。支持向量机等算法能够更好地运用云层的纹理信息,提高云检测的效果,比光谱法更具有普适性,但需要人为选取特征。

自 2016 年起,深度学习开始被应用于云检测的研究^[8-9]。经典的卷积神经网络由卷积层、池化层和

全连接层等组合而成,需要对大量标注过的数据进行训练,属于监督机器学习算法。陈洋等^[10]以主成分分析作为预处理手段,使用支持向量机对卷积神经网络提取到的特征进行分类,完成对资源三号卫星图像的云检测,但该方法步骤较多,且神经网络较浅,在较大数据集上容易过拟合。高琳等^[11]应用基于膨胀卷积的神经网络完成对资源三号遥感影像的端到端云识别,但文中所用数据集较小,该算法在大型数据集的表现有待验证。Zi 等^[12]应用卷积神经网络的变种网络(PCANet)与支持向量机,将光谱阈值法与机器学习中的简单线性聚类算法相结合,完成对 Landsat 8 图像的云检测,准确率达到 93%,其主要缺点是推断过程需分步进行,无法完成端到端的处理。

本文针对传统方法普适性较差、需要人工选取特征及无法端到端检测的问题,提出了一种基于残差网络 ResNet34^[13]的改进型 U-Net^[14]网络。通过引入残差模块提高 U-Net 的网络表达能力与泛化能力,在 Landsat 8 云覆盖评价(L8 CCA)数据集^[15]上较原版 U-Net 和传统云检测算法在检测精度上有较大提高。

2 理论方法

2.1 基于残差网络的 U-Net 网络

图 1 所示为本文方法的实验流程。将 L8 CCA 数据集划分为训练集、验证集和测试集,训练集与验

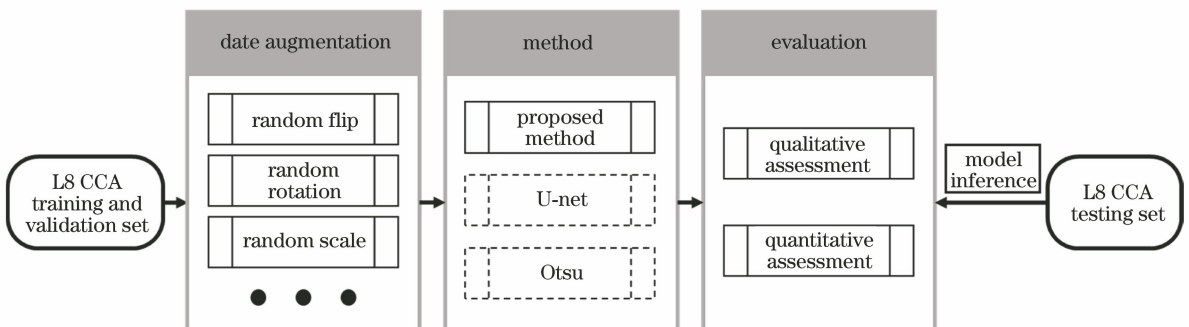


图 1 本文方法框图

Fig. 1 Block diagram of proposed method

证集经过数据增强后输入网络进行训练,将测试集输入训练收敛后的网络模型进行定性、定量评估,并与原版 U-Net 和传统方法 Otsu 算法的结果进行对比。

本文所采用的网络是结合残差网络与 U-Net 模型的基于编码-解码架构的全卷积网络,如图 2 所

示。输入数据经过由残差模块组成的编码器后生成多种尺度的高维特征向量;在解码阶段,网络不断对特征图进行上采样,并与编码器生成的相应尺度的特征向量组合;最终网络输出一个包含每个像素所属类别的置信度的二维特征图。置信度最大的类别即为该像素的预测类别。

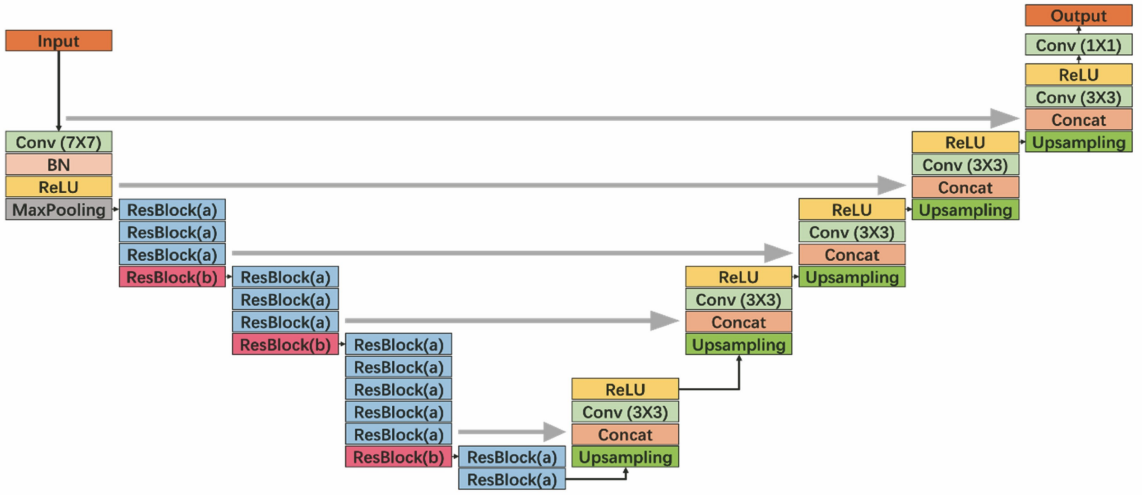


图 2 基于残差网络的 U-Net 结构图

Fig. 2 Architecture of U-Net based on residual network

U-Net^[14]最初被应用于生物医学影像的分割,因其轻量级、易训练、更适合小数据集等特点,在医学诊断、遥感等领域也有广泛的应用。U-Net 的主体由四次编码过程和四次解码过程组成,四次编码、解码之间有对应的跨层连接。编码器包括两层 3×3 卷积层、一层非线性函数 ReLU 及用于下采样的 2×2 的最大池化层。解码器包括两层 3×3 卷积层、一层 ReLU 及用于上采样的 2×2 反卷积。编码器与解码器间的跨层连接使 U-Net 能够同时提取到图像的高层及低层语义,提高分割效果^[16]。

随着深度学习网络的不断加深,研究者们发现过深的神经网络会在训练中产生梯度消失问题,从而导致模型正确率下降^[13]。为了解决这个问题,He 等^[13]在 2016 年提出了深度残差网络,即 ResNet。ResNet 的核心是残差模块,图 3(a)为一个残差模块,其公式为

$$y = F(x) + x, \quad (1)$$

式中: x 和 y 代表残差模块的输入和输出; $F(x)$ 为待学习的残差函数。图 3(b)为 ResNet34 中用于下采样的残差模块。

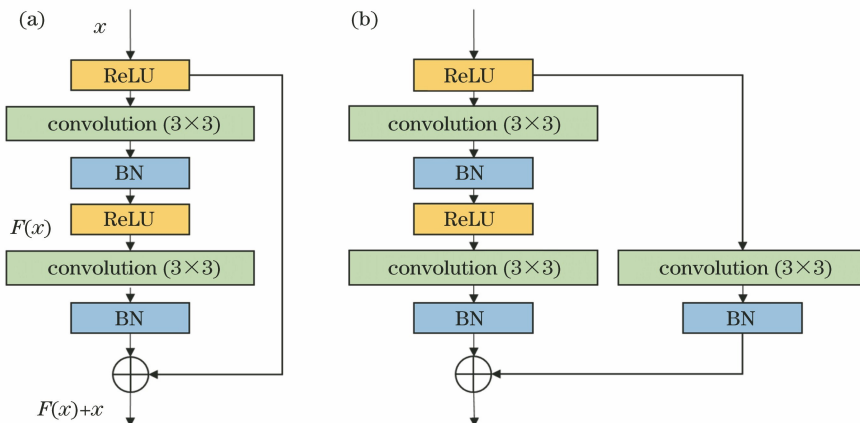


图 3 ResNet34 残差模块图。(a)特征提取残差模块;(b)下采样残差模块图

Fig. 3 ResNet34 residual block diagram. (a) Residual block for feature extraction; (b) residual block for down-sampling

为了提高 U-Net 的云检测精度,增强网络的鲁棒性和泛化能力,受 U-Net 中的跨层连接和残差网络中的短连接的相似设计的启发,本文将两种网络结合提出一种基于残差网络的 U-Net 模型。将残差模块应用于语义分割的编码器,代替原版 U-Net 中双卷积与最大池化的编码器设计,并将改进后的网络用于遥感图像的云检测。

本文将 ResNet34 的四次下采样的结果连接至解码器的上采样,设计出了基于残差网络改进的 U-Net。改进后的网络仍为编码-解码架构,其中编码器以 ResNet34 的组合卷积模块代替原版 U-Net 中

的双卷积编码模块。每个编码器中所包含的残差模块具体数量见表 1。将经由卷积模块操作后的特征跨层连接至对应解码器。为了避免过拟合,提高网络的泛化能力,在优化器中加入 L2 正则化。由于池化层在不断降采样过程中可能会导致特征的丢失,本文使用图 3(b)的残差模块代替最大池化层。而原解码器中的上采样操作使用反卷积操作,本文中将其改进为更易训练的双线性插值方法。最终网络结构如图 2 所示,具体结构参数见表 2。可训练参数由原版 U-Net 的 31390786 个增加至 53344258,提高了网络的表达能力及泛化能力。

表 1 ResNet34 结构参数

Table 1 Structural parameters of ResNet34

Layer name	Conv1	Conv2_x	Conv3_x	Conv4_x	Conv5_x
Output size	112×112	56×56	28×28	14×14	7×7
Information of blocks	Conv 7×7,64	Max_pool 3×3 $\begin{bmatrix} \text{Conv } 3 \times 3, 64 \\ \text{Conv } 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} \text{Conv } 3 \times 3, 128 \\ \text{Conv } 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} \text{Conv } 3 \times 3, 256 \\ \text{Conv } 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} \text{Conv } 3 \times 3, 512 \\ \text{Conv } 3 \times 3, 512 \end{bmatrix} \times 3$

表 2 基于残差网络的 U-Net 结构参数

Table 2 Parameters of ResNet-based U-Net

Layer name	Input size	Information of blocks	Output size
Encoder1	224×224×3	$\begin{bmatrix} \text{Conv}(7 \times 7) \\ \text{BatchNorm, ReLU} \\ \text{Max_pool}(3 \times 3) \end{bmatrix}$	112×112×64
Encoder2	112×112×64	Conv2_x	56×56×128
Encoder3	56×56×128	Conv3_x	28×28×256
Encoder4	28×28×256	Conv4_x	14×14×512
Bridge	14×14×512	Conv5_x	7×7×1024
Decoder1	7×7×1024	$\begin{bmatrix} \text{Up-sampling} \\ \text{Concat} \\ \text{Conv}(3 \times 3), \text{ReLU} \end{bmatrix}$	14×14×512
Decoder2	14×14×512	$\begin{bmatrix} \text{Up-sampling} \\ \text{Concat} \\ \text{Conv}(3 \times 3), \text{ReLU} \end{bmatrix}$	28×28×256
Decoder3	28×28×256	$\begin{bmatrix} \text{Up-sampling} \\ \text{Concat} \\ \text{Conv}(3 \times 3), \text{ReLU} \end{bmatrix}$	56×56×128
Decoder4	56×56×128	$\begin{bmatrix} \text{Up-sampling} \\ \text{Concat} \\ \text{Conv}(3 \times 3), \text{ReLU} \end{bmatrix}$	112×112×64
Decoder5	112×112×64	$\begin{bmatrix} \text{Up-sampling} \\ \text{Concat} \\ \text{Conv}(3 \times 3), \text{ReLU} \end{bmatrix}$	224×224×64
Output	224×224×64	Conv(1×1)	224×224×2

2.2 评价指标

本文选择像素精度 (PA)、平均像素精度

(mAP)及平均交并比(mIoU)作为评价指标^[18]。像素精度即分类正确的像素点数和所有的像素点数的

比例,平均像素精度即所有类别分类正确的像素点数和该类的所有像素点数的比例的均值,平均交并比即所有类别的 IoU 的均值,IoU 为交并比,即一个类别与预测结果与真值交集和并集之比。计算公式分别为

$$A_{\text{PA}} = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}}, \quad (2)$$

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

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

式中: k 均为类别; p_{ii} 为真正即被正确分为正的像素; p_{ij} 为假正; p_{ji} 为假负。

3 实验与分析

3.1 数据与参数设置

使用 L8 CCA 数据集中分辨率为 30 m 的可见光波段数据,并将波段 4、3、2 依次作为 RGB 通道合成真彩色图像。由于 L8 CCA 中的图像大小不一,本文将图像及真值统一剪裁为 4480×4480 大小。为了有效地进行训练,本文将图像再次剪裁为 224×224 大小的小块图像,即 patch。最终选择 84 张图像,生成 30800 张小块图像,其中 28000 张作为训练集,2800 张作为验证集,另有 7 张 4480×4480 大小的图像作为测试集。训练集经过包括旋转、放缩、翻转、Gamma 变换、亮度变换等数据增强操作后输入模型。

硬件环境为型号 E5-2620 v4 英特尔至强系列 CPU,32 GB 内存,显存大小为 11 GB 的英伟达 GeForce 1080 Ti 显卡;操作系统为乌班图。所有方法均使用 Python 3.6 进行代码编写,并使用 1.3.1 版本的 PyTorch^[18] 深度学习框架进行训练及推断。使用 Dice 损失^[19] 与二分类交叉熵^[20] 的平均值作为损失函数,公式分别为

$$L_{\text{Dice}} = 1 - \frac{2|P \cap T|}{|P| + |T|}, \quad (5)$$

$$L_{\text{Bce}} = - \sum_i T_i \ln(P_i) + (1 - T_i) \ln(1 - P_i), \quad (6)$$

$$L = (L_{\text{Bce}} + L_{\text{Dice}}) / 2, \quad (7)$$

式中: P 为网络预测结果; T 为真值。

本文使用经数据增强处理的 28000 张小块图像进行了 70 个周期的训练。批大小设置为 16,学习率的初值被设置为 0.001,并应用 Adam^[21] 优化器进行优化,使学习率以 0.5 的下降系数每 10 个周期更新一次。

3.2 结果与分析

图 4 为 70 个训练周期内的损失和评价指标的变化。其中,图 4(a)为网络在训练集和验证集上的损失变化,图 4(b)为在验证集的像素精度、平均像素精度、平均交并比变化。由图 4 可知,损失随着周期的增加逐渐降低并趋于稳定,且评价指标随着周期的增加逐渐提高并趋于稳定,最终显示为逐步收敛的模型。35 周期前,训练损失及模型在验证集上的评价指标均有较大起伏,其原因可能为数据增强操作:训练集数据均经过数据增强后输入模型,但一个周期结束后的验证阶段并未进行数据增强,所以训练前期的模型作用于真实数据会出现振荡。而在 50 周期后损失及指标变化均趋于平缓,所以在 50 周期后可以认为模型已经收敛。最终,本文选择第 70 周期的模型进行推断测试。

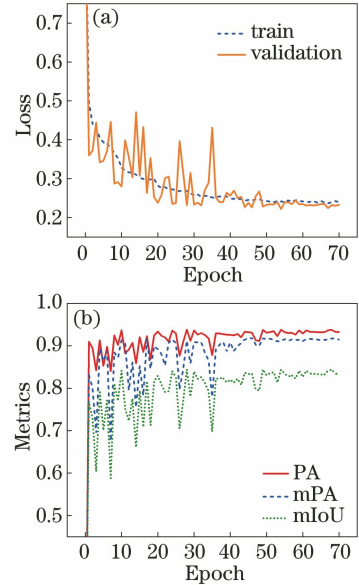


图 4 训练过程。(a) 损失随周期变化曲线;(b) 评价指标随周期变化曲线

Fig. 4 Training procedure. (a) Loss change curves versus the number of epochs; (b) metrics change curves versus the number of epochs

图 5 为本文方法、原版 U-Net 与传统方法 Otsu 算法^[22] 的结果对比,由上至下分别为 LC81590362014051LGN00、LC80290372013257LGN00 和 LC80200462014005LGN00 的场景。可以看出,

本文方法不论在总体结果[图 5 第一行中方框]还是在薄云[图 5 第二行中方框]及碎云[图 5 第三行中方框]的检测上均有更好表现。虽然 Otsu 算法在简单的云检测图像中应用较多,但在复杂环境下无法准确检测云层(如具有高反射地物的场景),并且

由于部分碎云的反射率较低,所以在碎云较多的图像中[图 5 第三行],Otsu 算法的效果较差。尽管原版 U-Net 在总体上的结果优于 Otsu 算法,但其网络较浅,无法有效地提取图像在薄云区域的特征,导致其在个别测试图像上的结果较差,逊于本文方法。

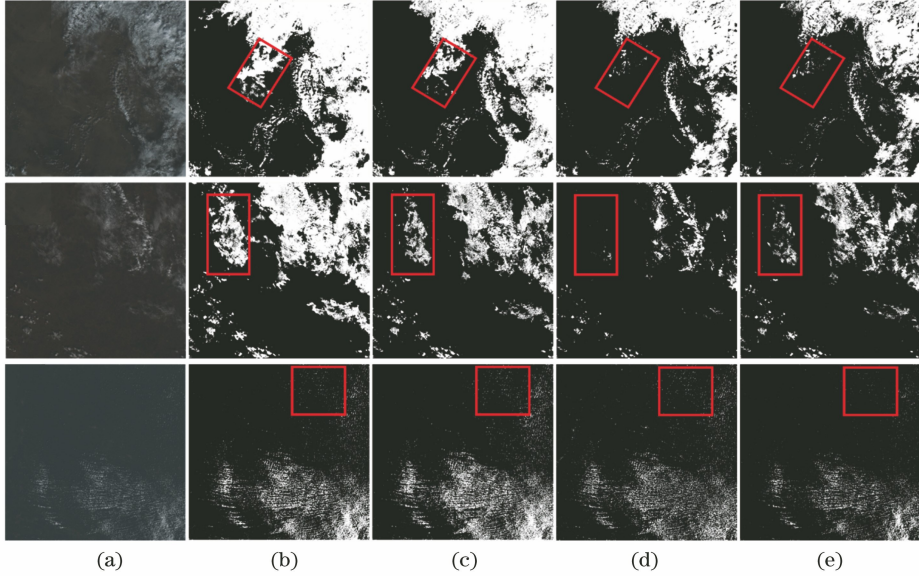


图 5 云检测示例。(a)输入图像;(b)云层的真值图像;(c)本文方法结果;(d) U-Net 结果;(e) Otsu 法结果

Fig. 5 Examples of cloud detection. (a) Input images; (b) ground truth cloud mask; (c) results of proposed method; (d) results of U-Net; (e) results of Otsu method

为进一步量化该方法的有效性,本文将测试集图像应用于该网络,其中推断时间为检测一张 4480×4480 大小的测试集图像的平均时间,最终测试的平均结果及其与原版 U-Net、传统 Otsu 算法的结果对比如表 3 所示,测试集各图像详细结果如表 4 所示。结果显示,基于残差网络改进后的 U-Net 在像素精度、平均像素精度及平均交并比上都有较大提升,其中总体像素精度为 93.33%,相较原版 U-Net 提高了 2.29%,较 Otsu 算法提高了 7.78%;平均像素精度为 93.43%,分别提升了 8.9%和 15.19%;平均交并比为 83.88%,分别提升了 7.85%和 16.84%。在运算时间方面,Otsu 算法由于无需进行大量计算故运算速度较快,但其检测效果较差;而相比原版 U-Net 算法,本文方法在获得更高精确度结果的同时,并未大幅增加运行时间。

表 3 不同云检测方法的结果对比

Table 3 Results comparison of different cloud detection methods

Method	PA / %	mPA / %	mIoU / %	Inference time / s
Otsu	85.55	78.24	67.04	1.0
U-Net	90.04	84.53	76.03	5.4
Proposed method	93.33	93.43	83.88	6.4

经过定性与定量的评价,实验结果显示本文方法优于传统云检测方法,同时能够弥补原版 U-Net 的缺点,在保持同量级运行时间的情况下,提升了检测的精度,实现端到端的云检测。

4 结 论

本文将基于深度学习的语义分割技术应用于云检测,提出了一种基于深度残差网络改进的 U-Net,并将网络在 Landsat 8 遥感图像上进行了验证。本文方法的像素精度达到 93.33%,平均像素精度达到 93.43%,平均交并比提升至 83.88%,相较 U-Net 及传统云检测方法 Otsu 算法有较大提升。实验结果显示,本文方法具有检测精度高、误检少的优势,并且在存在薄云及碎云的场景有着优于原版 U-Net 和传统方法的表现,证明了本文方法的有效性。

基于语义分割技术的云检测还有很大的潜力,本文方法还存在不足,如训练及推断时间较长,模型较大。后续研究将引入剪枝、量化、知识蒸馏等方法进行模型压缩,进而提高网络推断速度,减少模型所需计算资源。

表4 测试集详细结果
Table 4 Detail results of test set

Scene ID	Detection algorithm	PA/%	mPA/%	mIoU /%
LC80200462014005LGN00	Otsu	95.00	76.27	73.59
	U-Net	96.44	86.46	82.32
	Proposed	96.59	95.66	85.32
LC80210072014236LGN00	Otsu	88.03	80.99	66.62
	U-Net	93.36	95.87	80.49
	Proposed	87.82	92.49	70.10
LC80310202013223LGN00	Otsu	77.50	78.56	62.50
	U-Net	92.58	92.87	86.19
	Proposed	89.74	89.37	81.18
LC80290372013257LGN00	Otsu	83.29	70.28	60.85
	U-Net	76.28	57.81	45.41
	Proposed	97.56	94.28	92.57
LC81390292014135LGN00	Otsu	90.31	81.99	76.14
	U-Net	92.73	86.98	82.15
	Proposed	95.12	95.48	88.77
LC81590362014051LGN00	Otsu	79.69	76.20	63.12
	U-Net	83.34	80.47	69.21
	Proposed	93.51	92.52	87.35
LC81620432014072LGN00	Otsu	85.01	83.41	66.43
	U-Net	95.54	91.24	86.45
	Proposed	92.99	94.23	81.85

参 考 文 献

- [1] Ju J C, Roy D P. The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally [J]. Remote Sensing of Environment, 2008, 112(3): 1196-1211.
- [2] Oishi Y, Ishida H, Nakamura R. A new Landsat 8 cloud discrimination algorithm using thresholding tests [J]. International Journal of Remote Sensing, 2018, 39(23): 9113-9133.
- [3] Liu X Y, Sun L, Yang Y K, et al. Cloud and cloud shadow detection algorithm for Gaofen-4 satellite data [J]. Acta Optica Sinica, 2019, 39(1): 0128001.
刘心燕, 孙林, 杨以坤, 等. 高分四号卫星数据云和云阴影检测算法 [J]. 光学学报, 2019, 39(1): 0128001.
- [4] Li Z W, Shen H F, Cheng Q, et al. Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 150: 197-212.
- [5] Ghasemian N, Akhoondzadeh M. Introducing two Random Forest based methods for cloud detection in remote sensing images [J]. Advances in Space Research, 2018, 62(2): 288-303.
- [6] Christodoulou C I, Michaelides S C, Pattichis C S. Multifeature texture analysis for the classification of clouds in satellite imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(11): 2662-2668.
- [7] Li P F, Dong L M, Xiao H C, et al. A cloud image detection method based on SVM vector machine [J]. Neurocomputing, 2015, 169: 34-42.
- [8] Chen Y, Fan R S, Bilal M, et al. Multilevel cloud detection for high-resolution remote sensing imagery using multiple convolutional neural networks [J]. ISPRS International Journal of Geo-Information, 2018, 7(5): 181.
- [9] Mateo-Garcia G, Gomez-Chova L, Camps-Valls G. Convolutional neural networks for multispectral image cloud masking [C] // 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), July 23-28, 2017. Fort Worth, TX. IEEE, 2017: 2255-2258.
- [10] Chen Y, Fan R S, Wang J X, et al. Cloud detection of ZY-3 satellite remote sensing images based on deep learning [J]. Acta Optica Sinica, 2018, 38(1): 0128005.
陈洋, 范荣双, 王竞雪, 等. 基于深度学习的资源三号卫星遥感影像云检测方法 [J]. 光学学报, 2018,

- 38(1): 0128005.
- [11] Gao L, Song W D, Tan H, et al. Cloud detection based on multi-scale dilation convolutional neural network for ZY-3 satellite remote sensing imagery [J]. *Acta Optica Sinica*, 2019, 39(1): 0104002.
高琳, 宋伟东, 谭海, 等. 多尺度膨胀卷积神经网络资源三号卫星影像云识别 [J]. *光学学报*, 2019, 39(1): 0104002.
- [12] Zi Y, Xie F Y, Jiang Z G. A cloud detection method for Landsat 8 images based on PCANet [J]. *Remote Sensing*, 2018, 10(6): 877.
- [13] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016. Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [14] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [M] // *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2015: 234-241.
- [15] Foga S, Scaramuzza P L, Guo S, et al. Cloud detection algorithm comparison and validation for operational Landsat data products [J]. *Remote Sensing of Environment*, 2017, 194: 379-390.
- [16] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015. Boston, MA, USA. IEEE, 2015: 3431-3440.
- [17] Garcia-Garcia A, Orts-Escolano S, Oprea S, et al. A survey on deep learning techniques for image and video semantic segmentation [J]. *Applied Soft Computing*, 2018, 70: 41-65.
- [18] Paszke A, Gross S, Chintala S, et al. Automatic differentiation in pytorch [EB/OL]. [2020-02-17]. <https://openreview.net/forum?id=BJJsrmfCZ>.
- [19] Milletari F, Navab N, Ahmadi S A. V-net: fully convolutional neural networks for volumetric medical image segmentation [C] // 2016 Fourth International Conference on 3D Vision (3DV), October 25-28, 2016. Stanford, CA, USA. IEEE, 2016: 565-571.
- [20] Zhang Z L, Sabuncu M R. Generalized cross entropy loss for training deep neural networks with noisy labels [C] // *Advances in neural information processing systems*, Montréal, Quebec, CANADA, Dec 2-8, 2018. Trier: dblp: 2018: 8778-8788.
- [21] Kingma D P, Ba J. Adam: a method for stochastic optimization [J]. *Computer Science*, 2014.
- [22] Otsu N. A threshold selection method from gray-level histograms [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62-66.