

基于语义指导和自适应卷积的遥感云检测算法^{*}

徐梓川 龚晓峰

(四川大学电气工程学院 成都 610065)

摘要: 遥感卫星数据云检测分割是遥感影像处理中的重要环节,为了解决目前碎云薄云检测精度较低的问题,提出了一种采用基于高阶语义解码和自适应卷积编码的云检测方法。这种方法针对云团和碎云薄云之间的空间分布联系,提出了自适应卷积编码器来提取云团之间的关联信息。然后,使用高阶语义指导模块来解码语义特征,指导高分辨率的云掩码图生成。此外,这种方法还设计了一种动态联合损失函数,该损失函数通过动态计算样本中的漏检误检像素来构建权重,以引导神经网络关注碎云薄云特征,从而提高整体精度。实验结果表明,提出的算法在遥感图像上云分割能力可以达到96.5%的精确度和88.1%的交并比,可以很好地检测碎云薄云。

关键词: 遥感图像;云检测;注意力机制;损失函数;深度学习

中图分类号: TP751.1 文献标识码: A 国家标准学科分类代码: 510.4050

Cloud detection algorithm for remote sensing images based on semantic-guided and adaptive convolution

Xu Zichuan Gong Xiaofeng

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Cloud detection of remote sensing satellite data is a crucial component in the processing of remote sensing images. To address the issue of low accuracy in detecting broken-clouds and thin-clouds, this paper proposes a novel cloud detection method that utilizes high-order semantic-guided decoding and adaptive convolutional encoding. The method leverages the spatial distribution relationship between the main cloud and broken-clouds by introducing an adaptive convolutional encoder to extract correlation information between the main cloud clusters. A high-order semantic-guided decoding module is then utilized to decode semantic features, thus restoring high-resolution cloud mask images. Moreover, a dynamic fusion loss function is designed to calculate the weight by dynamically computing the missed and wrong pixels in the prediction, guiding the network to focus on broken-clouds and thin-clouds, features, thereby enhancing the overall accuracy. Experimental results demonstrate that the proposed algorithm achieves an accuracy of over 96.5% and an intersection over union of over 88.1%, effectively detecting broken-clouds and thin-clouds.

Keywords: remote sensing image; cloud detection; attention mechanism; loss function; deep learning

0 引言

遥感卫星的云检测主要依赖于图像的像素值和光谱特征分析,以识别和区分图像中的云层,并据此获取更准确的地表影像数据。遥感卫星云检测技术的主要目的是识别遥感图像中的云层,为研究者提供更准确的数据和信息。在军事领域,遥感卫星云检测技术可以用于目标探测、军事情报和战术决策等方面。通过使用遥感卫星云检测技术,军方可以实时监测敌方部队的活动情况,提高战术作战效率。

在民用领域,遥感卫星云检测技术主要用于气象预报、环境保护等领域。在气象预报方面,遥感卫星云检测技术可以帮助气象部门准确地预报天气情况,减轻自然灾害对人们生活的影响。在环境保护方面,遥感卫星云检测技术可以帮助监测大气污染、海洋污染等情况,从而及时采取措施保护环境。随着遥感技术的发展,研究者获取地面影像的能力得到显著提升。然而遥感卫星极易受到天气因素的影响,其中最明显的一个因素就是云层。云层的出现会直接阻挡视线,遮盖目标物体,使得其应用价值大大降低。云检

识别是使用遥感卫星图像的关键步骤,精确的云识别可以帮助研究者处理遥感卫星数据,大大提高研究者的处理效率。

在早期云检测方法中传统的图像处理方法占据着主导地位^[1]。云检测的方法大致分为两类:基于波段阈值和纹理信息的方法和基于深度学习的方法。基于波段阈值和纹理信息的方法需要专家特征来建立专家特征库。基于波段阈值和纹理信息的云检测方法最早是 Zhu 等^[2-3]提出的 FMask 算法,通过多波段的阈值来进行云检测。康一飞等^[4]利用高斯混合模型来自适应计算出遥感图像的灰度阈值,分割前景和后景来实现云检测分割。Otus 算法和 K-means 算法^[5]都是通过计算云层与相邻环境之间的相似性进行检测,难以区分复杂地表环境,对地表噪声敏感,鲁棒性差。这类方法数学原理清晰,计算成本低,但往往只利用图像的浅层信息,在处理复杂背景时,抗噪性能差,容易混淆。

基于深度学习的方法不需要人工的干预和介入,在检测过程中具备强大的鲁棒性,能够提取到遥感图像中的深层特征,而且随着深度的提高,越能够提取到抽象且具有全局性的信息,特征的表征能力就越强。其中最具代表的便是基于卷积神经网络的 CNN 模型^[6],Li 等^[7]提出 MSCFF 算法利用多尺度卷积进行特征融合提取,实现了高分辨率遥感图像的云分割。Guo 等^[8]提出的 CDNetV2 使用高级语义指导模块将高级语义送入解码模块,指导云语义的解码,在云雪分割的场景下取得了更加优异的效果。Wieland 等^[9]提出的 MS-UNet 基于改进的 UNet 模型结构,实现了对遥感图像中多种目标的高精度检测。目前基于 Transformer 的神经网络算法^[10-12]在多个图像处理的数据集中展现了出色的能力,Transformer 模型在多尺度信息建模以及全局上下文交互层面上具有显著的优点。Xie 等^[13]提出的 SegFormer 算法将 Transformer 引入到图像分割中,在多个图像分割数据集中取得了最好的性能。

对于云检测来说,其物理特性使得检测难以进行。例

如小面积的碎云在特征提取过程中极易被忽略,云雪共存的场景极易误识别雪为云。碎云漏检和高亮地表反射物误检易受样本数量所带来的影响。

为解决上述问题,本文构建了自适应多尺度卷积编码模块(adaptive multi-scale convolutional block, AMSCB),该模块利用 CNN 提取局部的纹理特征并建立有效的自注意力,再利用 Transformer 范式来将局部特征融入于全局进行交互,帮助模型更好的预测云的空间位置。当前的注意力解码模块^[14-17]往往都只用到了自注意力的部分,忽略了高阶语义信息的指导作用,针对这个问题本文提出了一种高阶语义指导解码模块 (high-order semantic-guided decoding module, HSDM),在解码过程中融合高级语义的特征信息,指导云掩膜的生成。最后利用动态联合损失函数(dyanmic fusion loss function, DFLoss)增加训练期间漏检误检碎云薄云数据的惩罚权重,减小训练样本所带来的影响,增强网络的碎云薄云检测能力。最后在数据集^[6]上实验,证明本文算法能有效提高遥感图像云检测的准确度。

1 模型算法介绍

1.1 总体架构

为了能够解决薄云、碎云检测难,下垫面影响大的问题,本文提出基于高阶语义指导解码和自适应卷积编码的云检测算法,其结构如图 1 所示。主要利用了自适应多尺度卷积编码模块来对原始遥感图像进行编码,然后利用高阶语义指导解码模块对特征图进行解码,生成最后的云掩码图。为了能够加强云团信息之间的联系,增强了信息交互尺度的通透性,利用自适应编码模块从不同的空间尺度对特征进行计算,适应不同大小的特征图,加强编码结构在信息提取过程中对空间位置信息的感知能力,提取出充足的语义信息。然后通过高阶语义指导解码模块将深层语义特征作为指导,来解码浅层的语义特征图得到更准确的预测。最后利用动态联合损失函数加强网络对薄云、碎云的感知能力,优化训练过程。

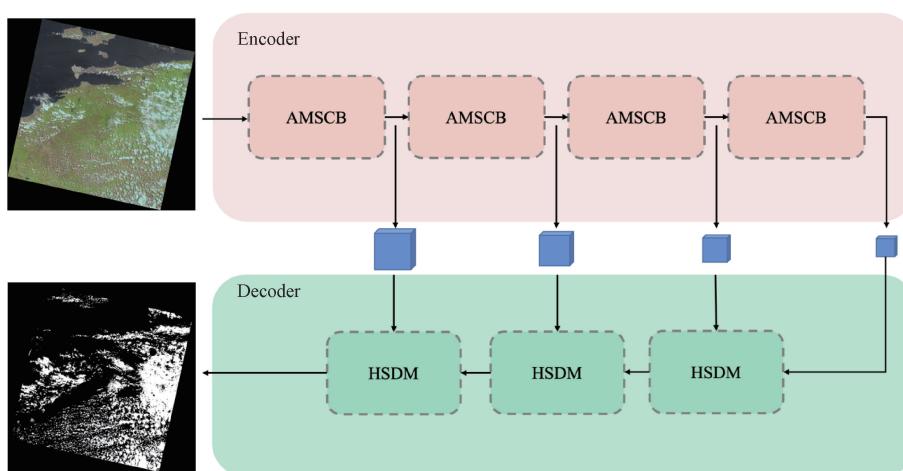


图 1 网络架构图

1.2 自适应多尺度卷积编码模块

在卫星遥感图像数据中,云团之间存在着位置关联,碎云、薄云与主要云团之间的空间分布存在着相应的关系,但 Transformer 自注意力模块在特征提取过程中难以处理局部特征,因此利用卷积提取像素局部关系特征的优点,设计了自适应多尺度卷积编码模块。该模块通过自适应大小的金字塔结构空洞卷积捕获不同尺度下的特征,增加神经网络感知空间位置感知能力,让网络可以有效的建模不同云块之间的像素关系。自适应多尺度卷积编码模块每一个阶段的结构如图 2 所示。

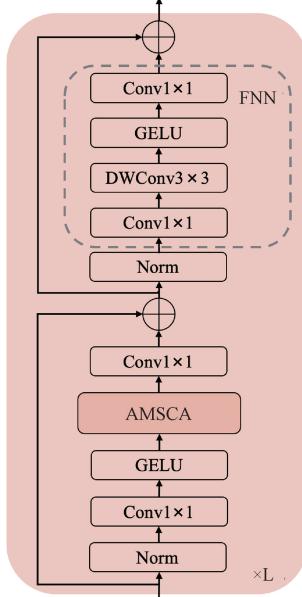


图 2 AMSCB 结构图

每一个阶段开始时,都先使用 Embedding 层来进行特征降维,其结构如图 3 所示。

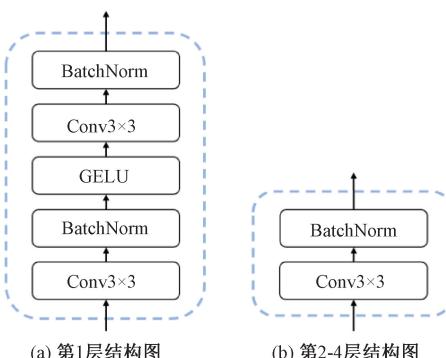


图 3 Embedding 网络结构图

在第一层结构如图 3(a)所示,使用两次卷积和 BN 逐步下采样,保证特征的细腻程度,并在两次操作中使用 GELU 激活函数来激活网络感知,最后得到下采样 1/4 的特征图。在其他层则利用图 3(b)的结构下采样 1/2 得到特征图。

在每一个模块中,先通过 Norm 层进行批归一化操作,然后送入自注意力计算,最后再经过 Norm 层归一化后送入前向反馈层(feed-forward network, FFN)中。前向反馈层如图 2 所示,由 3×3 大小的深度可分离卷积和 1×1 大小的卷积构成,其目的是对特征信息进行进一步提取。自注意力计算主要用到的是自适应多尺度卷积注意力 (adaptive multi-scale convolution attention, AMSCA),其结构如图 4 所示。

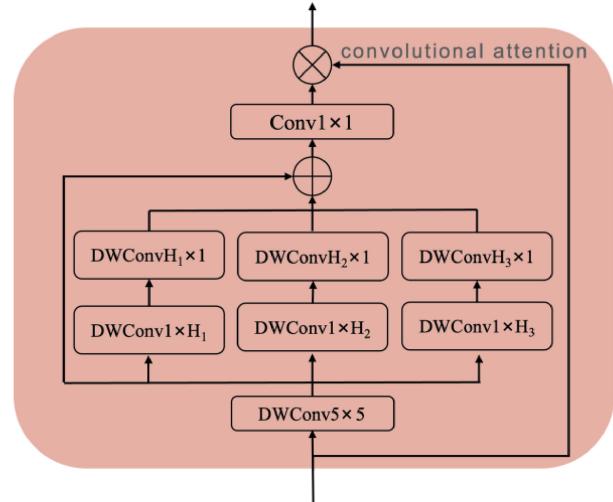


图 4 AMSCA 模块结构图

它首先利用一个大小为 5×5 的深度可分离卷积来提取特征信息,然后使用多路条形可分离卷积来计算不同尺度下上下文关系,生成自注意力权重,最后与原来的输入相乘,在数学上其表达式如式(1)和(2)所示。

$$\text{attention} = \text{Conv} \left(\sum_{i=0}^3 \text{Scale}_i (\text{DWConv}(F)) \right) \quad (1)$$

$$\text{Out} = \text{attention} \times F \quad (2)$$

其中, F 代表输入的特征, DWConv 代表深度可分离卷积, Conv 代表卷积, 其中 Scale_i 表示 i 个分支。在每个分支中选用了两个深度方向的长条形卷积来近似具有大核的卷积。针对大小为 (H, W) 的特征图每个分支的内核大小分别设置为 $H/2+1, H/4+1$ 和 $H/8+1$, 这样可以让主要云团和碎云之间可以找到充足的关联,而不过度关注更远距离的干扰。

主干网络共有四层,原输入图像分辨率大小 (H, W) , 在每一层中特征图分辨率为 $(H/4, W/4), (H/8, W/8), (H/16, W/16)$ 和 $(H/32, W/32)$, 其中各层重复次数为 3、3、12、3, 经过主干网络提取后各层通道数为 64, 128, 256, 512。

1.3 高阶语义指导解码模块

在传统的解码模块中,常常使用空间和通道注意力机制,这些方法利用自注意力机制来进行空间和通道上的信息建模,在同一维度上计算各特征的权重分配,但是这样的话会忽略掉高阶语义特征的指导作用。高阶语义中蕴

含着更高级的纹理空间语义指导信息, 在语义特征解码的过程中有着不可忽略的作用。

为了能够利用高阶语义特征中的指导作用, 加强特征图在空间和通道纬度上的信息交流, 本文设计了高阶语义指导模块。为了能够捕获空间和通道上的双重信息, 构建两种纬度上信息解码路径, 高阶语义指导解码模块需要输入高阶特征图 $X_{n+1} \in \mathbb{R}^{2C \times \frac{H}{2} \times \frac{W}{2}}$ 和低阶特征图 $X_n \in \mathbb{R}^{C \times H \times W}$, 其结构如图 5 所示。

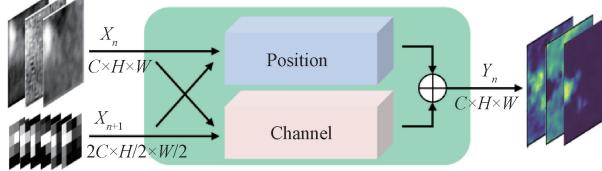


图 5 高阶语义指导解码模块示意图

在空间语义解码时, 主要利用高阶特征图在空间位置上的信息, 指导低阶特征图的位置信息解码, 其过程如图 6 所示。

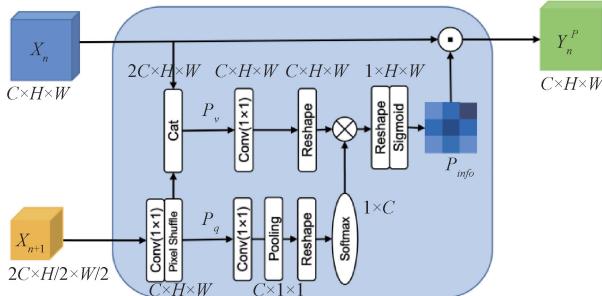


图 6 空间信息解码结构图

首先将 X_{n+1} 特征增强再重组采样成 $X_q \in \mathbb{R}^{C \times H \times W}$ 作为 Query, 其计算如式(3)所示。

$$X_q = F_s(\text{Conv}(X_{n+1})) \quad (3)$$

其中, F_s 表示采样重组。然后利用卷积提取 X_q 中的语义信息, 经过池化操作并转置后生成 $P_q \in \mathbb{R}^{1 \times C}$, 其计算如式(4)所示。

$$P_q = F_{sm}(\theta(F_p(\text{Conv}(X_q)))) \quad (4)$$

其中, F_p 表示全局池化操作, θ 表示 Reshape 操作, F_{sm} 表示 Softmax 函数。然后将 X_n 和 X_q 的特征信息拼接成 $X_v \in \mathbb{R}^{2C \times H \times W}$ 作为 Value, 其计算如式(5)所示。

$$X_v = \text{Cat}(X_q, X_n) \quad (5)$$

其中, Cat 表示拼接操作。利用卷积融合 X_v 中的语义信息并转置后生成 $P_v \in \mathbb{R}^{C \times HW}$, 其计算如式(6)所示。

$$P_v = \theta(\text{Conv}(X_v)) \quad (6)$$

将 P_v 与 P_q 相乘获得 $P_{qv} \in \mathbb{R}^{1 \times HW}$, 其计算如式(7)所示。

$$P_{qv} = P_q \times P_v \quad (7)$$

P_{qv} 经过转置和 Sigmoid 函数激活后即得到 $P_{info} \in$

$\mathbb{R}^{1 \times H \times W}$ 的空间特征信息, 其计算如式(8)所示。

$$P_{info} = \sigma(\theta(P_{qv})) \quad (8)$$

其中, σ 是 Sigmoid 函数。最后将 P_{info} 与低阶特征 X_n 相乘, 得到空间特征解码的特征图 Y_n^P 。

$$Y_n^P = X_n \cdot P_{info} \quad (9)$$

在空间信息解码分支中, 更加关注空间纹理信息, 指导神经网络在空间纹理位置上的注意力权重分布。

在通道语义解码时, 利用高阶特征图指导低阶特征图对各通道上的信息进行感知, 其过程如图 7 所示。

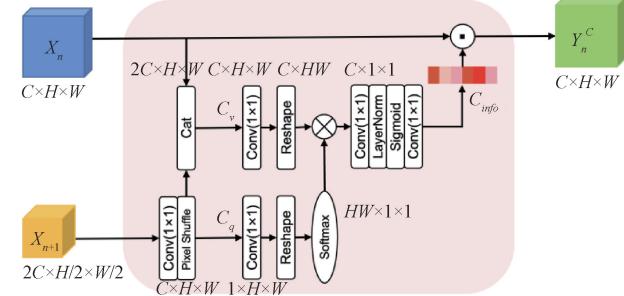


图 7 通道信息解码结构图

首先将 X_{n+1} 特征增强再重组采样成 $X_q \in \mathbb{R}^{C \times H \times W}$ 作为 Query, 其计算如式(10)所示。

$$X_q = F_s(\text{Conv}(X_{n+1})) \quad (10)$$

然后 X_q 经过卷积提取语义信息并转置后生成 $C_q \in \mathbb{R}^{HW \times 1 \times 1}$, 其计算如式(11)所示。

$$C_q = \theta(\text{Conv}(X_q)) \quad (11)$$

然后将 X_n 和 X_q 的特征信息拼接成 $X_v \in \mathbb{R}^{2C \times H \times W}$ 作为 Value, 其计算如式(12)所示。

$$X_v = \text{Cat}(X_q, X_n) \quad (12)$$

利用卷积融合 X_v 语义信息并转置后生成 $C_v \in \mathbb{R}^{C \times HW}$, 其计算如式(13)所示。

$$C_v = \theta(\text{Conv}(X_v)) \quad (13)$$

将 C_v 与 C_q 相乘获得 $C_{vq} \in \mathbb{R}^{C \times 1 \times 1}$, 其计算如式(14)所示。

$$C_{vq} = C_q \times C_v \quad (14)$$

经过转置和 Sigmoid 函数激活后即得到 $C_{info} \in \mathbb{R}^{C \times 1 \times 1}$ 通道特征信息, 其计算如式(15)所示。

$$C_{info} = \text{Conv}(\sigma(\text{LN}(\text{Conv}(C_{vq})))) \quad (15)$$

其中, LN 表示 LayerNorm, 最后 C_{info} 与低阶特征 X_n 相乘得到通道特征解码的特征图 Y_n^C 。

$$Y_n^C = X_n \cdot C_{info} \quad (16)$$

在通道语义解码分支中, 更加关注各通道中的语义信息, 指导神经网络在不同通道上的注意力权重分布。最后将空间特征和通道特征相互叠加, 获得在高阶语义指导下解码的特征信息。

$$Y_n = Y_n^P + Y_n^C \quad (17)$$

HSDM 在解码的过程中不断加强低阶语义的特征信

息,还能有效地融合不同纬度上的特征信息,使网络获得高效的解码能力,从而提升算法在云检测上的精度。

1.4 动态联合损失函数

遥感图像云检测中薄云、碎云检测一直是科研工作者解决的重点问题。薄云易受下垫面的干扰,影响薄云的精度检测。碎云数量少、且远离主要云团、易受周围像素的影响,导致检测精度的下降。因此设计了动态联合损失函数来解决这一问题,损失函数如式(18)所示。

$$L_{DF} = \sum_{i=1}^n [\omega_i (\alpha_{ce} L_{ce} + \alpha_{dice} L_{dice})] \quad (18)$$

L_{DF} 为动态联合损失函数, n 为批量的大小, L_{ce} 为交叉熵损失函数, L_{dice} 为骰子损失函数, ω_i 为动态权重。 L_{ce} 是分类问题上最经典的损失函数。遥感图像云分割任务中,往往单张训练图片样本中要么全为非云,要么全为云,在训练过程中存在着样本的分布不均的现象。 L_{dice} 被证明可以很好的解决样本分布不均衡的情况,通过将损失函数联合在一起使得神经网络在训练过程中更加平滑。考虑到薄云碎云部分在数据集中往往只有一小部分,在训练过程中极易被忽略,因此设计了动态感知漏检误检的权重 ω_i 来感知检测结果中误检和漏检的云在整体中的比重。当网络的检测结果中存在漏检和误检的像素时,网络检测错误的惩罚将按比例增加,加强网络对漏检误检的感知,动态权重 ω_i 的定义如式(19)所示。

$$\omega_i = \text{line} \left(\sigma \left(\frac{\sum_{h=0}^H \sum_{w=0}^W |P_{h,w} - L_{h,w}|}{H \times W} \right) \right) \quad (19)$$

$P_{h,w}$ 表示预测结果在 (h,w) 处的像素值, $L_{h,w}$ 表示标签在 (h,w) 处的像素值, $\sum_{h=0}^H \sum_{w=0}^W |P_{h,w} - L_{h,w}|$ 表示在这张图片中误检和漏检的像素总数,通过 Sigmoid 函数激活后线性映射得到权重 ω_i 。

2 实验与结果

2.1 实验数据

本文的实验数据是来自 Lantast-8 卫星的多光谱的影像数据^[6],包含 3 个可见波段和近红外线波段,在数据集中有林地、雪山和水体等多种下垫面,丰富的下垫面可以很好提升网络模型的实用性。

为了构建数据集本文舍弃了近红外波段,仅选用了可见光波段。数据集中包含 95 幅卫星遥感图像,将数据集进行划分训练集:验证集:测试集为 6:2:2。为了方便训练,本文将训练集和验证集中的图像分割成 384×384 的大小,最后训练集中包含 26 300 幅图像,验证集中包含 8 400 幅图像,最后在测试集中测试网络的云检测能力。

2.2 评价指标

单一精度评估指标难以准确完整的体现网络模型的实际性能,因此选用多种评估指标:交并比(intersection-over-union, IoU)、总体精度(overall accuracy, OA)和 F_1 -

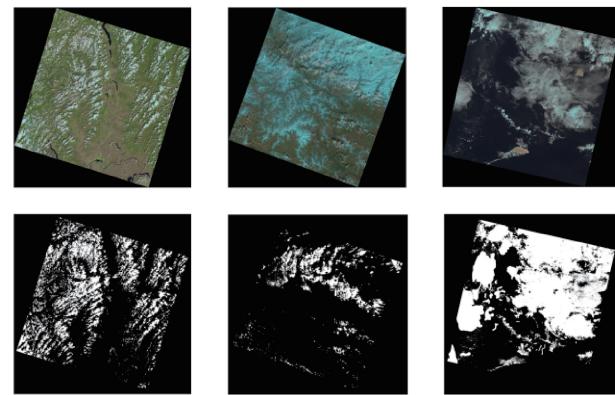


图 8 数据集

score 三种语义分割评价指标对云检测方法进行评价。

$$IoU = \frac{TP}{TP + FP + FN} \quad (20)$$

$$OA = \frac{TP + FN}{TP + FP + FN + TN} \quad (21)$$

$$F_1\text{-score} = \frac{2TP}{2TP + FP + FN} \quad (22)$$

在上述公式中 TP 代表原本为云且预测为云的数量, FP 代表原本为非云且预测为云的数量, FN 代表原为云且预测为非云的数量, TN 代表原本为非云且预测为非云的数量。 IoU 是类别为云的预测值和真实值之间的重合度,反映模型的预测能力, OA 为被预测正确的像素在整体中的占比,反映模型的整体准确率。 F_1 -score 是精确度和召回率的调和平均值,反映这两个指标的综合情况。通过上述精度评估指标,可以对云检测任务当中的分类精度进行衡量,判断出网络模型的综合性能。

2.3 实验环境配置

为了保证实验的可靠性,每个模型都使用相同的训练策略。批量大小设置为 4,网络优化方法使用 AdamW,该优化方法计算高效,内存使用很少。初始学习率为 0.000 01,学习率使用多步衰减策略,每两轮学习率衰减为前者的 0.9,让网络以较高的初始学习率进行快速地学习,并且在网络优化迭代的后期阶段逐步降低学习率。实验在 NVIDIA RTX 3080 GPU 12 G 上利用 Pytorch 实现。

2.4 实验结果

为验证神经网络在各项指标下的能力,选取 SegFormer、MS-UNet、CDNetV2 和 CloudNet^[7] 等算法进行对比。在进行比较时,首先将原始图像顺序切分成 384×384 的大小,再使用训练好的网络模型进行预测。

首先进行定量分析,在测试集中测试不同算法模型的评价指标,实验结果如表 1 所示。

从对比实验结果中可以发现,SegFormer 算法性能最低。本文算法相较于主流的 MS-UNet、CDNetV2 和 CloudNet 在各项指标均有一定程度的提升。 IoU 、 OA 和 F_1 -score 相较于 MS-UNet 分别提高了 8.41%, 2.65% 和

表 1 不同算法的性能对比

算法模型	IoU	OA	F_1 -score
SegFormer	0.774 6	0.922 1	0.873 0
MS-UNet	0.797 4	0.938 7	0.887 2
CDNetV2	0.833 4	0.942 9	0.909 1
CloudNet	0.785 0	0.964 8	0.879 2
本文算法	0.881 5	0.965 2	0.937 0

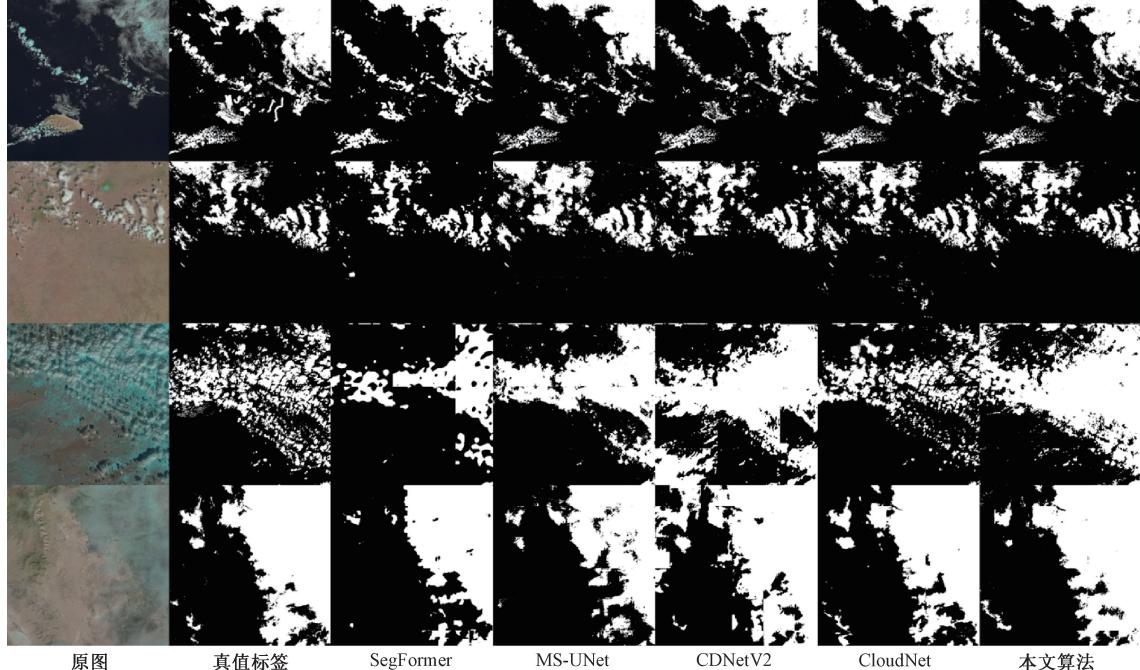


图 9 不同算法实验结果图

检的情况。CloudNet 能够很好的区分云和雪之间的特征差异,但是忽视了较多的碎云薄云,存在着漏检的情况。本文算法在不同场景下有很好的适应性,能够有效地区别云和下垫面,在实验结果中漏检和误检情况最少。此外本文算法还具备更好的碎云薄云检测能力,其效果如图 10 所示。

从图 10(a)可以观测到红色方框中心存在着一片碎云,SegFormer、MS-UNet、CDNetV2 和 CloudNet 均未能预测出这一片较小的云。本文算法准确的预测出其位置和形状,展现出本文算法在碎云检测中更能捕获云层之间的关联关系。原始图像在黄色方框中存在着一块薄云,可以看到在本文算法中薄云的纹理细节更加丰富,展现出本文算法能够更完整的识别出云层覆盖。本文算法在碎云检测上更加容易,并且能够更好的探测薄云。

为了能够更加有效的分析本文提出网络的易用性,计算对比网络所消耗的性能,对实验中的各算法进行测试。实验结果如表 2 所示。在表 2 中 Params 为参数量,代表了模型的参数大小,FPS 为单位时间处理的图像个数,代表了模型的计算速度。本文算法在 AMSCB 编码过程中

4.98%,相较于 CDNetV2 分别提高了 4.81%,2.23% 和 2.79%,相较于 CloudNet 分别提高了 9.65%,0.04% 和 5.78%。由此证明本文算法的有效性。对各算法在测试集中的表现进行定性分析,对比结果如图 9 所示。

在图 9 中的对比结果中可以发现,SegFormer 算法容易混淆地表的高亮物体,存在着明显的漏检情况。MS-UNet 和 CDNetV2 算法在检测过程中能够较好的识别下垫面与云之间的区别,但在云雪共存的场景下仍然存在误

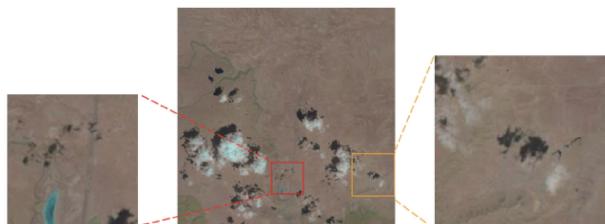
使用轻量化的深度可分离卷积来提取特征,它可以在保持性能的同时,大大降低算法的计算量。在 HSDM 解码过程中利用 1×1 大小的卷积核来充分融合特征信息,它能够以很低的成本对特征进行降维,减小算法的计算量。因此本文算法在适当的参数量下,取得了最快的计算速度。

2.5 消融实验

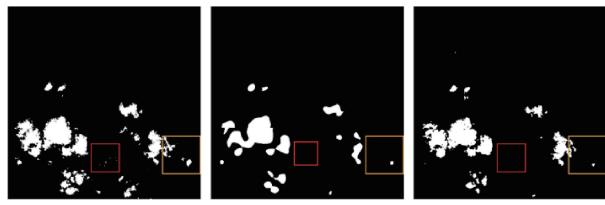
为了验证本文算法的各个模块在总体网络中发挥的作用,本文构造不同的消融实验网络。在 AblationNetV1 中去除 AMSCB,在 AblationNetV2 中去除 HSDM,在 AblationNetV3 中去除 DFLoss,其结构如表 3 所示。

在相同的实验环境下,对不同的模块进行消融实验,其表现结果如表 4 和如图 11 所示。

可以发现本文算法在各项指标中均具有更好的效果,表 4 第 1 行和图 11(a)表明缺失自适应多尺度卷积编码模块后,编码器难以提取云特征和掌握云分布规律,算法在整体精度上下降了约 2%。表 4 第 2 行和图 11(b)表明缺失高阶解码模块后,在解码过程中无法利用深层特征图中的高级语义来加强云的特征还原,算法在整体精度上下降了 2.5%。表 4 第 3 行和图 11(c)表明缺失动态联合损失



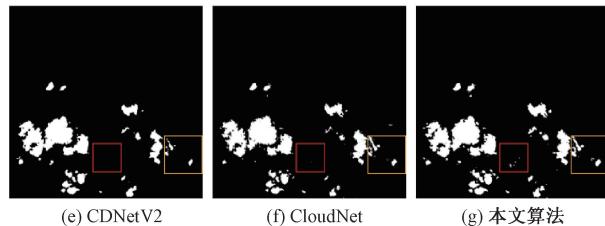
(a) 原图



(b) 真值标签

(c) SegFormer

(d) MS-UNet



(e) CDNetV2

(f) CloudNet

(g) 本文算法

图 10 不同算法实验结果局部效果图

表 2 不同算法的参数对比

算法模型	Params (M)	FPS(1/S)
SegFormer	43.68	38
MS-UNet	34.33	29
CDNetV2	67.08	40
CloudNet	36.47	31
本文算法	41.16	44

表 3 消融实验算法模型结构

算法模型	AMSCB	HSDM	DFLoss
AblationNetV1	×	√	√
AblationNetV2	√	×	√
AblationNetV3	√	√	×
本文算法	√	√	√

表 4 消融实验结果

算法模型	IoU	OA	F_1 -score
AblationNetV1	0.836 0	0.944 8	0.910 7
AblationNetV2	0.823 8	0.940 8	0.903 4
AblationNetV3	0.867 5	0.956 0	0.929 1
本文算法	0.881 5	0.965 2	0.937 0

函数后,在训练过程中感知碎云薄云能力降低,算法在整体精度上下降了 0.9%。表 4 第 4 行和图 10(g)表明,本文提出的各模块能够很好的相互配合,表现出优异的云分割能力。由此证明本文算法的各个模块均对于云检测分割

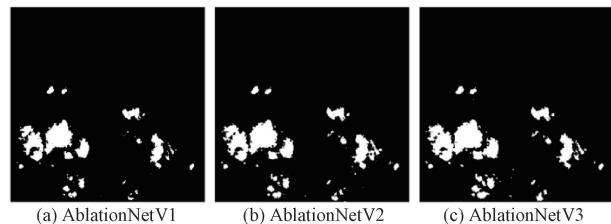
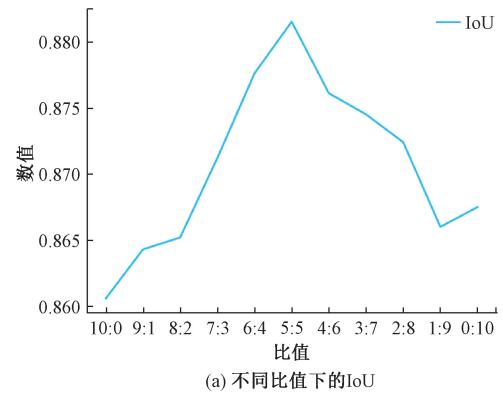


图 11 消融实验结果图

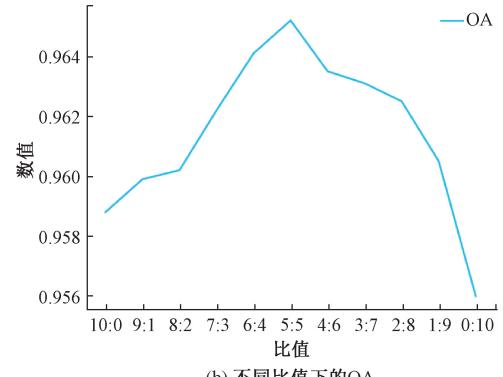
效果的提升是有效的。

2.6 不同比值损失函数的影响

在该小结探索交叉熵损失函数和骰子损失函数的比例对云检测能力的影响。取 $\alpha_{ce} : \alpha_{dice}$ 比值为 0 : 10, 1 : 9, 2 : 8, 3 : 7, 4 : 6, 5 : 5, 6 : 4, 7 : 3, 8 : 2, 9 : 1 和 10 : 0 进行实验,结果如图 12 所示。



(a) 不同比值下的IoU



(b) 不同比值下的OA

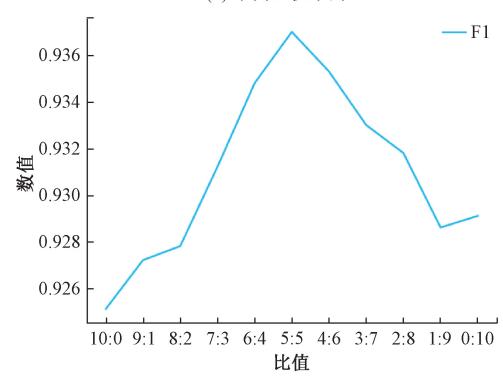


图 12 不同比值时网络的性能

当比值为 0:10 和 10:0 时评价指标中 OA、IoU 和 F_1 -score 相对于其他情况更低, 证明仅使用单独的损失函数时性能相对较差。当比值为 5:5 时, 评价指标 OA、IoU 和 F_1 -score 最大。因此可以得出结论在 $\alpha_{ce}:\alpha_{dice}$ 的比值取 5:5 时网络的综合性能更加优异, 故本文选取此比例作为损失函数的比例权值。

3 结 论

高效的云检测是对遥感卫星图像进行充分利用的前提。本文为解决云检测算法中薄云、碎云检测难的问题, 构造了自适应多尺度卷积编码模块, 提取云团的空间分布特征, 然后利用高阶语义指导解码模块, 有效的利用云特征图进行还原, 解码出高质量的云掩膜图, 最后利用动态联合损失函数帮助神经网络感知样本中的碎云和薄云, 在数据集上取得了 96.5%的整体精度和 88.1%的交并比, 并且能够更好的捕获遥感图像中的薄云、碎云, 提高卫星遥感图像分析利用的质量。

参考文献

- [1] 刘子力, 杨家俊, 王文静, 等. 遥感图像云检测方法综述[J]. 中国空间科学技术, 2023, 43(1): 1-17.
- [2] ZHU Z, WOODCOCK C E. Object-based cloud and cloud shadow detection in Landsat imagery [J]. Remote Sensing of Environment, 2012, 118: 83-94.
- [3] ZHU Z, WANG S, WOODCOCK C E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images [J]. Remote Sensing of Environment, 2015, 159: 269-277.
- [4] 康一飞, 潘励, 孙明伟, 等. 基于高斯混合模型法的国产高分辨率卫星影像云检测[J]. 武汉大学学报(信息科学版), 2017, 42(6): 782-788.
- [5] 李玉功. K-means 聚类算法的改进及其在彩色图像分割中的应用[D]. 合肥: 安徽大学, 2018.
- [6] MOHAJERANI S, SAEEDI P. Cloud-Net: An end-to-end cloud detection algorithm for Landsat 8 imagery [C]. IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2019: 1029-1032.
- [7] LI Z, SHEN H, 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.
- [8] GUO J, YANG J, YUE H, et al. CDnetV2: CNN-based cloud detection for remote sensing imagery with cloud-snow coexistence [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(1): 700-713.
- [9] WIELAND M, LI Y, MARTINIS S. Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network[J]. Remote Sensing of Environment, 2019, 230: 111203.
- [10] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[J]. ArXiv Preprint, 2020, ArXiv:2010.11929.
- [11] GUO M H, LU C Z, HOU Q, et al. Segnext: Rethinking convolutional attention design for semantic segmentation [J]. ArXiv Preprint, 2022, ArXiv: 2209.08575.
- [12] 丁志江, 李丹, 马志程, 等. 基于 Transformer 的车道线分割算法研究[J]. 电子测量与仪器学报, 2022, 36(10): 227-234.
- [13] XIE E, WANG W, YU Z, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30.
- [15] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7794-7803.
- [16] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.
- [17] LIU H, LIU F, FAN X, et al. Polarized self-attention: Towards high-quality pixel-wise regression[J]. ArXiv Preprint, 2021, ArXiv:2107.00782.

作者简介

徐梓川(通信作者), 硕士研究生, 主要从事深度学习, 遥感图像处理研究。

E-mail:765026654@qq.com

龚晓峰, 教授, 主要从事检测技术与自动化装置等研究