

## 基于 RDA-Net 模型的遥感影像云与云阴影检测方法

张晨<sup>1</sup>, 张秀再<sup>1,2\*</sup>, 杨昌军<sup>3</sup><sup>1</sup>南京信息工程大学电子与信息工程学院, 江苏 南京 210044;<sup>2</sup>南京信息工程大学江苏省大气环境与装备技术协同创新中心, 江苏 南京 210044;<sup>3</sup>中国气象局中国遥感卫星辐射测量和定标重点开放实验室国家卫星气象中心, 北京 100081

**摘要** 针对目前大多数云与云阴影检测方法容易产生误检、边缘细节丢失严重以及检测不够精确的问题, 提出一种基于双注意力卷积神经网络模型(RDA-Net)的遥感影像云与云阴影检测方法。模型中引入双注意力模块可以有效捕获全局特征的依赖关系, 使用递归残差模块可以避免深层网络出现退化, 改进空洞空间金字塔池化模块在不改变特征图尺寸的前提下可以提取图像的多尺度特征。首先对遥感影像数据集进行预处理并制作对应的标签, 然后利用高分一号 WFV 遥感影像数据集进行训练和测试。实验结果表明, 所提方法有效提高云与云阴影的检测精度, 在复杂条件下仍能获得较好的云与云阴影的边缘细节。

**关键词** 遥感; 双注意力; 云与云阴影检测; 递归残差; 改进空洞空间金字塔池化

中图分类号 TP751.1; P426.5

文献标志码 A

doi: 10.3788/LOP202158.2028005

## Remote Sensing Image Cloud and Cloud Shadow Detection Method Based on RDA-Net Model

Zhang Chen<sup>1</sup>, Zhang Xiuzai<sup>1,2\*</sup>, Yang Changjun<sup>3</sup><sup>1</sup> School of Electronics and Information, Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, China;<sup>2</sup> Jiangsu Province Atmospheric Environment and Equipment Technology Collaborative Innovation Center, Nanjing University of Information Science & Technology, Nanjing, Jiangsu 210044, China;<sup>3</sup> National Satellite Meteorological Center, Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, China Meteorological Administration, Beijing 100081, China

**Abstract** Aiming at the problems that most of the current cloud and cloud shadow detection methods are prone to misdetection, serious edge detail loss and insufficient detection accuracy, a remote sensing image cloud and cloud shadow detection method based on the dual attention convolutional neural network model (RDA-Net) is proposed. In the model, the dual attention module is introduced to can effectively capture the dependence of global features, the recursive residual module is used to avoid degradation of the deep network, and the improved atrous spatial pyramid pooling module can extract multi-scale features without changing the size of the feature map. First, the remote sensing image dataset is preprocessed and made the corresponding labels, and then the Gaofen-1 WFV remote sensing image dataset is used for training and testing. Experimental results show that the proposed method can effectively improve the detection accuracy of cloud and cloud shadow, and can still obtain better edge details of cloud and cloud shadow under complex conditions.

**Key words** remote sensing; dual attention; cloud and cloud shadow detection; recursive residual; improved atrous spatial pyramid pooling

**OCIS codes** 280.4788; 010.1615; 100.3008; 100.4996

收稿日期: 2020-11-24; 修回日期: 2020-12-18; 录用日期: 2021-01-02

基金项目: 国家自然科学基金青年科学基金(11504176, 61601230)、江苏省高校自然科学基金项目(13KJA510001)、江苏省自然科学基金青年基金(BK20141004)

通信作者: \*zxzhering@163.com

# 1 引言

随着遥感技术的快速发展,遥感影像在气象探测、水文监测和军事侦察等领域发挥着越来越重要的作用。地球表面有 67% 的面积被云覆盖<sup>[1]</sup>,云与云阴影的存在限制了基于卫星遥感影像的地物信息提取能力,严重影响了遥感影像的质量和遥感数据的可用性。因此,云与云阴影的检测是遥感影像处理过程中最关键的一个步骤,同时也为气象、农业和自然灾害预测等领域提供有力的科技保障<sup>[2]</sup>。

目前,云检测方法可以大致分为基于阈值的检测方法、基于纹理谱及其空间特征的检测方法和基于机器学习的检测方法。针对某种特定的遥感器,基于阈值的检测方法可以展现出较好的性能,但该方法对阈值的敏感程度较高,而且忽略了较多的空间信息<sup>[3]</sup>,当部分像元被云覆盖时,容易产生误检。Rossow 等<sup>[4]</sup>假设观测辐射值只来自云和晴空,将可见光波段 6 和红外波段 11 的辐射值进行比较以测试像元的辐射值,当像元的辐射值大于晴空时,则说明该像元为云像元,但该方法容易产生误判,导致检测结果不准确。基于纹理谱及其空间特征的检测方法的关键在于选择纹理特征,所选择的特征要精准反映云和下垫面的差异,该方法具有一定的检测效果,但遥感影像中云的种类较多,所以仅利用纹理特征进行检测,效果提升有限。李超炜等<sup>[5]</sup>对多光谱影像中云层的亮度、纹理和频率三个特征展开研究,统计云层与下垫面在各个方面的特征差异,根据分析结果利用支持向量机对特征进行分类,从而实现云检测,但该方法的复杂性较高,容易产生误检。

近年来,随着深度学习在语义分割领域的深入研究,基于深度学习的遥感影像云与云阴影检测方法的应用日益广泛<sup>[6]</sup>。陈洋等<sup>[7]</sup>采用主成分分析对卷积神经网络进行无监督的预训练,通过该网络来获得影像特征,将获得的特征输入支持向量机中进行分类,从而完成对国产资源三号卫星遥感影像的云检测。Dröner 等<sup>[8]</sup>使用基于 U-Net 的网络结构进行遥感影像云检测。Xu 等<sup>[9]</sup>提出了一种新的光学遥感影像云与云阴影检测框架,即 DeepMask 框架,利用残差网络 ResNet 生成像素级的云与云阴影掩模。相比于传统的机器学习方法,上述方法可获得更好的云检测效果,但对于云边缘检测而言能力较弱,导致细节丢失较多。

针对现有方法对云与云阴影检测不精确以及细

节丢失过多的问题,提出一种双注意力<sup>[10]</sup>卷积神经网络模型 RDA-Net,使用该模型对高分一号 WFV (Wide Field of View) 遥感影像进行像素级云与云阴影检测。该模型是根据编解码双路卷积神经网络改进而来的,主要由编码器、解码器和下边路组成。使用递归残差模块<sup>[11]</sup>、位置注意力模块<sup>[10]</sup>、通道注意力模块<sup>[10]</sup>、改进空洞空间金字塔池化模块<sup>[12]</sup>和多尺度特征融合技术<sup>[13]</sup>实现对云与云阴影的精确检测,并获得精准的边缘细节。

## 2 原理与方法

### 2.1 RDA-Net 模型

语义分割是一项对视觉图像进行像素级精准分类的任务,将图像中每个像素分为不同的语义类别,同一类别的像素分成同一类别标签,即相同类别的物体颜色标签相同。RDA-Net 是根据编解码双路卷积神经网络改进而来的,除了编码器和解码器这两条路径以外,还增加了一条下边路来进行网络模型的深度监督。同时引入多尺度特征融合技术将不同尺度的特征进行融合,用来提取多尺度和高层次的空间和语义信息。由于云层的薄厚程度和形状差异较大,而且下垫面类型复杂多样,特别是在雪和霾的环境下,大多数检测方法的检测难度较大。因此,提出一种基于 RDA-Net 模型的遥感影像云与云阴影检测方法。RDA-Net 模型主要由递归残差模块 (ResBlock)、位置注意力模块 (PAM)、通道注意力模块 (CAM)、改进空洞空间金字塔池化 (R-ASPP) 模块、卷积层、上采样层和特征融合层组成,各部分的作用如下。

1) 递归残差模块:随着网络模型深度的加深,网络的性能会随之增强,为此可以拟合更加复杂的特征输入,但同时也会出现某些问题而导致浅层网络的学习能力下降,从而限制深层网络的学习。为了学习高度复杂的特征,本研究引入 ResBlock,如图 1(a) 所示,其中  $\oplus$  为特征相加符号。每个 ResBlock 包含三个递归残差单元 (ResUnit)<sup>[11]</sup>,如图 1(b) 所示。递归残差单元由批量归一化层、激活函数 ReLU 和卷积层组成。为了使低层特征与高层特征相融合,将递归残差模块中的每个递归残差单元都与第一个卷积层进行跳跃连接以降低网络的训练难度。残差单元  $P_t$  可表示为

$$P_t = F(P_{t-1}, \sigma) + P_0, \quad (1)$$

式中:  $t=1, 2, 3, \dots, T_0$ , 其中  $T_0$  表示递归残差模块中递归残差单元的数目;  $P_{t-1}$  和  $P_t$  分别表示第  $t$

个递归残差单元的输入和输出;  $P_0$  表示递归残差模块中第一个卷积层的输出结果;  $F(\cdot)$  表示残差函

数;  $F(P_{t-1}, \sigma)$  表示学习到的残差映射, 其中  $\sigma$  为权值。

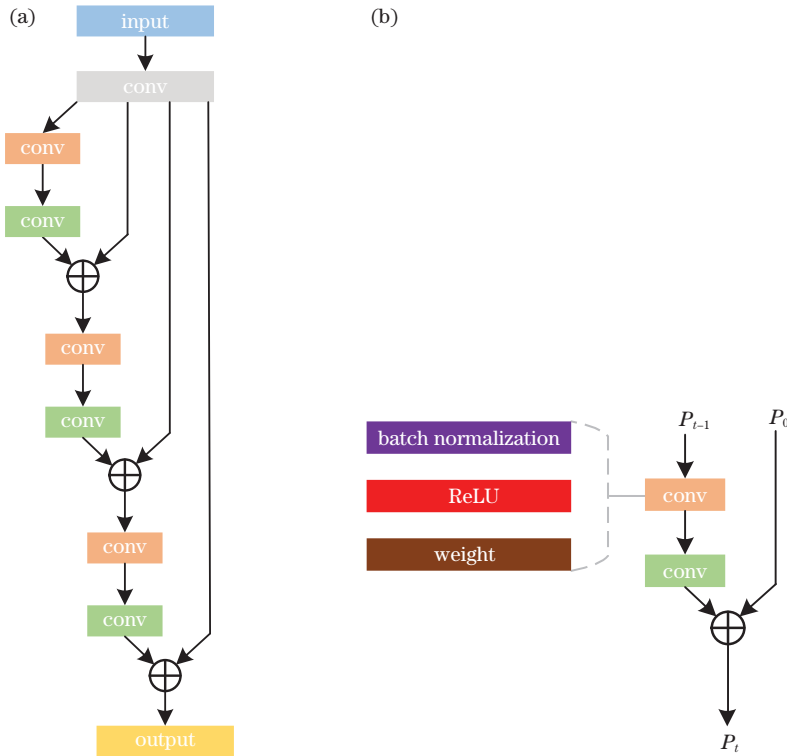


图 1 递归残差模块与递归残差单元的结构。(a) 递归残差模块; (b) 递归残差单元

Fig. 1 Structure of ResBlock and ResUnit. (a) ResBlock; (b) ResUnit

2) 位置注意力模块: 引入 PAM 的目的是捕获像素间更丰富的上下文信息, 将这些上下文信息编码到局部特征中以凸显关键特征的位置。PAM 的结构如图 2 所示, 其中  $\otimes$  为矩阵相乘符号。将获取的特征图  $A \in \mathbb{R}^{C_0 \times H_0 \times W_0}$  应用在大小为  $1 \times 1$  的卷积层上, 可以得到特征图  $B \in \mathbb{R}^{C_0 \times H_0 \times W_0}$ 、 $C \in \mathbb{R}^{C_0 \times H_0 \times W_0}$  和  $D \in \mathbb{R}^{C_0 \times H_0 \times W_0}$ , 其中  $C_0$  为通道数,  $H_0$  为高,  $W_0$  为宽。将  $B$ 、 $C$  和  $D$  的维度重塑为  $\mathbb{R}^{C_0 \times N_0}$ , 其中  $N_0$  为像素数量,  $N_0 = H_0 \times W_0$ ; 将  $C$

的转置矩阵与  $B$  进行矩阵相乘, 并将运算结果应用在 Softmax 层上, 可以得到空间注意力图谱  $Z \in \mathbb{R}^{N_0 \times N_0}$ 。空间注意力图谱  $Z$  可表示为

$$|Z| = z_{ji} = \frac{\exp(B_i \cdot C_j)}{\sum_{i=1}^{N_0} \exp(B_i \cdot C_j)}, \quad (2)$$

式中:  $z_{ji}$  表示特征图中第  $i$  个像素位置与第  $j$  个像素位置之间的相关性, 相关性越大, 特征表示越相似, 反之亦然。

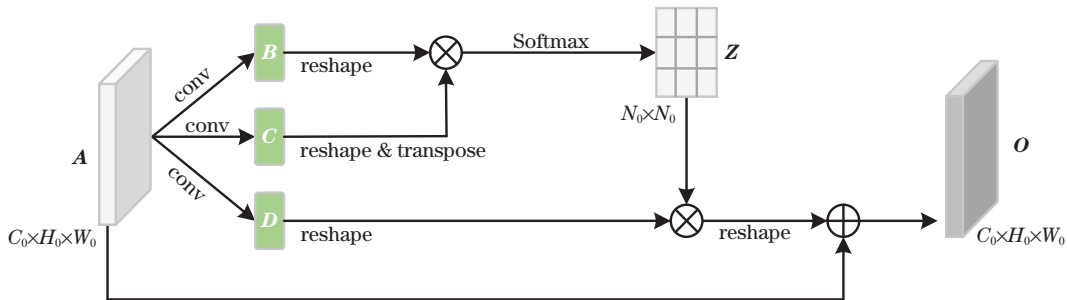


图 2 PAM 的结构

Fig. 2 Structure of PAM

同时, 将特征图  $D$  与空间注意力图谱  $Z$  的转置矩阵进行矩阵相乘, 并将计算结果的维度重塑为

$\mathbb{R}^{C_0 \times H_0 \times W_0}$  可以得到特征  $M \in \mathbb{R}^{C_0 \times H_0 \times W_0}$ 。特征  $M$  可表示为

$$|\mathbf{M}| = m_j = \sum_{i=1}^{N_0} (z_{ji} \cdot D_i). \quad (3)$$

将特征  $\mathbf{M}$  乘以位置尺度系数  $\delta$  并与特征图  $\mathbf{A}$  进行元素求和,可以得到输出  $\mathbf{O} \in \mathbb{R}^{C_0 \times H_0 \times W_0}$ 。 $\mathbf{O}$  表示所有位置特征与原始特征的加权和,包含了全局上下文信息,可选择性地聚合上下文以凸显关键特征的位置,从而提高云与云阴影的检测精度。输出  $\mathbf{O}$  可表示为

$$|\mathbf{O}| = o_j = \delta m_j + A_j. \quad (4)$$

$\delta$  的初始值为 0,随着训练的进行输出  $\mathbf{O}$  会逐步分配更多的权重。

3) 通道注意力模块:引入 CAM 的目的是建

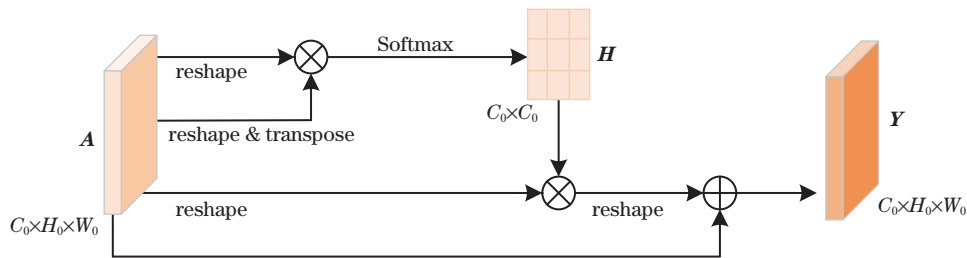


图 3 CAM 的结构

Fig. 3 Structure of CAM

同时,将特征图  $\mathbf{A}$  与通道注意力图谱  $\mathbf{H}$  的转置矩阵进行矩阵相乘,并将运算结果的维度重塑为  $\mathbb{R}^{C_0 \times H_0 \times W_0}$  可以得到特征  $\mathbf{N} \in \mathbb{R}^{C_0 \times H_0 \times W_0}$ 。特征  $\mathbf{N}$  可表示为

$$|\mathbf{N}| = n_v = \sum_{u=1}^{C_0} (h_{vu} \cdot A_u). \quad (6)$$

将特征  $\mathbf{N}$  乘以通道尺度系数  $\varphi$  并与特征图  $\mathbf{A}$  进行元素求和,可以得到输出  $\mathbf{Y} \in \mathbb{R}^{C_0 \times H_0 \times W_0}$ 。 $\mathbf{Y}$  为所有通道特征与原始特征的加权和,其建立了特征映射之间的长期语义依赖关系,有助于提高特定语义的特征表示,从而提高云与云阴影检测结果的完整性。输出  $\mathbf{Y}$  可表示为

$$|\mathbf{Y}| = y_v = \varphi n_v + A_v. \quad (7)$$

$\varphi$  的初始值为 0,随着训练的进行输出  $\mathbf{Y}$  会逐步分配更多的权重。

4) 空洞卷积与改进空洞空间金字塔池化模块:

在卷积神经网络中,一般通过扩大感受野来增强语义信息的提取能力。扩大感受野的作用主要是增大卷积核的尺寸和增加网络的层数。然而,随着卷积核尺寸的增大和网络层数的加深,空间信息的提取能力减弱,计算的复杂度增大,训练时长增加<sup>[14]</sup>。因此,本研究引入不同扩张率的空洞卷积对特征图

立通道之间的相互依赖性,改善特定语义的特征表示。CAM 的结构如图 3 所示,将获取的特征图  $\mathbf{A} \in \mathbb{R}^{C_0 \times H_0 \times W_0}$  的维度重塑为  $\mathbb{R}^{C_0 \times N_0}$ ,将  $\mathbf{A}$  与  $\mathbf{A}$  的转置矩阵进行矩阵相乘,并将运算结果应用在 Softmax 层上可以得到通道注意力图谱  $\mathbf{H} \in \mathbb{R}^{C_0 \times C_0}$ 。通道注意力图谱  $\mathbf{H}$  可表示为

$$|\mathbf{H}| = h_{uv} = \frac{\exp(A_u \cdot A_v)}{\sum_{u=1}^{C_0} \exp(A_u \cdot A_v)}, \quad (5)$$

式中: $h_{uv}$  表示特征图中第  $u$  个通道与第  $v$  个通道之间的相关性。

进行特征提取。

在空洞卷积改变扩张率来增大或者减小感受野的同时可保持卷积核和特征图的尺寸不变。一个扩张率为 2 和卷积核尺寸为  $3 \times 3$  的空洞卷积,其感受野的大小与卷积核尺寸为  $5 \times 5$  的卷积感受野相同,但前者参数数量更少,同时保留了遥感影像的细节信息。空洞卷积对应的感受野  $\eta$  可表示为

$$\eta = (k - 1) \times r + 1, \quad (8)$$

式中: $r$  表示扩张率; $k$  表示卷积核尺寸。当  $r = 1$  时,该卷积为普通卷积。

为了提高遥感影像解析的性能、更好地捕获输入特征的上下文多尺度信息以及提高云与云阴影的检测精度,本研究设计一种 R-ASPP 模块,结构如图 4 所示。首先通过逐步增大扩张率来获得更全面的背景信息;然后通过减小扩张率来捕获局部信息以增强细节特征的提取能力。该模块包含 5 个独立分支,其中三个空洞卷积分支由 4 个不同扩张率的空洞卷积层组成,卷积核的尺寸均为  $3 \times 3$ ,扩张率分别为  $(3, 4, 6, 2)$ 、 $(4, 6, 12, 3)$ 、 $(6, 18, 12, 4)$ ;一个卷积分支由一个卷积核尺寸为  $1 \times 1$  的卷积层组成;一个池化分支由平均池化层、 $1 \times 1$  卷积层和上采样层组成;最后将 5 个分支所输出的特征图进行特征融合。

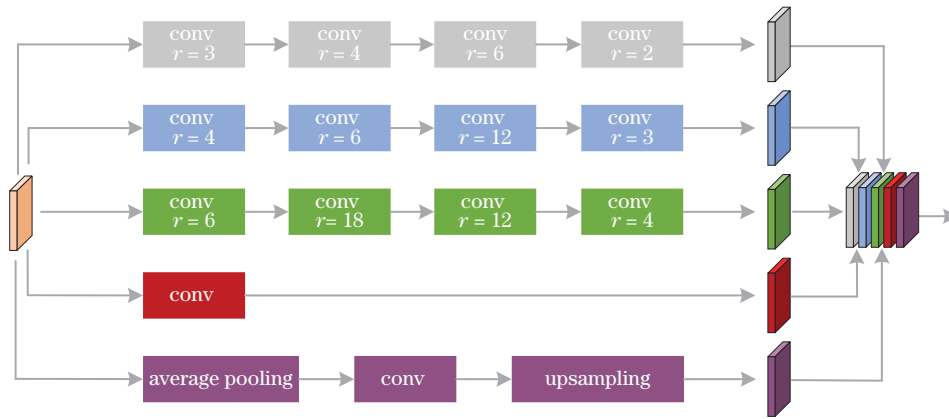


图 4 R-ASPP 的结构

Fig. 4 Structure of R-ASPP

RDA-Net 的结构如图 5 所示, 其由编码器路径、解码器路径和一条下边路组成。该模型共有 8 个 ResBlock、5 个双注意力模块 (DaBlock) 和一个 R-ASPP 模块, 包含步长为 1、不改变特征图尺寸的卷积层, 以及步长为 2、改变特征图尺寸的卷

积层。ResBlock 分为 4 种, 分别为 ResBlock 1、ResBlock 2、ResBlock 3 和 ResBlock 4, 每个 ResBlock 均使用步长为 1 和卷积核尺寸均为  $3 \times 3$  的卷积层, 卷积通道数分别设为 16、32、64 和 128。

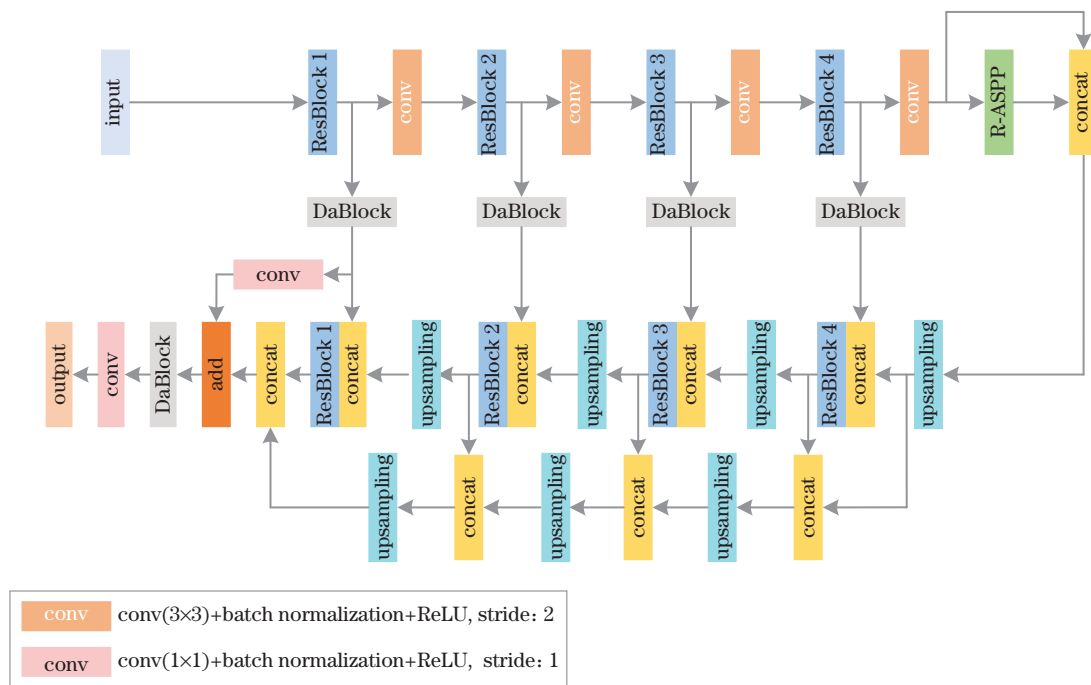


图 5 RDA-Net 的结构

Fig. 5 Structure of RDA-Net

编码器路径使用步长为 2 和卷积核尺寸为  $3 \times 3$  的卷积层来代替池化层进行下采样运算, 可以将特征图的尺寸减半。为了获得上下文多尺度信息, 将最后一个下采样运算的输出应用在 R-ASPP 模块上。为了获得全局特征的依赖关系, 将递归残差模块的输出应用在 DaBlock 上。解码器路径采用双线性插值法对特征图进行上采样运算, 可以恢复特征图的尺寸。在最后一个卷积层后加入比率为 0.5

的 Dropout 操作, 可以防止网络出现过拟合的情况。

下边路由上采样层和特征融合层组成, 其输出的特征可以与解码器路径的输出特征进行融合。将最后一个双注意力模块的输出经过  $1 \times 1$  的卷积层后得到的特征输入 Softmax 损失层中, 使用 Softmax 函数来判断特征中每个像元的类别概率, 从而实现端到端的像素级语义分割, 可以将遥感影像中的像元分为云像元、云阴影像元和下垫

面像元。同时计算损失值,本研究采用的损失函数  $L(\theta)$  可表示为

$$L(\theta) = -\frac{1}{R} \left\{ \sum_{r=1}^R \sum_{q=1}^Q 1\{f_r = q\} \text{lb} \left[ \frac{\exp(\theta_q^T \mathbf{g}_r)}{\sum_{q=1}^Q \exp(\theta_q^T \mathbf{g}_r)} \right] \right\}, \quad (9)$$

式中: $R$  表示每张图像中像素点的个数; $Q$  表示类别数,本文研究的是三分类语义分割,所以  $Q = 3$ ;  $f_r$  表示类别标签; $\theta$  表示模型参数, $\theta_q^T$  表示  $\theta_q$  的转置; $\mathbf{g}_r$  表示输入图像的像素观测向量  $\mathbf{G} = [\mathbf{g}_1 \ \mathbf{g}_2 \ \mathbf{g}_3 \ \cdots \ \mathbf{g}_R]$  中的元素; $1\{\cdot\}$  表示显函数。

### 2.2 U-Net 的改进模型

U-Net<sup>[15]</sup> 被应用在生物医学图像分割领域,网络结构由对称的编码器-解码器组成,其对较小的数据集可实现高精度的语义分割。为了验证 RDA-

Net 模型对云与云阴影检测的有效性,使用残差模块来代替 U-Net 编码器中的卷积层<sup>[16]</sup>,即将 U-Net 改进为 RU-Net, RU-Net 的结构如图 6 所示。模型的编码器部分由残差模块和下采样层组成,使用步长为 2 和卷积核尺寸为  $3 \times 3$  的卷积层进行下采样,残差模块包含步长为 1 和卷积核尺寸为  $3 \times 3$  的卷积层及跳跃连接,密集的跳跃连接在卷积层间进行级联可以更好地融合低层特征与高层特征,卷积层后加入批量归一化层和激活函数 ReLU。解码器部分由残差模块和上采样层组成,采用双线性插值法进行上采样。编码器与解码器之间通过矩阵级联的方式进行特征融合,可以使网络融合更多的低层特征。最后一个残差模块输出的特征经过  $1 \times 1$  卷积层后,利用 Softmax 函数来判断特征中每个像元的类别概率。

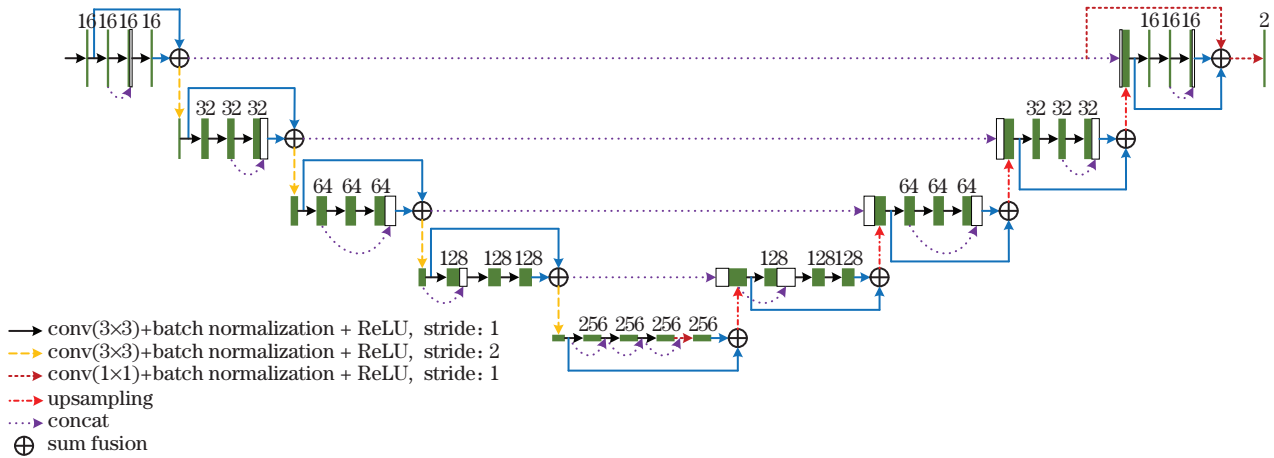


图 6 RU-Net 的结构

Fig. 6 Structure of RU-Net

## 3 实验设置与评价

### 3.1 实验环境

本文提出的方法是在 Windows 10 64 位操作系统中使用 Python 进行代码编写,在 Keras 深度学习框架中<sup>[17]</sup> 设计模型来实现。实验的硬件环境为 Intel Core i7 9700K 处理器(CPU),32 G 的内存, NVIDIA GeForce RTX 2080 Ti(11 G)图形处理器(GPU)。

### 3.2 实验数据的选取及预处理

#### 3.2.1 实验数据的选取

选取的遥感影像来自高分一号 WFV 云与云阴影数据集<sup>[18]</sup>。WFV 是高分一号卫星中最关键的组件之一,其包含 4 个空间分辨率为 16 m 和时间分辨率为 4 天的集成相机,每个相机有 4 个多光谱波段,

观测幅宽达到 800 km。该数据集包含 108 张遥感影像,基本涵盖了不同类型的下垫面。根据遥感影像制作相应的标签,标签由专家手动标记,将其转换为灰度图像,使用灰度值分别为 0(黑色)、128(灰色)和 255(白色)来代表下垫面像元、云阴影像元和云像元。为了提高语义分割的效率,将数据集中遥感影像及标签的大小统一转为  $1024 \text{ pixel} \times 1024 \text{ pixel}$ ,从中选取 87 张作为训练集与验证集,21 张作为测试集。实验中使用的遥感影像块及其对应的标签如图 7 所示。

#### 3.2.2 实验数据的预处理

通常在模型训练阶段需要大量的参数,若使模型在这些参数下正确工作则需要大量的数据进行训练,而实际情况中遥感影像数据集包含的数据量较少,而且尺寸较大,因此需要对遥感影像及其标签进行旋

转、变换亮度、注入噪声、变换饱和度和变换颜色和裁剪等数据增强处理。选取一张数据增强后的高分一号 WFV 遥感影像,数据增强后的效果如图 8 所示。

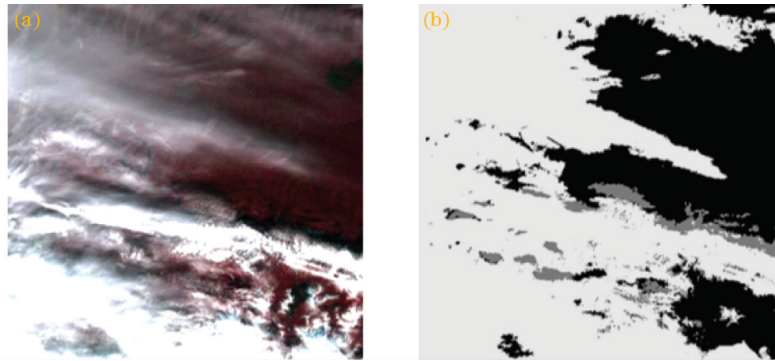


图 7 实验数据集。(a)遥感影像块;(b)标签

Fig. 7 Experimental dataset. (a) Remote sensing image block; (b) label

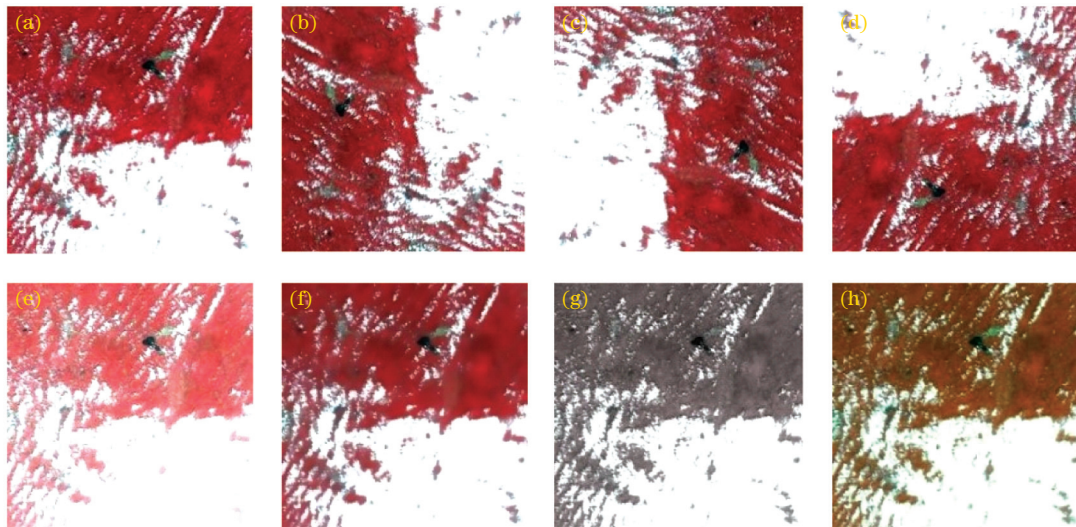


图 8 实验数据增强的效果。(a)原图;(b)垂直旋转;(c)水平旋转;(d)水平垂直旋转;  
(e)变换亮度;(f)注入噪声;(g)变换饱和度;(h)变换颜色

Fig. 8 Enhanced effects of experimental data. (a) Original image; (b) vertical rotation; (c) horizontal rotation; (d) horizontal and vertical rotation; (e) transformation of brightness; (f) injection of noise; (g) transformation of saturation; (h) transformation of color

### 3.3 实验细节

对于训练阶段,训练开始前需对输入图像进行归一化处理,可以加快模型收敛和优化的速度。实验使用 Adam 优化器对参数进行优化,学习率  $\alpha = 0.0001$ ,衰减率  $\beta_1$  分别为 0.900 和 0.999,批量大小取经验值,设为 8。当迭代 127 次附近时,损失函数值达到较低值并保持小幅浮动。

在测试阶段直接调用保存好的最优模型,对测试集中尺寸为  $1024 \text{ pixel} \times 1024 \text{ pixel}$  的遥感影像进行像素级的云与云阴影检测。为了验证 RDA-Net 模型的有效性,将其与其他模型进行比较。云与云阴影检测的实验流程如图 9 所示。

将增强后的高分一号 WFV 遥感影像及其标签裁剪成 12528 张尺寸为  $256 \text{ pixel} \times 256 \text{ pixel}$  的图像,其中 10100 张作为训练集,2428 张作为验证集。

此外,加入消融实验来探究 R-ASPP 模块和 DaBlock 对 RDA-Net 模型云检测结果的影响。实验包括移除其中一种模块以及同时移除所有模块,实验过程中使用同一数据集并设置相同的实验参数。将总体精度  $O_{\text{Accuracy}}$  作为评价指标,其指被正确检测的类别像元数与总的类别像元数的比值,可表示为

$$O_{\text{Accuracy}} = \frac{V}{V + V'}, \quad (10)$$

式中: $V$  表示每个类别像元被正确检测的范围; $V'$  表示每个类别像元被误检的范围。总体精度与迭代次数的关系如图 10 所示。

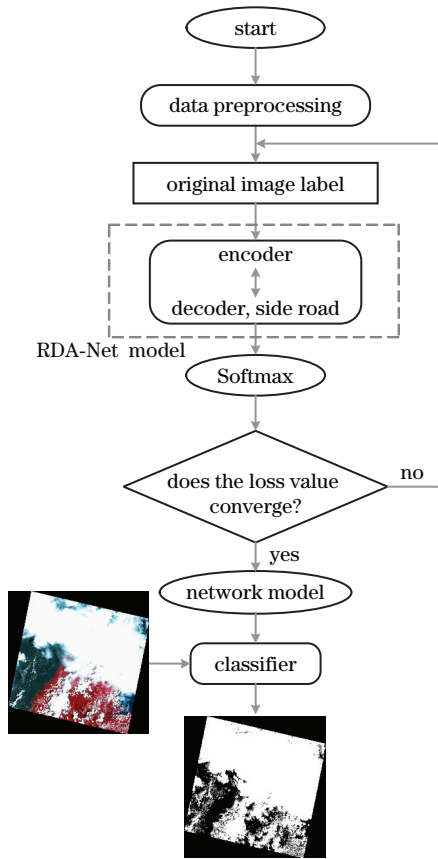


图 9 云与云阴影检测的实验流程

Fig. 9 Experimental flow of cloud and cloud shadow detection

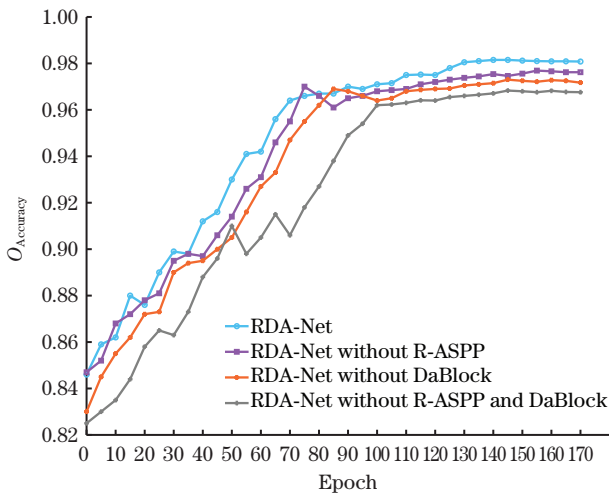


图 10 总体精度与迭代次数的关系

Fig. 10 Relationship between overall accuracy and number of iterations

### 3.4 评价指标

为了定量评价模型的检测性能,使用精确率( $P_{\text{Precision}}$ )、准确率( $A_{\text{Accuracy}}$ )、召回率( $R_{\text{Recall}}$ )、 $F_1$ 值和均交并比  $M_{\text{MIoU}}$  5 种语义分割定量评价指标进行评价。5 种语义分割定量评价指标可分别表示为

$$P_{\text{Precision}} = \frac{S}{S + E}, \quad (11)$$

$$A_{\text{Accuracy}} = \frac{S + W}{S + R + E + W}, \quad (12)$$

$$R_{\text{Recall}} = \frac{S}{S + R}, \quad (13)$$

$$F_1 = 2 \times \frac{R_{\text{Recall}} \times P_{\text{Precision}}}{R_{\text{Recall}} + P_{\text{Precision}}}, \quad (14)$$

$$M_{\text{MIoU}} = \frac{S}{S + W + E}, \quad (15)$$

式中: $S$  表示正确检测为云像元的范围; $W$  表示正确检测为非云像元的范围; $R$  表示将云像元误检为非云像元的范围; $E$  表示将非云像元误检为云像元的范围。当计算云像元的各个评价指标时,将云阴影像元划为下垫面像元;当计算云阴影像元的各个评价指标时,将云像元划为下垫面像元。

### 3.5 云检测结果分析

#### 3.5.1 定性分析

将所提方法与 5 种典型的基于深度学习的方法进行云检测结果的定性分析,5 种方法分别为  $K$ -means<sup>[19]</sup>、RU-Net、FCN-8s (Fully Convolutional Network-8s)<sup>[20]</sup>、SegNet (Segmentation Network)<sup>[21]</sup> 和 DeepLab<sup>[22]</sup>。图 11 为 6 种方法对高分一号 WFV 遥感影像进行云检测结果的视觉对比。选取两张下垫面和云形状不同的遥感影像,其中 image 1 的原始影像包含大量的厚云、薄云、碎云和河流,容易产生混淆,image 2 的原始影像包含厚云、薄云和冰雪,若进行精确的云检测,则难度较大。从图 11 可以看到,FCN-8s 方法在云检测过程中丢失了大量细节,由于该方法只经过三次上采样操作,所以只能检测出云区的大致位置,故无法很好地获得高层语义信息; $K$ -means 方法对厚云区域的检测效果比较明显,对薄云区域的检测效果较差,容易受到冰雪等高亮下垫面的影响;SegNet 方法增加了上采样操作,所以检测结果优于 FCN-8s 方法和  $K$ -means 方法,但仍会出现过度分割和细节丢失过多的情况;DeepLab 方法得到的云检测结果较好,但存在云边界处检测不够精细的问题;RU-Net 方法由于使用残差模块和多尺度特征融合技术,所以取得很好的像素级云检测结果,但在某些边缘区域,检测效果仍不精确;RDA-Net 方法的像素级云检测效果最佳,可以很好地区分云区域与冰雪等高亮的下垫面,可以精确检测出细小的云边缘区域以及较难检测的薄云区域。



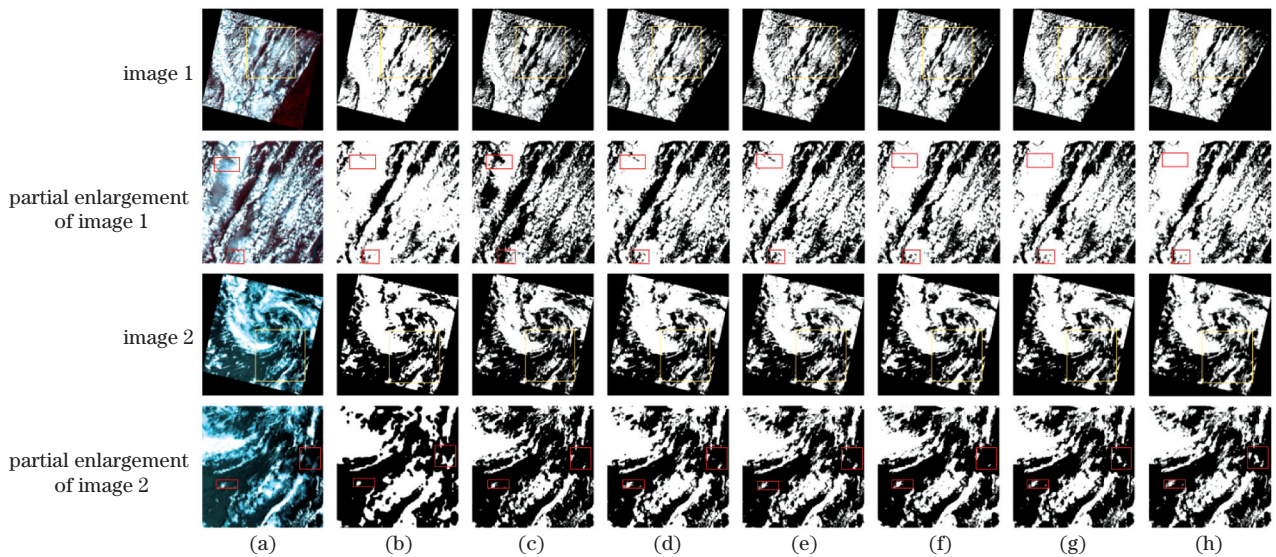


图 11 高分一号 WFV 遥感影像云在不同方法下的检测结果对比。(a)原始影像;(b)FCN-8s 方法;(c)K-means 方法;(d)SegNet 方法;(e)DeepLab 方法;(f)RU-Net 方法;(g)RDA-Net 方法;(h)云标签

Fig. 11 Comparison of detection results of Gaofen-1 WFV remote sensing image cloud under different methods.

(a) Original images; (b) FCN-8s method; (c) K-means method; (d) SegNet method; (e) DeepLab method; (f) RU-Net method; (g) RDA-Net method; (h) cloud tags

### 3.5.2 定量分析

将所提方法与上述 5 种典型的基于深度学习的方法进行云检测结果的定量分析。使用  $P_{\text{Precision}}$ 、 $A_{\text{Accuracy}}$ 、 $R_{\text{Recall}}$ 、 $F_1$  和均交并比  $M_{\text{MIoU}}$  5 种语义分割定量评价指标对 6 种方法进行评价。表 1 为 6 种方法对 21 张高分一号 WFV 遥感影像测试集进行云检测得到的定量评价指标均值。从表 1 可以看到, K-means 方法的云检测效果最差, 因为该方法容易受到高亮下垫面的影响, 所以评价指标均较低; FCN-8s 方法的准确率和召回率较低, 检测过程中丢失了大量细节; SegNet 方法和 DeepLab 方法的精确率和准确率相对较低, 云边界处的检测不够精细; RU-Net 方法取得很好的云检测结果, 但 RU-Net 方法的整体评价指标低于 RDA-Net 方法; 所提的 RDA-Net 方法的云检测效果最好, 在整体定量评价指标上较其他方法有明显优势, 精确率可

达 94.74%, 准确率达到 97.82%, 召回率达到 93.69%,  $F_1$  值达到 0.9421, 均交并比达到 0.8790, 综合性能明显优于其他云检测方法。

表 1 不同方法的云检测定量比较结果

Table 1 Quantitative comparison results of cloud detection by different methods

Method	$P_{\text{Precision}} / \%$	$A_{\text{Accuracy}} / \%$	$R_{\text{Recall}} / \%$	$F_1$	$M_{\text{MIoU}}$
FCN-8s	90.25	84.54	86.38	0.8827	0.7606
K-means	76.42	84.17	72.63	0.7448	
SegNet	90.31	93.03	90.72	0.9051	0.7953
DeepLab	92.66	94.86	92.05	0.9235	0.8019
RU-Net	93.80	<b>97.93</b>	92.94	0.9337	0.8375
RDA-Net	<b>94.74</b>	97.82	<b>93.69</b>	<b>0.9421</b>	<b>0.8790</b>

### 3.6 云阴影检测结果分析

#### 3.6.1 定性分析

将所提的方法与 RU-Net 方法进行云阴影检测结果的定性分析。图 12 为两种方法对高分一号

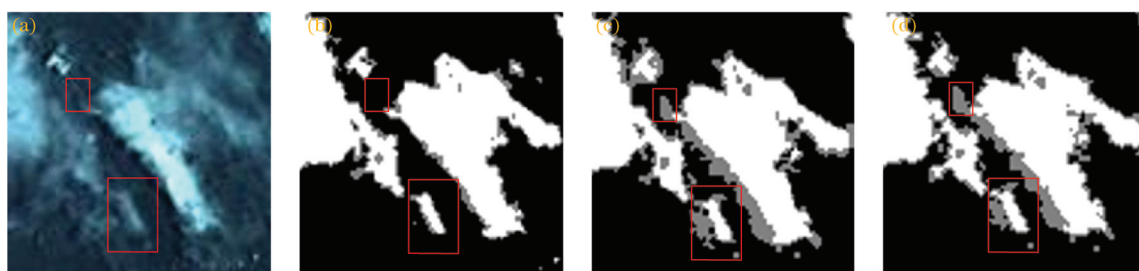


图 12 高分一号 WFV 遥感影像云阴影在两种方法下的检测结果视觉对比。

(a)原始影像;(b)RU-Net 方法;(c)RDA-Net 方法;(d)云与云阴影标签

Fig. 12 Visual comparison of cloud shadow detection results of Gaofen-1 WFV remote sensing image under two methods.

(a) Original image; (b) RU-Net method; (c) RDA-Net method; (d) cloud and cloud shadow labels

WFV 遥感影像进行云阴影检测结果的视觉对比。由于原始影像中包含大量的薄云,为此精确检测出薄云的阴影难度较大。云与云阴影标签中的白色区域表示云,灰色区域表示云阴影,黑色区域表示下垫面。从图 12 可以看到,RU-Net 方法对云阴影的检测效果较差,细节丢失严重,无法精确检测薄云阴影;RDA-Net 方法对像素级云阴影的检测可以取得较好的效果,能够检测出薄云阴影,检测结果优于 RU-Net 方法,而且保留更多的云阴影边缘细节,但与云检测效果相比,云阴影检测效果弱于云检测效果。

### 3.6.2 定量分析

将所提方法与 RU-Net 方法进行云阴影检测结果的定量分析,使用  $P_{\text{Precision}}$ 、 $A_{\text{Accuracy}}$ 、 $R_{\text{Recall}}$ 、 $F_1$  和  $M_{\text{MIOU}}$  5 种语义分割定量评价指标进行评价。表 2 为两种方法对 21 张高分一号 WFV 遥感影像测试集进行云阴影检测得到的定量评价指标均值。从表 2 可以看到,RDA-Net 方法的各评价指标均高于 RU-Net 方法,精确率可达 85.25%,准确率达到 96.04%,召回率达到 80.38%, $F_1$  值达到 0.8274;相较于 RU-Net 方法使用的普通残差结构,RDA-Net 方法引入了递归残差模块并与双注意力模块和改进空洞空间金字塔池化模块结合,所以综合性能明显优于 RU-Net 方法,可有效对云阴影进行检测。

表 2 不同方法的云阴影检测定量比较结果

Table 2 Quantitative comparison of cloud shadow detection by different methods

Method	$P_{\text{Precision}} / \%$	$A_{\text{Accuracy}} / \%$	$R_{\text{Recall}} / \%$	$F_1$	$M_{\text{MIOU}}$
RU-Net	74.67	90.38	68.73	0.7158	0.8375
RDA-Net	<b>85.25</b>	<b>96.04</b>	<b>80.38</b>	<b>0.8274</b>	<b>0.8790</b>

## 4 结 论

本研究将基于深度学习的语义分割方法应用于遥感影像云与云阴影检测,提出一种基于 RDA-Net 的遥感影像云与云阴影检测方法,该方法结合了递归残差模块、双注意力模块、改进空洞空间金字塔池化模块和多尺度特征融合技术。递归残差模块可以避免深层网络出现退化的问题。双注意力模块可以有效捕捉全局特征的依赖关系,从而扩大云、云阴影区域与下垫面区域之间的差异,帮助神经网络对云与云阴影区域进行聚焦,提高大面积云区域的分割完整度以及碎云和云阴影的检测精度。改进空洞空间金字塔池化模块在不增加参数的前提下增大感受

野,可以提取图像的多尺度特征。多尺度特征融合技术将不同尺度的特征进行融合,从而提取多尺度和高层次的空间和语义信息。实验结果表明,RDA-Net 方法可有效对遥感影像云与云阴影进行像素级检测,能够获得较高的检测精度,当下垫面为混淆度极高的雪和霾等背景物时,仍能取得较好的检测结果。

由于云阴影检测易受到山脉和地面物体阴影的影响,容易出现云阴影误检的现象,因此云阴影的检测精度低于云。今后的工作将继续优化模型并应用包含更多云阴影的数据集以加强对云阴影的训练,进一步提高遥感影像云与云阴影的检测效果。

## 参 考 文 献

- [1] King M D, Platnick S, Menzel W P, et al. Spatial and temporal distribution of clouds observed by MODIS onboard the terra and aqua satellites [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(7): 3826-3852.
- [2] Skakun S, Vermote E F, Roger J C, et al. Validation of the LaSRC cloud detection algorithm for Landsat 8 images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(7): 2439-2446.
- [3] 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.
- [4] Rossow W B, Schiffer R A. Advances in understanding clouds from ISCCP [J]. Bulletin of the American Meteorological Society, 1999, 80(11): 2261-2287.
- [5] Li C W, Deng X P, Zhao H C. Cloud detection algorithm for remote sensing images based on multiple feature extraction [J]. Digital Technology and Application, 2017(9): 124-127.  
李超伟, 邓新蒲, 赵昊宸. 基于多特征提取的遥感影像云检测算法研究 [J]. 数字技术与应用, 2017(9): 124-127.
- [6] Xie F Y, Shi M Y, Shi Z W, et al. Multilevel cloud detection in remote sensing images based on deep learning [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(8): 3631-3640.
- [7] 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.
- [8] Drönner J, Korfhage N, Egli S, et al. Fast cloud segmentation using convolutional neural networks [J]. *Remote Sensing*, 2018, 10(11): 1782.
- [9] Xu K, Guan K Y, Peng J, et al. DeepMask: an algorithm for cloud and cloud shadow detection in optical satellite remote sensing images using deep residual network [EB/OL]. (2019-11-09) [2020-11-15]. <https://arxiv.org/abs/1911.03607>.
- [10] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3141-3149.
- [11] Tai Y, Yang J, Liu X M. Image super-resolution via deep recursive residual network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2790-2798.
- [12] Song T Q, Li J X, Zhang X Y. Building recognition in high-resolution remote sensing image based on deep learning [J]. *Computer Engineering and Applications*, 2020, 56(8): 26-34.  
宋廷强, 李继旭, 张信耶. 基于深度学习的高分辨率遥感图像建筑物识别[J]. *计算机工程与应用*, 2020, 56(8): 26-34.
- [13] Jiang H, Zhang C, Wu M. Pedestrian detection based on multi-scale fusion features [C] // 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC), August 22-24, 2018, Guiyang, China. New York: IEEE Press, 2018: 329-333.
- [14] Xi Z H, Hou C Y, Yuan K P, et al. Super-resolution reconstruction of accelerated image based on deep residual network [J]. *Acta Optica Sinica*, 2019, 39(2): 0210003.  
席志红, 侯彩燕, 袁昆鹏, 等. 基于深层残差网络的加速图像超分辨率重建[J]. *光学学报*, 2019, 39(2): 0210003.
- [15] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [M] // Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [16] Zhang J Q, Li X Y, Li L Y, et al. Landsat 8 remote sensing image based on deep residual fully convolutional network [J]. *Laser & Optoelectronics Progress*, 2020, 57(10): 102801.  
张家强, 李潇雁, 李丽圆, 等. 基于深度残差全卷积网络的 Landsat 8 遥感影像云检测方法[J]. *激光与光电子学进展*, 2020, 57(10): 102801.
- [17] Manaswi N K. Understanding and working with Keras [M] // Manaswi N K. *Deep learning with applications using python*. Berkeley: Apress, 2018: 31-43.
- [18] Li Z W, Shen H F, Li H F, et al. Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery [J]. *Remote Sensing of Environment*, 2017, 191: 342-358.
- [19] Wu C, Wang H W, Wang Z Q, et al. Zero-shot classification for remote sensing scenes based on locality preservation [J]. *Acta Optica Sinica*, 2019, 39(7): 0728001.  
吴晨, 王宏伟, 王志强, 等. 基于局部保持的遥感场景零样本分类算法[J]. *光学学报*, 2019, 39(7): 0728001.
- [20] 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. New York: IEEE Press, 2015: 3431-3440.
- [21] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [22] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.