

# 基于改进 M 型卷积网络的 RGB 彩色遥感图像云检测

胡敬锋<sup>2</sup>, 张秀再<sup>1,2</sup>, 杨昌军<sup>3\*</sup>

<sup>1</sup>南京信息工程大学江苏省大气环境与装备技术协同创新中心, 江苏 南京 210044;

<sup>2</sup>南京信息工程大学电子与信息工程学院, 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** RGB 彩色图像中云没有明显的颜色分布与纹理模式, 导致云检测易产生误检且细节丢失严重。针对这一问题, 提出一种改进的 M 型卷积网络(RM-Net)模型, 实现端到端的像素级语义分割。对原始数据集进行增强, 并标注对应的像素级标签。利用空洞空间金字塔池化, 在不丢失信息的前提下提取图像多尺度特征, 并结合残差单元使网络不易出现退化。利用编码器模块与左路径提取图像全局上下文信息, 利用解码器模块与右路径恢复图像空间分辨率, 根据融合后的特征判别每个像元的类别概率, 将其输入分类器进行像素级的云和非云分割。对 Landsat8 和高分一号 WFV RGB 彩色图像进行训练和测试, 实验结果表明本文方法在不同条件下能很好地检测云边缘细节, 并取得较高精度的云阴影检测, 由此证明本文方法具有较好的泛化性与稳健性。

**关键词** 图像处理; 云检测; 遥感图像; M 型卷积网络; 空洞空间金字塔池化; 残差单元; 云阴影检测

中图分类号 TP751.1

文献标识码 A

doi: 10.3788/LOP56.162804

## Cloud Detection of RGB Color Remote Sensing Images Based on Improved M-Net

Hu Jingfeng<sup>2</sup>, Zhang Xiuzai<sup>1,2</sup>, Yang Changjun<sup>3\*</sup>

<sup>1</sup>*Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China;*

<sup>2</sup>*School of Electronic and Information Engineering, Nanjing University of Information Science and 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** Cloud detection is prone to error and considerable loss of details because clouds do not have obvious color distribution and texture pattern in RGB color images. Therefore, this study proposes an improved M-Net model called the RM-Net model to achieve end-to-end pixel-level semantic segmentation. An original dataset is enhanced and a corresponding pixel-level label is marked. Multi-scale image features are extracted without losing data via atrous spatial pyramid pooling, and residual units are combined to make the network resistant to degradation. Global context informations of the images are extracted using the encoder module and the left path. The spatial resolutions of the images are restored using the decoder module and the right path. Each pixel's category probability is determined based on fused features, and pixel-level cloud and non-cloud segmentation are performed using the input classifier. When training and testing Landsat8 and GaoFen-1 WFV RGB color images, experimental results show that the proposed method can well detect cloud edge details under various conditions and achieve high-precision cloud shadow detection, thus demonstrating that the proposed method has high generalization and robustness.

**Key words** image processing; cloud detection; remote sensing images; M-Net; atrous spatial pyramid pooling;

收稿日期: 2019-01-27; 修回日期: 2019-03-04; 录用日期: 2019-04-09

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

\* E-mail: yangcj@cma.gov.cn

residual units; cloud shadow detection

OCIS codes 100.2000; 100.4996; 100.3008; 010.1615

## 1 引言

随着遥感技术的发展,遥感影像已广泛用于资源利用、环境检测及气象探测等工作领域。全球有效云量(云发射率加权的云量)约为  $0.50 \pm 0.05$ <sup>[1]</sup>, 准确地分离遥感图像中的有云像元和无云像元已成为遥感数据处理的重要环节,同时也为天气预报与自然灾害预测提供最可靠的信息源。因此,有效的云检测是对遥感影像后续分析与利用的前提。

目前,多种云检测方法被提出。基于阈值的云检测方法<sup>[2-4]</sup>主要通过提取每个像素的各种光谱特征,然后使用一个或多个阈值确定云掩模。该方法主要利用低层次的光谱信息,空间信息涉及相对较少,对下垫面和云覆盖范围的敏感性较高<sup>[5]</sup>,这决定了云检测工作的复杂性,导致产生误检。随着大数据应用推广及机器学习技术的快速发展,云检测中这种复杂性工作可由计算机来实现,从而极大节省成本,所以基于机器学习的云检测方法的应用越来越广泛<sup>[6]</sup>。Xu 等<sup>[7]</sup>利用决策树提取云边界。Li 等<sup>[8]</sup>利用亮度和纹理特征训练支持向量机(SVM)分类器检测云。Hu 等<sup>[9]</sup>将计算机视觉与随机森林方法相结合,得到云覆盖掩模图。其中,随着深度学习在图像识别<sup>[10]</sup>、姿态估计<sup>[11]</sup>、自然语言处理<sup>[12]</sup>等方面的快速发展,神经网络强大的识别能力得到了广泛的认可,基于深度学习的云检测方法也引起了人们的关注。Shi 等<sup>[13-14]</sup>利用超像素分割和全卷积神经网络(FCNN)对 Quickbird、谷歌地球图像和 Landsat8 图像中的云进行检测。陈洋等<sup>[15]</sup>采用卷积网络对资源三号遥感影像进行云检测。这些方法与传统机器学习方法相比,可实现更高精度的云检测,但对云的边缘检测较为粗糙,导致云的细节丢失严重。

针对现有方法在遥感图像云检测中易产生误检及边缘细节检测精度不够高的问题,提出一种深度语义分割网络 RM-Net,对 Landsat8 和高分一号 WFV RGB 彩色图像进行像素级云检测。该网络是对 M 型卷积网络(M-Net)<sup>[16]</sup>的改进,其框架主要由编码器、解码器和 2 个边路径组成。该网络在分割网络中引入避免网络训练退化的残差单元<sup>[17]</sup>,增加采样密度(在不增加计算量的前提下)的空洞空间金字塔池化(ASPP)<sup>[18]</sup>,并利用 Adam 算法对网络

进行优化。实验结果表明,对 M-Net 的改进可进一步提高其性能,在复杂条件下仍能取得较高精度的云检测。

## 2 原理与方法

在 RGB 彩色遥感图像中,云通常是白色和明亮的,易与大多地面物体区分。但是,由于云的厚度和形态各不相同,下垫面地物种类繁多,常存在容易混淆的物体(如白色的建筑物、冰、雪等),且与长波长传感器(即红外、多光谱)相比,RGB 色彩波段对散射条件更为敏感<sup>[19]</sup>,故仅利用图像的空间信息进行云识别,难度较大。因此,从目标分割和分类的角度解决这个问题可得到更直观的结果,还可避免目前大多数云检测方法依赖于特定传感器的弊端。本文将 RGB 彩色图像与其对应的标签输入网络进行训练和测试,并得到最终的云检测结果。

### 2.1 RM-Net 分割网络

在语义分割中,低层特征包含大量空间信息,高层特征包含大量语义信息。若要取得好的分割结果,有必要同时使用低层次的细节和高层次的语义信息。如 U-Net<sup>[20]</sup>,可将低层次精细的细节补偿至高层次的语义特征中。M-Net<sup>[16]</sup>是基于 U-Net 改进而来,除了具有 U-Net 中的编码器和解码器路径外,还增加了两条侧路径来提供深度监督功能。M-Net 采样过程中会丢失较多信息,易造成对薄云和含有高亮度下垫面的云误检。针对这一问题,提出一种改进的 M 型卷积网络模型(RM-Net),其主要由多个残差单元、卷积层、反卷积层与 concat 层组成。

1) 残差单元:神经网络的性能随着其深度的加深而提高,但同时会引起网络的退化,而残差神经网络<sup>[17]</sup>可很好地解决这一问题,其方法主要是在堆叠层上采取残差学习,每个残差单元都可以用一般形式表示为

$$x_{i+1} = f[g(x_i) + H(x_i, Y_i)], \quad (1)$$

式中: $x_i$  和  $x_{i+1}$  分别是第  $i$  个残差单元的输入和输出; $H(x_i, Y_i)$  代表学习到的残差映射; $Y_i$  为权值; $g(x_i)$  为恒等映射; $f(\cdot)$  为激活函数。当输入和输出的维度相同时,可以采用恒等映射,此时  $g(x_i) = x_{i+1}$ ;若维度不同时,需使用  $1 \times 1$  的卷积恢复维度。卷积单元和残差单元的对比如图 1 所示,其由批量归一化(BN)、激活函数 ReLU 和卷积层(Conv)多重组

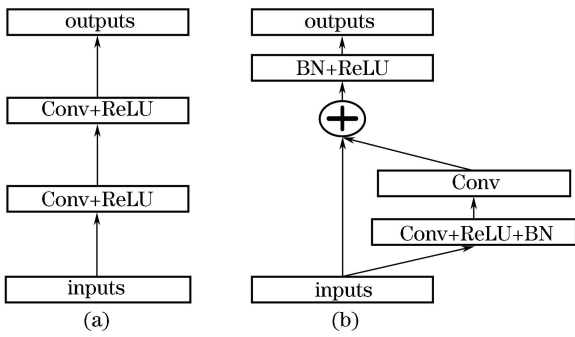


图1 卷积单元与残差单元。(a)前向传播卷积单元；  
(b)残差单元

Fig. 1 Convolution unit and residual unit. (a) Forward propagation convolution unit; (b) residual unit

合而成。残差中低层与高层网络之间的跳跃连接将促进信息传播而不退化,使网络更易于训练。

2) 卷积层:在图像处理中,大尺寸的图像在网络中权重数量较大,造成难以对传统的神经网络进行训练。而图像的空间联系是局部的,距离越近的像素相关性越强。因而,每个神经元不需要对全局图像进行感知,只需对图像进行局部感知。且由于图像某一部分的统计特性与其他部分相同,所以对于图像中的不同部分可使用同样的学习特征。含有卷积层的网络中每个神经元的权重个数为卷积核的大小,这样可极大地减少权重的数量。在分割网络中,卷积层用于提取输入的特征和增大感受野,高层特征包含大量空间信息,低层特征包含大量语义信息,多层的卷积从低级特征中提取更复杂的特征。假设每个卷积层的输入为  $H \times W \times C$  的特征图,其中  $H$  与  $W$  分别为特征图的高与宽,  $C$  为通道数,  $W_j, b_j$  与  $X_j$  分别为第  $j$  层卷积核的权重、偏置项与输入特征图,则第  $j$  层特征图的输出  $Z_j$  与其大小  $F'$  可分别表示为

$$Z_i = f(W_j \times X_j + b_j), \quad (2)$$

$$F' = \text{floor}[(F + 2P - k)/s] + 1, \quad (3)$$

式中:  $f(\cdot)$  为非线性激活函数 ReLU, 即  $f(x) = \max(0, x)$ , 除最后一个卷积层外, 在每个卷积层之后都应用 ReLU 和批量归一化层(BN);  $F$  为输入特征图的大小;  $P$  为 Padding;  $k$  与  $s$  分别为卷积核的大小与步长;  $\text{floor}(\cdot)$  表示向下取整。

3) 扩张卷积与 ASPP: 在语义分割中, 增大感受野有利于提取语义信息, 但会弱化空间信息的提取, 而扩张卷积很好地解决了这一问题。扩张卷积与普通卷积相比, 卷积核大小和输出的特征图大小一样, 但扩张卷积具有更大的感受野(卷积核在图像上看到的区域更大)。一个扩张率为 2 的  $3 \times 3$  卷积

核, 其感受野与  $5 \times 5$  的卷积核相同, 但参数数量仅为 9 个, 是  $5 \times 5$  卷积核参数数量的 36%。扩张卷积对应的感受野  $R_{\text{RF}}$  表达式为

$$R_{\text{RF}} = (k - 1) \times r + 1, \quad (4)$$

式中:  $r$  为扩张率, 当  $r = 1$  时为普通卷积。

跳跃连接将底层特征与高层特征融合, 同时使两个堆叠的卷积层获得更大的感受野。假设  $k_1$  与  $k_2$  为两个堆叠的卷积核大小, 新的感受野  $R'_{\text{RF}}$  的表达式为

$$R'_{\text{RF}} = k_1 + k_2 - 1. \quad (5)$$

具有不同扩张率的 ASPP 能有效地捕获多尺度信息, 但当等价后的感受野大于特征图大小时, 会造成滤波器的退化<sup>[21]</sup>。由于本文使用的训练图像大小为  $256 \text{ pixel} \times 256 \text{ pixel}$ , 编码器最后一个残差单元接收的特征图大小为  $32 \text{ pixel} \times 32 \text{ pixel}$ , 所以在编码器最后一个残差单元使用 3 个  $3 \times 3$  扩张率依次为 3, 6, 9 的递增的扩张卷积层, 并结合多个跳跃连接, 使采样更加密集。根据(4)式和(5)式, 本文使用的 ASPP 中最大的感受野为 31 (小于该卷积层输入的特征图大小), 从而最大化提取有效特征。在 ASPP 后使用  $1 \times 1$  的卷积层可将通道数减少到 128, 以满足维度匹配。

4) 反卷积层: 反卷积是卷积的逆过程, 可将输入的特征图恢复为原输入图像的大小, 实现像素级的语义分割。假设反卷积层的卷积核大小为  $k'$ , 则恢复的特征图大小  $O'$  可表示为

$$O' = (O - 1)s' + k' - 2P', \quad (6)$$

式中:  $O$  为输入反卷积层的特征图大小;  $s'$  为反卷积的步长;  $P'$  为 Padding。

5) concat 层: 在信息提取的过程中, 不同的卷积层提取不同的空间信息与语义信息, 而 concat 层可将两个及以上同等大小的特征图进行拼接, 从而达到特征融合、促进信息传播的作用。

RM-Net 深层网络架构如图 2 所示, Conv 为卷积运算, ReLU 为激活函数, BN 为批量归一化层, stride 为卷积运算的步长, Deconv 为反卷积运算, eltwise sum 为求和运算, concat 为特征融合运算, 其由 4 个滤波器路径组成, 即编码与解码路径和 2 条边缘路径。该网络共有 31 层和 8 个残差单元, 包含不改变特征图大小、步长为 1 的卷积层, 以及改变特征图大小、步长为 2 的卷积层与反卷积层。在卷积滤波器的级联中, 丰富的跳跃连接将底层特征与高层特征融合, 以结合语义信息和空间信息, 促进信息的反向传播而不会出现退化, 使网络能更好地学习特征。

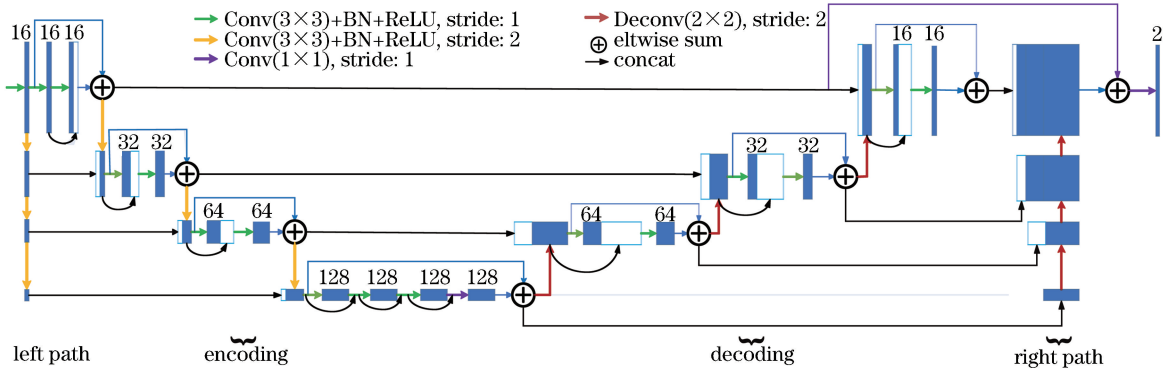


图2 RM-Net网络结构

Fig. 2 RM-Net network structure

在分割网络中,编码器从RGB彩色图像中提取稳健抽象表示,解码器根据编码器的响应计算像素级的分割和分类掩码。使用步长为2的卷积层代替池化层使特征图减半,每个卷积层的特征通道数都随着步长为2的卷积和反卷积而变化。由于测试集图像为1000 pixel×1000 pixel,在经过连续4个下采样后,第1次上采样时恢复的特征图大于第3个下采样后的特征图,无法完成特征融合,导致不能对测试集图像进行云检测。因此,利用包含不同扩张率的ASPP替换M-Net中最后1个池化层和编码器的最后2个卷积层,在避免信息丢失与不增加计算量的前提下,使扩张率进一步增大,同时像素采样更加密集,从而增大感受野并融合多尺度特征。解码路由2×2的反卷积层和2个级联的步长为1的3×3卷积层重复组成。

在最后1个卷积层之后添加概率为0.5的dropout函数可防止网络过拟合。对于2条边路径,左路径由3个步长为2的3×3卷积层组成,其输出端连接到对应的编码层输入端;右路由反卷积层和concat层组成,并融合解码层的输出。最后1个残差单元提取的特征经过1×1的卷积核后进入SoftmaxWithLoss层,其中Softmax函数用于计算每个像素类分布的概率值,并利用Softmax函数的输出计算Loss值。Loss函数 $J(\theta)$ 的定义为

$$J(\theta) = -\frac{1}{M} \left[ \sum_{m=1}^M \sum_{k=1}^K 1\{y_m = k\} \log \frac{\exp(\theta_k^T h_m)}{\sum_{k=1}^K \exp(\theta_k^T h_m)} \right], \quad (7)$$

式中: $k$ 为标签图中像素的类别值; $T$ 为转置; $y_m$ 为类别标签,有 $K$ 个不同取值,将遥感图像分割成云与非云区域属于二分类问题,所以此处 $K=2$ ; $h_m$

为输入图像像素观测向量; $\theta$ 为模型参数; $M$ 为图像的像素数; $1\{\}$ 为显性函数。训练数据集依次经过神经网络、隐藏层,最后传播到输出层,此过程是分割网络的前向传播。由于分割网络的输出结果与实际结果之间存在误差,计算该误差并将其从输出层向隐藏层反向传播,直至传播到输入层,此过程是分割网络的反向传播。在反向传播过程中,根据误差调整网络权重参数值,不断迭代上述过程,直至网络收敛。

## 2.2 DDCN分割网络

为与本文方法进行对比,将结合多尺度特征进行像素级云检测的深度卷积神经网络(DCN)<sup>[22]</sup>改进为DDCN网络,结构如图3所示。考虑到计算机显存、计算速度以及便于对比等方面原因,将DCN方法中使用的最小与最大通道数由64,4096分别改为8,256。在融合除Conv2\_1外的6个不同分支的特征时,分别使用步长为2,4,8,8,8的反卷积替换卷积,使特征图恢复为原输入图片大小,起到端到端语义分割的作用,并无需将标签图像压缩为原图大小的1/8,避免信息的丢失而不利于有效地监督学习。图3中,卷积层中的参数3×3×64,代表卷积核的大小为3,通道数为64;3×3×128(2)括号内的数值代表扩张卷积的扩张率;池化层参数3×3/2与反卷积层参数Deconv1/2中反斜杠后的数值代表步长;C为concat层。将融合的多尺度特征输入输出为2的1×1卷积层,再通过Softmax函数计算每个像素类分布概率。

## 3 数据集的选取与增强

选取的遥感图像来自L8 SPARCS云和云阴影验证数据集<sup>[23]</sup>。此数据集像素级的云掩模根据11个波段的Landsat8数据集获得,其groundtruth的



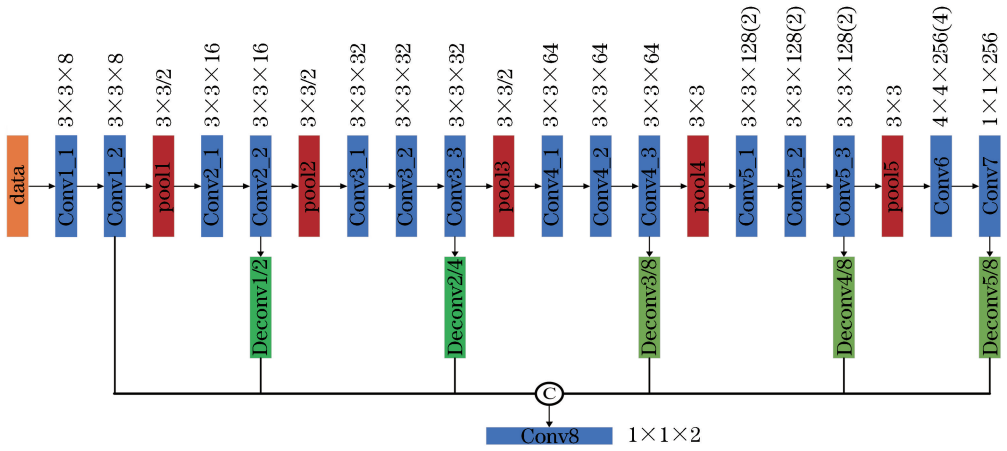


图 3 DDCN 网络结构

Fig. 3 DDCN network structure

准确性足够高,涵盖大的薄云、海洋上空云和下垫面为冰/雪等多种情况,具有较好的代表性。80 张影像含有 7 个手动标记的类,包括云、云阴影、淹没区、冰/雪、水体、水面阴影和陆地,每张图片大小为  $1000 \text{ pixel} \times 1000 \text{ pixel}$ 。从中选取 64 张图像作为训练集与验证集,16 张图像作为测试集。

为验证本文云检测方法的泛化性,选取中国高分 1 号云和云阴影覆盖数据集<sup>[24]</sup>中 RGB 彩色图像进行云检测实验。高分 1 号卫星上的宽视场相机(WFV)成像系统的空间分辨率为 16 m,共有 4 个多光谱波段,该数据集有 108 张影像,

包含全球不同的土地覆盖类型,所有相关的掩模都对云作了标记。为降低图像分割对计算机性能的要求以及提高图像分割的效率,将数据集中的图像统一转化为  $1000 \text{ pixel} \times 1000 \text{ pixel}$ ,从中选取 86 张图像作为训练集与验证集,22 张图像作为测试集。

在训练阶段,由于原始数据少且尺寸过大,且考虑到 GPU 内存的局限性、计算速度及分割方法的时效性等,通过翻转、饱和度调整、亮度调整、颜色调整和加噪等操作增强图像。选取一张增强后的 Landsat8 图像,其效果如图 4 所示。

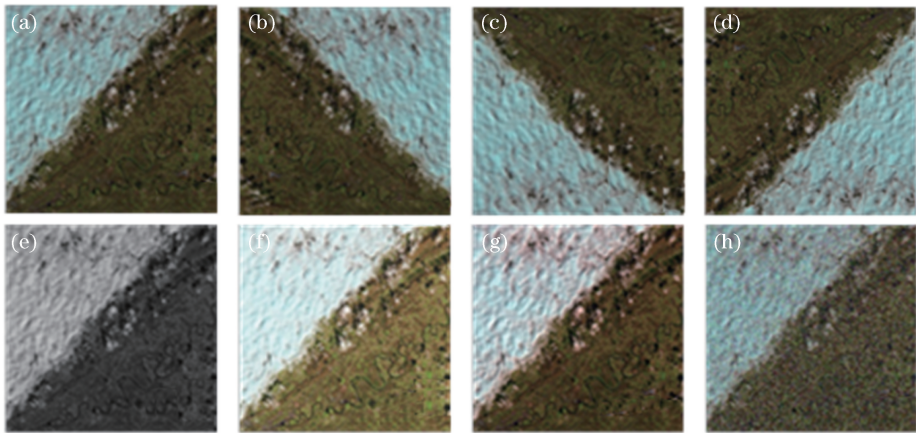


图 4 数据集增强。(a)原图;(b)水平翻转;(c)垂直翻转;(d)水平垂直翻转;(e)饱和度调整;(f)亮度调整;(g)颜色调整;(h)加噪

Fig. 4 Dataset enhancement. (a) Original image; (b) horizontal flip; (c) vertical flip; (d) horizontal and vertical flips; (e) saturation adjustment; (f) brightness adjustment; (g) color adjustment; (h) add noise

将增强后的 Landsat8 原始图像裁剪成大小为  $256 \text{ pixel} \times 256 \text{ pixel}$  的 4608 张子图像,其中 4000 张作为训练集,608 张作为验证集。将增强后的高

分 1 号 WFV 原始图像同样裁剪成  $256 \text{ pixel} \times 256 \text{ pixel}$  的子图像,训练集与验证集图像数量分别为 5000 张与 1192 张。

## 4 实验与评价

### 4.1 实验环境

实验采用 Ubuntu16.04 下接口为 Python 的深度学习框架 Caffe<sup>[25]</sup>, 硬件环境为处理器 Intel i7 8700, 内存为 16G, 图像处理器 (GPU) 为 GeForce GTX 1080, 显存为 8G 的计算机。

### 4.2 实验步骤

1) 数据集预处理: 计算训练集在 RGB 三个维度的均值, 在训练与调用模型时, 将训练集、验证集和测试集分别减去该均值 (遵循模型训练过程, 仅从训练集中获取信息), 可提高训练与测试的速度与精度。

2) 制作标签: 云掩模图像由相关专家手动标记, 再将其转化为灰度图像。由于同一类灰度值是变化的, 需要进行灰度值固定, 将掩模图像中每个像元对应的是“云”还是“非云”, 分别用灰度值 1 和 0 表示, 形成与原图像对应的二值掩模图像, 用于监督网络的训练。

3) 实验参数设置: 训练时批次大小取经验值为 8, 即每迭代 1 次处理 8 张图像。在参数优化阶段, 梯度优化使用 Adam 算法, 设置动量大小  $\beta_1 = 0.9$  和  $\beta_2 = 0.999$  对学习率进行衰减以防止出现过拟合现象, 将学习率设置为 0.0001, 当代价函数损失收敛且趋于平稳时, 停止训练, 共迭代  $12 \times 10^4$  次。

4) 测试: 调用训练生成的最优模型对测试集  $1000 \text{ pixel} \times 1000 \text{ pixel}$  的图像进行像素级云检测, 云区用白色表示, 非云区用黑色表示, 并与其他方法作定性和定量比较。采用交叉验证的方式, 即利用未参与训练的图像对 6 种云检测方法进行精度测试。

为验证加入的残差单元与 ASPP 对云检测的有效性, 将 RM-Net 方法使用的残差单元与 ASPP 分别移除和同时移除, 再通过使用相同的数据集并设置相同的实验参数进行训练。训练迭代的次数与总准确度  $O_{\text{Overall accuracy}}$  的关系如图 5 所示, 总准确度定义为

$$O_{\text{Overall accuracy}} = \frac{p}{p + p'}, \quad (8)$$

式中:  $p$  为每个类别的像元被正确检测的个数;  $p'$  为每个类别的像元被错误检测的个数。

由图 5 可知, 分割网络中加入残差单元与 ASPP 均可改善网络的性能, 且同时加入这两种结构可进一步提升网络的性能, 验证了本文所构建网

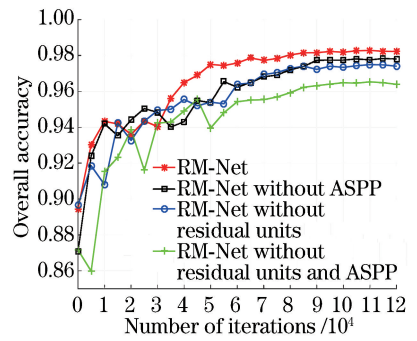


图 5 Overall accuracy 曲线

Fig. 5 Overall accuracy curves

络的合理性。

### 4.3 云检测方法比较

#### 4.3.1 定性比较

图 6 为采用 k-means、CNN + SP<sup>[13]</sup>、FCN2s<sup>[26]</sup>、M-Net<sup>[16]</sup>、DDCN 和 RM-Net 等 6 种方法对 Landsat8 图像进行云检测结果的视觉比较。2 张 Landsat8 图像云的形态与下垫面各不相同, 选取各自对应的子区域以达到更好的视觉对比。图 6 (a) 中的原图包含薄云、厚云和部分碎云, 半透明的薄云区域难以判别; 图 6 (c) 中的原图既有云也有雪/冰 (图中蓝色区域为雪/冰), 易产生混淆; 图 6 (b)、(d) 分别为图 6 (a)、(c) 的子区域, 原图中蓝色区域为雪/冰, 白色或半透明白色区为云, 其余为晴空区; 黑白图中白色为云, 黑色为晴空区。由图 6 可知, k-means 方法可以有效识别厚云, 但对薄云识别效果较差, 且易把明亮的地物识别为云, 如含有雪/冰覆盖的下垫面。CNN + SP 方法将 CNN 与超像素聚类相结合, 即利用超像素区域作为输入块, 单独输出每个块, 可以克服“盐和胡椒”噪声的干扰, 但容易受到超像素分割的影响而产生初始误差, 无法获得像素级的检测精度, 易漏检碎云。FCN2s 方法在采样过程丢失较多特征, 导致云检测结果比较模糊和平滑, 在云边界上丢失了很多细节。M-Net 方法得到的云检测结果优于 FCN2s 方法, 但对薄云分割结果较差。DDCN 方法利用多尺度特征取得了较好的像素级分割结果, 但在薄云和云边缘区域分割结果不理想。RM-Net 方法可更好地提取云的特征, 即使在复杂的条件下 (下垫面高亮、存在大量薄云等), 云的形态仍能较为完整地提取, 位置准确, 生成的云掩模较为精细, 总体更接近于标签图像 (groundtruth)。

图 7 为采用 k-means、CNN + SP<sup>[13]</sup>、FCN2s<sup>[26]</sup>、M-Net<sup>[16]</sup>、DDCN 和 RM-Net 等 6 种方

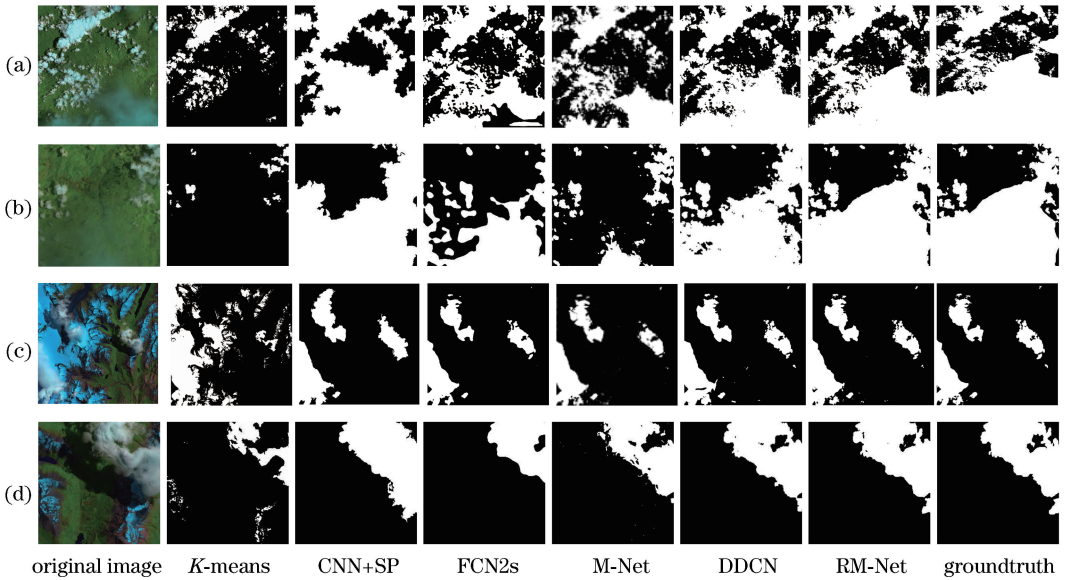


图 6 6 种方法对 Landsat8 图像云检测结果视觉对比

Fig. 6 Visual comparison of Landsat8 image cloud detection results obtained by six methods

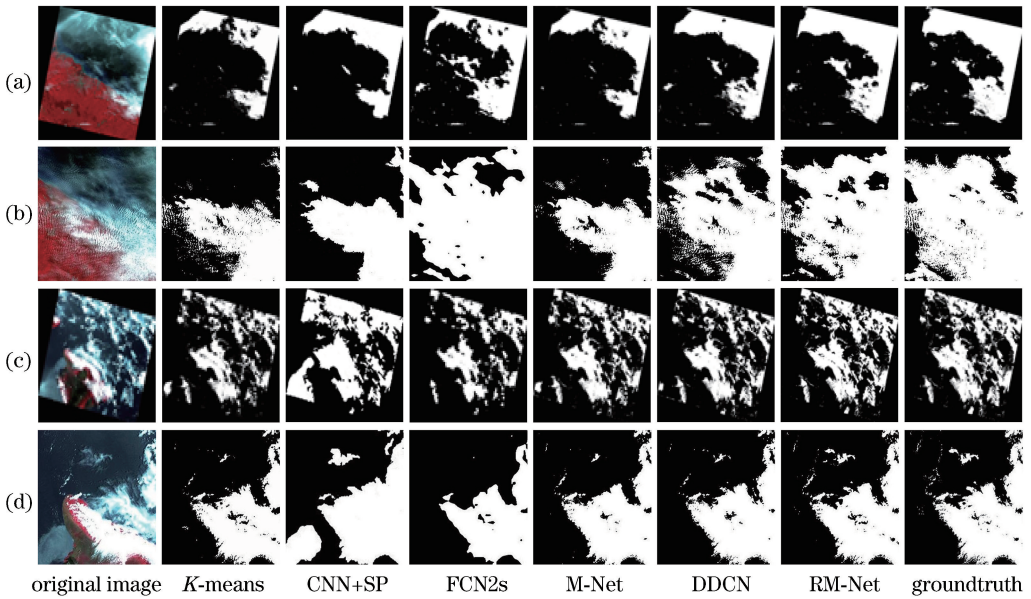


图 7 6 种方法对高分 1 号 WFV 图像云检测结果视觉对比

Fig. 7 Visual comparison of GaoFen-1 WFV image cloud detection results obtained by six methods

法对高分 1 号 WFV 图像检测结果。图 7(a)中的原图含有大量半透明的薄云和厚云；图 7 (c)中的原图含有薄云、厚云和雪；图 7 (b)和(d)分别为图 7 (a)和(c)的子区域。原图中蓝色区域为雪/冰，白色或半透明白色区域为云，其余为晴空区；黑白图中白色为云，黑色为晴空区。由图 7 可知，云检测结果视觉对比类似于图 6，由于高分 1 号 WFV 图像中的云分布相对复杂以及图像压缩等方面的原因，6 种方法都存在部分云误检，但 RM-Net 方法的云检测性能

相对更好，检测结果总体更接近于标签图像，证明了所提方法的泛化性。

#### 4.3.2 定量比较

用 4 个指标定量评价云检测方法的有效性，即精确率  $P_{\text{Precision}}$ 、召回率  $R_{\text{Recall}}$ 、准确率  $A_{\text{Accuracy}}$  和调和均值  $F_{1\text{score}}$ 。评价指标定义表达式分别为

$$P_{\text{Precision}} = \frac{C_{CC}}{C_{CC} + N_{NC}}, \quad (9)$$

$$R_{\text{Recall}} = \frac{C_{CC}}{C_{CC} + C_{CN}}, \quad (10)$$



$$A_{\text{Accuracy}} = \frac{C_{\text{CC}} + N_{\text{NN}}}{C_{\text{CC}} + C_{\text{CN}} + N_{\text{NC}} + N_{\text{NN}}}, \quad (11)$$

$$F_{\text{1score}} = \frac{2 \times P_{\text{Precision}} \times R_{\text{Recall}}}{P_{\text{Precision}} + R_{\text{Recall}}}, \quad (12)$$

式中： $C_{\text{CC}}$ 是正确检测为云像元的个数； $C_{\text{CN}}$ 是云像元误检测为非云像元的个数； $N_{\text{NC}}$ 是非云像元误检测为云像元的个数； $N_{\text{NN}}$ 是正确检测为非云像元的个数。

表1为6种方法对16张Landsat8测试图像中的云检测作定量评价时4个定量评价指标的平均值。从表1中可知，k-means方法的4个定量评价指标相对最差，因为该方法易将高亮的地物误检为云，且不能有效识别薄云，所以准确率较低。CNN+SP方法的精确率相对较低，会将部分非云误检为云。FCN2s方法和M-Net方法的召回率较低，存在少量云漏检情况。DDCN方法与RM-Net方法都取得了较好的检测性能，RM-Net方法在测量指标上优于DDCN方法。

表2为6种方法对22张高分1号WFV图像中的云检测作定量评价时4个定量评价指标的平均值。由表2可知，云检测方法的定量评价与表1类似，RM-Net方法的综合性能相对更优。

表1 6种方法对Landsat8图像云检测结果的定量比较

Table 1 Quantitative comparison of Landsat8 image cloud detection results obtained by six methods

Method	$P_{\text{Precision}}$	$R_{\text{Recall}}$	$A_{\text{Accuracy}}$	$F_{\text{1score}}$
k-means	0.8366	0.6585	0.8396	0.7369
CNN+SP	0.8605	0.9025	0.9064	0.8704
FCN2s	0.9293	0.8734	0.9243	0.9005
M-Net	<b>0.9432</b>	0.9091	0.9673	0.9258
DDCN	0.9322	0.9283	0.9728	0.9302
RM-Net	0.9334	<b>0.9509</b>	<b>0.9816</b>	<b>0.9418</b>

表2 6种方法对高分1号WFV图像云检测结果的定量比较

Table 2 Quantitative comparison of GaoFen-1 WFV image cloud detection results obtained by six methods

Method	$P_{\text{Precision}}$	$R_{\text{Recall}}$	$A_{\text{Accuracy}}$	$F_{\text{1score}}$
k-means	0.7499	0.7154	0.8356	0.7322
CNN+SP	0.8413	0.8635	0.8914	0.8523
FCN2s	0.9019	0.8935	0.9238	0.8976
M-Net	0.9307	0.9039	0.9590	0.9132
DDCN	<b>0.9316</b>	0.9273	0.9654	0.9294
RM-Net	0.9265	<b>0.9353</b>	<b>0.9762</b>	<b>0.9309</b>

#### 4.4 云与云阴影检测

由于遥感图像中云与云阴影总是成对出现，因此云阴影的检测对遥感图像的应用也具有重要意

义。挑选Landsat8和高分一号WFV RGB含云与云阴影图像各20张，用来训练、验证与测试的图像分别为12张，4张与4张，扩充后的训练集为900张，验证集为252张，大小为256 pixel×256 pixel。将标签图像中每个像元对应的“云”、“云阴影”和“晴空”分别用2,1,0表示，用于监督网络训练。最后一个卷积层输出设为3，即实现云、云阴影和晴空的三分类。

图8为本文方法与DDCN(最后一层卷积层输出为3)方法对Landsat8和高分一号WFV RGB图像云和云阴影检测的视觉对比，groundtruth中白色为云，灰色为云阴影，黑色为晴空区。由图8可看出，两种云检测方法都能有效地检测云及云阴影。与云阴影检测结果相比，云检测结果更加准确。

表3为对两种遥感图像(8张)中云与云阴影检测作定量评价时4个定量评价指标的平均值，即云阴影精确率 $P'_{\text{Precision}}$ 、云阴影召回率 $R'_{\text{Recall}}$ 、准确率 $A'_{\text{Accuracy}}$ 和云阴影调和均值 $F'_{\text{1score}}$ ，在计算云阴影像元的精确率与召回率时，把云像元识别为晴空像元。从表3中可看出，RM-Net方法4个评价指标均高于DDCN方法，表明RM-Net方法可同时对云和云阴影进行有效的检测。

## 5 结论

利用改进的M型卷积网络(RM-Net)方法对RGB彩色遥感图像进行像素级云检测，该方法结合了残差网络、空洞空间金字塔池化与M-Net的优点。残差单元使网络在加深时不易出现退化，丰富的跳跃连接可促进网络的前向传播和反向传播，空洞空间金字塔池化以及堆叠的卷积层使网络在不增加参数及不弱化空间信息的前提下增大感受野，从而提取更多的特征，为像素级预测提供多尺度信息。实验中定性和定量分析表明，融合多特征的DDCN方法与RM-Net方法可对多种复杂条件下的RGB彩色遥感图像的云进行更为有效的检测，在云的边缘生成精细的掩模，误检率较低，并能对云阴影进行较为有效的检测，具有较高的检测精度与很好的泛化性。此外，RM-Net方法中丰富的跳跃连接以及包含不同扩张率的ASPP可使网络提取更多有效特征，所以实验结果优于DDCN方法和其他对比方法。

运用深度学习方法可显著提高遥感图像云检测的速度与精度，为遥感图像的准确应用起到了一定的促进作用。但本文云检测方法还存在许多不足，



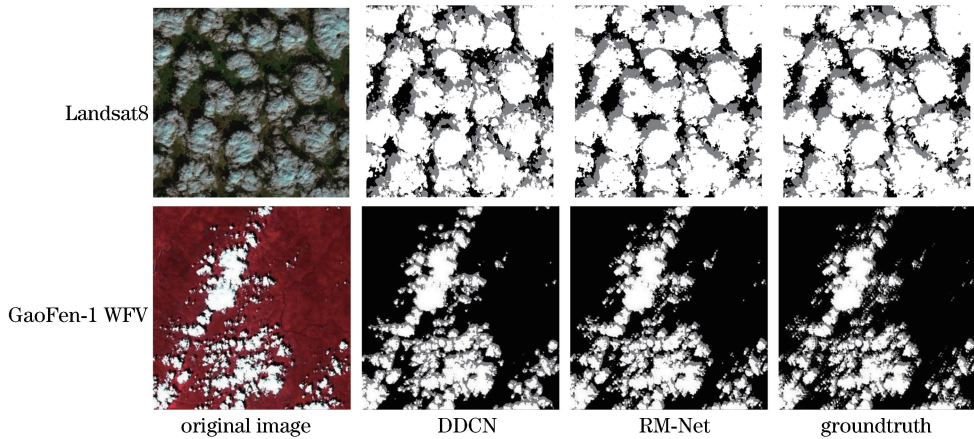


图8 云和云阴影检测结果

Fig. 8 Detection results of cloud and cloud shadow

表3 云与云阴影检测结果定量比较

Table 3 Quantitative comparison of detection results of cloud and cloud shadow

Method	$P'_{Precision}$	$R'_{Recall}$	$A'_{Accuracy}$	$F'_{Score}$
DDCN	0.7734	0.7116	0.9356	0.7412
RM-Net	<b>0.8657</b>	<b>0.7942</b>	<b>0.9703</b>	<b>0.8284</b>

如没有将厚云与薄云区分开,影像中存在部分云阴影误检和少量薄云漏检等。下一步工作将完善分割网络并对更多类型的数据集进行实验,进一步提高云与云阴影检测精度。

## 参 考 文 献

- [1] Stubenrauch C, Rossow W B, Kinne S, *et al.* GEWEX cloud assessment: a review [C]. AIP Conference Proceedings, 2013, 1531(1): 404-407.
- [2] Cihlar J, Howarth J. Detection and removal of cloud contamination from AVHRR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 1994, 32(3): 583-589.
- [3] Wang Q, Sun L, Wei J, *et al.* Improvement of universal dynamic threshold cloud detection algorithm and its application in high resolution satellite [J]. Acta Optica Sinica, 2018, 38(10): 1028002.  
王权, 孙林, 韦晶, 等. 动态阈值云检测算法改进及在高分辨率卫星上的应用 [J]. 光学学报, 2018, 38(10): 1028002.
- [4] Yang C J, Xu J M, Zhao F S. Application of time series in FY2C cloud detection [J]. Journal of Atmospheric and Environmental Optics, 2008, 3(5): 377-391.  
杨昌军, 许健民, 赵凤生. 时间序列在 FY2C 云检测中的应用 [J]. 大气与环境光学学报, 2008, 3(5): 377-391.
- [5] 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.
- [6] Tan K, Zhang Y J, Tong X. Cloud extraction from Chinese high resolution satellite imagery by probabilistic latent semantic analysis and object-based machine learning [J]. Remote Sensing, 2016, 8(11): 963.
- [7] Xu L N, Niu R Q, Fang S H, *et al.* Cloud detection based on decision tree over Tibetan Plateau with MODIS data [J]. Proceedings of SPIE, 2013, 8921: 89210G.
- [8] 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.
- [9] Hu X Y, Wang Y, Shan J. Automatic recognition of cloud images by using visual saliency features [J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(8): 1760-1764.
- [10] Ma Y J, Li X Y, Song X F. Traffic sign recognition based on improved deepconvolution neural network [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121009.  
马永杰, 李雪燕, 宋晓凤. 基于改进深度卷积神经网络的交通标志识别 [J]. 激光与光电子学进展, 2018, 55(12): 121009.
- [11] Liang L Y, Zhang T T, He W. Head pose estimation on multi-scale convolutional neural network [J]. Laser & Optoelectronics Progress, 2019, 56(13): 131003.  
梁令羽, 张天天, 何为. 多尺度卷积神经网络的头部姿态估计 [J]. 激光与光电子学进展, 2019, 56(13): 131003.

- 131003.
- [12] Young T, Hazarika D, Poria S, *et al.* Recent trends in deep learning based natural language processing [J]. *IEEE Computational Intelligence Magazine*, 2018, 13(3): 55-75.
- [13] Shi M Y, Xie F Y, Zi Y, *et al.* Cloud detection of remote sensing images by deep learning [C] // 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), July 10-15, 2016, Beijing, China. New York: IEEE, 2016: 701-704.
- [14] Mohajerani S, Krammer T A, Saeedi P. A cloud detection algorithm for remote sensing images using fully convolutional neural networks [C] // 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), August 29-31, 2018, Vancouver, BC, Canada. New York: IEEE, 2018: 18304815.
- [15] 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.
- [16] Mehta R, Sivaswamy J. M-net: a convolutional neural network for deep brain structure segmentation [C] // 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), April 18-21, 2017, Melbourne, VIC, Australia. New York: IEEE. 2017: 437-440.
- [17] 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. New York: IEEE, 2016: 770-778.
- [18] 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.
- [19] Cheng Q, Shen H F, Zhang L P, *et al.* Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 92: 54-68.
- [20] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [M] // Navab N, Hornegger J, Wells W, *et al.* Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. cham: springer, 2015, 9351: 234-241.
- [21] Yang M K, Yu K, Zhang C, *et al.* DenseASPP for semantic segmentation in street scenes [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 3684-3692.
- [22] Zhan Y J, Wang J, Shi J P, *et al.* Distinguishing cloud and snow in satellite images via deep convolutional network [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(10): 1785-1789.
- [23] Hughes M, Hayes D. Automated detection of cloud and cloud shadow in single-date landsat imagery using neural networks and spatial post-processing [J]. *Remote Sensing*, 2014, 6(6): 4907-4926.
- [24] 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.
- [25] Jia Y Q, Shelhamer E, Donahue J, *et al.* Caffe: convolutional architecture for fast feature embedding [C] // Proceedings of the ACM International Conference on Multimedia, November 3-7, 2014, Orlando, Florida, USA. New York: ACM, 2014: 675-678.
- [26] Kamran S A, Sabbir A S. Efficient yet deep convolutional neural networks for semantic segmentation [J/OL]. (2017-07-28) [2019-01-02]. <https://arxiv.org/abs/1707.08254>.