

doi: 10.6046/zrzyyg.2021128

引用格式: 刘广进,王光辉,毕卫华,等. 基于 DenseNet 与注意力机制的遥感影像云检测算法[J]. 自然资源遥感,2022,34(2): 88–96. (Liu G J, Wang G H, Bi W H, et al. Cloud detection algorithm of remote sensing image based on DenseNet and attention mechanism[J]. Remote Sensing for Natural Resources, 2022, 34(2): 88–96.)

基于 DenseNet 与注意力机制的遥感影像云检测算法

刘广进^{1,2}, 王光辉^{1,2}, 毕卫华³, 刘慧杰², 杨化超¹

(1. 中国矿业大学环境与测绘学院, 徐州 221116; 2. 自然资源部国土卫星遥感应用中心, 北京 100048; 3. 皖北煤电集团有限责任公司, 宿州 234002)

摘要: 遥感影像云检测是遥感影像处理过程中的第一步, 针对传统的云检测算法小块薄云检测效果差的问题, 该文提出了一种融合注意力机制的密集连接网络遥感影像云检测方法。首先, 将自然资源部国土卫星遥感应用中心提供的影像人工勾取云矢量并制作云标签, 再将其进行顺序裁剪、色彩抖动、旋转等预处理, 以增广样本量; 然后, 将预处理后的遥感影像及其标签一并输入到以 DenseNet 作为编码器与解码器的神经网络中, 编码器与解码器之间加入级联的空洞卷积模块以增大感受野, 双注意力机制与全局上下文建模模块以抑制一些无关的细节信息; 最后, 经过实验验证表明其精确率可以达到 95% 以上, 交并比可以达到 91% 以上, 较传统云检测算法有较大提高, 可以很好地提取小块薄云。

关键词: 云检测; DenseNet; 双注意力机制; 全局上下文建模模块; 空洞卷积

中图分类号: TP 751.1 **文献标志码:** A **文章编号:** 2097-034X(2022)02-0088-09

0 引言

遥感卫星的出现使得我们获取地表影像数据的能力大大提高, 凭借其全天时、大范围、时效性, 使得遥感技术广泛应用于农业生产、地质防灾、气象预报、军事预警、地理国情等领域。但是遥感影像的获取过程中云的存在是常见现象, 由于云的遮挡, 使得地面信息无法被卫星获取到, 给后续影像分析解译带来了困难^[1-2]。所以, 遥感影像云检测是首要而且必要的一环, 云检测的精度与速度对分析解译有直接影响。

云检测的方法大致可以分为 3 种: 手工勾绘方法、基于波段阈值及纹理信息的检测方法与基于深度学习的检测方法^[3]。其中第一种方法为最传统的方法, 效率较低, 本文不再赘述。第二种方法中基于波段阈值的方法需要人为监督, 调整参数才可以达到较优的检测结果, 基于纹理的检测算法需要后续对错误分类进行剔除。基于深度学习的云检测过程中不需要人工干预, 实现了端到

端的检测。基于波段阈值的云检测方法最早是 Zhu 等^[4]提出 FMask 算法, 通过设定波段阈值来实现 Landsat 卫星影像的云检测。在基于纹理的云检测方法中, Otsu 算法^[5]和 K-means 聚类法^[6]等都是基于邻域的相似性进行检测, 对噪声比较敏感, 鲁棒性较差。康一飞等^[7]利用高斯混合模型自适应获取影像的灰度阈值, 从而分离 ZY-3 影像的前景与背景, 实现云的检测, 但对地表上较亮的地物的误判比较严重; 仇一帆等^[8]使用 CFMask 算法检测 Landsat8 影像的云及云阴影, 然后用其结果代替人工勾绘的云标签参与深度学习训练, 再用训练出来的模型检测新图像的云及云阴影, 结果发现其检测结果比原始标签精度更高; 栗旭升等^[9]利用灰度共生矩阵提取图像的纹理特征, 使用纹理特征与光谱特征结合, 再使用支持向量机的方法对高分一号卫星影像进行云检测; 徐启恒等^[10]使用超像素分割方法与卷积神经网络相结合, 采用半监督的方式与传统的 Otsu 算法相比, 云检测精度有较大的提高; 刘云峰等^[11]将双重注意力机制模型与全卷积神经网络模型相对比, 发现双

收稿日期: 2021-04-23; 修订日期: 2021-07-14

基金项目: 国家重点研发计划项目“集成北斗/Galileo/LiDAR/倾斜摄影的智慧城市三维场景重建关键技术研究”(编号: 2017YFE0119600)资助。

第一作者: 刘广进(1998-), 男, 硕士研究生, 研究方向为遥感影像云检测。Email: 1538868186@qq.com。

通信作者: 王光辉(1982-), 男, 高级工程师, 研究方向为遥感影像信息提取与变化检测。Email: wanggh@lasac.cn。

注意力机制模型用于云检测更为准确;张家强等^[12]对 Unet 网络结构进行改进,编码器与残差结构相结合,提高了模型的泛化能力,比传统的 Unet 模型检测精度更优;张永宏等^[13]提出解码器不仅在 Unet 模型的编码器中引入残差模块,而且将密集连接模块融入解码器中,可以很好地检测出薄云大量的碎云;康超萌^[14]采用自适应超像素分割算法与支持向量机相结合的方法进行云检测,可以很好地检测出薄云,但在数据量大的情况下,效率较低;张晨等^[15]针对云与云阴影的误检现象,使用 ResNet 作为编码器与解码器,跳跃连接部分引入双重注意力机制,编码器与解码器之间使用改进的空洞金字塔池化模块以提取图像的多尺度特征,有效提高了云检测的精度。

随着深度学习算法的不断进步,越来越多的遥感问题都可以通过深度学习的方法得到有效解决。基于上述研究,本文提出一种融入注意力机制的密集连接网络,以解决深度学习算法检测小块云朵效果差的问题。首先,借鉴 D-LinkNet 网络^[16]的编码器解码器思想,使用 DenseNet^[17]作为编码器与解码器以训练到更深层的网络结构,能够提取更多特征;然后,中间层引入自注意力机制 GCNet-Block^[18]可以更好地提取上下文信息,引入双注意力模块^[19]增加有用信息的特征级别,降低无用信息的特征级别,使网络能够注意到有用的特征;最后,经过级联的空洞卷积结构,使网络在不改变特征图分辨率的前提下,增大感受野,更加有利于捕获影像的全局信息,将编码器阶段得到的特征图与解码器阶段上采样得到的特征图进行跳跃拼接,实现了对编码器阶段得到的特征的复用。

1 融入注意力机制的密集连接网络

1.1 DenseNet 结构

DenseNet 网络是 2017 年由 Huang 等^[20]提出, DenseNet 借鉴了 He 等^[21]提出 ResNet 的跳跃连接以减轻梯度消散现象,从而训练到更深层网络的思想,直接将任意 2 层进行跳跃连接,以求最大化减轻梯度消散问题,从而训练到更深的网络。ResNet 网络模型的公式为:

$$x_l = H(x_{l-1}) + x_{l-1}, \quad (1)$$

式中: x_{l-1} 为第 $l-1$ 层输出的特征图; H 为对特征图进行卷积、批归一化、激励等操作; x_l 为第 l 层输

出的特征图。

ResNet 模型结构如图 1 所示。

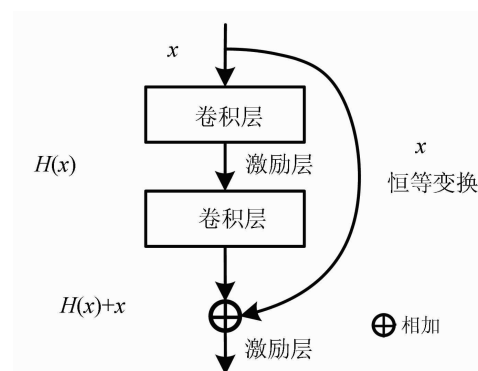


图1 ResNet-Block 结构

Fig.1 ResNet-Block structure

而 DenseNet 不同的是其互相连接所有的层,具体来说就是每个层都会接受其前面所有层作为其额外的输入,密集连接模块(densely connected block)结构如图 2 所示,其公式为:

$$x_l = H(x_0, x_1, \dots, x_{l-1}), \quad (2)$$

式中 $[x_0, x_1, \dots, x_{l-1}]$ 表示前 $l-1$ 层的输出的特征图的拼接。

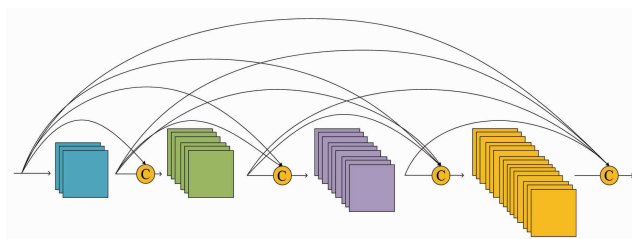


图2 密集连接模块

Fig.2 Densely connected block

虽然通过跳跃拼接的方式保留了原始的特征,减轻了梯度传播过程中消散现象,但随着网络层数越来越深,通道数也越来越大,参数量也越来越多,从而也难以训练到更深的网络。为此 DenseNet 还包含了一个重要的转换模块(transition block),如图 3 所示,用在密集连接模块之后,将得到的特征图的通道数减小为原来的一半。但仅仅这样还是会大大增加通道数,所以在密集连接模块中每次拼接之前都加入一个瓶颈(bottleneck)结构,将其特征图的通道数减小为增长率,这样就可以大幅度减小通道数。则经过一个密集连接模块之后的特征数就可以表示为:

$$C' = C + gn, \quad (3)$$

式中: C' 为经过密集连接模块之后的通道数; C 为经过密集连接模块之前的通道数; g 为通道数的增长率; n 为层数量。

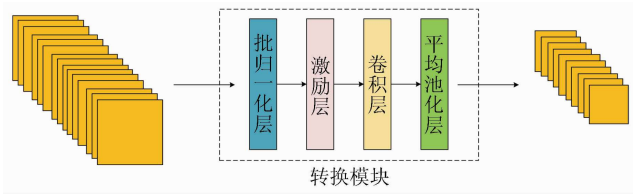


图3 转换模块

Fig.3 Transition block

通过密集连接模块后经过转换模块 (transition block) 可将其通道数降为原始的一半, 并进行下采样减半尺寸大小, 从而更加简化了计算量, 提升计算效率。

1.2 注意力机制

1.2.1 双注意力机制

随着计算机视觉的发展, 各种注意力机制相继被提出, 从而使计算机关注有用信息, 忽略不感兴趣的信息。注意力机制分为通道注意力机制与位置注意力机制。

通道注意力机制最早提出的是 SENet^[22], 通过压缩激励 (squeeze - exciting) 的方式对每个通道进行加权, 实现对信息量大的通道重点关注, 信息量少的通道关注力度更小。然而, 双注意力机制中的通道注意力模块不像 CBAM (convolutional block attention module)^[23] 中的通道注意力机制还需手工设计感知机, 更加方便了应用。双注意力机制的通道注意力模块如图4所示。

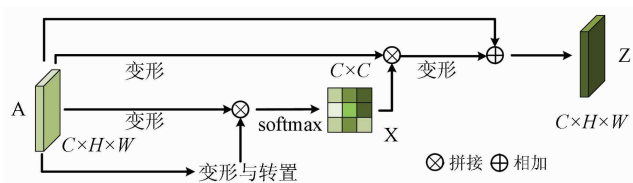


图4 通道注意力模块

Fig.4 Channel attention model

通道注意力模块的具体流程如下: 首先将特征图 A 变形为 $C \times N$ (其中 $N = H \times W$) 大小, 再转置成 $N \times C$ 大小, 与另一个只变形的特征图矩阵相乘得到 $C \times C$ 大小的特征图, 再经过 softmax 层, 得到不同通道的权值特征图 X, 再与变形过后的 A 进行矩阵相乘得到 $C \times N$ 大小的特征图, 再变形为 $C \times H \times W$ 大小的特征图, 最后与原始的 A 进行矩阵相加, 得到通道加权后的特征图 Z。

位置注意力模块的最早应用是 CBAM 中的空间注意力 (spatial attention model, SAM) 模块, 其首先将原始的特征图分别对同一位置不同通道进行全局最大池化与全局平均池化, 分别生成通道数为 1 的特征图, 然后将两者拼接起来, 继而通过一个卷积核尺寸为 7, 填充尺寸为 3 的卷积将拼接起来的特征

图的通道数降为 1, 从而达到获取其位置权重的目的。

而双注意力机制中的位置注意力模块则是使用通过自相关矩阵的变换得到位置权重, 可以更注意全局特征。其模型结构如图5所示。

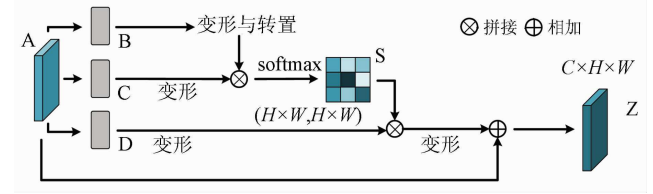


图5 位置注意力模块

Fig.5 Position attention model

其结构与通道注意力模块类似, 区别是其进行矩阵相乘时得到的特征图大小为 $(H \times W, H \times W)$, 将所有通道信息进行了压缩, 然后进行 softmax, 得到各个位置的权值 S。

1.2.2 全局上下文建模模块

双注意力机制可以融合通道注意力机制与位置注意力机制, 虽然其加入了自相关矩阵大大改善了注意力陷入局部的问题, 但充分利用全局的上下文信息是有必要的。所以在进行融合双重注意力机制之前, 先将全局的上下文信息提取出来。

为了解决此问题, 最早提出的是 NLNet (non - local net)^[24], 它利用自我注意机制建立远程依赖, 使网络能够更全面地理解图像, 从而不会使计算机局部感知图像。对于每个查询点, 首先计算该点与所有点的点对关系, 得到注意力图, 然后对所有点的特征进行加权求和, 得到与查询端相关的全局特征, 最后将全局特征加入到每个查询点的特征中, 完成远程依赖关系的建模 (图6)。

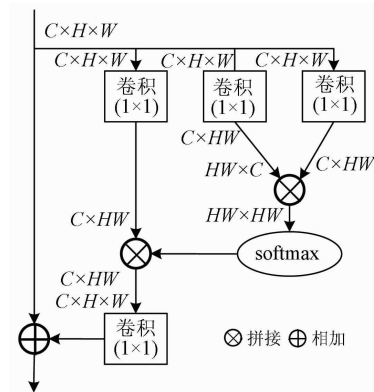


图6 非局部神经网络

Fig.6 Non - local neural networks

但是由于 NLNet 内部有较多的矩阵变换运算, 导致计算量大大增加, 所以极大地限制了 NLNet 的使用, 为此 GCNet 提出了简化 NLNet 版本, 即 Simple - NLNet, 并结合了 SENet 计算量少的优点, 从而不仅

有效地建立远程依赖,又节省了网络结构的计算量。
GCNet 的 GC - Block 模块如图 7 所示。

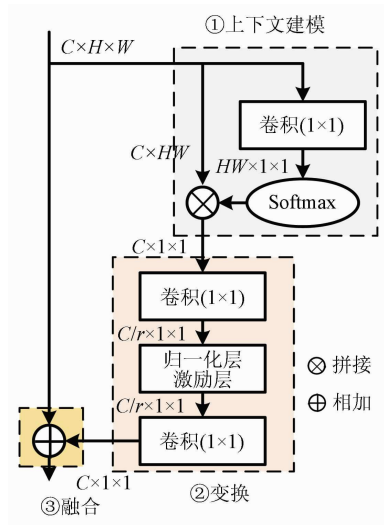


图 7 全局上下文建模模块

Fig. 7 Global context block

对于 GCNet 其可以分为 3 个步骤: 首先是需要使用全局平均池化进行上下文建模 (context modeling), 原始大小为 $C \times H \times W$ 一方面先将其变形成 $C \times HW$ 大小的特征图, 另一方面使用 1×1 卷积降低通道数为 1, 生成 $1 \times H \times W$ 大小的特征图, 再变形成 $HW \times 1 \times 1$ 大小的特征图, 经过 softmax 层后与 $C \times HW$ 大小的特征图进行矩阵相乘, 得到 $C \times 1 \times 1$ 大小的特征图, 至此每个通道的特征图的大小都变为 1×1 , 完成了上下文建模。第二步捕获通道间的依赖, 先将 $C \times 1 \times 1$ 大小的特征图经过一个瓶颈结构, 降低通道数, 图中 r 为通道数减少的倍数, 从而减小了计算量, 然后加入归一化层提高泛化能力, 使用 ReLU 函数进行激励, 得到非线性关系, 最后还原

通道数。第三步是将原始的特征图与变换后的特征图使用相加操作进行融合。

1.3 空洞卷积模块

感受野表示单个像素点在上层卷积网络特征图中的区域范围大小, 感受野越大, 则表示通过卷积得到一个像素点所使用的上层特征图的像素个数越多, 即所使用到的特征越多。感受野的大小决定了网络能否获取特征图的全局信息。扩大感受野最简单的方法就是增大卷积核的大小, 但是卷积核大小的增大一定会使计算量增加, 所以随着 Deeplab V1^[25] 提出的空洞卷积, 即扩大卷积核并在卷积核中使用部分用 0 来填充, 从而可以简单方便地解决计算量增加的问题。如图 8 所示, 特征图分为同等的 6 份, 第一层为分别使用 1, 2, 4, 8, 16 大小空洞率进行空洞卷积, 后面 5 层类似, 最后将空洞卷积后的特征图与最后一层的原始特征图相加, 从而可以增大感受野。

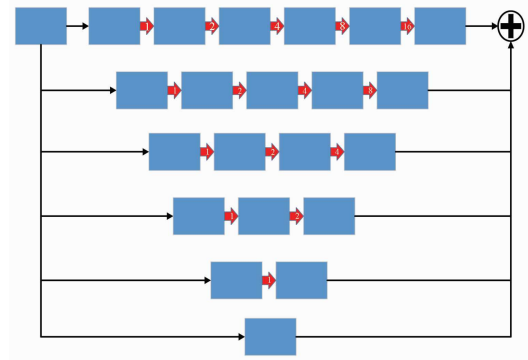


图 8 空洞卷积模块

Fig. 8 Dilated - convolution block

1.4 融合注意力机制的密集连接网络

在上述研究的基础上, 本文提出融合注意力机制的密集连接网络, 具体模型结构如图 9 所示。

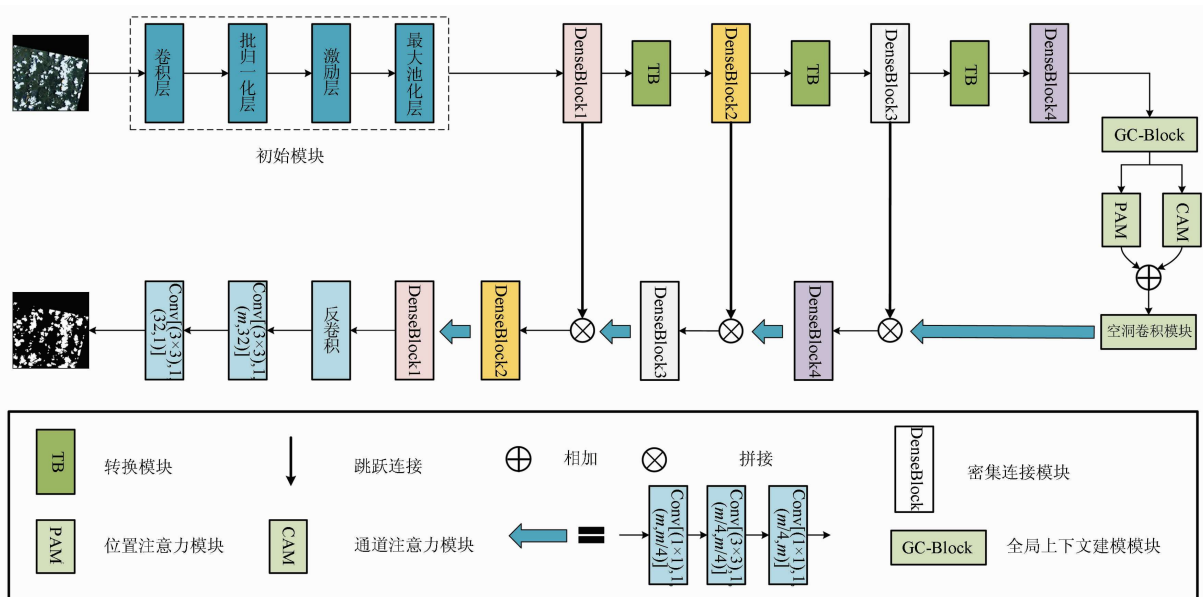


图 9 融入注意力机制的密集连接网络

Fig. 9 Attention DBlock densely connected networks

1) 编码器阶段: 将裁剪好的 3 通道影像经过一个初始模块, 初始模块由 4 个部分组成, 第一部分为卷积核大小为 7, 填充大小为 3 的卷积层以不改变特征图的大小, 输出通道数为 64 以减弱网络对影像失真影响; 第二部分为批归一化层, 其作用加速模型收敛以及一定程度上缓解深层网络的梯度弥散问题; 第三部分为激励层, ReLU 可以增加网络的非线性, 并使一部分神经元的输出变为 0, 减少了参数之间的相互依赖, 防止过拟合; 第四部分为最大池化层, 最大池化层将原始特征图的宽高尺寸减半, 减少了参数量, 并一定程度上增加了网络的非线性关系。然后分别经过 4 个密集连接模块与转换模块提取特征, 密集连接模块有效地减弱了梯度消散问题, 转换模块有效地减少了参数量, 使网络不至于“参数爆炸”。其中 4 个密集连接模块的层数量分别为 6, 8, 12, 16。

2) 中间层: 首先通过 GC-Block 提取特征图的上下文信息, 双注意力机制融合了位置注意力与通道注意力信息, 使计算机更好地注意需要的特征, 再将结果送入空洞卷积模块中使其在不改变特征图分辨率的前提下, 增大感受野, 进一步提取特征图的全局特征。

3) 解码器阶段: 解码器部分首先上采样使特征图的通道数减小为刚通过编码器阶段 DenseBlock3 的特征图的通道数, 图中 m 为当前层特征图的通道数; 然后与其进行跳跃拼接, 从而特征图通道数变为原来的 2 倍, 跳跃拼接可以实现特征复用, 保留了更多的原始特征; 经过 DenseBlock 之后, 通道数增加, 再通过上采样降低通道数, 直到最后将特征图变为通道数为 1 的二值图。

2 实验与分析

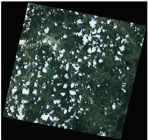
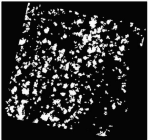
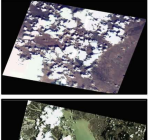
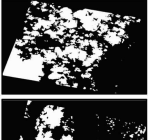
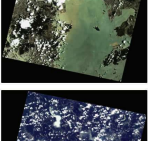
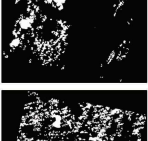
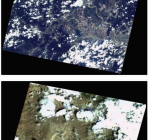
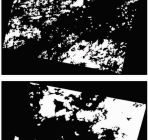
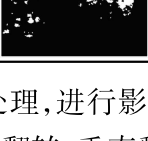
2.1 实验数据准备与参数设置

本文的实验均在 ubuntu16.04, CUDA 10.1, NVIDIA UNIX 64 核的 pytorch1.2.0 深度学习框架下进行的。

本次实验的数据来自自然资源部国土卫星遥感应用中心提供的国产卫星遥感影像, 其中包括高分一号、高分二号、高分六号、资源卫星影像, 由于我国中东部地区云量相对较小, 所以大部分选取新疆、西藏及内蒙古等边疆地区的遥感影像, 小部分选取中东部地区的遥感影像, 共 596 景, 比例关系约为 6:4。标签制作为手工勾绘云矢量, 后使用自制 python 脚

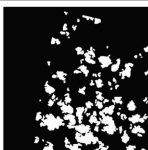
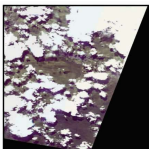
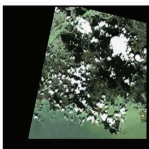
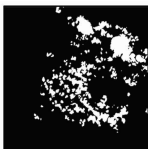
本将其转化为二值图标签。原始标签及其真值标签部分数据如表 1 所示。

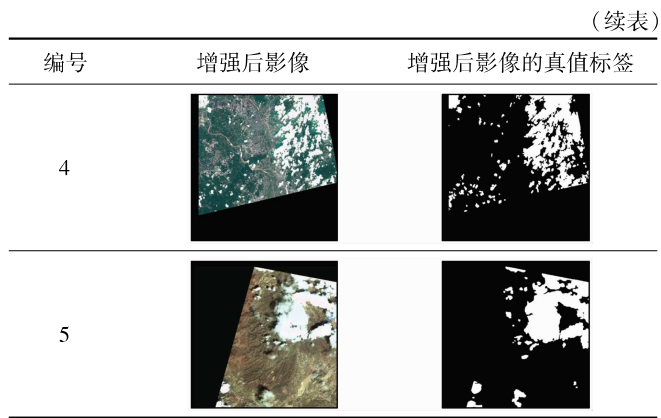
表 1 原始影像及其真值标签
Tab.1 Original image and its truth tag

编号	原始影像	真值标签	影像类型
1			草地
2			荒地
3			水域
4			城镇
5			山地

训练之前首先经过预处理, 进行影像数据的扩充, 方法有: 色彩抖动、水平翻转、垂直翻转、顺序裁剪、旋转等。首先进行顺序裁剪, 将原始的影像与标签裁剪成 512 像素 × 512 像素大小的影像共 2 468 张, 再使用色彩抖动、平移、旋转、尺度缩放、水平翻转、垂直翻转进行扩充数据集, 共得到 12 340 张, 按照 8:2 的比例分为训练集 9 872 张与测试集 2 468 张。预处理的目的是扩充数据量, 对数据进行增广, 防止过拟合, 提高模型的泛化能力。增强后的影像及其真值标签见表 2。

表 2 增强之后的影像及真值标签
Tab.2 Enhanced image and its truth tag

编号	增强后影像	增强后影像的真值标签
1		
2		
3		



$$BCELoss(P,GT) = - \sum_{i=1}^{W'} \sum_{j=1}^{H'} [gt_{ij} \ln p_{ij} + (1 - gt_{ij}) \ln (1 - p_{ij})] \quad , \quad (5)$$

式中： GT 为标签影像 (ground true)； P 为预测影像 (predicet mask)； B 为批大小； W' 为影像的宽度； H' 为高度； gt_{ij} 为标签影像在 i,j 位置的像素值； p_{ij} 为预测影像在 i,j 位置的像素值。

2.3 评价指标

这里选取的指标为交并比 IoU 、召回率 $recall$ 以及精确率 $precision$ 。其计算方式分别为：

$$IoU = \frac{TP}{TP + FP + FN} \quad , \quad (6)$$

$$recall = \frac{TP}{TP + FN} \quad , \quad (7)$$

$$precision = \frac{TP}{TP + FP} \quad , \quad (8)$$

式中： TP 为真阳性，表示原本为云预测为云的数量； TN 为真阴性，表示原本为非云预测为非云的数量； FP 为假阳性，表示原本为非云错误预测为云的数量； FN 为假阴性，表示原本为云错误预测为非云的数量。其中 3 个指标越高表示精度越高。

2.4 云检测结果分析

由于影像不是多光谱影像，不能利用以热红外波段实施云检测的 FMask 算法进行检测，所以传统方法分别使用 Otsu 阈值算法、Otsu 多阈值算法及 K - means 聚类算法进行云检测。深度学习算法分别使用 SegNet, Unet, D - LinkNet50，未加注意力机制的本文方法 (D - DenseNet) 及本文方法 (AD - DenseNet) 对增强后的影像进行训练。首先进行定量分析，其中在训练集上的不同算法模型评价指标表现如表 3 所示，验证集上的不同算法模型的评价指标表现如表 4 所示，从表 3—4 中可以看出，传统算法的精度较深度学习算法的精度低，而在深度学习算法中，无论在训练集上还是在验证集上的召回率、交并比以及精确率都

参数设置：batchsize 为 4，步长为 20，使用的优化器为 Adam，初始学习率为 0.001，自动调整学习率，每次调整变为原来的 1/2，直到损失值 7 次不再下降为止。

2.2 损失函数

由于这是针对二分类问题，所以使用 BCELoss 作为损失函数。BCELoss 是二分类问题中优秀的损失函数，其公式为：

$$BCELoss = \sum_{i=1}^B BCELoss(P_i,GT_i) \quad , \quad (4)$$

为本文方法精度最高。加上注意力机制后交并比增长约一个百分点，召回率与精确率也都有小幅度增长。

表 3 训练集上的不同算法模型评价结果

Tab.3 Evaluation results of different algorithm models on training sets

算法模型	Recall	Precision	IoU
Otsu 阈值法	0.865 706	0.341 597	0.324 407
Otsu 多阈值法	0.684 431	0.833 832	0.602 283
K - means 聚类法	0.838 212	0.385 039	0.358 404
SegNet	0.908 152	0.936 953	0.855 852
Unet	0.945 295	0.954 285	0.904 343
D - LinkNet50	0.942 009	0.948 958	0.896 580
D - DenseNet	0.947 085	0.955 037	0.906 659
AD - DenseNet	0.948 931	0.957 636	0.910 701

表 4 验证集上的不同算法模型评价结果

Tab.4 Evaluation results of different algorithm models on the validation sets

算法模型	Recall	Precision	IoU
Otsu 阈值法	0.859 295	0.389 704	0.366 328
Otsu 多阈值法	0.702 932	0.868 563	0.635 349
K - means 聚类法	0.841 581	0.423 713	0.392 414
SegNet	0.923 405	0.926 182	0.860 104
Unet	0.946 430	0.958 225	0.908 927
D - LinkNet50	0.948 262	0.949 005	0.902 285
D - DenseNet	0.951 485	0.958 911	0.904 211
AD - DenseNet	0.953 866	0.961 321	0.918 611

图 10 为 AD - DenseNet 算法的损失值与交并比随着训练轮数 $epoch$ 的变化，其中一个 $epoch$ 表示全部数据训练一次。一共训练了 120 个 $epoch$ ，训练集与测试集上的交并比均高于 0.91，并且最后都达到收敛。图 11 为 AD - DenseNet 算法的精确率与召回率随着 $epoch$ 的变化曲线，可以发现最后均收敛，并且精确率都在 0.95 以上。不同算法云检测结果如图 12 所示。

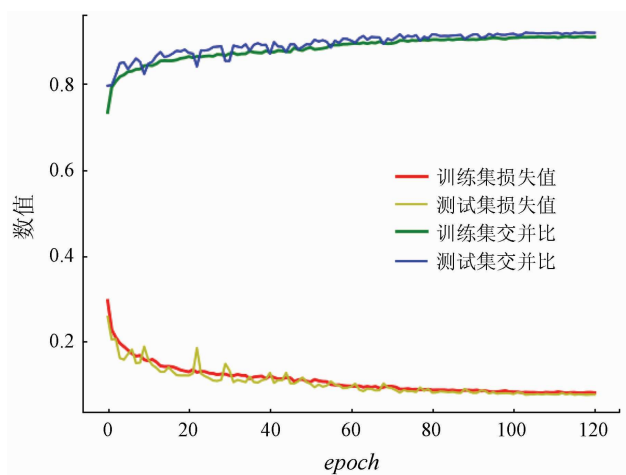


图 10 AD-DenseNet 算法损失值与交并比随 epoch 的变化曲线
Fig. 10 Change curve of AD-DenseNet algorithm loss and IoU with epoch

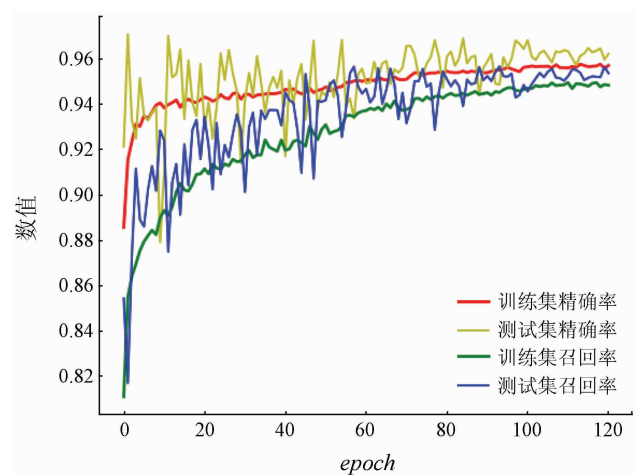


图 11 AD-DenseNet 算法精确率与召回率随 epoch 的变化曲线
Fig. 11 Change curve of AD-DenseNet algorithm precision and recall with epoch

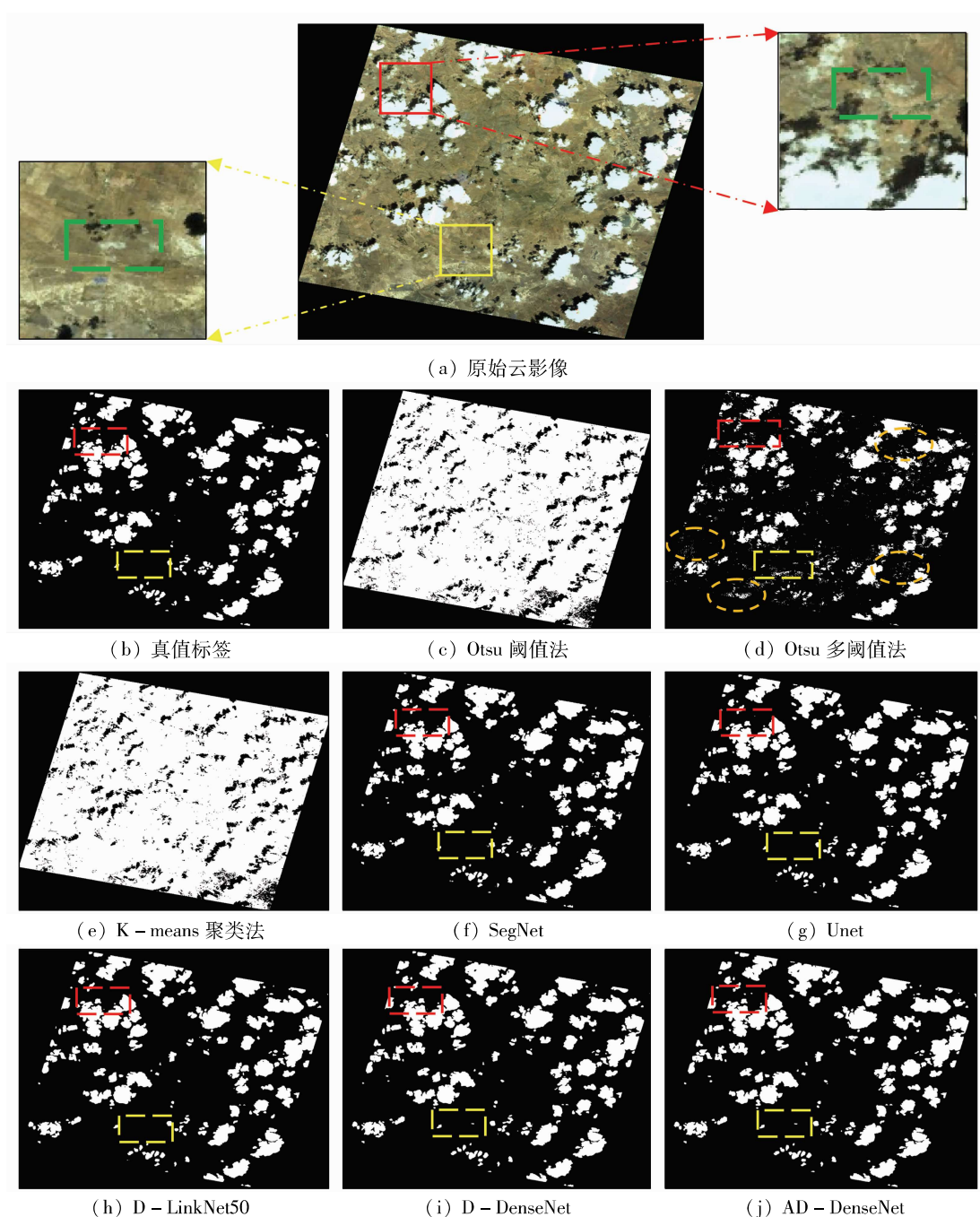


图 12 不同算法的云检测结果对比
Fig. 12 Comparison of cloud detection results of different algorithms

首先使用传统方法 Otsu 阈值法、Otsu 多阈值法及 K-means 聚类法对原始影像进行云检测,结果如图 12(c)~(e)所示,从中可以看到 Otsu 阈值法与 K-means 聚类法检测效果明显不佳,由于其对噪声鲁棒性较差,Otsu 多阈值法可以较好地检测出云的范围,但是从图 12(d)中橙色椭圆虚线框中可以看到有明显的错检现象。对于深度学习方法,分别使用 SegNet, Unet, D-LinkNet50, D-DenseNet 及 AD-DenseNet 进行预测测试集中的影像,首先,将原始影像进行顺序切分为 512 像素 × 512 像素大小的小影像,再使用训练好的网络模型进行预测,最后按照规则进行还原成原始影像大小。纵观图 12 的(f)~(j),其大体上没什么差别,预测精度都较高,但对于小块云朵的检测却不是如此。从图 12(a)中可以看到,原始影像在红色框中有 2 片较小的云,由于其后方有明显的云的阴影,所以判定其为云,真值标签只勾绘出了一片云,SegNet, Unet, D-LinkNet 及 D-DenseNet 均未预测出 2 片小块云,但 AD-DenseNet 算法却准确预测出其位置及形状。图 12(a)中黄色框中,有一块较小的薄云,放大之后可清晰看到其云阴影,其中 SegNet, Unet 和 D-LinkNet 均未预测出此小块云,D-DenseNet 与 AD-DenseNet 正确预测出其位置及形状。

3 结论

本文针对传统的深度学习算法不能很好地检测出小块云的问题,提出了一种融合注意力机制的密集连接网络(attention DBlock densely connected networks, AD-DenseNet)。

1) 首先从编码器与解码器结构出发,不再使用通用的 ResNet-Block 结构作为骨干网络,使用 DenseNet-Block 作为编码器与解码器,从而可以训练到更深层的网络,提取到更多的影像特征。

2) 中间层使用 GC-Block 提取全局特征,防止网络陷入局部,引入双注意力机制,使网络更加注意有用信息,抑制无关信息。

3) 中间层加入 DBlock 结构,使网络在不改变分辨率的情况下增大感受野,以提取全局特征。

从实验结果来看,交并比可以达到 0.91 以上,精确率可以达到 0.95 以上,本文方法可以很好地检测出小块云朵,较传统方法有所提高。

但是对于雪的误判还存在问题,这将是下一步需要研究的方向。

参考文献 (References):

[1] Zhang Y, Rossow W B, Lacis A, et al. Calculation of radiative flu-

xes from the surface to top of atmosphere based on ISCCP and other global data sets; Refinements of the radiative transfer model and the input data [J]. *Journal of Geophysical Research: Atmospheres*, 2004, 109(d19): 105-115.

[2] 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*, 2007, 112(3): 1196-1211.

[3] Xie F, Shi M Y, Shi Z, 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: 3631-3640.

[4] Zhu Z, Woodcock C E. Object-based cloud and cloud shadow detection in Landsat imagery [J]. *Remote Sensing of Environment*, 2012, 118: 83-94.

[5] Otsu N. A Threshold selection method from gray-level histograms [J]. *IEEE Transactions on Systems Man and Cybernetics*, 2007, 9(1): 62-66.

[6] 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.

Li Y G. Improvement of K-means clustering algorithm and its application in color image segmentation [D]. Hefei: Anhui University, 2018.

[7] 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.

Kang Y F, Pan L, Sun M W, et al. Gaussian mixture model based cloud detection for Chinese high resolution satellite imagery [J]. *Geomatics and Information Science of Wuhan University*, 2017, 42(6): 782-788.

[8] 仇一帆, 柴登峰. 无人工标注数据的 Landsat 影像云检测深度学习算法 [J]. *国土资源遥感*, 2021, 33(1): 102-107. doi: 10.6046/gtzyyg. 2020090.

Qiu Y F, Chai D F. A deep learning method for Landsat image cloud detection without manually labeled data [J]. *Remote Sensing for Land and Resources*, 2021, 33(1): 102-107. doi: 10.6046/gtzyyg. 2020090.

[9] 栗旭升, 刘玉峰, 陈冬花, 等. 结合图像特征的支持向量机高分一号云检测 [J]. *国土资源遥感*, 2020, 32(3): 55-62. doi: 10.6046/gtzyyg. 2020. 03. 08.

Li X S, Liu Y F, Chen D H, et al. Cloud detection based on support vector machine with image features for GF-1 data [J]. *Remote Sensing for Land and Resources*, 2020, 32(3): 55-62. doi: 10.6046/gtzyyg. 2020. 03. 08.

[10] 徐启恒, 黄滢冰, 陈 洋. 结合超像素和卷积神经网络的国产高分辨率遥感影像云检测方法 [J]. *测绘通报*, 2019(1): 50-55.

Xu Q H, Huang Y B, Chen Y. Cloud detection for Chinese high resolution remote sensing imagery using combining superpixel with convolution neural network [J]. *Bulletin of Surveying and Mapping*, 2019(1): 50-55.

[11] 刘云峰, 杨 珍, 韩 骁, 等. 国产高分辨率遥感卫星影像云检测方法分析 [J]. *测绘通报*, 2020(11): 66-70.

Liu Y F, Yang Z, Han X, et al. Research on cloud detection method of domestic high-resolution satellite images [J]. *Bulletin of Surveying and Mapping*, 2020(11): 66-70.

[12] 张家强, 李潇雁, 李丽圆, 等. 基于深度残差全卷积网络的 Landsat8 遥感影像云检测方法 [J]. *激光与光电子学进展*, 2020, 57(10): 364-371.

Zhang J Q, Li X Y, Li L Y, et al. Landsat8 remote sensing image

- based on deep residual fully convolutional network[J]. *Laser and Optoelectronics Progress*, 2020, 57(10): 364–371.
- [13] 张永宏, 蔡朋雁, 陶润喆, 等. 基于改进 U-Net 网络的遥感图像云检测[J]. *测绘通报*, 2020(3): 17–20.
Zhang Y H, Cai P Y, Tao R Z, et al. Cloud detection for remote sensing images using improved U-Net[J]. *Bulletin of Surveying and Mapping*, 2020(3): 17–20.
- [14] 康超萌. 基于神经网络的国产高分光学遥感图像云检测[D]. 西安: 中国科学院大学(中国科学院西安光学精密机械研究所), 2018: 20–27.
Kang C M. Cloud detection in domestic high-resolution remote sensing image based deep neural networks[D]. Xi'an: University of Chinese Academy of Sciences, 2018: 20–27.
- [15] 张晨, 张秀再, 杨昌军. 双注意力 RDA-Net 遥感影像云与云阴影检测方法[J/OL]. *激光与光电子学进展*, 2021: 1–19. <https://kns.cnki.net/kcms/detail/31.1690.TN.20210222.1114.023.html>.
Zhang C, Zhang X Z, Yang C J. Dual attention RDA-Net for cloud and cloud shadow detection of remote sensing image[J/OL]. *Laser and Optoelectronics Progress*, 2021: 1–19. <https://kns.cnki.net/kcms/detail/31.1690.TN.20210222.1114.023.html>.
- [16] Zhou L C, Zhang C, Wu M. D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018: 182–186.
- [17] 彭博. 基于深度学习的遥感图像道路信息提取算法研究[D]. 成都: 电子科技大学, 2019: 45–46.
Peng B. Research on road information extraction of remote sensing image based on deep learning[D]. Chengdu: University of Electronic Science and Technology of China, 2019: 45–46.
- [18] Cao Y, Xu J R, Stephen L, et al. GCNet: Non-local networks meet squeeze-excitation networks and beyond[C]//IEEE CVF International Conference on Computer Vision (ICCV), 2019: 1971–1980.
- [19] Fu J, Liu J, Tian H, et al. Dual attention network for scene segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2019: 3146–3154.
- [20] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 2017.
- [21] He K M, Zhan X Y, Ren S Q. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.
- [22] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 7132–7141.
- [23] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]//15th European Conference on Computer Vision (ECCV), 2018, 11211: 3–19.
- [24] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7794–7803.
- [25] 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.

Cloud detection algorithm of remote sensing image based on DenseNet and attention mechanism

LIU Guangjin^{1,2}, WANG Guanghui^{1,2}, BI Weihua³, LIU Huijie², YANG Huachao¹

(1. School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China; 2. Land Satellite Remote Sensing Application Center, MNR, Beijing 100048, China; 3. Wanbei Coal and Electricity Co. Ltd., Suzhou 234002, China)

Abstract: The cloud detection of remote sensing images is the first step in the process of remote sensing image processing. To address the problem that the traditional cloud detection algorithm has a poor effect on the detection of small and thin clouds, this paper proposes a cloud detection method for densely connected network remote sensing images based on the attention mechanism. First, cloud vectors are manually checked from the images provided by the Land Satellite Remote Sensing Application Center of the Ministry of Natural Resources and cloud labels are made, and the images are preprocessed by sequential clipping, color jitter, rotation, etc. to enlarge the sample size. Then, the pre-processed remote sensing images and their labels are fed into a neural network with DenseNet as the encoder and decoder, and a cascaded atrous convolution module is added between the encoder and decoder to increase the receptive field, and a dual attention mechanism and a global context modeling module are added to suppress some irrelevant detailed information. Finally, the experimental results showed that the accuracy rate could reach 95% and the intersection over union could reach 91%, which are big improvements over the traditional cloud detection algorithm, and this method performs well in extracting small and thin clouds.

Keywords: cloud detection; DenseNet; dual attention mechanism; global context modeling module; atrous convolution