

基于多尺度特征的地基云图分类检测算法

孙继飞, 贾克斌

引用本文

孙继飞, 贾克斌. 基于多尺度特征的地基云图分类检测算法[J]. 计算机科学, 2024, 51(6A): 230400041-6.

SUN Jifei, JIA Kebin. Classification and Detection Algorithm of Ground-based Cloud Images Based on Multi-scale Features [J]. Computer Science, 2024, 51(6A): 230400041-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[融合注意力机制的DeeplabV3+ 服装图像分割方法](#)

Clothing Image Segmentation Method Based on Deeplabv3+ Fused with Attention Mechanism
计算机科学, 2024, 51(6A): 230900153-7. <https://doi.org/10.11896/jsjcx.230900153>

[方向感知金字塔聚合网络的道路中心线提取](#)

Direction-aware Pyramidal Aggregation Network for Road Centerline Extraction
计算机科学, 2024, 51(6A): 230400101-7. <https://doi.org/10.11896/jsjcx.230400101>

[基于双流卷积神经网络的稻米缺陷分割](#)

Rice Defect Segmentation Based on Dual-stream Convolutional Neural Networks
计算机科学, 2024, 51(6A): 230600107-8. <https://doi.org/10.11896/jsjcx.230600107>

[基于注意力机制与密集邻域预测的轻量化图像语义分割](#)

Lightweight Image Semantic Segmentation Based on Attention Mechanism and Densely AdjacentPrediction
计算机科学, 2024, 51(6A): 230300204-8. <https://doi.org/10.11896/jsjcx.230300204>

[基于图卷积神经网络的点云语义分割综述](#)

Review of Point Cloud Semantic Segmentation Based on Graph Convolutional Neural Networks
计算机科学, 2024, 51(6A): 230400196-7. <https://doi.org/10.11896/jsjcx.230400196>

基于多尺度特征的地基云图分类检测算法

孙继飞 贾克斌

北京工业大学信息学部 北京 100024

先进信息网络北京实验室 北京 100024

(kidsun@emails.bjut.edu.cn)

摘要 地基云的自动识别方法和技术为气象分析中的云状识别和云量估计任务提供了重要的手段和依据。然而,对这两种任务的研究往往独立,互不相干,导致地基云图的分类与分割技术无法有效地结合使用。特别是当云图中出现多类云状时,现有技术难以按不同云类分别划分区域并进行云量计算。为了解决这一问题,提出用基于深度学习的语义分割方法实现对地基云图的按类分割。首先,构建了地基云图语义分割数据集 GBCSS,该数据集包含 3000 幅云图,共计 11 个类别。在此基础上,提出了一种基于 U 型神经网络的改进方案 UNet-PPM 作为地基云图语义分割模型。为了增强网络对云的轮廓特征提取能力,引入了金字塔池化模块。该模块提取并聚合了不同尺度的图像特征,提升了网络获取全局信息的能力。最后,将设计的网络在 GBCSS 上进行了训练以及评估,其在测试集上达到了 91.5% 的像素准确率。与 U-Net 相比,UNet-PPM 在像素准确率上有 5.4% 的提升,表明该网络对云的轮廓特征提取的能力更强,以及语义分割应用在地基云图中的可行性。

关键词: 地基云图;语义分割;云图数据集;全卷积网络;金字塔池化模块

中图分类号 TP183;TP751

Classification and Detection Algorithm of Ground-based Cloud Images Based on Multi-scale Features

SUN Jifei and JIA Kebin

Faculty of Information Technology, Beijing University of Technology, Beijing 100024, China

Beijing Laboratory of Advanced Information Networks, Beijing 100024, China

Abstract Clouds constantly contribute significantly to climate change in addition to having a short-term impact on local temperatures. To study local cloud details, the ground-based observation is used caused by its ability of cloud image capture in high temporal and spatial resolution. The research on automatic identification of ground-based clouds is primarily focused on two areas: cloud classification and cloud detection. Traditionally, both of them are regarded as separate and unrelated tasks. Cloud classification are independent of the segmentation, and most segmentation techniques focus on binary segmentation. This making it difficult to segment regions by different cloud types when the cloud image contains multiple classes of clouds. To address this problem, this paper proposes a semantic segmentation method based on deep learning for the combination of two tasks. First, it constructs the ground-based cloud image semantic segmentation(GBCSS) dataset, which contains 3000 cloud images with a total of 11 types. All images are resized to a square format of 256×256 pixels. Then, an improved scheme based on U-shaped neural networks is designed as the semantic segmentation model for ground-based cloud images. The pyramid pooling module is combined for extracting and aggregating image features at different scales. This module improves the network's ability to obtain global information. The developed network UNet-PPM achieves 91.5% pixel accuracy on average on the test set after being trained and assessed on GBCSS. Our suggested enhanced method outperforms the U-Net, Deeplabv3+, DANet and BiSeNetv2 in terms of pixel accuracy. Experiment results show that the pyramid pooling module contributes a lot to extract cloud contour features and restrain the overfitting problem. Our work show the feasibility of semantic segmentation application in cloud image automatic observation.

Keywords Ground-based cloud image, Semantic segmentation, Cloud image dataset, Fully convolutional network, Pyramid pooling module

1 引言

云作为一种重要的天气现象,时刻影响着大气辐射、水文循环以及全球气候^[1]。不同的水汽条件和大气运动会生成不

同形态和强度的云。借助这一点,人们会观察云的生消演变以预测雨、雪、雹、雷电等天气活动。因此,云的观测一直都是一项重要而又困难的任务。目前,云的成像数据根据不同的观测手段可分为卫星云图和地基云图。卫星云图的拍摄尺度

基金项目:北京市自然科学基金(4212001)

This work was supported by the Beijing Natural Science Foundation, China(4212001).

通信作者:贾克斌(kebinj@bjut.edu.cn)

范围大,但其局部区域的分辨率低,不利于对云进行精细识别。而地基云图具有更高的时空分辨率,包含了云的形状、结构、纹理等细节,对局部区域内的天气预测起到关键作用。过去,对地基云图的分析往往需要经验丰富的观测员从图像中目测云体信息。而近年来随着计算机视觉技术的不断发展,研究人员开始设计能够从地基云图中自动提取云体信息的技术。与人工观测相比,自动化观测的推理速度更快,且允许高频稠密观测;观测结果准确可定量,不受主观意识干扰;观测手段更容易操控,观测成本更低。

国内外对地基云图的自动化观测任务的研究主要聚焦于云状信息与云量信息的提取,即地基云图分类任务和分割任务。一般而言,云图分类任务的输出是单个或多个云状名称标签,如图 1(a)所示。在卷积神经网络(Coavalutional Neural Network,CNN)大范围使用之前,该任务的研究重心被放在了不同云状的特征提取上。例如,Zhu 等^[2]提取了地基云图中云的纹理特征、颜色特征以及形状特征,通过 K 最近分类器对积状云、层状云、卷云进行了分类,其分类器的平均准确率为 83.9%。之后,基于 CNN 设计的 Alexnet^[3]取得了 2012 年 ImageNet 图像分类竞赛的冠军。由于 CNN 无需人工设计特征提取器,且泛用性强,因此 CNN 开始被广泛用于解决各领域的图像分类问题,其中就包括地基云图分类任务。例如,文献[4]提出的一种结合了注意力机制的 CNN 在包含 7 类云的 NREL(National Renewable Energy Laboratory)地基云图数据集上具有 95.6% 的准确率,体现了 CNN 在地基云图分类任务中的优势所在。而云图分割任务则是将地基云图的每一个像素分类为云或者背景,从而得到一幅二值分割图,如图 1(b)所示。过去,最常用的分割法为阈值分割法。例如,Saito 等^[5]提出了基于 RGB 通道定义的天空指数,以此为阈值去分割云层与天空,达到了 93.6% 的像素准确率。2015 年,Long 等^[6]提出全卷积网络(Fully Convolutional Network,FCN)架构,实现了端到端的语义分割。于是,越来越多的 FCN 架构被用于地基云图分割任务中。例如,Zhang 等^[7]提出了针对地基云图分割的 LGCSegNet,其在 HBMCD_GT 数据集上达到了 96.83% 的像素准确率。

然而,目前提出的地基云图分类与分割任务的解决方案仍存在局限性。实际观测中,一幅地基云图可能存在多种不同类型的云状,而目前的地基云图分类算法大多只能识别出云量占比最大的云状。另一方面,地基云图分割算法也只能将云图分割为云和背景两类,并不能在同一幅图存在多种云状时按类分割。我们认为,这样的处理方式不能满足如今精准预测复杂多变的天气的需求。气象分析中,更加高效的方法在面对含有多种云状的地基云图输入时,应该能够按照不同的云状类别划分,输出不同云状各自的分布情况,并通过精确各类云状的分割统计出云量情况,如图 1(c)所示。然而目前单一云图的分类和分割算法并不能解决该问题,而国内外对这种多云状分割的研究工作也鲜有报道^[8]。

为此,本文提出了一种针对地基云图语义分割的网络模型 UNet-PPM,该网络具有与一般视觉场景下的语义分割类似的效果,实现了对地基云图的按多种云类分割。从另一个角度看,UNet-PPM 既输出了云量信息,又输出了云状信息,

因此我们的工作将云图分类与云图分割这两个互不关联的任务统一了起来。此外,针对目前尚未存在公开的可用于地基云图语义分割的数据集的问题,本文构建了 GBCSS(Ground-Based Cloud Semantic Segmentation)数据集。该数据集包含 3000 幅云图,依照世界气象组织的标准将云分为 10 类,每幅云图都带有对应的像素级标签,可用于地基云图语义分割的模型训练及评估。

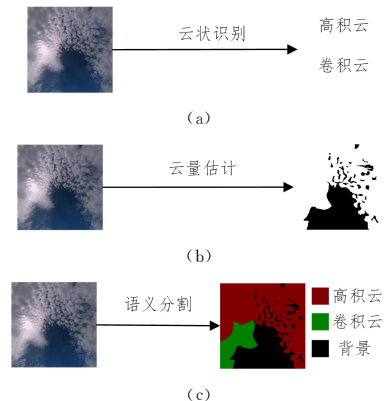


图 1 地基云自动化观测任务样例

Fig. 1 Examples of automatic observation for ground-based cloud

2 地基云图数据集

任何基于监督学习训练的神经网络都离不开大量的标注数据。通常,训练数据越多,标注越精确,则模型的准确率越高^[9]。尽管目前有一些公开的与地基云图相关的数据集^[10-12],但这些数据集的标签不是仅包含云状类别,就是仅区分了云和背景区域。它们只能用于上文提到的分类或者二值分割任务,不能满足地基云图语义分割数据集的要求。为此,我们构建了一个全新的基于地基云图的语义分割数据集 GBCSS。

GBCSS 数据集中的地基云图由华云升达(北京)气象科技有限责任公司的全天空成像仪拍摄后经矫正筛选得到。为了对图像进行像素级标注,首先要确定分类标准。世界气象组织在云的分类系统中将云分为 3 族 10 属 28 类^[13],而在实际观测中,研究人员通常会按照云属区分云状类别,也就是将云划分为 10 类:高积云、高层云、积雨云、卷积云、卷云、卷层云、积云、雨层云、层积云、层云。本文也将延续这一划分标准,将 GBCSS 数据集中的云按上述 10 类云属划分。再加上 1 类背景(天空、太阳等非云物体),GBCSS 最终被设定包含 11 种云状类别,每一种类别都对应一种标注颜色。表 1 列出了 GBCSS 数据集的细节。然后,每一张云图都在专家的监督下被标注,得到一张对应的标注图像。最后,针对部分云属占比较少的问题,我们对部分数据进行了数据增广,以减小不同云状样本数量的差距。GBCSS 数据集共包含 3000 张带有标签的地基云图,图像的像素分辨率为 256×256。

图 2 给出了 GBCSS 数据集中的一些样本。与现有的地基云图公开数据集^[11,14-16]相比,GBCSS 数据集的云属齐全,数据量大;且数据源自气象站,标注过程由气象专家监督。因此该数据集的参考价值高,为后续网络的训练及评估奠定了基础。

表1 GBCSS数据集

Table 1 Details of GBCSS dataset

种类	简写	标签值(R,G,B)	特征
高积云	Ac	128,0,0	密集灰色云束,轮廓呈绒毛状
高层云	As	0,128,0	不透明的灰色厚云,布满天空
积雨云	Cb	128,128,0	厚雷云,轮廓不清晰
卷积云	Cc	0,0,128	白亮小片的云束,呈细鳞片状
卷云	Ci	128,0,128	白色透亮,纤维状薄云
卷层云	Cs	0,128,128	能透光的薄云,呈丝缕状
积云	Cu	128,128,128	云块厚重,边缘清晰,如拳头一般
雨层云	Ns	64,0,0	暗灰色雨云,常布满天空
层积云	Sc	192,0,0	松散的灰色云,常呈波状分布
层云	St	64,128,0	均匀的灰色云,覆盖面积大
无云	No	0,0,0	不属于任何云

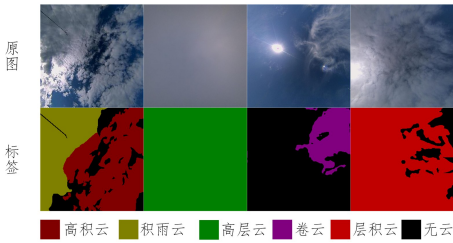


图2 GBCSS数据集样本

Fig. 2 Samples of GBCSS dataset

3 基于多尺度特征提取的 UNet-PPM

本文提出的地基云图语义分割模型 UNet-PPM 使用了编码器-解码器思想作为网络的基本框架,并引入了金字塔池化模块(Pyramid Pooling Module, PPM)以捕获网络浅层信息中的多尺度特征。因此,本节首先会简述编码器-解码器以及金字塔池化模块的原理和特点,然后再详述本文提出的网络结构。

3.1 相关工作

2015年,Ronneberger等^[17]提出了基于全卷积网络的 U-Net 分割模型。U-Net 的主要贡献在于其完全对称的编解码器结构,以及将同一阶段中下采样与上采样的特征图拼接起来的跳连接(Skip Connection)方式。这保证了最终恢复的特征图能够融合更多的浅层特征以及多尺度特征。

金字塔池化模块^[18]的结构如图3所示。对于输入的特征图,该模块首先使用4种不同尺度的池化操作,得到4种尺寸的特征图;然后对所有的特征图都添加 1×1 的卷积操作以减少特征图的通道数;接着使用双线性插值对特征图进行上采样,使得它们的尺寸与原始特征图保持相同;最后将这些特征图以及输入特征图在维度通道上拼接在一起,得到最终的金字塔全局池化特征图。这种多尺度池化并联的结构聚合了不同区域的上下文信息,提升了网络获取全局信息的能力。

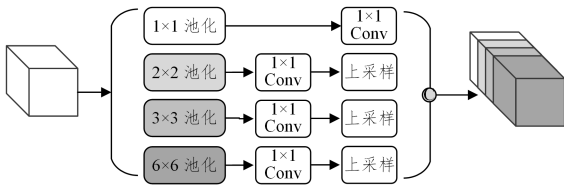


图3 金字塔池化模块

Fig. 3 Pyramid pooling module

3.2 语义分割 UNet-PPM 模型

U-Net 是语义分割任务中的常用模型,但在地基云的识别中,它容易陷入云体局部的特征提取而忽视对整体轮廓的

关注。本文认为,这个问题与 U-Net 的跳连接方式有关。在 U-Net 网络的解码器末端,特征图经过3次卷积以及激活函数,输出最终的预测图。该特征图由两个更小的特征图拼接得到。其中的特征图 A 由原始图经过4次下采样与4次上采样后得到。而需要注意的是,特征图 B 仅由原始图经过两次卷积后得到,这导致特征图 B 仅具有极为浅层的特征信息。我们认为这样的跳连接操作没能充分发挥深度神经网络强大的提取特征能力。因此,我们使用了 PPM。这个模块对每一个即将被拼接的下采样特征图进行了一次多尺度特征提取,以补充上采样过程中融合特征图的特征信息。

本文提出的 UNet-PPM 网络结构示意图由图4所示。其中的编码器由5个卷积层和4个池化层组成。表2的上半部分记录了每个卷积层的卷积核大小以及数量。受4次池化的影响,解码器最终输出的特征图尺寸为原输入图像的 $1/16$ 。此外,编码器还复制了前4个卷积层的输出并将它们输入 PPM 中,用多尺度池化的方法提取这些特征图的全局信息,用于网络的解码部分。解码器由4个卷积层和4个上采样层组成,每一层的结构在表2的下半部分详细说明。其中,每个卷积层的输入都由上一个卷积层经上采样输出的特征图,以及对应层级 PPM 的输出特征图拼接得到。这样做能够将深层特征和浅层特征结合在一起,保留更多的全局信息。解码器的输入经过4次上采样后,其特征图尺寸被还原至与网络输入图像同样大小。最后使用 1×1 的卷积将特征图的通道数融合至11,得到最终的预测图。

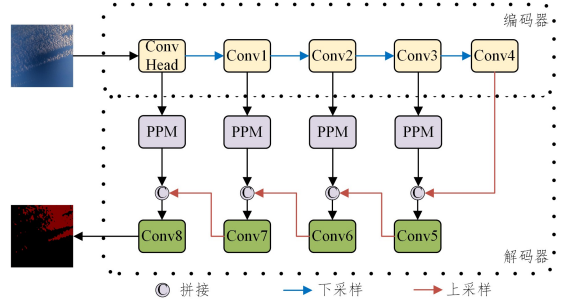


图4 UNet-PPM 示意图

Fig. 4 Overview of UNet-PPM

表2 UNet-PPM 的网络结构

Table 2 Architecture of UNet-PPM

网络层	特征图输出尺寸	网络结构
ConvHead	256×256	$[3 \times 3, 64]$
		$[3 \times 3, 64]$
Conv1	128×128	$[3 \times 3, 128]$
		$[3 \times 3, 128]$
Conv2	64×64	$[3 \times 3, 256]$
		$[3 \times 3, 256]$
Conv3	32×32	$[3 \times 3, 512]$
		$[3 \times 3, 512]$
Conv4	16×16	$[3 \times 3, 1024]$
		$[3 \times 3, 1024]$
Conv5	32×32	$[3 \times 3, 1024]$
		$[3 \times 3, 512]$
Conv6	64×64	$[3 \times 3, 512]$
		$[3 \times 3, 256]$
Conv7	128×128	$[3 \times 3, 256]$
		$[3 \times 3, 128]$
Conv8	256×256	$[3 \times 3, 128]$
		$[3 \times 3, 64]$
		$[1 \times 1, 11]$

4 实验结果与分析

4.1 实验环境及超参数设置

实验所涉及到的神经网络的搭建、训练及测试都基于 PyTorch 1.13.1 深度学习框架。GPU 使用一张 GeForce RTX 4090, 其显存大小为 24GB。

网络训练则使用在 GBCSS 中标记的 3000 幅地基云图作为数据集。其中训练集含有 2700 幅图像, 测试集含有 300 幅图像, 训练集与测试集的比例为 9:1。所有输入图像的分辨率都为 256×256 。网络训练批次设为 32, 迭代期为 100 次, 学习率为 0.001, 损失函数使用交叉熵函数。

本文采用像素准确率 (Pixel Accuracy, PA) 评估网络模型在语义分割工作上的准确度。对于语义分割, PA 表示一次预测中, 正确预测的像素数量占总预测像素数量的

比例, 可用式 (1) 表示:

$$PA = \frac{\sum_{i=1}^k p_{ii}}{\sum_{i=1}^k \sum_{j=1}^k p_{ij}} \quad (1)$$

其中, k 表示数据集中的类别总数, p_{ij} 表示类别为 i 的像素被预测为类别 j 的数量, p_{ii} 表示类别为 i 的像素被预测为类别 i 的数量。

4.2 实验结果分析

本文在 GBCSS 数据集上将 UNet-PPM 与几种常见的用于语义分割的网络, 包括 U-Net, DeepLabv3+^[19], DANet^[20], BiSeNetv2^[21], 进行了对比, 实验结果如表 3 所列。我们的方法的总 PA 为 91.5%, 与 U-Net, DeepLabv3, DANet, BiSeNetv2 相比分别提高了 5.4%, 5.9%, 4.5%, 5.6%, 达到了最优。

表 3 网络性能比较

Table 3 Comparison of network performance

网络	PA/%											PA/%	GFLOPs
	Ac	As	Cb	Cc	Ci	Cs	Cu	Ns	Sc	St	No		
U-Net	78.4	84.7	95.3	61.2	57.2	66.3	86.8	97.8	81.9	83.4	95.4	86.1	48.37
DeepLabv3+	85.7	61.8	94.1	48.5	64.5	84.4	82.4	96.6	92.3	83.6	95.8	85.6	7.55
DANet	76.2	83.5	96.8	62.3	64.7	87.2	83.2	99.9	82.4	80.0	93.1	87.0	128.43
BiSeNetv2	81.2	94.3	95.2	55.8	62.9	76.0	60.8	97.9	84.8	86.5	90.3	85.9	2.36
UNet-PPM	96.7	97.3	99.3	61.0	77.0	84.1	82.1	99.9	90.1	82.0	95.9	91.5	118.20

图 5 给出了各个网络在测试集上的可视化样例。对于第一幅图, U-Net, DeepLabv3+, DANet 都将部分积雨云的云体误判为积云。虽然 BiSeNetv2 没有出现云体误判的问题, 然而

与 UNet-PPM 相比, 该网络没能精准划分出云的轮廓, 导致分割效果不佳。从样例的分析可以看出, UNet-PPM 与其他网络相比能够准确区分外观相似的云体, 并分割出完整的轮廓。

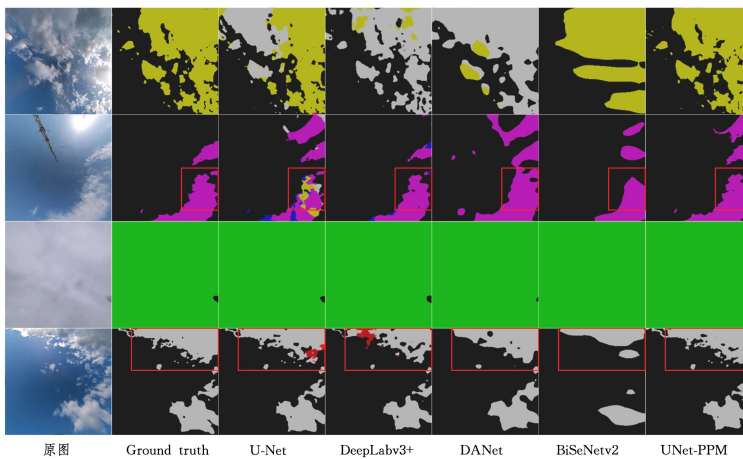


图 5 地基云图语义分割结果示例

Fig. 5 Examples of ground-based cloud semantic segmentation

将上述网络在训练过程中的 Loss 随迭代增加的收敛情况进行对比, 结果如图 6 所示。可以看出, 本文提出的 UNet-PPM 的收敛速度最快, 最终的 Loss 值最低, 说明 UNet-PPM 网络在 GBCSS 数据集上的分割效果最优。

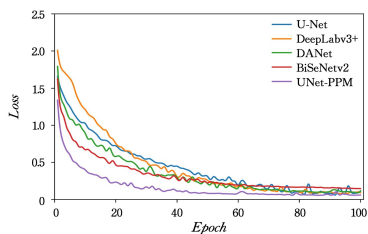


图 6 Loss 曲线

Fig. 6 Loss curve

图 7 分别展现了 5 个网络在训练过程中对训练集和测试集的准确率变化曲线。如果一个网络对已有数据 (训练集) 的拟合效果远好于未知数据 (测试集), 则说明该网络出现了过拟合现象。由于检测一个模型的表现就是通过观察它在未知场景下的预测能力, 所以网络在训练时应该避免或者缓解过拟合问题。以 U-Net 为例, 在 epoch 达到大约 20 的时候, 网络在测试集的准确率就停止了增长, 而训练集的准确率却一直增加。这说明网络在之后的训练中仅仅在学习训练集的特例, 而无法将提取到的特征泛化至一般的测试集上, 因此 U-Net 在第 20 epoch 后的训练存在严重的过拟合问题。类似的现象也出现在 DANet 和 BiSeNetv2 中。而相比之下, UNet-PPM 在大约第 40 epoch 后, 其训练集和测试集的准确率都不再增长, 并且最终在

测试集上的准确率最高。这说明所提方法的特征提取能力更强,极大地缓解了过拟合的问题。

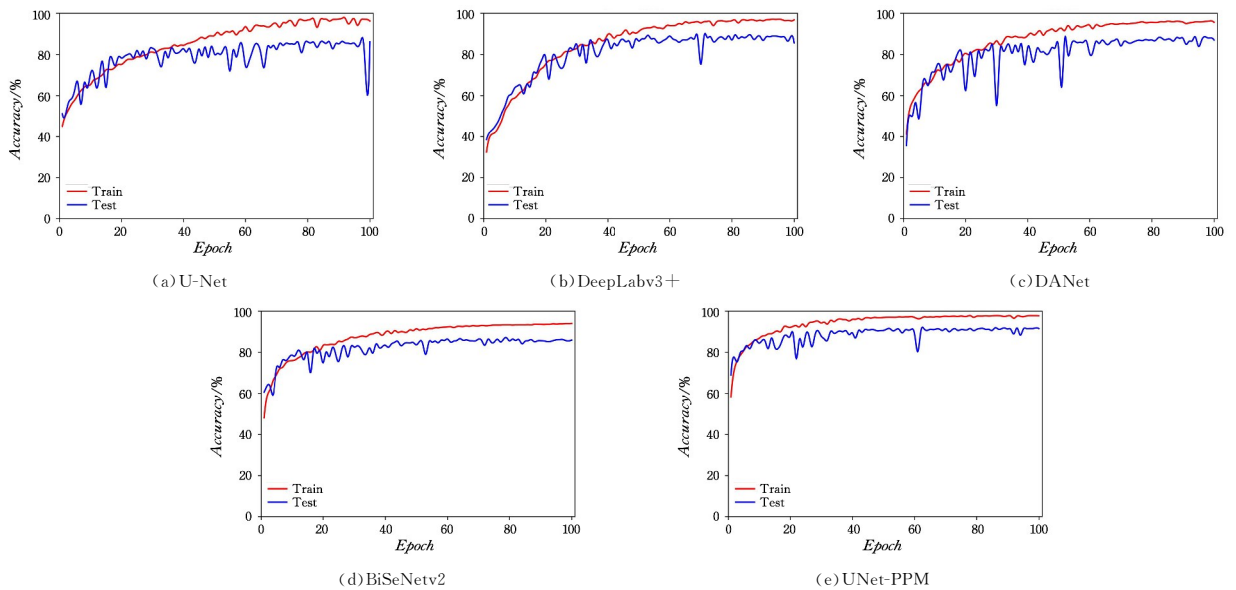


图7 网络在训练集/测试集上的表现

Fig. 7 Performance of each network on training and test set

结束语 在地基云自动化观测任务中,云状信息与云量信息是人们关心的两种云的主要属性。然而,以往的相关研究都将获取两种信息对应的分类任务与分割任务独立对待,导致当观测视野内同时出现多种云状时,无法根据不同的云状分别计算云量。因此,本文提出用语义分割将云状识别与云量估计两个任务结合起来,用深度学习实现地基云图语义分割任务。本文的主要贡献如下:

1)针对目前没有地基云图语义分割公开数据集这一问题,构建了GBCSS数据集。该数据集包含3000张地基云图,每一幅都带有像素级标签,用于后续的有监督网络训练与测试。

2)提出了一个基于深度学习的地基云图语义分割模型UNet-PPM。该模型利用金字塔池化模块对U型网络进行了改进,在GBCSS上的测试集的准确率达到91.5%。

由于目前国内外对地基云图的语义分割的相关研究很少,我们认为在这个领域还有很多改进工作可以做。

1)使用高质量的地基云图扩充GBCSS数据集,以进一步提升网络性能。目前,该数据集仍存在类别不均衡的问题。例如,不考虑背景的前提下,10类云属中占比最多的两类云为高层云(9.94%)以及雨层云(9.87%),而最低的两类云为积云(3.60%)和卷云(3.81%)。这使得网络难以提取低占比云属的特征。尽管数据增强能够在一定程度上缓解类别不均衡的问题,但是用该方法扩充的数据存在大量噪声,减少了网络的推理性能^[22]。因此,引入更多的真实数据仍然是解决样本不均衡问题的最优手段。

2)尽管人们可以借助成像设备获取大量的地基云图,但由于地基云图的语义标注工程量大,导致带标签的数据很少。为此可引入半监督学习^[23-24],在大量无标签数据的分布中提取信息,使得深度网络的训练能够摆脱大量带标签样本的依赖。

3)基于卷积神经网络的语义分割算法虽然精确度高,但是其参数量大、运算复杂度高的缺陷导致这类算法在实际部署时会受到存储空间和算力资源的限制。例如嵌入式设备的

处理器性能差,内存小且功耗受限,使得高精度的网络模型难以在这类平台上实时运行。因此,网络轻量化成为了一个重要的探索方向。在保证精度不变的前提下,通过知识蒸馏^[25]、剪枝^[26]等手段减少模型的参数量,将轻量级网络推向实际应用。

参考文献

- [1] ZHAO C, CHEN Y, LI J, et al. Fifteen-year statistical analysis of cloud characteristics over China using Terra and Aqua Moderate Resolution Imaging Spectroradiometer observations [J]. *International Journal of Climatology*, 2019, 39(5): 2612-2629.
- [2] ZHU B, YANG J, LV W T, et al. Ground-based visible cloud image classification method based on KNN algorithm [J]. *Journal of Applied Meteorological Science*, 2012, 23(6): 721-728.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [4] ZHU W, CHEN T, HOU B, et al. Classification of ground-based cloud images by improved combined convolutional network [J]. *Applied Sciences*, 2022, 12(3): 1570.
- [5] SAITO M, IWABUCHI H. Cloud discrimination from sky images using a clear-sky index [J]. *Journal of Atmospheric and Oceanic Technology*, 2016, 33(8): 1583-1595.
- [6] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015: 3431-3440.
- [7] ZHANG X, JIA K B, LIU J, et al. Segmentation Technology of Ground-Based Cloud Image for Lightweight [J]. *Measurement & Control Technology*, 2022, 41(9): 37-43.
- [8] YE L, CAO Z, XIAO Y, et al. Supervised fine-grained cloud detection and recognition in whole-sky images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(10): 7972-7985.

- [9] BANKERT R L. Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network[J]. *Journal of Applied Meteorology and climatology*, 1994, 33(8): 909-918.
- [10] DEV S, LEE Y H, WINKLER S. Multi-level semantic labeling of sky/cloud images[C]// 2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015: 636-640.
- [11] DEV S, LEE Y H, WINKLER S. Categorization of cloud image patches using an improved texton-based approach[C]// 2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015: 422-426.
- [12] DEV S, LEE Y H, WINKLER S. Color-based segmentation of sky/cloud images from ground-based cameras[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 10(1): 231-242.
- [13] COHN S A. A new edition of the international cloud atlas[J]. *WMO Bulletin*, Geneva, World Meteorological Organization, 2017, 66: 2-7.
- [14] ZHANG J, LIU P, ZHANG F, et al. CloudNet: Ground-based cloud classification with deep convolutional neural network[J]. *Geophysical Research Letters*, 2018, 45(16): 8665-8672.
- [15] ZHUO W, CAO Z, XIAO Y. Cloud classification of ground-based images using texture-structure features[J]. *Journal of Atmospheric and Oceanic Technology*, 2014, 31(1): 79-92.
- [16] XIE W, LIU D, YANG M, et al. SegCloud: A novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation[J]. *Atmospheric Measurement Techniques*, 2020, 13(4): 1953-1961.
- [17] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference*, Munich, Germany. Springer International Publishing, 2015: 234-241.
- [18] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 2881-2890.
- [19] CHEN L C, ZHU Y, PAPANDEOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018: 801-818.
- [20] 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.
- [21] YU C, GAO C, WANG J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation [J]. *International Journal of Computer Vision*, 2021, 129: 3051-3068.
- [22] GONG C, WANG D, LI M, et al. Keepaugument: A simple information-preserving data augmentation approach [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 1055-1064.
- [23] LIU J W, LIU Y, LUO X L. Semi-Supervised Learning Methods [J]. *Chinese Journal of Computers*, 2015, 38(8): 1592-1617.
- [24] CHEN M, DU Y, ZHANG Y, et al. Semi-supervised learning with multi-head co-training[C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. 2022: 6278-6286.
- [25] HEO B, KIM J, YUN S, et al. A comprehensive overhaul of feature distillation[C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019: 1921-1930.
- [26] LIU Z, LI J, SHEN Z, et al. Learning efficient convolutional networks through network slimming[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2017: 2736-2744.



SUN Jifei, born in 1993, master. His main research interests include ground-based cloud images processing and semantic segmentation.



JIA Kebin, born in 1962, Ph.D, professor, Ph.D supervisor. His main research interest is image and video processing.