引文格式：胡文超,解志刚,叶周园,等. 利用弱监督预训练CNN模型的遥感影像云检测方法[J]. 地理空间信息, 2024, 22(12): 25-29. doi:10. 3969/j.issn. 1672-4623.2024.12.006

# 利用弱监督预训练CNN模型的遥感影像云检测方法

胡文超[1,2]，解志刚[1]，叶周园[2]，刘亚萍[3*]
（1. 32022部队，湖北 武汉 430074；2. 武汉大学 遥感信息工程学院，湖北 武汉 430079；
3. 中交第二公路勘察设计研究院有限公司，湖北 武汉 430056）

**摘　要：**近年来，深度学习在遥感影像云检测方面得到广泛应用，但该方法依赖大规模高质量的人工标注样本。因此，借鉴弱监督学习和迁移学习思想，提出了一种基于弱监督预训练CNN模型的云检测方法。首先采用快速云检测方法生成大规模弱监督样本，将其代替人工标注样本对CNN模型进行预训练；然后利用少量人工标注样本对所得预训练模型进行精调；最后利用训练所得模型对影像进行云检测。实验结果表明，该方法总体精度达到96.00%，F1分数为87.70%，在维持较好检测精度的前提下显著降低了人工标注成本，在工程实践中具有较好的应用前景。

**关键词：**云检测；弱监督；CNN；遥感影像；迁移学习
**中图分类号：**P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）12-0025-05

# Cloud Detection Method for Remote Sensing Image Based on Weakly Supervised Pre-training CNN Model

HU Wenchao[1,2], XIE Zhigang[1], YE Zhouyuan[2], LIU Yaping[3]
(1. Troops of PLA 32022, Wuhan 430074, China;
2. School of Remote Sensing Information Engineering, Wuhan University, Wuhan 430079, China;
3. CCCC Second Highway Consultants Co., Ltd., Wuhan 430056, China)

**Abstract:** In recent years, deep learning has been widely used in cloud detection of remote sensing images, but it relies on large-scale high-quality manually labeled samples. Based on weakly supervised learning and transfer learning, we proposed a cloud detection method based on weakly supervised pre-training CNN model in this paper. Firstly, we used a fast detection method to generate a large-scale weakly supervised sample, which was used to pre-train the CNN model instead of manually labeled samples. Then, we used a small number of manually labeled samples to fine-tune the pre-training model. Finally, we used the obtained model to conduct cloud detection of images. The experimental results show that the overall accuracy of this method reaches 96.00%, with F1-score of 87.70%. On the premise of maintaining ideal detection accuracy, it significantly reduces manual labeling costs, presenting good application prospects in industrial practice.

**Key words:** cloud detection, weakly supervision, CNN, remote sensing image, transfer learning

随着遥感技术的高速发展，高分辨率遥感影像已广泛应用于资源探测、气象监测、环境保护和军事侦察等众多领域，为社会经济建设和人们工作生活提供了重要信息支撑。但由于光学遥感卫星成像时经常会受到云层遮挡，造成“云污染”。据统计，地球表面约66%以上区域经常被云覆盖[1]。云噪声会对遥感数据的后续处理和应用产生不利影响，它们的检测是大多数分析和应用中的不可或缺的初始步骤和重要部分[2]。

近年来，国内外学者在云检测方面开展了大量卓有成效的研究，特别是随着卷积神经网络（convolutional neural networks，CNN）的兴起和计算机运算能力的大幅提升，深度学习在图像分割领域展现出强大能力，基于深度学习的方法成为了遥感影像云检测方面的研究热点。如郑凯[3]等基于ResNet50骨干网络，提出了一种结合ResNet和优化的DeepLabV3+的全卷积神经网络，实现对天绘一号遥感影像的云雪检测，在稳健性、检测率和速度方面都有着优秀的表现。杨剑[4]等在高分一号影像的云检测任务中，使用了基于代价敏感学习的方法，克服了云和背景样本不均衡造成的模型泛化能力差问题，有效提升了模型的云识别能力。大量实验表明，基于CNN模型的云检测方法检测效果一般优于传统方法，但它最大的瓶颈是

收稿日期：2023-10-20。
第一作者简介：胡文超（1987—），硕士，工程师，主要从事摄影测量、遥感测绘方面的科研和工程技术工作，E-mail：feilanxyz@163.com。
通信作者：刘亚萍（1990—），硕士，高级工程师，注册测绘师，研究方向为倾斜摄影测量、激光雷达扫描测量、数据处理和误差分析等，E-mail：415865178@qq.com。

需要大规模高质量的训练数据作支撑，而通过人工获取训练样本成本巨大，这在一定程度上限制了CNN在影像云检测方面的应用。为了解决CNN过度依赖人工标注样本这一问题，有研究学者提出了弱监督学习[5-6]的方法，即利用不完全、不准确的弱标签数据代替人工标注数据来训练模型，减少人工标注样本的使用。如仇一帆[7]提出了一种无人工标注数据的云和云阴影深度学习方法，即使用常规云检测算法CFMask检测出的云及其阴影图像替代人工标注数据用于训练CNN模型，再利用所得模型对待测影像进行云及阴影检测，实验结果总体精度达到85.55%，与常规CFMask方法相比精度更高。吴代强[8]等利用Landsat8带有标签的公开数据集和基于阈值的高概率样本选取方法构建样本集，训练出随机森林模型对高分五号进行云检测，与常规方法相比，避免了人工标注数据的使用，显著提高了卫星数据的处理效率。

基于以上内容，本文借鉴计算机视觉领域的弱监督学习和迁移学习思想，设计了一种基于弱监督预训练CNN模型的云检测方法。首先利用快速云检测方法获取粗略云检测结果，然后将其作为弱监督样本输入CNN模型进行预训练，再利用少量的人工标注样本进行精调获取最终训练模型，实现对遥感影像的高效云检测。该方法旨在探索在使用少量人工标注样本的情况下，通过一定策略训练出具有实用价值的云检测模型，从而有效降低人工标注成本，提高云检测效率。

## 1 研究数据

### 1.1 原始数据

本文选取天绘一号卫星不同时期在不同地区获取的多光谱影像数据开展研究。基于样本均衡全面的原则，云影像包括积云、层云、卷云、碎云等不同类型，下垫面涵盖山地、林地、耕地、荒地、沙漠、湖泊、冰雪等。共计选取了34景不同下垫面含不同云类型的遥感卫星影像数据。影像数据大小为6 000×6 000像素，由蓝（0.43～0.5 μm）、绿（0.52～0.61 μm）、红（0.61～0.6 μm）、近红外（0.76～0.90 μm）4个通道（RGBNir）组成，空间分辨率为10 m。

### 1.2 数据集构建

#### 1.2.1 数据预处理

为了提高原始影像的可识别性，对影像数据进行了预处理，主要包括*DN*值辐射校正、影像平滑和影像增强等，以去除噪声，突出影像中的轮廓显示，为后续影像分割做准备。

#### 1.2.2 样本制作

NDVI法[9]是一种典型的基于光谱特征的云检测方法。而*D*阈值法[10]是基于NDVI法改进的一种方法，该方法定义了一种阈值*D*。

$$D=\frac{|\mathrm{NDVI}|^{b}}{\beta^{2}}$$

式中，$\beta$为可见光红色波段的像素值；$b$为光谱特征空间中两类样本的最佳分界线斜率。利用同一地区多时相影像中的无云或少云影像，通过最小二乘直线拟合可获取$b$值，再通过公式计算出*D*值，最后利用该阈值对有云影像进行云检测。

*D*阈值法算法相对简单，云检测速度较快，且可达到一定的检测精度。本文采用*D*阈值法快速生成粗略的云掩膜图像，形成大规模弱监督样本，并通过人工标记方式制作实验所需的少量人工标注样本。

#### 1.2.3 影像切分

单景卫星影像图幅比较大，不适合直接输入CNN模型进行训练，因此将单景影像和对应云掩膜图像（均为6 000×6 000像素）切分为384×384像素互不重叠的影像块，共获得8 704张影像块和对应的掩膜图像。

#### 1.2.4 数据增强

为了提高训练模型的泛化能力，本文使用数据扩充技术对数据集进行了扩充，主要使用几何平移的方式，包括水平翻转、旋转和缩放，对数据进行了随机增强。

#### 1.2.5 数据集划分

数据集经过数据增强后，根据要求划分为训练集、验证集和测试集，共包含由34景影像切割而成的8 704组样本。其中24景数据按照80%和20%的比例随机分配为训练集和测试集，保证了样本的随机性和泛化性。剩余的10景用于构建测试集。

## 2 云检测方法

本文实验流程如图1所示。

### 2.1 Cloud-Net网络

随着深度学习在计算机视觉领域的发展，多种语义分割CNN模型被开发，应用于遥感影像分类问题。Cloud-Net[11]是基于U-Net[12]改进的云检测算法，在U-Net的基础上对内部卷积块进行了优化，每个块中均包含卷积、连接、池化等操作，并融入残差网络的思想，提高了模型对云特征的敏感性。它能够更充分提取图像的全局特征和多层次特征来恢复更精确的特征图，且不需要任何超像素分割等复杂的预处理步骤，

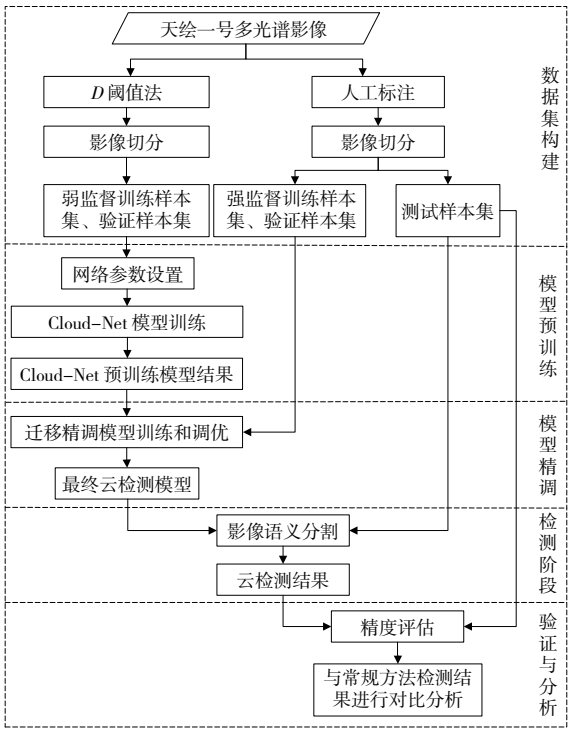



图1　云检测实验方案技术流程

已在Landsat8影像上得到成功验证。本文将基于文献[13]中的Cloud-Net语义分割模型搭建应用于天绘一号卫星影像云检测的深度网络模型。

### 2.2　预训练与精调

迁移学习可以智能地应用以前学到的知识来更快或更好地解决新问题[13]，在许多现实问题中得到成功应用。预训练和精调是迁移学习中的2个重要部分。预训练是指利用已有的数据集对某个网络模型进行训练，得到一个预训练模型成果的过程。而精调是指将预训练过的模型作用于自己的数据集，并使参数适应自己数据集的过程。在CNN领域中，由于获取大规模数据集成本高昂，利用少量样本训练模型又容易造成过拟合等原因，从头开始训练整个CNN网络的情况比较少。常见的操作是先依托一个已有的大型数据集预训练出一个模型，然后使用该模型作为类似任务的初始化设置或固定的特征提取器，这样既克服了对大规模数据集的依赖，又能较快达到理想的效果，在解决实际问题中发挥了重要作用。受此启发，本文探索将预训练和精调训练策略用于天绘一号卫星影像云检测的模型训练中，即先利用大量的弱监督样本对Cloud-Net进行预训练，再利用少量的人工标注样本对预训练模型进行精调。

### 2.3　参数设置与网络训练

本文中模型的图像通道数设置为4，输入图像大小为384×384像元。Cloud-Net语义分割模型在进行训练之前，网络权重被初始化为[-1，1]之间的均匀随机分布。模型训练的初始学习率设置为$10^{-4}$，学习衰减率设置为0.7，容忍度为15，该策略一直持续到学习率达到$10^{-7}$或迭代数达到500次为止，批量大小设置为16。在优化Cloud-Net模型的过程中，使用了Adam梯度下降法，解决了损失函数的优化问题，同时使用了soft Jaccard损失函数[10]，用于计算误差，使模型训练更加准确。模型输出为一个384×384×2的概率图，用于预测图像上每个像元所属类别的概率，根据具有最大概率的类别输出每个像元的检测结果，生成最终的云掩膜图像。

### 2.4　精度评价

为了对影像的云检测精度进行定量评价，本文采用了精确率（Precision）、召回率（Recall）、总体精度（overall accuracy，$OA$）和$F1$分数（F1-score）作为评估指标。

精确度是被预测为云的像元中真正为云的概率，召回率是正确预测为云的像元占所有被预测为云的像元的比例。总体精度是是检测结果中预测正确的像元数量占影像所有像元数量的比例。$F1$分数值为精确率和召回率的加权平均值，用于平衡两个评价指标，能够更均衡地评价本文方法的综合性能。以上评价指标的计算公式分别为：

$$\text{Precision}=\frac{TP}{TP+FP}$$

$$\text{Recall}=\frac{TP}{TP+FN}$$

$$OA=\frac{TP+TN}{TP+TN+FP+FN}$$

$$F_1=\frac{2\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$$

式中，$TP$为正确检测为云的像元数；$FP$为非云像元中错检为云的像元数；$FN$为云像元中错检为非云的像元数，$TN$为非云像元中正确检测为非云的像元数。

## 3　实验结果分析

模型训练使用了预训练和精调策略，在预训练阶段使用的训练数据为24景弱监督样本（6 144组，包含影像块和对应的云掩膜图像），精调阶段使用的训练数据为6景人工标注样本（1 536组），精调使用的6景影像从预训练使用的24景影像中选取，弱监督样本与人工标注样本比例为4∶1，实验样本配置方案如图2所示。为了评估本文方法的有效性，对云检测结果进行定性分析和定量分析，并将本文方法的检测结果与全监督Cloud-Net方法、NDVI检测法和$D$阈值法检测结果进行对比分析，探讨本文方法的实用性。

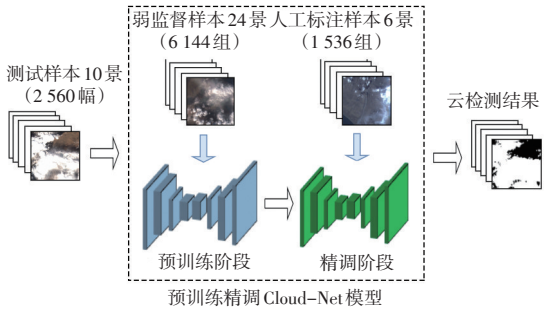



图2　实验样本配置方案

### 3.1　定性分析

从实验检测结果中选取了5个的典型场景，对不同方法的云检测结果进行视觉对比，如图3所示，图3a、e场景的下垫面依次为山地、城镇、林地、荒地、湖泊。每一行从左至右依次为原始影像、groud truth、NDVI法检测结果、*D*阈值法检测结果、全监督Cloud-Net检测结果和本文方法检测结果。原始影像为真彩色合成，groud truth为通过人工标注的云掩膜标签，检测结果和groud truth中的云和非云像元分别用白色和黑色标示，检测结果如图3所示。

通过对检测结果进行目视比对，可以看出，对于山地、城镇、林地、荒地、湖泊5种场景，本文能够较准确地定位出云层在影像中的位置，比较清楚地识别云层的边缘轮廓，对于图3b、c中的积云以及图3d、e中的层云部分，本文方法能够取得较准确的检测结果，与常规全监督Cloud-Net方法相比，结果非常接近。对于图3a、e的薄云部分，本文方法依然可以检测出大部分的云像元，虽然丢失了少量细节，但生成的云掩膜图像与全监督Cloud-Net方法检测结果基本一致，表现均优于NDVI法和*D*阈值法。

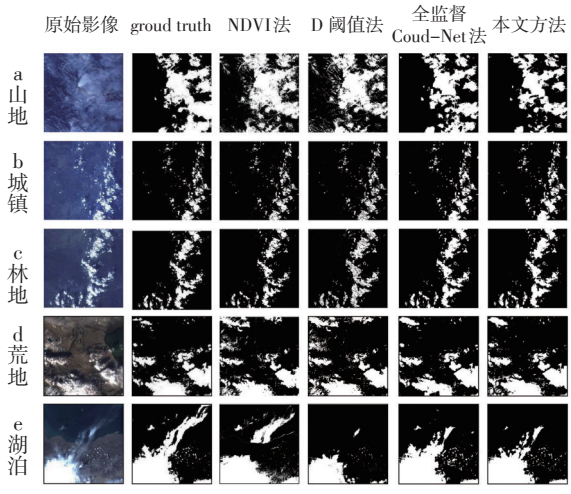



图3　云检测结果视觉对比

### 3.2　定量分析

本文采用2.4中的4种指标对实验结果进行精度评价，并将本文方法与其他方法的检测精度进行对比，结果如表1和图4所示。可以看出，本文方法虽然只使用了少量的人工标注样本，云检测召回率仍达到88.98%，总体精度达到96%，与全监督Cloud-Net方法检测结果相比非常接近（分别略低于0.71%和0.49%）；精确率为84.53%，*F*1分数为86.70%，与全监督CloudNet方法相比也只是小幅降低（分别降低了3.26%和1.82%）。与常规NDVI检测法、*D*阈值法检测精度相比，除*D*阈值法的精确率外，其他各项精度指标均占优，达到了主流云检测方法的水平，证明基于弱监督预训练CNN语义分割模型进行遥感影像云检测的思路是可行的。

表1　测试样本集云检测定量评价/%

| 方法 | Precision | Recall | *OA* | *F*1 |
|---|---|---|---|---|
| NDVI检测法 | 76.19 | 79.06 | 88.28 | 77.60 |
| *D*阈值法 | 87.60 | 74.33 | 90.87 | 80.31 |
| 全监督Cloud-Net法 | 87.79 | 89.27 | 96.49 | 88.52 |
| 本文方法 | 84.53 | 88.98 | 96.00 | 86.70 |

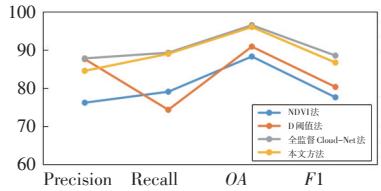



图4　不同方法测试样本集云检测精度对比/%

为了进一步分析云检测工作效率，对本文方法和全监督Cloud-Net方法的检测精度、人工标注工作量和耗时进行了统计和量化对比，如表2所示。可以看出，对同样的测试样本进行云检测，本文方法与全监督Cloud-Net方法相比，在云检测精度基本接近（总体精度和*F*1分数分别略降低于0.49%和1.82%）的情况下，人工标注工作量和时间消耗显著降低（人工标注工作量和耗时分别大幅降低了75%和55.3%）。表明本文方法在维持较好云检测精度的同时，大幅降低了人工标注成本和缩短了工作周期，显著提高了工程效率。

表2　测试样本集云检测结果总体对比

| 方法 | 检测精度/% | | | | 人工标注工作量/景 | 耗时/h |
|---|---|---|---|---|---|---|
| | Precision | Recall | *OA* | *F*1 | | |
| 全监督Cloud-Net方法 | 87.79 | 89.27 | 96.49 | 88.52 | 24景 | 969 |
| 本文方法 | 84.53 | 88.98 | 96.00 | 86.70 | 6景 | 433 |
| 二者对比 | ↓3.26 | ↓0.29 | ↓0.49 | ↓1.82 | ↓75 | ↓55.3 |

## 4　结　论

云检测是遥感影像工程应用中不可或缺的一项基

础工作，本文针对常规深度学习云检测方法依赖大规模人工标注数据集的问题，提出了一种基于弱监督预训练CNN模型的新方法，并将其与其他代表性典型方法进行了实验比较。实验结果表明，本文方法能够在维持较好检测精度的前提下大幅降低人工标注工作量，为降低训练数据获取成本、提高云检测工作效率提供了符合工程快速应用的有效路径。本文方法在模型训练策略上仍有改进空间，下一步将针对人工标注样本和弱监督样本的比例组合方案继续开展深入研究，进一步提高云检测精度和工程效率。

## 参考文献


[1] Zhang Y, Rossow W B. Calculation of Radiative Fluxes 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

[2] Arvidson T, Gasch J, Goward S N. Landsat 7's long-term Aacquisition Plan: An Innovative Approach to Building a Global Imagery Archive[J]. Remote Sensing of Environment, 2001, 78(1-2):13-26

[3] 郑凯, 李建胜, 杨戬峰, 等. 天绘一号卫星遥感影像云雪检测的ResNet与DeepLabV3+综合法[J]. 测绘学报, 2020, 49(10):1 343-1 353

[4] 杨剑,颜源,李乐林. 基于代价敏感学习的遥感影像云检测方法[J]. 地理空间信息,2021,19(8):54-57

[5] Zhou Z. A Brief Tntroduction to Weakly Supervised Learning[J]. National Science Review,2018,5(1):44-53

[6] Deselaers T, Alexe B, Ferrari V. Weakly Supervised Localization and Learning with Generic Knowledge[J]. International Journal of Computer Vision,2012,100(3):275-293

[7] 仇一帆, 柴登峰. 无人工标注数据的Landsat影像云检测深度学习方法[J]. 国土资源遥感, 2021, 33(1):102-107

[8] 吴代强, 何涛. 基于Landsat8样本数据库的高分五号影像云识别[C]//第七届高分辨率对地观测学术年会论文集, 2020:440-450

[9] 李民赞, 韩东海, 王秀. 光谱分析技术及其应用[M]. 北京: 科学出版社, 2006

[10] Girolamo L D, Davies R. The Image Navigation Cloud Mask for the Multiangle Imaging Spectro Radiometer (MISR) [J]. Journal of Atmospheric and Oceanic Technology, 1995, 12(6):1 215-1 228

[11] Mohajerani S, Saeedi P. Cloud-Net:An End-to-end Cloud Detection Aalgorithm for Landsat8 Imagery[C]// IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2019:1 029-1 032

[12] Ronneberger O, Fischer P, Brox T. U-net:Convolutional CNnetworks for Biomedical Imagesegmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention, Springer, Cham, 2015: 234-241

[13] Zhang X X, Li H X, Cheng C, et al. Transfer Learning Based 3D Fuzzy Multivariable Control for an RTP System[J]. Applied Intelligence, 2020, 50(3):812-829


（上接第24页）


ing Remote Sensing Data:A Machine Learning Approach[J]. Advances in Water Resources,2010,33(1):69-80

[11] Jamei M, Ali M, Karbasi M, et al. A High Dimensional Features-based Cascaded Forward Meural Network Coupled with MVMD and Boruta-gbdt for Multi-step Ahead Forecasting of Surface Soil Moisture[J]. Engineering Applications of Artificial Intelligence,2023,120:105 895

[12] Wu K, Rodriguez G A, Zajc M, et al. A New Drone-borne GPR for Soil Moisture Mapping[J]. Remote Sensing of Environment,2019,235:111 456

[13] Malajner M, Gleich D, Planinsic P. Soil Type Characterization for Moisture Estimation Using Machine Learning and UWB-time of Flight Measurements[J]. Measurement, 2019,146:537-543

[14] Chen T, He T, Benesty M, et al. Xgboost:Extreme Gradient Boosting[J]. R Package Version 0. 4-2,2015,1(4):1-4

[15] Zhang Y, Liang S, Ma H, et al. Generation of Global 1 km Dai Soil Moisture Product from 2000 to 2020 Using Ensemble Learning[J]. Earth System Science Data Discussions, 2023:1-37

[16] Carmona P, Climent F, Momparler A. Predicting Failure in the US Banking Sector:An Extreme Gradient Boosting Approach[J]. International Review of Economics & Finance, 2019,61:304-323