

面向植被覆盖的土地利用类型分类方法

赵鲁康, 朱习军

青岛科技大学, 信息科学技术学院, 山东 青岛

收稿日期: 2024年4月17日; 录用日期: 2024年5月24日; 发布日期: 2024年5月31日

摘要

随着高光谱遥感图像和神经网络的应用, 基于高光谱遥感图像的植被覆盖类型分类进一步的发展。然而由于高光谱遥感图像具有更加丰富的光谱信息, 传统分类方法不能很好地同时提取光谱特征和空谱特征。为了解决上述问题, 本文以U-Net模型为框架, 融合了3D-CNN和多尺度特征提取模块, 提出了一种基于三维卷积和多尺度特征融合的神经网络模型, 使得网络在进行特征提取和分类的过程中, 更好的融合光谱信息、空间信息、全局信息和细节信息, 最终使模型的分类准确度得以提升。

关键词

多尺度特征, 三维卷积, 高光谱遥感图像, 植被覆盖

Land Use Type Classification Method for Vegetation Coverage

Lukang Zhao, Xijun Zhu

School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao Shandong

Received: Apr. 17th, 2024; accepted: May 24th, 2024; published: May 31st, 2024

Abstract

With the application of hyperspectral remote sensing images and neural networks, the classification of vegetation cover types based on hyperspectral remote sensing images has been further developed. However, since hyperspectral remote sensing images have more abundant spectral information, traditional classification methods cannot extract spectral features and spatial-spectral features at the same time. In order to solve the above problems, this paper takes the U-Net model as the framework, integrates the 3D-CNN and multi-scale feature extraction module, and proposes a neural network model based on three-dimensional convolution and multi-scale feature fusion.

In the process of feature extraction and classification, the network better integrates spectral information, spatial information, global information and detail information, and finally improves the classification accuracy of the model.

Keywords

Multi-Scale Features, 3D-CNN, Hyperspectral Remote Sensing Images, Vegetation Coverage

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

高光谱遥感图像(Hyperspectral Remote Sensing Image, HSI) [1]在土地利用分类领域具有广泛的应用前景,通过获取地球表面物体的连续光谱信息[1],高光谱遥感图像(HSI)能够提供丰富的光谱特征,从而实现对不同土地利用类型的准确分类。随着高光谱遥感图像(HSI)获取技术的发展,结合高光谱数据进行土地利用类型分类成为了研究的热点。

传统的土地利用分类方法往往基于单一或少数光谱波段,难以捕捉土地利用类型的细微差异。而高光谱遥感图像提供了丰富的光谱信息,能够更准确地区分不同的土地利用类型。通过充分利用高光谱图像的光谱特征,可以提高土地利用分类的精度和准确性。

通过植被覆盖土地类型分类的研究,我们可以用来支持环境保护和生态恢复。高光谱遥感图像的植被覆盖类型分类可以为环境保护和生态恢复提供支持。通过对高光谱图像中不同植被覆盖类型的识别和分析,可以评估生态系统的健康状况、生物多样性水平[2]以及植被覆盖程度[3]等。这有助于制定相应的环境保护策略和生态恢复计划,促进生态环境的改善和可持续发展。其次,还促进植被监测和管理。高光谱遥感图像的植被覆盖类型分类可以为植被监测和管理提供重要信息,包括植被种类、分布、健康状况等,有助于及时发现植被变化、评估生态环境和制定合理的植被保护措施。不仅如此,通过被覆盖土地类型分类的研究还可以为灾害监测和应急响应提供支持。植被覆盖类型分类技术可以用于自然灾害监测和应急响应,有助于指导灾害监测和救援工作,降低灾害对生命和财产的损失。

为了有效地利用高光谱遥感图像,研究人员开始开发各种光谱信息提取技术。这些技术包括光谱指数计算(如归一化差异植被指数[4]、水体指数[5]等)、主成分分析[6]、线性无关分解[7]等。这些方法能够从高光谱数据中提取有关地物光谱特征的关键信息。同时,高光谱遥感图像的数据量庞大,对数据处理和分析提出了挑战。为了降低数据维度和提高分类效率,研究人员开发了一系列数据处理方法,如特征选择、特征降维(如主成分分析和小波变换[8])、特征融合等。为了实现高准确性的城市土地利用分类,研究人员结合高光谱遥感图像和分类算法进行了深入研究。传统的分类方法包括最大似然分类[9]、支持向量机(SVM) [10]、决策树[11]等,但这些方法难以充分利用高光谱数据的丰富光谱信息和空间上下文关系。

近年来,随着深度学习技术的快速发展,基于高光谱遥感图像的土地分类也取得了显著的进展。深度学习模型如卷积神经网络(CNN) [12]和循环神经网络(RNN) [13]被广泛应用于土地利用分类任务中。这些模型能够自动学习地物的特征表示,并在分类过程中捕捉地物之间的复杂关系,从而提高分类的准确性和鲁棒性[14]。此外,还有一些新兴的技术应用于高光谱遥感图像的土地利用分类中。例如,基于图卷积网络(GCN) [15]的方法可以有效地利用地物之间的空间关系。同时,多源数据融合技术结合高光谱遥感

图像与其他数据源(如激光雷达数据、热红外图像等)进行融合, 提供更全面、准确的土地利用分类结果。

2. 基于三维卷积和多尺度特征融合的神经网络模型

2.1. 基于三维卷积和多尺度特征融合的神经网络模型设计

本文以 U-Net 网络模型为基础, 从高光谱遥感图像的特点出发, 设计了一种基于三维卷积和多尺度特征融合的神经网络模型(Neural network model based on 3D convolution and multi-scale feature fusion, MS-3DNet), 其具体结构见图 1。

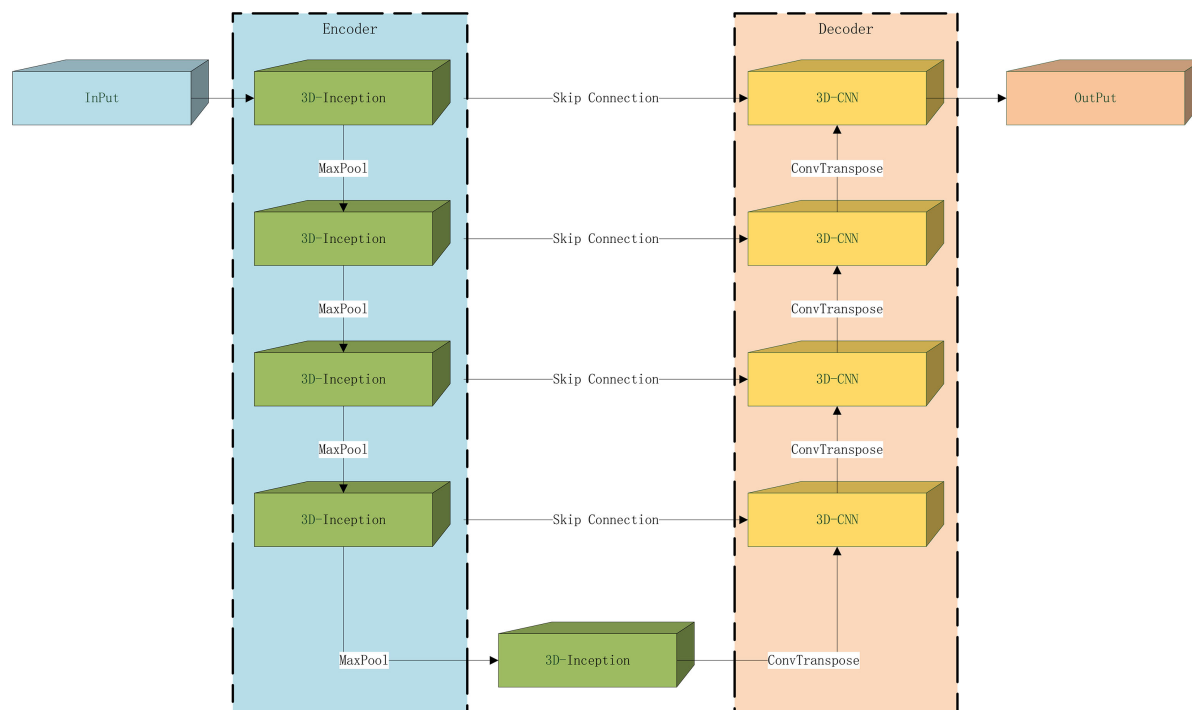


Figure 1. Structure diagram of neural network model based on 3D-multi-scale feature fusion

图 1. 基于三维卷积和多尺度特征融合的神经网络模型结构图

在图 1 中, 高光谱遥感图像首先通过 Encoder 模块进行特征提取, 在特征提取阶段采用三维多尺度特征提取模块(3D-Inception)进行特征提取, 提取的特征立方体经过最大化池化(MaxPool)后输入到下一层 3D-Inception, 同时特征体等待通过跳跃连接输入到同层的 Decoder 中进行特征融合。通过 3D-Inception 模块进行特征提取, 不仅可以通过三维卷积同时提取特征图的光谱特征和空间特征, 还可以通过多尺度的感受野提取不同尺度的特征。

2.2. 三维卷积

三维卷积[16]在处理三维数据时, 与二维卷积相比, 三维卷积考虑了数据的空间维度, 即在三个维度上进行滤波操作, 这使得三维卷积能够更好地捕获数据的立体特征。由于三维卷积在三个维度上进行操作, 因此能够更好地捕获数据中的立体特征。与二维卷积类似, 三维卷积也具有参数共享的特性, 即在整个输入数据的不同位置上使用相同的卷积核进行滤波操作。这有助于减少模型的参数数量, 降低过拟合的风险, 并提高模型的泛化能力。三维卷积可以通过多层叠加的方式提取数据的立体特征, 从而构建更复杂的特征表示。通过堆叠多个三维卷积层和池化层, 可以逐渐提取数据中的抽象和复杂的立体特征。

三维卷积最重要的特征便是采用了三维卷积核, 使其能够在特征立方体三个维度上同时提取特征。具体结构见图 2。

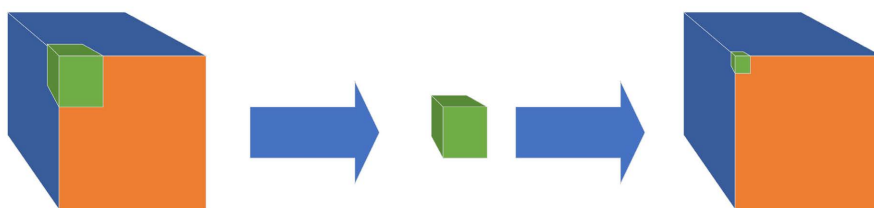


Figure 2. Schematic diagram of 3D convolution
图 2. 三维卷积示意图

2.3. 基于多尺度特征提取的编码器结构设计

由于植被覆盖分类研究中的某些地物可能具有相似的光谱特征, 如不同种类的作物可能具有相似的反射谱, 这会增加农业用地类型分割的难度。但是由于在空间尺寸上植被覆盖地块通常具有较大的空间分辨率, 空间信息对于网络模型具有较大作用, 为了提高对空间信息的提取, 本文提出的基于三维卷积和多尺度特征融合的神经网络模型(Networks based on 3D-CNN and Different Scales, DS-3DNet)对编码器进行了设计, 即采用不同尺度的卷积核感受野对空间特征进行提取, 即 Inception 模块[17]进行特征提取。

具体来说, 编码器部分采用 $1 \times 1 \times 1$ 、 $3 \times 3 \times 3$ 、 $5 \times 5 \times 5$ 、MaxPool 四种不同的感受野, 具体结构见图 3。

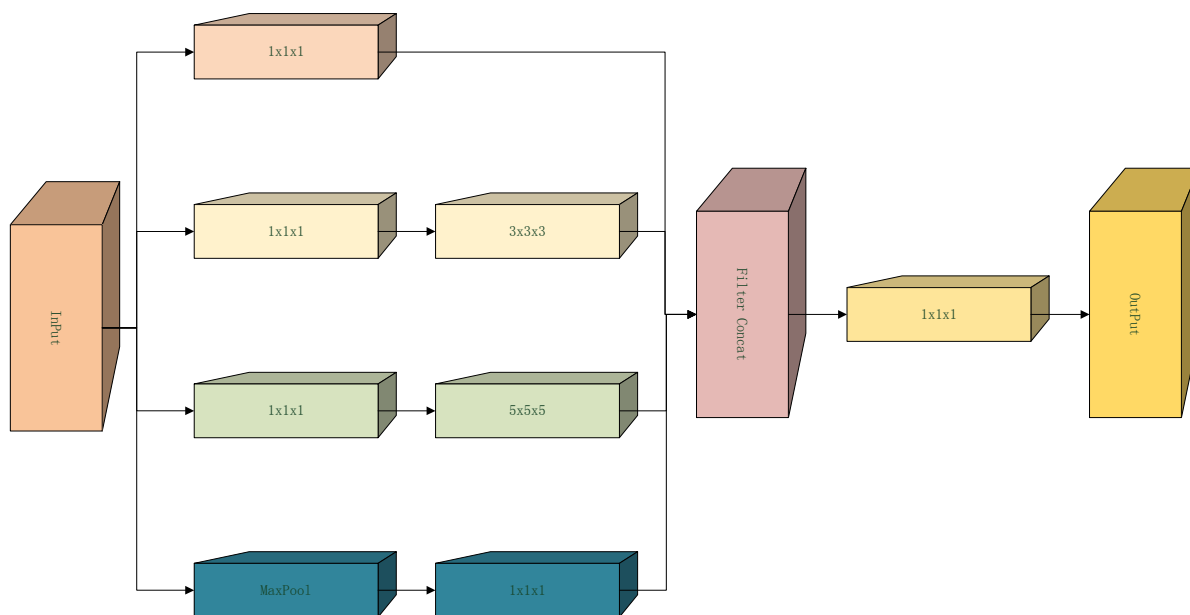


Figure 3. Structure diagram of a multi-scale feature extraction encoder
图 3. 多尺度特征提取编码器结构图

2.4. 基于 U-Net 结构的解码器结构设计

U-Net 网络架构通过不同深度的特征图融合, 既能够关注到深层特征图的全局信息, 又能关注到浅层特征图的细节信息, 适合语义分割[18]等任务, 因此参考了 U-Net 网络解码器的设计, 将不同深度的特

征图通过转置卷积[19]逐步恢复空间分辨率, 并通过跳跃连接将同层次的编码器提取的特征图进行融合, 使得网络模型加强对全局信息和细节信息的同时关注, 提升网络的性能, 其解码器部分每层结构见图 4。

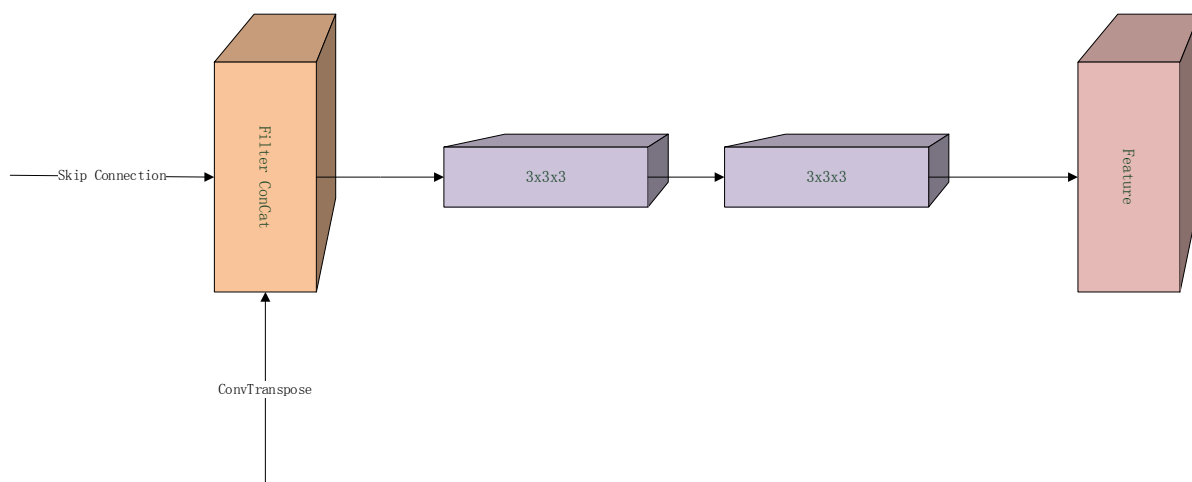


Figure 4. Encoder structure diagram
图 4. 编码器结构图

经过 4 层编码器进行处理后, 图像恢复至原来的空间分辨率, 输入到输出层进行分类处理。

3. 实验结果与分析

本文基础实验环境如下: CPU 6xE5-2680 V4; GPU NVIDIA GeForce RTX A4000, 显存 16.9 GB; 内存 32 GB; PyTorch v2.0.1, Python v3.10.12, Linux 操作系统。

3.1. 数据集介绍

IndianPines 数据集是使用美国航空航天局(NASA)的 AVIRIS (Airborne Visible/Infrared Imaging Spectrometer)传感器采集的高光谱图像数据。采集自美国印第安纳州西北部地区, 主要用于研究农田覆盖和土地利用。

IndianPines 数据集图像尺寸为 145×145 像素。图像覆盖了一个农村地区, 包括田地、森林、道路等地物。IndianPines 数据集包含了 224 个连续的光谱波段, 其中 20 个波段收到噪声的影响而被去除, 涵盖了可见光、红外线和近红外线等波段。这些光谱波段的数据被用来描述不同地物的光谱特征。

IndianPines 数据集包含 145×145 像素, 包含 16 个类别。其中分布最多的种类是 Soybean-mintill, 数量达 2455 个, 占比 23.95%; 分布最少的类别是 Oats, 数量仅 20, 占比 0.20%。为了更清楚直观地观察 IndianPines 数据集中的样本类别的分布, 假彩图和标签图见图 5。

3.2. 评价指标

3.2.1. 总体精度(OA)

总体精度(Overall Accuracy) [20]是指分类模型在所有样本上的分类正确率, 即被正确分类的样本数占总样本数的比例。以二分类问题为例, 总体精度计算公式为:

$$\text{OverallAccuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

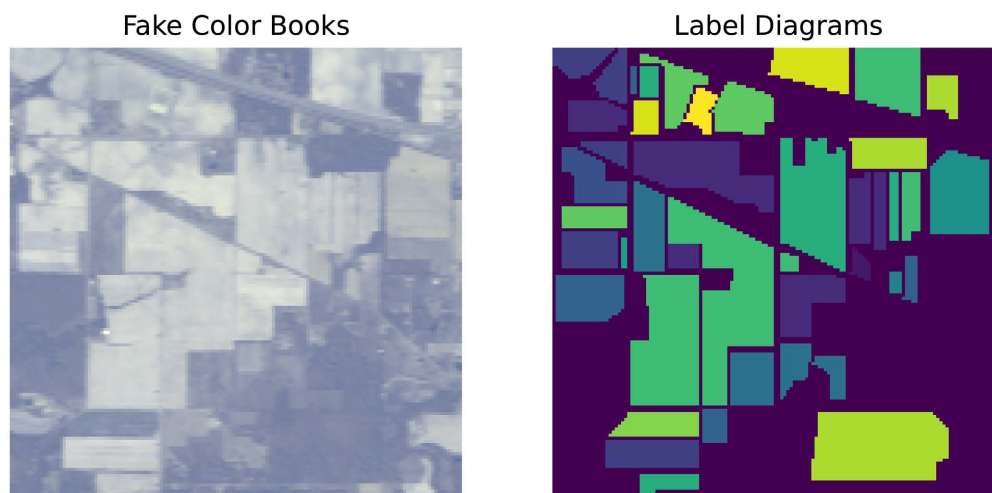


Figure 5. Indian Pines dataset
图 5. Indian Pines 数据集

其中: TP——预测为正类且实际为正类的样本数。

TN——预测为负类且实际为负类的样本数。

FP——预测为正类但实际为负类的样本数。

FN——预测为负类但实际为正类的样本数。

3.2.2. 平均精度(AA)

平均精度(Average Accuracy) [20]通常用于多类别分类任务, 是每个类别的分类准确率的平均值。与总体精度不同, 平均精度考虑了每个类别的分类准确率, 因此更能反映模型在不同类别上的分类表现。

仍以二分类问题为例, 计算平均精度的步骤如下:

首先, 计算每个类别的精确率其计算公示如下:

$$\text{Accuracy}_0 = \frac{\text{TN}}{\text{TN} + \text{FN}} \quad (2)$$

$$\text{Accuracy}_1 = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

然后, 将每个类别的分类指标值相加, 并除以类别数, 得到平均精度。

$$\text{AverageAccuracy} = \frac{1}{N} \sum_{i=0}^{N-1} \text{Accuracy}_i \quad (4)$$

3.2.3. Kappa 系数

Kappa 系数(Cohen's Kappa coefficient) [20]是一种用于评估分类模型性能的统计量, 它考虑了模型预测与实际观察之间的一致性, 而不仅仅是预测的准确率。Kappa 系数通常用于衡量分类模型在多类别分类问题上的性能。Kappa 系数的计算基于混淆矩阵, 其公式如下:

$$\text{Kappa} = \frac{P_0 + P_e}{1 - P_e} \quad (5)$$

其中, P_0 表示观察到的准确率, 即模型预测与实际观察之间的一致性, P_e 表示随机准确率, 即模型预测与实际观察之间的一致性的期望值。Kappa 系数的取值范围为[-1, 1], 其含义如下:

$\text{Kappa} = 1$: 完美一致, 模型的预测与实际观察完全一致。

$\text{Kappa} = 0$: 随机一致, 模型的预测与实际观察的一致性与随机预测的一致性相同。

$\text{Kappa} < 0$: 一致性低于随机水平, 模型的预测与实际观察的一致性低于随机预测的一致性。

$\text{Kappa} > 0$: 一致性高于随机水平, 模型的预测与实际观察的一致性高于随机预测的一致性。

Kappa 系数的计算涉及两个参数: 观察到的准确率(P_0)和随机准确率(P_e), 对于一个具有 K 个类别的分类问题, 混淆矩阵的大小为它们的计算方法如下:

观察到的准确率(P_0)可以通过混淆矩阵中的对角线元素之和除以总样本数来计算:

$$P_0 = \frac{\sum_{i=1}^K n_{ii}}{\sum_{i=1}^K \sum_{j=1}^K n_{ij}} \quad (6)$$

其中, n_{ij} 表示混淆矩阵中第 i 行第 j 的元素。

随机准确率(P_e)表示模型的预测与实际观察之间的一致性的期望值, 即在没有模型的情况下, 预测与实际观察之间的一致性。计算方法如下:

$$P_e = \frac{\sum_{i=1}^K n_{i+} \times n_{+i}}{\left(\sum_{i=1}^K n_{i+}\right)^2} \quad (7)$$

3.3. 对比实验

在对比实验中, 使用 IP 公开数据集进行实验比较, 以评估本文提出的基于三维卷积和多尺度特征融合的神经网络模型(Neural network model based on 3D convolution and multi-scale feature fusion, MS-3DNet)与其他文献提出的方法之间的差异。

实验过程中, 选取数据集中 20% 的样本作为训练集和验证集(训练集和验证集比例为 1:1), 其余样本用作测试集, 评估指标采用总体准确度(OA)、平均准确度(AA)、 Kappa 系数, 实验结果见表 1、图 6、图 7。

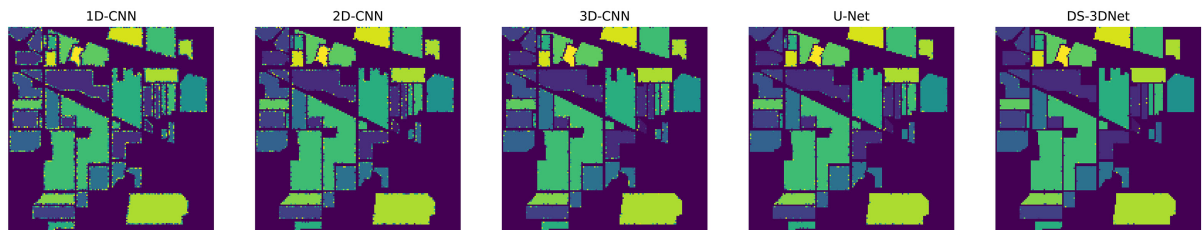


Figure 6. Classification results of different network models in IP datasets

图 6. 不同网络模型在 IP 数据集的分类结果

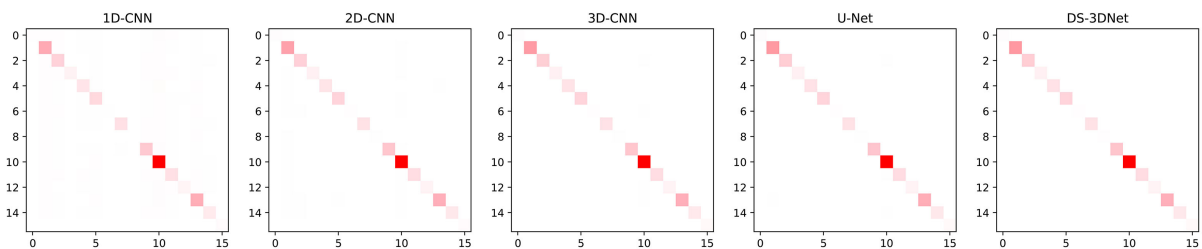


Figure 7. Heatmap of the classification results of different models

图 7. 不同模型分类结果热力图

根据表 1 可知, 本文提出的基于三维卷积和多尺度特征融合的神经网络模型(DS-3DNet)算法在评价指标平均精度(AA)、总体精度(OA)、Kappa 系数分别为 98.20%、98.50%、97.95%, 各项评价指标均高于相关文献提及到算法。

Table 1. IP dataset classification results

表 1. IP 数据集分类结果

评价指标	1D-CNN	2D-CNN	3D-CNN	U-Net	DS-3DNet
OA	83.67%	92.15%	93.25%	95.22%	98.20%
AA	76.28%	90.37%	94.13%	95.12%	98.50%
Kappa	80.46%	92.18%	94.14%	94.53%	97.95%

3.4. 消融实验

3.4.1. 三维卷积模块

为了验证三维卷积模块的对实验结果产生影响, 本文采用原始 U-Net 模型、采用三维卷积的 U-Net 模型、原始 ResNet 模型[21]、采用三维卷积的 ResNet 模型、原始 FCN 模型[21]和采用三维卷积的 FCN 模型进行消融实验, 实验结果见表 2、图 8、图 9。

Table 2. Statistics of ablation experiment results for the 3D convolution module

表 2. 针对三维卷积模块的消融实验结果统计

Model	OA (%)	AA (%)	Kappa (%)
U-Net	95.74	94.66	95.15
U-Net + 3D-CNN	96.80	96.18	96.36
ResNet	92.52	89.88	91.51
ResNet + 3D-CNN	92.89	90.46	91.93
FCN	94.45	91.75	93.69
FCN + 3D-CNN	95.19	93.28	94.53

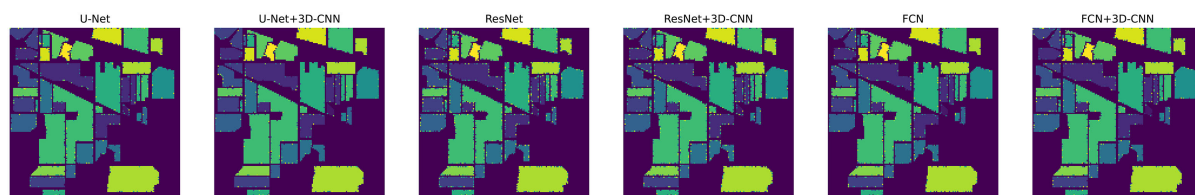


Figure 8. Segmentation results of ablation experiments of different models

图 8. 不同模型消融实验分割结果

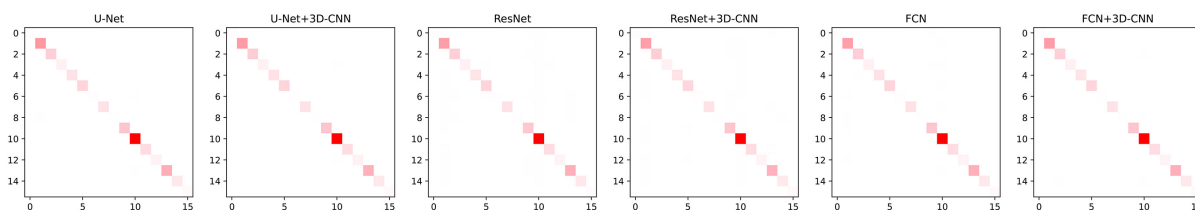


Figure 9. Heat map of the classification results of different models of ablation experiments

图 9. 不同模型消融实验分类结果热力图

通过表 2 针对三维卷积模块的消融实验结果统计可知, 三维卷积模块在处理高光谱遥感图像的过程中, 可以更好的利用光谱信息, 提升模型的分类的精度性能, 通过引入三维卷积模块, U-Net 模型的总体精度(OA)提升 1.06%, 平均精度(AA)提升 1.52%, Kappa 系数提升 1.21%; ResNet 模型的总体精度(OA)提升 0.37%, 平均精度(AA)提升 0.58%, Kappa 系数提升 0.42%; FCN 模型的总体精度(OA)提升 0.74%, 平均精度(AA)提升 1.53%, Kappa 系数提升 0.84%。

3.4.2. 基于不同尺度特征提取(Multiscale Feature Extraction, MFE)的编码器模块

为了验证基于不同尺度特征提取的编码器模块对实验结果产生的影响, 本文采用原始 U-Net 模型、采用基于不同尺度特征提取的编码器模块的 U-Net 模型、原始 ResNet 模型、采用基于不同尺度特征提取的编码器模块的 ResNet 模型、原始 FCN 模型和采用基于不同尺度特征提取的编码器模块的 FCN 模型进行消融实验, 实验结果见表 3、图 10、图 11。

Table 3. Statistics of ablation experimental results based on multi-scale feature extraction

表 3. 针对基于不同尺度特征提取的编码器模块的消融实验结果统计

Model	OA (%)	AA (%)	Kappa (%)
U-Net	95.74	94.66	95.15
U-Net + MFE	96.25	96.13	95.94
ResNet	92.52	89.88	91.51
ResNet + MFE	93.24	93.16	93.40
FCN	94.45	91.75	93.69
FCN+MFE	94.15	94.33	94.71

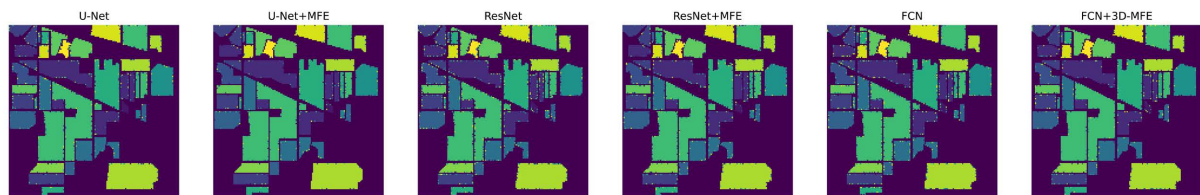


Figure 10. Segmentation results of ablation experiments of different models

图 10. 不同模型消融实验分割结果

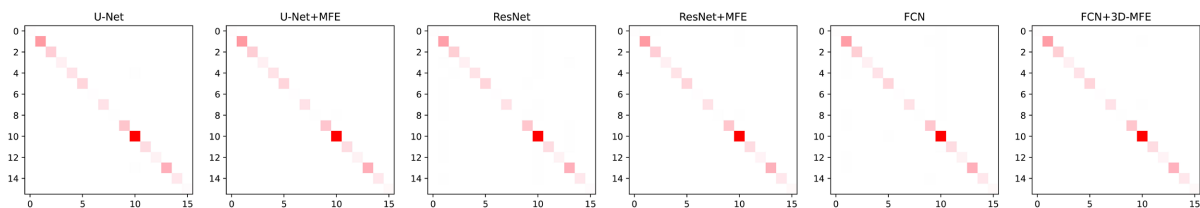


Figure 11. Heat map of the classification results of different models of ablation experiments

图 11. 不同模型消融实验分类结果热力图

通过表 3 可知, 基于不同尺度特征提取的编码器模块在处理高光谱遥感图像的过程中, 可以更好的发掘空间信息, 提升模型的分类的精度性能, 通过引入基于不同尺度特征提取的编码器模块, U-Net 模型的总体精度(OA)提升 1.03%, 平均精度(AA)提升 1.01%, Kappa 系数提升 1.41%; ResNet 模型的总体精度(OA)提升 1.89%, 平均精度(AA)提升 1.79%, Kappa 系数提升 1.22%; FCN 模型的总体精度(OA)提升 0.43%, 平均精度(AA)提升 0.91%, Kappa 系数提升 1.05%。

3.4.3. 基于 U-Net 结构的解码器模块

为了验证基于不同尺度特征提取的编码器模块对实验结果产生的影响, 本文针对提出的基于三维卷积和多尺度特征融合的神经网络模型(DS-3DNet)中的跳跃连接[22]设计了消融实验, 分别是不保留跳跃连接的 DS-3DNet 模型、保留最后 1 层的跳跃连接的 DS-3DNet 模型、保留最后 2 层的跳跃连接的 DS-3DNet 模型、保留最后 3 层的跳跃连接的 DS-3DNet 模型、保留全部跳跃连接的 DS-3DNet 模型。实验结果见表 4、图 12、图 13。

Table 4. Statistics of ablation experimental results based on multi-scale feature extraction
表 4. 针对基于不同尺度特征提取的编码器模块的消融实验结果统计

Model	OA (%)	AA (%)	Kappa (%)
No Skip Connections	86.14	86.55	85.67
Retain the last Skip Connections	91.44	90.25	90.31
Retain the last Two Skip Connections	95.32	94.87	94.76
Retain the last Three Skip Connections	97.11	96.52	96.56
Retain All Skip Connections	98.20	98.50	97.95



Figure 12. Segmentation results of ablation experiments of different models
图 12. 不同模型消融实验分割结果

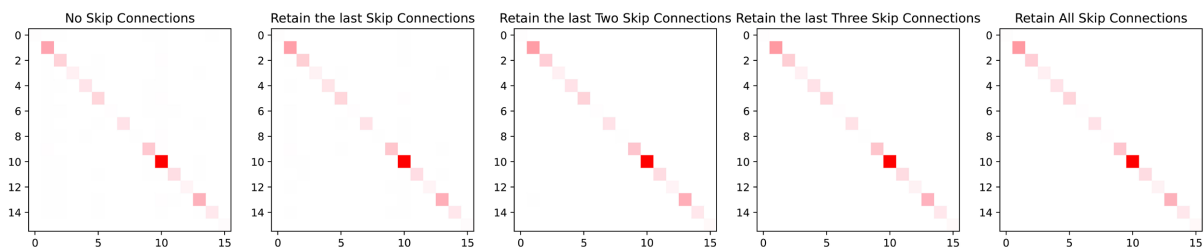


Figure 13. Heat map of the classification results of different models of ablation experiments
图 13. 不同模型消融实验分类结果热力图

通过表 4 可以看出, 不保留跳跃连接 DS-3DNet 模型和保留全部跳跃连接的 DS-3DNet 模型在模型分类的准确度上存在较大差异, 两者在总体精度(OA)、平均精度(AA)、Kappa 系数等评价指标上的差值分别为: 12.06%、11.95%、12.28%, 差值较大, 提升比较明显, 同时可以看出每增加一层跳跃连接, 总体精度(OA)提升分别为 5.30%、3.88%、1.79%、1.09%, 平均精度(AA)提升分别为 3.70%、4.62%、1.65%、1.98%, Kappa 系数提升分别为 4.64%、4.45%、1.80%、1.39%, 通过分析可以看出随着跳跃连接的增加, 跳跃连接对于模型性能提升的影响逐步降低。

4. 结束语

由于传统卷积神经网络随着网络深度的加深, 模型更加注重全局信息, 导致局部细节丢失, 而且对

于高光谱遥感图像来说, 因其具有更加丰富的光谱信息, 传统的 1D-CNN、2D-CNN 不能很好的同时兼顾光谱信息和空间信息。同时在面向植被覆盖利用类型的高光谱遥感图像分割场景时, 由于植被类型通常拥有相似的光谱特征, 不易通过光谱信息进行区分, 而在空间特征对于高光光谱遥感图像分割具有较大的增益特征, 为了解决上述问题, 本文以 U-Net 网络模型为主要框架融合了 3D-CNN 网络模型和不同尺度特征提取融合的网络模型, 提出了基于三维卷积和多尺度特征融合的神经网络模型(DS-3DNet), 使得网络在进行特征提取和分类的过程中, 更好的融合光谱信息、空间信息、全局信息和细节信息, 使得网络在分类精度方面有了一定的提升。

参考文献

- [1] 杜培军, 夏俊士, 薛朝辉, 等. 高光谱遥感影像分类研究进展[J]. 遥感学报, 2016, 20(2): 236-256.
- [2] 马克平. 试论生物多样性的概念[J]. 生物多样性, 1993, 1(1): 20-22.
- [3] 程红芳, 章文波, 陈锋. 植被覆盖度遥感估算方法研究进展[J]. 国土资源遥感, 2008, 1(1): 16-21.
- [4] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. 遥感学报, 2005, 9(5): 589-595.
- [5] 虞晓芬, 傅玳. 多指标综合评价方法综述[J]. 统计与决策, 2004(11): 119-121.
- [6] Abdi, H. and Williams, L.J. (2010) Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2, 433-459. <https://doi.org/10.1002/wics.101>
- [7] Singer, A. and Coifman, R.R. (2008) Non-Linear Independent Component Analysis with Diffusion Maps. *Applied and Computational Harmonic Analysis*, 25, 226-239. <https://doi.org/10.1016/j.acha.2007.11.001>
- [8] 张静远, 张冰, 蒋兴舟. 基于小波变换的特征提取方法分析[J]. 信号处理, 2000, 16(2): 156-162.
- [9] 骆剑承, 王钦敏, 马江洪, 等. 遥感图像最大似然分类方法的 EM 改进算法[J]. 测绘学报, 2002, 31(3): 234-239.
- [10] Cherkassky, V. and Ma, Y. (2004) Practical Selection of SVM Parameters and Noise Estimation for SVM Regression. *Neural Networks*, 17, 113-126. [https://doi.org/10.1016/S0893-6080\(03\)00169-2](https://doi.org/10.1016/S0893-6080(03)00169-2)
- [11] 栾丽华, 吉根林. 决策树分类技术研究[J]. 计算机工程, 2004, 30(9): 94-96.
- [12] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [13] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1-6, 26.
- [14] 纪守领, 杜天宇, 邓水光, 等. 深度学习模型鲁棒性研究综述[J]. 计算机学报, 2022, 45(1): 190-206.
- [15] 徐冰冰, 岑科廷, 黄俊杰, 等. 图卷积神经网络综述[J]. 计算机学报, 2020, 43(5): 755-780.
- [16] 张开放, 苏华友, 窦勇. 一种基于混淆矩阵的多分类任务准确率评估新方法[J]. 计算机工程与科学, 2021, 43(11): 1910-1919.
- [17] 张珂, 冯晓晗, 郭玉荣, 等. 图像分类的深度卷积神经网络模型综述[J]. 中国图象图形学报, 2021, 26(10): 2305-2325.
- [18] Lu, S., Ding, Y., Liu, M., et al. (2023) Multiscale Feature Extraction and Fusion of Image and Text in VQA. *International Journal of Computational Intelligence Systems*, 16, Article No. 54. <https://doi.org/10.1007/s44196-023-00233-6>
- [19] 徐辉, 祝玉华, 甄彤, 等. 深度神经网络图像语义分割方法综述[J]. 计算机科学与探索, 2021, 15(1): 47-59.
- [20] 李荣陆. 文本分类及其相关技术研究[D]: [博士学位论文]. 上海: 复旦大学, 2005.
- [21] 李炳臻, 刘克, 顾佼佼, 等. 卷积神经网络研究综述[J]. 计算机时代, 2021(4): 8-12, 17. <https://doi.org/10.16644/j.cnki.cn33-1094/tp.2021.04.003>
- [22] 周涛, 董雅丽, 霍兵强, 等. U-Net 网络医学图像分割应用综述[J]. 中国图象图形学报, 2021, 26(9): 2058-2077.