

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2021)08-1994-15

论文引用格式: Cheng S Y, Xiong Y J, Yao Y and Li Q L. 2021. Hyperspectral image classification using an inverted feature pyramid network with U-Net. *Journal of Image and Graphics* 26(08): 1994-2008(程嵩阳 熊玉洁 姚瑶 李庆利. 2021. 结合倒置特征金字塔和 U-Net 的高光谱图像分类. *中国图象图形学报* 26(08): 1994-2008) [DOI: 10. 11834/jig. 210194]

结合倒置特征金字塔和 U-Net 的高光谱图像分类

程嵩阳¹ 熊玉洁^{1,2*} 姚瑶¹ 李庆利²

1. 上海工程技术大学电子电气工程学院, 上海 201620; 2. 华东师范大学上海市多维度信息处理重点实验室, 上海 200241

摘要: 目的 地物分类是对地观测研究领域的重要任务。高光谱图像具有丰富的地物光谱信息,可用于提升遥感图像地物分类的准确度。如何对高光谱图像进行有效的特征提取与表示是高光谱图像分类应用的关键问题。为此,本文提出了一种结合倒置特征金字塔和 U-Net 的高光谱图像分类方法。方法 对高光谱数据进行主成分分析(principal component analysis, PCA)降维,获取作为网络输入的重构图像数据,然后使用 U-Net 逐层提取高光谱重构图像的空间特征。与此同时,利用倒置的特征金字塔网络抽取相应层级的语义特征;通过特征融合,得到既有丰富的空间信息又有较强烈的语义响应的特征表示。提出的网络利用注意力机制在跳跃连接过程中实现对背景区域的特征响应抑制,最终实现了较高的地物分类精度。结果 分析了 PCA 降维方法和输入数据尺寸对分类性能的影响,并在 Indian Pines、Pavia University、Salinas 和 Urban 数据集上进行了对比实验。本文方法在 4 个数据集上分别取得了 98.91%、99.85%、99.99% 和 87.43% 的总体分类精度,与支持向量机(support vector machine, SVM)等相关算法相比,分类精度高出 1%~15%。结论 本文提出一种结合倒置特征金字塔和 U-Net 的高光谱图像分类方法,可以应用于有限训练样本下的高光谱图像分类任务,并在多个数据集上取得了较高的分类精度。实验结果表明倒置特征金字塔结构与 U-Net 结合的算法能够高效地实现高光谱图像的特征提取与表示,从而获得更精细的分类结果。

关键词: 高光谱图像分类; 稀少样本; 倒置特征金字塔网络(IFPN); U-Net; 特征融合

Hyperspectral image classification using an inverted feature pyramid network with U-Net

Cheng Songyang¹, Xiong Yujie^{1,2*}, Yao Yao¹, Li Qingli²

1. School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;

2. Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China

Abstract: Objective Terrain classification is an important research task in the field of earth observation using remote sensing technology. The hyperspectral image has rich spectral information; thus, it can be applied to the classification of remote sensing image. With the rapid development of the hyperspectral technology, the hyperspectral remote sensing image processing and analyzing technology has attracted wide attention of academia. The hyperspectral images have dozens or even

收稿日期: 2021-03-23; 修回日期: 2021-05-11; 预印本日期: 2021-05-18

* 通信作者: 熊玉洁 xiong@sues.edu.cn

基金项目: 国家重点研发计划项目(2019YFC1509202); 国家自然科学基金项目(62006150); 上海青年科技英才扬帆计划项目(19YF1418400); 上海市多维度信息处理重点实验室开放课题基金资助项目(2020MIP001); 中央高校基本科研业务费专项资金资助

Supported by: National Key Research and Development Program of China (2019YFC1509202); National Natural Science Foundation of China (62006150); Shanghai Young Science and Technology Talents Sailing Program(19YF1418400)

hundreds of continuous narrow spectral bands compared with the traditional panchromatic band and multi-spectral remote sensing image, which provides detailed spectral and spatial feature information. Accordingly, these images have been widely used in various aspects, such as precision agriculture, city planning, and military defense. Hyperspectral images have high dimensional data, and redundancy and noise exist; thus, transformed data must be utilized for image processing. In the application of hyperspectral image classification, the manner by which to effectively represent the features of hyperspectral image is the most critical step in current studies. In this work, we propose an approach for hyperspectral image classification by using an inverted feature pyramid network and U-Net. **Method** The dimension of the hyperspectral remote sensing image data is high. Principal component analysis (PCA) method plays a significant role in transforming useful information in the images to the most important k characteristic, thus reducing the amount of data and enhancing the data features. After PCA, the data are segmented and collected by means of sliding window. The surrounding area of each pixel is defined as a patch, which is regarded as the input of the proposed network. The category of the pixel is the ground truth label. In the first stage, U-Net is used to extract spatial features of hyperspectral image at the pixel level. The left side of the network is the contraction path, which corresponds to the encoder part of the classic encoder-decoder. The right side of the network is the extension path, which can be regarded as a decoder. The feature maps in the extension path are the result of combining two parts of a feature map along two dimensions, making the acquired features more visible. In the first part, the feature maps from the same layer of contraction path and the feature maps from the upper layer of extension path are simultaneously fed to the attention mechanism. The feature region of this part has a higher weight value. The second part is obtained by deconvolution of the feature graph from the upper layer of the extension path. In a layered way, these feature maps with rich spatial information are fused with feature maps containing rich semantic information obtained by inverted feature pyramid network layers. Therefore, the obtained feature maps have reliable spatial and strong semantic information. Finally, the weight value of the effective features in the image is increased, and the region of irrelevant background is suppressed owing to the attention mechanism. Thus, the classification result of hyperspectral image is acquired. **Result** We conduct experiments to evaluate the effectiveness of the proposed method and attempt to investigate the influence of PCA retained principal component number and the size of input data for the performance of classification. We conduct contrast experiments on four publicly available hyperspectral image datasets to demonstrate the performance of the proposed method: Indian Pines, Pavia University, Salinas, and Urban. Experimental results show that the proposed method for hyperspectral image classification is effective, and the best PCA retained principal component numbers are 3, 20, 10, and 3. Meanwhile, the best input sizes of the proposed model are 64, 32, 32, and 64. We obtain 98.91%, 99.85%, 99.99%, and 87.43% overall classification accuracy rates, 98.07%, 99.39%, 99.09%, and 78.30% average classification accuracy rates, and 0.987, 0.998, 0.999, and 0.831 Kappa values for the four hyperspectral image datasets, respectively, which are higher than those of the other classification algorithms. **Conclusion** Hyperspectral images are capable of accurately presenting the rich terrain information contained in the specific region with the help of hundreds of continuous and subdivided spectral bands; however, useless information exists in each spectral band. The mechanism by which to effectively extract the key terrain information from the data of hyperspectral images and utilize them for classification is the most important and difficult problem. We propose to combine U-Net and the inverted pyramid network for hyperspectral image classification. First, we reduce the dimension of hyperspectral image data with the help of PCA method. We adopt the method of sliding window to build patches after the data dimension is reduced. These patches are fed into the model. U-Net is regarded as the backbone of the proposed network, and it aims to extract the characteristics of a hyperspectral image. Then, the rich characteristics of the spatial information are fused with the features from the inverted pyramid network. Subsequently, the abundant spectral and spatial information is obtained. The utilization of attention mechanism allows the model to effectively focus on spectral and spatial information and reduce the influence of signal-to-noise to classification performance. Experimental results show that the proposed method can be applied to hyperspectral image classification tasks with limited training samples and achieve good classification results. The classification accuracy of a hyperspectral image can also be improved by properly handling the input data. In our future work, we will attempt to investigate the manner by which to make the model's structure less complex while maintaining high hyperspectral image classification performance with less training data samples.

Key words: hyperspectral image classification; small samples; inverted feature pyramid network (IFPN); U-Net; feature fusion

0 引言

随着高光谱成像技术的不断发展,基于高光谱遥感图像的处理与分析技术得到了学术界和工业界的广泛重视。与传统的全色波段和多光谱遥感图像相比,高光谱图像具有几十甚至数百个连续的窄光谱波段,能够提供地物详细的光谱和空间特征信息。因此,其在城市规划、精准农业和军事国防等领域得到了广泛应用(肖亮等 2020)。但高光谱图像数据维度高、数据量大,并且存在冗余性和噪声,如何准确地利用光谱空间特征进行图像处理与分类是当前研究的重点与难点。

遥感高光谱图像波段数目大,信息复杂度高,从原始高光谱数据有效地进行特征提取至关重要。同时考虑光谱和空间联合信息的手工特征提取方法成为遥感图像分类领域研究的热点。为了表示复杂的空间光谱特征,Luo 等人(2019)提出了一种基于空间光谱信息、判别信息和超图学习的空间光谱超图判别分析特征学习算法,该算法利用最优投影矩阵来提取高光谱图像的空间光谱特征,以获得更好的分类精度。Jia 等人(2020)提出了一种基于级联超像素正则化 Gabor 特征融合的高光谱图像分类方法,替代了分别反映局部区域表面物质强度和变化的 Gabor 幅值和相位特征。Peng 等人(2019)提出了一种基于最大似然估计的高光谱图像联合稀疏表示方法,在空间邻域存在非均匀像素的情况下,该方法具有较好的分类效果。He 等人(2019)提出了一种通过构建多尺度协方差图的手工特征提取方法,该方法不仅增加了训练样本所表达的信息量,而且自然地将原始场景中包含的光谱和空间信息整合在一起,提高了高光谱图像的分类能力。然而,手工特征提取的方法往往关注局部特征,且通常是将数据的光谱和空间特征分开提取分析,导致有用信息的丢失,难以获得更好的全局特征对图像进行分类。利用卷积神经网络对高光谱图像进行深度特征提取的算法得到了广泛关注。Paoletti 等人(2019a)提出了一种新的基于光谱空间胶囊网络的卷积神经网络(convolutional neural networks, CNN),使其专门用于遥感高光谱数据高精度分类,同时显著降低网络结

构的复杂性。Paoletti 等人(2019b)同时针对高光谱图像数据设计了另一种 CNN 结构,致力于改善光谱空间特征的信息。该模型利用基于残差的方法逐步增加所有卷积层的特征维度,形成金字塔状,使得高光谱空间属性的多样性可以跨层增加,从而提高网络的性能。虽然卷积神经网络可以挖掘高光谱数据中更深、更为抽象的特征,但面对数据的高复杂度与信息冗余特点,模型存在过拟合的可能。针对此类问题,Yu 等人(2021)提出了一个带有反馈注意机制的空间光谱密集 CNN 框架,将频谱-空间特征以紧凑连接的方式组合起来,通过两个独立的密集 CNN 网络提取足够的信息。Zhu 等人(2021)以残差网络为基准学习光谱和空间特征,提出了端到端的残差光谱空间注意力网络,使用软注意机制实现光谱波段和空间信息的选择和增强。Haut 等人(2019)将注意力机制与残差网络结构结合在一起,表征数据中包含的光谱空间信息,提高了模型的特征提取能力。

在高光谱图像特征得到有效提取之后,如何进行准确的分类是研究重点,为此,许多研究者从分类器的角度提出了有效的方案。分类器可分为光谱空间分类器和光谱分类器。Zhang 等人(2018)提出了一种基于光谱空间分类器思想的分类网络,该网络利用不同区域的输入对语义情境感知表示进行编码,具有更强的判别能力。Jia 等人(2018)提出了一种基于 3 维 Gabor 小波的相位编码和汉明距离匹配框架,该方法利用平行于光谱轴方向的 Gabor 相位特征与象限比特编码方案,采用归一化的汉明距离匹配方法确定两个样本的相似性,使用最近邻分类器进行像素级识别,从而进行高光谱图像的分类,是一种基于光谱分类器的分类方法。此外,如何有效地使用分类器对提取的特征进行分类也得到了研究。Liu 等人(2020)提出将广义张量回归的方法用于高光谱图像的分类。该方法利用张量表示的优点,将一种简单有效的分类器扩展至其张量版本,然后利用不同模式的识别信息增强模型的表达能力。李玉等人(2019)将熵信息测度和类间距离测度引入 K-means 聚类函数,通过最小化目标函数得到最优分类结果。刘洋等人(2016)基于像元的空间连续性,创新性地使用 L2 稀疏编码的最小误差和编码

系数进行高光谱图像分类。

在实际研究和遥感应用中,只有少量标记的数据可用于模型训练,如何利用有限的标记数据实现高光谱图像分类也是一个具有挑战性的问题。Wu 和 Prasad(2018)提出了一种用于高光谱图像分类的半监督深度学习方法,该方法利用高光谱图像中有限的标记数据和丰富的未标记数据,能够有效地训练深度神经网络。刘万军等人(2019)针对小样本高光谱图像数据分类精度低的问题提出了一种复杂度较低的变维卷积神经网络模型。冉琼等人(2018)以超像元作为图像分类的最小单元,利用子空间投影算法对数据进行降维处理,与非子空间投影算法相比,在有限训练样本条件下具有更高的分类精度。

传统的高光谱分类方法大多依赖于手工特征,在特征提取的过程中会丢失数据的细节信息。基于深度学习的方法可以同时提取低层次和高层次的语义特征,从而显著提高高光谱图像分类的精度。但基于深度学习的高光谱图像分类方法在有限训练样本的情况下,分类性能会受到样本特征不足以及 Hughes 现象的影响。通常,解决此类问题的有效方法是减少高光谱数据的维度,去除原始数据中的冗余信息。为此,本文采用主成分分析方法对光谱通道数较多的高光谱数据进行降维处理,在降低数据量的同时保留丰富的光谱特征。使用结合倒置特征金字塔和 U-Net 的网络对降维后的数据进行光谱和空间特征信息的提取,同时引入注意力机制抑制背景区域的特征响应,实现对地物区域的精确分类。为验证本文方法的有效性,使用 4 个高光谱数据集进行实验。实验结果表明,本文提出的分类算法具有更高的高光谱图像地物分类精度。

1 本文方法

1.1 数据预处理

高光谱遥感图像波段多,数据维度高,并且不同波段之间存在较高的相关性,导致数据在光谱维度上存在冗余。本文利用主成分分析(principal component analysis, PCA)方法,通过线性变换将数据变换到一个新的坐标系中以减少数据集的维度,同时保持数据集对方差贡献最大的特征。PCA 变换步骤如下:

1) 需降维处理的样本观测数据矩阵为 X , X 中的每一行表示一个波段的数据,每一列表示不同波段中图像对应的像元数据;

2) 求出其协方差矩阵 $\text{cov}(X)$;

3) 根据特征方程式

$$(\lambda I - \text{cov}(X))\mu = 0 \quad (1)$$

求出 $\text{cov}(X)$ 的特征值 λ 。式中, I 为单位矩阵, μ 为特征向量;

4) 求出各个特征值对应的特征向量并组成特征矩阵 P ;

5) 根据变换矩阵 $A = U^T$, 得到 PCA 变换表达式 $PX = U^T X$, 其中 U 是由 $\text{cov}(X)$ 按由大到小排列的特征值对应的特征向量 μ 组成的。组成特征矩阵 P 的变量 m 即称为第 m 主成分。

经过 PCA 变换,高光谱数据由原始数据的高维降至低维,移除了冗余特征,有利于减少网络模型的训练参数,加快模型的收敛速度。

1.2 网络结构

本文算法整体框架如图 1 所示。整个网络结构由 3 部分组成: U-Net 的收缩路径、扩展路径以及倒置特征金字塔网络(inverted feature pyramid network, IFPN)。高光谱图像在收缩路径部分经过两次卷积和 ReLU 激活函数后得到通道数为 $n(n = c, 64, 128, 256, 512, 1024, k)$ 的特征图,每层的特征图通过倒置特征金字塔网络,分别与尺寸不同、通道数相同的特征图进行融合,所得到的特征图与扩展路径中对应的上一层特征图同时进入注意力机制模块以获取目标区域权重值较高的特征图,在扩展路径中与经过上采样的特征图融合后经过卷积操作和激活函数,最终获得地物种类的分类结果。

1.2.1 U-Net 模型

通过网络结构图可以看出,整个框架呈 U 型排列,该网络的左侧是收缩路径,相当于经典编码—解码器中的编码器部分。与原始 U-Net 模型不同的是,网络输入切割好的高光谱图像数据,先经过两次带有填充层的 $3 \times 3 \times 3$ 卷积操作,填充层的作用在于保证图像进行卷积操作的输入尺寸和输出尺寸保持不变,每次卷积运算后都带有一个 ReLU 激活函数。随后采用步长为 2 的 2×2 池化操作进行下采样,从而依次获取不同层的特征向量。每进行一次下采样操作,特征通道的数量不发生变化但特征图的尺寸减为原来的一半。池化的过程表示为

$$O = \frac{n + 2p - f}{s} + 1 \quad (2)$$

式中, O 是输出特征图的大小尺寸, n 是输入特征图的大小尺寸, p 是填充层参数, f 是卷积核大小, s 是步长数值。

该网络的右侧是扩展路径, 相当于解码器。扩展路径中的特征拼接是由两部分的特征映射进行同维度叠加。特征拼接的左半部分是由同层收缩路径的特征图与上一层扩展路径的特征图分别通过注意力机制得到的特征映射, 该部分的特征区

域具有较高的权重值; 特征拼接的右半部分是由上一层扩展路径的特征图进行反卷积所得, 反卷积公式为

$$O = s \times (n - 1) + f - 2p \quad (3)$$

特征图的通道数量通过拼接操作增加一倍, 再经过反卷积操作进行上采样, 所得特征图的尺寸大小逐层增加一倍, 通道数逐层减为原来的一半。多次重复该步骤可以确保输出图像与输入图像具有相同的分辨率。输出的最后一层采用 1×1 卷积进行分类操作, 待分类的地物类别数为 k 。

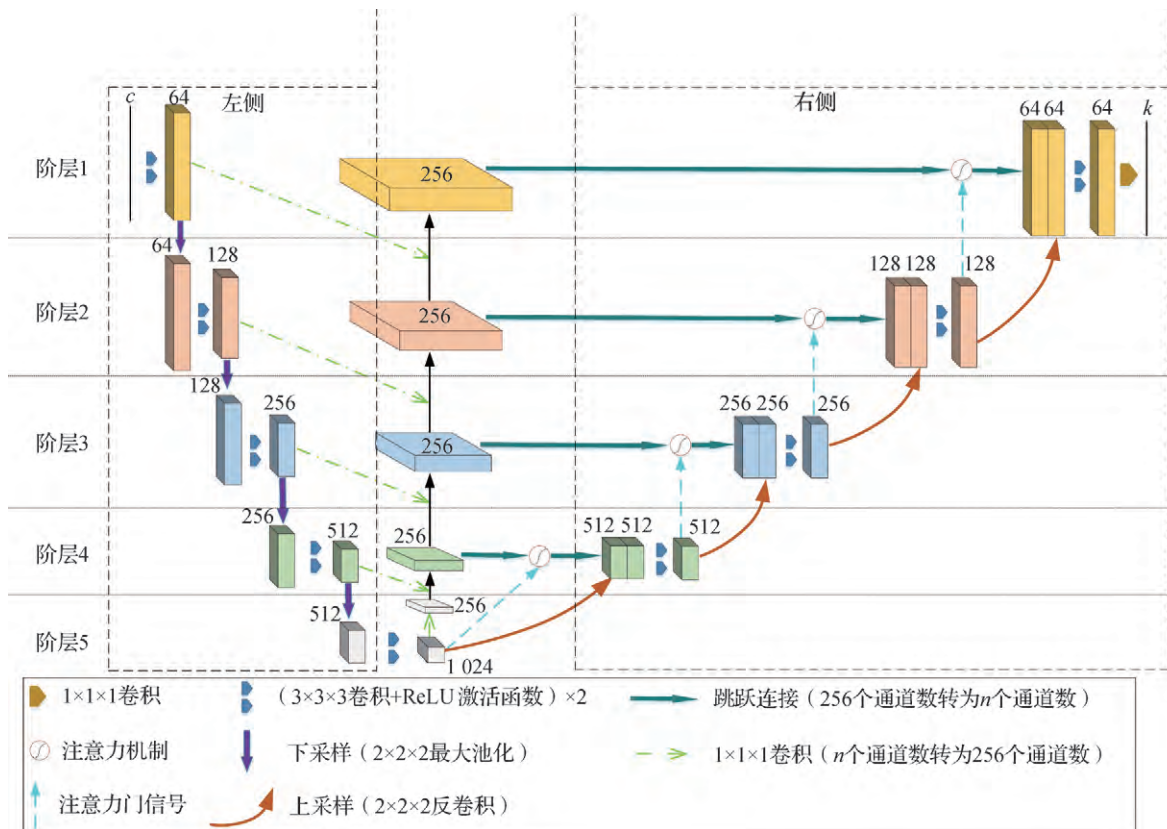


图1 结合倒置特征金字塔和 U-Net 的结构图

Fig.1 The architecture of U-Net combined with IFPN

1.2.2 倒置特征金字塔网络

本文所使用的主网络是 U-Net 的收缩路径网络结构, 整体网络可以分成 3 个部分: 自上而下的收缩路径网络、自底向上的特征金字塔网络和横向融合。倒置特征金字塔网络 (IFPN) 中联结这 3 个部分的横向融合结构图如图 2 所示。

自上而下的收缩路径是卷积网络的前馈计算, 它所构成的特征层级是由尺寸不同、通道数不同的特征映射组成。在收缩路径中, 每层得到的特征图进行 $1 \times 1 \times 1$ 卷积操作, 由于不同层的特征图的通

道数是不相同的, 故通过升维和降维处理使其与特征金字塔网络层中对应特征图的通道数保持一致。对于阶层 $N (N = 1, 2, 3, 4, 5)$ 的卷积输出, 本文将表示为 $x_i (i = 1, 2, 3, 4, 5)$ 。

自底而上的特征金字塔网络是由最底层的特征图先进行 $1 \times 1 \times 1$ 卷积和 2 倍上采样操作, 2 倍上采样之前的特征图可以表示为 $p_i (i = 1, 2, 3, 4, 5)$, 再与每一层通过 $1 \times 1 \times 1$ 的卷积输出 x_i 进行融合所组成, 输出表示为 $o_i (i = 1, 2, 3, 4)$ 。本文设定倒置特征金字塔网络特征图采用的固定通道数为

256. 自底而上的特征金字塔网络路径通过对空间上更抽象但语义更强的高层特征图进行上采样来生成高分辨率的特征,经过这样的结构, o_i 可以学来自 x_i 更深层的语义信息。最终特征映射集称为 $\{o_2, o_3, o_4, o_5\}$,分别对应于特征图 $\{x_2, x_3, x_4, x_5\}$ 。同层对应的两类特征图具有相同的尺寸。

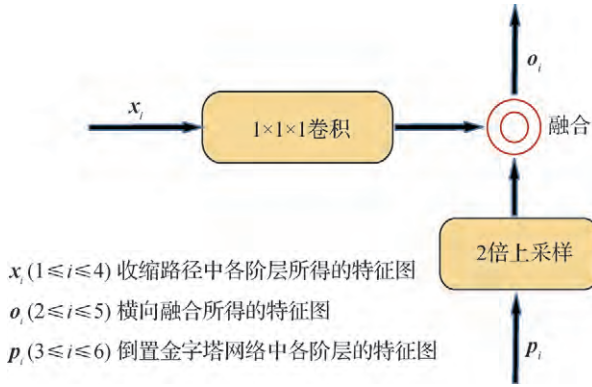


图2 IFPN 横向融合结构图

Fig.2 The structure of IFPN horizontal fusion

通过横向连接自底向上特征金字塔网络的路径,可以增强高层特征的表达。每个横向连接结构中,自上而下的收缩路径和自底向上的特征金字塔路径的特征图具有相同的尺寸。然后通过元素相

加,将上采样映射与相应的自底向上映射合并。通过迭代,生成最终的分辨率图。最后,本文在每个合并的图上分别附加一层跳跃连接进行 $3 \times 3 \times 3$ 卷积操作来消除上采样的混叠效应,将其生成的最终特征映射送入注意力机制模块。

1.2.3 注意力机制模块

在数据处理过程中,注意力机制的加入可以使模型关注图像中的有效信息并且忽略图像中的无关信息。其核心思想是对编码器的所有输出进行加权组合,然后再输入到当前位置的解码器中以影响解码器的输出。通过对编码器的输出进行加权,不仅可以实现输入与输出相对应,还能够利用更多原始数据的上下文信息。本网络所使用的注意力模块结构如图3所示。对于不同的阶层 N , x 表示同层倒置金字塔网络所得到的特征映射, g 是阶层 $N+1$ 中的特征门信号,目的是控制不同空间位置中特征的重要性。 x 和 g 分别进行 $1 \times 1 \times 1$ 卷积操作使通道数量相同且保持图像的大小不变,然后叠加作用于ReLU激活函数,最后通过 $1 \times 1 \times 1$ 的卷积和Sigmoid结合生成一个由0到1组成的权重值,即注意力系数 α , α 在目标区域取得较大值,在背景区域中取得较小值,有助于提高图像分类的精度。

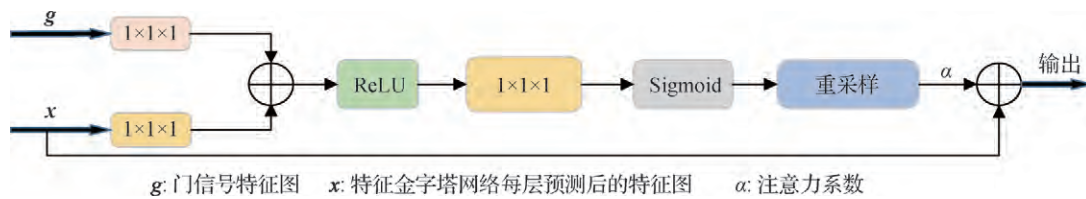


图3 注意力机制模块结构

Fig.3 The structure of attention module

2 实验数据集概述

为了有效地评估本文网络模型的性能,本文选取 Indian Pines、Pavia University、Salinas 和 Urban 这4个高光谱遥感数据集进行地物分类实验。这4种带有标注的数据集分别被随机分为训练集和测试集,以比例1:9来模拟小样本数据测试。

2.1 Indian Pines 数据集

Indian Pines 高光谱数据集制作于1992年,使

用 AVIRIS(airborne visible infrared imaging spectrometer) 光谱仪对美国某印度松树区域进行成像。如图4所示,图像尺寸为 145×145 像素,包含200个波段,成像波长为 $[0.4, 2.5] \mu\text{m}$ 。该图像包含16类地物,共计10249个样本像素。

2.2 Pavia University 数据集

Pavia University 数据集是由 ROSIS-03 (reflective optics system imaging spectrometer-03) 光谱仪于2013年对意大利帕维亚城所成像的高光谱数据。如图5所示,图像包含103个波段,分辨率为 1.3 m ,图像尺寸为 610×340 像素,波长范围为 $[0.43,$

0.86] μm 。该数据集除去背景像素外,共包含9类地物的42 776个像素。

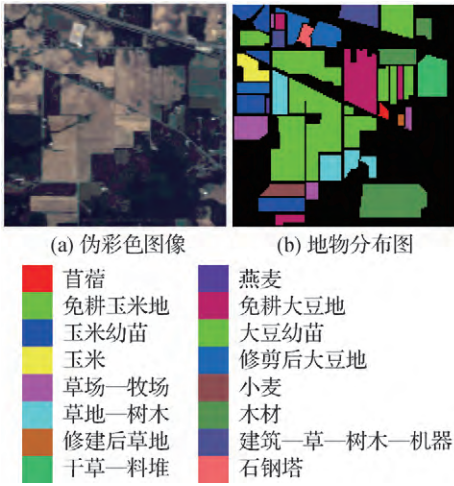


图4 Indian Pines 图像

Fig.4 Indian Pines image

((a) pseudo color image; (b) ground truth map)

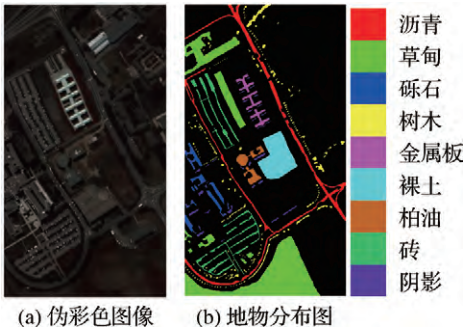


图5 Pavia University 图像

Fig.5 Pavia University image

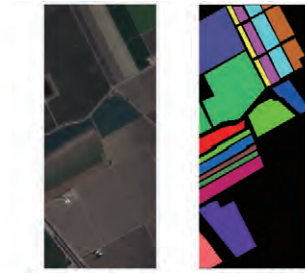
((a) pseudo color image; (b) ground truth map)

2.3 Salinas 数据集

Salinas 数据是由 AVIRIS 成像光谱仪对美国加利福尼亚州 Salinas 山谷所成图像,其空间分辨率为 3.7 m。如图 6 所示,图像包含 204 个波段,尺寸大小为 512 \times 217 像素,包含 56 975 个背景像素,可用于地物分类的像素共 16 类,数量为 54 129 个。

2.4 Urban 数据集

Urban 数据集是由 HYDICE(hyperspectral digital imagery collection experiment) 传感器对美国德克萨斯州某城郊住宿区所成的像,其空间分辨率为 4 m。如图 7 所示,该数据集图像尺寸大小为 307 \times 307 像素,包含 162 个波段,包含 6 类不同的地物,共 94 249 个可用于分类的像素。



(a) 伪彩色图像 (b) 地物分布图

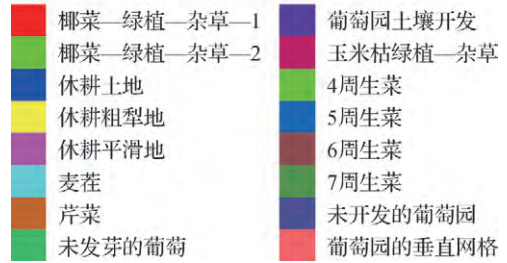


图6 Salinas 图像

Fig.6 Salinas image

((a) pseudo color image; (b) ground truth map)

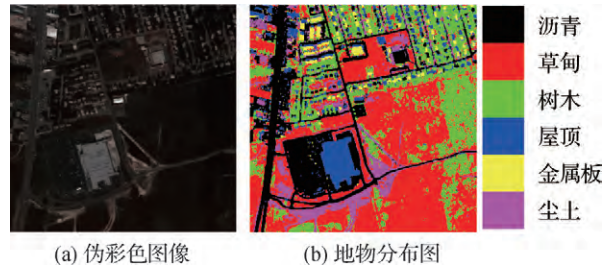


图7 Urban 图像

Fig.7 Urban image

((a) pseudo color image; (b) ground truth map)

3 实验和分析

实验采用深度学习框架 Pytorch(1.4.0) 进行模型搭建,CPU 为 Intel Core i7-9700K,GPU 选用 NVIDIA GTX 2080 进行加速,内存为 16 GB。实验分为 4 个部分进行,第 1 和第 2 部分研究 PCA 降维方法以及网络输入尺寸大小对实验结果的影响,确定最优参数。第 3 部分在第 1 和第 2 部分基础上进行,并与其他方法比较,检验本文模型针对小样本高光谱图像的分类效果;对于难以观察分类效果的 Urban 数据集,列出了不同实验条件下的地物分类结果。第 4 部分分析了不同实验条件下的模型训练时间。模型评估指标采用总体分类精度(overall accuracy ,OA)、平均分类精度(average accuracy ,

AA) 以及 Kappa 系数。

3.1 PCA 保留主成分数量对实验结果的影响

经过 PCA 变换,高光谱数据由原始数据的数百维降至 n 维。本文设定变换后的 n 维数据至少需含有原图 90% 的特征信息,经分析及验证,并考虑每个数据集分类的难度,对于较难分类的 Indian Pines 数据集和 Urban 数据集,将 n 设为 3、10、20 和 30;对于 Pavia University 数据集,将 n 设为 3、10 和 20;对于易于分类的 Salinas 数据集,将 n 设为 3 和 10。

由表 1 所示,当 PCA 保留主成分数量为 3 时,本文模型在 Indian Pines 数据集取得了最高的 OA 值和 Kappa 系数值,与此同时,该模型的总体分类精度也达到最高。图 8 展示 PCA 保留主成分数量对 Indian Pines 数据集分类结果的影响。不难看出,对于图像左上方较难分类的草地—牧场及周围区域,使用 PCA 方法将高光谱数据降成 3 维时的分类效果最好。

表 1 Indian Pines 数据集上 PCA 保留主成分数量对实验分类结果的影响

Table 1 Influence of PCA retained principal component number on classification results on Indian Pines dataset

评估指标	参数设置			
	PCA = 3	PCA = 10	PCA = 20	PCA = 30
OA	98.52	97.54	95.63	98.01
AA	96.40	96.58	88.39	97.67
Kappa	98.31	97.18	95.01	97.72

注:加粗字体表示每行最优值。

对于 Pavia University 数据集,由表 2 所示,当 PCA 保留主成分数量为 20 时,OA 值和 Kappa 值略高一些,但此时的 AA 值偏低。图 9 表明 PCA 保留主成分数量对 Pavia University 数据集分类结果的影响。从图 9 中可看出,此时模型分类效果与真值图像几乎相同。

由表 3 所示,对于 Salinas 数据集,PCA 保留主成分数量为 10 时,本文模型分类性能最好,接近 100%。图 10 显示出了 PCA 保留主成分数量对 Salinas 数据集分类结果的影响。从图 10 中观察可看出,当主成分数量为 3 或者 10 时,本文模型均能取得很好的分类效果。

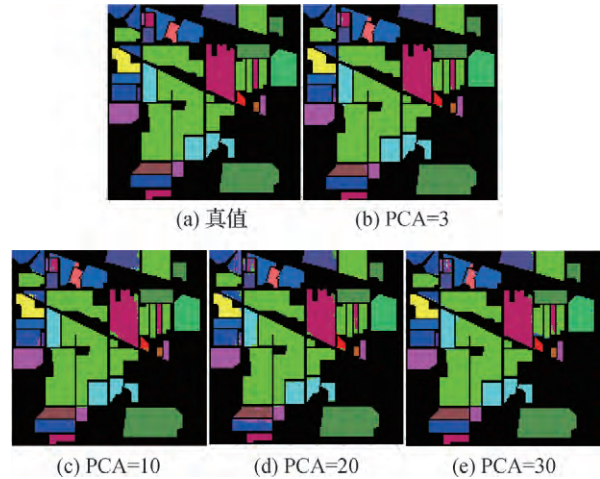


图 8 PCA 保留主成分数量对 Indian Pines 数据集分类结果的影响

Fig. 8 Influence of PCA retained principal component number on classification results of Indian Pines dataset((a) ground truth; (b) PCA = 3; (c) PCA = 10; (d) PCA = 20; (e) PCA = 30)

表 2 Pavia University 数据集上 PCA 保留主成分数量对实验分类结果的影响

Table 2 Influence of PCA retained principal component number on classification results on Pavia University dataset

评估指标	参数设置		
	PCA = 3	PCA = 10	PCA = 20
OA	99.81	99.82	99.85
AA	99.46	99.46	99.39
Kappa	99.74	99.75	99.81

注:加粗字体表示每行最优值。

由表 4 可看出,对于 Urban 数据集,PCA 保留主成分为 3 时,模型分类性能的 3 个评价指标均为最优。图 11 的结果显示 PCA 保留主成分为 3 时,模型对高光谱图像中的大多数地物类别都能保持相对较高的正确识别率,此时的分类结果图中具有更多的平滑区域。

3.2 网络输入样本大小的影响

实验采用的数据集是固定区域的高光谱图像,为充分利用图像的信息,本文在数据经过 PCA 降维之后采用滑动窗口的方法对数据进行分割和收集。在每个像素周围截取 $m \times m \times n$ 的区域作为网络输入样本,像素所在位置的类别为真值标签,其中 m 为截取的像素块的尺寸, n 为数据经过降维之后的维度。为探究网络输入数据尺寸对实验效果的影

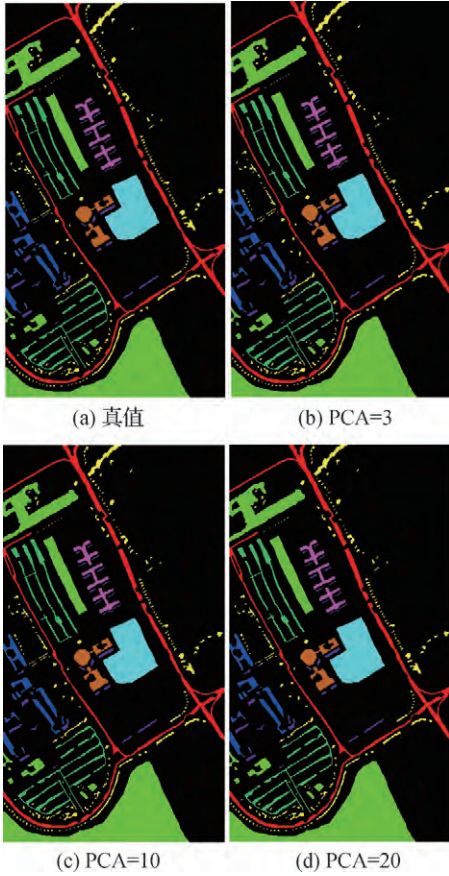


图9 PCA保留主成分数量对Pavia University数据集分类结果的影响

Fig. 9 Influence of PCA retained principal component number on classification results of Pavia University dataset ((a) ground truth; (b) PCA = 3; (c) PCA = 10; (d) PCA = 20)

表3 Salinas数据集上PCA保留主成分数量对实验分类结果的影响

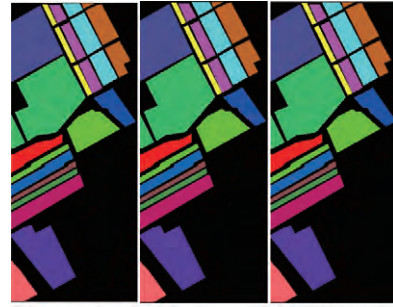
Table 3 Influence of PCA retained principal component number on classification results on Salinas dataset

评估指标	参数设置	
	PCA = 3	PCA = 10
OA	99.96	99.99
AA	99.96	99.99
Kappa	99.96	99.99

注:加粗字体表示每行最优值。

响 本文选取 16 32 64 作为 m 的参数值。

由表5可以看出,当网络输入尺寸为64时,模型的3个评估指标在Indian Pines数据集上均取得了最高的精度,说明此时模型分类性能最优。图12列出了模型输入尺寸对Indian Pines数据集分



(a)真值 (b) PCA=3 (c) PCA=10

图10 PCA保留主成分数量对Salinas数据集分类结果的影响

Fig. 10 Influence of PCA retained principal component number on classification results of Salinas dataset ((a) ground truth; (b) PCA = 3; (c) PCA = 10)

表4 Urban数据集上PCA保留主成分数量对实验分类结果的影响

Table 4 Influence of PCA retained principal component number on classification results on Urban dataset

评估指标	参数设置			
	PCA = 3	PCA = 10	PCA = 20	PCA = 30
OA	86.67	85.80	84.92	85.05
AA	81.23	78.43	79.72	79.65
Kappa	82.22	80.84	79.89	79.95

注:加粗字体表示每行最优值。

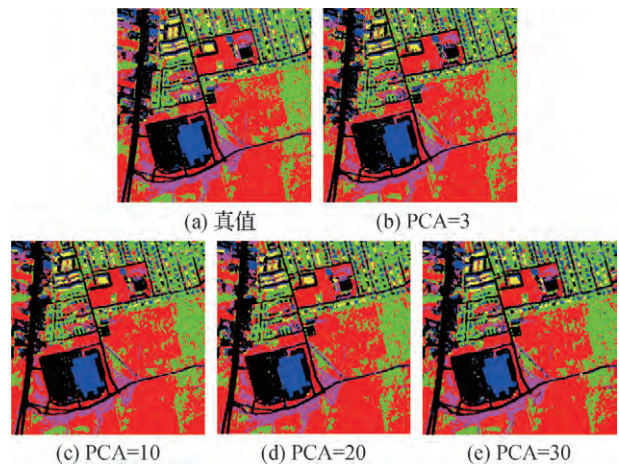


图11 PCA保留主成分数量对Urban数据集分类结果的影响

Fig. 11 Influence of PCA retained principal component number on classification results of Urban dataset((a) ground truth; (b) PCA = 3; (c) PCA = 10; (d) PCA = 20; (e) PCA = 30)

类结果的影响。从预测分类效果图可得到,当网络

表 5 Indian Pines 数据集上网络输入样本大小对实验分类结果的影响

Table 5 Influence of input size on classification results of Indian Pines dataset

评估指标	参数设置		
	size = 16	size = 32	size = 64
OA	97.61	98.52	98.91
AA	93.85	96.41	98.07
Kappa	97.26	98.31	98.75

注: 加粗字体表示每行最优值。

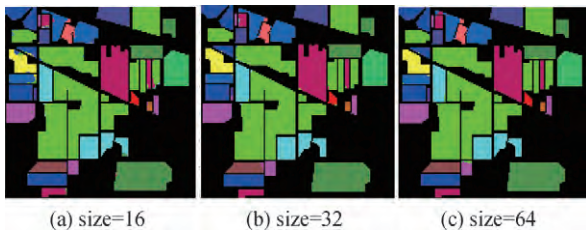


图 12 网络输入样本大小对 Indian Pines 数据集分类结果的影响

Fig. 12 Influence of input size on classification results of Indian Pines dataset((a) size = 16; (b) size = 32; (c) size = 64)

输入尺寸设置为 64 时, 预测错误的像素数量最少。

由表 6 可得, 当网络输入尺寸为 32 时, 本文模型在 Pavia University 数据集上的总体分类精度和 Kappa 系数值达到了最高值。图 13 为模型输入尺寸对 Pavia University 数据集分类结果的影响。从图 13 中可观察到, 虽然 3 种参数值对应的分类效果都接近真值, 但网络输入尺寸为 32 时预测错误率最低。

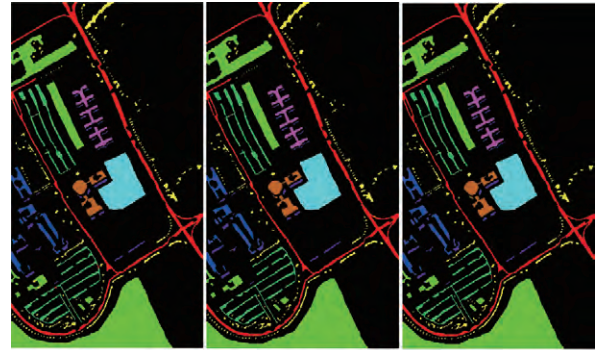
表 6 Pavia University 数据集上网络输入样本大小对实验分类结果的影响

Table 6 Influence of input size on classification results of Pavia University dataset

评估指标	参数设置		
	size = 16	size = 32	size = 64
OA	99.79	99.82	99.68
AA	99.57	99.41	99.21
Kappa	99.73	99.76	99.58

注: 加粗字体表示每行最优值。

从表 7 可以看出, 当网络输入样本大小为 32 和



(a) size=16 (b) size=32 (c) size=64

图 13 网络输入样本大小对 Pavia University 数据集分类结果的影响

Fig. 13 Influence of input size on classification results of Pavia University dataset((a) size = 16; (b) size = 32; (c) size = 64)

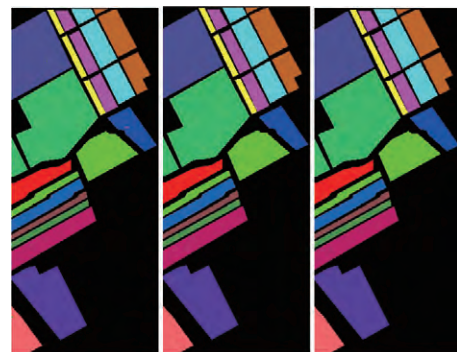
64 时, 模型在 Salinas 数据集上取得了同样的高准确率分类结果。并且从图 14 的显示效果对比来看, 参数设定为 32 和 64 均能产生极接近真值图像的分类效果。

表 7 Salinas 数据集上网络输入样本大小对实验分类结果的影响

Table 7 Influence of input size on classification results of Salinas dataset

评估指标	参数设置		
	size = 16	size = 32	size = 64
OA	99.96	99.98	99.98
AA	99.94	99.97	99.97
Kappa	99.95	99.98	99.98

注: 加粗字体表示每行最优值。



(a) size=16 (b) size=32 (c) size=64

图 14 网络输入样本大小对 Salinas 数据集分类结果的影响

Fig. 14 Influence of input size on classification results of Salinas dataset((a) size = 16; (b) size = 32; (c) size = 64)

由表 8 可看出,当网络输入样本大小为 64 时,模型在 Urban 数据集上取得了较好的总体地物分类结果。从图 15 可看出,虽然样本输入大小为 64 时的预测结果相比于真值标签仍有较大差距,但此时模型的错分区域没有样本输入大小为 16 和 32 时错分的区域大。

表 8 Urban 数据集上网络输入样本大小对实验分类结果的影响

Table 8 Influence of input size on classification results of Urban dataset

评估指标	参数设置		
	size = 16	size = 32	size = 64
OA	86.67	85.64	87.43
AA	81.23	81.75	78.30
Kappa	82.22	80.94	83.14

注:加粗字体表示每行最优值。

表 9 本文模型的分类结果

Table 9 Classification results of proposed method

评估指标	数据集			
	Indian Pines	Pavia University	Salinas	Urban
OA	98.79 ± 0.084	99.79 ± 0.040	98.99 ± 0.004	87.01 ± 0.297
AA	97.82 ± 0.189	99.39 ± 0.021	98.98 ± 0.008	78.47 ± 0.294
Kappa	98.61 ± 0.094	99.73 ± 0.052	98.98 ± 0.008	82.65 ± 0.346

3.1 节和 3.2 节的结果表明对于 Indian Pines、Pavia University、Salinas 和 Urban 数据集,当 PCA 维度分别选取 3、20、10、3 和输入尺寸分别选取 64、32、32、64 时的分类结果最优。由此可得,对于不同的数据集,选择不同的 PCA 保留主成分数量和网络输入样本大小能够一定程度上影响分类精度。现选择基于小样本的高光谱图像分类方法与本文实验结果进行比较。表 10—表 12 分别列出了其他方法在 Indian Pines、Pavia University、Salinas 数据集上的分类结果。表 13 针对 Urban 数据集分类结果图相对于其他 3 个数据集结果图不容易观察的问题,列出了不同 PCA 保留主成分数量和不同网络输入样本大小条件下每类地物的分类结果。本文实验数据集的划分方式与其他方法保持一致,采用每类地物的分类精度、图像的整体精度、图像的平均精度和

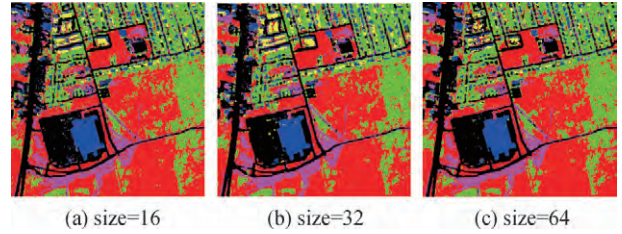


图 15 网络输入样本大小对 Urban 数据集分类结果的影响
Fig. 15 Influence of input size on classification results of Urban dataset ((a) size = 16; (b) size = 32; (c) size = 64)

3.3 对比结果分析

本文模型在 4 个数据集上的平均分类精度和标准差见表 9,每个实验在不同的数据集重复 5 次,对结果进行平均以消除随机因素对结果的影响。实验结果显示 Urban 数据集的分类结果不如另外 3 个数据集,最主要的原因是 Urban 高光谱图像的拍摄区域位于城市郊区,地物类别少且在空间分布上较为分散,图像空间信息的可利用性不高。

Kappa 系数依次进行对比。

从表 10 可得,相比于其他方法,本文方法的 OA、AA 和 Kappa 系数最高,在 Indian Pines 数据集上取得了最好的分类效果,证明了本文方法的有效性。免耕玉米地、大豆幼苗和石钢塔类别的分类结果分类精度偏低,但在样本总数较少的修剪后草地和燕麦类别分类结果上,本文方法结果远优于其他方法,说明本文网络对于小样本类别的地物特征能够进行充分提取,从而获得更好的分类性能。

表 11 显示了 Pavia University 数据集上本文方法与其他方法的比较结果。不同于表 10,表 11 中的分类结果数据更优。其主要原因在于 Pavia University 数据集地物类别种类较少,且每一类别样本数目较多。本文的实验结果仍优于其他方法,达到了 99.85% 的总体精度,证明了本文方法在地物种

表 10 在 Indian Pines 数据集上不同方法的分类精度
Table 10 The categorized accuracy of different methods on Indian Pines dataset

地物种类	样本总数	方法						
		SVM	OMP	NRS	KCRT	DKCRT	DV-CNN	本文
苜蓿	46	42.32	47.20	52.68	72.32	71.34	64.29	100.00
免耕玉米地	1 428	78.77	65.36	48.42	80.63	83.34	83.28	98.66
玉米幼苗	830	63.72	47.26	25.40	68.95	73.35	77.64	99.76
玉米	237	72.37	40.56	35.54	63.33	75.21	68.22	98.75
草场—牧场	483	89.55	85.10	88.18	92.89	93.43	93.56	99.79
草地—树木	730	96.64	91.41	95.25	98.07	98.71	96.80	99.86
修建后草地	28	67.00	26.20	54.00	64.80	63.80	80.77	100.00
干草—料堆	478	98.92	98.70	99.67	99.73	99.78	99.77	100.00
燕麦	20	35.28	29.44	33.61	50.83	50.83	33.33	100.00
免耕大豆地	972	70.53	48.73	28.10	76.11	84.87	85.94	99.08
大豆幼苗	2 455	86.17	70.52	98.14	83.24	82.95	90.90	98.12
修剪后大豆地	593	80.48	51.74	35.28	82.07	89.50	75.47	98.79
小麦	205	97.66	94.35	95.27	99.24	99.29	92.97	100.00
木材	1 265	96.59	93.33	97.59	95.62	96.54	94.91	100.00
建筑—草—树木—机器	386	52.44	48.40	59.03	63.63	61.04	79.60	99.74
石钢塔	93	87.80	88.04	81.67	86.67	87.38	86.55	88.78
OA/%		82.60	69.87	70.81	83.78	86.06	87.60	98.91
AA/%		76.01	64.15	64.25	79.93	81.91	81.15	98.07
Kappa		0.800	0.654	0.653	0.815	0.841	0.858	0.987

注: 加粗字体为每行最优值; 对比方法的实验数据来自于刘万军等人(2019)研究。对比方法中: 正交匹配追踪(orthogonal matching pursuit,OMP)、最近邻正则化子空间分类器(nearest regularized subspace,NRS)、KCRT(kernel collaborative representation with Tikhonov regularization)、DKCRT(discriminative kernel CRT)、变维卷积神经网络(dimensionality-varied convolutional neural network,DV-CNN)。

表 11 在 Pavia University 数据集上不同方法的分类精度
Table 11 The categorized accuracy of different methods on Pavia University dataset

地物种类	样本总数	方法						
		SVM	OMP	NRS	KCRT	DKCRT	DV-CNN	本文
沥青	6 631	91.91	70.29	93.01	83.29	86.49	98.73	99.97
草甸	18 649	98.51	92.60	98.75	95.92	97.98	99.67	100.00
砾石	2 099	33.26	57.49	58.59	67.19	67.86	91.80	99.76
树木	3 064	85.80	82.54	78.53	84.09	83.78	98.19	99.03
金属板	1 345	98.89	99.77	98.85	99.47	99.22	100.00	100.00
裸土	5 029	49.98	58.37	53.25	75.88	74.02	97.92	100.00
柏油	1 330	14.87	54.33	41.92	74.53	75.86	94.65	99.40
砖	3 682	89.99	63.09	84.78	79.28	86.50	95.96	99.62
阴影	947	99.83	83.62	91.08	99.95	99.74	95.78	98.81
OA/%		84.26	78.97	85.96	87.45	89.29	98.28	99.85
AA/%		73.59	73.57	77.64	84.40	85.72	96.97	99.39
Kappa		0.785	0.718	0.808	0.833	0.856	0.977	0.998

注: 加粗字体为每行最优值; 其他方法的实验数据来自于刘万军等人(2019)研究。

类较少的高光谱分类数据集上具有良好的分类性能。

表 12 为 Salinas 数据集上本文方法与其他方法的分类结果比较。此数据集虽然地物种类较多,但同类物体排列紧凑,不同类物体较分散,易于区分。

本文方法在此数据集上取得了 99.99% 的总体分类精度,优于所列其他方法。实验结果证明了对于此数据集的每一种类别,本文方法均能充分提取其特征,达到近乎 100% 的分类准确率。

表 12 在 Salinas 数据集上不同方法的分类精度
Table 12 The categorized accuracy of different methods on Salinas dataset

地物种类	样本总数	方法						
		SVM	SR	KSR	SOMP	ScSPM	K-SVD	本文
花椰菜—绿植—杂草—1	2 009	99.50	99.72	99.61	100.00	100.00	100.00	100.00
花椰菜—绿植—杂草—2	3 726	99.97	99.34	99.28	99.79	100.00	100.00	100.00
休耕地	1 976	99.66	97.58	97.47	97.02	100.00	100.00	100.00
休耕粗犁地	1 394	99.84	99.52	99.52	96.18	99.28	97.77	99.64
休耕平滑地	2 678	98.67	98.26	98.18	93.90	99.54	99.63	99.93
麦茬	3 959	99.86	99.75	99.69	100.00	100.00	100.00	100.00
芹菜	3 579	99.75	99.84	99.78	99.94	100.00	100.00	100.00
未发芽的葡萄	11 271	85.95	87.67	89.73	97.84	96.46	99.33	100.00
葡萄园土壤开发	6 203	99.84	99.73	99.70	99.96	100.00	100.00	100.00
玉米—枯绿植—杂草	3 278	98.07	96.81	96.75	98.68	99.80	99.86	100.00
4 周生菜	1 068	99.69	98.23	98.02	98.02	100.00	100.00	100.00
5 周生菜	1 927	99.94	100.00	99.88	93.31	100.00	100.00	100.00
6 周生菜	916	99.64	99.15	98.91	90.18	100.00	99.39	100.00
7 周生菜	1 070	98.23	96.37	96.16	98.23	99.69	99.38	100.00
未开发的葡萄园	7 268	76.98	67.85	67.82	88.03	90.61	99.24	100.00
葡萄园的垂直网格	1 807	99.51	99.45	99.32	99.39	100.00	100.00	100.00
OA/%		93.65	92.48	92.42	96.83	97.94	99.65	99.99
AA/%		97.19	96.21	96.10	96.90	99.09	99.66	99.99
Kappa		0.929	0.934	0.933	0.965	0.987	0.997	0.999

注:加粗字体为每行最优值;对比方法的实验数据来自于 Zhang 等人(2016)研究;对比方法中:稀疏表示(sparse representation, SR)、KSR(kernel sparse representation)、同步正交匹配追踪(simultaneous orthogonal matching pursuit, SOMP)、ScSPM(spatial pyramid based on matching sparse coding)、K-SVD(K-singular value decomposition)。

对于 Urban 数据集,本文针对分类结果图不容易观察的问题列出并分析各个实验条件下的分类结果。由表 13 可看出,网络输入尺寸大小为 64 时,虽然 6 类地物的分类结果中只有草甸和屋顶的分类精度高,但此时模型预测效果最好,其主要原因在于模型对样本总数高的类别预测效果好。

3.4 效率实验

高光谱遥感图像数据量庞大,算法效率是基于

深度学习的高光谱影像分类研究的重点关注之处。本文模型在 4 个数据集上的每个 epoch 实际训练所用时间见表 14。由表 14 可看出,数据预处理后,PCA 保留主成分数量对训练速度几乎无影响,原因在于网络的输入通道数在卷积中转换为固定的通道数,对模型参数量影响较小。而网络输入样本尺寸增大势必会引起模型参数量增加,使得训练时间增长。

表 13 Urban 数据集图像分类精度
Table 13 Classification accuracy on Urban dataset

地物种类	样本总数	参数设置						
		PCA = 3	PCA = 10	PCA = 20	PCA = 30	size = 16	size = 32	size = 64
沥青	18 570	87.13	87.69	83.89	87.15	87.13	87.60	87.54
草甸	35 198	91.15	86.39	90.05	86.86	91.15	90.56	91.16
树木	22 468	85.89	85.03	84.45	85.27	85.89	87.54	84.59
屋顶	6 821	81.33	81.97	71.98	82.97	81.33	78.38	85.60
金属板	2 436	65.71	84.58	77.63	79.46	65.71	53.65	75.60
尘土	8 756	79.26	83.59	80.26	74.58	79.26	75.36	82.95
OA/%		86.67	85.80	84.92	85.05	86.67	85.64	87.43
AA/%		81.23	78.43	79.72	79.65	81.23	81.75	78.30
Kappa		82.22	80.84	79.89	79.95	82.22	80.94	83.14

注: 加粗字体为每行最优值。

表 14 本文模型在 4 个数据集上的训练时间
Table 14 Training time of proposed method on four dataset

数据集	参数设置						
	PCA = 3	PCA = 10	PCA = 20	PCA = 30	size = 16	size = 32	size = 64
Indian Pines	3	3	3	3	1	3	8
Pavia University	10	10	10	-	12	13	31
Salinas	14	14	-	-	7	15	44
Urban	12	12	12	12	11	27	77

注 “-”表示未在该参数设置下对数据集进行实验。

4 结 论

高光谱图像在数百个连续且细分的光谱波段中能对目标区域所含的丰富地物信息准确地呈现。但各波段所含的信息存在冗余性,如何有效地从高光谱图像数据中提取关键的地物特征信息用于分类是高光谱影像研究的重点与难点。

针对上述问题提出了结合倒置特征金字塔和 U-Net 的卷积神经网络。首先,通过主成分分析方法对高光谱图像数据进行降维处理,采用滑动窗口方法对降维后的数据进行像元构建,并将像元作为模型的输入。本文模型使用 U-Net 作为骨干网络提取高光谱图像的地物特征。然后,这些富含空间信息的特征图以分层的形式与倒置的特征金字塔网络逐层进行信息融合,从而得到丰富的光谱和空间信息。注意力机制的引入可以提高模型对于有效光谱和空间信息的关注度,同时降低背景区域对模型分类性能的影响,有利于高光谱图像的分类。实验结

果证明了本文方法可以充分提取关键的地物特征进行准确地分类,是一种有效的高光谱图像分类方法。在未来的研究中,将着重考虑更少训练数据样本的情况下,如何使模型在降低复杂度的同时保持良好的高光谱图像分类性能。

参考文献(References)

- Haut J M, Paoletti M E, Plaza J, Plaza A and Li J. 2019. Visual attention-driven hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10): 8065-8080 [DOI: 10.1109/TGRS.2019.2918080]
- He N J, Paoletti M E, Haut J M, Fang L Y, Li S T, Plaza A and Plaza J. 2019. Feature extraction with multiscale covariance maps for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 755-769 [DOI: 10.1109/TGRS.2018.2860464]
- Jia S, Lin Z J, Deng B, Zhu J S and Li Q Q. 2020. Cascade superpixel regularized gabor feature fusion for hyperspectral image classification. *IEEE Transactions on Neural Networks and Learning Systems*, 31(5): 1638-1652 [DOI: 10.1109/TNNLS.2019.2921564]

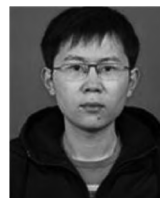
- Jia S, Shen L L, Zhu J S and Li Q Q. 2018. A 3-D gabor phase-based coding and matching framework for hyperspectral imagery classification. *IEEE Transactions on Cybernetics*, 48(4): 1176-1188 [DOI: 10.1109/TCYB.2017.2682846]
- Li Y, Zhen C, Shi X and Zhao Q H. 2019. Hyperspectral image classification algorithm based on entropy weighted K-means with global information. *Journal of Image and Graphics*, 24(4): 630-638 (李玉, 甄畅, 石雪, 赵泉华. 2019. 基于熵加权 K-means 全局信息聚类的高光谱图像分类. *中国图象图形学报*, 24(4): 630-638) [DOI: 10.11834/jig.180502]
- Liu J J, Wu Z B, Xiao L, Sun J and Yan H. 2020. Generalized tensor regression for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 58(2): 1244-1258 [DOI: 10.1109/TGRS.2019.2944989]
- Liu W J, Yin X, Qu H C and Liu L M. 2019. Dimensionality-varied convolutional neural network for improving the classification performance of hyperspectral images with small-sized labeled samples. *Journal of Image and Graphics*, 24(9): 1604-1618 (刘万军, 尹岫, 曲海成, 刘腊梅. 2019. 提高小样本高光谱图像分类性能的变维卷积神经网络. *中国图象图形学报*, 24(9): 1604-1618) [DOI: 10.11834/jig.180693]
- Liu Y, Ji X F and Wang Y Y. 2016. Classification of hyperspectral image based on double L2 sparse coding. *Journal of Image and Graphics*, 21(12): 1707-1715 (刘洋, 姬晓飞, 王杨扬. 2016. 基于双重 L2 稀疏编码的高光谱图像分类. *中国图象图形学报*, 21(12): 1707-1715) [DOI: 10.11834/jig.20161215]
- Luo F L, Du B, Zhang L P and Tao D C. 2019. Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image. *IEEE Transactions on Cybernetics*, 49(7): 2406-2419 [DOI: 10.1109/TCYB.2018.2810806]
- Paoletti M E, Haut J M, Fernandez-Beltran R, Plaza J, Plaza A, Li J and Pla F. 2019a. Capsule networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(4): 2145-2160 [DOI: 10.1109/TGRS.2018.2871782]
- Paoletti M E, Haut J M, Fernandez-Beltran R, Plaza J, Plaza A J and Pla F. 2019b. Deep pyramidal residual networks for spectral-spatial hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 740-754 [DOI: 10.1109/TGRS.2018.2860125]
- Peng J T, Li L Q and Tang Y Y. 2019. Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images. *IEEE Transactions on Neural Networks and Learning Systems*, 30(6): 1790-1802 [DOI: 10.1109/TNNLS.2018.2874432]
- Ran Q, Yu H Y, Gao L R, Li W and Zhang B. 2018. Superpixel and subspace projection-based support vector machines for hyperspectral image classification. *Journal of Image and Graphics*, 23(1): 95-105 (冉琼, 于浩洋, 高连如, 李伟, 张兵. 2018. 结合超像元和子空间投影支持向量机的高光谱图像分类. *中国图象图形学报*, 23(1): 95-105) [DOI: 10.11834/jig.170201]
- Wu H and Prasad S. 2018. Semi-supervised deep learning using pseudo labels for hyperspectral image classification. *IEEE Transactions on Image Processing*, 27(3): 1259-1270 [DOI: 10.1109/TIP.2017.2772836]
- Xiao L, Liu P F and Li H. 2020. Progress and challenges in the fusion of multisource spatial-spectral remote sensing images. *Journal of Image and Graphics*, 25(5): 851-863 (肖亮, 刘鹏飞, 李恒. 2020. 多源空-谱遥感图像融合方法进展与挑战. *中国图象图形学报*, 25(5): 851-863) [DOI: 10.11834/jig.190620]
- Yu C Y, Han R, Song M P, Liu C Y and Chang C I. 2021. Feedback attention-based dense CNN for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*: #3058549 [DOI: 10.1109/TGRS.2021.3058549]
- Zhang M M, Li W and Du Q. 2018. Diverse region-based CNN for hyperspectral image classification. *IEEE Transactions on Image Processing*, 27(6): 2623-2634 [DOI: 10.1109/TIP.2018.2809606]
- Zhang X R, Liang Y L, Zheng Y G, An J L and Jiao L C. 2016. Hierarchical discriminative feature learning for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 13(4): 594-598 [DOI: 10.1109/LGRS.2016.2528883]
- Zhu M H, Jiao L C, Liu F and Yang S Y and Wang J N. 2021. Residual spectral-spatial attention network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(1): 449-462 [DOI: 10.1109/TGRS.2020.2994057]

作者简介



程嵩阳, 1996年生, 男, 硕士研究生, 主要研究方向为模式识别与智能系统。

E-mail: M020119109@sues.edu.cn



熊玉洁, 通信作者, 男, 讲师, 硕士生导师, 主要研究方向为模式识别与智能系统。

E-mail: xiong@sues.edu.cn

姚瑶, 女, 硕士研究生, 主要研究方向为模式识别与智能系统。E-mail: y54yaoyao@163.com

李庆利, 男, 教授, 主要研究方向为多维成像技术及系统、多维影像的人工智能识别分析方法、高光谱遥感及交叉学科应用、模式识别与智能系统。E-mail: qlili@cs.ecnu.edu.cn