

文章编号: 1673-5196(2024)02-0087-09

# 基于改进 SE-Net 和深度可分离残差的高光谱图像分类

王 燕\*, 王振宇

(兰州理工大学 计算机与通信学院, 甘肃 兰州 730050)

**摘要:** 针对目前常见的用于高光谱图像分类的卷积神经网络参数数量多, 训练时间长, 对样本数量依赖性大的问题, 提出一种适用于有限训练样本条件下基于改进压缩激活网络和深度可分离残差的分类网络 MDSR&SE-Net. 首先使用主成分分析对原始高光谱图像进行通道降维, 然后通过三维卷积神经网络连接多特征残差结构, 同时嵌入改进的 SE 模块提取高光谱图像的空间和光谱细节特征, 最后将提取到的特征数据输入 Softmax 分类器激活分类. 为了使网络更加轻量, 通过在残差结构中使用深度可分离卷积和引入全局平均池化减少参数数量. 实验结果显示, 使用有限训练样本在三种常见高光谱数据集上总体分类精度均达到 99% 以上.

**关键词:** 高光谱图像; 深度可分离卷积; 残差网络; 压缩激活网络

**中图分类号:** TP751.1 **文献标志码:** A

## Hyperspectral image classification based on improved SE-Net and depth-separable residuals

WANG Yan, WANG Zhen-yu

(College of Computer and Communication, Lanzhou Univ. of Tech., Lanzhou 730050, China)

**Abstract:** In response to the challenges posed by convolutional neural network (CNNs) commonly used for hyperspectral image classification, namely, their high parameter count, extended training times, and sensitivity to sample quantity, a classification network MDSR&SE-Net based on improved squeeze and excitation network and depth-separable residuals was proposed for limited training samples. First, the principal component analysis is employed in this model to reduce the dimension of the original HSI. Then, the multi-feature residual structure is connected by 3D convolutional neural network, and the spatial & spectral details of hyperspectral images are extracted by embedding the improved squeeze and excitation block. Finally, the extracted feature information is input into Softmax classifier to activate classification. To further lightweight the network, the number of parameters is reduced by using the depth separable convolution in the residual structure and introducing global average pooling. Experimental results show that overall accuracy of the three common hyperspectral data sets with the limited training samples are above 99%.

**Key words:** hyperspectral image; depth separable convolution; residual network; SE Net

高光谱图像(hyperspectral image, HSI)是一种使用多平台成像光谱仪以几十至几百个连续的光谱通道对目标区域的同时成像. 因此, 高光谱图像在拥有清晰空间信息的同时, 还包含着极为丰富的光谱

信息, 真正做到了“图谱合一”, 这一特性使得高光谱图像成为图像研究的热门领域. 目前 HSI 广泛应用于农业评估<sup>[1]</sup>、食品检测<sup>[2]</sup>、生物医学<sup>[3]</sup>和目标检测<sup>[4]</sup>等领域.

虽然 HSI 拥有丰富的感知信息, 但是关于它的应用仍然存在一些挑战, 如图像数据量庞大, 相邻波段之间存在相关性, 造成数据冗余; 图像中的混合像素分辨率较低, 往往会出现“同谱异物”和“同物异谱”现象; HSI 分类研究使用的数据集依靠大量人工

收稿日期: 2022-03-04

基金项目: 国家自然科学基金(61863025), 甘肃省重点研发计划-工业类(18YF1GA060)

通讯作者: 王 燕(1971-), 女, 甘肃泾川人, 教授.

Email: wangyan@lut.edu.cn

标记,数量有限且代价高昂.目前大多数分类模型均依靠较大的训练集划分来获取良好的分类结果,对样本标记的依赖性大.

近年来,用于 HSI 的分类算法主要包括支持向量机(support vector machine, SVM)<sup>[5]</sup>、逻辑回归(logistic regression, LR)、最大似然分类(maximum likelihood classification, MLC)<sup>[6]</sup>、K 近邻算法(K-nearest neighbors, KNN)和 K 均值算法等.上述算法虽然实现起来较为简单,运算速率快,但是只能提取图像的浅层特征信息,分类精度较低<sup>[7]</sup>.深度学习算法能够从原始数据中自动提取多层特征信息,目前在计算机视觉领域得到了大量应用,如卷积神经网络(convolutional neural networks, CNN)<sup>[8]</sup>、循环神经网络(recurrent neural networks, RNN)<sup>[9]</sup>、深度置信网络(deep belief networks, DBN)<sup>[10]</sup>和堆栈自编码器(stacked auto encoder, SAE)<sup>[11]</sup>.其中 CNN 是目前 HSI 分类使用最多的深度学习算法.

CNN 主要由输入层、卷积层、池化层和全连接层(fully connected layer, FC)组成.根据卷积核尺寸和计算维度的不同,分为一维卷积神经网络(1D-CNN)、二维卷积神经网络(2D-CNN)和三维卷积神经网络(3D-CNN).对于高光谱图像分类任务,1D-CNN 只能关注到图像光谱维度的特征,2D-CNN 虽然能够提取空间维度的特征,但是对光谱维度特征的提取能力不足,而 3D-CNN 可以通过三维卷积核的三维滑动同时提取 HSI 的光谱维特征和空间维特征.实验证明,单纯使用 1D-CNN 或 2D-CNN 的分类模型,由于不能充分感知 HSI 丰富的光谱和空间信息,性能表现通常较为普通.因此, Hu 等<sup>[12]</sup>提出了一种使用 1D-CNN 和 2D-CNN 的双通道空洞卷积神经网络模型(DCD-CNN),分别提取 HSI 中的光谱特征和空间特征. Roy 等<sup>[13]</sup>提出了一种 3D-CNN 结合 2D-CNN 的深层混合网络来解决 HSI 的分类问题,该网络采用 3D-CNN 提取 HSI 的空-谱综合特征,2D-CNN 进一步提取更深层的抽象特征,取得了良好的分类结果.虽然深层网络具有更强的特征学习能力和特征表达能力,但是网络深度的增加会产生性能退化的问题,例如识别精度快速达到饱和,随即开始下降. Zhong 等<sup>[14]</sup>提出了一种空间-光谱残差网络(SSRN),该网络每隔一层卷积就添加一个跳跃连接,保证了梯度流的回传,避免了上述模型经常遇到的问题,但为了保证模型的分类精度,其在实验过程中划分了较多训练样本.

目前,深度学习中的注意力机制(attention mechanism, AM)广泛应用于计算机视觉研究领域.

注意力机制通过模拟人类视觉神经系统快速扫描全局视觉图像,捕获大脑重点关注的目标区域,对该区域投入更多的注意力资源,以获取更加详细的目标细节信息,同时也抑制了其他无用的信息. Hu 等<sup>[15]</sup>提出了一种压缩激活网络(squeeze and excitation network, SE-Net),通过显式建模通道之间的相互依赖性,自适应地校准了通道维度的特征信息.但是 SE-Net 仅通过全局平均池化(global average pooling, GAP)策略获取输入特征的通道描述向量,感受野相对单一,并且缺乏对所得描述向量进行重新校准的机制.因此,保证准确获取 HSI 通道的关键权重信息变得尤为重要.

针对上述问题,本文提出了一种适用于有限训练样本条件下基于改进 SE-Net 和深度可分离残差的分类网络 MDSR&SE-Net. 本文的主要工作为:提出了两种特征融合残差结构块,通过在结构块中使用深度可分离卷积(depth-wise separable convolution, DS-Conv)代替二维卷积(two dimension convolution, 2D-Conv)使模型更加轻量化,同时在结构块尾部嵌入改进的 SE 模块以标示不同特征通道的关注度.使用 3D-CNN 串联特征融合残差结构作为主干网络架构,通过引入 GAP 连接 FC 层,在不丢失特征信息的前提下进一步降低参数数量,减少训练开销并且加快模型收敛.选取有限标签数据作为训练样本,使用 MDSR&SE-Net 最终获得了较高的分类精度.

## 1 相关工作

### 1.1 2D-CNN & 3D-CNN

2D-CNN 能够通过卷积核的感受野有效提取 HSI 目标像素邻域范围内的空间信息.二维卷积核的大小是针对图像的尺寸设计的,它的通道数必须与图像的通道数保持一致,在图像的整个通道上作卷积运算.2D-CNN 的数学描述为

$$v_{l,i}^{x,y} = \sigma \left( \sum_m \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} k_{l,i,m}^{h,w} v_{l-1,m}^{x+h,y+w} + b_i \right) \quad (1)$$

式中:  $\sigma(\cdot)$  表示激活函数;  $m$  表示第  $i-1$  层特征图的数量;  $H_l$  和  $W_l$  分别表示 2D 卷积核的尺寸;  $k_{l,i,m}^{h,w}$  表示第  $l$  层第  $m$  张特征图位置为  $(h, w)$  的第  $i$  个卷积核;  $v_{l-1,m}^{x+h,y+w}$  表示第  $l-1$  层第  $m$  张特征图位置为  $(x+h, y+w)$  的神经元数值;  $b_i$  表示偏置.

由于 2D-CNN 的运算性质可能使原图像关键信息产生丢失,致使最终分类精确度得不到提升. 3D-CNN 可以利用三维卷积核在 HSI 的空间维度和光谱维度滑动,同时提取图像的空间信息和光谱

信息. 基于该优势, 3D-CNN 已大量运用于各种深度学习模型中. 3D-CNN 的数学描述为

$$v_{l,i}^{x,y,z} = \sigma \left( \sum_m \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} \sum_{r=0}^{R_l-1} k_{l,i,m}^{h,w,r} v_{l-1,m}^{x+h,y+w,z+r} + b_{l,i} \right) \quad (2)$$

式中:  $\sigma(\cdot)$  表示激活函数;  $m$  表示第  $i-1$  层特征图的数量;  $H_l$ 、 $W_l$  和  $R_l$  表示三维卷积核的尺寸;  $k_{l,i,m}^{h,w,r}$  表示第  $l$  层第  $m$  张特征图位置为  $(h, w, r)$  的第  $i$  个卷积核;  $v_{l-1,m}^{x+h,y+w,z+r}$  表示第  $l-1$  层第  $m$  张特征图位置为  $(x+h, y+w, z+r)$  的神经元数值;  $b_{l,i}$  表示偏置.

### 1.2 深度可分离卷积

DS-Conv 与 2D-Conv 有所区别, 它将 2D-Conv 分解成 depth-wise 卷积和 point-wise 卷积两个步骤. 在 depth-wise 卷积中, 卷积核的每一个通道与输入特征图的对应通道作卷积运算, 通道之间的特征信息不进行融合, 而 point-wise 卷积利用  $1 \times 1$  卷积实现了特征图不同通道之间的线性结合.

从本质上讲, DS-Conv 就是 2D-Conv 在通道方向上的分解. 由于两种卷积最后可以得到相同维度的特征信息, 因此网络中的两种卷积可被视为是等效的. 相比于 2D-Conv, DS-Conv 的优势是大幅降低了模型的计算量. 两种卷积类型的计算量之比为

$$\frac{D_K D_K H D_F D_F + H N D_F D_F}{D_K D_K H N D_F D_F} = \frac{1}{N} + \frac{1}{D_K^2} \quad (3)$$

式中:  $D_K$  表示卷积核的尺寸;  $H$  表示输入特征图的通道数;  $D_F$  表示输入特征图的空间尺寸;  $N$  表示输出特征图的通道数.

### 1.3 残差结构

He 等<sup>[16]</sup>提出的 ResNet 通过构造残差结构解决了因网络层数加深而导致的性能退化问题. 其主要工作是在卷积层中添加跳跃连接, 使梯度流回传更加流畅, 隐层权重可以快速迭代更新, 保证了每一个残差结构块都能够学习到前层输出的特征信息.

瓶颈残差(bottleneck residual)结构是 ResNet 中的一种特殊残差结构. 首部使用  $1 \times 1$  卷积将前层输出的高维数据进行压缩, 降低结构内的参数数量, 在尾部再次使用  $1 \times 1$  卷积进行升维操作以保证输入输出数据尺寸一致, 最后借助跳跃连接与前层输出数据进行残差计算. 瓶颈残差结构输入与输出的数学描述为

$$H(x) = f(x) + x \quad (4)$$

式中:  $x$  表示前层输入;  $f(x)$  表示残差函数;  $H(x)$  表示残差输出.

其中残差函数  $f(x)$  可表示为

$$f(x) = \sigma \{ \sigma [ \sigma (xw_1 + b_1) w_2 + b_2 ] w_3 + b_3 \} \quad (5)$$

式中:  $\sigma(\cdot)$  表示激活函数;  $w_1$ 、 $w_2$ 、 $w_3$  表示卷积核;  $b_1$ 、 $b_2$ 、 $b_3$  表示偏置.

## 2 本文方法

### 2.1 MDSR&SE-Net 模型

本文提出的 MDSR&SE-Net 主要由 3D-CNN 特征提取单元和 DS-residual&SE 特征提取单元组成, 如图 1 所示. 该网络模型首先使用主成分分析(principal component analysis, PCA)<sup>[17]</sup> 去除原始高光谱图像的冗余信息; 然后将降维后的图像分割成  $M^{P \times P \times C}$  的图像块输入网络进行训练, 其中  $P$  表示邻域范围,  $C$  表示通道数量.

3D-CNN 特征提取单元由两层三维卷积堆叠组成, 用于提取高光谱图像的空间和光谱特征信息. 在该单元中, 首先对输入数据进行批量归一化(batch normalization, BN) 处理, 然后通过尺寸分别为  $K^{7 \times 7 \times 5}$  和  $K^{5 \times 5 \times 5}$  的卷积核学习图像的多维特征. 其中两层卷积均使用线性整流函数 ReLU 对输出特征数据进行激活. 最后利用窗口大小为 3、步长为 1 的最大池化层对输出数据进行下采样, 作为下一层网络的输入数据.

将 3D-CNN 特征提取单元得到的特征数据输入 DS-residual&SE 特征提取单元, 进一步提取更深层维度特征. 该单元由一个窗口大小为 2、步长为 1 的最大池化层连接两组特征融合残差块组成, 为了进一步提高分类精度, 本网络在每组特征融合残差块尾部嵌入了改进的 SE 模块, 对输出数据的通道赋予权值以强化有效特征信息. 最后通过 GAP 将输出的特征数据转换成一维向量, 连接包含 128 个神经元组成的 FC 层, 经过 Softmax 分类器得到最终像素的分类结果.

在 MDSR&SE-Net 中, 添加 L2 正则化和随机失活率 dropout 避免网络模型出现过拟合; 使用交叉熵作为损失函数, 同时使用遵循动量与自适应梯度思想的 Adam 优化器加速模型收敛. 为了能够提取到图像的边缘信息, 同时保证前向传播的一致性, 网络模型中的所有卷积层都设置有零填充.

### 2.2 特征融合残差结构

本文借鉴瓶颈残差结构的思想设计出两种特征融合残差模块. 同时, 为了使网络模型更加轻量化, 结构内部选择采用 DS-Conv, 如图 2 所示.

深度可分离残差块-a 首先使用  $1 \times 1$  卷积对前层特征数据进行压缩; 然后分别使用核尺寸为  $K^{7 \times 7}$

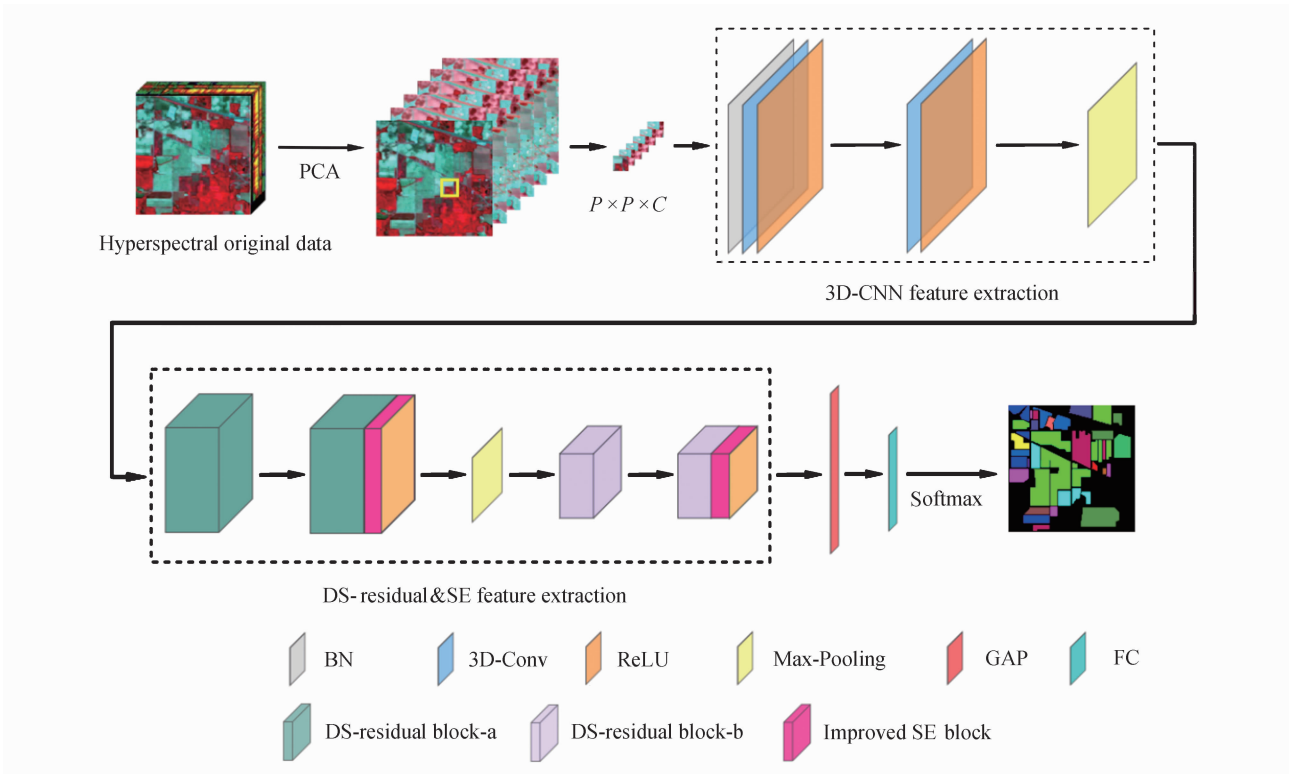


图 1 MDSR&SE-Net 结构  
Fig. 1 MDSR&SE-Net structure

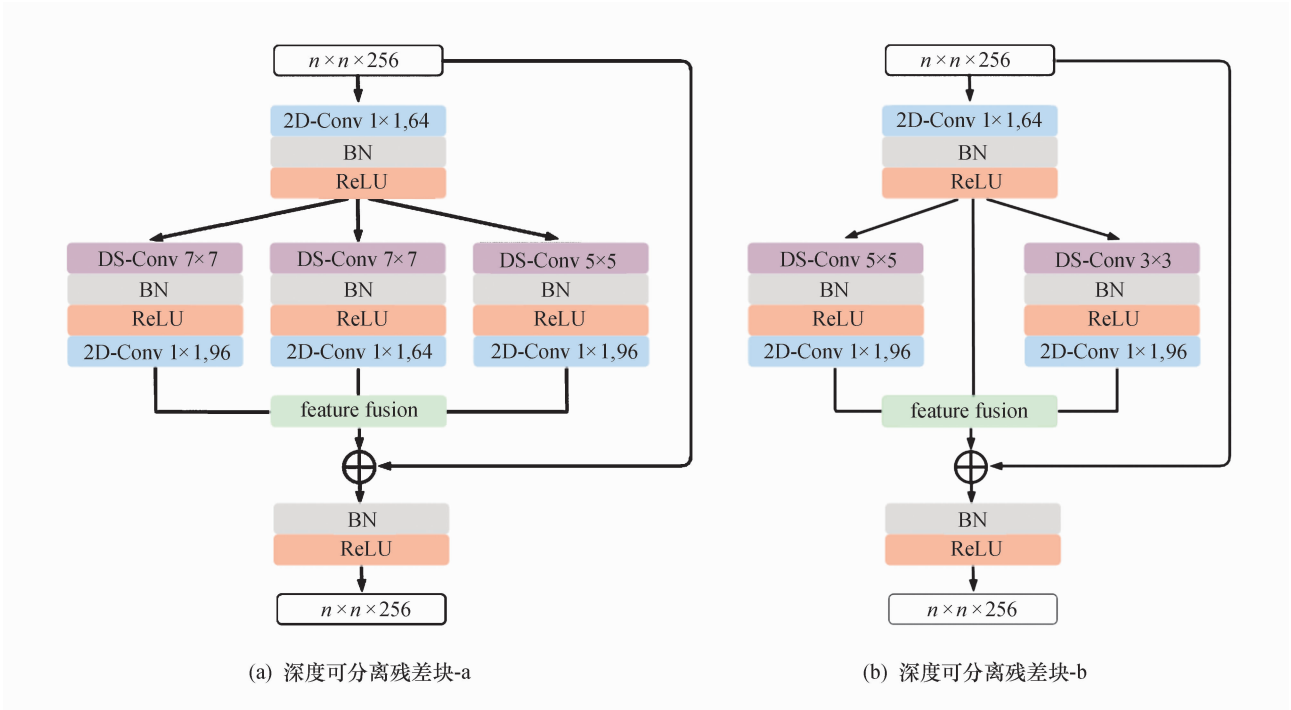


图 2 特征融合残差结构  
Fig. 2 Feature fusion residual block

和  $K^{5 \times 5}$  两种不同感受野的卷积核对压缩后的数据进行特征提取, 输出特征数据分别为  $F_1^{n \times n \times 96}$ 、 $F_2^{n \times n \times 64}$  和  $F_3^{n \times n \times 96}$ , 再将所得特征数据进行融合, 得到一个  $F^{n \times n \times 256}$  的特征数据; 最后借助跳跃连接与

前层输出特征数据进行残差计算。  
在深度可分离残差块-b 中, 同样使用  $1 \times 1$  卷积对深度可分离残差块-a 输出的特征数据进行压缩, 然后分别使用核尺寸为  $K^{5 \times 5}$  和  $K^{3 \times 3}$  的卷积操

作进一步提取空间特征信息,再将其与前层  $1\times 1$  卷积所得特征数据进行融合;最后借助跳跃连接与前层输出特征数据进行残差计算.

由于 DS-residual&SE 特征提取单元中添加了池化操作对特征数据进行下采样处理,为了提高模型提取空间特征的有效性,同时避免由于过多边缘填充带来的细节丢失问题,深度可分离残差块-b 相较于深度可分离残差块-a,选择采用感受野更为精细的卷积核组合,该结构在做到多尺度特征提取的同时,模型计算量增加较小.

表 1 是在 epoch 设为 50 的条件下两种卷积类型在 MDSR&SE-Net 中的性能对比,可以看出 DS-Conv 在保证网络模型最终分类精度的同时参数数量与训练时间明显减少.

表 1 不同卷积类型参数数量和性能对比  
Tab. 1 Comparison of the performance and parameters of different convolution

#	卷积类型	参数数量	训练时间/s	OA/%
Indian Pines	2D-Conv	2 265 730	417	99.13
	DS-Conv	614 786	403	99.12
Salinas	2D-Conv	2 265 730	1 484	99.54
	DS-Conv	614 786	1 432	99.54
Pavia University	2D-Conv	2 264 827	1 145	99.36
	DS-Conv	613 883	1 097	99.49

2.3 改进的 SE 模块

原始 SE 模块使用 GAP 对输入特征进行压缩抽取,然后通过激活操作输出包含通道赋值的权重信息. 本文在原始 SE 模块的基础框架中进行了改进,提出一种改进的 SE 模块,如图 3 所示.

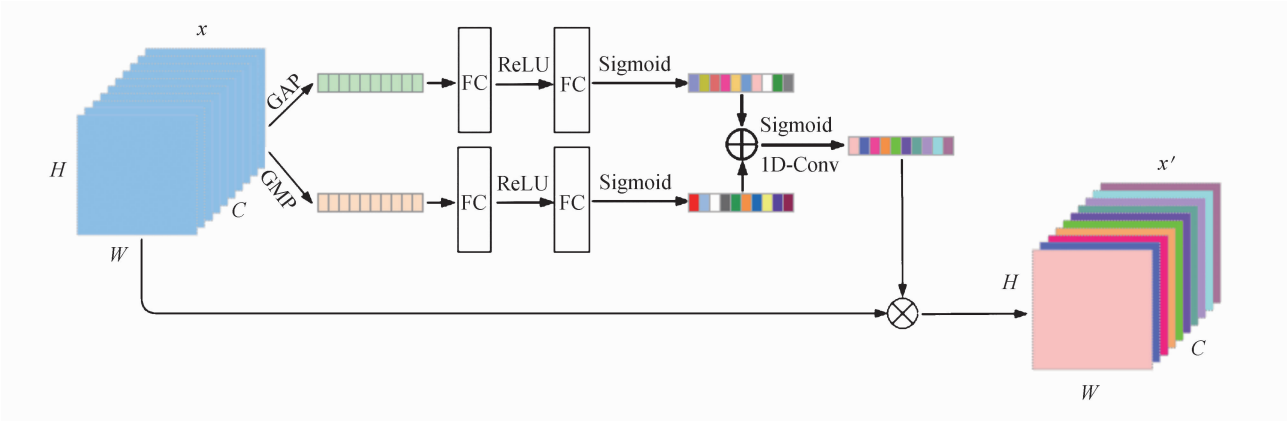


图 3 改进的压缩激活模块  
Fig. 3 Improved squeeze and excitation block

在改进的 SE 模块中,压缩操作通过 GAP 和全局最大池化(global max pooling,GMP)实现,将深度可分离残差块输出数据  $x$  的特征通道压缩成两组通道描述向量,两组向量具有不同侧重点的全局感受野.将压缩操作得到的一维向量通过激励操作提取每个特征通道之间的相关性.具体地,激活操作将 GAP 和 GMP 操作得到的两组向量分别通过采用瓶颈结构的 FC 层进行神经关联,使用 ReLU 和 Sigmoid 函数激活.然后将两组向量进行融合后使用一维卷积进一步提取通道相关性,赋予每个通道特定的权重用以标示不同通道的关注度,最终输出一个显示重要信息的特征数据  $x'$ .

3 实验与分析

本文所有实验在 Intel(R) Xeon(R) Silver 4116 @2.10 GHz,NVIDIA Tesla T4,128 GB 内存硬件支持下运行.基于 Windows10 系统使用 Python3.7.6 的 Tensorflow2.1 框架实现.

3.1 实验数据与评价指标

实验中使用常见的三个高光谱图像数据集,分别是 Indian Pines(IP)、Salinas(SA)和 Pavia University(PU).其中 IP 是由 AVIRIS 对美国印第安纳州一块区域的成像,截取了其中大小为  $145\times 145$  的图像进行标注,拥有 200 个波段,可分类像素 10 249 个,包含小麦、玉米、大豆、森林等 16 种类别标签.PU 是使用光学光谱成像仪对意大利帕维亚大学的成像,尺寸大小为  $610\times 340$ ,拥有 103 个波段,可分类像素 42 776 个,包含树、牧场、道路、砖块等 9 种类别标签.SA 是 AVIRIS 对美国萨利纳斯山谷的成像,该图像的尺寸大小为  $512\times 217$ ,拥有 204 个波段,可分类像素 54 129 个,包含芹菜、玉米等 16 种类别标签.三个数据集中,IP 样本数量最少,空间分辨率最低,因此该数据集的分类难度最大.

评价指标分别使用整体分类精度(overall accuracy,OA)、平均分类精度(average accuracy,AA)和 Kappa 系数(KA)来表示网络的分类精度.

为了能够验证 MDSR&SE-Net 在有限训练样本条件下依然可以获得较高的分类精度,对 IP、PU 和 SA 三个数据集选取有限数据作为训练样本,具体样本划分见表 2~4。

表 2 IP 数据集中训练样本与测试样本的划分情况  
Tab. 2 Segmentation of training samples and test samples in IP dataset

#	地物类别	训练样本数	测试样本数
1	Alfalfa	5	41
2	Corn-notill	143	1 285
3	Corn-mintill	83	747
4	Corn	24	213
5	Grass-pasture	48	435
6	Grass-trees	73	657
7	Grass-pasture-mowed	3	25
8	Hay-windrowed	48	430
9	Oats	2	18
10	Soybean-notill	97	875
11	Soybean-mintill	245	2 210
12	Soybean-clean	59	534
13	Wheat	20	185
14	Woods	126	1 139
15	Buildings-Grass-Trees-Drives	39	347
16	Stone-Steel-Towers	9	84
Total		1 024	9 225

表 3 PU 数据集中训练样本与测试样本的划分情况  
Tab. 3 Segmentation of training samples and test samples in PU dataset

#	地物类别	训练样本数	测试样本数
1	Asplalt	298	6 333
2	Meadows	839	17 810
3	Gravel	94	2 005
4	Tree	138	2 926
5	Painted metal sheets	60	1 285
6	Bare Soil	226	4 803
7	Bitumen	60	1270
8	Self-blocking bricks	166	3 516
9	Shadows	43	904
Total		1 924	40 852

3.2 参数选定

输入 MDSR&SE-Net 的图像邻域尺寸 patch 是一个重要的超参数,若 patch 设置过大,则会增加网络模型的计算消耗;而 patch 设置过小,将会造成某些部分空间特征的丢失,影响最终的分类准确度。学习率 lr 的确定同样重要,如果简单将学习率 lr 设置过大,虽然可以显著加快模型的训练速度,但是极有可能出现模型直接越过最低损失值的情况,最终导致损失值出现震荡,模型无法收敛;相反,学习率 lr 设置过小,模型收敛会变得缓慢,由于隐层权重更

表 4 SA 数据集中训练样本与测试样本的划分情况  
Tab. 4 Segmentation of training samples and test samples in SA dataset

#	地物类别	训练样本数	测试样本数
1	Bro-Coil_green_weeds_1	70	1 939
2	Bro- Coil_green_weeds_2	130	3 596
3	Fallow	69	1 907
4	Fallow_rough_plow	49	1 345
5	Fallow_smooth	94	2 584
6	Stubble	139	3 820
7	Celery	125	3 454
8	Grapes_untrained	394	10 877
9	Soil_vinyard_develop	217	5 986
10	Corn_senesced_green_weeds	115	3 163
11	Lettuce_romaine_4wk	37	1 031
12	Lettuce_romaine_5wk	68	1 859
13	Lettuce_romaine_6wk	32	884
14	Lettuce_romaine_7wk	38	1 032
15	Vinyard_untrained	254	7 014
16	Vinyard_vertical_trellis	63	1 744
Total		1 894	52 235

新变慢,同样存在得不到最低损失值的可能. 因此对于不同的分类模型,确定与之匹配的超参数是影响其分类精度的关键. 此外选择恰当的随机失活率 dropout 和训练轮数 epoch 以防止模型过度训练. 图 4 分别显示出在三个不同的数据集上,超参数 patch、lr、epoch 和 dropout 对最终分类结果的影响. 实验结果显示,在 IP 数据集中 patch 设置为 13 取得了最高准确度,而在 SA 和 PU 数据集中 patch 设置为 11 取得了最高准确度;将 lr、epoch 和 dropout 分别确定为 0. 002、50 和 0. 2;L2 正则化参数  $\gamma$  设置为常用的 0. 001.

3.3 特征融合残差块组合

不同的特征融合残差块组合在对三个数据集进行分类时存在着性能差异(见表 5~7). 当仅设置 4 个深度可分离残差块-a 或 4 个深度可分离残差块-b 结构时,训练时间和三种评价指标均没有达到最优;而当设置 2 组深度可分离残差块-a 串联深度可分离残差块-b 结构时,虽然三种评价指标有所提升,但是训练时间有所增加;当选择设置 2 个深度可分离残差块-a 串联 2 个深度可分离残差块-b 结构时,训练时间和分类精度表现为最优.

3.4 改进 SE 模块的实验性能

为了验证本文提出的改进 SE 模块对特征通道的加权处理能力优于原始 SE 模块,同时说明将 SE 模块嵌入原主干网络 MDSR-Net 后,网络模型最终的分类性能有所提升,使用不同的主干网络设置了消融实验,实验结果如表8所列. 相较于MDSR-Net,



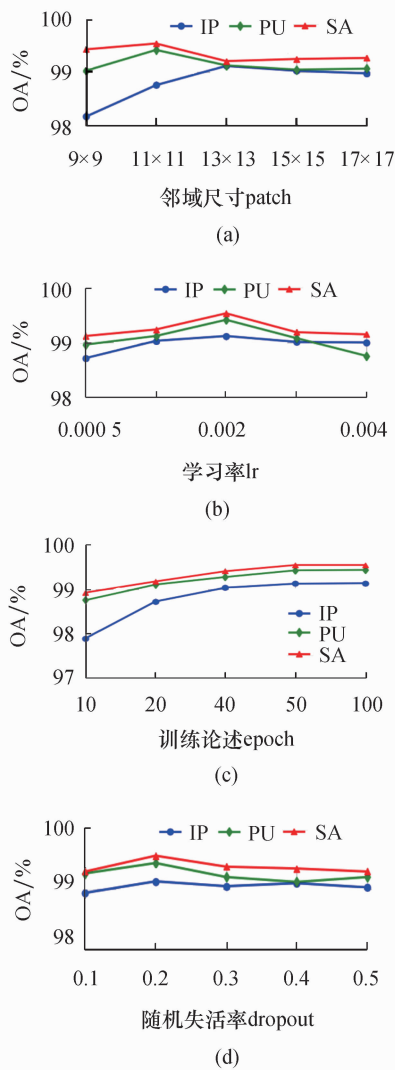


图 4 在使用不同 patch、lr、epoch 和 dropout 的性能对比  
Fig. 4 Performance with different patch, lr, epoch and dropout

表 5 不同特征融合残差块组合在 IP 中的性能表现  
Tab. 5 Performance of different combination of residual blocks in IP

Residual block 组合	参数数量	训练时间/s	OA/%
[Block-a]×4	634 882	438	99.08
[Block-b]×4	594 690	363	98.97
[[Block-a]+[Block-b]]×2	614 786	410	99.08
[Block-a]×2+[Block-b]×2	614 786	403	99.12

表 6 不同特征融合残差块组合在 SA 中的性能表现  
Tab. 6 Performance of different combination of residual blocks in SA

Residual block 组合	参数数量	训练时间/s	OA/%
[Block-a]×4	634 882	1 533	99.43
[Block-b]×4	594 690	1 357	99.10
[[Block-a]+[Block-b]]×2	614 786	1 449	99.46
[Block-a]×2+[Block-b]×2	614 786	1 432	99.54

表 7 不同特征融合残差块组合在 PU 中的性能表现  
Tab. 7 Performance of different combination of residual blocks in PU

Residual block 组合	参数数量	训练时间/s	OA/%
[Block-a]×4	633 979	1 236	99.27
[Block-b]×4	593 787	995	99.02
[[Block-a]+[Block-b]]×2	613 883	1 094	99.30
[Block-a]×2+[Block-b]×2	613 883	1 097	99.49

表 8 SE 模块和改进 SE 模块在网络模型中的精度对比  
Tab. 8 Precision compare between SE module and improved SE module

数据集	分类精度	MDSR-Net	MDSR&SE*-Net	MDSR&SE-Net
IP	OA/%	98.98	99.06	99.12
	AA/%	98.86	98.95	99.17
	KA×100	98.56	98.67	98.99
SA	OA/%	99.30	99.39	99.54
	AA/%	99.47	99.60	99.65
	KA×100	99.32	99.37	99.49
PU	OA/%	99.19	99.28	99.49
	AA/%	99.27	99.30	99.36
	KA×100	99.10	99.19	99.33

添加原始 SE 模块的 MDSR&SE\*-Net 分类精度有小幅增加,而添加改进 SE 模块的 MDSR&SE-Net 则有较为明显的精度提升.

3.5 实验结果与分析

本文的对比实验在训练样本划分和训练轮数相同的前提下进行. 选择使用经典的 3D-CNN<sup>[18]</sup>、SSRN<sup>[14]</sup>、Cluster-CNN<sup>[19]</sup> 和 JSSAN<sup>[20]</sup> 四种模型与本文 MDSR&SE-Net 进行精度对比,以验证其分类性能. 本文所有实验设置为运行 20 次,在所得 20 组实验数据中选择精度最高的 10 组,并取其平均值,五种模型的分类精度如表 9 所列.

在 IP 数据集中,MDSR&SE-Net 和其他四种分类模型中性能最好的 JSSAN 相比,OA、AA 和 KA 分别提升了 0.21%、0.63% 和 0.23;而在 SA 与 PU 数据集中,相较于 JSSAN、MDSR&SE-Net 取得的 OA 分别提升了 0.11%和 0.19%,AA 分别提升了 0.25%和 0.08%,KA 分别提升了 0.14 和 0.13.

图 5~7 为五种模型在三个数据集上取得的分类预测图. 可以看出,3D-CNN 在训练样本有限的情况下出现了明显的错误分类;SSRN、Cluster-CNN 和 JSSAN 的分类结果优于 3D-CNN,但依然在地物边缘区域存在较多的误差;而 MDSR&SE-Net 利用 3D-CNN 对高光谱图像初始空-谱特征数据进行了提取,由于其设置于主干网络前端,因此可以避免过

表 9 五种分类模型在不同数据集中的分类精度

Tab. 9 The accuracy of five classification models in different data sets

数据集	Accuracy	3D-CNN	SSRN	Cluster-CNN	JSSAN	Proposed
IP	OA/%	93.23	97.95	98.08	98.91	99.12
	AA/%	95.51	97.23	98.46	98.54	99.17
	KA×100	92.30	97.69	97.33	98.76	98.99
SA	OA/%	97.06	99.07	99.12	99.43	99.54
	AA/%	98.56	98.66	99.10	99.40	99.65
	KA×100	96.03	98.93	98.99	99.35	99.49
PU	OA/%	96.16	99.06	98.75	99.30	99.49
	AA/%	97.85	98.75	98.96	99.28	99.36
	KA×100	94.75	98.14	98.60	99.20	99.33

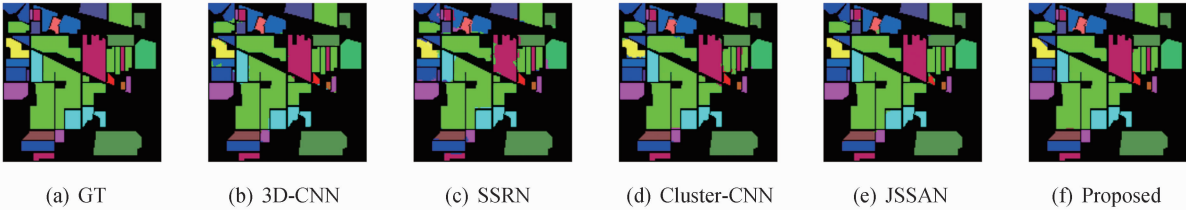


图 5 五种分类模型获得的 IP 预测图与真实图的对比

Fig. 5 Comparison between the predicted and truth maps of IP obtained by five classification models

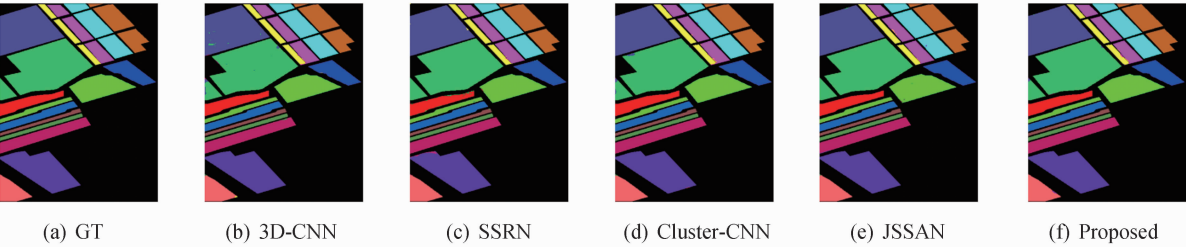


图 6 五种分类模型获得的 SA 预测图与真实图的对比

Fig. 6 Comparison between the predicted and truth maps of SA obtained by five classification models

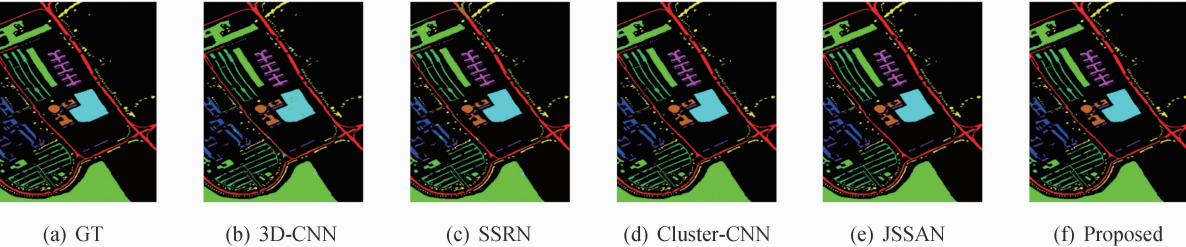


图 7 五种分类模型获得的 PU 预测图与真实图的对比

Fig. 7 Comparison between the predicted and truth maps of PU obtained by five classification models

多通道信息的丢失. 得益于多尺度残差结构及通道注意力机制, 通过对前层输出特征图的多次整合, 可极大程度地关注特征数据块更富价值的通道特征, 同时能够充分学习输出数据更为抽象的空间特征. 实验显示, 使用 AM 的分类模型总体分类性能优于传统 CNN 模型, 本文所提出的 MDSR&SE-Net 整体分类精度相较于其他四种模型有明显提升.

4 结语

目前高光谱图像分类模型存在参数数量庞大, 结构冗余复杂, 训练时间较长, 并且需要依赖较多的训练集划分来获得不错的分类结果的问题, 本文提出 MDSR&SE-Net 进行了针对性的优化. 该网络结构具有参数数量少, 训练时间短, 同时在有限训练样



本条件下依然可以取得较高的分类精确度,但是,其网络结构与分类精度还有进一步改进和优化的空间.在今后的研究中,将考虑结合更加合理的数据预处理算法以及引入混合深度可分离卷积进一步提高网络模型的分类能力.

参考文献:

[1] MA H Q, HUANG W J, DONG Y Y, *et al.* Using UAV-based hyperspectral imagery to detect winter wheat fusarium head blight [J]. *Remote Sensing*, 2021, 13(15):3024.

[2] NTURAMBIRWE J F, PEROLD W J, OPARA U L, *et al.* Classification learning of latent bruise damage to apples using shortwave infrared hyperspectral imaging [J]. *Sensors*, 2021, 21(15):4990.

[3] YU R, REN L L, LUO Y Q, *et al.* Early detection of pine wilt disease in *Pinus tabulaeformis* in North China using a field portable spectrometer and UAV-based hyperspectral imagery [J]. *Forest Ecosystems*, 2021, 8(1):44.

[4] XIE W Y, ZHANG J Q, LEI J, *et al.* Self-spectral learning with GAN based spectral-spatial target detection for hyperspectral image [J]. *Neural Networks*, 2021, 142:375-387.

[5] REZA S M, HASAN G. A probabilistic SVM approach for hyperspectral image classification using spectral and texture features [J]. *International Journal of Remote Sensing*, 2017, 38(15):4265-4284.

[6] KOHEI A. Maximum likelihood classification based on classified result of boundary mixed pixels for high spatial resolution of satellite images [J]. *International Journal of Advanced Computer Science and Applications*, 2020, 11(9):24-30.

[7] 颜铭靖, 苏喜友. 基于三维空洞卷积残差神经网络的高光谱影像分类方法 [J]. *光学学报*, 2020, 40(16):163-172.

[8] JOHN R A, KANDASAMY J, VIGNESH T, *et al.* Development of computer vision for inspection of bolt using convolutional neural network [J]. *Materials Today: Proceedings*, 2021, 45(7):6931-6935.

[9] LAKHAL M Y, CEVIKALP H, ESCALERA S. *et al.* Recurrent neural networks for remote sensing image classification [J]. *IET Computer Vision*, 2018, 12(7):1040-1045.

[10] ATIF M, LIN M T. Multiple deep-belief-network based spectral-spatial classification of hyperspectral images [J]. *Tsinghua Science and Technology*, 2019, 24(2):183-194.

[11] 康 妙, 计科峰, 冷祥光, 等. 基于栈式自编码器特征融合的 SAR 图像车辆目标识别 [J]. *雷达学报*, 2017, 6(2):167-176.

[12] 胡 丽, 单 锐, 王 芳, 等. 基于双通道空洞卷积神经网络的高光谱图像分类 [J]. *激光与光电子学进展*, 2020, 57(12):356-362.

[13] ROY S K, KRISHNA G, DUBEY S R, *et al.* HybridSN: exploring 3D-2D CNN feature hierarchy for hyper spectral image classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(2):277-281.

[14] ZHONG Z L, LI J, LUO Z M, *et al.* Spectral-spatial residual network for hyperspectral image classification: a 3-D deep learning framework [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(2):847-858.

[15] HU J, SHEN L, SUN G, *et al.* Squeeze-and-excitation networks [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2018:7132-7141.

[16] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016:770-778.

[17] YUE J, ZHAO W Z, MAO S J, *et al.* Spectral-spatial classification of hyperspectral images using deep convolutional neural networks [J]. *Remote Sensing Letters*, 2015, 6(6):468-477.

[18] CHEN S T, JIN M, DING J. Hyperspectral remote sensing image classification based on dense residual three-dimensional convolutional neural network [J]. *Multimedia Tools and Applications*, 2021, 80(2):1859-1882.

[19] YAO W, LIAN C, BRUZZONE L. Cluster-CNN: clustering-based feature learning for hyperspectral image classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(11):1991-1995.

[20] LI L, YIN J H. Joint spatial-spectral attention network for hyperspectral image classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(10):1816-1820.