

结合动态卷积和三重注意力机制的高光谱图像分类

王爱丽¹, 刘美红¹, 薛冬¹, 吴海滨^{1*}, 赵蓝飞¹, 岩堀祐之²

¹哈尔滨理工大学测控技术与通信工程学院黑龙江省激光光谱技术及应用重点实验室, 黑龙江 哈尔滨 150080;

²中部大学计算机科学学院, 日本 爱知 487-8501

摘要 考虑到高光谱图像训练样本数量有限及高光谱维度对分类精度的影响, 提出了一种结合动态卷积和三重注意力机制(TA)的高光谱分类算法。首先, 采用主成分分析(PCA)去除光谱冗余, 并将处理后的数据输入改进的残差网络中。然后, 在残差网络中引入动态卷积, 利用动态卷积核提取深度精细化特征, 并利用TA模型实现跨维度信息交互, 关注更重要的高光谱空间-光谱特征, 降低无用信息的影响。最后, 使用Softmax全连接层实现对高光谱图像的分类。在Pavia University, Kennedy Space Center, Salinas 3个公开数据集上与其他6种分类算法进行了比较, 实验结果表明, 所提算法的分类效果最优, 总体分类精度分别达到了97.49%、94.21%、98.65%。

关键词 机器视觉; 高光谱图像分类; 残差网络; 动态卷积; 三重注意力机制; 跨维度信息

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP202259.1015011

Hyperspectral Image Classification Combined Dynamic Convolution with Triplet Attention Mechanism

Wang Aili¹, Liu Meihong¹, Xue Dong¹, Wu Haibin^{1*}, Zhao Lanfei¹, Iwahori Yuji²

¹Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application, College of measurement and control technology and communication Engineering, Harbin University of Science and Technology, Harbin 150080, Heilongjiang, China;

²Department of Computer Science, Chubu University, Aichi 487-8501, Japan

Abstract Considering the limited number of training samples of hyperspectral images and the influence of high spectral dimension on classification accuracy, a novel hyperspectral classification algorithm combined dynamic convolution with triplet attention mechanism (TA) is proposed. First, principal component analysis (PCA) is used to remove spectral redundancy. Then, the processed data are input into the modified residual network. Second, dynamic convolution is introduced in the residual network to extract deep and refined features. The TA model is used to interact with cross-dimensional information to focus on the extremely important hyperspectral spatial-spectral features and reduce the impact of useless information. Finally, the Softmax fully connected layer is used to realize the classification of hyperspectral images. The results demonstrate that the classification effect of the proposed algorithm outperforms six other classification algorithms on three public datasets of Pavia University, Kennedy Space Center, and Salinas. Furthermore, the overall classification accuracy of the proposed algorithm reaches 97.49%, 94.21%, and 98.65% on three datasets, respectively.

Key words machine vision; hyperspectral image classification; residual network; dynamic convolution; triplet attention mechanism; cross-dimensional information

收稿日期: 2021-08-16; 修回日期: 2021-08-24; 录用日期: 2021-09-25

基金项目: 国家自然科学基金(61671190)、JSPS科学基金(#20K11873)、黑龙江省普通高校专项资金(LGYC2018JC050)

通信作者: *woo@hrbust.edu.cn

1 引言

高光谱图像具有图谱合一且光谱信息丰富、光谱分辨率高、成像波段数量多、在某一光谱段范围内连续成像等优点,在土地覆盖分析^[1]、环境监测^[2]和军事侦察^[3]等众多任务上有着广泛应用前景,因此关于高光谱图像高效分类的研究成为了重点。

深度学习已经成为高光谱图像分类领域的研究新热点,相较于传统的机器学习算法诸如支持向量机(SVM)、K-近邻法(K-NN)、朴素贝叶斯、决策树等,基于深度学习的高光谱图像分类算法能够自动获取图像的高级特征,使得分类模型能更好地表征数据的特点,提高分类精度。残差网络(ResNet)在精度和收敛等方面都展现出了很好的特性,如Zhong等^[4]使用具有不同深度和宽度的深度残差网络进行高光谱图像分类,王爱丽等^[5]提出了一种以残差网络架构为基础的胶囊网络和空洞卷积相结合的激光雷达(LiDAR)数据分类算法,李勇等^[6]提出了一种基于改进残差密集网络的重建算法,这些算法均取得了较好的结果。张怡卓等^[7]提出了一种使用残差分层融合网络对高光谱图像的地物进行分类的方法,该方法可以自动提取更丰富的光谱空间信息,解决了网络层数增加带来的负担,但是该方法仍然存在耗时较长的问题。针对高光谱图像中存在的训练样本少、光谱维度高的问题,郑姗姗等^[8]提出了一种改进的多尺度三维残差网络模型,该模型在提取图像深层局部空间-光谱联合特征的同时降低了参数量,加速了模型的收敛速度。韩彦岭等^[9]提出了一种基于残差网络特征融合的高光谱图像分类方法,该方法解决了高光谱图像分类问题中深度特征提取不足的问题,获得了较好的分类性能。

上述研究表明,堆积残差单元构建更好更复杂的网络虽然提升了模型的性能,但是静态卷积的使用,固定的人为设置的参数不利于高光谱图像特征的提取,因此Chen等^[10]设计了一种动态卷积,该卷积可以动态地聚合多个并行卷积核并自适应地调整卷积参数,在不增加网络深度或宽度的情况下改善了模型的性能,在ImageNet和COCO数据集上的实验结果表明,该动态卷积显著提高了特征表达能力。Wei等^[11]使用动态混合分组卷积对图像进行分类,该方法有效减小了卷积的计算量和卷积模型的尺寸。Zhu等^[12]提出了一种动态显著目标检测网

络,该网络采用两步训练策略端到端地训练动态网络,提高了识别精度。

与此同时,在大量的网络模型训练过程中,卷积层对特征的关注度并不是相同的,因此引入注意力机制有区别地处理不同的特征也是高光谱图像分类的一个研究热点^[13]。魏祥坡等^[14]设计了一种残差通道注意力网络,在残差网络中将注意力机制和卷积层结合,重新调整特征通道,促进信息传递。Wang等^[15]将挤压-激励注意力模块嵌入残差网络进行高光谱图像分类,嵌入注意力模块的残差网络能自适应地学习不同光谱波段和相邻像素的权值,有效地抵抗噪声干扰,提高了分类效果。Ma等^[16]提出了一种双分支多注意力机制网络,该网络使用卷积块注意模块集成频谱和空间信息,减少了两类特征之间的干扰,提高了分类精度。张祥东等^[17]提出了一种结合扩张卷积与注意力机制的三维-二维串联卷积神经网络模型,该模型有效提高了高光谱训练样本在有限条件下的分类精度。王欣等^[18]提出了一种改进的DenseNet联合空谱注意力机制的分类方法,在空间维度引入空间注意力机制,在光谱维上引入了通道注意力机制。上述方法在进行高光谱图像分类时,空间注意模型和通道注意模型都是完全孤立的,三重注意力机制(TA)^[19]的出现很好地解决了这个问题,Cui等^[20]还提出了一种新的双三重注意力网络,交互的分支应用于光谱维度时,构建通道维度与空间维度的跨维度信息交互,应用于空间维度时,也同时考虑了与通道的相关性,实现了高光谱图像的高精度分类。

本文提出了一种结合动态卷积与三重注意力机制的高光谱分类算法(DTAResNet)。首先,将经过主成分分析(PCA)降维的高光谱图像送入残差网络,并将残差网络中的恒等残差块的后面两层卷积替换为动态卷积,通过多个并行的卷积核集成为一个动态核,提高了特征表达能力,减少了额外参数的计算,实现了深度特征的提取,避免了单一特征信息的不完整表达。其次,在动态卷积层后引入了TA模块,实现了通道维度和空间维度之间的跨维度交互,提高了网络在训练过程中对重要特征的提取能力。最后,在3个公开的高光谱数据集上,将所提算法与其他6种算法进行了对比实验,实验结果表明,在训练样本有限的情况下,所提算法的平均分类精度最高。

2 所提算法

所提结合动态卷积和三重注意力机制的高光谱分类算法整体流程如图 1 所示。首先使用 PCA 降维去除光谱冗余。考虑到深度学习中网络层数的加深会导致梯度消失、梯度爆炸的现象,分类模型采用残差网络作为基本框架;动态卷积的卷积核不同于常规卷积中的单一核,而是集成多个并行的卷积核作为一个动态核,动态卷积可以提升

模型表达能力而且无需提升网络深度与宽度,因此将恒等残差块中后面两个卷积层替换为动态卷积来充分挖掘高光谱图像的深层特征信息,从而进一步提高分类精度;此外,为了关注更重要的特征,在动态卷积层后加入三重注意力机制模块,实现跨维度信息交互,将注意力投入重点区域,提取更具有价值的空间信息和光谱信息,减少无用信息的干扰;最后利用基于 Softmax 的全连接层(FC)进行分类。

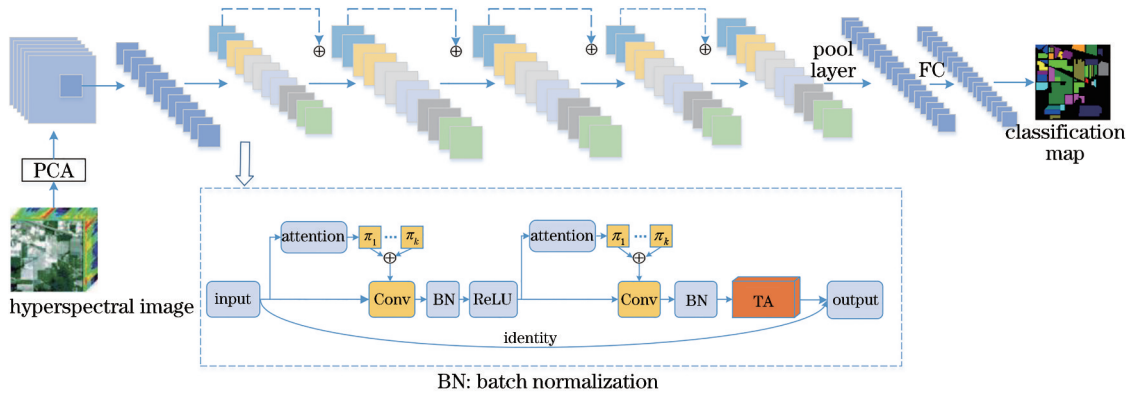


图 1 结合动态卷积和三重注意力机制的高光谱图像分类算法流程图

Fig. 1 Flow chart of hyperspectral image classification algorithm combined dynamic convolution with triple attention mechanism

2.1 残差网络

卷积神经网络在图像分类、目标检测等领域都取得了较好的效果,随着深度学习领域的不断发展,越来越多的优秀模型相继被提出,且研究发现:并不是网络结构越深,模型的效果越好,相反随着模型层数的加深,会出现梯度消失、梯度爆炸等现象。ResNet 的出现解决了这一问题,其主要思想是在输入通道和输出通道之间增加具有联系的通道——“shortcut connection”,输入信息可以直接到达输出,避免了信息的遗漏,同时训练网络时只需要上面输出的残差,原理如图 2 所示。

图 2 中, x 表示输入, $H(x)$ 表示输出, $F(x)$ 表示残差映射,学习形成残差函数,最终残差单元的输出 $H(x) = F(x) + x$,这种残差函数更容易优化,也能使网络层数大大加深。残差单元的计算过程为

$$F(x) = W_2 \sigma(W_1 x), \quad (1)$$

式中: σ 代表非线性函数 ReLU; W_1 和 W_2 分别为权重层 1 和权重层 2 的权重。然后通过一个 shortcut 和第 2 个 ReLU 获得输出 $H(x)$ 。

$$H(x) = F(x, \{W_i\}) + x. \quad (2)$$

当需要对输入和输出维数进行变化时(比如改变通道数目),可以在进行 shortcut 操作时对 x 进行

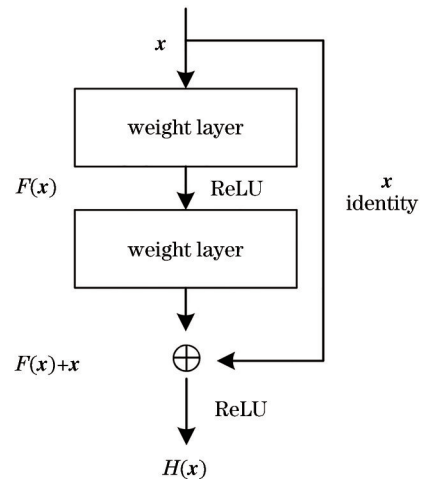


图 2 ResNet 的残差单元结构图

Fig. 2 Residual unit structure diagram of ResNet

一个线性变换 W_s , 即

$$H(x) = F(x, \{W_i\}) + W_s x. \quad (3)$$

通常单一层的残差块并不能起到提升作用,因此往往两层以上的残差块才能解决网络性能退化问题。

2.2 动态卷积网络

动态卷积是为了在网络结构与计算消耗中找到一个平衡,在不增加网络的深度或宽度的情况下增加模型的表达能力,即根据输入图像自适应地调

整卷积的参数。动态卷积在每层使用的不是一个卷积核,而是根据多个并行卷积核进行动态聚合,调整每个卷积核的权重,针对性地选择合适的参数进行特征的提取。一般的动态感知器如图 3 所示。

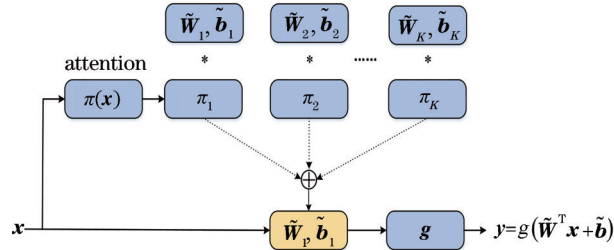


图 3 动态感知器

Fig. 3 Dynamic perceptron

从图 3 可以看出,输出结果

$$y = g(\tilde{W}^T x + \tilde{b}), \quad (4)$$

$$\tilde{W} = \sum_{k=1}^K \pi_k(x) \tilde{W}_k, \quad (5)$$

$$\tilde{b} = \sum_{k=1}^K \pi_k(x) \tilde{b}_k, \quad (6)$$

$$\text{s.t. } 0 \leq \pi_k(x) \leq 1, \sum_{k=1}^K \pi_k(x) = 1, \quad (7)$$

式中: W 、 b 、 g 分别表示权重、偏置和激活函数; π_k 表示注意力权值,并不是固定的,而是随着输入的变化而变化。 π_k 包含注意力权值计算和动态权值融合,即

$$O(\tilde{W}^T x + \tilde{b}) \gg O(\sum \pi_k \tilde{b}_k) + O[\pi(x)], \quad (8)$$

式中: $O(\cdot)$ 表示感知器的计算量。

与动态感知器相似,动态卷积也具有 K 个卷积核,如图 4 所示。在动态卷积之后接 BN 层和 ReLU 层,给其中某个层设置 K 个尺度和通道数相同的

核,通过各自的注意力权重进行融合,从而得到该层的卷积核参数。同时在注意力层中首先进行全局平均池化,得到全局空间特征,再通过两个全连接层映射到 K 维度,最后进行 Softmax 归一化,这样得到的注意力权重就可以分配给该层的 K 个核。原本的一个固定的卷积核,现在变为可以根据输入进行动态选取的卷积核,显著提高了特征表达能力。

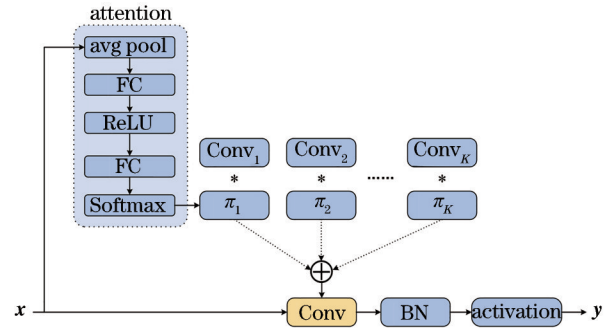


图 4 动态卷积

Fig. 4 Dynamic convolution

2.3 TA 模型

注意力机制^[21]通常是一个额外的神经网络,能够选择输入的某些部分或给不同部分分配不同的权重,从而能够在大量信息中筛选出重要的信息。所使用的 TA 模型如图 5 所示,该模型通过使用三支结构捕获跨维度交互信息来计算注意力权重。对于输入张量,TA 通过旋转操作,使用残差变换建立维度间的依存关系,并对通道间和空间信息进行编码。前 2 个分支分别用来捕获通道维度 C 和空间维度 W 或 H 之间的跨通道交互,permutation 采用旋转操作来建立通道维度和任意一个空间维度之间

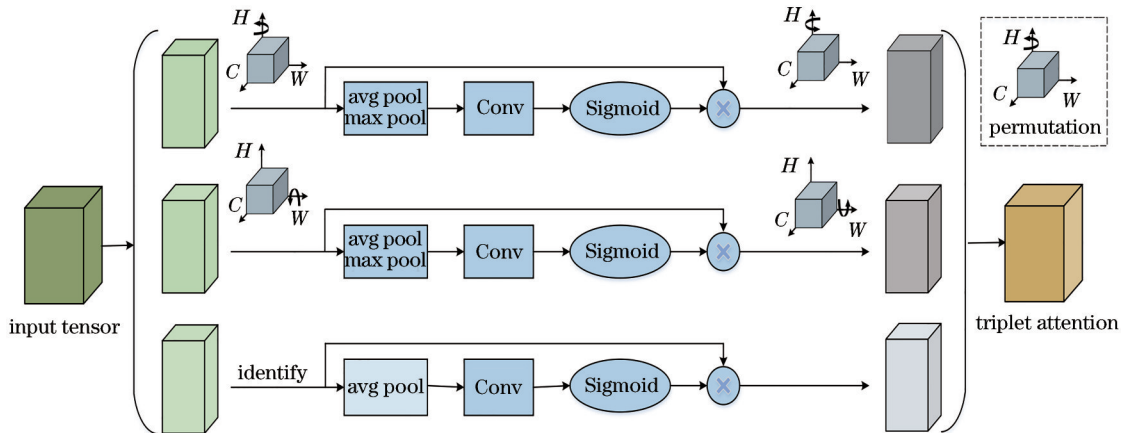


图 5 TA 原理图

Fig. 5 TA schematic diagram

的连接,第3分支用于获取传统的空间注意力权重。

TA模型的架构如图6所示,对于一个给定的大小为 $C \times H \times W$ 的输入,具体实现过程如下:

1) 第1分支为通道 C 和空间 H 维度交互捕获分支,输入特征经过跨维度相互作用(permute)后变为 $W \times H \times C$ 维度特征,接着在 W 维度上进行 Z-Pool, Z-Pool 是最大池化和平均池化,输出为 $2 \times H \times C$ 特征,然后经过 7×7 卷积层和 BN 层,最后经过跨维度相互作用变为 $C \times H \times W$ 维度特征。

2) 第2分支为通道 C 和空间 W 维度交互捕获分支,输入特征先经过跨维度相互作用变为 $H \times C \times W$ 维度特征,接着在 H 维度上进行 Z-Pool,然后经过 7×7 卷积层和 BN 层,最后经过跨维度相互作用再变为 $C \times H \times W$ 维度特征。

3) 第3分支为空间注意力计算分支,输入特征经过 channel pool、 7×7 卷积层、BN 层,最后由 Sigmoid 激活函数生成空间注意力权重。

4) 最后对3个分支输出特征相加求平均。

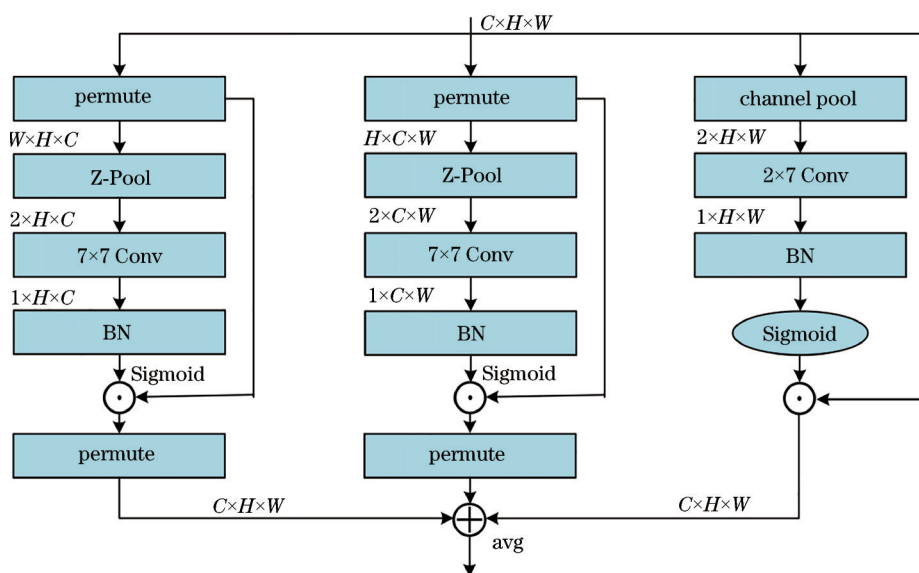


图6 TA架构图

Fig. 6 Architecture diagram of TA

3 实验结果及分析

3.1 实验数据集描述

为了验证所提算法的有效性,在 Pavia University、

Salinas 和 Kennedy Space Center 3 个公开的高光谱数据集上进行了分类实验,数据集参数信息如表 1 所示。分类结果评价指标为总体精度(OA)、平均精度(AA)和统计 Kappa 系数(K)。

表 1 实验数据集参数

Table 1 Experimental dataset parameters

Parameter	Dataset		
	Pavia University	Kennedy Space Center	Salinas
Sensor	ROSIS	AVIRIS	AVIRIS
Size /(pixel×pixel)	610×340	512×614	512×217
Resolution /m	1.3	18	3.7
Spectral band	103	176	204
Land-cover	9	13	16
Total sample pixel	42776	5211	54129

图 7 分别为 Pavia University、Kennedy Space Center 和 Salinas 的假彩色图和真值图。

3.2 实验平台参数设置

所有实验均是在配备 Intel(R) Xeon(R) 4208

CPU@2.10 GHz 处理器、Nvidia GeForce RTX 2080 Ti 显卡的系统上运行的。为了减小实验误差,从训练集中随机抽取有限样本进行训练,epoch 设置为 200,批大小设置为 32,取 10 次实验的平均值作

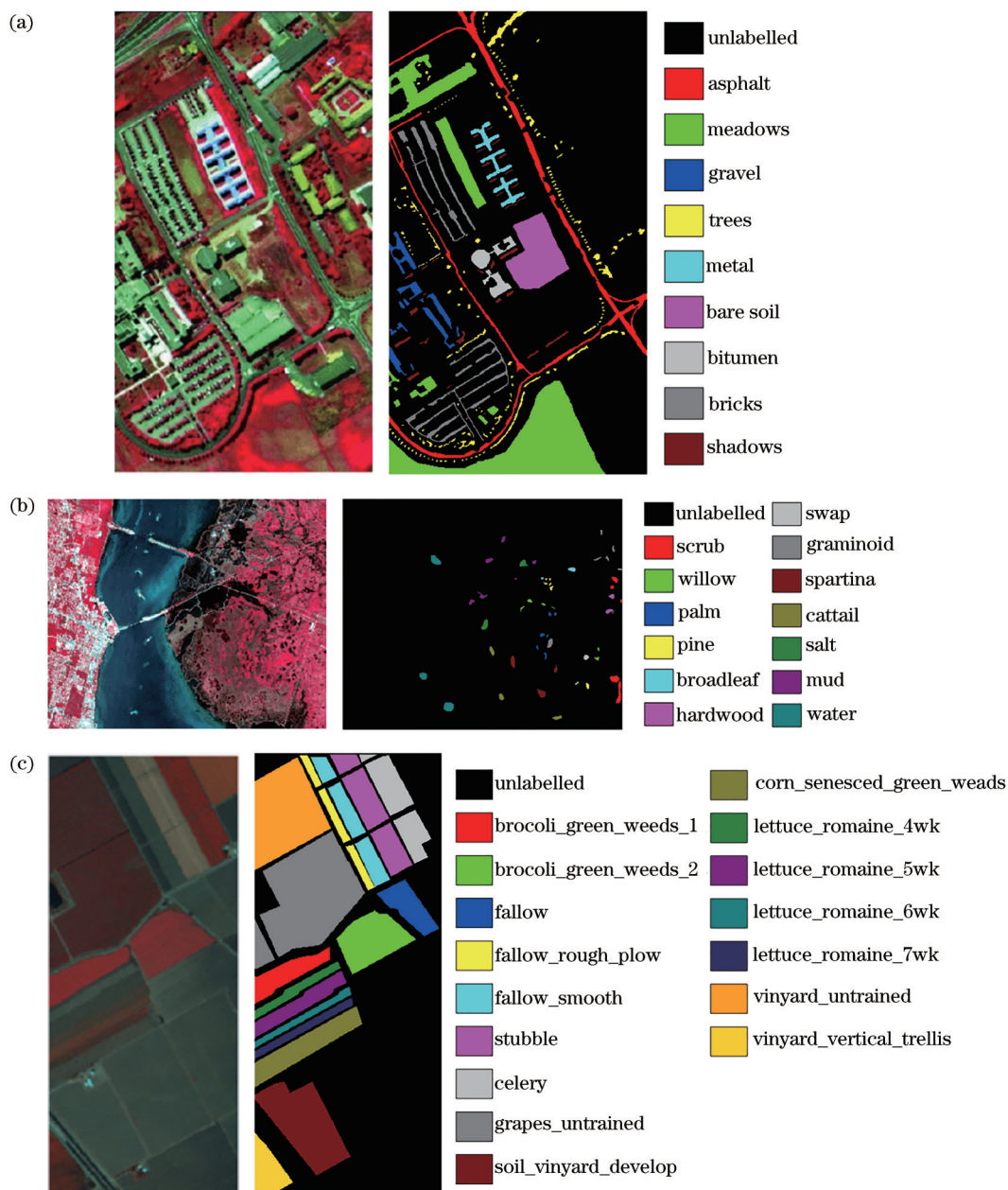


图 7 各数据集假彩色图和真值图。(a) Pavia University; (b) Kennedy Space Center; (c) Salinas

Fig. 7 False color map and ground truth map of each dataset. (a) Pavia University; (b) Kennedy Space Center; (c) Salinas

为最终结果。网络训练过程中,为了处理大量的可训练参数,使用 Adam 优化器学习三维频谱空间滤波器的权重,采用余弦退火调整学习率,以余弦函数为周期,并在每个周期最大值时重新设置学习率。初始学习率为 0.001,以 15 为周期,15 个 epoch 之后通过自动增加学习率跳过局部最优来重新设置。

所提算法以 ResNet-34 作为主要特征提取网络,ResNet-34 由 Conv1、Conv2_x、Conv3_x、Conv4_x、Conv5_x 组成,各部分参数设置如表 2 所示。图像输入尺寸为 9×9 ,最终 Conv5_x 输出图像大小为 $1 \times$

1。首先将处理好的大小为 $9 \times 9 \times \text{band}$ 的数据输入残差网络的普通卷积层中,其卷积核大小为 7×7 ,步长 stride 为 2, padding 为 3,该层用于对图像进行初始特征提取。然后将得到的大小为 $64 \times 5 \times 5$ 的特征图输入动态卷积层中,动态卷积的卷积核为 3×3 ,步长 stride 为 1, padding 为 1,而且动态卷积的卷积核是将注意力和权重相乘,得到 batch size 融合后的卷积核。本实验组的动态卷积核由 4 个 3×3 的卷积核组成,输出 $64 \times 3 \times 3$ 的特征图,最后输入三重注意力机制。

表 2 ResNet-34 的参数设置

Table 2 Parameter settings of ResNet-34

Layer name	Output size	ResNet-34 parameter setting
Conv1	5 × 5	7 × 7, 64, stride 2
		3 × 3 max pool, stride 2
Conv2_x	3 × 3	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$
Conv3_x	2 × 2	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$
Conv4_x	1 × 1	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$
Conv5_x	1 × 1	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$

3.3 分类结果分析

主要对比了所提算法与径向基函数支持向量机算法(RBF-SVM)^[22]、扩展形态轮廓支持向量机算法(EMP-SVM)^[23]、深度卷积神经网络(DCNN)^[24]、生成对抗网络(GAN)^[25]、残差网络(ResNet)^[26]、金字塔残差网络(PyResNet)^[27]对高光谱数据集的分类性能。各个数据集的样本被随机分为两个互不交叉的训练集和测试集,训练样本的数量也对高光谱图像分类精度有着重要影响,以Pavia University为例,训练样本分别为总样本的2%、4%、6%、8%、10%、12%时不同分类算法的总体分类精度如图8所示。从图8可以看出,一开始随着训练样本的增加,分类精度也在增加,当训练样本为10%的时候,OA基本达到了最高,之后趋于平缓,甚至出现了下降的趋势。因此,实验中训练集、测试集和验证集分别为总

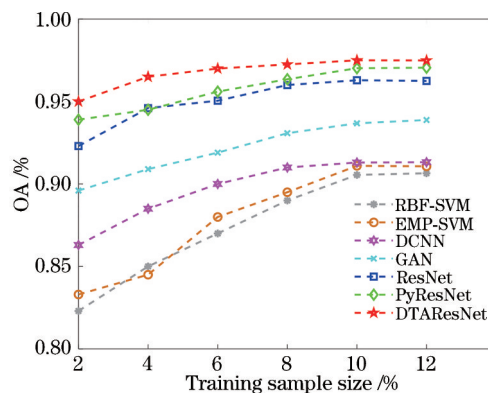


图 8 不同分类算法的总体分类精度对比分析

Fig. 8 Comparative analysis of overall classification accuracy of different classification algorithms

样本的10%、80%和10%。

表3~5为不同算法对3个数据集不同地物的分类精度,其中粗体数据表示最优结果。从表中可以看出,在Pavia University数据集中,所提算法的OA最高,达到了 $97.49 \pm 0.24\%$,与RBF-SVM、EMP-SVM、DCNN、GAN、ResNet、PyResNet对比,分别提高了6.84、6.42、6.17、3.61、1.00、0.44个百分点。Salinas数据集上的分类精度如表5所示,除“Grapes_untrained”“Vinyard_untrained”外,所提算法在其他14类样本上都达到了98%以上的分类精度,整体表现优于其他所有算法,验证了动态卷积和TA模块的引入能更好地增强特征判别能力,且能聚焦高光谱图像的关键场景区域,较大幅度提高分类精度。

表 3 Pavia University 数据集的分类结果

Table 3 Classification results of Pavia University dataset

Class	RBF-SVM	EMP-SVM	DCNN	GAN	ResNet	PyResNet	DTAResNet
Asphalt	88.51 ± 2.44	89.06 ± 1.33	90.59 ± 0.62	94.97 ± 2.09	95.59 ± 3.43	95.55 ± 0.25	96.35 ± 0.64
Meadows	88.86 ± 3.52	88.12 ± 0.23	89.26 ± 1.05	97.62 ± 1.96	97.10 ± 2.06	98.91 ± 0.71	99.16 ± 0.32
Gravel	78.51 ± 2.93	78.65 ± 3.05	78.89 ± 2.65	89.53 ± 3.62	87.53 ± 5.27	93.81 ± 4.60	91.09 ± 1.07
Trees	86.17 ± 1.28	88.95 ± 0.53	89.05 ± 1.45	96.72 ± 0.85	99.03 ± 0.33	99.35 ± 0.28	99.57 ± 0.18
Metal	94.11 ± 2.36	93.23 ± 1.29	94.55 ± 0.67	97.94 ± 0.99	98.56 ± 1.36	99.67 ± 0.17	99.81 ± 0.13
Bare Soil	90.11 ± 0.62	90.13 ± 0.54	90.23 ± 1.23	96.34 ± 1.43	98.35 ± 1.06	98.39 ± 0.18	99.05 ± 0.62
Bitumen	82.13 ± 2.52	81.66 ± 3.31	83.69 ± 2.82	98.73 ± 1.12	99.29 ± 0.51	92.31 ± 0.88	99.41 ± 0.39
Bricks	82.38 ± 2.57	83.05 ± 1.61	83.57 ± 2.91	95.01 ± 1.05	94.61 ± 0.50	89.44 ± 4.66	89.96 ± 3.07
Shadows	93.29 ± 0.55	95.26 ± 0.56	94.68 ± 0.46	97.39 ± 0.97	99.39 ± 0.58	98.96 ± 0.56	98.42 ± 1.45
OA / %	90.65 ± 1.32	91.07 ± 0.85	91.32 ± 1.22	93.88 ± 2.28	96.49 ± 1.78	97.05 ± 0.45	97.49 ± 0.24
AA / %	87.11 ± 2.09	87.57 ± 1.38	88.28 ± 1.54	95.82 ± 1.56	96.61 ± 1.68	96.27 ± 1.36	96.98 ± 1.57
100K	88.92 ± 1.46	88.72 ± 1.44	89.43 ± 1.09	92.93 ± 1.94	95.31 ± 2.41	96.22 ± 0.11	96.67 ± 0.32

表 4 Kennedy Space Center 数据集的分类结果
Table 4 Classification results of Kennedy Space Center dataset

Class	RBF-SVM	EMP-SVM	DCNN	GAN	ResNet	PyResNet	DTAResNet
Scrub	92.11±2.32	93.42±1.38	92.48±2.21	95.29±0.45	96.00±1.75	96.24±0.43	96.63±2.65
Willow	82.25±1.52	83.66±2.65	83.96±2.82	80.83±2.42	92.82±2.80	93.24±0.54	97.68±0.49
Palm	85.37±3.64	85.69±3.21	73.52±4.98	89.43±1.96	88.67±2.91	87.62±1.72	89.98±0.76
Pine	60.08±4.33	62.43±5.63	60.22±8.52	83.02±2.43	76.80±3.46	86.01±1.91	86.02±0.64
Broadleaf	62.56±5.98	64.42±4.65	61.09±5.90	96.15±1.53	78.50±1.63	77.55±2.67	76.15±1.53
Hardwood	65.38±3.45	67.65±5.10	64.24±5.24	95.82±1.91	79.43±2.93	86.15±1.24	96.80±0.30
Swap	63.52±5.66	65.56±6.36	65.95±7.24	95.16±0.85	75.88±2.90	84.70±1.68	82.68±1.85
Graminoid	71.52±3.84	73.28±2.98	73.60±3.22	75.78±2.10	96.10±1.75	96.17±0.39	96.52±0.13
Spartina	81.56±4.55	85.32±3.55	86.94±2.84	95.23±2.24	93.93±3.30	94.28±0.54	96.82±2.24
Cattail	90.78±1.84	93.25±1.22	93.52±0.98	98.98±0.38	96.77±1.34	99.30±0.05	99.48±0.38
Salt	93.65±1.46	95.38±2.01	95.91±0.55	96.06±0.93	99.51±0.48	99.54±0.06	99.98±0.01
Mud	90.35±2.19	91.01±2.58	89.39±1.20	96.37±1.32	97.09±0.95	96.30±0.47	92.44±1.91
Water	99.26±0.24	99.31±0.32	99.84±0.04	99.09±0.14	99.65±0.05	99.28±0.18	99.85±0.33
OA / %	80.65±3.08	81.97±3.27	81.04±2.90	93.56±0.98	89.93±2.54	94.01±1.08	94.93±1.83
AA / %	79.87±3.16	81.57±3.20	80.05±3.52	92.09±1.44	90.11±2.02	92.03±0.91	93.16±1.00
100K	79.93±3.45	80.78±2.96	79.67±3.78	92.85±1.56	88.86±4.96	93.16±2.04	94.35±1.98

表 5 Salinas 数据集的分类结果

Table 5 Classification results of Salinas dataset

Class	RBF-SVM	EMP-SVM	DCNN	GAN	ResNet	PyResNet	DTAResNet
Brocoli_green_weeds_1	83.42±1.56	96.16±0.28	97.75±0.22	85.71±1.23	99.25±0.05	99.90±0.06	99.95±0.02
Brocoli_green_weeds_2	92.19±0.85	99.27±0.02	97.19±0.14	87.50±1.05	99.35±0.20	99.95±0.03	99.99±0.01
Fallow	90.66±1.94	80.45±1.78	78.39±0.12	80.46±2.13	99.35±0.48	99.03±0.03	99.97±0.02
Fallow_rough_plow	95.33±1.92	98.34±0.35	99.07±0.08	99.90±0.01	97.49±1.80	99.24±0.05	98.93±1.38
Fallow_smooth	86.52±2.64	94.56±1.32	98.84±0.23	90.35±1.06	99.32±0.43	99.61±0.01	99.71±0.23
Stubble	97.62±0.25	99.45±0.09	99.86±0.55	99.34±0.33	99.46±0.02	99.76±0.06	99.98±0.01
Celery	95.22±0.74	97.36±0.65	99.09±0.29	99.90±0.03	99.38±0.01	99.96±0.01	99.99±0.01
Grapes_untrained	84.35±2.61	84.38±0.46	94.96±1.23	90.95±1.08	96.39±0.47	95.45±1.28	96.61±1.23
Soil_vinyard_develop	98.45±0.45	98.74±1.06	99.69±0.02	99.92±0.03	99.47±0.02	99.88±0.02	99.97±0.02
Corn_senesced_green_weeds	80.55±4.69	91.98±0.91	99.25±0.15	98.23±0.65	99.70±0.05	99.83±0.01	99.72±0.15
Lettuce_roumaine_4wk	85.55±3.13	90.73±3.22	92.19±1.68	97.56±0.84	98.71±1.25	99.15±0.09	99.23±1.08
Lettuce_roumaine_5wk	96.04±0.69	99.79±0.10	99.89±0.19	99.04±0.42	99.52±0.24	99.80±0.01	99.90±0.07
Lettuce_roumaine_6wk	98.54±0.43	98.22±0.65	94.78±2.09	80.62±2.56	99.68±0.15	99.06±0.32	99.63±0.29
Lettuce_roumaine_7wk	86.33±1.88	96.23±0.24	95.85±0.67	80.21±2.91	99.45±0.54	99.70±0.24	99.33±0.37
Vinyard_untrained	66.78±6.66	63.89±6.32	93.85±1.89	79.99±0.59	94.04±1.52	94.79±2.17	95.58±1.89
Vinyard_vertical_trellis	83.72±4.84	79.78±5.12	99.95±0.02	91.08±0.72	99.20±0.68	98.68±0.33	99.92±0.02
OA / %	87.79±2.65	89.74±1.43	96.38±0.76	95.00±1.45	98.25±0.28	98.18±0.38	98.65±0.13
AA / %	88.83±2.21	91.83±1.41	96.29±0.60	91.30±0.98	98.73±0.48	98.99±0.29	99.30±0.40
100K	88.93±1.97	88.96±2.05	95.66±0.93	93.31±1.36	98.05±0.31	97.97±0.04	98.17±0.23

图 9~11 分别为高光谱数据的真值图和各种分类算法的假彩色图。传统的 RBF-SVM 和 EMP-

SVM 算法分类结果的噪点较多,如图 9 中部分交通用地误分为草地,地物的分类精度相对较低。

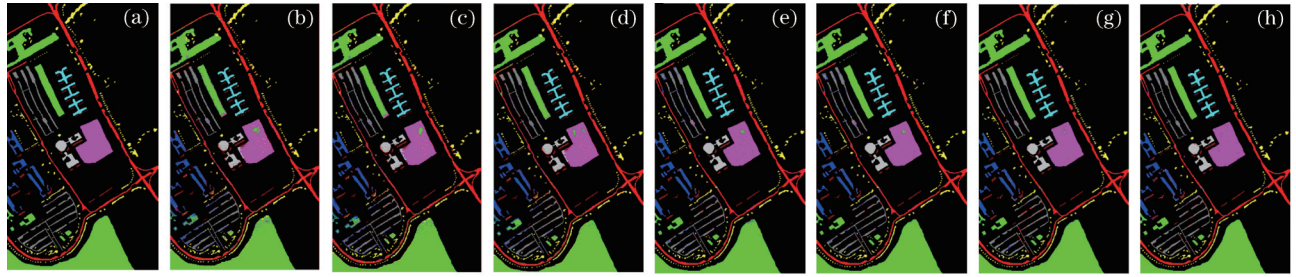


图 9 Pavia University 数据集的分类结果图。(a)真值图; (b) RBF-SVM; (c) EMP-SVM; (d) DCNN; (e) GAN; (f) ResNet; (g) PyResNet; (h) DTAResNet

Fig. 9 Classification results of Pavia University dataset. (a) Ground truth; (b) RBF-SVM; (c) EMP-SVM; (d) DCNN; (e) GAN; (f) ResNet; (g) PyResNet; (h) DTAResNet

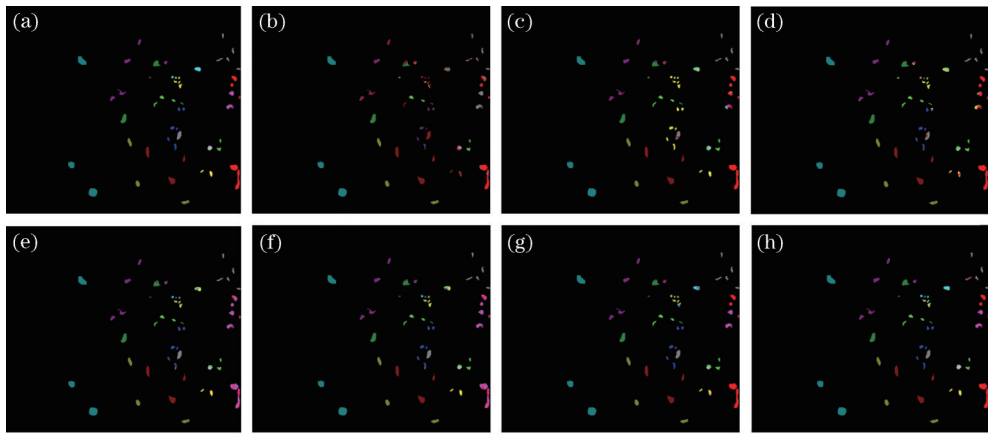


图 10 Kennedy Space Center 数据集的分类结果图。(a)真值图; (b) RBF-SVM; (c) EMP-SVM; (d) DCNN; (e) GAN; (f) ResNet; (g) PyResNet; (h) DTAResNet

Fig. 10 Classification results of Kennedy Space Center dataset. (a) Ground truth; (b) RBF-SVM; (c) EMP-SVM; (d) DCNN; (e) GAN; (f) ResNet; (g) PyResNet; (h) DTAResNet

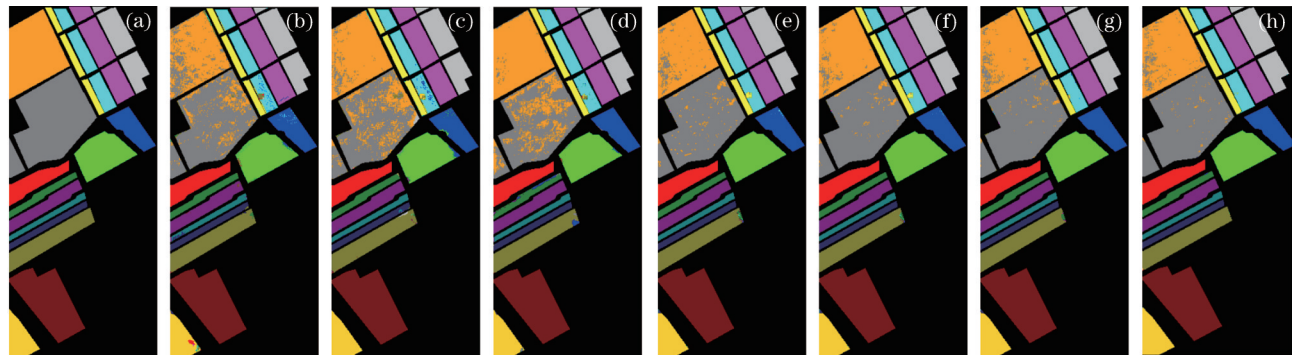


图 11 Salinas 数据集的分类结果图。(a)真值图; (b) RBF-SVM; (c) EMP-SVM; (d) DCNN; (e) GAN; (f) ResNet; (g) PyResNet; (h) DTAResNet

Fig. 11 Classification results of Salinas dataset. (a) Ground truth; (b) RBF-SVM; (c) EMP-SVM; (d) DCNN; (e) GAN; (f) ResNet; (g) PyResNet; (h) DTAResNet

ResNet 和 PyResNet 的分类效果与 SVM、DCNN 和 GAN 相比有所改善,但是仍有部分误分现象,而所提算法更加充分地利用了各卷积层及特征图,分类效果有了更大的提升,也消除了块状的误分类现象,保护了边缘信息。同时,从表 3~5 可以看到,所

提算法对大部分地物种类的分类精度达到了 96% 以上,有不少类甚至高达 99%。实验结果表明,所提算法在 3 种数据集上均可以有效提取更精细化特征,跨维度信息交互关注更重要特征,从而提高分类精度。

表 6 以 Pavia university 数据集为例统计了不同网络模型的训练时间。从表 6 可以看出: SVM 的训练时间虽短, 但精度不高; 而 ResNet 和 PyResNet 由于添加了短连接和金字塔卷积, 训练时间要比

DCNN 更长一些; DTAResNet 中的动态卷积虽然引入了少量额外的计算开销, 但是显著增加了特征表达能力, 因此时间上轻微的劣势是可以接受的, 也表明了所提算法的有效性。

表 6 不同分类算法训练时间

Table 6 Training time of different classification algorithms

unit: min

Algorithm	RBF-SVM	EMP-SVM	DCNN	GAN	ResNet	PyResNet	DTAResNet
Time	5.63	9.01	12.85	10.39	17.37	17.65	18.57

4 结 论

提出了一种结合动态卷积和三重注意力机制的高光谱分类算法。Pavia University、Kennedy Space Center 和 Salinas 数据集上进行的一系列实验结果表明, 加入了动态卷积和三重注意力机制的 DTAResNet 能够提高各数据集的总体精度、平均精度、Kappa 系数以及大部分类别的精度, Salinas 数据集中有两类更是高达 99.99%, 验证了所提算法在高光谱图像分类中的良好分类效果。虽然所提算法对高光谱图像分类效果较好, 精度有所提升, 但是计算成本仍然较大, 接下来将进一步优化模型, 增强残差模块对卷积特征映射的识别能力。

参 考 文 献

- [1] Bhosle K, Musande V. Evaluation of deep learning CNN model for land use land cover classification and crop identification using hyperspectral remote sensing images[J]. Journal of the Indian Society of Remote Sensing, 2019, 47(11): 1949-1958.
- [2] 张兵. 高光谱图像处理与信息提取前沿[J]. 遥感学报, 2016, 20(5): 1062-1090.
Zhang B. Advancement of hyperspectral image processing and information extraction[J]. Journal of Remote Sensing, 2016, 20(5): 1062-1090.
- [3] Liu F S, Wang Q. A sparse tensor-based classification method of hyperspectral image[J]. Signal Processing, 2020, 168: 107361.
- [4] Zhong Z L, Li J, Ma L F, et al. Deep residual networks for hyperspectral image classification[C]// 2017 IEEE International Geoscience and Remote Sensing Symposium, July 23-28, 2017, Fort Worth, TX, USA. New York: IEEE Press, 2017: 1824-1827.
- [5] 王爱丽, 张宇泉, 吴海滨, 等. 基于空洞卷积胶囊网络的激光雷达数据分类[J]. 中国激光, 2021, 48(11): 1110003.
Wang A L, Zhang Y X, Wu H B, et al. LiDAR data

classification based on dilated convolution capsule network[J]. Chinese Journal of Lasers, 2021, 48(11): 1110003.

- [6] 李勇, 金秋雨, 赵怀慈, 等. 基于改进残差密集网络的高光谱重建[J]. 光学学报, 2021, 41(7): 0730001.
Li Y, Jin Q Y, Zhao H C, et al. Hyperspectral image reconstruction based on improved residual dense network[J]. Acta Optica Sinica, 2021, 41(7): 0730001.
- [7] 张怡卓, 徐苗苗, 王小虎, 等. 残差网络分层融合的高光谱地物分类[J]. 光谱学与光谱分析, 2019, 39(11): 3501-3507.
Zhang Y Z, Xu M M, Wang X H, et al. Hyperspectral image classification based on hierarchical fusion of residual networks[J]. Spectroscopy and Spectral Analysis, 2019, 39(11): 3501-3507.
- [8] 郑姗姗, 刘文, 单锐, 等. 一种改进多尺度三维残差网络的高光谱图像分类方法[J]. 计算机工程, 2020, 46(12): 215-221.
Zheng S S, Liu W, Shan R, et al. A hyperspectral image classification method based on improved multi-scale three-dimensional residual network[J]. Computer Engineering, 2020, 46(12): 215-221.
- [9] 韩彦岭, 崔鹏霞, 杨树瑚, 等. 基于残差网络特征融合的高光谱图像分类[J]. 国土资源遥感, 2021, 33(2): 11-19.
Han Y L, Cui P X, Yang S H, et al. Classification of hyperspectral image based on feature fusion of residual network[J]. Remote Sensing for Land & Resources, 2021, 33(2): 11-19.
- [10] Chen Y P, Dai X Y, Liu M C, et al. Dynamic convolution: attention over convolution kernels[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11027-11036.
- [11] Wei H, Wang Z L, Hua G X. Dynamically Mixed Group Convolution to Lighten Convolution Operation [C]// 2021 4th International Conference on Artificial

- Intelligence and Big Data, May 28-31, 2021, Chengdu, China: IEEE Press, 2021: 203-206.
- [12] Zhu Y, Cheng T Y, Tang H, et al. DANet: dynamic salient object detection networks leveraging auxiliary information[J]. IEEE Access, 2021, 9: 92070-92082.
- [13] Hang R L, Li Z, Liu Q S, et al. Hyperspectral image classification with attention-aided CNNs[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(3): 2281-2293.
- [14] 魏祥坡, 余旭初, 管凌霄. 利用残差通道注意力网络的高光谱图像分类[J]. 测绘科学技术学报, 2019, 36(2): 161-166, 172.
Wei X P, Yu X C, Guan L X. Hyperspectral image classification using residual channel attention network [J]. Journal of Geomatics Science and Technology, 2019, 36(2): 161-166, 172.
- [15] Wang L, Peng J T, Sun W W. Spatial-spectral squeeze-and-excitation residual network for hyperspectral image classification[J]. Remote Sensing, 2019, 11(7): 884.
- [16] Ma W P, Yang Q F, Wu Y, et al. Double-branch multi-attention mechanism network for hyperspectral image classification[J]. Remote Sensing, 2019, 11(11): 1307.
- [17] 张祥东, 王腾军, 朱劭俊, 等. 基于扩张卷积注意力神经网络的高光谱图像分类[J]. 光学学报, 2021, 41(3): 0310001.
Zhang X D, Wang T J, Zhu S J, et al. Hyperspectral image classification based on dilated convolutional attention neural network[J]. Acta Optica Sinica, 2021, 41(3): 0310001.
- [18] 王欣, 樊彦国. 基于改进 DenseNet 联合空谱注意力机制的高光谱图像分类[J]. 激光与光电子学进展, 2022, 59(2): 0210014.
Wang X, Fan Y G. Hyperspectral image classification based on modified DenseNet joint spatial spectrum attention mechanism[J]. Laser & Optoelectronics Progress, 2022, 59(2): 0210014.
- [19] Misra D, Nalamada T, Arasanipalai A U, et al. Rotate to attend: convolutional triplet attention module [C]//2021 IEEE Winter Conference on Applications of Computer Vision, January 3-8, 2021, Waikoloa, HI, USA. New York: IEEE Press, 2021: 3138-3147.
- [20] Cui Y, Yu Z K, Han J C, et al. Dual-triple attention network for hyperspectral image classification using limited training samples[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 5504705.
- [21] Wang F, Tax D M J. Survey on the attention based RNN model and its applications in computer vision [EB/OL]. (2016-01-25) [2021-06-02]. <https://arxiv.org/abs/1601.06823>.
- [22] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines [J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(8): 1778-1790.
- [23] Gu Y F, Liu T Z, Jia X P, et al. Nonlinear multiple kernel learning with multiple-structure-element extended morphological profiles for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(6): 3235-3247.
- [24] 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.
- [25] Zhu L, Chen Y S, Ghamisi P, et al. Generative adversarial networks for hyperspectral image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(9): 5046-5063.
- [26] Song W W, Li S T, Fang L Y, et al. Hyperspectral image classification with deep feature fusion network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(6): 3173-3184.
- [27] Paoletti M E, Haut J M, Fernandez-Beltran R, et al. Deep pyramidal residual networks for spectral-patial hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(2): 740-754.