计算机科学与探索
Journal of Frontiers of Computer Science and Technology
ISSN 1673-9418,CN 11-5602/TP

《计算机科学与探索》网络首发论文

题目：　快速 3D-CNN 结合深度可分卷积对高光谱图像分类
作者：　王燕，梁琦
网络首发日期：　2021-04-29
引用格式：　王燕，梁琦．快速 3D-CNN 结合深度可分卷积对高光谱图像分类．计算机科学与探索. https://kns.cnki.net/kcms/detail/11.5602.TP.20210429.1136.004.html

网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

网络首发时间：2021-04-29 14:11:52
网络首发地址：https://kns.cnki.net/kcms/detail/11.5602.TP.20210429.1136.004.html

快速 3D-CNN 结合深度可分卷积对高光谱图像分类

王燕+，梁琦

兰州理工大学 计算机与通信学院，兰州 730050

+ 通信作者 E-mail: wangyan@lut.cn

摘 要: 针对卷积神经网络在高光谱图像特征提取和分类的过程中，存在空谱特征提取不充分以及网络层数太多引起的参数量大、计算复杂的问题，提出快速 3D-CNN（3-Dimension convolutional neural networks）结合深度可分离卷积(depthwise separable convolution，DSC)的轻量型卷积模型。该方法首先利用增量主成分分析(incremental Principal component analysis，IPCA)对输入的数据进行降维预处理；其次将输入模型的像素分割成小的重叠的三维小卷积块，在分割的小块上基于中心像素形成地面标签，利用三维核函数进行卷积处理，形成连续的三维特征图，保留空谱特征。用 3D-CNN 同时提取空谱特征，然后在三维卷积中加入深度可分离卷积对空间特征再次提取，丰富空谱特征的同时减少参数量，从而减少计算时间，分类精度也有所提高。所提模型在 Indian Pines、Salinas Scene 和 University of Pavia 公开数据集上验证，并且同其它经典的分类方法进行比较。实验结果表明，该方法不仅能大幅度节省可学习的参数，降低模型复杂度，而且表现出较好的分类性能，其中总体精度(overall accuracy，OA)、平均分类精度(average accuracy，AA)和 Kappa 系数均可达 99%以上。

关键词: 高光谱图像分类；空谱特征提取；3D-CNN；深度可分离卷积；深度学习

文献标志码： A **中图分类号：** TP751.1

Fast 3d-cnn combined with depth separable convolution for hyperspectral image classification

WANG Yan+, LIANG Qi

School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China

Abstract: Three dimensional convolutional neural network can extract spectral and spatial features at the same time, but due to the problem of spectral dimension and volume, the calculation is complex. In the process of feature extraction and classification of hyperspectral image by convolution neural network, there are some problems, such as insufficient space spectrum feature extraction, too many network layers, large amount of parameters and complex calculation. A fast 3d-cnn combined with depth separable convolution is proposed the lightweight 3D convolution model of convolution (DSC) is used to classify hyperspectral images. Firstly, incremental principal component analysis is used Secondly, the pixels of the input model are divided into small overlapped three-dimensional small convolution blocks, on which the ground label is formed based on the center pixel, and the three-dimensional kernel

基金项目： 基于大数据的“人肉搜索”建模、分析与控制，国家自然科学基金（61863025）；无人机关键技术研究，甘肃省重点研发计划-工业类（18YF1GA060）。

This work was supported by the Modeling, analysis and control of "human flesh search" based on big data Foundation of China (61863025), the Research on Key Technologies of UAV Foundation of China (18YF1GA060).

function is used for convolution processing to form a continuous three-dimensional feature map, while preserving the spatial spectrum features. 3d-cnn is used to extract spatial spectral features of pixels at the same time, and then depth separable convolution is added to 3D convolution to extract spatial features again, which enriches spatial spectral features and reduces the amount of parameters,so as to reduce the operation time and improve the classification accuracy.The proposed model is validated on Indian pines, Salinas scene and University of Pavia public datasets, and compared with other classical classification methods. The experimental results show that this method can not only save learning parameters and reduce the complexity of the model, but also show good classification performance. The overall accuracy OA, average classification accuracy AA and kappa coefficient can reach more than 99%, and the confusion matrix also shows good classification performance.

Key words: hyperspectral image (HSI) classification; spatial spectrum feature extraction; 3-Dimension convolutional neural networks(3D-CNN); depth wise separable convolution; deep learning

高光谱图随着高光谱遥感技术的发展，新的高光谱传感器能够同时采集光谱特征和空间特征的连续图像[1]。高光谱图像含有丰富的地物特征，包括光谱特征和空间特征。因此，在农业、环境监测、城市规划和军事侦察等领域有着广泛的应用[2]。因为高光谱图像具有三维立体图像的属性，因此为了更好的获得地物信息，可以充分利用空间和光谱的联合特征进行图像分类[3]。但是高光谱图像在描述了丰富的地物细节信息的同时，也出现了数据之间的高度相关，引起数据的大量冗余问题[4]，增加了计算的复杂度。因此，减少光谱数据间的冗余量，降低数据维度并且提取高光谱图像的空谱联合特征成为高光谱图象分类中的首要任务[5-6]。

深度学习算法的提出，让高光谱图像的分类方法有了新的进展。深度学习算法被研究者广泛运用到高光谱图像的分类中，取得了良好的研究成果。2014 年，将深度学习网络 SAE 运用到高光谱图像的分类中，并提出了一种融合光谱特征和空间特征的深度学习模型，得到了较高的分类精度[7]，相继越来越多的深度学习模型被研究者应用。2015 年，将深度置信网络(deep belief network，DBN)模型引入到高光谱图像分类中，同时结合了主成分分析的方法(principle component analysis，PCA)对数据进行降维预处理。对特征进行分层学习和采用逻辑回归的方法对高光谱图像的空谱特征进行提取，取得了良好的分类效果[8]。2015 年，第一次将卷积神经网络(CNN)模型应用到高光谱图像分类中，但是建立的 CNN 模型只能进行光谱特征的提取[9]。2016 年提出了一种基于 CNN 的深度特征提取方法，并建立了一种基于三维卷积神经网络的深度有限元模型，以提取高光谱遥感图像的空谱特征，获得较高的分类精度[10]。2017 年，提出了频谱空间残差网络(SSRN)，SSRN 中的残差剩余块使用恒等映射来连接其他的 3D 卷积层，便于梯度的反向传播，同时提取更深层次的光谱特征，缓解了其他深度学习模型存在的精度下降现象[11]。2019 年，通过自适应降维，提出了一种用于频谱空间 HSIC 的半监督三维卷积神经网络(CNN)来解决维数诅咒问题[12]。这些研究成果表明基于深度学习的方法在高光谱图像分类方面取得了一定的成果。但是基于深度模型的方法通常存在过拟合的现象，这是因为在利用深度模型的方法进行训练的时候需要大量带标签的数据，但是高光谱图像带标签的样本不足。因此，为了尽量避免这样的问题，需要合适的卷积模型，既能充分发挥卷积神经网络的巨大优势，又能减少可学习的参数，从而缓解过拟合问题以及对于训练样本数据量的需求。现有的卷积模型较复杂，网络的参数量大，带来计算复杂的问题。需要有更加轻量型的卷积网络来满足计算时间、效率以及内存的要求。

为了解决上述问题，提出了一种新的快速 3D-CNN 结合深度可分离卷积的方法。该方法首先将高光谱立方体图像分割成相同大小的重叠三维小立方体。利用三维卷积核函数对这些小立方体进

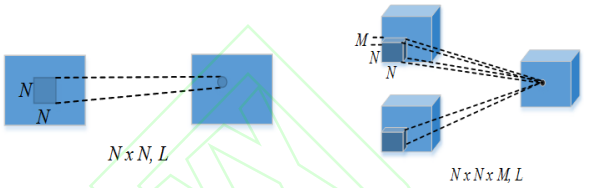
行处理，生成多个连续波段的三维特征图，以保留空间和光谱的联合信息，为特征提取过程提供丰富的空谱信息。使用增量主成分分析(IPCA)对数据进行预处理，获得比较重要的波段信息。然后将预处理的数据输入到 3D-CNN 模型中，在三维卷积对空谱特征同时提取后，加入深度可分离卷积对空间特征再次进行提取以获得更加丰富的空间特征信息，最后采用 Soft-max 分类器进行分类。相对于其他 2D/3D-CNN 模型，所提模型涉及的参数更少，同时运算速率也相对较高，可以获得较高的分类精度。本文的模型实验对比了几种典型的 HSIC 方法。最后的实验和比较结果表明，本文所提模型方法性能优于其他比较的方法。

1 基础理论

1.1 三维卷积神经网络

早期 CNN 模型的建立是针对二维数据设计的，可以对二维形状进行很好的识别，在目标识别[13]、图像分割[14]等方面有很好的应用效果，但是对三维立体形状进行识别分析时就有一定的局限性。研究的高光谱图像不同于一般的二维图像，它是由空间维和光谱维共同构成，形成三维的数据立方体。在空间维也包含大量的信息，二维卷积神经网络(2D-CNN)对高光谱图像的空间特征可以较好的提取，但是不利于同时提取像元的光谱和空间特征[15]。利用卷积神经网络对高光谱图像进行分类时，首先要对数据进行降维、去噪等预处理操作，但是降维操作可能会引起高光谱图像中细节信息的丢失，对高光谱图像分类造成影响。

针对二维卷积神经网络对三维高光谱数据的信息利用不足的问题，可引入三维卷积神经网络对高光谱图像的空谱特征同时提取。三维卷积神经网络(3D-CNN)与二维卷积神经网络(2D-CNN)的网络结构非常类似，两类结构都是由基本的卷积层和池化层组成。关键的不同之处在于 3D-CNN 结构采用的是 3D 卷积核来对图像进行卷积操作。如图 1 所示 2D-CNN 和 3D-CNN 卷积操作的示例，N×N 代表卷积核大小，三维比二维多了 *M* 代表的光谱维度大小，*L* 为卷积层输出通道。3D-CNN 在空间维和光谱维上同时进行运算，从而同时提取图像的空谱特征。不会单独的提取某一类特征引起特征提取不足，导致分类结果不理想。



(a) 2D-CNN　　(b) 3D-CNN

Fig.1　The algorithm 2D-CNN and 3D-CNN

图 1　2D-CNN 和 3D-CNN 算法

其中，三维卷积神经网络的卷积核在长、宽和通道三个方向上移动，计算神经网络第 *i* 层第 *j* 个特征图在(x,y,z)处的点值 *V* 计算公式如(1)式所示：

$$V_{i,j}^{xyz} = f\left(\sum_{m}\sum_{l=0}^{Li-1}\sum_{w=0}^{Wi-1}\sum_{h=0}^{Hi-1} W_{i,j,m}^{lwh} \cdot V_{i-1}^{(x+l)(y+w)(z+h)} + b_{i,j}\right) \quad (1)$$

式(1)中：*m* 表示第 *i-1* 层中与当前特征图相连的特征图；与表示卷积核的长度和宽度，表示卷积核在光谱维度上的尺寸；*W* 代表与 *i-1* 层相连的第 *m* 个特征图的连接权值；*b* 表示第 *i* 层第 *j* 个特征图的偏置，*f* 为激活函数。

1.2 深度可分离卷积

深度可分离卷积[16](DSC)是普通二维卷积神经网络的一种变换形式，可以替换普通二维卷积神经网络[17]。核心思想就是将普通 *N* 个通道为 *M* 的卷积拆分成 1 个通道为 *M* 的卷积，该卷积进行单通道滤波操作，区别于普通卷积滤波后通道相加和 *N* 个 1×1×M 的卷积。如图 2(a)所示，为普通的卷积神经网络，由卷积层，批归一化操作和激活函数构成的。图 2(b)所示为深度可分离卷积，它是由一个 3×3 卷积核大小的深度可分离的卷积层，批归一化，激活函数和 1×1 卷积核大小的卷积层，批归一化，激活函数构成。即分为深度卷积(depthwise convolution)和逐点卷积(pointwise convolution)两部分[18]。

在对多个输入通道执行常规 2D 卷积中，卷积核的通道数与输入的通道数一致。

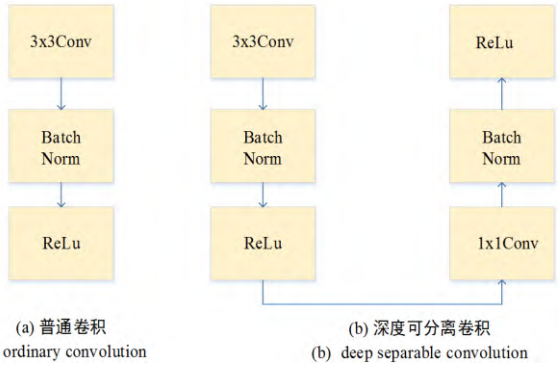


Fig.2 (a) The difference between ordinary convolution and (b) deep separable convolution

图 2 (a)普通卷积和(b)深度可分离卷积的区别

混合所有通道来产生最后的一个输出，深度卷积将输入特征映射的每个通道分别卷积，捕获每个通道的空间特征。逐点卷积集成所有提取的空间特征并学习输入特征映射的信道相关信息，对获得的特征图进行类似于普通卷积的通道融合操作。可以在损失精度不多的情况下降低参数量和计算量。

2 本文模型

2.1 模型介绍

通过上述分析，本文所提出模型的网络架构如图 3 所示。其中包括一个输入层、三个三维卷积层、两个深度可分离卷积层，分别进行深度卷积和逐点卷积、全连接层包括一个 Flatten 平滑层、两个 Dense 层以及两个 Dropout 层，防止过拟合，卷积层均采用 ReLU 激活函数进行非线性映射。ReLU 激活函数比传统的 Sigmoid 函数和 Tanh 函数收敛速度更快。ReLU 激活函数的形式为：

$$f(x) = \max(0, x) \qquad (2)$$

最后通过 Soft-max 分类器进行高光谱图像特征的分类。每个层的输入输出维度、参数大小如表 1 所示。

模型中 3D-CNN 卷积核大小为：3D_conv_layer1 = 8×3×3×7×1，其中 $K_1^1 = 3$，$K_2^1 = 3$，$K_3^1 = 7$。3D_conv_layer2 = 16×3×3×5×8，其中，$K_1^2 = 3$，$K_2^2 = 3$，$K_3^2 = 5$。3D_conv_layer3 = 32×3×3×3×16，其中，$K_1^3 = 3$，$K_2^3 = 3$，$K_3^3 = 3$，最后将三维的输出特征 reshape 成二维数据，提取高光谱图像的空间特征，添加两个深度可分离卷积层 Separable_conv2d_1_layer4 = 3×3×64，Separable_conv2d_1_layer5 = 1×1×128。为了增加空间光谱特征图的数量，在平化层之前部署了 3 个三维卷积层，高光谱图像的空间信息确定空间维上相邻像素间的空域特征，并且空域特征可以弥补谱域特征的不足，应用空域特征增加光谱空间的特征，提升高光谱图像的分类精确度。因此在三维卷积层后增加两个深度可分离卷积层，在可以减少参数的同时增加空间特征，提取更丰富的空谱特征，确保模型能够在没有损失的情况下区分不同波段的空间信息。提出的快速 3D-CNN 和 DSC 组合模型的总参数(即可调权值)为 377408，比单独使用快速 3D-CNN 的参数少了约一半多。卷积的填充方式选择零填充，不需要进行批处理归一化和数据增强。

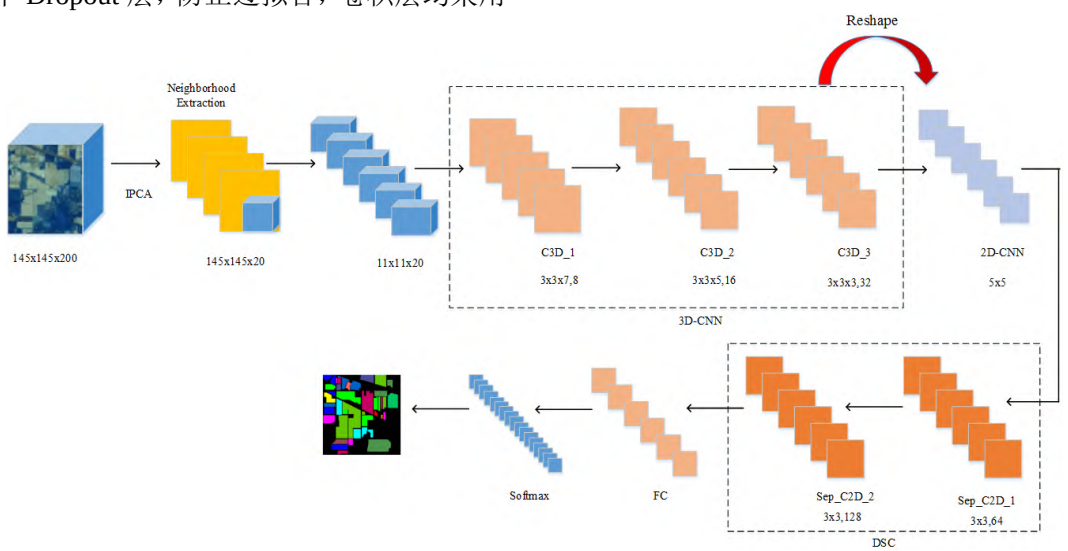


Fig.3 The Network framework

图 3 网络框架

Table 1 Parameters of model on IP dataset with Window Size of 11×11

表 1 模型在 Window Size 大小为 11x11 的 IP 数据集上的参数

Layer(type)	Output Shape	Parameter#
input_1(Input Layer)	(11,11,20,1)	0
Conv3d_1(Conv3D)	(9,9,14,8)	512
Conv3d_2(Conv3D)	(7,7,10,16)	5776
Conv3d_3(Conv3D)	(5,5,8,32)	13856
Reshape_1(Reshape)	(5,5,256)	0
Separable_conv2d_1(separable)	(3,3,64)	18752
Separable_conv2d_1(separable)	(3,3,128)	8384
Flatten_1(Flatten)	(128)	0
Dense_1(Dense)	(256)	295168
Dropout_1(Dropout)	(256)	0
Dense_2(Dense)	(128)	32896
Dropout_1(Dropout)	(128)	0
In total,377408 trainable parameters are required		
Train on 4304 sample,validate on 1845 samples		

2.2 IPCA 降维与数据预处理

高光谱图像的像素表现出高类间相似性、高类内变异性，这样的问题对于任何分类模型都需要密集的处理。为了克服上述问题，需要对数据进行预处理操作。原始的 PCA 方法会把所有的数据一次性的放入内存中，这在大数据集的情况下有可能会遇到问题，因此提出使用增量主成分分析(IPCA)对数据进行预处理。增量主成分分析能够避免直接计算协方差矩阵,并且不需要一次性获得全部图像数据,它采用增量的学习方式,使用每次新获得的图像数据逐步对主元估计值进行迭代更新，训练集可以分成一个个小批量，一次给 IPCA 算法输入一个。

假设输入向量序列为 $u'(t)$=1，2，...，第 n 幅图像输入时均值为 $m(n)=\frac{1}{n}\sum_{t=1}^{n}u'(t)$,协方差矩阵为:

$$A(n)=\frac{1}{n}\sum_{t=1}^{n}\left[u'(t)-m(n)\right]\left[u'(t)-m(n)\right]^{T}=\frac{1}{n}\sum_{i=1}^{n}u(t)u(t)^{T} \quad (3)$$

这里 $u(t)=u'(t)-m(n)$，u(n)的第 i 个特征值和特征向量的计算公式为 $\lambda_i x_i(n)=A(n)x_i[n]$，其中 $x_i(n)$ 为第 n 输入时待求的第 i 个特征向量，λ_i 为对应的特征值。IPCA 算法为了加快迭代的速度，整个迭代是对特征值和特征向量的乘积 $\lambda_i x_i$ 进行的。

IPCA 减少了减少空谱带间的冗余，获得比较重要的波段信息，同时保持空间维度的完整性。在图 3 中，IPCA 将光谱波段从 200 减小到 20，而空间维数保持不变。同时三维立方体分割成小的重叠的三维空间小立方块，传递给模型进行特征提取，并在这些小立方体块上基于中心像素形成地面标签。

2.3 3D 卷积结合深度可分离的卷积过程

图 4 为三维空间光谱特征学习框架。该部分由三个三维卷积层、ReLU 激活函数组成，对高光谱图像的光谱和空间特征同时提取。网络的输入数据大小为 11×11×20，第一层卷积核的大小为 3×3×8，经过两层三维卷积操作后，输出为 32 个 5×5×8 大小的特征映射。在完成 3D 卷积操作后，再次进行空间特征提取，利用 reshape 进行 3D 到 2D 变形。为了学习后期二维空间的输出特征，将三维特征重构为大小为 5×5 的 32 个二维特征图，只需要研究二维空间特征，与三维卷积相比减少了网络参数和操作成本。

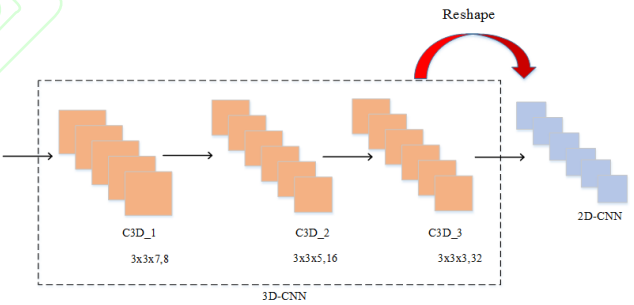


Fig.4 3D spectral-spatial feature learning

图 4 3D 光谱-空间特征学习

图 5 为基于深度可分卷积的二维空间特征学习。采用深度可分离卷积对输出的二维特征进行提取，在不引入额外参数的情况下更好的提取空间特征。与传统的二维卷积不同的是，深度可分卷积在保持信道独立的情况下进行空间卷积，然后进行深度卷积。通过特征重塑后，网络输入数据为 256 个特征映射，大小为 5×5。SeparableConv2D 实现整个深度分离卷积过程，即深度方向的空间卷积和输出通道混合在一起的逐点卷积。第一步是深度卷积运算，用 64 个 3×3 大小的卷积核对输入数据进行卷积运算，得到了 64 个通道的特征图，每个卷积核只对输入层的一个通道进行卷积。第二步是逐像素点

卷积运算。用 128 个 1×1 大小的卷积核在这 64 个特征图上进行卷积运算，将不同通道的信息进行融合。经过 1×1 的卷积，显著地减小了尺寸深度。经上层输出的 64 个通道，1×1 卷积将把这些通道嵌入到单个通道中。1×1 卷积之后，添加 Batch Normalization 批量正则化提高模型泛化能力，添加非线性激活函数 ReLu，非线性允许网络学习更复杂的功能。同时每个卷积核与输入图像进行卷积，得到一个空间特征映射。深度可分卷积不仅减少了网络中的参数量和计算量，而且提高了网络训练速度，降低了 HSI 分类中过拟合的几率。使用填充来保证输出特征映射的大小与输入的大小相同。

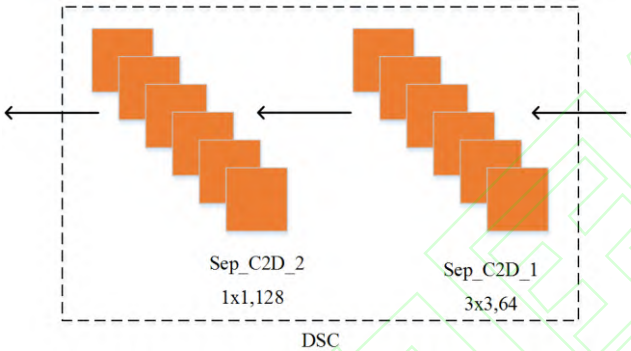


Fig.5 DSC spatial feature learning

图 5 DSC 空间特征学习

3 实验及结果

3.1 实验数据集及实验描述

3.1.1 高光谱数据集

Indian Pines(IP)是由 AVIRIS 传感器收集的印第安纳州西北部印第安纳农林高光谱试验地的图像，该图像由 145×145 个像素组成，其中 220 个光谱带，范围为 0.2~0.4m，空间分辨率很好。将其中 20 个噪声波段移除，该数据集共 16 类地物。

Pavia University(PU)帕维亚大学数据集在意大利北部的帕维亚收集使用反射光学系统成像光谱仪(ROSIS)光学传感器。PU 数据集由 610×610 空间和 103 个光谱波段组成，该数据集包含 9 类地物。

Salinas(SA)数据集由机载可视红外成像光谱仪(AVIRIS)对美国加利福尼亚州的 Salinas 山谷进行成像，空间分辨率为 3.7m 其包含 512×217 个地物像素和 224 个光谱通道，共有 200 个光谱通道可用于分类，该数据集包含 16 类地物。图 6 所示是三个数据集的伪彩色图。

(a)IP (b)PU (c)SA

Fig. 6 Pseudo color maps of three datasets

图 6 三个数据集的伪彩色图

3.1.2 实验细节描述

在实验过程中经过不断的测试调整，将 batch size 设为 256，epoch 迭代次数设置为 50，采用 Adam 优化器对网络进行学习训练，初始学习率设置为 0.001(同时设置 decay=1e-06 每次参数更新后学习率衰减)进行实验。采用 ReLu 函数作为激活函数，以提高计算效率并加快函数收敛速度。分别随机选取训练数据量为 60%，测试数据量为 40%的 IP、PU 和 SA 数据集进行实验。为了实验的公平，对不同的数据集在输入体积的三维 patch 中提取了相同的空间维度，例如 IP、SA 和 UP 的空间维度均为 11×11×20。采用 OA、AA 和 Kappa(K)系数和混淆矩阵来评价分类性能。其中 OA 是用来评价所有样本的分类正确率，AA 是每一类分类的精度，Kappa 系数是一种常用来计算分类精度的方法，代表分类与完全随机的分类产生错误减少的比例。混淆矩阵是分别统计分类模型归错类，归对类的观测值个数，然后把结果放在一个表里展示出来。

3.1.3 实验环境描述

实验环境为 Intel(R) Xeon(R) Silver 4116 CPU @ 2.10GHz ,内存为 128GB 的 PC 机，具体程序由 Pycharm2019 编写，在 Windows10 系统下基于 python3.7 的 Tensorflow2.0 框架实现。在所有的实验中，初始测试集/训练集按 40%/60%的比例划分，将训练样本(占总样本的 60%)再按 30%/70%的比例划分为训练集和验证集。

3.2 实验结果及分析

3.2.1 不同降维结果下的分类效果

维度变换是将现有数据降低到更小的维度，尽量保证数据信息的完整性。一般情况下，降维会造

成原始数据信息的损失，降维后的数据要尽可能的保留原始数据的信息，本文利用增量主成分分析对数据进行降维，通过给定需要原始数据信息的百分比来确定选择前 75 个主成分，降到 20 维，但是如果需要保留更多的原始数据的信息就会造成降维后的维数仍然很多，导致分类效果不明显。表 2 基于 IP 数据集进行实验分析，在其他的超参数均不变化的情况下，将降维参数 numComponents 作为单一变量，降维到不同大小的维度时对分类效果的影响。从表中可以看到在降维到 20 的情况下，Kappa、OA、AA 三个分类精度指标均最高，因此在保证其他参数不变的情况下，本文利用 IPCA 将数据降维到 20 维是最适合的情况，同时在实验过程中发现 IPCA 对数据的预处理速度较快。

Table 2　Classification accuracy of different dimensionality reduction based on IP dataset

表 2　基于 IP 数据集的不同降维大小下的分类精度

NumComponents	15	20	25	30
Kappa×100	99.11	**99.61**	98.77	99.47
OA	99.21	**99.65**	98.92	99.53
AA	97.84	**99.78**	98.91	99.67

3.2.2 不同空间维度大小对分类的影响

在深度卷积神经网络中，输入图像的尺寸越大，模型卷积参数的数量就越大，计算复杂度也会越高。另外，如果输入图像的大小过小，网络接收的可用字段也会过小，无法获得良好的分类结果。表 3 表示不同空间邻域大小对所提模型性能的影响。设置空间维度大小为 9×9×20，11×11×20，13×13×20，17×17×20，23×23×20，25×25×20，在三个数据集上分别通过实验得到不同窗口的训练时间以及 Kappa、OA、AA 的分类精度，训练时间高度依赖于网速和可用内存以及模型参数量。对 OA、AA、K 的精度分析时，可以得出结论，随着空间维度的逐渐增大，IP 数据集基本呈增长趋势，PU 数据集在 23×23 大小的时候出现下降，但整体还是在增长。SA 数据集的分类精度较稳定，当空间输入大小达到 11×11 时，分类精度开始缓慢变化，11×11 的窗口大小在精度和时间上对于 IP、PU、SA 三个数据集是足够的，而在 13×13、17×17、23×23 和 25×25 的空间维度大小下几乎是相同的。通过实验可以看到在空间维度不断增大的过程中，模型的各精度指标会显著增加，但同时参数量增加，计算时间增多。该方法主要是在快速 3D-CNN 模型的基础上做的改进，为了比较的公平性，网络超参数不变，维度大小也选取 11×11×20 的情况。通过比较可以看到模型参数量减少，同时各个分类指标的精度也相对较高，训练时间减少。

Table 3　Impact of spatial window size on our proposed model in three datasets

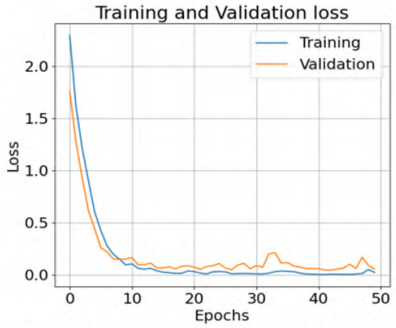
表 3　三个数据集上空间维度大小对模型性能的影响

Window Size	IP				PU				SA			
	K	OA	AA	Tr_time(s)	K	OA	AA	Tr_time(s)	K	OA	AA	Tr_time(s)
9x9	99.08	99.19	98.90	18.65	99.90	99.92	99.89	67.73	99.89	99.90	99.92	199.06
11x11	99.30	99.39	99.60	45.58	99.96	99.97	99.97	91.42	99.95	99.96	99.94	231.19
13x13	99.78	99.74	99.78	36.65	100	100	100	146.50	99.98	99.98	99.99	305.56
17x17	99.66	99.70	99.84	62.16	100	100	100	326.47	99.98	99.99	99.99	408.13
23x23	99.74	99.78	99.76	147.68	98.24	98.67	98.11	538.34	100	100	100	1026.73
25x25	99.66	99.70	99.86	172.65	99.96	99.97	99.88	954.98	99.97	99.97	99.99	1471.46

3.2.3 分类损失率和准确率

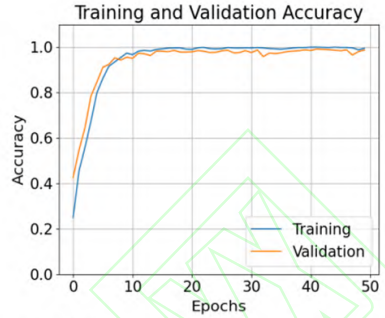
实验通过训练和验证损失率以及准确率来分析网络模型的稳定性以及拟合度。主要在 IP 数据集上进行实验验证。图 7 的曲线表示 IP 数据集在窗口大小为 9×9 的情况下的训练和验证集损失率和准确率分类效果，和 11×11 大小下的曲线图相比较，模型的拟合度不高。文中主要针对 11×11 大小的情况进行训练。图 8 的曲线可以看到，从左边 a 图可以看到模型在 epochs 达到 15 左右的时候就开始收敛，训练集和验证集的损失率已经接近于 0。右边 b 图 epochs 达到 15 左右时也开始收敛，训练集

和验证集达到将近 100%的准确率。从两个图中可以看出训练集和验证集的曲线基本吻合，不存在较大的震荡现象，模型的拟合度相当高。通过实验可以看到所提模型相对稳定，收敛速度非常快，分类精确度较高。



（a）Training and Validation loss

（a）训练和验证损失率

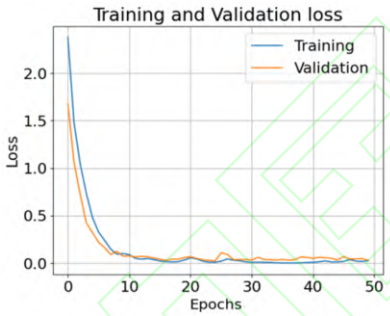


(b)Training and Validation accuracy

(b)训练和验证准确率

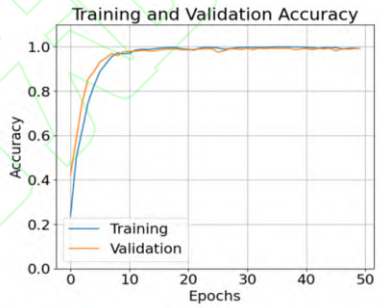
Fig.7 Loss rate (a) and accuracy rate (b) of space size 9 × 9

图 7 空间大小为 9 × 9 的损失率(a)和准确率(b)



（a）Training and Validation loss

（a）训练和验证损失率



(b)Training and Validation accuracy

(b)训练和验证准确率

Fig.8 Loss rate (a) and accuracy rate (b) of space size 11 × 11

图 8 空间大小为 11 × 11 的损失率(a)和准确率(b)

3.2.4 深度可分离卷积的影响

通过比较实验来测试深度可分卷积的影响。在三维卷积层加入深度可分离卷积，形成比较模型，与快速 3D-CNN 的模型进行对比[19]。其他参数设置与 Fast-3D-CNN 一致。IP、PU 和 SA 数据集的初始测试集/训练集按 40%/60%的比例划分，将训练样本(占总样本的 60%)再按 30%/70%的比例划分为训练集和验证集。对不同的数据集提取相同的空间维数 11×11×20。图 9-11 表示所提模型和 Fast-3D-CNN 模型的分类图对比。从图中可以看出在其他条件都相同的情况下，加入深度可分离卷积的 3D-CNN 的分类效果更好。可以看到 Fast-3D-CNN 方法最终的分类结果图中含有大量的斑点，并且存在小区域内错分的情况。而本文方法的分类结果图则含有非常少量的斑点，且在同质的小区域内相对平滑，并且好几类地物几乎完全正确分类。由此可见，该方法具有较好的分类效果。

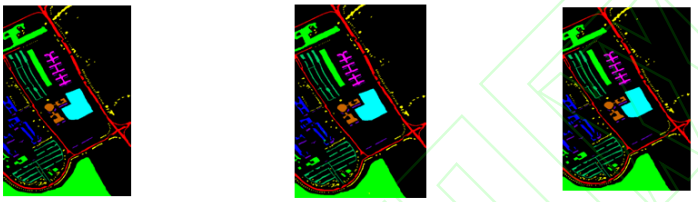
(a)IP Figure datum map
(a)IP 地物基准图

(b)Fast-3D-CNN(98.56%)
(b)快速三维卷积神经网络(98.56%)

(c)Proposed(99.68%)
(c)所提模型(99.68%)

Fig.9 Comparison of classification graphs of two models on IP dataset

图 9 在 IP 数据集上两个模型的分类图对比



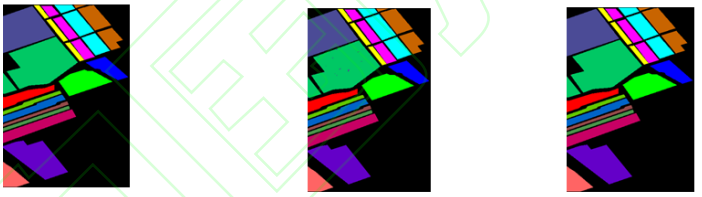
(a)PU Figure datum map
(a)PU 地物基准图

(b)Fast-3D-CNN(99.71%)
(b)快速三维卷积神经网络(99.71%)

(c)Proposed(99.90%)
(c)所提模型(99.90%)

Fig.10 Comparison of classification graphs of two models on PU dataset

图 10 在 PU 数据集上两个模型的分类图对比



(a)SA Figure datum map
(a)SA 地物基准图

(b)Fast-3D-CNN(99.39%)
(b)快速三维卷积神经网络(99.39%)

(c)Proposed(99.98%)
(c)所提模型(99.98%)

Fig.11 Comparison of classification graphs of two models on SA dataset

图 11 在 SA 数据集上两个模型的分类图对比

3.2.5 不同方法下的实验性能比较

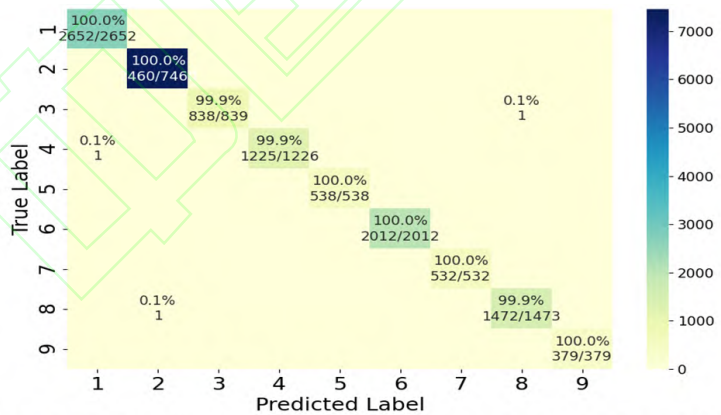
为了验证所提网络模型方法的正确性和有效性，最后将所提卷积网络的方法和传统的卷积模型进行对比，2D-CNN[20]、3D-CNN[21]、Multi-scale-3D-CNN[22]、和 Hybrid SN[23]方法进行对比。为了保证实验的公平性，将所有对比网络中的超参数都设置相同，例如将输入数据降维到 20 维，空间维度大小设置为 11×11×20，epoch 周期为 50，batch size 为 256。同前面实验一样，分别从 Indian Pines、Salinas scene 和 Pavia University 三个数据集上随机选取 60%训练数据，40%测试数据进行验证，并且重复实验 5 次，最后取这 5 次的平均值。表 4 为不同方法下的实验结果。从表中可以看出，提出的方法与其他方法中分类性能最好的方法 Hybrid SN 相比较，对于 Indian Pines 数据集，其 OA 高出 1.86%，AA 高出 2.11%，Kappa 系数高出 2.34。对于 Salinas scene 数据集，其 OA 高出 1.89%，AA 高出 1.16%，Kappa 系数高出 2.09。对于 Pavia University 数据集，其 OA 高出 1.53%，AA 高出 1.9%，Kappa 系数高出 2.02。可以看出在快速 3D-CNN 的基础上结合深

度可离卷积具有较好的分类效果。图 12 表示 PU 数据集的混淆矩阵，可以看到 PU 数据集中大部分地物的分类精度达到 100%，如 Asphalt，Meadows 和 Painted metal sheets 等，只有个别地物被错分，如将 0.1%的 Gravel 错分为 Self-Blocking Bricks,表现出较明显的分类效果。图 13 表示不同方法下高光谱图像分类图的效果，从效果图可以看到所提方法的优势。由此可见，所提及的方法，结合深度可分卷积的快速 3D-CNN 模型对高光谱图像的分类具有较好的分类效果。可以看到错分的类很少，在混淆矩阵中表现明显。

Table 4 Comparison of experimental performance under different methods

表 4 不同方法下的实验性能对比表

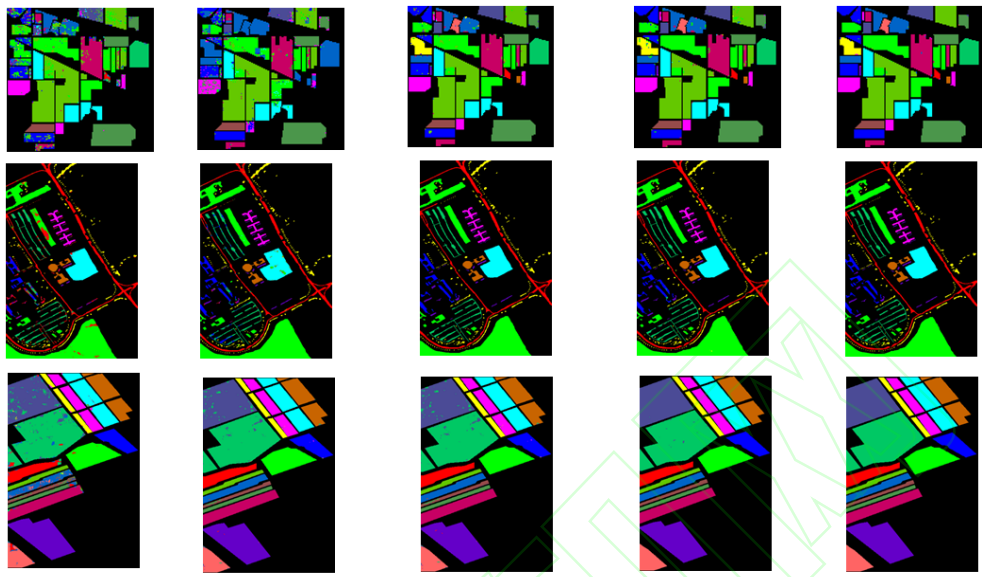
Methods		2D-CNN	3D-CNN	Multi-scale-3D-CNN	Hybrid SN	Proposed
Indian Pines	OA(%)	80.27	82.62	81.39	97.75	**99.61**
	AA(%)	68.32	76.51	75.22	97.54	**99.65**
	Kappa×100	75.26	79.25	81.20	97.44	**99.78**
Salinas scene	OA(%)	96.34	85.00	94.20	98.06	**99.96**
	AA(%)	94.36	89.63	96.66	98.80	**99.94**
	Kappa×100	95.93	83.20	93.61	97.85	**99.95**
Pavia University	OA(%)	96.63	96.34	95.95	98.40	**99.97**
	AA(%)	94.84	97.03	97.52	97.89	**99.97**
	Kappa×100	95.53	94.90	93.40	97.89	**99.96**



类别 1=Asphalt 2=Meadows 3=Gravel 4=Trees 5=Painted metal sheets 6=Bare Soil 7=Bitumen 8=Self-Blocking Bricks 9=Shadows。

Fig.12 Confusion matrix on PU dataset

图 12 在 PU 数据集上的混淆矩阵



(a)2D-CNN (b)3D-CNN (c)Multi-scale-3D-CNN (d)Hybrid SN (e)Proposed

Fig.13 classification effect of different methods on three datasets

图 13 不同方法在在三个数据集上的分类效果图

4 结束语

本文提出的快速 3D-CNN 结合深度可分离卷积的高光谱图像分类方法，首先利用 IPCA 对高光谱原始图像进行降维预处理，降低冗余频谱，减少了图像波段的数量同时保持空间维度的完整性；利用三维卷积神经网络同时提取光谱和空间特征，然后引入深度可分离卷积，设计了新的卷积层 DSC 层。该层充分发挥了深度可分离卷积对空间特征提取的优势，能大幅度节省可学习的参数；最后基于两种卷积方式设计了快速 3D-CNN 和深度可分离卷积结合的网络框架。实验表明，该方法不仅在有限的标签样本下表现出了较好的分类性能，而且同基于标准卷积层的模型相比，大大降低了模型复杂度，减少可学习参数的同时节省了内存空间。

将提出的模型方法和其他传统的卷积神经网络的方法相比较，分类的性能较好，但是本文中还存在许多不足之处。例如如何设计更加完善的深度卷积网络模型，解决网络梯度下滑的问题，这将成为下一步的研究重点。

参考文献:

[1]Chen B. Hyperspectral image processing and information extraction[J].Journal of remote sensing,2016,1062-1090.

张兵.高光谱图像处理与信息提取前沿[J].遥感学报，2016, 1062-1090.本文提出的 3D-2D CNN 结合光谱-空间注意力的高

[2]Zhang C M, Mu Y K, Yan T Y. Development and Prospect of Hyperspectral[J]. Remote Sensing Technology, 2018, 171(3): 108-118.

张淳民，穆廷魁，颜廷昱. 高光谱遥感技术发展与展望.航天返回与遥感[J],2018,171(3):108-118.

[3]Zhao W, Du S. Spectral–spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach[J].IEEE Transactions on Geoscience and Remote Sensing,2016,54(8):4544-4554.

[4]Liu Y F, Li X R, Feng Y M, et al.Representativeness and Redundancy-Based Band Selection for Hyperspectral Image Classification[J].International Journal of Remote Sensing, 2021,42(9):3534-3562.

[5]Yu H L,Xie S J,Zou J J. Classification of 3D CNN hyperspectral remote sensing images with standard fractional dimension reduction[J].Computer engineering and Application,2021,57(4):169-175.

佘海龙，解山娟，邹静洁.标准分数降维的 3D-CNN 高光谱遥感图像分类[J].计算机工程与应用,2021,57(4):169-175.

[6]Yu C, Han R, Song M ,et al. A simplified 2d-3d cnn architecture for hyperspectral image classification based on spatial-spectral fusion[J].IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, PP(99): 1-1.

[7]CHEN Y, LIN Z, ZHAO X. Deep learning-based classification of hyperspectral Data[J].IEEE Journal of Selected Topics in Applied Earth Observations& Remote Sensin, 2014, 7(6): 2094-2107.

[8]Zhao X, Chen, Y S Jia. Spectral-spatial classification of hyperspectral data based on deep belief network[J].IEEE journal of selected topics in applied earth observations and remote

sensing,2015,8(6):2381-2392.
[9]Wei H, Yangyu H , Li W , et al. Deep convolutional neural networks for hyperspectral image classification[J].Journal of Sensors,2015,2015:1-12.
[10] Chen Y , Jiang H , Li C ,et al. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J].IEEE Transactions on Geoscience & Remote Sensing, 2016,54(10):6232-6251.
[11] Zhong Z , Li J, Luo Z,et al. Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework[J].IEEE Transactions on Geoscience & Remote Sensing.2017,1-12.
[12] Sellami A,Farah M,Farah I R ,et al. Hyperspectral imagery classification based on semi-supervised 3-d deep neural network and adaptive band selection[J].Expert Systems with Application,2019,129(SEP),246-259.
[13]Han X , Jiang T , Zhao Z, et al. Target Recognition Based on Deep Convolution Neural Network[J].Research on Remote Sensing Image&International Journal of Pattern Recognition and Artificial Intelligence.2019,34(2).
[14]Chen L C, Papandreou G, Kokkinos I,et al.Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
[15]Fang L, Liu G, Li S, et al.Hyperspectral image classification with squeeze multibias network[J].IEEE Transactions on Geoscience and Remote Sensing,2019,57(3):1291-1301.
[16] Ma S, Liu W, Cai W ,et al. Lightweight deep residual cnn for fault diagnosis of rotating machinery based on depthwise separable convolutions[J].IEEE Access,2019,7, 57023-57036.
[17]Cao Y K,Gui L Y.Design of lightweight time convolution network based on deep separable convolution[J].Computer Engineering,2020.
曹渝昆,桂丽媛.基于深度可分离卷积的轻量级时间卷积网络设计[J].计算机工程,2020.
[18]Khan Z Y, Niu Z. Cnn with depthwise separable convolutions and combined kernels for rating prediction[J].Expert Systems with Applications,2020, 114528.
[19] Ahmad M , Khan A M , Mazzara M , et al. A Fast and Compact 3-D CNN for Hyperspectral Image Classification[J]. IEEE Geoscience and Remote Sensing Letters, 2021:1-5.
[20] Ben Hamida A, Benoit A, Lambert P, et al.3-D deep learning approach for remote sensing image classification[J].IEEE Transactions on Geoscience and Remote Sensing, 2018,56(8):4420-4434.
[21]Chen S, Jin M, Ding J. Hyperspectral remote sensing image classification based on dense residual three-dimensional convolutional neural network[J]. Multimedia tools and applications. 2021;2020,80(2):1859-82.
[22] Chen H , He M, Li B. Multi-scale 3D deep convolutional neural network for hyperspectral image classification[C]//in IEEE International Conference on Image Processing (ICIP), 2017, 3904-3908.
[23]Swalpa, Kumar, Roy. HybridSN: Exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 17(2): 277-281.

王燕（1971-），女，甘肃泾川，硕士，教授,CCF 会员,主要研究方向为模式识别与人工智能。
WANG Yan, born in 1971, M.S.professor. CCF member. Her research interests include Pattern recognition and Artificial intelligence.

梁琦（1996-），女，甘肃静宁，硕士，主要研究方向为模式识别与人工智能。
LIANG Qi, born in 1996, M.S. Her research interests include Pattern recognition and Artificial intelligence.