

基于三维空洞卷积残差神经网络的高光谱影像分类方法

颜铭靖, 苏喜友*

北京林业大学信息学院, 北京 100083

摘要 高光谱影像是典型的高维数据,在光谱维和空间维都包含了大量信息。针对高光谱影像分类时光谱维数据量巨大的特点,提出一种基于三维空洞卷积残差神经网络的高光谱影像分类方法。该方法以高光谱像元立方体作为数据输入,使用三维卷积核同时提取高光谱数据的空间维和光谱维特征,并通过在卷积核中引入空洞结构,在不增加网络参数量和减少数据特征的情况下提高卷积核的感受野,从而提高神经网络的分类精度。该方法利用残差结构避免了由网络层数加深导致的梯度消失问题,最终使用 Softmax 分类器完成高光谱像元的分类工作。实验结果表明:所提方法在 Indian Pines 和 Salinas 数据集上分别取得了 97.303% 和 97.236% 的总体分类精度,与各对照组相比具有更好的分类效果,由此证明所提方法可以提升高光谱影像的分类性能。

关键词 遥感; 高光谱影像分类; 空洞卷积; 三维卷积; 残差神经网络

中图分类号 TP751.1

文献标志码 A

doi: 10.3788/AOS202040.1628002

Hyperspectral Image Classification Based on Three-Dimensional Dilated Convolutional Residual Neural Network

Yan Mingjing, Su Xiyou*

School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China

Abstract Hyperspectral image is typical high-dimensional data which contains abundant information in both spectral and spatial dimensions. In this paper, a hyperspectral image classification method based on three-dimensional dilated convolutional residual neural network is proposed for characterizing large amounts of data in the spectral dimension during hyperspectral image classification. In this method, hyperspectral pixel cubes were applied as input data. Further, a three-dimensional convolutional kernel was used to simultaneously extract the spectral and spatial characteristics of hyperspectral data. Then the receptive field of the convolutional kernel was enhanced without adding network parameters or reducing data features by introducing a dilated structure in the convolutional kernel. Thus, the classification accuracy of the neural network was improved, avoiding the problem of gradient disappearance caused by the deepening of network layers using a residual structure. Finally, the Softmax classifier was used to complete the classification of hyperspectral pixels. Results show that this method obtained an overall classification accuracy of 97.303% and 97.236% on the Indian Pines and Salinas datasets, respectively, exhibiting a better classification effect than other control groups. Thus, the proposed method can improve the classification performance of hyperspectral images.

Key words remote sensing; hyperspectral image classification; dilated convolution; three dimensional convolution; residual neural network

OCIS codes 280.4991; 100.4145; 110.2960; 300.6170

1 引 言

遥感卫星影像的应用一直都是世界各国重点研究的方向,在各应用方向中,遥感影像分类是其中非

常重要的一支。在现阶段,随着空间高分辨率和光谱高分辨率遥感影像的出现,遥感影像分类技术得到了快速发展。高光谱遥感影像的像元可以提供大量的空间维与光谱维信息,从而大大提高了遥感影

收稿日期: 2020-03-06; 修回日期: 2020-04-15; 录用日期: 2020-05-15

基金项目: 北京农业信息技术研究中心开放课题(KF2018W004)

* E-mail: suxiyou@126.com

像分类的精度^[1]。

遥感影像的自动分类方法一般可以分为三大类,即监督分类、非监督分类、半监督分类。常用的监督分类方法有 Logistic 回归(LR)、支持向量机(SVM)、最小距离分类(MDC)、最大似然分类(MLC)^[2-3];非监督分类方法有 K-Means 均值算法、迭代自组织的数据分析法(ISODATA)等^[4]。上述方法均基于遥感影像中的浅层特征进行学习和分类,依据浅层特征的线性或非线性处理结果来区分不同的地物。上述方法的优势是实现起来比较简单、计算量小、方便部署,劣势是分类精度不高。随着遥感影像传感器和计算机技术的飞速发展,基于浅层特征的分类方法逐渐失去优势,同时,基于浅层特征的方法不能学习到高光谱遥感影像中更复杂的信息^[5],因此以卷积神经网络(CNN)为代表的深度学习成为遥感影像地物分类的新趋势。

滕文秀等^[6]首次将深度学习的概念应用于高光谱影像分类,使用堆叠式自动编码器(SAE)提取高光谱影像的光谱维和空间维信息,取得了良好的效果。Zhong 等^[7]使用包含三维卷积层的光谱空间残差网络(SSRN)提取高光谱影像立方体特征,通过引入残差结构来缓解由网络深度增加引起的梯度消失问题;在高光谱影像分类中,该方法的分类精度较传统方法有了很大提升。此外,还有研究者提出了一种基于门控循环单元(GRU)的级联递归神经网络(RNN)模型,并通过增加若干卷积层来提取空间维和光谱维的特征^[8]。

三维结构的高光谱影像通常同时包含空间信息以及丰富的光谱信息。单纯基于光谱特征或基于空间特征的分类方法都不能完全发挥高光谱影像的优势^[9]。在对三维数据进行处理时,最小距离法和

SVM 等基于光谱的分类方法很难获取空间维的信息,而二维卷积神经网络(2D-CNN)不能很好地获取光谱维信息。与前述两类方法相比,将空间信息和光谱信息相结合的三维卷积神经网络(3D-CNN)在进行高光谱图像分类时更符合高光谱影像的特点,有利于提高光谱影像分类的精度。同时,为解决高光谱影像光谱维数据量较大的问题,本文引入了空洞卷积结构,该结构能在不增加网络参数的情况下扩增卷积层的感受野,避免了采用池化层等传统方法扩增时引起的特征丢失问题。该思路最初见于医学图像处理领域处理图像数量巨大的核磁共振影像^[10]。本文考虑到医学影像与高光谱影像数据结构具有一定的相似性,受其启发,提出了一种基于三维空洞卷积残差结构的卷积神经网络高光谱影像分类方法,以提高高光谱影像地物分类的精度。

2 研究方法

2.1 3D-CNN

CNN 最初被用于二维数据的处理与分析,在目标识别、图像分割等方面具有良好的应用效果。但是,传统 CNN 使用的卷积核通常是二维的,在处理高光谱影像这种三维结构的高维数据时就需要引入大量参数,在处理标注信息较少的高光谱影像时,有可能产生过拟合现象。同时,高光谱遥感影像在空间维也包含有大量信息,2D-CNN 的卷积核结构不利于同时提取像元的空间信息与光谱信息^[11]。为对上述问题进行优化,可引入 3D-CNN。如图 1 所示,3D-CNN 不是将空间信息和光谱信息作为独立的特征进行提取,而在空间维和光谱维上同时进行运算,可以同时提取空谱特征。

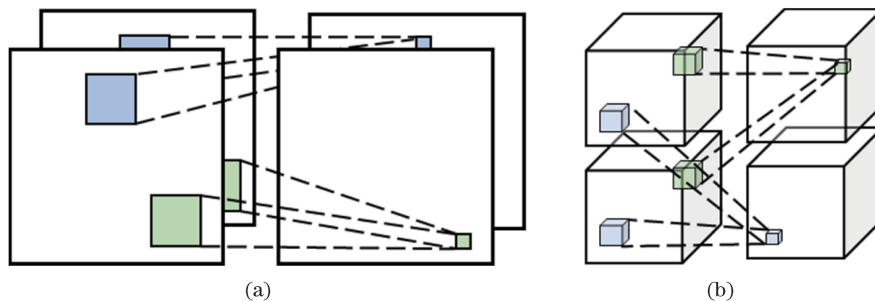


图 1 2D-CNN 与 3D-CNN 示意图。(a) 2D-CNN; (b) 3D-CNN

Fig. 1 Two-dimensional and three-dimensional convolution network diagrams. (a) 2D-CNN; (b) 3D-CNN

3D-CNN 的卷积核会在三个方向上移动,计算神经网络第 i 层第 j 个特征图在 (x, y, z) 处的点值 V 的计算公式为

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

式中： m 表示第 $i-1$ 层中与当前特征图相连的特征图； L_i 与 W_i 表示卷积核的长度和宽度； H_i 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； $b_{i,j}$ 表示第 i 层第 j 个特征图的偏置； f 为激活函数。3D-CNN 最初由 Ji 等^[12] 提出并应用于人体运动识别，目前在视频分析和医疗影像处理领域应用得较多。3D-CNN 与 2D-CNN 最大的差别是 3D-CNN 的卷积核会在三维方向上进行卷积操作，可以获取到光谱维信息，适合处理高光谱影像数据这样的三维数据。

2.2 空洞卷积

对于高光谱影像，线性组合中的两个波段的图层距离通常较远。以植被识别与监测中常用的归一化植被指数 (NDVI) 为例，在机载可见光/红外成像光谱仪 (AVIRIS) 高光谱卫星的遥感影像数据中，红色波段 (655.56 nm) 与近红外波段 (819.48 nm) 间隔 19 个图层，需要使用 3 个以上长度为 3 的卷积核才能覆盖上述两个图层，进而提取到 NDVI 特征。若要通过 3D-CNN 以较少的卷积层数量学习到上述光谱特征，则需要扩大卷积层的感受野。扩大感受野的常用方法是扩大卷积核或添加池化层 (pooling)，虽然较大的三维卷积核有利于神经网络获取更多的光谱上下文信息，但会增加网络的计算负荷，而使用池化层则会牺牲一部分特征。在此情况下，引入空洞卷积层替代传统池化层，可以在减少数据损失的同时扩大卷积层的感受野^[13]。

空洞卷积在原始卷积核的基础上，通过在值与值之间插入权重为 0 的行与列来扩大卷积核的感受野，如图 2 所示。空洞卷积使用膨胀系数 r 来控制插入权重为 0 的行和列的数量，这样得到的卷积核的感受野的长、宽、高均为 $L + (L-1)(r-1)$ ， L 代表原卷积核的边长。当原始 3D 卷积核为 $3 \times 3 \times 3$ 时，其单核感受野范围与卷积核尺寸一致，而转化为空洞 3D 卷积核则可以使单核感受野增至 $5 \times 5 \times 5$ 。当多个卷积层组合时，空洞卷积核的优势就会更加明显，原始 $3 \times 3 \times 3$ 的 3D 卷积核经过 3 层叠加，最终的感受野为 $7 \times 7 \times 7$ ， $r=2$ 的空洞卷积核则为 $13 \times 13 \times 13$ ，从而使得在处理高光谱数据时，可以更频繁地感受到更多的光谱特征。

2.3 残差结构

与浅层神经网络相比，深层 CNN 具有更多的非线性映射结构，可完成结构复杂的函数逼近，获得的特征更加抽象，提取到的语义信息更加完整，适合处理数据量大的高光谱影像数据^[14]。然而，神经网络

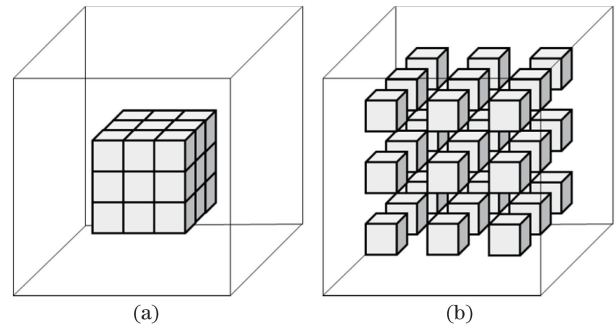


图 2 普通卷积核与空洞卷积核示意图。

(a) 普通卷积核；(b) 空洞卷积核 ($r=2$)

Fig. 2 Normal and dilated convolution kernel diagrams.

(a) Normal kernel; (b) dilated kernel ($r=2$)

层次的加深会带来梯度消失的问题，靠前部分的隐藏层权值更新缓慢或者更新停滞，导致分类的准确率不再继续增加，出现随着层次增加准确率反而下降的情况。为解决上述问题，He 等^[15] 提出了残差网络 (ResNet)，其中的残差结构可以有效解决由神经网络深度增加而导致的梯度下降问题。

残差结构最大的特点就是采用跳层连接，这使得其相较于普通神经网络具有更好的学习性能。残差模块输入与输出的关系可以表示为

$$H(x) = f(x) + x, \quad (2)$$

式中： $H(x)$ 表示计算结果； x 表示模块输入； $f(x)$ 表示残差学习函数。残差结构可根据具体需要对 $f(x)$ 进行相应的改变，可将 (2) 式扩充为

$$y = f(x, \{W_i\}) + x, \quad (3)$$

式中： x 与 y 分别对应残差模块的输入与输出； $f(x, \{W_i\})$ 表示待训练的残差映射，输入 x 和残差映射 f 的维度需保持一致。在省略偏差的情况下，可将残差映射写成 $f(x, \{W_i\}) = W_2 \sigma(W_1 x)$ ，其中 σ 为激活函数，在这里表示 ReLU 激活函数。

若输入 x 和残差映射 f 的维度不一致，则需要跳层连接上增加一个线性投影 W_s 来使维度相同，即

$$y = f(x, \{W_i\}) + W_s x. \quad (4)$$

在残差模块训练过程中，输出是由输入与其矩阵变化的结果相加得到的，并未引入新的参数，网络的参数量不会变化，在不影响反向传播过程的同时能够加快模型的训练速度并提高训练效果。

3 网络结构与参数分析

3.1 网络结构

普通的 CNN 通常适用于二维平面数据，在处理高光谱影像这种高维数据时，不能很好地同时提

取到空间维和光谱维特征。针对此问题,本文提出了一种包含空洞卷积层和残差连接的 3D-CNN,用它对高光谱影像进行地物分类。进入神经网络后,在网络前段进行浅层特征提取,并去除部分噪声,在网络中段进行深度特征的提取,提取出深度特征后,在网络后段降低数据量和进一步去除噪声,最终使用 Softmax 分类器完成像元块数据的地物分类。

残差特征学习部分包含两个空洞卷积结构块,每个结构块包含 3 个卷积层。为深入探讨空洞卷积层对神经网络的影响,本文从感受野覆盖范围和形式以及累加线性特征进行非线性处理两方面进行研究。本文假设了如图 3 所示的空洞卷积层与普通卷积层的 7 种排列组合方式以及残差结构块内部激活函数的 2 种分布策略。

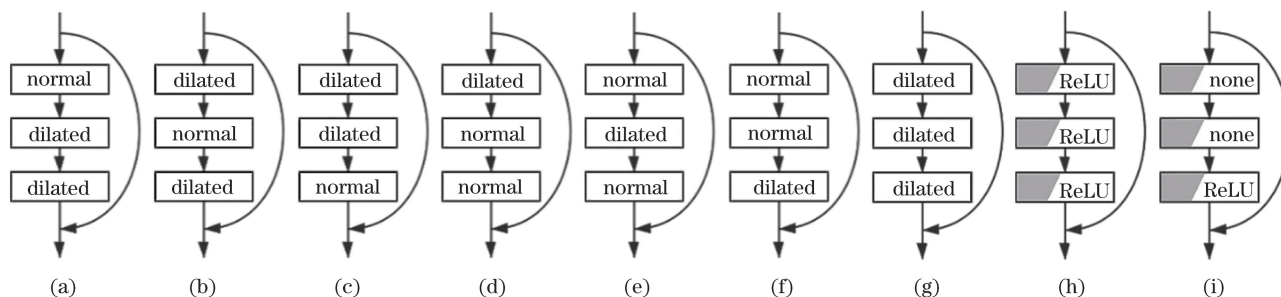


图 3 空洞卷积层与普通卷积层的 7 种排列组合以及内部激活函数的 2 种分布策略。(a) 方式 1; (b) 方式 2; (c) 方式 3; (d) 方式 4; (e) 方式 5; (f) 方式 6; (g) 方式 7; (h) 分布策略 I; (i) 分布策略 II

Fig. 3 Seven permutation and combination types of dilated and normal convolutional layers and two activation function distribution strategies. (a) Type 1; (b) type 2; (c) type 3; (d) type 4; (e) type 5; (f) type 6; (g) type 7; (h) distribution strategy I; (i) distribution strategy II

不同空洞卷积结构块在训练时均采用相同的训练参数和训练集,并选用总体精度(OA)、平均精度(AA)和 Kappa 系数(Kappa)作为模型精度的评价

指标。表 1 为不同空洞卷积层与普通卷积层排列组合策略的训练结果评价,表 2 为不同结构块内部激活函数分布策略的训练结果评价。

表 1 采用第 1 种激活函数分布策略(I)的排列组合的分类精度

Table 1 Classification accuracy of permutation and combinations with the first active function distribution strategy (I) %

| Structure type | Indian Pines | | | | Salinas | | | |
|----------------|--------------|--------|--------|---------|---------|--------|--------|---------|
| | Kappa | OA | AA | OA-mean | Kappa | OA | AA | OA-mean |
| Type 1- I | 95.852 | 96.976 | 92.041 | | 95.797 | 96.986 | 94.524 | |
| Type 2- I | 95.676 | 96.851 | 91.499 | 96.942 | 95.534 | 96.794 | 94.748 | 96.823 |
| Type 3- I | 95.885 | 97.000 | 91.383 | | 95.388 | 96.690 | 94.694 | |
| Type 4- I | 95.155 | 96.480 | 90.501 | | 95.862 | 97.028 | 95.027 | |
| Type 5- I | 94.811 | 96.232 | 89.872 | 96.505 | 95.808 | 96.991 | 94.772 | 96.943 |
| Type 6- I | 95.609 | 96.801 | 91.559 | | 95.554 | 96.810 | 94.666 | |
| Type 7- I | 95.680 | 96.855 | 91.335 | 96.855 | 95.320 | 96.644 | 94.192 | 96.644 |

由表 1 与表 2 可知:双空洞卷积层与单普通卷积层组合(Type 1、Type 2、Type 3)的平均总体精度整体上高于单空洞卷积层与双普通卷积层组合(Type 4、Type 5、Type 6)的平均总体精度,但全空洞卷积层组合的总体精度整体上低于双普通卷积层组合的平均总体精度。双空洞卷积层与单普通卷积层的组合以残差块为激活单元进行激活(分布策略 I),激活的效果整体上优于逐层进行激活的策略(分布策略 II)。

虽然空洞卷积结构能够扩大卷积层的感受野,

但是单纯增加空洞卷积层的效果并不能取得最优的感受野分布效果。由图 4 可以看出:纯空洞卷积层组合的感受野虽然面积较大,而且空洞卷积层的特性使得部分数据位置的访问频次很大,但有较多的数据位置并没有被访问到;空洞卷积层与普通卷积层组合的感受野的访问频次分布得更加均匀,虽然感受野面积略小于前者,但在感受野范围内不存在未访问的数据位置。综合卷积核分布效果可知:纯空洞卷积层组合的感受野跨度为 13,在 3 种分布策略中最大,但其有效访问像元数量仅为 49;一空洞

表 2 采用第 2 种激活函数分布策略(Ⅱ)的排列组合的分类精度

Table 2 Classification accuracy of different permutations and combinations with the second active function distribution strategy (Ⅱ)

| Structure type | Indian Pines | | | | Salinas | | | |
|----------------|--------------|--------|--------|---------|---------|--------|--------|---------|
| | Kappa | OA | AA | OA-mean | Kappa | OA | AA | OA-mean |
| Type 1-Ⅱ | 95.976 | 97.066 | 91.912 | | 95.804 | 96.989 | 94.793 | |
| Type 2-Ⅱ | 95.698 | 96.866 | 91.800 | 96.973 | 95.883 | 97.045 | 94.903 | 96.944 |
| Type 3-Ⅱ | 95.866 | 96.987 | 91.956 | | 95.532 | 96.796 | 94.432 | |
| Type 4-Ⅱ | 95.775 | 96.921 | 91.839 | | 95.455 | 96.741 | 94.262 | |
| Type 5-Ⅱ | 94.798 | 96.220 | 90.179 | 96.689 | 95.718 | 96.928 | 94.737 | 96.856 |
| Type 6-Ⅱ | 95.783 | 96.927 | 91.543 | | 95.676 | 96.898 | 94.682 | |
| Type 7-Ⅱ | 95.618 | 96.808 | 91.351 | 96.808 | 95.457 | 96.739 | 94.568 | 96.739 |

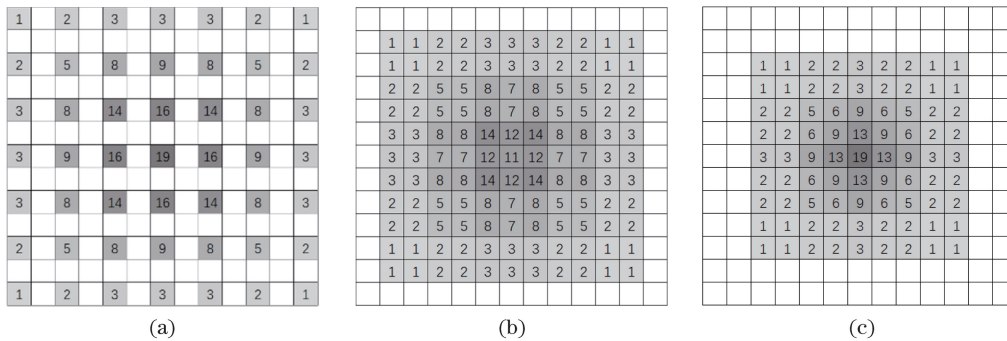


图 4 不同卷积层组合的感受野分布。(a)空洞参数分布为(2,2,2);(b)空洞参数分布为(1,2,2);(c)空洞参数分布为(1,1,2)

Fig. 4 Receptive field's distributions of different convolution combinations. (a) Dilation parameter distribution is (2, 2, 2); (b) dilation parameter distribution is (1, 2, 2); (c) dilation parameter distribution is (1, 1, 2)

二卷积组合的感受野跨度为 9,在 3 种分布策略中最小,其有效访问像元数量为 81;二空洞一卷积组合的感受野跨度为 12,仅比纯空洞卷积组合小 1,但有效访问像元数量为 121,综合性能最好。表 1 与

表 2 的分类精度结果也支持这一结论。

综合比较表 1 与表 2 的分类精度,最终选取 Type 1-Ⅱ 的空洞卷积结构块组成残差特征学习部分。CNN 整体结构如图 5 所示。

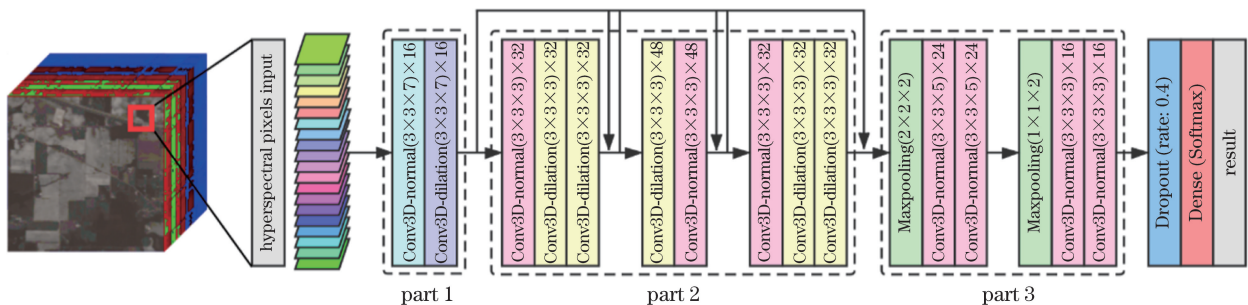


图 5 网络结构

Fig. 5 Network structure

该网络分为三个主体部分:part 1 包含 2 个卷积核大小为 $3 \times 3 \times 7$ 、通道数为 16 的 3D 卷积层,其中 1 个为空洞卷积层;part 2 包含 8 个卷积核大小为 $3 \times 3 \times 3$ 、通道数分别为 32 和 48 的 3D 卷积层,其中有

5 个为空洞 3D 卷积层;part 3 包含 2 个 Maxpooling 层,2 个卷积核大小为 $3 \times 3 \times 5$ 、通道数为 24 的 3D 卷积层,以及 2 个卷积核大小 $3 \times 3 \times 3$ 、通道数为 16 的 3D 卷积层,所有卷积层的卷积模式均为 same,并在

归一化之后使用 ReLU 函数进行激活。在数据通过神经网络后,为避免过拟合,在结果分类前添加参数为 0.4 的 Dropout 层,最终分类由全连接层和 Softmax 分类器完成。

本网络输入的数据为 $7 \times 7 \times N$ 的像元块, N 为像元块的光谱维长度,数据标签为像元块空间几何中心像元所属的地类。为模拟实际的应用场景,训练数据集和验证数据集中均混入了未进行地物类别标注的背景像元,训练集和验证集的像元数量均为总像元数量的 50%,训练集中包含 20% 的已标记地物类别的像元,验证集中包含 80% 的已标记地物类别的像元。

实验总流程可以分为训练网络和验证网络两大部分,训练网络包含制作训练数据和训练神经网络两部分。在训练网络阶段,首先在原图像中按比例抽取已标记和未标记的高光谱像元立方体,进行随机混序排列,然后将它们依次输入到网络模型中进行特征提取,最终使用 Softmax 分类器进行分类。

在验证网络阶段,先将验证数据依次输入到已训练好的网络模型中,提取出分类特征,采用 Softmax 分类器进行分类预测,然后使用精度评价指标验证网络的有效性。

3.2 参数设置与分析

在空洞 CNN 中,膨胀系数(dilation rate)是一个很重要的参数,代表卷积核中有效行列之间不参与计算的行列的个数。扩大因子增大会使卷积核的覆盖范围变大,从而扩大感受野,使卷积层能够学习到更广的特征,但也会使一部分细节信息被忽视掉。

在实验中,选取 2~7 共 6 个扩大因子,对其影响进行分析。考虑到数据集中各地类的像元数量极不平衡,最终选用 OA、AA 和 Kappa 系数作为模型精度的评价指标,其中,OA 表征对总体分类结果的评价,AA 表征各类别自身分类精度的均值,Kappa 系数表征模型分类结果与参考结果的一致程度。空洞卷积层中不同膨胀系数对应的精度结果如图 6 所示。

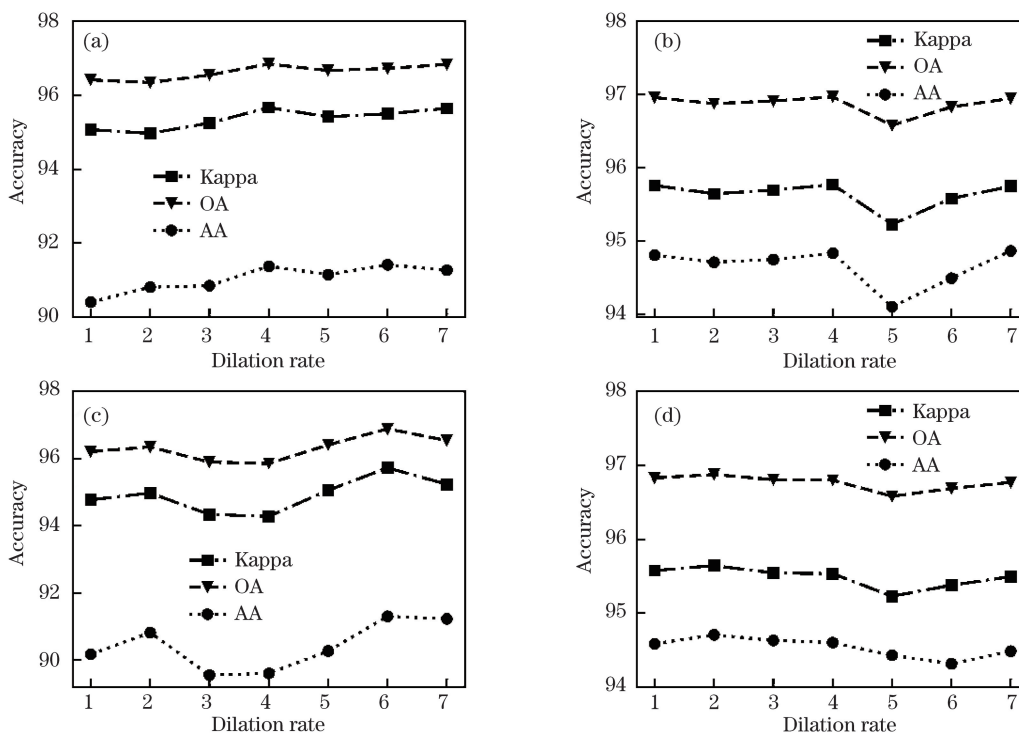


图 6 两个数据集中空洞参数(膨胀系数)对应的精度。(a) Indian Pines 光谱维;(b) Salinas 光谱维;(c) Indian Pines 空间维;(d) Salinas 空间维

Fig. 6 Corresponding precision of dilation rate in two datasets. (a) Indian Pines spectral dimension; (b) Salinas spectral dimension; (c) Indian Pines spatial dimension; (d) Salinas spatial dimension

实验结果表明,在空间维,空洞参数虽然可以发挥一定作用,但当空洞参数超过 2 之后,空间维空洞结构的增大反而会带来负面影响,即卷积核尺寸大于数据块的空间维尺寸,影响分类精度;在光谱维,

空洞参数与总体精度 OA 之间不是线性关系,随着空洞参数增大,总体精度 OA 先小幅下降再上升,然后再下降。当空洞参数开始增大时,卷积层感受野扩张范围不足,同时特征损失增加,导致总体精度

OA 下降;当空洞结构增大到一定范围时,感受野的扩张使卷积层能学习到更多的特征,从而弥补了空洞结构所导致的特征损失,使分类精度 AA 升高;随着空洞参数继续增大,由于光谱维数据有限,感受野的增大并不能学习到更多的特征,而空洞结构的特征损失继续增大,从而导致分类精度 AA 再次下降。综合两个数据集的实验结果,本文选取光谱维为 4、空间维为 2 的空洞参数。

4 实验结果与分析

为评价本文所提神经网络结构的分类效果,选取 Indian Pines 以及 Salinas 两个典型的高光谱遥感数据集进行地物分类实验。

Indian Pines 数据集是由 AVIRIS 于 1992 年对美国印第安纳州一块实验地进行成像得到的遥感图像,空间分辨率为 20 m,包含 145×145 个地物像素和 224 个光谱通道,去掉水分吸收较强的光谱通道(第 104~108 通道,第 150~163 通道,第 220 通道)后,共有 200 个光谱通道可用于分类。该数据集包含 16 个地物类别和 1 个背景类别,其中包含 10249 个已标记地物类别的像元和 10776 个背景像元。

Salinas 数据集是由 AVIRIS 于 1998 年对美国加利福尼亚州的 Salinas 山谷进行成像的遥感图像,空间分辨率为 3.7 m,包含 512×217 个地物像素和 224 个光谱通道,去掉水分吸收较强的光谱通道(第 108~112 通道,第 154~167 通道,第 224 通道)后,共有 200 个光谱通道可用于分类。该数据集包含 16 个地物类别和 1 个背景类别,其中包含 54129 个已标记地物类别的像元和 56975 个背景像元。

图 7 为上述两个地区的伪彩色合成图像。

实验所用的计算机配置为 Intel(R) Core(TM) i7-7700HQ CPU, NVIDIA GeForce GTX 1070, 16 GB 内存,在 Windows 10 系统下基于 Python 3.71 的 Keras 框架实现。实验数据使用主成分分析(PCA)法进行预处理,在保留 99.9%原始特征的条件对光谱维进行压缩。在网络中使用 Adam 优化器进行网络训练,初始学习率为 0.001,设置损失函数的稳定容忍值为 4,每次触发使学习率降低到其当前值的 1/10。网络各卷积层采用 He normal^[16]进行初始化,激活函数选用 ReLU。

为验证本文提出的网络模型的有效性,选取了 5 个网络模型对照组,包括经典的 SVM 分类算法、相同网络层次结构的 2D-CNN、Res-3DCNN^[17]、

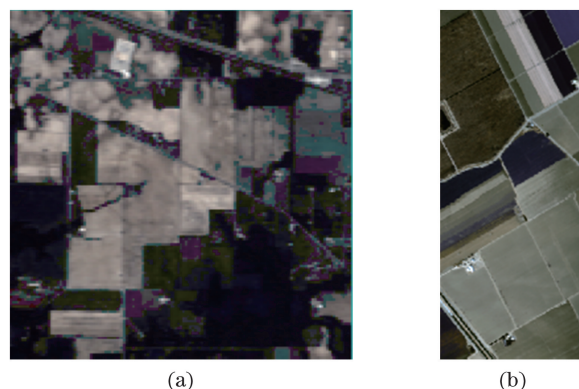


图 7 数据集伪彩色合成图。(a) Indian Pines;
(b) Salinas

Fig. 7 Pseudo-color composite images of two datasets.

(a) Indian Pines; (b) Salinas

M3D-DCNN^[18]、相同网络层次结构的 3D-CNN。在相同网络层次结构的对照组中,所有的 CNN 对照组只在卷积和池化层属性上有差异,层数和位置关系均保持一致,并使用相同的训练与验证数据集以及相同的网络训练参数进行训练。所有对照实验的结果均为 20 次实验结果的均值。表 3 和表 4 分别为 5 个网络模型对照组与本文所提网络模型 Dilated-3D-CNN 在 Indian Pines 数据集与 Salinas 数据集上的分类精度。

由表 3 和表 4 可知:在进行高光谱像元分类时,CNN 经过了多层次的特征学习与筛选,其分类精度和 Kappa 系数普遍优于传统机器学习算法 SVM;包含 3D 卷积结构的 Res-3DCNN、M3D-DCNN、3D-CNN 与 Dilated-3D-CNN 的分类精度与 Kappa 系数均高于 2D-CNN,这表明 3D 结构在高光谱像元分类中可以提取到更多的特征,具有一定优势;Dilate-3D-CNN 的 Kappa 系数与分类精度均高于 3D-CNN;在网络参数量一致的情况下,Kappa 系数和总体分类精度 OA 在 Indian Pines 和 Salinas 数据集上分别提升了 0.927 和 0.669 个百分点、0.582 和 0.416 个百分点,这表明在卷积层中加入空洞结构可以提高 CNN 的分类性能。

图 8 和图 9 为各个模型在 Indian Pines 和 Salinas 数据集上的分类结果。由于数据经过了 PCA 处理,故而属于线性分类的 SVM 的分类效果明显逊于神经网络。3D-CNN 与 Dilated-3D-CNN 的分类效果如上文所述,错分的情况要少于其他模型。将二者进行对比后可发现,3D-Dilated-CNN 的分类边界更加平滑,已标记像元与背景像元的错分现象更少,更贴近真实地物分布。

表 3 各网络模型在 Indian Pines 数据集上的分类精度

Table 3 Classification accuracy of each network model in Indian Pines dataset

%

| Class name | Classification accuracy | | | | | |
|------------------------------|-------------------------|--------|-----------|----------|---------------|----------------|
| | SVM | 2D-CNN | Res-3DCNN | M3D-DCNN | 3D-CNN | Dilated-3D-CNN |
| Background | 69.123 | 98.797 | 98.406 | 99.240 | 99.376 | 99.653 |
| Alfalfa | 24.348 | 52.174 | 73.587 | 69.130 | 82.065 | 82.065 |
| Corn-notill | 61.583 | 85.007 | 84.352 | 88.235 | 93.092 | 94.492 |
| Corn-mintill | 40.126 | 80.624 | 86.640 | 89.329 | 92.948 | 94.555 |
| Corn | 29.473 | 77.426 | 84.325 | 88.270 | 93.418 | 92.806 |
| Grass-pasture | 73.113 | 83.013 | 83.046 | 86.115 | 90.971 | 91.766 |
| Grass-trees | 75.993 | 89.370 | 89.411 | 92.493 | 94.945 | 96.075 |
| Grass-pasture-mowed | 23.750 | 58.036 | 70.893 | 70.714 | 78.750 | 81.786 |
| Hay-windrowed | 85.471 | 97.406 | 97.416 | 97.699 | 98.180 | 98.441 |
| Oats | 22.750 | 49.250 | 47.000 | 72.000 | 75.250 | 87.000 |
| Soybean-notill | 55.710 | 84.767 | 84.268 | 88.971 | 93.897 | 95.307 |
| Soybean-mintill | 71.536 | 91.440 | 89.668 | 93.733 | 96.449 | 97.132 |
| Soybean-clean | 40.780 | 72.686 | 75.941 | 86.169 | 89.771 | 92.119 |
| Wheat | 84.293 | 93.317 | 94.756 | 95.683 | 96.585 | 97.390 |
| Woods | 55.486 | 86.644 | 86.067 | 92.518 | 95.636 | 95.945 |
| Buildings-Grass-Trees-Drives | 28.455 | 58.503 | 63.965 | 77.532 | 87.946 | 90.303 |
| Stone-Steel-Towers | 33.441 | 77.796 | 81.667 | 86.237 | 90.860 | 89.570 |
| Kappa | 54.791 | 88.317 | 88.514 | 92.541 | 95.377 | 96.304 |
| OA | 64.636 | 91.721 | 91.820 | 94.622 | 96.634 | 97.303 |
| AA | 51.496 | 78.603 | 81.848 | 86.710 | 91.185 | 92.730 |

表 4 各网络模型在 Salinas 数据集上的分类精度

Table 4 Classification accuracy of each network model in Salinas dataset

%

| Class name | Classification accuracy | | | | | |
|---------------------------|-------------------------|--------|-----------|----------|---------------|----------------|
| | SVM | 2D-CNN | Res-3DCNN | M3D-DCNN | 3D-CNN | Dilated-3D-CNN |
| Background | 70.502 | 98.295 | 98.365 | 98.491 | 98.672 | 98.806 |
| Brocoli-green-weeds-1 | 94.556 | 95.768 | 91.594 | 86.573 | 95.888 | 96.374 |
| Brocoli-green-weeds-2 | 95.456 | 98.332 | 97.772 | 98.552 | 98.355 | 98.395 |
| Fallow | 44.098 | 86.180 | 73.336 | 86.282 | 89.653 | 90.567 |
| Fallow-rough-plow | 60.740 | 82.123 | 80.289 | 86.596 | 88.238 | 90.585 |
| Fallow-smooth | 47.043 | 87.957 | 87.580 | 89.402 | 91.423 | 91.356 |
| Stubble | 98.964 | 96.419 | 95.926 | 87.176 | 96.583 | 96.638 |
| Celery | 91.562 | 96.462 | 96.159 | 97.399 | 97.122 | 97.837 |
| Grapes-untrained | 83.636 | 92.165 | 92.618 | 93.300 | 95.896 | 96.643 |
| Soil-vinyard-develop | 64.362 | 93.482 | 92.085 | 93.865 | 95.142 | 96.079 |
| Corn-senesced-green-weeds | 79.297 | 92.466 | 90.876 | 92.527 | 95.597 | 95.165 |
| Lettuce-romaine-4wk | 71.196 | 94.180 | 93.153 | 94.094 | 95.781 | 95.058 |
| Lettuce-romaine-5wk | 31.651 | 95.916 | 90.497 | 94.714 | 97.738 | 98.093 |
| Lettuce-romaine-6wk | 22.858 | 86.096 | 71.331 | 90.291 | 91.745 | 92.651 |
| Lettuce-romaine-7wk | 52.255 | 86.670 | 83.944 | 88.426 | 89.789 | 90.792 |
| Vinyard-untrained | 38.987 | 85.982 | 83.519 | 87.330 | 93.218 | 95.129 |
| Vinyard-vertical-trellis | 97.662 | 95.990 | 96.006 | 96.231 | 96.068 | 96.388 |
| Kappa | 61.951 | 93.267 | 92.013 | 93.226 | 95.567 | 96.149 |
| OA | 70.697 | 95.185 | 94.324 | 95.178 | 96.820 | 97.236 |
| AA | 67.343 | 92.028 | 89.121 | 91.838 | 94.524 | 95.091 |

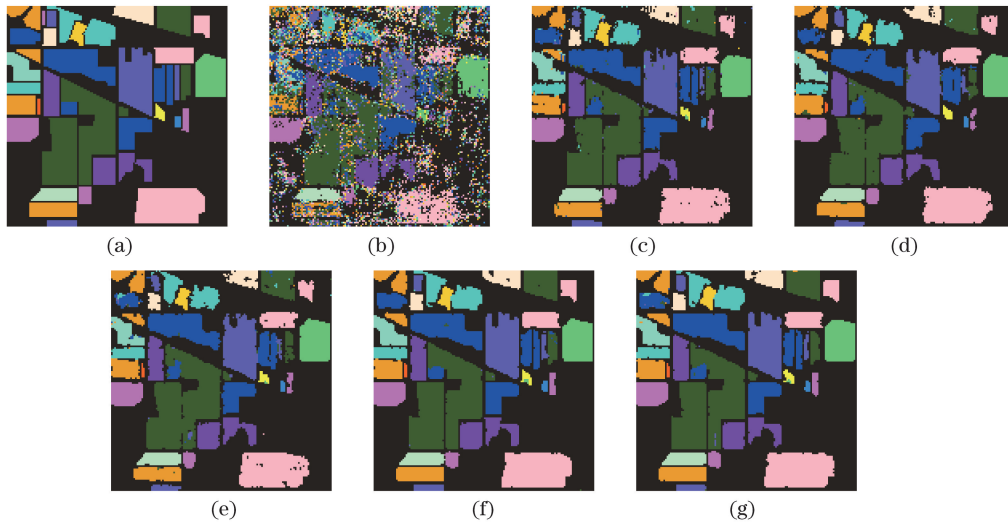


图 8 各网络模型在 Indian Pines 数据集上的分类结果。(a) 真值图;(b) SVM;(c) 2D-CNN;(d) Res-3DCNN;
(e) M3D-DCNN;(f) 3D-CNN;(g) Dilated-3D-CNN

Fig. 8 Classification images of different network models in Indian Pines dataset. (a) True value image;
(b) SVM; (c) 2D-CNN; (d) Res-3DCNN; (e) M3D-DCNN; (f) 3D-CNN; (g) Dilated-3D-CNN

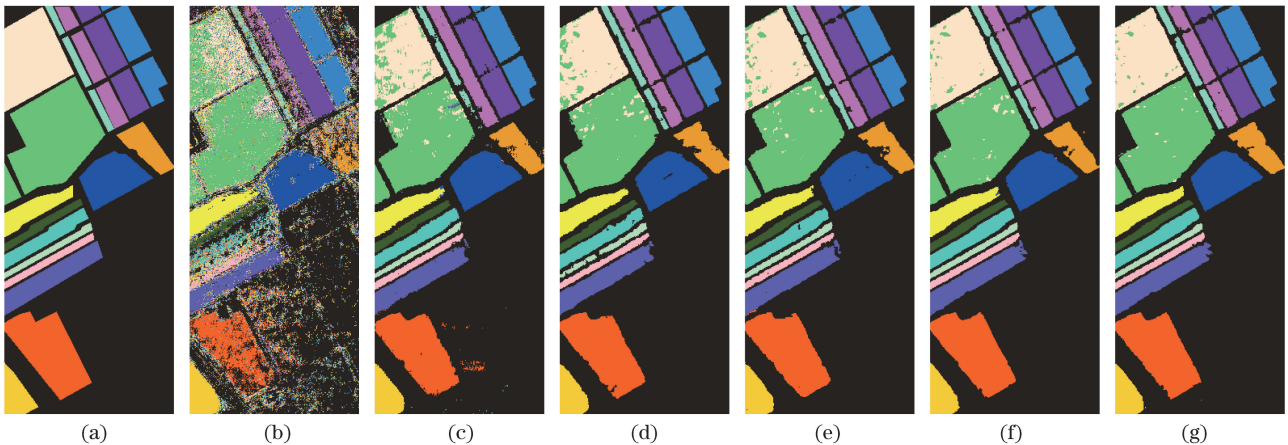


图 9 各网络模型在 Salinas 数据集上的分类结果。(a) 真值图;(b) SVM;(c) 2D-CNN;(d) Res-3DCNN;
(e) M3D-DCNN;(f) 3D-CNN;(g) Dilated-3D-CNN

Fig. 9 Classification images of different network models in Salinas dataset. (a) True value image;
(b) SVM; (c) 2D-CNN; (d) Res-3DCNN; (e) M3D-DCNN; (f) 3D-CNN; (g) Dilated-3D-CNN

5 结 论

本文研究了 CNN 在高光谱影像分类中的应用,设计了一种基于三维空洞卷积残差神经网络的分类模型 Dilated-3D-CNN,该模型通过在网络中引入空洞结构来扩大卷积层的感受野,提高地物分类的精度。在 Indian Pines 和 Salinas 数据集上,将本文所提网络模型与其他 5 种网络模型进行对比,结果表明,本文设计的网络模型具有最好的分类性能,这说明合适的空洞结构可以在网络参数量不变的基础上提高分类精度。

基于三维空洞卷积残差神经网络的高光谱影像

分类方法可以很好地同时提取空间维特征和光谱维特征,为处理三维高光谱影像数据提供了新思路。在实际应用中,由于现有神经网络程序框架的优化问题,空洞结构仍会给网络带来一定负担。随着框架的优化与发展,三维卷积层和空洞结构将在高光谱影像分类领域具有更大的潜力,如何使用二者来提高分类精度和效率还需要进一步研究。

参 考 文 献

- [1] Meher S K. Semisupervised self-learning granular neural networks for remote sensing image classification[J]. Applied Soft Computing, 2019, 83: 105655.
- [2] Li Y F, Lin H. Multi-spectral remote sensing image

- classification of ground coverage based on CNN[J]. *Microprocessors*, 2019, 40(1): 43-48.
- 李玉峰, 林辉. 基于 CNN 的多光谱遥感图像地物覆盖分类[J]. *微处理机*, 2019, 40(1): 43-48.
- [3] Liu P, Choo K K R, Wang L Z, et al. SVM or deep learning? A comparative study on remote sensing image classification[J]. *Soft Computing*, 2017, 21(23): 7053-7065.
- [4] Zhao C X, Qian L X. Comparative study of supervised and unsupervised classification in remote sensing image [J]. *Journal of Henan University (Natural Science Edition)*, 2004, 34(3): 90-93.
- 赵春霞, 钱乐祥. 遥感影像监督分类与非监督分类的比较[J]. *河南大学学报(自然科学版)*, 2004, 34(3): 90-93.
- [5] Cao L L, Li H T, Han Y S, et al. Application of convolutional neural networks in classification of high resolution remote sensing imagery [J]. *Science of Surveying and Mapping*, 2016, 41(9): 170-175.
- 曹林林, 李海涛, 韩颜顺, 等. 卷积神经网络在高分遥感影像分类中的应用[J]. *测绘科学*, 2016, 41(9): 170-175.
- [6] Teng W X, Wang N, Chen T S, et al. Deep adversarial domain adaptation method for cross-domain classification in high-resolution remote sensing images [J]. *Laser & Optoelectronics Progress*, 2019, 56(11): 112801.
- 滕文秀, 王妮, 陈泰生, 等. 基于深度对抗域适应的高分辨率遥感影像跨域分类[J]. *激光与光电子学进展*, 2019, 56(11): 112801.
- [7] 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.
- [8] Mou L C, Ghamisi P, Zhu X X. Deep recurrent neural networks for hyperspectral image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3639-3655.
- [9] Dai X Y, Xue W. Hyperspectral remote sensing image classification based on convolutional neural network[C]//37th Chinese Control Conference. New York: IEEE Press, 2018: 10373-10377.
- [10] Wolterink J M, Leiner T, Viergever M A, et al. Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease[M]// *Reconstruction, Segmentation, and Analysis of Medical Images*. Cham: Springer International Publishing, 2017: 95-102.
- [11] Zhang H K, Li Y, Jiang Y N. Deep learning for hyperspectral imagery classification: the state of the art and prospects[J]. *Acta Automatica Sinica*, 2018, 44(6): 961-977.
- 张号逵, 李映, 姜晔楠. 深度学习在高光谱图像分类领域的研究现状与展望[J]. *自动化学报*, 2018, 44(6): 961-977.
- [12] Ji S W, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 221-231.
- [13] Feng B W, Lü X Q, Gu Y, et al. Brain tumor segmentation using three-dimensional parallel CNN combined with dilated convolution[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141009.
- 冯博文, 吕晓琪, 谷宇, 等. 结合空洞卷积的三维并行 CNN 用于脑肿瘤分割[J]. *激光与光电子学进展*, 2020, 57(14): 141009.
- [14] Sun Z J, Xue L, Xu Y M, et al. Overview of deep learning [J]. *Application Research of Computers*, 2012, 29(8): 2806-2810.
- 孙志军, 薛磊, 许阳明, 等. 深度学习研究综述[J]. *计算机应用研究*, 2012, 29(8): 2806-2810.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 27-30 June 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [16] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//2015 IEEE International Conference on Computer Vision (ICCV). 7-13 Dec. 2015, Santiago, Chile. New York: IEEE Press, 2015: 1026-1034.
- [17] Ding J, Chen S T. Hyper-spectral remote sensing image classification based on residual 3D convolutional neural network[J]. *Laser Journal*, 2019, 40(12): 45-52.
- 丁杰, 陈苏婷. 基于残差三维卷积神经网络的高光谱遥感图像分类[J]. *激光杂志*, 2019, 40(12): 45-52.
- [18] He M Y, Li B, Chen H H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification[C]//2017 IEEE International Conference on Image Processing (ICIP). 17-20 Sept. 2017, Beijing, China. New York: IEEE Press, 2017: 3904-3908.