

深度学习在高光谱图像分类领域的研究现状与展望

张号逵¹ 李映¹ 姜晔楠¹

摘要 高光谱图像 (Hyperspectral imagery, HSI) 分类是高光谱遥感对地观测技术的一项重要内容, 在军事及民用领域都有着重要的应用. 然而, 高光谱图像的高维特性、波段间高度相关性、光谱混合等使得高光谱图像分类面临巨大挑战. 近年来, 随着深度学习新技术的出现, 基于深度学习的高光谱图像分类在方法和性能上得到了突破性的进展, 为其研究提供了新的契机. 本文首先介绍了高光谱图像分类的背景、研究现状及几个常用的数据集, 并简要概述了几种典型的深度学习模型, 最后详细介绍了当前的一些基于深度学习的高光谱图像分类方法, 总结了深度学习在高光谱图像分类领域中的主要作用和存在的问题, 并对未来的研究方向进行了展望.

关键词 深度学习, 高光谱图像分类, 卷积神经网络, 栈式自编码网络, 深度置信网络

引用格式 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. 自动化学报, 2018, 44(6): 961–977

DOI 10.16383/j.aas.2018.c170190

Deep Learning for Hyperspectral Imagery Classification: The State of the Art and Prospects

ZHANG Hao-Kui¹ LI Ying¹ JIANG Ye-Nan¹

Abstract Hyperspectral imagery (HSI) classification occupies an important place in the earth observation technology of hyperspectral remote sensing, and it is widely used in both military and civil fields. However, due to HSI's characteristics including high dimensionality in data, high correlation between spectrum and mixing in spectrum, HSI classification faces great challenges. In recent years, as new deep learning technology emerges, the HSI classification methods based on deep learning have achieved some breakthroughs in methodology and performance and provided new opportunities for the research of HSI classification. In this paper, we review the research background, actuality of HSI classification technologies and several common datasets. Then, we provide a brief overview of several typical deep learning models. Finally, we introduce some deep learning based HSI classification methods in detail, summarize the main function and existing problems of deep learning in HSI classification, and present some prospects for future work.

Key words Deep learning, hyperspectral imagery (HSI) classification, convolutional neural network (CNN), stacked autoencoder, deep belief network

Citation Zhang Hao-Kui, Li Ying, Jiang Ye-Nan. Deep learning for hyperspectral imagery classification: the state of the art and prospects. *Acta Automatica Sinica*, 2018, 44(6): 961–977

经过 20 世纪后半叶的发展, 遥感技术在理论、技术及应用方面发生了重大的变化. 其中, 高光谱

图像 (Hyperspectral imagery, HSI) 遥感技术无疑是遥感技术重大变化中十分重要的一个方面. 高光谱图像遥感技术, 利用高光谱传感器/成像光谱仪以数十至数百个连续细分的波段对目标区域同时成像, 将图像与光谱结合, 同时获取空间及光谱信息, 得到以像素为单位的高光谱图像^[1-2]. 该图像含有大量的光谱信息, 能够采集目标在电磁波谱的紫外、可见光、近红外和中红外区域的辐射或反射信息^[1]. 而且, 随着高光谱成像系统的进一步发展, 高光谱图像采集到的信息将更加细致, 更加丰富, 一方面, 空间分辨率会更高; 另一方面, 光谱分辨率将显著提高^[1]. 与多光谱遥感等普通遥感图像相比, 高光谱图像波段多, 光谱波段覆盖范围广, 信息量巨大, 形成的三维数据块能够有效地反映成像目标的信息, 并有效地结合了空间信息和光谱信息^[3]. 因此, 高光谱遥感对地观测技术被广泛的应用在很多不同的

收稿日期 2017-04-14 录用日期 2017-09-23
Manuscript received April 14, 2017; accepted September 23, 2017

国家重点研发计划项目 (2016YFB0502502), 预研领域基金课题 (614023804016HK03002), 陕西省国际科技合作计划项目 (2017KW-006), 西北工业大学博士论文创新基金 (CX201816) 资助

Supported by National Key Research and Development Program (2016YFB0502502), Foundation Project for Advanced Research Field (614023804016HK03002), Shaanxi International Scientific and Technological Cooperation Project (2017KW-006), and Sponsored by Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (CX201816)

本文责任编辑 张军平

Recommended by Associate Editor ZHANG Jun-Ping

1. 西北工业大学计算机学院陕西省语音与图像信息处理重点实验室 西安 710129

1. Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, School of Computer Science, Northwestern Polytechnical University, Xi'an 710129

领域^[4-5], 例如矿业^[6]、天文学^[7]、化学成像^[8]、农业^[9]、环境科学^[10-11]、荒地火灾跟踪和生物威胁检测^[12]等. 随着高光谱成像仪的进一步发展, 高光谱图像含有的信息量会更加巨大, 高光谱图像的应用范围也会更加广泛. 不同场合的应用, 越来越大的数据量对高光谱遥感对地观测技术也提出了更复杂的要求. 高光谱图像分类技术是高光谱遥感对地观测技术的一项重要内容, 其具体任务是对高光谱图像中的每一个像素所代表的目标进行分类. 然而由于高光谱图像存在高维特性, 谱间相似性及混合像素, 高光谱图像分类技术面临一系列的挑战. 首先, 由于高光谱图像成像的光谱波段范围广, 波段窄, 波段数量多. 一方面高光谱图像相邻波段之间相关性较大, 存在较高的信息冗余^[1]; 另一方面, 对高光谱图像蕴含的丰富的谱信息直接进行处理所需计算量巨大, 对数据降维又需要选取合适的降维方法^[1], 由于高光谱图像类型多种多样, 不同的传感器采集到的数据也有所不同, 选取合适的降维方法是十分困难的; 其次, 高光谱图像中存在异物同谱及同谱异物问题, 导致数据结构呈高度非线性, 极大地增加了分类的难度, 一些基于统计模式识别的分类模型难以处理原始的高光谱图像数据. 再次, 与纳米级别的光谱分辨率相比, 高光谱图像的空间分辨率较低. 高光谱图像中存在一些混合像素, 即一个像素中存在多个不同类别的目标, 这对于高光谱图像分类, 也带来了一定的难度^[13-14]. 最后, 高光谱图像训练样本较少. 对采集到的高光谱图像进行标记十分困难, 实地勘测需要大量的人力物力, 并且需要耗费大量的时间, 高光谱图像的采集也会受到天气等环境因素的影响^[15-16]. 在早期的高光谱图像分类技术中, 高光谱图像分类方法仅仅利用了高光谱图像中丰富的光谱信息, 没有更深入地挖掘数据内在的信息. 例如, 距离分类器^[17]、K 近邻分类器^[18]、最大似然分类器^[19]、罗杰斯特回归^[20]等. 这些方法大多都会受到 Hughes 现象的影响^[21], 即在训练数据有限的情况下, 当数据维度过高, 分类的精度会大幅度下降. 针对这一问题, 研究者们提出了一些特征提取的方法^[22-24] 和判别学习的方法^[25]. 特征提取的方法例如主成分分析^[23]、独立成分分析^[24]、线性判别分析^[26-28], 将高维数据映射到一个低维空间, 同时保留类别信息. 判别学习的方法例如支持向量机^[25]、多类别的罗杰斯特回归^[26]、人工免疫网络^[29]等, 通过学习得到高维空间中的数据到类别的映射关系. 近几年, 高光谱图像分类方法中引入了高光谱图像的空间信息, 这一类方法简称为基于空谱联合特征的高光谱图像分类方法^[30-32]. 空间信息的引入, 显著提高了高光谱图像分类精度. 同样地, 为了有效联合空间信息和光谱信息完成高光谱图像分

类任务, 研究者们提出了一些能够提取低维度的空谱联合特征的特征提取方法^[33-34]. 无论是对于基于谱信息的分类方法来说, 还是对基于空谱信息的分类方法来说, 特征提取都是必不可少的, 其分类精度很大程度上依赖于提取到的特征. 然而, 高光谱图像类别复杂多样, 很难选择一种特征提取的方法能适用于所有类型的高光谱数据. 深度学习为特征提取提供了理想的解决方案, 针对多种多样的数据, 深度学习模型可以自主地从数据中学习特征, 舍弃了研究人员预先设计特征, 然后根据预先设计的特征提取规则从数据中提取特征的模式. 2006 年 Hinton 等提出逐层训练的思想^[35], 使得深度学习引起了学术界的广泛关注. 2012 年 Krizhevsky 等利用两个 GPU 训练了一个卷积神经网络 (Convolutional neural network, CNN) 对 Imagenet 数据库上的数据进行分类并取得了非常高的分类精度^[36], 这一成果在图像分类, 目标检测等领域掀起了深度学习研究的热潮. 近几年, 随着图像分类、目标检测等领域掀起了深度学习的热潮, 一些深度学习网络模型, 包括卷积神经网络 (CNN)^[36]、深度置信网络 (Deep belief network, DBN)^[35]、栈式自编码网络 (Stacked autoencoder, SAE)^[37] 等也被逐步引入到高光谱图像分类领域, 并取得了突破性的进展. 针对相关学术及技术进展情况, 本文主要综述及分析了基于深度学习的高光谱图像分类的最新进展.

1 高光谱图像分类问题及常用数据集

1.1 高光谱图像的特点

高光谱遥感技术, 通过采集不同目标反射或辐射出不同波长的电磁波, 得到以像素为单位的遥感影像, 在采集过程中, 会采集这些目标对上百个不同的连续波段的反射或辐射信息. 从波段强度的角度出发, 可以将高光谱图像看成一个三维的数据块, 以空间位置建立 Y 轴和 X 轴, 以波段强度关系建立 Z 轴. 如图 1 所示:

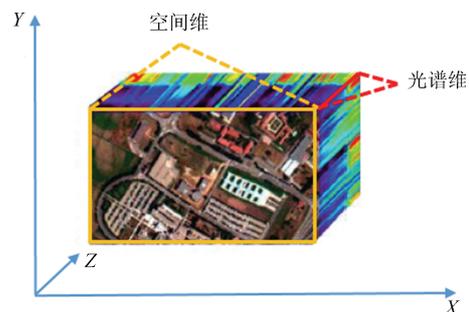


图 1 高光谱图像示意图

Fig. 1 Illustration of HSI

假设一个高光谱图像空间上含有 $M \times N$ 个像

素, 光谱维上含有 K 个波段, 则高光谱图像是由 K 个空间尺寸为 $M \times N$ 的二维图像组成的一个三维图像. 对于空间位置为 (i, j) 的一个像素点 $P(i, j)$, 其光谱特征为 $(p_1(i, j), p_2(i, j), \dots, p_k(i, j))$, 其中下标 $1, 2, \dots, k$ 表示波段编号.

1.2 高光谱图像分类问题的研究难点

高光谱图像中丰富的光谱信息为目标分类提供有力的依据的同时, 也为高光谱图像的分类引入了一些困难. 1) 高光谱图像数据量庞大, 导致高光谱图像相关处理的计算量很大, 如何快速精确地对待分类像素进行分类是研究高光谱图像分类时需要重点考虑的一个问题. 2) 比起纳米级别的光谱分辨率, 高光谱图像的空间分辨率较低, 图像中存在一些混合像素, 混合像素的存在给高光谱图像的精确分类, 也带来了一定的难度. 3) 由于高光谱图像成像的光谱波段范围广、波段窄、波段数量多, 导致相邻波段之间相关性较大, 高光谱图像中会存在较高的信息冗余, 对分类精度有影响, 如何降低冗余信息对高光谱图像分类的影响是非常重要的. 最后, 高光谱图像训练样本较少, 对于参数量比较多的分类方法来说容易出现过拟合问题, 如何在少量样本的情况下, 训练出高效的分类模型对高光谱图像分类来说十分重要.

从特征提取的角度来讲, 深度学习的相关模型为高光谱图像分类中存在的这些问题提供了一种解决思路, 然而, 将深度学习应用到高光谱图像分类中来也面临一些困难. 1) 数据维度问题, 与普通的自然图像不同, 高光谱图像数据呈现三维结构, 光谱信息十分丰富, 空间信息相对较少, 将深度学习应用到高光谱图像分类中的第一步是必须设计出适合高光谱图像结构, 能够有效利用高光谱图像丰富光谱信息及有限的空间信息的网络结构. 2) 样本量相对较少的问题, 深度学习的各种模型通常包含大量的参数, 这些参数是需要通过训练确定具体值的, 而训练过程依赖大量的有标注的样本数据. 3) 由于成像原理不同, 高光谱图像与其他图像的数据特点也有所

不同, 深度学习相关模型应用在其他图像方面时所积累的一些经验不一定适用于高光谱图像分类.

1.3 高光谱图像分类常用数据集

针对高光谱图像分类问题所提出的各种算法, 通常需要在一些高光谱图像数据集上进行验证, 表1中罗列了一些高光谱图像分类中常用的数据集. 其中 AVIRIS 指的是美国宇航局喷气推进实验室 (Jet Propulsion Laboratory, JPL) 的机载可见光、红外成像光谱仪 (Airborne visible/infrared imaging spectrometer, AVIRIS). ROSIS 指的是德国航空航天局的反射光学系统成像光谱仪 (Reflective optics system imaging spectrometer, ROSIS). 数据大小指的是高光谱图像空间维度上像素的个数, 样本量指的是高光谱图像的所有像素中有人工标记类别的像素的个数, 空间分辨率指的是每个像素实际对应的空间大小. 以数据集 Indian Pines 为例, 该数据集空间维度上一共有 145×145 个像素, 每个像素的实际覆盖面积为 $20 \times 20 \text{ m}^2$. 该数据中一共包含了 16 类不同的地物目标, 一共有 10 249 个像素有人工标注的类别.

2 深度学习典型的结构模型和开源工具

深度学习是机器学习的一个分支, 起源于人工神经网络. 相比传统的人工神经网络, 深度学习模型的层数更深, 抽象能力更强. 深度学习典型的结构模型包括卷积神经网络 (CNN)^[36]、基于自编码网络 (Autoencoder, AE) 的栈式自编码网络 (SAE)^[37]、基于受限玻尔兹曼机 (Restricted Boltzmann machine, RBM) 的深度置信网络 (DBN)^[35]、递归神经网络 (Recurrent neural networks, RNN)^[38]. 目前在高光谱图像分类领域有所应用的主要是 CNN、SAE 及 DBN 三种深度学习模型. 本节将概述这三种深度学习模型的主要思想和方法.

2.1 CNN 模型

与传统的神经网络需要以特征向量作为输入数

表1 高光谱图像分类常用数据集

Table 1 Several common datasets of HSI classification

数据	Indian Pines	Salinas	Kennedy Space Center	Pavia Center	Pavia University	Botswana
采集时间	1992 年	1992 年	1996 年	2001 年	2001 年	2001 年
采集地点	印第安纳州	加利福尼亚州	佛罗里达	意大利北部	意大利北部	奥卡万戈三角洲
采集设备	AVIRIS	AVIRIS	AVIRIS	ROSI	ROSI	Hyperion
光谱覆盖范围 (μm)	0.4×2.5	0.4×2.5	0.4×2.5	0.43×0.86	0.43×0.86	0.4×2.5
数据大小 (像素)	145×145	512×217	512×614	1096×492	610×340	1476×256
空间分辨率 (m)	20	3.7	18	1.3	1.3	30
波段数	224	224	224	115	115	242
去噪后波段数	200	204	176	102	103	145
样本量	10 249	54 129	5 211	7 456	42 776	3 248
类别数	16	16	13	9	9	14

据不同, CNN 可以直接输入二维的图像进行处理. CNN 在计算机视觉领域的应用十分广泛, 例如人体行为识别^[39]、图像分类^[36, 40-41]、目标检测^[42-43]、目标跟踪^[44], 从单一图像中进行景深估计^[45]、超分辨率重建^[46-47]. CNN 主要包含三个部分: 卷积层 (Convolution layer)、池化层 (Pooling layer)、全连接层 (Fully connected layer). 全连接层的输出可以看作是卷积网络抽取到的特征, 具体结构如图 2 所示, 图中 C 代表卷积层, P 代表池化层, F 表示全连接层. 在卷积层, 利用多个共享的卷积核对输入图像进行卷积操作 (同一个卷积核对输入图像在不同位置进行卷积操作时, 卷积核的参数不改变) 得到多张特征图, 每个特征图的维度会依照卷积核的尺寸大小有一定的降低. 在池化层, 利用取最大值池化及取平均值池化等操作对输入特征进一步压缩, 经过池化操作, 特征的空间尺寸会减半, 池化层可以保证特征对旋转的不变性并降低特征维度. 网络的输入数据在经历多个卷积层及池化层之后, 最后经过全连接层得到输出. 网络训练时采用反向传播 (Back propagation, BP) 算法^[36] 逐层反向传递误差, 之后利用随机梯度下降 (Stochastic gradient descent, SGD) 算法^[36] 更新各层的卷积核和偏置.

随着数据的维度和数量的增加, CNN 处理的具体任务的相关难度也在提升. 以“卷积层-池化层-全连接层”结构的 CNN 模型为基础, 学者们陆续提出了一些非常有效地改进的结构和训练方法. 在模型结构方面, 改进的模型主要有: Network in network (NIN)^[48]、GoogleNet^[49]、VGGNet^[50]、SPPNet^[40]、Residual networks (ResNet)^[51] 及 Fully convolutional networks (FCN)^[52] 等. 其中, NIN 在传统的卷积层中嵌入了一个微小网络, 将传统的卷积层改进为多层感知卷积层 (Multilayer perceptron convolution layer) 增强了卷积层的计算拟合能力. 同时, 为了解决全连接层参数过多的问题, 提出了全局均值池化操作 (Global average pooling).

GoogleNet 提出了 Inception module, 每个 Inception module 会包含多个分支, 最终整个网络通过堆叠多层 Inception module 得到 Inception module 的各种版本还在持续更新中. VGGNet 采用空间尺寸为 3×3 的卷积核构建了相对较深的卷积神经网络. SPPNet 提出了空间金字塔池化层 (Spatial pyramid pooling), 在全连接层之前加入空间金字塔池化层, 在网络结构不变的情况下, 能够对不同尺度的图像进行处理. ResNet 通过堆砌 Residual unit 构建了非常深的卷积神经网络, 最深的网络已经达到一千多层, 其中 Residual unit 采用了 Shortcut connection 结构. FCN 中, 舍弃了全连接层, 引入了反卷积层, 网络能够接受任意大小的输入图像, 主要应用在图像分割方面. 在网络训练方面, 提出的技巧主要有 Data augmentation^[36]、Drop-connect^[53]、Drop-out^[36]、Local response normalization^[36]、Batch normalization^[54]、Pre-training 及 Fine-tuning^[55] 等. 其中 Data augmentation 是针对训练数据进行扩充, 可以有效地改善训练数据不足时导致的过拟合问题, 主要包括裁剪、旋转、缩放、翻转等操作. Drop-connect 具体操作是在训练过程中, 以预先设定的概率舍弃部分连接权. Drop-out 具体操作是在训练过程中, 以预先设定的概率将部分节点输出值置零. Local response normalization 和 Batch normalization 都是对数据进行归一化操作, 前者是局部范围内进行归一化, 后者是计算整个 Batch 里边的数据的均值及方差并进行归一化操作. Pre-training 与 Fine-tuning 对应, 前者可以用来初始化网络, 具体指的是可以在目标任务数据以外的含有大量有标签数据的数据集上训练得到一个网络, 之后利用该网络的参数初始化针对特定任务的网络. 后者指的是利用 Pre-training 初始化网络之后, 在任务目标的数据集上对网络进行微调, 以达到最优的效果.

2.2 SAE 模型

SAE 是由多个 AE 堆建起来的, 结构如图 3 所

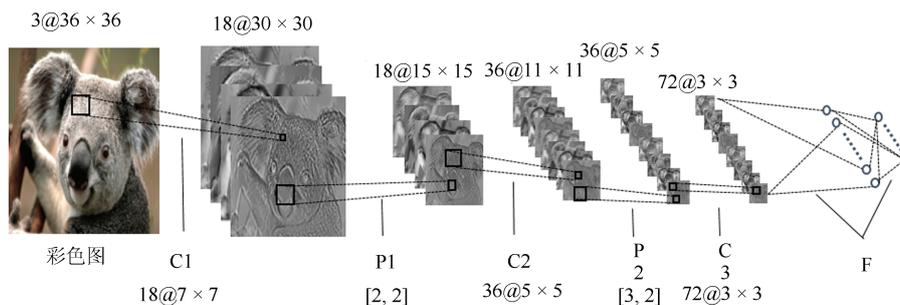


图 2 CNN 网络结构

Fig. 2 The structure of CNN

示. 网络前向运算时, 输入数据从底层 AE 输入网络, 逐层运算最后得到网络的输出. 网络的训练过程按顺序分为预训练 (Pre-training) 和微调 (Fine-tuning) 两个阶段. 预训练阶段, 从底层开始每个 AE 单独训练, 训练的目标是最小化 AE 的输入层和重构层的差异, AE 无监督训练结束后移除重构层的部分, 并将隐层的输出数据作为高一层 AE 的输入数据并对高一层 AE 进行训练, 重复该过程直到网络整体达到预定的深度. 预训练阶段, SAE 是逐层训练的, 且不需要用到任何标签信息, 属于无监督学习 (Unsupervised learning). 微调阶段, 以预训练阶段得到的各层的权值及偏置作为 SAE 网络的初始权重和偏置, 以数据的标签 (标注信息) 作为监督信号, 利用 BP 算法计算各层误差, 并利用 SGD 算法完成各层权重及偏置的更新.

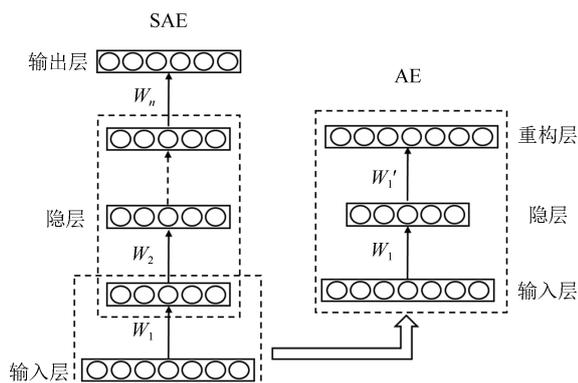


图 3 SAE 网络结构

Fig. 3 The structure of SAE

2.3 DBN 模型

类似于 SAE, DBN 是由多个 RBM 逐层构建起来的, 结构如图 4 所示, 正向运算过程与 SAE 网络的正向运算过程类似, 训练过程也分为预训练和微调两个阶段. 预训练阶段, 从底层开始每个 RBM 单独训练, 训练目标为最小化 RBM 网络能量, 在可见层和隐层之间构建稳定的映射关系, 训练采

用无监督的对比散度算法 (Contrastive divergence, CD)^[56]. 低层的 RBM 训练完成之后, 将其隐层的输出作为高一层 RBM 的输入, 再训练高一层的 RBM, 重复该过程, 直到构建出合适的网络规模. 微调阶段, 将预训练阶段确定的各层的权值及偏置作为初始化的权值及偏置, 以数据标签为指导信号计算网络误差, 利用 BP 算法计算各层误差, 最后利用 SGD 算法完成各层权值及偏置的更新.

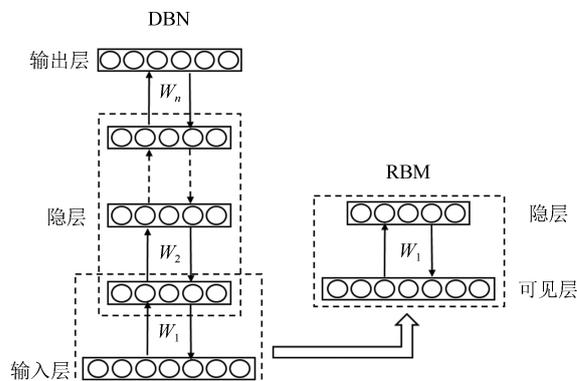


图 4 DBN 网络结构

Fig. 4 The structure of DBN

2.4 典型开源工具简介

近年来, 深度学习发展十分迅速, 各种优秀的开发工具也层出不穷, 其中包括 Tensorflow、Theano、Pytorch、Caffe 及 CNTK 等. 适用于深度学习开发的编程语言也有很多种, 包括 Python、C++、Go、Matlab 等. 依照开发工具所支持的编程语言, 表 2 总结了目前的一些应用得比较广泛的开发工具.

3 深度学习在高光谱图像分类领域的研究现状

高光谱图像同时包含了空间信息和光谱信息, 光谱信息的分辨率很高, 一般都达到了纳米级别, 空

表 2 几种主流的深度学习开发工具

Table 2 Several mainstream development tools of deep learning

工具	机构	支持语言	官网
Tensorflow	Google	Python/C++/Go/Java	https://www.tensorflow.org/
Theano	U Montreal	Python	http://deeplearning.net/software/theano/
Pytorch	Facebook	Python	http://pytorch.org/
Caffe	BVLC	C++/Python/Matlab	http://caffe.berkeleyvision.org/
CNTK	Microsoft	C++/Python/C	http://cntk.codeplex.com/
Matconvnet	/	Matlab	http://www.vlfeat.org/matconvnet/
MXNet	DMLC	Python/C++/R/Julia/Scala/Go/Matlab/JavaScript	http://mxnet.io/index.html
Torch	Facebook	Lua	http://torch.ch/
Deeplearning4J	DeepLearning4J	Java/Scala	https://deeplearning4j.org/

间分辨率相对较低. 高光谱图像分类问题是一个像素级别的分类问题, 具体而言就是对图像中的每一个像素进行分类确认像素的类别. 高光谱图像分类大体上可以分为无监督分类和有监督分类 (包含半监督分类) 两种. 无监督分类指的是在没有预先标注数据标签的情况下对高光谱图像进行分类 (聚类), 主要思想是依照能代表像素特点的特征信息 (空间信息、谱信息及特征等) 将相似的像素归为一类. 有监督的分类指的是在有预先标注数据作为监督信号的情况下对高光谱图像进行分类, 主要思想是利用有标注的数据学习像素特征信息与像素类别之间的内在关系, 然后利用这种关系对没有标注的数据进行分类, 确定像素类别. 目前基于深度学习的高光谱图像分类主要属于有监督分类.

3.1 CNN 在高光谱图像分类中的应用

CNN 在图像分类、目标检测、场景分割等方面应用的十分广泛, 并取得了非常显著的成果. 与 SAE 及 DBN 相比, CNN 是目前在高光谱图像分类领域应用最多的深度学习网络模型. 基于 CNN 的高光谱图像分类方法主要可以分为三类: 基于谱特征的分类方法、基于空间特征的分类方法及基于空谱联合特征的分类方法.

3.1.1 基于谱特征的分类方法

高光谱图像同时具有空间信息和光谱信息, 其中, 光谱分辨率较高且光谱信息十分丰富. 高光谱图像数据呈三维结构, 从每一个像素所在的空间位置都可以提取出一条由谱信息组成的一维谱向量, 仅利用该一维谱向量的信息来确定像素的类别即为基于谱特征的分类. 传统的高光谱图像分类方法中, 有大量的基于谱特征的分类方法, 例如距离分类器^[17]、K 近邻分类器^[18]、最大似然分类器^[19]等. 在这些传统的基于谱特征的分类方法中, 分类模型所利用的谱特征一般是原始的从像素位置提取到的光谱信息或者是经过一定的预先设定的特征提取方法从原始光谱信息提取出来的谱特征. 而基于 CNN 的高光谱图像谱特征分类则是利用一维 CNN

(1D-CNN) 自主地从原始的光谱信息中提取深度光谱特征并完成高光谱图像的分类, 具体如图 5 所示.

将部分标注数据 (标注了像素所属类别) 的谱信息提取出来作为 1D-CNN 的输入, 类别标签作为监督信号对 1D-CNN 进行训练, 网络的损失函数一般通过多元罗杰斯特回归^[57]来计算, 并通过 SGD 等算法对网络的权值进行迭代更新, 最终利用训练好的网络对图像中所有像素进行分类得到分类结果. 传统的 CNN 都是处理具有二维结构的图像的二维卷积神经网络 (2D-CNN), 例如 AlexNet^[36]、VggNet^[50] 及 GoogLeNet^[49] 等. 二维卷积操作与一维卷积操作的主要区别在于卷积操作和池化操作的维度, 二维卷积操作使用二维的卷积核在二维数据上进行卷积操作, 具体公式为

$$\text{map}_{l,j}^{x,y} = f \left(\sum_m \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} k_{l,j,m}^{h,w} \text{map}_{(l-1),m}^{(x+h),(y+w)} + b_{l,j} \right) \quad (1)$$

式中, $k_{l,j,m}^{h,w}$ 表示第 l 层第 j 个卷积核在 (h, w) 位置的具体值, 该卷积核与 $(l-1)$ 层网络中第 m 个特征图相连. H_l 和 W_l 分别表示卷积核的高和宽, $b_{l,j}$ 表示第 l 层第 j 个特征图的偏置. $\text{map}_{(l-1),m}^{(x+h),(y+w)}$ 表示第 $(l-1)$ 层中第 m 个特征图在 $(x+h, y+w)$ 位置的具体值. $\text{map}_{l,j}^{x,y}$ 分别代表第 l 层第 j 个特征图在 (x, y) 位置的输出数据. 一维卷积操作使用一维的卷积核在一维的特征向量上进行卷积操作, 其表达式为

$$v_{l,j}^x = f \left(\sum_m \sum_{h=0}^{H_l-1} k_{l,j,m}^h v_{(l-1),m}^{(x+h)} + b_{l,j} \right) \quad (2)$$

其中, $k_{l,j,m}^h$ 表示第 l 层第 j 个卷积核在位置 h 处的具体值, 该卷积核与第 $(l-1)$ 层网络中第 m 个特征向量相连. H_l 表示一维卷积核的长度. 与卷积操作对应, 二维池化操作池化块是二维的, 一维池化操作, 池化块是一维的. 2D-CNN 卷积网络和

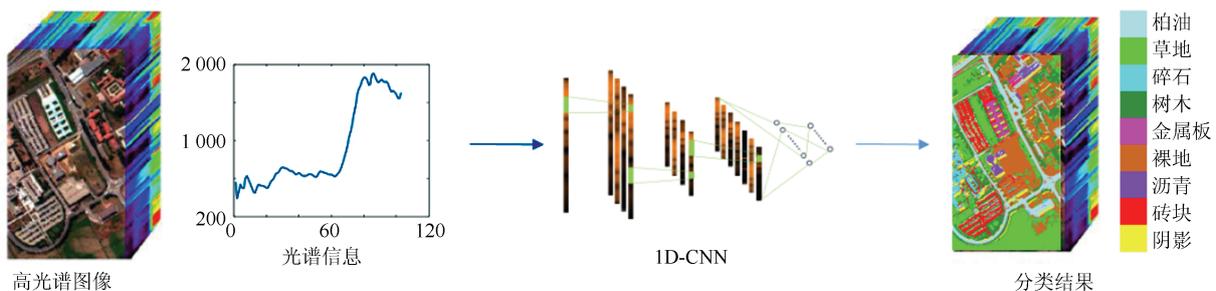


图 5 基于 1D-CNN 的高光谱图像谱特征分类

Fig. 5 Spectral feature classification of HSI based on 1D-CNN

1D-CNN 的激励函数, 全连接部分及训练方法都是相同的。

CNN 发展十分迅速, 网络的具体结构灵活多变. 为了防止过拟合, 梯度扩散等问题, 在设计网络具体结构时, 研究者们经常会采用很多技巧. 文献 [58] 采用了含有一层卷积层一层池化层两层全连接层的 1D-CNN 进行高光谱图像分类, 网络的激励函数选取了双曲正切函数, 分类部分选择了 Softmax^[59]. 文献 [60] 设计了含有一层卷积层, 两层全连接层的 1D-CNN, 利用 Batch normalization^[54] 对卷积层的输出进行了标准化操作. 利用 Batch normalization 使得网络模型在训练过程中可以选择较大的学习率, 提升了网络训练的速度. 此外, 为了防止过拟合, 文献 [60] 中的模型还在第一层全连接层中使用了 Dropout^[36] 并采用 PReLU (Parametric rectified linear unit) 函数^[61] 作为激励函数.

3.1.2 基于空间特征的分类方法

空间信息, 也称为上下文信息, 在高光谱图像分类中, 合理地利用空间信息能很好地提升分类精度. 在利用光谱信息进行高光谱图像分类时, 通常只考虑待分类像素的光谱信息, 没有很好地考虑像素和像素之间的依赖关系. 像素和像素之间的依赖关系指的是位置相近的像素之间的类别很大程度上会有一些的关联, 例如, 有一些空间位置上联通的像素被分类为停车场, 那么这些像素中间出现的光谱信息与金属光谱信息非常相似的像素很有可能是汽车. 如果一个像素周边很多像素都是草地, 那么位于中间的像素很可能也是草地^[62]. 基于空间特征的高光谱图像分类方法, 简单来说就是输入的信息不再是从单个像素位置所提取的光谱信息, 而是从一个邻域范围内提取出来的光谱信息. 提取一个待分类像素的空间信息通常的做法是, 先对高光谱图像进行压缩, 再从压缩之后的数据上, 提取以待分类像素为中心的一个邻域范围内的所有信息. 这样的空间信息提取做法会保留待分类像素的空间信息, 但是会损失一定的光谱信息, 本文中将其称为基于空间特征的分类方法. 在基于 CNN 的高光谱图像空间特

征分类方法中, 提取出来的邻域信息会保持其原始的空间结构, 之后利用 CNN 对提取出来的信息进行更深层次加工, 提取出更深层次的空间特征, 最后利用提取出来的深层次的空间特征完成高光谱图像分类, 具体框架如图 6 所示.

CNN 最初的模型便是应用在二维结构的数据上, 高光谱图像的三维结构是将传统的 2D-CNN 应用到高光谱图像分类方面的最大阻碍. 2D-CNN 的卷积层需要对每层的二维特征图进行卷积, 针对每一特征图都需要一个卷积核 (网络参数), 直接将 2D-CNN 应用到三维高光谱图像上时, 由于高光谱图像有上百个波段 (输入数据包含上百个二维特征图), 2D-CNN 将需要很多的参数. 高光谱图像标注数据相对较少, 在参数过多的情况下很容易过拟合, 而且, 过大的网络规模也会增加运算成本.

针对这一问题, 研究者们通常会利用数据降维的方法在光谱维上对高光谱图像数据进行降维. 例如文献 [63] 采用了 R-PCA (Principle component analysis) 对高光谱图像进行降维, 保留前 10 或 30 主成分, 并利用一个有三层卷积层的 CNN 从空间尺寸为 5×5 的邻域提取空间特征并完成了高光谱图像的分类. 文献 [64] 采用 PCA 降维, 并保留前三主成分, 之后利用 CNN 从空间尺寸为 42×42 的邻域提取空间特征, 与文献 [63] 方法不同的是, 文献 [64] 所采用的 CNN 有两层池化层且所选取的邻域尺寸更大.

先对高光谱图像进行光谱维度上的压缩降维, 之后直接利用 CNN 从待分类像素的邻域范围内提取空间特征完成分类, 是比较简单直观的基于 CNN 的高光谱图像空间特征分类方法. 除此之外, 研究者们还提出了很多其他方法, 将 CNN 与其他模型相结合提取空间特征. 文献 [65] 首先利用 PCA 降维将高光谱图像的光谱维压缩到三维, 之后再利用拉普拉斯金字塔变换得到多种空间尺寸的压缩数据, 最后每种尺寸都采用 CNN 提取特征. 文献 [66] 利用 PCA 对高光谱图像进行压缩, 再利用属性剖面 (Attribute profiles, APs)^[67] 对数据进一步加工, 最终利用 CNN 从空间尺寸为 42×42 的邻域提取特

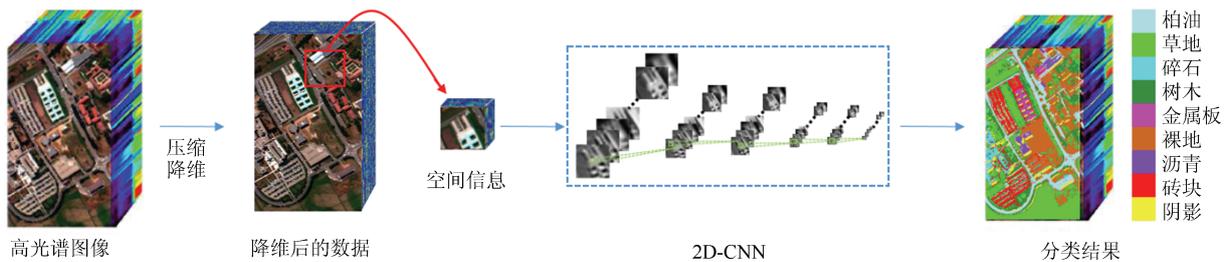


图 6 基于 2D-CNN 的高光谱图像谱特征分类

Fig. 6 Spatial feature classification of HSI based on 2D-CNN

征完成分类. 文献 [68] 在利用 CNN 从经过 PCA 降维后的高光谱图像数据中提取空间特征之后, 利用稀疏编码对 CNN 提取出来的空间特征进行了进一步的加工, 最后基于稀疏编码得到的特征完成分类任务. 文献 [69] 也应用了 2D-CNN 网络来完成高光谱图像分类, 但是 2D-CNN 并不是用来提取空间特征, 而是用来对高光谱图像进行重建, 通过重建提升了不同类别的可区分性, 以此促进高光谱图像的分类精度.

3.1.3 基于空谱联合特征的分类方法

三维结构的高光谱图像数据, 既包含了空间信息, 又包含了丰富的光谱信息. 基于光谱特征或基于空间特征都没有充分发挥高光谱图像的优势, 与这两种方式相比, 利用空间信息和谱信息相结合的方式对高光谱图像分类更符合高光谱的特点, 也更有利于提高高光谱图像分类的精度. 基于 CNN 的高光谱图像空谱联合分类方法在空间信息和光谱信息的联合上主要有两种方式, 一种是基于深度网络模型分别提取空间特征和光谱特征, 之后再

将光谱特征和空间特征结合起来完成分类, 如图 7 所示; 另一种方式是同时提取空间特征和光谱特征, 通常是利用三维结构的模型例如三维卷积网络 (3D-CNN) 直接对三维结构的数据块进行处理, 如图 8 所示.

基于 CNN 分别提取谱特征和空间特征的分类方法中, 空间特征的提取一般是通过降维方法将高光谱图像降维, 之后利用 2D-CNN 从一定领域内提取空间特征. 谱特征的提取方法比较多, 可以利用深度学习的模型也可以利用非深度学习的模型. 文献 [70–71] 均利用 2D-CNN 提取从降维的高光谱图像中提取空间特征, 利用 1D-CNN 从光谱向量中提取谱特征, 最后通过全连接层及 Softmax 完成分类. 为了缓解标注数据过少引起过拟合, 文献 [71] 利用翻转、旋转操作将训练数据扩充到原来的 6 倍. 文献 [72] 利用含有空间金字塔池化 (Spatial pyramid pooling)^[40] 层的 2D-CNN 提取空间特征, 利用 SAE 网络提取谱特征. Zhao 等^[73] 对局部鉴别嵌入 (Local discriminant embedding, LDE)^[74] 进行了改进, 提出了平衡局部鉴别嵌入 (Balanced local discriminant embedding, BLDE) 算法并利用该算

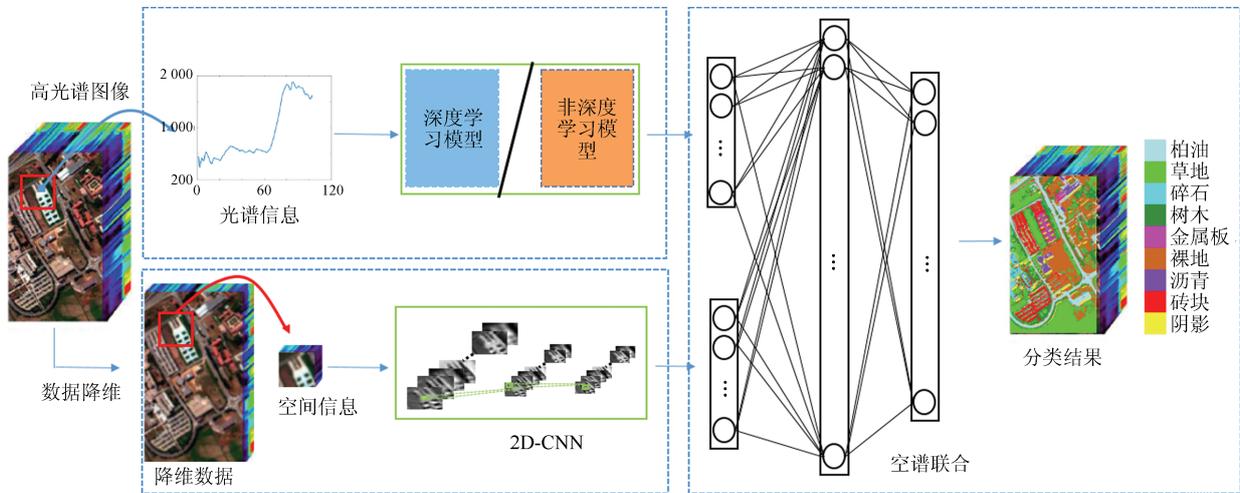


图 7 基于谱特征和空间特征分别提取的分类流程图

Fig. 7 The flow chart of classification based on extracting spectral and spatial feature separately

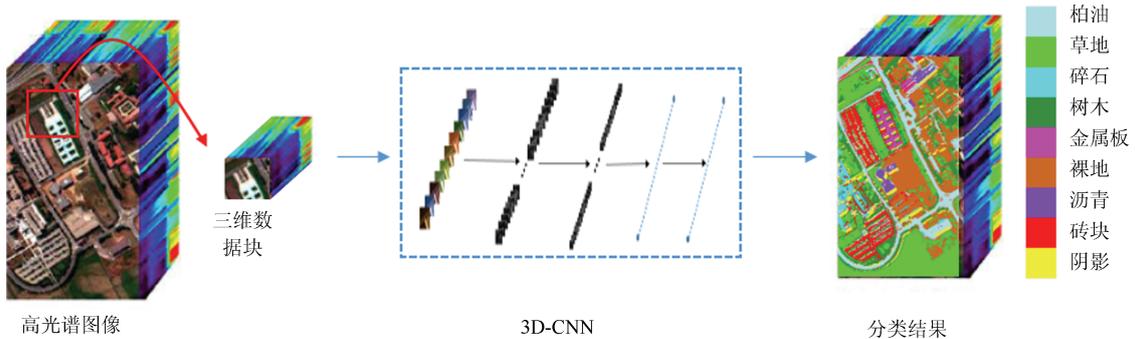


图 8 基于 3D-CNN 同时提取谱特征和空间特征的分类流程图

Fig. 8 The flow chart of classification based on 3D-CNN extracting spectral and spatial feature simultaneously

法提取了光谱特征. 最终联合 BLEED 算法提取的光谱特征和 2D-CNN 提取的空间特征, 完成了高光谱图像分类. 基于 CNN 同时提取高光谱图像空谱联合特征时, 高光谱数据不必通过 PCA 等降维方法进行压缩, 利用合适的 CNN 模型直接对三维的数据块进行特征提取. 这些方法所使用的 CNN 模型可以分为 3D-CNN、2D-CNN 及 1D-CNN.

文献 [75–76] 均采用了 3D-CNN, 在未降维的高光谱图像数据上, 以待分类像素为中心, 提取一定邻域范围内的所有像素的谱信息, 组成三维的数据块, 利用 3D-CNN 对三维的数据块进行特征提取. 文献 [76] 用一种结构比较紧凑, 网络规模较小的 3D-CNN 从空间尺寸为 5×5 的数据块中提取深度空谱联合特征. 文献 [75] 利用规模较大, 结构较为复杂的 3D-CNN 从空间尺寸为 27×27 的数据块中提取深度空谱特征. 与 2D-CNN 相比, 3D-CNN 的卷积核是三维的, 而且卷积的方向从空间上的二维扩展到了空间加光谱的三维, 如图 9 所示:

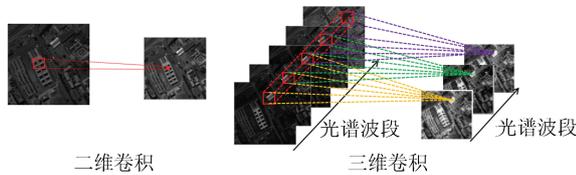


图 9 二维卷积和三维卷积

Fig. 9 2D convolution operation and 3D convolution operation

三维卷积操作中卷积层第 l 层, 第 j 个特征立方体 (x, y, z) 位置的具体值可以通过如下公式计算:

$$v_{ij}^{xyz} = f \left(\sum_m \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} \sum_{r=0}^{R_l-1} k_{ljm}^{hwr} v_{(l-1)m}^{(x+h)(y+w)(z+r)} + b_{lj} \right) \quad (3)$$

其中, R_l 是三维卷积核在光谱维的尺寸, H_l 和 W_l 分别是三维卷积核的高和宽. j 是当前层卷积核的编号, k_{ljm}^{hwr} 是第 l 层第 j 个卷积核位于 (h, w, r) 位置的具体值, 该卷积核与第 $(l-1)$ 层第 m 个特征立方体相连.

文献 [77–78] 均采用了 2D-CNN 来同时提取空谱特征. 具体而言, 文献 [77] 首先利用空间尺寸分别为 3×3 和 1×1 的三维卷积核 (卷积核在光谱维度上的尺寸与数据波段数保持一致) 对提取的三维数据块进行二维卷积操作 (只在空间上进行卷积, 光谱方向上无卷积). 再将两种不同尺寸的卷积核提取出来的特征进行连接组成新的三维数据 (Inception module), 最后利用空间尺寸为 1×1 的三维卷

积核进行后续的卷积操作. 该方法在网络结构上借鉴了 ResNet 的结构^[51], 利用了残差学习 (Residual learning) 中所提出的计算模块. 为了防止标注数据过少引起过拟合, 作者还利用水平翻转及沿对角线翻转操作将训练数据扩充到原来的 4 倍. 文献 [78] 中, Slavkovikj 等从未降维的高光谱数据中提取空间尺寸为 3×3 的数据立方体, 然后将该数据立方体重新排列成矩阵, 在第一层卷积层中采用二维卷积核沿谱方向进行卷积操作, 之后再用两层一维卷积核沿谱方向进行卷积操作, 卷积层之后连接了两层全连接层.

3.2 SAE 在高光谱图像分类中的应用

SAE 网络由 AE 模块堆建而成, 根据 AE 模块的不同呈现出来的 SAE 网络会有所不同, 例如对 AE 模块有稀疏性要求时, SAE 网络也会变成稀疏的 SAE. AE 通常是采用无监督的形式进行训练的, 根据具体训练方式的不同, 堆建起来的 SAE 在性能方面也会有所不同. 基于 SAE 网络的高光谱图像分类方法大体也可以分为基于谱特征的分类方法, 基于空间特征的分类方法及基于空谱联合特征的分类方法, 如图 10 所示.

基于 SAE 进行高光谱图像谱特征分类时, 通常将从待分类像素位置提取出来的光谱向量作为网络的输入数据, 利用 SAE 提取更深层次, 更抽象的深度谱特征, 之后基于深度谱特征完成分类任务. 文献 [79, 31] 均采用了 SAE 提取深度谱特征, AE 在训练的过程中添加了约束条件, 从输入层到隐层及从隐层到重构层的连接权是转置的关系, 网络最后调优阶段选择了交叉熵 (Cross entropy) 作为损失函数.

文献 [80] 中, Ma 等采用了 SAE 对高光谱图像进行了基于谱特征的分类, 但是对传统 SAE 中的损失函数进行了改进, 传统的损失函数一般都定义为分类误差, 通过优化损失函数使得分类误差最低, 而该文将优化目标函数定义为在特征空间中同类像素距离更近, 不同类像素距离更远. 文献 [81] 及文献 [82] 均采用了栈式去噪自编码网络 (Stacked denoising autoencoder, SDAE)^[83] 提取谱特征, 并进行基于谱特征的分类. 文献 [82] 在利用谱特征分类之后, 还利用谱特征的相似性和空间位置关系对原图进行超像素分割, 对每个超像素进行投票, 依照超像素里边多数像素点的类别确定该超像素的类别, 得到最终分类结果.

SDAE 是 SAE 的一种改进形式, SDAE 是由去噪自编码模块 (Denoising autoencoder, dAE)^[79] 构成的. 训练方法及构建方法均与 SAE 相同, 不同的地方是 dAE 中在从输入层到隐层时, 会对输入数

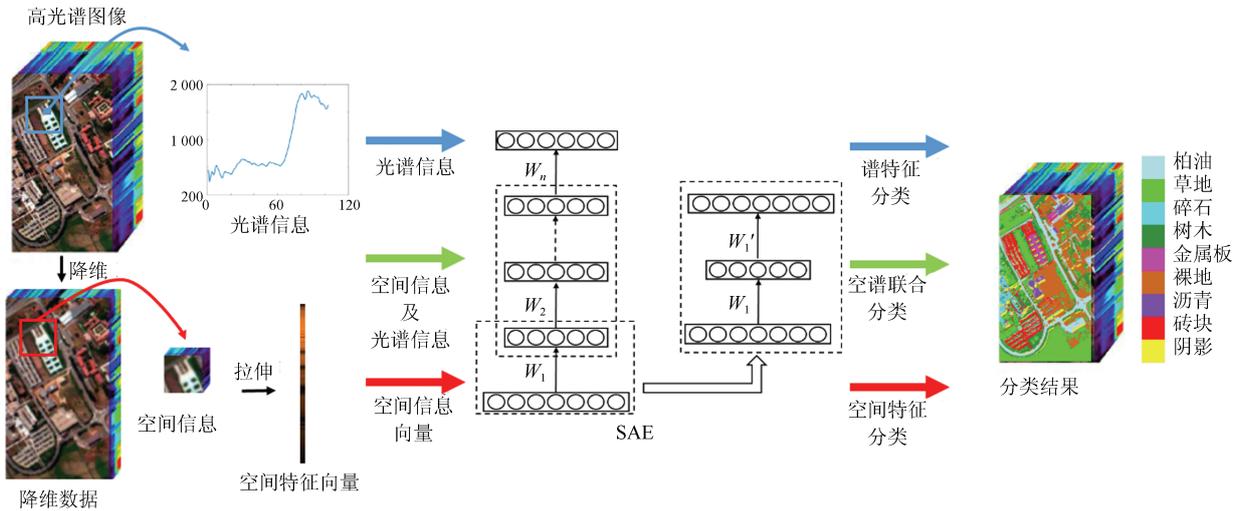


图 10 基于 SAE 的高光谱图像分类
Fig. 10 HSI classification based on SAE

据加入噪声 (以一定概率将数据层节点置零), 具体如图 11 所示.

基于 SAE 进行高光谱图像空间特征分类时, 主要做法是先对高光谱图像进行降维, 之后以待分类像素为中心提取邻域范围内的所有信息, 并将提取出的信息拉成向量, 作为 SAE 的输入数据. SAE 不能直接处理空间上呈二维结构的数据块, 为了满足 SAE 对输入格式的要求, 需要将数据拉伸成一维的向量, 这点与基于 CNN 的高光谱图像空间特征分类方法区别较大.

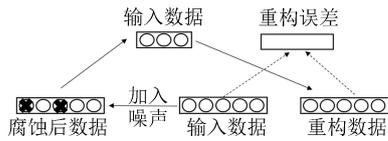


图 11 去噪自编码示意图
Fig. 11 The illustration of dAE

利用 SAE 提取空间特征时也需要对高光谱图像进行光谱维的降维, 不降维直接将待分类像素一定邻域范围内的所有信息都提取出来作为 SAE 的输入数据会导致输入数据维度过高^[80]. SAE 的隐层虽然没有强制性地限制隐节点的个数, 但是隐节点的个数不能过分小于输入数据的维度, 否则会导致从输入层到隐层丢失太多信息. 因此输入数据维度过高还可能导致网络整体参数过多, 不利于网络的收敛及训练. 降维的方法比较多, 文献 [79, 31] 均是利用 PCA 降维, 将高光谱图像分别压缩到 3 维和 4 维, 之后从分类像素的邻域内提取空间信息, 利用 SAE 进一步提取深度空间特征. 在预训练过程, 限制 AE 的隐层到重建层的连接权输入层到隐层连接权的转置. 在调优过程, 选择了交叉熵作为准则函数.

基于 SAE 进行高光谱图像空谱联合分类时, 通常的做法是分别提取空间信息和光谱信息, 最后合并得到空谱联合信息. 从待分类像素的空间位置提取出的光谱向量, 在高光谱图像压缩图上从邻域提取空间信息并将空间信息拉成向量, 合并光谱向量和由空间信息拉成的向量. 将合并得到的信息输入 SAE 中, 借助 SAE 提取深度空谱特征并完成分类.

SAE 网络的输入数据必须是特征向量, 因此在空谱联合分类的方法中, 无论是光谱信息还是空间信息都需要以特征向量的形式输入到网络中去. 光谱信息本身就是向量的形式无需此操作, 而空间信息本身是二维或三维的结构, 因此需要向量化. 直接将邻域提取出来的空间信息拉成向量是最简单的方法, 文献 [31, 84] 中空间特征提取部分就是如此. 文献 [85] 采用了 SDAE 网络提取空谱特征, 而且在训练过程中加入了相似性归一化项, 即原始的比较相似的光谱数据在 SDAE 的隐层中应该比较相似, 同时考虑了编码后再解码的数据恢复的质量; 然后, 采用协同表示进行分类; 最后, 再利用图分割的方法对协同表示分类结果进行分割得到最终分类结果. 文献 [86] 中, Tao 等首先利用 PCA 降维将高光谱图像压缩到一维, 然后以待分类像素为中心, 提取 3×3 、 5×5 及 7×7 三个邻域范围的信息一同作为空间信息, 并利用栈式稀疏自编码 (Stacked sparse autoencoder, SSAE)^[87] 对其进行特征提取. 文献 [88] 也考虑了多尺度的空间信息, 而光谱信息是从降维之后的图像中提取出来的. 文献 [89] 首先利用 PCA 降维将高光谱图像压缩到 50 维, 之后利用三维 Gabor 小波变换从降维后的图像中提取特征, 最后再将光谱向量与小波变换提取到的特征联接起来, 利用 SAE 提取更深层次的空谱特征. 文献 [90] 两

次引入空间信息, 提取特征过程中, 利用 SAE 提取特征时, 输入数据是一定邻域范围内的空间信息. 分类过程中, 在分类之前将特征图一定邻域范围内的数据提取出来进行信息融合, 最后利用分类器对融合的特征进行分类.

除了利用 SAE 分别提取谱特征和空间特征, 之后将空谱特征结合进行分类的方法之外, 还有一些利用 SAE 进行高光谱图像分类的方法. 文献 [91] 中将 AE 网络与 CNN 结合在一起, 首先, 提取以待分类像素为中心的一个大邻域块 (7×7) 内的数据; 之后, 在大邻域块内提取小邻域块 (4×4) 的数据, 利用 SAE 对小邻域块的数据进行编码, 将编码的结果作为卷积核对大邻域块进行卷积操作来提取空谱特征; 最后, 利用 Softmax 分类器进行分类. Ma 等 [92] 提出了一种半监督的基于 SAE 的分类方法, 利用局部判断及全局判断两种机制扩充训练样本集, 其中全局判断及最终的分类器都应用了 SAE 网络.

3.3 DBN 在高光谱图像分类中的应用

DBN 与 SAE 类似, 都是逐层堆建起来的网络模型, 堆建 SAE 用的是 AE 模块, 而堆建 DBN 采用的是 RBM 模块. 与 CNN 及 SAE 相比, DBN 在高光谱图像分类方面应用的较少. 同样的, 基于 DBN 的高光谱图像分类方法主要可以分为三种不同的方法, 基于谱特征的分类方法 [93-94]、基于空间特征的分类方法 [95] 及基于空谱联合特征的分类方法 [32, 96], 如图 12 所示.

与基于 SAE 的高光谱图像分类方法类似, 在利用 DBN 进行高光谱图像谱特征分类时, 主要做法是利用 DBN 对从待分类像素位置采集到的光谱信息进行更深层次的特征提取, 之后利用深度特征完

成分类. He 等 [93] 提出了一种新型的模块并将该模块堆建起来构成深度网络 (Deep stacking network, DSN), 每层模块包含输入层、隐层和输出层. 输入层与隐层之间的连接权可以通过对比散度算法初始化或者随机初始化, 然后固定下来, 输出层直接就是类别标签. 隐层与输出层间的连接权可以通过计算伪逆得到, 该过程不依赖于 SGD 算法, 最终堆建起来的网络在末尾添加罗杰斯特回归分类器后, 整体利用 SGD 算法进行调优. Zhong 等 [94] 考虑到在小样本条件下, 基于 DBN 进行高光谱图像分类时, DBN 中存在大量的死节点 (不响应的节点) 和过分活跃节点 (响应频率非常高的节点), 对传统的 DBN 进行了改进, 在 DBN 预训练及调优阶段引入了多样性先验项 (Diversity promoting prior). 通过引入该项增强了连接权的多样性, 减少了死节点和过分活跃的节点的数量, 提高了分类的效率.

基于 DBN 的高光谱图像空间特征分类方法与基于 SAE 的高光谱图像空间特征分类方法十分类似. 文献 [95] 中, Li 等首先利用 PCA 将原始高光谱图像压缩, 保留前 3 个主成分, 之后提取 7×7 邻域内的数据作为输入, 利用 DBN 网络进行特征提取及分类. 文献 [32, 96] 均是利用 DBN 分别提取光谱特征与空间特征, 之后将光谱特征和空间特征连接形成空谱特征, 再基于空谱特征完成分类任务, 分类整体框架与文献 [31] 中的框架基本一致, 而文献 [32] 中的方法在 DBN 的训练过程中引入了稀疏限制、高斯噪声等.

3.4 分类方法的比较

3.4.1 评价指标

评价一个高光谱图像分类方法的优劣时通常需要计算一些定量的指标, 包括: 总分类精度 (Overall

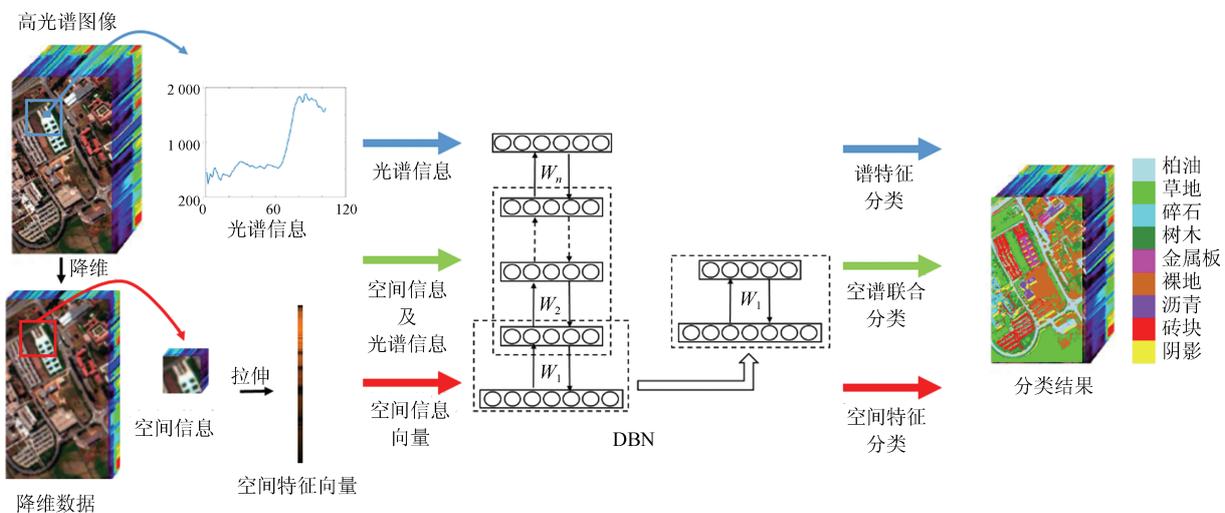


图 12 基于 DBN 的高光谱图像分类
Fig. 12 HSI classification based on DBN

accuracy, OA), 平均分类精度 (Average accuracy, AA) 和 Kappa 统计量 (Kappa statistic, K) 三个指标. 总分类精度表示整个测试数据集中正确分类的像素个数与整个测试集像素个数的比, 平均分类精度表示各个类别分类精度的平均值. Kappa 系数是一个测定分类结果与实际结果吻合程度的指标, 它不仅考虑了被正确分类的像素, 同时也考虑了各种漏分 (某些目标没有被分到应该属于的类别) 和错分 (某一类目标被错误的分到其他类别) 情况, 计算公式为

$$K = \frac{N \sum_{i=1}^n C_{ii} - \sum_{i=1}^n C_{i+} C_{+i}}{N^2 - \sum_{i=1}^n C_{i+} C_{+i}} \quad (4)$$

上述公式中 N 表示像素总个数, C_{ij} 表示混淆矩阵^[97] 的第 i 行第 j 列的元素, C_{i+} 、 C_{+i} 分别表示第 i 行和第 i 列总像素个数, n 表示类别数.

3.4.2 方法对比

无论是基于 CNN、SAE 还是 DBN 的高光谱图像分类方法, 均包含了基于谱特征分类, 基于空间特征分类, 基于空谱联合分类三大类. 表 2 中, 以 Pavia University 为测试数据对比了基于 CNN、SAE 及 DBN 的高光谱分类方法的分类精度. 每种方法均包含基于谱特征分类, 基于空间特征分类和基于空谱联合分类三种方法. 在基于 CNN 的分类方法中训练集与测试集的比例为 1 : 9^[75] (有标签的数据中抽取 10% 作为训练数据, 剩余的 90% 作为测试数据), 在基于 DBN 的分类方法中训练集与测试集的比例为 1 : 1^[32], 在基于 SAE 的分类方法中训练集, 验证集及测试集的比例依次为 6 : 2 : 2^[31].

依照目前现有方法的分类结果, 在采用相同的深度学习模型的情况下, 基于空谱联合特征的分类

方法精度最高, 基于空间特征的分类方法次之, 基于谱特征分类的方法精度相对最差. 高光谱图像中存在异物同谱及同谱异物等问题, 仅仅利用光谱特征, 难以有效解决这些问题, 从而也会影响分类的结果. 在利用空间信息进行分类时, 当待分类的像素对应的地物目标出现同谱异物或异物同谱等问题时, 邻域的像素所提供的信息 (上下文信息) 对区分待分类像素的类别会起到至关重要的作用. 例如当周边的像素覆盖的地物目标均是草地时, 中心的像素覆盖的地物目标很大概率上也是草地. 但是, 在基于空间特征的分类方法中, 需要对数据进行压缩, 而压缩的过程会损失一定量的光谱信息, 从而影响了最终的分类结果. 基于深度学习的空谱联合特征分类方法中, 既利用了光谱信息, 又利用了空间信息, 充分发挥了高光谱图像的数据优势, 因此取得了优于前两种分类方法的效果. 表 4 及图 13 中, 以 Indian Pines 为测试数据, 在训练集与测试集 1 : 1 的实验条件下, 对比了分别基于 SAE、DBN、2D-CNN、3D-CNN 及双通道卷积神经网络 (Dual channel convolutional neural network, DC-CNN) 的 5 种高光谱图像空谱联合分类的方法.

图 13 中, (a) 对应 Indian Pines 的伪彩色图; (b) 对应人工标记的 16 个类别; (c) 对应基于 SAE 的分类方法的分类结果, 其中 OA = 93.98%; (d) 对应基于 DBN 的分类方法的分类结果, 其中 OA = 95.91%; (e) 对应基于 2D-CNN 的分类方法的分类结果, 其中 OA = 95.97%; (f) 对应基于 3D-CNN 的分类方法的分类结果, 其中 OA = 99.07%; (g) 对应基于 DC-CNN 的分类方法的分类结果, 其中 OA = 99.92%. 从实验结果可以看出, 基于 CNN 的分类方法, 包括 2D-CNN、3D-CNN、DC-CNN 分类精度相对较高. 实际上, 统计分析现有的基于深度学习的高光谱图像分类方法的分类效

表 3 Pavia University 分类结果

Table 3 The classification results of Pavia University

模型	谱特征			空间特征			空谱联合特征		
	OA (%)	AA (%)	K ($\times 100$)	OA (%)	AA (%)	K ($\times 100$)	OA (%)	AA (%)	K ($\times 100$)
CNN ^[71]	92.28	92.55	90.37	94.04	97.52	92.43	99.54	99.77	99.56
SAE ^[31]	95.14	94.01	93.70	98.12	97.32	97.55	98.52	97.82	98.07
DBN ^[32]	96.42	95.09	95.30	98.62	97.95	98.19	99.05	98.48	98.75

表 4 Indian Pines 分类结果

Table 4 The classification results of Indian Pines

模型	SAE ^[31]	DBN ^[32]	2D-CNN ^[64]	3D-CNN ^[76]	DC-CNN ^[71]
OA (%)	93.98	95.91	95.97	99.07	99.92
AA (%)	93.81	94.20	93.23	98.66	99.57
K ($\times 100$)	93.13	95.34	95.40	98.93	99.91

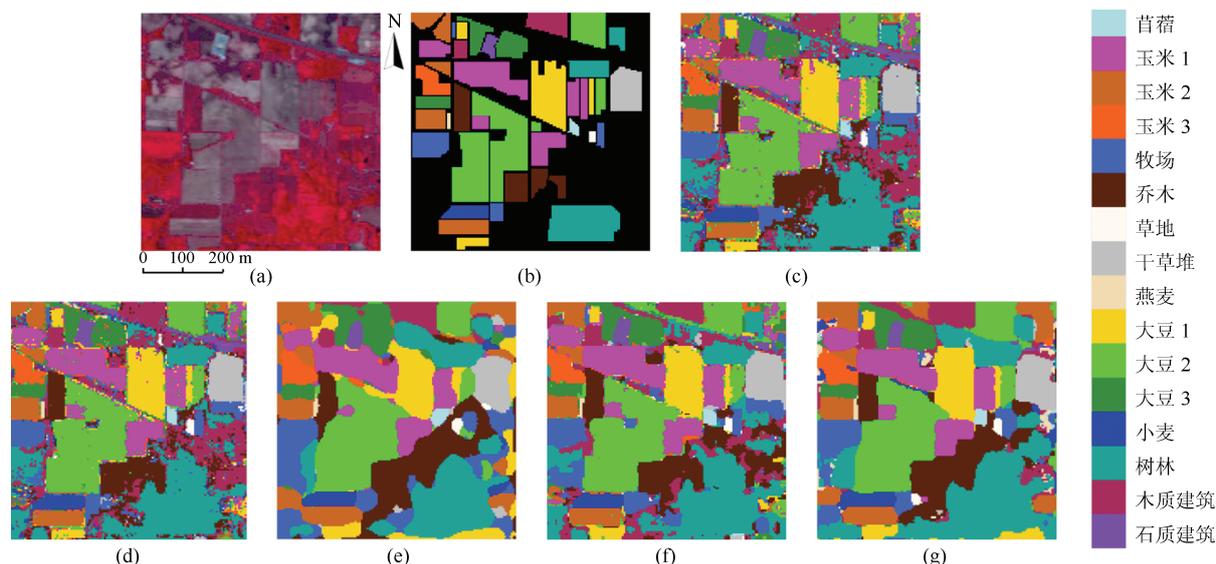


图 13 Indian Pines 分类结果图

Fig. 13 The classification of Indian Pines

果及提出时间,可以发现,基于 SAE 和 DBN 的高光谱图像分类方法发展较早,基于 CNN 的各类高光谱图像分类方法提出较晚.但是就目前发展现状而言,基于 CNN 的各类高光谱图像分类方法发展最快,论文的数量最多.整体而言,分类效果较基于 SAE 和 DBN 的高光谱图像分类也更好.

4 总结与展望

本文简要回顾了高光谱图像分类的研究背景,简要介绍了目前在高光谱图像分类领域应用较为广泛几个数据集和三种深度学习网络模型 CNN、SAE、DBN.在此基础上,总结了 CNN、SAE、DBN 这三种深度学习模型在高光谱图像分类领域的研究现状.高光谱图像同时含有空间信息和光谱信息,应用范围非常广,而且随着遥感技术的发展,高光谱图像的质量会有进一步的提升,因此高光谱图像的应用前景十分广阔.从已有的基于深度学习的高光谱图像分类方法来看,该类基于深度学习的分类方法发展迅速,而且潜力巨大.利用深度学习模型能够有效地从高维的、信息量大的高光谱图像中自主提取出有利于提高分类精度的鲁棒的深度特征.而且,利用深度学习的相关模型也能有效地结合并利用高光谱图像中的空间信息及光谱信息,能充分发挥高光谱图像同时蕴含空间信息及光谱信息的优势.

目前主要有两个问题阻碍着基于深度学习的高光谱图像分类方法进一步发展.首先,小样本问题,深度学习模型训练过程中需求大量有效标注数据而目前高光谱图像标注数据相对较少.深度学习虽然

具有很强的特征提取能力,但是网络的训练过程是有监督的,即使是 SAE 或 DBN,在网络整体调优阶段依然是有监督的.有监督的训练过程就需要一定高质量的标注数据,网络规模较大时,需求的标注数据可能更多.然而,在目前已公开的高光谱图像数据中,有人工标注的标签的数据相对较少.其次,空谱联合特征提取问题.利用空谱联合的策略无疑是更能充分发挥高光谱图像同时含有空间信息和光谱信息的优势的,而且空谱联合也是解决异物同谱及同谱异物等问题的一种有效手段.然而从上文的综述中可以看出,基于深度学习的高光谱图像空谱联合特征提取方法非常多,各有优缺点,有效地空谱特征对提高高光谱图像分类精度是十分重要的.因此,在后续的工作中,基于深度学习的高光谱图像分类方法应该考虑以下几点研究方向.

1) 在深度学习模型本身结构方面,模型中参数数量较多是模型训练过程中需要大量标注数据的主要原因,考虑设计新型的深度学习模型,通过构建稀疏结构的模型或设计更加有效的激励函数来降低模型中冗余的参数.

2) 在模型的训练方面,目前深度学习模型训练方法主要是随机梯度下降算法及一些改进的随机梯度下降算法.在处理特定的数据集时,直接利用该算法对随机初始化的网络模型训练时,为使网络收敛,通常需要大量的标注数据.考虑将模型的训练过程分为预训练及调优两个过程,预训练可以在其他的标注数据较多的高光谱图像数据集上进行,之后再结合迁移学习,One-shot learning^[98]及 Zero-shot learning^[99]等思想,将预训练好的模型移植到特定

任务的数据集上, 利用少量样本进行微调, 最后完成分类任务.

3) 在标注样本数据较少方面, 一方面可以考虑将主动学习与半监督等方法与深度学习的方法相结合. 首先利用少量标注样本对网络模型进行初步训练, 之后再利用主动学习及半监督的算法从未标注的数据中自动标注部分数据, 结合已有的标注数据和自动标注的新数据进一步的训练网络模型; 另一方面, 考虑将生成式对抗网络 (Generative adversarial network, GAN)^[100] 引入到高光谱图像分类中来. 对每个类别都建立一个生成式对抗网络, 利用该网络生成符合该类别特点的新的样本数据, 以此增加该类别的标注数据的数据量, 建立具有大量标注数据的数据集, 最后利用该数据集对深度学习网络模型进行训练.

4) 在空谱联合特征提取方面, 基于深度学习提取空谱联合特征时, 需要考虑的问题主要是, 空间邻域大小, 网络的规模, 结构及复杂程度. 现有的方法, 无论采用哪种深度网络模型, 主要依靠多次对比实验分析选择合适的参数. 如何根据高光谱图像本身的空间分辨率, 光谱分辨率及采集场景等信息, 指导空间邻域大小的选取, 网络规模的设计等是值得关注的一个问题.

5) 从分类结果上来说, 高光谱图像分类与自然图像处理问题中的语义分割非常类似. 但是数据特点等有所不同, 深度学习中有大量的效果非常显著的语义分割的方法^[52, 101], 目前我们暂时还没有发现有基于深度学习的语义分割的方法应用在高光谱图像分类上的文章, 在后续的工作中, 这是一个非常值得考虑的方向.

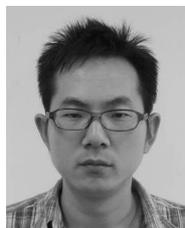
References

- 1 Bioucas-Dias J M, Plaza A, Camps-Valls G, Scheunders P, Nasrabadi N, Chanussot J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, 2013, **1**(2): 6–36
- 2 Benediktsson J A, Chanussot J, Moon W M. Very high-resolution remote sensing: challenges and opportunities [point of view]. *Proceedings of the IEEE*, 2012, **100**(6): 1907–1910
- 3 Fauvel M, Tarabalka Y, Benediktsson J A, Chanussot J, Tilton J C. Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 2013, **101**(3): 652–675
- 4 Landgrebe D. Hyperspectral image data analysis. *IEEE Signal Processing Magazine*, 2002, **19**(1): 17–28
- 5 Nasrabadi N M. Hyperspectral target detection: an overview of current and future challenges. *IEEE Signal Processing Magazine*, 2014, **31**(1): 34–44
- 6 Van der Meer F. Analysis of spectral absorption features in hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 2004, **5**(1): 55–68
- 7 Hege E K, O'Connell D, Johnson W, Bastay S, Dereniak E L. Hyperspectral imaging for astronomy and space surveillance. In: *Proceedings of the 2004 Volume 5159, Imaging Spectrometry IX*. San Diego, USA: SPIE, 2004. 380–391
- 8 Gowen A A, O'Donnell C P, Cullen P J, Downey G, Frias J M. Hyperspectral imaging-an emerging process analytical tool for food quality and safety control. *Trends Food Science & Technology*, 2007, **18**(12): 590–598
- 9 Lacar F M, Lewis M M, Grierson I T. Use of hyperspectral imagery for mapping grape varieties in the Barossa Valley, South Australia. In: *Proceedings of the 2001 IEEE International Geoscience and Remote Sensing Symposium*. Sydney, Australia: IEEE, 2001. 2875–2877
- 10 Stuffer T, Förster K, Hofer S, Leipold M, Sang B, Kaufmann H et al. Hyperspectral imaging An advanced instrument concept for the EnMAP mission (environmental mapping and analysis programme). *Acta Astronautica*, 2009, **65**(7-8): 1107–1112
- 11 Malthus T J, Mumby P J. Remote sensing of the coastal zone: an overview and priorities for future research. *International Journal of Remote Sensing*, 2003, **24**(13): 2805–2815
- 12 Plaza A, Du Q, Chang Y L, King R L. High performance computing for hyperspectral remote sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2011, **4**(3): 528–544
- 13 Plaza A, Martinez P, Perez R, Plaza J. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles. *Pattern Recognition*, 2004, **37**(6): 1097–1116
- 14 Samat A, Li J, Liu S C, Du P J, Miao Z L, Luo J Q. Improved hyperspectral image classification by active learning using pre-designed mixed pixels. *Pattern Recognition*, 2016, **51**: 43–58
- 15 Imani M, Ghassemian H. Band clustering-based feature extraction for classification of hyperspectral images using limited training samples. *IEEE Geoscience and Remote Sensing Letters*, 2014, **11**(8): 1325–1329
- 16 Li F, Xu L L, Siva P, Wong A, Clausi A. Hyperspectral image classification with limited labeled training samples using enhanced ensemble learning and conditional random fields. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, **8**(6): 2427–2438
- 17 Kumar S, Ghosh J, Crawford M M. Best-bases feature extraction algorithms for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 2001, **39**(7): 1368–1379
- 18 Samaniego L, Bardossy A, Schulz K. Supervised classification of remotely sensed imagery using a modified k-NN technique. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, **46**(7): 2112–2125
- 19 Ediriwickrema J, Khorram S. Hierarchical maximum-likelihood classification for improved accuracies. *IEEE Transactions on Geoscience and Remote Sensing*, 1997, **35**(4): 810–816
- 20 Foody G M, Mathur A. A relative evaluation of multi-class image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, **42**(6): 1335–1343
- 21 Hughes G. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 1968, **14**(1): 55–63
- 22 Villa A, Benediktsson J A, Chanussot J, Jutten C. Hyperspectral image classification with independent component discriminant analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, **49**(12): 4865–4876

- 23 Prasad S, Bruce L M. Limitations of principal components analysis for hyperspectral target recognition. *IEEE Geoscience and Remote Sensing Letters*, 2008, **5**(4): 625–629
- 24 Li W, Prasad S, Fowler J E, Bruce L M. Locality-preserving dimensionality reduction and classification for hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, **50**(4): 1185–1198
- 25 Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, **42**(8): 1778–1790
- 26 Li J, Bioucas-Dis J M, Plaza A. Hyperspectral image segmentation using a new Bayesian approach with active learning. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, **49**(10): 3947–3960
- 27 Liao W Z, Pizurica A, Scheunders P, Philips W, Pi Y G. Semisupervised local discriminant analysis for feature extraction in hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, **51**(1): 184–198
- 28 Li W, Prasad S, Fowler J E, Bruce L M. Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 2011, **8**(5): 894–898
- 29 Zhong Y F, Zhang L P. An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, **50**(3): 894–909
- 30 Sun L, Wu Z B, Liu J J, Xiao L, Wei Z H. Supervised spectral-spatial hyperspectral image classification with weighted Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, **53**(3): 1490–1503
- 31 Chen Y S, Lin Z H, Zhao X, Wang G, Gu Y F. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, **7**(6): 2094–2107
- 32 Chen Y S, Zhao X, Jia X P. Spectral-spatial classification of hyperspectral data based on deep belief network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, **8**(6): 2381–2392
- 33 Zhang L P, Zhang L F, Tao D C, Huang X. Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, **51**(1): 242–256
- 34 Zhang L F, Zhang Q, Zhang L P, Tao D C, Huang X, Du B. Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding. *Pattern Recognition*, 2015, **48**(10): 3102–3112
- 35 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527–1554
- 36 Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, USA: Curran Associates, 2012. 1097–1105
- 37 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, **313**(5786): 504–507
- 38 Zaremba W, Sutskever I, Vinyals O. Recurrent neural network regularization [Online], available: <http://arxiv.org/abs/1409.2329>, July 7, 2017.
- 39 Zhu Yu, Zhao Jiang-Kun, Wang Yi-Ning, Zheng Bing-Bing. A review of human action recognition based on deep learning. *Acta Automatica Sinica*, 2016, **42**(6): 848–857 (朱煜, 赵江坤, 王逸宁, 郑兵兵. 基于深度学习的人体行为识别算法综述. *自动化学报*, 2016, **42**(6): 848–857)
- 40 He K M, Zhang X Y, Ren S Q, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1904–1916
- 41 Luo Jian-Hao, Wu Jian-Xin. A survey on fine-grained image categorization using deep convolutional features. *Acta Automatica Sinica*, 2017, **43**(8): 1306–1318 (罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. *自动化学报*, 2017, **43**(8): 1306–1318)
- 42 Girshick R. Fast R-CNN. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1440–1448
- 43 Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA: IEEE, 2014. 580–587
- 44 Guan Hao, Xue Xiang-Yang, An Zhi-Yong. Advances on application of deep learning for video object tracking. *Acta Automatica Sinica*, 2016, **42**(6): 834–847 (管皓, 薛向阳, 安志勇. 深度学习在视频目标跟踪中的应用进展与展望. *自动化学报*, 2016, **42**(6): 834–847)
- 45 Liu F Y, Shen C H, Lin G S. Deep convolutional neural fields for depth estimation from a single image. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 5162–5170
- 46 Dong C, Chen C L, He K M, Tang X O. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(2): 295–307
- 47 Sun Xu, Li Xiao-Guang, Li Jia-Feng, Zhuo Li. Review on deep learning based image super-resolution restoration algorithms. *Acta Automatica Sinica*, 2017, **43**(5): 697–709 (孙旭, 李晓光, 李嘉锋, 卓力. 基于深度学习的图像超分辨率复原研究进展. *自动化学报*, 2017, **43**(5): 697–709)
- 48 Lin M, Chen Q, Yang S C. Network in network [Online], available: <https://arxiv.org/abs/1312.4400>, July 7, 2017
- 49 Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 1–9
- 50 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [Online], available: <https://arxiv.org/abs/1409.1556>, July 7, 2017
- 51 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770–778
- 52 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 3431–3440
- 53 Wan L, Zeiler M, Zhang S X, LeCun Y, Fergus R. Regularization of neural networks using DropConnect. In: Proceeding of the 30th International Conference on Machine Learning. Atlanta, USA: IMLS, 2013. 1058–1066

- 54 Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceeding of the 32th International Conference on Machine Learning. Lille, France: IMLS, 2015. 448–456
- 55 Erhan D, Bengio Y, Courville A, Manzagol P, Vincent P, Bengio S. Why does unsupervised pre-training help deep learning. *Journal of Machine Learning Research*, 2010, **11**: 625–660
- 56 Hinton G E. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 2002, **14**(8): 1771–1800
- 57 Li J, Bioucas-Dias J M, Plaza A. Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, **48**(11): 4085–4098
- 58 Hu W, Huang Y Y, Wei L, Zhang F, Li H C. Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, 2015, **2015**: Article No. 258619
- 59 Liao J G, Chin K V. Logistic regression for disease classification using microarray data: model selection in a large p and small n case. *Bioinformatics*, 2007, **23**(15): 1945–1951
- 60 Mei S H, Ji J Y, Bi Q Q, Hou J H, Du Q, Li W. Integrating spectral and spatial information into deep convolutional Neural Networks for hyperspectral classification. In: Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Beijing, China: IEEE, 2016. 5067–5070
- 61 He K M, Zhang X Y, Ren S Q, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1026–1034
- 62 Tang Y Y, Lu Y, Yuan H L. Hyperspectral image classification based on three-dimensional scattering wavelet transform. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, **53**(5): 2467–2480
- 63 Makantasis K, Karantzas K, Doulamis A, Doulamis N. Deep supervised learning for hyperspectral data classification through convolutional neural networks. In: Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Milan, Italy: IEEE, 2015. 4959–4962
- 64 Yue J, Zhao W Z, Mao S J, Liu H. Spectral-spatial classification of hyperspectral images using deep convolutional neural networks. *Remote Sensing Letters*, 2015, **6**(6): 468–477
- 65 Zhao W Z, Guo Z, Yue J, Zhang X Y, Luo L Q. On combining multiscale deep learning features for the classification of hyperspectral remote sensing imagery. *International Journal of Remote Sensing*, 2015, **36**(13): 3368–3379
- 66 Aptoula E, Ozdemir M C, Yanikoglu B. Deep learning with attribute profiles for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 2016, **13**(12): 1970–1974
- 67 Dalla Mura M, Benediktsson J A, Waske B, Bruzzone L. Morphological attribute profiles for the analysis of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, **48**(10): 3747–3762
- 68 Liang H M, Li Q. Hyperspectral imagery classification using sparse representations of convolutional neural network features. *Remote Sensing*, 2016, **8**(2): Article No. 99
- 69 Li Y S, Xie W Y, Li H Q. Hyperspectral image reconstruction by deep convolutional neural network for classification. *Pattern Recognition*, 2017, **63**: 371–383
- 70 Yang J X, Zhao Y Q, Chan J C W, Yi C. Hyperspectral image classification using two-channel deep convolutional neural network. In: Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Beijing, China: IEEE, 2016. 5079–5082
- 71 Zhang H K, Li Y, Zhang Y Z, Shen Q. Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network. *Remote Sensing Letters*, 2017, **8**(5): 438–447
- 72 Yue J, Mao S J, Li M. A deep learning framework for hyperspectral image classification using spatial pyramid pooling. *Remote Sensing Letters*, 2016, **7**(9): 875–884
- 73 Zhao W Z, Du S H. Spectral-spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(8): 4544–4554
- 74 Chen H T, Chang H W, Liu T L. Local discriminant embedding and its variants. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 846–853
- 75 Chen Y S, Jiang H L, Li C Y, Jia X P, Ghamisi P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(10): 6232–6251
- 76 Li Y, Zhang H K, Shen Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, 2017, **9**(1): Article No. 67
- 77 Lee H, Kwon H. Contextual deep CNN based hyperspectral classification. In: Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium. Beijing, China: IEEE, 2016. 3322–3325
- 78 Slavkovic V, Verstockt S, De Neve W, Van Hoecke S, Van de Walle R. Hyperspectral image classification with convolutional neural networks. In: Proceedings of the 23rd ACM International Conference on Multimedia. Brisbane, Australia: ACM, 2015. 1159–1162
- 79 Lin Z H, Chen Y S, Zhao X, Wang G. Spectral-spatial classification of hyperspectral image using autoencoders. In: Proceedings of the 9th International Conference on Information, Communications and Signal Processing (ICICSP). Tainan, China: IEEE, 2013. 1–5
- 80 Ma X R, Wang H Y, Geng J, Wang J. Hyperspectral image classification with small training set by deep network and relative distance prior. In: Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Beijing, China: IEEE, 2016. 3282–3285
- 81 Xing C, Ma L, Yang X Q. Stacked denoise autoencoder based feature extraction and classification for hyperspectral images. *Journal of Sensors*, 2016, **2016**: Article No. 3632943
- 82 Liu Y Z, Cao G, Sun Q S, Siegel M. Hyperspectral classification via deep networks and superpixel segmentation. *International Journal of Remote Sensing*, 2015, **36**(13): 3459–3482
- 83 Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P A. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, 2010, **11**: 3371–3408
- 84 Zhang L P, Zhang L F, Du B. Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 2016, **4**(2): 22–40
- 85 Ma X R, Wang H Y, Geng J. Spectral-spatial classification of hyperspectral image based on deep auto-encoder. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, **9**(9): 4073–4085

- 86 Tao C, Pan H B, Li Y S, Zou Z R. Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification. *IEEE Geoscience and Remote Sensing Letters*, 2015, **12**(12): 2438–2442
- 87 Shin H C, Orton M R, Collins D J, Doran S J, Leach M O. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(8): 1930–1943
- 88 Wang L Z, Zhang J B, Liu P, Choo K K R, Huang F. Spectral-spatial multi-feature-based deep learning for hyperspectral remote sensing image classification. *Soft Computing*, 2017, **21**(1): 213–221
- 89 Li J M, Bruzzone L, Liu S C. Deep feature representation for hyperspectral image classification. In: Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Milan, Italy: IEEE, 2015. 4951–4954
- 90 Ma X R, Geng J, Wang H Y. Hyperspectral image classification via contextual deep learning. *EURASIP Journal on Image and Video Processing*, 2015, **2015**(1): Article No. 20
- 91 Han X B, Zhong Y F, Zhang L P. Spatial-spectral classification based on the unsupervised convolutional sparse autoencoder for hyperspectral remote sensing imagery. In: Proceedings of the 2016 ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Prague, Czech Republic: ISPRS, 2016. 25–31
- 92 Ma X R, Wang H Y, Wang J. Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, **120**: 99–107
- 93 He M Y, Li X H, Zhang Y F, Zhang J, Wang W G. Hyperspectral image classification based on deep stacking network. In: Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Beijing, China: IEEE, 2016. 3286–3289
- 94 Zhong P, Gong Z Q, Schönlieb C. A diversified deep belief network for hyperspectral image classification. In: Proceedings of the 2016 ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Prague, Czech Republic: ISPRS, 2016. 443–449
- 95 Li T, Zhang J P, Zhang Y. Classification of hyperspectral image based on deep belief networks. In: Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP). Paris, France: IEEE, 2014. 5132–5136
- 96 Le J H, Yazdanpanah A P, Regentova E, Muthukumar V. A deep belief network for classifying remotely-sensed hyperspectral data. In: Proceedings of the 11th International Symposium on Visual Computing. Las Vegas, USA: Springer, 2015. 682–692
- 97 Thompson W D, Walter S D. A reappraisal of the kappa coefficient. *Journal of Clinical Epidemiology*, 1988, **41**(10): 949–958
- 98 Vinyals O, Blundell C, Lillicrap T, Kavukcuoglu K, Wierstra D. Matching networks for one shot learning. In: Proceedings of the 2016 Advances in Neural Information Processing Systems. Barcelona, Spain: Curran Associates, 2016. 3630–3638
- 99 Socher R, Ganjoo M, Sridhar H, Bastani O, Manning C D, Ng A Y. Zero-shot learning through cross-modal transfer. In: Proceedings of the 2013 Advances in Neural Information Processing Systems. Lake Tahoe, USA: Curran Associates, 2013. 935–943
- 100 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: Curran Associates, 2014. 2672–2680
- 101 Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention. Munich, Germany: Springer, 2015. 234–241



张号逵 西北工业大学计算机学院陕西省语音与图像信息处理重点实验室博士研究生。2016年于西北工业大学计算机学院陕西省语音与图像信息处理重点实验室获得计算机应用技术专业硕士学位。主要研究方向为图像处理, 深度学习, 高光谱图像分类。E-mail: hkzhang1991@mail.nwpu.edu.cn

(ZHANG Hao-Kui Ph.D. candidate at Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, the School of Computer Science, Northwestern Polytechnical University. He received his master degree in computer application technology from Shaanxi Provincial Key Laboratory of Speech and Image Information Processing in 2016. His research interest covers image processing, deep learning, and hyperspectral image classification.)



李映 西北工业大学计算机学院教授。2002年于西安电子科技大学雷达信号处理国家重点实验室电路与系统专业获博士学位。主要研究方向为图像处理, 计算智能, 信号处理。本文通信作者。E-mail: lybyp@nwpu.edu.cn

(LI Ying Professor at the School of Computer Science. She received her Ph.D. degree in electrical circuit and system, from the National Key Laboratory of Radar Signal Processing, Xidian University in 2002. Her research interest covers image processing, computation intelligence, and signal processing. Corresponding author of this paper.)



姜晔楠 西北工业大学计算机学院陕西省语音与图像信息处理重点实验室博士研究生。2017年于东北林业大学获得计算机科学与技术专业硕士学位。主要研究方向为高光谱图像分类, 深度学习。E-mail: ynjiang@mail.nwpu.edu.cn

(JIANG Ye-Nan Ph.D. candidate at Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, the School of Computer Science, Northwestern Polytechnical University. She received her master degree in computer science from Northeast Forestry University in 2017. Her research interest covers hyperspectral image processing and deep learning.)