分类号： 密级：

U D C ： 编号：

工程硕士学位论文

基于卷积神经网络及集成学习的高光谱遥感图像分类方法研究

|  |  |
| --- | --- |
| 硕士研究生 | ：周星 |
| 指导教师 | ：王立国　教授 |
| 学科、专业 | ：电子与通信工程 |
| 论文主审人 | ： 教授 |

哈尔滨工程大学

2019年3月

分类号： 密级：

U D C ： 编号：

工程硕士学位论文

基于卷积神经网络及集成学习的高光谱遥感图像分类方法研究

|  |  |
| --- | --- |
| **硕士研究生** | ：周星 |
| **指导教师** | ：王立国　教授 |
| **学位级别** | ：工程硕士 |
| **学科、专业** | ：电子与通信工程 |
| **所在单位** | ：信息与通信工程学院 |
| **论文提交日期** | ：2019年03月 |
| **论文答辩日期** | ：2019年03月 |
| **学位授予单位** | ：哈尔滨工程大学 |

Classified Index:

U.D.C:

A Dissertation for the Degree of M. Eng

Research on the Semi-Supervised Slassification of Hyperspectral Images with Clustering Algorithm

|  |  |
| --- | --- |
| **Candidate:** | Zhou Xing |
| **Supervisor:** | Prof. Wang Liguo |
| **Academic Degree Applied for:** | Master of Engineering |
| **Specialty:** | Electronic and Communication Engineering |
| **Date of Submission:** | March, 2019 |
| **Date of Oral Examination:** | March, 2019 |
| **University:** | Harbin Engineering University |

哈尔滨工程大学

学位论文原创性声明

本人郑重声明：本论文的所有工作，是在导师的指导下，由作者本人独立完成的。有关观点、方法、数据和文献的引用已在文中指出，并与参考文献相对应。除文中已注明引用的内容外，本论文不包含任何其他个人或集体已经公开发表的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

作者（签字）：

日期： 年 月 日

哈尔滨工程大学

学位论文授权使用声明

本人完全了解学校保护知识产权的有关规定，即研究生在校攻读学位期间论文工作的知识产权属于哈尔滨工程大学。哈尔滨工程大学有权保留并向国家有关部门或机构送交论文的复印件。本人允许哈尔滨工程大学将论文的部分或全部内容编入有关数据库进行检索，可采用影印、缩印或扫描等复制手段保存和汇编本学位论文，可以公布论文的全部内容。同时本人保证毕业后结合学位论文研究课题再撰写的论文一律注明作者第一署名单位为哈尔滨工程大学。涉密学位论文待解密后适用本声明。

本论文（□在授予学位后即可 □在授予学位12个月后 □解密后）由哈尔滨工程大学送交有关部门进行保存、汇编等。

作者（签字）： 导师（签字）：

日期： 年 月 日 年 月 日

# 摘 要

随着科学的不断发展，遥感技术正逐渐渗透于国家经济、生态环境、国防科技等诸多方面，并且发挥着不可替代的作用。高光谱影像分类技术是当前遥感科学的主要研究方向之一，其目的在于确定遥感影像中的地物类别。较高的分类精度有助于人们更好地理解地物分布，因此寻求高分类精度的图像分类算法有着重要的意义。当前，分类算法众多，然而因为高光谱数据的特殊性，高光谱影像分类算法依然存在着众多问题。本文基于当前分类算法存在的问题，提出两种高光谱影像分类算法。本文的研究工作描述如下：

针对传统高光谱影像分类仅利用光谱信息的问题，本文提出一种简单的空-谱信息提取方法，有效地采集到光谱空间联合信息。考虑到当前树形集成方法在数据科学竞赛中的突出表现，本文首次将LightGBM算法用于高光谱影像分类中。最后，联合该空-谱信息提取方法和LightGBM算法，提出一种新型高光谱影像分类算法，并通过两组实验验证了利用该算法可以极大提高分类精度。

传统高光谱影像空-谱分类算法严重依赖领域内知识、只能获取“浅”层特征、需要人工合成特征；基于1D卷积神经网络的高光谱影像分类算法不能利用数据的空间信息；基于2D卷积神经网络的高光谱影像分类算法，为了避免产生过大的卷积核，必须对原始数据进行降维，损失了一定程度的光谱信息。鉴于此，本文提出一种新型基于3D-CNN的高光谱影像分类算法。该算法构建一种通用的高光谱影像分类框架，利用3D滤波器从三个方向同时进行卷积操作，并且通过设置合适的卷积步长使得一部分卷积层可以在卷积操作同时起到池化的作用，提升了网络的执行效率。最后，通过两组实验验证了本章算法有着较高的分类表现。

关键词：高光谱影像；聚类；分类；卷积神经网络；LightGBM

# ABSTRACT

With the continuous development of science, remote sensing technology is gradually infiltrating into the national economy, ecological environment, national defense science and technology, and plays an irreplaceable role. Hyperspectral image classification technology is one of the main research directions of remote sensing science, and its purpose is to determine the feature types in remote sensing images. Higher classification accuracy helps people better understand the distribution of features, so it is important to seek image classification algorithms with high classification accuracy. At present, there are many classification algorithms. However, due to the particularity of hyperspectral data, there are still many problems in hyperspectral image classification algorithms. Based on the problems of current classification algorithms, two hyperspectral image classification algorithms are proposed. The research work of this paper is described as follows:

Aiming at the problem of using only spectral information in traditional hyperspectral image classification, this paper proposes a simple method for extracting spatial-spectral information, which effectively collects joint information of spectral space. Considering the outstanding performance of the current tree integration method in the data science competition, the LightGBM algorithm is used for hyperspectral image classification for the first time. Finally, a new hyperspectral image classification algorithm is proposed by combining the space-spectral information extraction method and LightGBM algorithm. It is verified by two sets of experiments that the classification accuracy can be greatly improved by using this algorithm.

The traditional hyperspectral image space-spectral classification algorithm relies heavily on domain knowledge, can only obtain "shallow" layer features, and requires artificial synthesis features; hyperspectral image classification algorithm based on 1D convolutional neural network can not utilize spatial information of data; based on 2D In order to avoid excessive convolution kernels, the hyperspectral image classification algorithm of convolutional neural networks must reduce the dimensionality of the original data, so a certain degree of spectral information is lost. In view of this, this paper proposes a new 3D-CNN based hyperspectral image classification algorithm. The algorithm constructs a general hyperspectral image classification framework, which uses a 3D filter to simultaneously perform convolution operations from three directions, and by setting an appropriate convolution step, a part of the convolutional layer can simultaneously serve as a pool in the convolution operation. The role of the role. Finally, two sets of experiments verify that the algorithm in this chapter has a high classification performance.

**Key words:** hyperspectral imagery; clustering; classification; convolutional neural network; LightGBM

目 录

[第1章 绪论 2](#_Toc533438528)

[1.1 课题研究背景及意义 2](#_Toc533438529)

[1.2 高光谱遥感数据特征 3](#_Toc533438530)

[1.3 国内外在该方向的研究现状及分析 5](#_Toc533438531)

[1.3.1 深度学习发展现状 5](#_Toc533438532)

[1.3.2 集成学习发展现状 6](#_Toc533438533)

[1.3.3 高光谱图像分类研究现状 7](#_Toc533438534)

[1.4 本文的研究内容与组织结构 8](#_Toc533438535)

[第2章 高光谱影像分类研究 10](#_Toc533438536)

[2.1 高光谱影像分类概述 10](#_Toc533438537)

[2.1.1有监督分类方法 10](#_Toc533438538)

[2.1.2无监督分类方法 11](#_Toc533438539)

[2.1.3半监督分类方法 12](#_Toc533438540)

[2.2 评价方法 14](#_Toc533438541)

[2.2.1混淆矩阵（Confusion matrix） 14](#_Toc533438542)

[2.2.2总体精度与平均精度 14](#_Toc533438543)

[2.2.2 Kappa系数 15](#_Toc533438544)

[2.3 实验数据 15](#_Toc533438545)

[2.3.1 Indian Pines高光谱影像 16](#_Toc533438546)

[2.3.2 University of Pavia高光谱影像 17](#_Toc533438547)

[2.3.3 SalinasValley高光谱影像 18](#_Toc533438548)

[2.4 仿真实验 19](#_Toc533438549)

[2.4.1基于SVM的高光谱遥感分类算法实验 20](#_Toc533438550)

[2.4.2仿真结果与分析 20](#_Toc533438551)

[2.5 本章小结 21](#_Toc533438552)

[第3章 基于树形集成学习的高光谱图像空-谱分类算法 22](#_Toc533438553)

[3.1 引言 22](#_Toc533438554)

[3.2 集成学习基本理论 22](#_Toc533438555)

[3.2.1 Boosting 23](#_Toc533438556)

[3.2.2 Boosting Tree 23](#_Toc533438557)

[3.2.3 *GBDT* 24](#_Toc533438558)

[3.2.4 *Bagging* 25](#_Toc533438559)

[3.2.5 随机森林 25](#_Toc533438560)

[3.3 实验算法 26](#_Toc533438561)

[3.3.1 LightGBM 26](#_Toc533438562)

[3.3.2 空间聚类方式的空谱信息获取 27](#_Toc533438563)

[3.3.2 基于LightGBM和K-Means的高光谱图像空-谱分类算法 28](#_Toc533438564)

[3.4 实验仿真结果与分析 29](#_Toc533438565)

[3.5 本章小结 36](#_Toc533438566)

[第4章 基于3D卷积神经网络的高光谱图像分类 37](#_Toc533438567)

[4.1 引言 37](#_Toc533438568)

[4.2 深度学习基本理论 38](#_Toc533438569)

[4.2.1 深度学习基本思想 38](#_Toc533438570)

[4.2.2 深度学习框架 40](#_Toc533438571)

[4.3 实验算法 41](#_Toc533438572)

[4.3.1 卷积神经网络原理 41](#_Toc533438573)

[4.3.2 3D-卷积神经网络 44](#_Toc533438574)

[4.4 基于3D-卷积神经网络的高光谱遥感图像分类 46](#_Toc533438575)

[4.5 仿真结果与分析 47](#_Toc533438576)

[4.6 本章小节 55](#_Toc533438577)

[结 论 57](#_Toc533438578)

[参考文献 59](#_Toc533438579)

[攻读硕士学位期间发表的论文和取得的科研成果 67](#_Toc533438580)

[致 谢 68](#_Toc533438581)

第1章 绪论

## 1.1 课题研究背景及意义

遥感（Remote Sensing，RS）是上世纪60年代发展而来的一门新兴技术，目前该技术已经演化为一门用于对地观测的综合性学科。“遥感”可以简单的理解为在不与物体进行物理接触的情况下获取有关物体或现象的信息，与地面观测正好相反。遥感对地观测通常指的是基于卫星或者飞机传感器技术来检测和分类地球上的物体，包括在表面上以及大气层和海洋中的物体[1]。它是人类获取资源环境动态信息的重要手段，是进行科学决策的重要依据，在国家家经济发展中具有重大意义[2]。近60年来，随着人类对地球环境的日益关注，以高分辨率、高动态的和系统化为特征的高新科技不断推动着遥感技术的发展。目前，众多发达国家诸如美、日、英、法、以及发展中国家中国、印度都在竞相发展以高光谱、高动态、高空间分辨率为导向的新式遥感对此观测技术。

当前，遥感对地观测技术的应用已经不局限于传统的土地规划、军事、气象预测、生态环境监控、资源勘探、城市规划、农作物生长检测等领域，随着社会和科技的不断发展，遥感对地观测技术不断融入我们的生活[3]。政府机关可以通过遥感影像轻易定位新建房屋以此避免偷税漏税，保险公司通过遥感影像可以观测受灾地形从而最小化损失，甚至遥感技术正在被用于提高太阳能发电效率等等。遥感技术正逐渐渗透于国家经济、生态环境、开发、国防等诸多方面，并发挥着不可替代的作用。然而，面对浩瀚的遥感影像数据，如何有效、快速、智能地进行遥感数据处理是当今遥感届急需解决的问题，这个问题事关遥感对地观测技术的利用价值、可靠性和普及程度等方法。因此，对于遥感技术研究加大投入力度至关重要。

根据卫星传感器探测的电磁波端不同，遥感对地观测技术被分为可见光遥感、红外遥感、微波遥感、紫外线遥感、多光谱和高光谱遥感[4]。其中高光谱遥感(Hyperspectral Remote Sensing, HRS)属于新型遥感技术，相较于传统的遥感技术收集了更多的重要信息。光谱传感器以10nm波长间隔对地物进行信息采集，探测地表物体的发射能量和反射能量，这样对于每个像素都会有一个连续的光谱曲线形成，据此同时获取了地表物体和光谱和地物分布信息，实现了“图谱合一”，这就是高光谱遥感技术。我们可以将高光谱遥感影像看为由一叠二维图像组成的立方体，每二维图像是该高光谱数据一个波段的所有信息，其中每个像素又包含多个波段的光谱信息，如图1.1所示。高光谱图像中每类地物具有不同的辐射强度，这表现为不同地物光谱曲线差别较大，相同地物光谱曲线更类似。基于此，可以通过光谱曲线的特性对高光谱图像进行分类。高光谱遥感技术是当前遥感技术发展的一大趋势，且越来越受到工业届和学术界的重视。如今高光谱遥感技术被广泛应用于地质调查、植被研究、大气科学研究、海洋研究、农业生产等领域。



图1.1 高光谱图像结构

当前对高光谱遥感技术的研究主要包括波段选择、高光谱图像分类、图像解混、超分辨率制图。其中，高光谱影像分类和识别问题是一个重要的分支，对于高光谱技术的发展有着至关重要的意义。高光谱影像分类建立在遥感图像分类的基础之上，考虑到高光谱图像的独有特性，对其中的像元采用特有的方法进行区分和确认。高光谱图像包含着丰富的光谱信息和空间信息，为分类算法在遥感图像分类的发展中带来了很大的机遇。然而，高光谱图像波段数量较多且互相重叠，影像中包含像元数目少，且空间相关性高，这样使得高光谱影像中隐藏着大量需要被充分挖掘的信息。因此，传统遥感图像分类技术不能用于高光谱影像分类中。因而，联合高光谱数据特点构建合适的分类模型用于高光谱影像分类十分必要。基于此，本文利用卷积神经网络和集成学习的优势，对高光谱影像的分类进行了相关研究。并基于提出的算法，设计构建了用于高光谱影像分类的实验系统软件。

## 1.2 高光谱遥感数据特征

相较于传统遥感图像，高光谱遥感图像除了包含二维平面信息，还包含这一维的谱信息，这样就构成一个蕴含大量空间信息和谱信息的高光谱遥感数据。如图1.2所示，这是通过AVIRIS获取的Indiana农田高光谱数据的数据立方体。对于获取的数据，我们可以考虑从多个角度来看解读，从光谱维度而言，每一个波段构成的图像可以看成一个二维空间分布图像；从空间维度而言，每个像素都可以看成一个连续且反应地物特征的光谱曲线。可以将高光谱遥感影像的特点归结为分类优劣两大类。

高光谱遥感影像不同于传统遥感影像识别优势在于：

(1) 光谱分辨率很高，有众多的波段，可以获取地物接近连续的光谱特征曲线，可以根据需要选择提取的波段来突出目标特征。

(2) 在同一空间分辨率下，光谱覆盖的范围更广，能够探测到地物更多对电磁波的响应特征。

(3) 波段数目众多，便于波段之间的相互校正。

(4) 定量化的连续光谱特性为地物光谱机理模型引入了图像分类提供了方便。

(5) 包含着丰富的空间、光谱和辐射信号，综合了多种信息。

高光谱遥感影像不同于传统遥感影像识别劣势在于：

(1) 数据量大，图像包含几十个和上百个波段，数据量是单波段遥感的几百倍；数据存在大量冗余，没有合适的方法，会影响分类精度。

(2) 对于高光谱图像分类一方面要求更高的光谱定标和反射率转换的精度，另一方面又因为成像机理复杂，数据量巨大因为难以对图像进行预处理，诸如几何校正、大气校正、光谱定标和反射率转换.

(3) 波段多而且波段间的相关性高，因此分类需要的训练样本数目大大增多，往往因训练样本不足导致得到的训练参数不可靠，即维数灾难。

(4) 常规的遥感处理模型和方法不适用于高光谱图像分类。如统计学分类模型的参数估计对于谱特征选择有很高的要求。



图1.2 Indiana农田高光谱图像立方体

## 1.3 国内外在该方向的研究现状及分析

由于高光谱数据在光谱分辨率上有着巨大的优势，因此迎来了各国学者专家的广泛关注。在本课题所涉及的方向和技术层面，一部分相关的论文资料和技术模型对论文的顺利开展有着很大的推动作用。接下来将从深度学习及卷积神经网络、集成学习和高光谱遥感图像分类技术三个方面介绍国内外研究现状和成果的分析。

### 1.3.1 深度学习发展现状

深度学习(Deep Learning, DL)作为机器学习算法中的新兴技术，其动机在于建立能用于模拟人脑进行分析学习的神经网络。其实质上是通过神经网络，对观测的数据进行分级特征表示，然后将低层次的特征逐步抽象为高级特征。

深度学习的主体构成是神经网络，而神经网络的最基本单元是神经元。传统神经网络的研究始于上世纪五、六十年代，最初被称为感知机，也被称单层神经网络。然而感知机只能用于简单的线性分类问题，如果加上一层网络，那么其功能会大大增强。直到1986年，Hinton[5]等人提出反向传播(Back Propagation, BP)算法，才解决双层网络需要较复杂计算度的问题。然而双层神经网络仍然存在训练时间长以及局部最优解求解困难等问题，神经网络的相关研究再次陷入谷底。直到2006年，Hinton[6]在*Science*上发表文章，首次提出深度置信网络(Deep Belief Network, DBN)的概念，这才解决了训练多层网络训练耗时长和最优解难以求解两个问题，自此深度学习再此被受到广大学者关注。

短短几年以来，深度学习模型不断被完善，逐渐成为大数据领域飞速发展的一种趋势。2012年，Andrew Ng教授带领团队完成Google Brain项目，使用16000个CPU核心，对包含10亿个节点的深度网络进行训练，使其进行无监督训练，在不提供如四肢长度、长相等人工标注的特征的前提下，使系统可以自行理解并区分‘猫’的类别。同一年的ImageNet竞赛中，Krizhevsky[7]等人训练了一个大型的深度卷积神经网络AlexNet，它包含了8个学习层：5个卷积层和3个全连接层, 该模型借助了GPU强大的计算能力、深层次的CNN网络，以及Dropout和随机梯度下降等技巧，最终赢得了2012年的ILSVRC，实现了top-5的错误率为15.3%的好成绩。2014年，Facebook人工智能研究院(FAIR)构建了一个9层、1.2亿个参数的深度神经网络，将人脸识别正确率提升到97.25%。2015年，微软亚洲研究院(MSRA)的Kaiming He等人[8]设计了一个152层的CNN网络，比之前的深度网络都深，但降低了复杂度且易于训练，使得分类错误率降到3.6%。以及2016年，Google人工智能部门DeepMind开发的AlphaGo智能系统，

以4:1的成绩击败了当前世界围棋冠军李世石，震惊全球。这些发展都向人们展示了深度学习的强大之处。

当前深度学习主要被应用于图像识别、语音识别、自然语音处理等领域，且都取得较好的结果。卷积神经网络(Convolutional Neural Network，CNN)相较于其余类型网络在图像识别领域，可以深层次提取图像抽象层次的信息，进而获得更高的识别精度。考虑到高光谱遥感数据光谱曲线的抽象度极高，符合CNN的学习特性以及前人将其用于遥感影像处理中取得出色的效果，因此深入探究CNN在高光谱遥感影像分类中的应用，十分有必要。

### 1.3.2 集成学习发展现状

集成学习(Ensemble Learning, EL)有时候被笼统的称为提升(Boosting)方法，它被广泛用于分类和回归任务中, 且取得极好的效果。提升方法基于这样一个思想：对于一个复杂任务来说，将多个专家的判断进行适当的综合所得出的判断，要比其中任何一个专家单独判断好[9]。其实，就是“三个臭皮匠顶一个诸葛亮”的道理。

集成学习起源于Kearns和Valiant提出的“强可学习”和“弱可学习”概念，即概率近似正确(probably approximately correct, PAC)[10]可学习理论。目前主流的集成学习可以分为提升(Boosting)、 袋装(Bagging)、堆(Stacking)方法。Boosting算法中最优代表的是Freund和Schapire[11]于1995年提出的Adaboost和Friedman等人[12]在2000年提出的Boosting Tree。Bagging是Breiman[13]于1996年提出，它并行式集成学习方法最著名的代表。2001年，Breiman[14]对Bagging进行扩展优化，提出随机森林算法(Random Forest, RF)。Stacking是一种组合多个模型的方法，主要注重分类器的组合，相比Boosting和Bagging使用较少。

集成学习未来发展趋势分为两大块：集成学习模型的优化和模型的并行化。处于大数据时代，数据来源各不相同，海量的异构数据已经成为大数据智能处理的瓶颈。如何处理多元异构的海量数据已经成为大数据领域急需解决的问题。集成学习具有融合多种及分类器，且可以通过高性能服务器构建大型并行计算系统的能力，因此可以肯定其即将在大数据领域发挥巨大的作用。

目前集成学习由于其具有较好的泛化性能，在诸多领域有着广阔的应用前景。多次Kaggle和阿里天池数据比赛中，基于集成学习的算法都取得最好的成绩。在北美Netflix

电影推荐算法比赛中，一组采用集成算法的队伍获得了第一名。此外，集成学习别广泛用于生物医学、多媒体影像、数据挖掘等领域。2016年，华盛顿大学T.Chen等人[15]提出Xgboost提升算法，引起了全球数据竞赛届的震惊，自此各种基于XGboost算法的策略模型被用于数据竞赛中。据统计，目前Kaggle竞赛中，XGboost凭借其灵活、可靠且准确、计算速度快的特点，在绝大多数回归和分类问题上都有着顶尖的表现。2017年，微软亚洲研究院G.Ke[16]等人提出LightGBM算法不断优化，以此挑战XGboost的霸主地位。相较于XGboost，LightGBM有着更快的训练速度和更高的效率、更低的内存占用、更高的准确率、大数据并行处理能力等优势。考虑到高光谱影像数据光谱曲线复杂且样本较少的特点，探究XGboost和LightGBM在高光谱遥感分类中的应用有着重要的意义。

### 1.3.3 高光谱图像分类研究现状

地物分类是遥感数据处理的主要内容，其主要依据是：同一类别的像元在光谱特征和空间特征上具有极高的相似度，不同类的地物其光谱特征和空间特征具有极大的差别。高光谱数据具有图谱合一的特点，因此利用高光谱数据进行遥感地物识别分类会有更好的可靠性。

目前，国内外学者对高光谱图像特性和高光谱图像分类算法进行大量深入的研究。大量的特征提取方法被提出并用于高光谱影像分类。目前高光谱分类方法大致分为以下几类：

根据像元类别标签的多少可以分为有监督分类、无监督分类和半监督分类。有监督分类指的是所有训练样本需要标签，相应其分类精度比较高，然而有监督分类需要去进行数据标注耗时耗力。对于无监督分类，其训练样本不需要标签，像元根据其自身相似性自发的聚合到一起，这样相同种类的像元就会聚合到一起，不同种类的像元就会想成不同簇，从而完成分类。高光谱遥感影像分类中，无监督分类精度一般不高，因此主要起着辅助作用。半监督分类则相当于有监督和无监督的一种折衷，只需要标记一部分样本，其余样本可以通过领域适配（Domain Adaption）和标签转递[4]（Lable Propagation）等方法来扩充。半监督分类更符合实际的应用场景，因此也是目前高光谱遥感影像分类的一个研究热点。

根据是否利用空间信息可以将高光谱影像分类分为利用空间和光谱结合的分类方法和只依靠光谱特性分类的方法。仅利用地物光谱特性分类的高光谱图像分类方法主要有光谱角匹配（Spectral Angel Mapping, SAM）[17]、光谱关联匹配（Spectral Correlation Mapping, SCM）[18]等等。然而，不同地物可能具有相似的光谱曲线，同类地物的光谱曲线可能存在大的差别。因此，出现了空间信息和谱信息结合的分类算法。当前结合空间信息的分类方式主要有利用马尔科夫随机场（Morkov Random Fields, MRFs）[19]提取空间信息的方法、利用数学形态特征（Morphological Profiles, MPs）[20]、利用稀疏表示(Sparse Representation, SR)[21]以及利用图像分割（Image Segementation）[22]等方法。

支持向量机（Support Vector Machine, SVM）具有严禁的数学理论依据、而且需要设置的参数少、计算速度快并且依据其核特性类可以很好的求解最优解、同时更适用于小样本数据。因此，基于SVM的算法被广泛用于高光谱遥感领域[23]-[30]。此外基于传统统计学习算法如决策树（Decision Tree）、贝叶斯分类器（Bayes）以及神经网络（Nerural Network）的高光谱遥感分类方法都被广泛研究。然而大多数分类算法都或多或少存在着分类困难、信息浪费、表意不清、欠缺逻辑等等问题，考虑到卷积神经网络可以挖掘图像深层次抽象信息的能力和XGboost、LightGBM在数据科学强大的学习能力以及在数据科学竞赛中的出色表现。并且已有众多基于深度学习的高光谱分类算法被提出[31]-[36]，且都取得极高的分类精度，因此深度学习以及新型的集成算法在高光谱遥感影像分类中的应用将会持续火热。

## 1.4 本文的研究内容与组织结构

近年来，深度学习凭借其强大的特征提出能力逐渐被用于高光谱遥感分类中，并有着出色的表现。新式集成算法凭借其快速的训练速度、易于调参等优点在世界各大数据竞赛中独占鳌头。本课题从提高高光谱影像分类精度出发，分别提出了基于传统特征提取算法的KC-LGM高光谱影像分类算法和基于深度学习特征提取方式的3D-CNN高光谱影像分类算法。论文的论文主要内容如图1.3所示，各章节内容安排如下：

第1章，阐述课题的研究背景及意义，介绍高光谱遥感数据的特征和深度学习、集成学习以及高光谱遥感影像分类的研究发展。

第2章，阐述高谱影像的分类流程与评价标准，介绍经典的分类算法和课题需要的高光谱数据集的相关信息，并利用介绍的分类算法对数据进行仿真实验。

第3章，介绍了集成学习的基本理论，提出一种结合聚类算法和LightGBM的高光谱影像分类算法，利用此算法进行了相关实验，验证了该算法相较于传统分类算法的优越性。

第4章，介绍了卷积神经网络相关理论，提出一种基于3D卷积神经网络的高光谱影像分类算法，相较于传统机器学习分类分类算法和传统深度学习分类算法，简化了图像特征提取方式的同时，提高了分类精度。



图1.3 论文主要研究内容

# 高光谱影像分类研究

## 2.1 高光谱影像分类概述

高光谱影像分类是高光谱遥感数据走向实际应用的一个重要环节。大量实验表明，绝大多数物质在波长20-40nm之内则可以被精准识别，高光谱数据的光谱分辨率一般是2-10nm，这样有效的避免了“同物异谱”的不可分状况，这正是高光谱影像数据可以进行地物精确分类识别的重要原因。当前，高光谱影像分类方法主要以计算机技术为支撑，联合数字图像处理技术和人工智能机器学习算法，进而对高光谱影像中感兴趣的未知地物进行识别分类[37]。根据训练集对数据标签的需求，可以将高光谱影像分类算法分为三大类：有监督分类算法、无监督分类算法、半监督分类算法[38]。

 高光谱影像分类的流程可以描述为：对于采集的高光谱数据，首先对其进行预处理，这是因为传感器在数据采集过程中，会受到外界气象条件及传感器内部噪音影像。常用的预处理主要包括辐射校正、大气校正和几何校正。其次，进行空间和光谱信息特征提取，并进行特征融合与优化。最后，将处理后的信息输入分类器，结果则会产生地物类别标签图。整个流程如图3.1所示。



图2.1 高光谱分类流程图

### 2.1.1有监督分类方法

有监督分类算法使用有标签样本中的先验信息训练分类器，以得到分类器的参数，然后利用完成训练的分类器对未知类别的数据进行分类。有监督分类最显著的特点就是分类精度较高，然而因为有监督分类要求所有训练样本有标签，可在实际情况下往往没有足够的先验信息，这样就造成了有监督分类方法的局限性。典型的监督分类主要包括最大似然判别分类、光谱角匹配、最小距离分类、SVM分类等等[39]。

最大似然判别分类[40](Maximun Likelihood, ML)基于传统统计学习中的朴素贝叶斯算法。ML算法可以快速地进行分类，目前已经被广泛应用于高光谱图像分类。其核心思想是根据训练样本的光谱相似度和贝叶斯准则来估计测试样本类别。

假设以高光谱图像包含个地物类别，分别为。为了便于处理，假设每个类别发生的概率分别为且均相等。依据贝叶斯准则，样本x属于某一类的概率可以表示为：

 (2-1)

其中，

 (2-2)

因为不影响判别效果且各类别的先验信息均相等，因此分类问题可以转化为比较概率大小的问题。通过数学求解，ML的判别函数如下：

， (2-3)

也可以表示为：

 (2-4)

其中指测试样本光谱向量，n为光谱波段数，为第i类均值向量，为第i类的协方差矩阵。当且时，如果则样本属于第i类，即：

 (2-5)

ML算法优点是非常容易理解，且便于编码。然而ML算法基于朴素贝叶斯学习与分类算法，因此假设用于分类的特征在类确定的条件下都是条件独立，这种假设必然会以一定程度的精度牺牲作为代价。高光谱数据的特征有特别高的维度，这使得ML算法运算时间很长。高光谱影响分类一般不会提供过多的训练样本，因此过分依赖样本统计特性的ML算法在样本不足时必然会严重影响高光谱影响分类的效果。

### 2.1.2无监督分类方法

无监督分类指的是在没有先验知识的情况下，仅仅依据影像本身的数据特征，利用这些数据代表的地物光谱特征的相似和相异特性来分类[41]。因为为监督算法不需要要先验知识，不需要现场标记且易于实现，因此被广泛应用于遥感影像分类中。然而此方法严重依赖相似性度量，对光谱和噪声变化极为敏感。样本数据以及类别间的不平均往往造成估计不准确的问题，这使得无监督分类方法一般难以取得较高的精度[42]。典型的无监督分类算法包括K均值聚类算法、ISODATA算法以及平行管道法等等[4]。

K均值聚类算法（K-means）是最典型的聚类算法[43]，主要通过迭代地更新各个聚类的聚类中心，直到系统完成最优迭代结果时停止。首先假设样本集为，该数据集可以通过聚类成为*k*个相互独立的簇，也就是被分为K个类别。其簇划分最小误差函数为

 (2-6)

其中是簇的均值向量。直观上来看上式刻画了簇内样本围绕着簇的均值向量的紧密程度，值越小则簇内样本相似度越高。

具体的K均值聚类算法步骤如下：

⑴ 将*N*个样本初始划分成*k*个聚类，其中，且分别计算每个类别的聚类中心和;

⑵ 随机选取一个测试样本**x**，且假设其在中；

⑶ 若，则返回步骤1，否则继续；

⑷ 计算如下隶属度函数：

 (2-7)

⑸ 若，则将**x**从聚类中移动到聚类中；

⑹ 更新各个聚类的中心向量；

⑺ 若连续迭代*N*次均保持不变，则算法停止，否则返回步骤2。

K-均值聚类算法的优点是当处理的数据集分布密集且各类别之间有较大区别时，分类效果较好。其欠缺之处是初始聚类中心的选取对算法性能的影响较大且在先验知识不足的条件下，不容易确定聚类的个数[44]。

### 2.1.3半监督分类方法

高光谱影像分类中存在着大量的无标签样本，通过发掘无标签样本的信息并结合先验信息，这种方式被称为半监督分类算法[45]。在遥感影像分类中，获取标签是一项耗时耗力的工作，然而无标签样本大量存在且较容易获，因此半监督分类被广泛应用。典型的无监督分类方法包括生成模型、协同训练模型、传导支持向量机、基于图的算法、自训练模型、多项逻辑回归等等[46]。

多项逻辑回归[47]（Multinomial Logistic Regression, MLR）可以利用后验概率将无监督信息融入到了分类器中，从而实现了半监督分类理念。在一个m类的有监督分类问题中，令表示有标签样本集，其中表示特征向量，表示相应的类别标签。在m类分类问题中，可以表示为这样的二进制向量，即当样本属于类别c的情况其对应的二进制类别向量中其他项均为0。在多项逻辑递归问题中，后验类别概率可以表示为：

 (2-8)

其中，表示类别c的权重向量。由于样本的后验概率之和为1，这使得类别的权重向量中存在一项冗余的，所以随意选取。这里只需考虑维向量且通过最大化对数似然估计对进行估计，如下：

 (2-9)

在已知先验概率的情况下进行最大后验估计，可以得到：

 (2-10)

事实上，当训练样本集可分的情况下是无界的，因此先验信息有着重要的作用。

虽然MLR算法是建立在线性分类器的基础之上，但是其可以通过核函数的形式将其扩张到非线性分类的情况下，也就是将原始的输入特征向量r进特定地转化，即

。

为了估计各个类别的概率，等价于对公式进行求解，其中的对数似然函数为：

 (2-11)

 (2-12)

先验概率使得中的大多元素为零，所以多项逻辑回归分类器的复杂性得到了较好的控制，同时也提升了多项逻辑回归分类器的再生性能。

式（2-10）中是非二次规划的且log又是非平滑的，这导致求解凸问题（2-10）具有一定的困难。我们通常使用最大-最小化框架将（2-10）转化为一系列的二次规划问题。其中，求解每个规划问题所产生的计算代价为，然而当特征向量的维数较高或图像中类别数较多时，MLR算法有着巨大的计算代价。因此通常使用变量分裂与增广拉格朗日逻辑递归算法对递归因子进行估计，该算法可以极大的减少求解过程中的计算代价。

然而，多项逻辑回归算法的性能非常依赖于所选模型[48]。若所选模型的概率分布与数据分布相符，那么利用无标签样本学习到准确模型参数的可能性较大，若所选模型的概率分布与数据分布不相符，那么学习到的模型参数也不准确从而导致模型学习结果恶化。

## 2.2 评价方法

当完成对高光谱影像的地物分类覆盖后，需要采用相关理论衡量分类结果的好坏。通常把监督图像即地物分布真实信息视为基准，将其与分类结果进行对比，从而完成对分类方法的评估。一般情况下，评估结果分类几类，分别为混淆矩阵、总体分类精度、平均分类精度、Kappa系数[49]。

### 2.2.1混淆矩阵（Confusion matrix）

混淆矩阵判断分类结果与实际地物覆盖信息是否一致，是高光谱遥感影像分类算法最基本的评价指标[38]。通过混淆矩阵，可以直观的看到地物在不同类别中的错分样本数目和在本类别中的正确分类样本数目。假设分类样本类别为n，那么混淆矩阵是一个n\*n的矩阵，如下式定义：

  (2-13)

上式中，任意一个元素m(i,j)表示第j个元素被分为第i个类别的样本数量，则类别i的错分误差为：

 (2-14)

其中表示第i类地物的总数。那么第j类地物的漏分误差为：

 (2-15)

对角线元素是被正确分类样本，该值越大表明该分类方法越好。

### 2.2.2总体精度与平均精度

总体分类精度（Overall Accuracy, OA），也被称为总体精度，用于描述被正确分类的样本占所有分类样本的比例[50]。针对某个地物类求总体分类精度，用于衡量分类方法对该类的分类效果：

 (2-16)

此外亦可以针对全部测试样本求其分类精度，用于衡量分类方法对遥感影像的总体表现：

 (2-17)

最后根据各类别的总体精度，对其求均值，用以计算平均精度（Averaged Accuracy, AA）:

 (2-18)

其中M为样本总数。

### 2.2.2 Kappa系数

Kappa系数（Kappa Coefficient, KC）使用离散分析法计算通过分类策略得到的地物分布结果与真实地物分布之间的相似程度，可以更好的评价分类精度的优劣[51]。当分类结果具有较大不确定性时，Kappa系数利用离散多远分析法，可以较好地应对分类精度过于依赖类别数和样本的问题，因此可以更好的反应分类方法的效果。一般来讲，Kappa系数越高分类精度就越好。当其大于0.75时，说明分类结果和测试样本具有高度一致性，分类方法好；如果其小于0.4，那么说明分类效果不好。如果介意两者之间，那么说明分类结果中等。其计算公式如下：

 (2-19)

## 2.3 实验数据

本文实验用到了四组数据，分别为AVIRIS[52]（Airborne Visible/Infrared Imaging Spectrtometer）获取的Indian Pines高光谱数据、Salinas Valley高光谱数据、KSC高光谱数据，以及 ROSIS（Reflective Optics System Imaging SPectrometer）获取的University of Pavia高光谱数据。选择这几组数据作为本课题的实验数据主要考虑到以下：

⑴ 皆是高光谱影像基准数据，包含的地物覆盖值相对准确，这样有利于重复实验与算法再现，当前这四组数据被广泛用于高光谱影像分类研究中。

⑵ 具有不同的空间分辨率和光谱分辨率，且通过AVIRIS和ROSIS两个不同的传感器获取，这样可以更客观全面的验证分类的有效性。

⑶ 有着不同的观测场景，能验证方向具有更好的泛化性能。Indian Pines主要包含植被覆盖信息，University of Pavia主要包含校园场景信息，Salinas Valley主要覆盖农耕地和种植园场景，KSC主要包含植被信息。

### 2.3.1 Indian Pines高光谱影像

Indian Pines数据[53]由AVIRIS机载成像光谱仪于1992年6月12日在美国印第安纳州西北部Purdue大学农场采集，数据集中每个波段中包含145\*145个像元，波段覆盖范围为400-2500nm。该数据初始有220个波段，排除掉水吸收和噪声等干扰后，可用波段数为200。该数据集共包含16个地物类别，其中三分之二是农田，其余主要是多年生植被，其样本数如表2-1 所示。图2.2给出此遥感数据集的示意图，其中图2.2(a)是假彩色合成遥感图，图2.2(b)给出其地物真值覆盖图。该数据空间分辨率低，且地物类型多为植被，光谱曲线相似度高，这使得该数据的分类任务有一定挑战性。

表 2.1 印第安那农田数据集地物类别

Table 2.1　Land-covers for the Indian Pines

|  |  |  |
| --- | --- | --- |
| 编号 | 地物类别 | 样本个数 |
| 1 | Alfalfa | 54 |
| 2 | Corn-notill | 1434 |
| 3 | Corn-min | 834 |
| 4 | Corn | 234 |
| 5 | Grass/Pasture | 497 |
| 6 | Grass/Trees | 747 |
| 7 | Grass/pasture-mowed | 26 |
| 8 | Hay-windrowed | 489 |
| 9 | Oats | 20 |
| 10 | Soybeans-notill | 968 |
| 11 | Soybeans-min | 2468 |
| 12 | Soybean-clean | 614 |
| 13 | Wheat | 212 |
| 14 | Woods | 1294 |
| 15 | Bldg-Grass-Tree-Drive | 380 |
| 16 | Stone-steeltowers | 95 |
|  |  |  |



(a) (b)



图2.2 印第安那农田数据集(a) 假彩色合成图像(b)参考地物

### 2.3.2 University of Pavia高光谱影像

帕维亚大学高光谱数据集[54]由ROSIS采集于意大利北部地区，帕维亚大学及其周边场景。数据集波段覆盖范围为430-860nm,其中每个波段包含610\*340个像元，共115个波段，去除水吸收和噪声等干扰后留下103个可用光谱波段，空间分辨率约为1.3m。

图2.3给出此遥感数据集的示意图，其中图2.3(a)是假彩色合成遥感图，图2.3(b)给出其地物真值覆盖图。此数据集共有9种地物，主要包含建筑、植被、卵石等，其样本数如表2.2所示。此数据虽然具有较高的空间分辨率，然而其光谱分辨率低，波段范围小且地物种类广泛，这对于研究高光谱影像分类有一定的代表性。

表2.2 帕维亚大学数据集合地物类别

Table 2.2　Land-covers for the University of Pavia

|  |  |  |
| --- | --- | --- |
| 编号 | 地物类别 | 样本个数 |
| 1 | Asphalt | 6641 |
| 2 | Meadows | 18649 |
| 3 | Gravel | 2099 |
| 4 | Trees | 3064 |
| 5 | Metal Sheets | 1345 |
| 6 | Soil | 5029 |
| 7 | Bitumen | 1330 |
| 8 | Bricks | 3682 |
| 9 | Shadows | 947 |



(a) (b)



图2.3 帕维亚大学数据集(a) 假彩色合成图像 (b) 参考地物

### 2.3.3 SalinasValley高光谱影像

Salinas Valley遥感数据[55]由AVIRIS传感器于1998年在美国加利福尼亚的萨利纳斯峡谷采集。此数据集波段覆盖范围为400-2500nm，原始数据包括224个波段。同Indiana Pines数据相比，该数据采集飞行高度低，光谱分辨率极高达到3.7m，且每一光谱波段包含512\*217个像素。去除水吸收和噪声等20个干扰波段，剩余204的光谱波段被用于分类。

图2.4给出此遥感数据集的示意图，其中图2.4(a)是假彩色合成遥感图，图2.4(b)给出其地物真值覆盖图。此数据有着极高的空间分辨率，且覆盖地物空间分布相对均匀，覆盖地区主要是农作物种植区，覆盖的16的地物类别主要为菜园、葡萄园和待耕种的土地等等，其样本数量如表2-3所示。表2.3显示了萨利纳斯峡谷地物类别：

表2.3 萨利纳斯峡谷数据集合地物类别

Table 2.3　Land-covers for the Salinas Valley

|  |  |  |
| --- | --- | --- |
| 编号 | 地物类别 | 样本个数 |
| 1 | Brocoli\_green\_weeds\_1 | 2009 |
| 2 | Brocoli\_green\_weeds\_2 | 3726 |
| 3 | Fallow | 1976 |
| 4 | Fallow\_rough\_plow | 1394 |
| 5 | Fallow\_smooth | 2678 |
| 6 | Stubble | 3959 |
| 7 | Celery | 3579 |
| 8 | Crapes\_untrained | 11271 |
| 9 | Soli\_binyard\_develop | 6203 |
| 10 | Corn\_senesced\_green\_weeds | 3278 |
| 11 | Lettuce\_romaine\_4wk | 1068 |
| 12 | Lettuce\_romaine\_5wk | 1927 |
| 13 | Lettuce\_romaine\_6wk | 916 |
| 14 | Lettuce\_romaine\_7wk | 1070 |
| 15 | Vinyard\_untrained | 7268 |
| 16 | Vinyard\_vertical\_trellis | 1807 |



1. (b)



图2.4 萨利纳斯峡谷数据集 (a)假彩色合成图 (b) 参考地物

## 2.4 仿真实验

高光谱遥感数据样本量少且光谱特征维数极高。支持向量机[56]（Support Vector Machine, SVM）有着数学理论严谨、运算速度快、参数少且易于调节等特征。因此，SVM被广泛用于高光谱遥感影像分类中，且有着相当好的效果。本节将简要介绍基于SVM的高光谱影像分类，并利用SVM分类算法测试实验环境的可行性。

### 2.4.1基于SVM的高光谱遥感分类算法实验

SVM是一种典型的二分类分类模型，通过一定的扩展可以将其很好的用于多分类中。其基本模型是定义在特征空间上的间隔最大的线性分类器，间隔最大使它有别于感知机；SVM还包括一些核技巧的优化，这使得它实质上又成为一个非线性分类器。SVM的学习策略是间隔最大化，可以理解为一个求解凸二次规划的问题，SVM的学习算法是求解凸二次规划的最优解算法[57]。

SVM学习方法包含构建由简至繁的模型，主要包括：线性可分支持向量机、线性支持向量机和非线性支持向量机。简单模型是复杂模型的基础，其本质是复杂模型的特殊情况。当训练数据线性可分时，通过硬间隔最大化，学习一个线性的分类器，即线性可分支持向量机，又称为硬间隔支持向量机；当训练数据近似线性可分时，通过软间隔最大化，也可以学习一个线性的分类器，即线性支持向量机，又称为软间隔支持向量机；当训练数据不可分时，通过核技巧以及软间隔最大化，学习非线性支持向量机[9]。

对于高光谱遥感数据，可以构建基于核方法的支持向量机分类算法来验证试验的可行性。本次实验构建基于线性核函数(Linear)和径向核函数(RBF)的支持向量机学习算法，用来学习高光谱影像的光谱特征，并通过Kappa系数和OA值衡量分类结果。本实验选用Indiana Pines遥感数据集进行测试，利用Python3.6语言和Pycharm集成开发环境进行仿真。

2.4.2仿真结果与分析

本实验中，选用原始Indiana Pines所有数据作为研究对象。由于样本的总数较多，我们在各类别中选择10%的样本作为训练数据，将其比率定义为训练样本率（Traing Data Rate, TDR），把剩余样本作为测试集合。

图2.5(a)-2.5(c)分别是分类标准图和利用RBF-SVM、Linear-SVM作为分类器的分类效果图，其展示将训练好的分类器用于全体样本测试的效果，其值比在测试集上的分类指标稍高。为了找到RBF-SVM分类器对Indiana Pines数据的最好分类效果对应的参数，实验中取RBF-SVM的惩罚系数分别位1、2、4、8、16，取gamma值分别为0.01、0.05、0.125、0.25、0.5，通过网格搜索方法进行实验。通过实验可知，当gamma=0.05,C=8时，RBF-SVM平均分类效果最好，OA值为79.7，Kappa系数为0.763。通过对比发现Linear-SVM性能稍差于RBF-SVM，平均OA值为79.0，平均Kappa系数为0.756。然而，针对此数据集，由于其中Oats以及Grass/pasture-mowed样本数太少，在利用RBF-SVM分类实验中，C的最优取值不太稳定，但是C为8的次数最多且精度最好，因此这里将8作为C的最优取值。此外，通过进行PCA降为后再送入支持向量机的一系列，发现效果欠佳，因此在此不予赘述。(有空在添加一个表格)



（a）原始地物分布图 （b） RBF-SVM分类图 OA=81.68 Kappa=0.79



（c）Linear-SVM分类图 OA=81.36 Kappa=0.78

图2.4 Indian Pines分类效果图

通过实验和以上分析，可以看出单一的SVM分类器不能满足高光谱遥感分类的要求。其表现为，精度不够高且不稳定。因此，研究将集成学习和深度学习的思想用于高光谱遥感影像数据有着重要的意义。

## 2.5 本章小结

本章首先概述高光谱遥感影像分类的理论，将高光谱影像分类步骤抽象为：预处理，特征提取，分类器训练三个步骤。同时，简要介绍了高光谱影像监督分类、无监督分类和半监督分类三种方法。其次，介绍了高光谱影像分类的评价标准和本课题涉及到的高光谱遥感数据，包括：Indian Pines, Salinas Valley, University of Pavia, KSC。最后，采用基于SVM的高光谱影像分类算法对Indian Pines数据集进行了分类实验，完整的完成了高光谱影像的分类流程。本章内容为后序内容提供理论支持。

# 第3章 基于树形集成学习的高光谱图像空-谱分类算法

## 3.1 引言

集成学习是当前机器学习和数据挖掘领域的主要研究方向之一。其本质不是一个单独的机器学习算法，而是通过构建多个有差异性的弱分类器，并基于相关策略对这些若干分类器结果进行整合，从而完成分类预测任务[57]。

近年来基于树的集成学习算法[58](Tree-Ensemble Alogrithm)在工业届中被广泛应用，并且取得了极好的效果。其主要原因在于：Tree-Ensemble Alogrithm相较于传统分类算法具有更好的模型可控性，因此不容易产生过拟合；工业届中的数据必然存在着一定的噪音，Tree-Ensemble Alogrithm通常具有更好的抗噪能力；此外对于样本特征多样性，Tree-Ensemble Alogrithm有着更好的适应性。高光谱遥感数据由高空中光谱仪的采集，因此存在着一定数量的误差，并且包含较高的光谱波段数。因此，考虑将Tree-Ensemble Alogrithm应用于高光谱遥感图像分类有着重要的意义。

当前有监督分类算法已经广泛用于传统数据，并且在高光谱图像分类中有着比较好的精度。然而传统有监督分类算法往往只利用了图像的光谱数据，这样对遥感图像中蕴含的空间信息是一种浪费。为了克服此问题，众多学者采用一定的分类框架将空间信息融入到高光谱图像分类中，显著的提升了分类器的性能。目前主要的空间特征提取方法包括：1）形态滤波特征提取；2）马尔科夫随机场特征提取；3）分割特征提取[4]。

鉴于上述内容，本章提出了基于树形集成学习的高光谱图像空-谱分类算法——KC-LGM高光谱分类算法，采用基于Tree-Ensemble Alogrithm的分类算法并结合K-means聚类算法，从空-谱两个方面更加全面的利用了高光谱像的空间信息。

## 3.2 集成学习基本理论

依据集成学习单个弱分类器的生成方式，目前集成学习方法大致可以分为两大类，即个体弱分类器存在强依赖关系、必须串行产生的序列化方法，其主要代表是Boost方法；个体分类器之间不存在强依赖关系、可以同时产生的并行化方法，其主要代表是Bagging和“随机森林”[57]。集成学习的基本结构如图3.1，本节将简要介绍集成学习相关理论算法。



图3.1 集成学习结构图

3.2.1 Boosting

Boosting是一簇可以将弱分类器提升成强分类器的算法[59]。其主要工作机制类似：首先通过初始训练集训练一个基分类器，再依据基分类器的表现对样本分布进行调整，使得之前被错分的样本后续得到更高的权值，再根据调整后的样本训练下一个基分类器;

不断重复执行，直到分类器数目达到制定值，最终将所有及分类器加权组合。Boosting算法族中最有代表性的算法是Adaboost[60]和GBDT(Gradient Boosting Decision Tree)。

3.2.2 Boosting Tree

Boosting实际上采用的是基分类器线性组合和前向分布算法的一个求解最优解过程，将决策树作为基函数的Boosting被称为Boosting Tree[61]。在分类问题中决策树是一个二叉分类树，而回归问题中决策树是一个二叉回归树。Boosting Tree的模型可以表示为决策树的加法模型：

  (3-1)

上式中， 表示单个决策树；M代表决策树数目；表示决策树的参数。

Boosting Tree采用前向分布算法，首先确定初始提升树，第m步的模型可以表示为：

 (3-2)

上式中表示当前模型，下一颗决策树的参数可以通过经验风险极小化方法确定，

 (3-3)

因为树的线性组合可以很好的拟合训练数据，即使数据输入输出之间关系很复杂，因此提升树是一个有效的分类器。

Boosting Tree的描述如下：

输入：训练数据集

输出：提升树

(1) 初始化

(2) 对m = 1,2,3…M

(a) 将损失函数的负梯度作为残差

(b) 拟合残差学习一个回归树，得到

(c) 更新

(3) 得到回归问题的提升树



### 3.2.3 *GBDT*

GBDT(Gradient Boosting Decision Tree)是机器学习中一种被广泛应用的分类算法。其本质是将Gradient Boosting算法作为Boosting Tree的优化策略，进而实现的一种Boosting Tree算法。对于Boosting Tree，当损失函数是平方损失和指数损失函数时，每一步的优化都很简单。但对于一般的损失函数，每一步的优化存在着一定难度。

基于此问题，Frediman提出Gradient Boosting算法[62]。本质上来讲，这是一种Gradient Descient算法的近似方法，其关键在于利用损失函数的负梯度在当前模型的值

 (3-4)

作为回归问题Boosting Tree算法中的残差近似值，进而拟合一个回归树。本章提出的高光谱遥感分类算法便是基于GBDT的一种变式算法LightGBM。

GBDT算法描述如下：

输入：训练数据集损失函数

；

输出：回归树。

(1) 初始化

 (3-5)

(2) 对

(a) 对, 计算

 (3-6)

(b) 对拟合一个回归树，得到第个树的叶子结点区域，

(c) 对, 计算

 (3-7)

(d) 更新

(3) 获得最后的回归树

 (3-8)

### 3.2.4 *Bagging*

欲得到泛化性能好强的集成，那么集成中的多个基学习器应该尽可能相互独立。然而“独立”在现实任务中无法做到，那么可以力求基学习器尽可能具有较大的差异。基于此，可以考虑对同一个数据集进行采样，生成多个不同的子集，再根据不同的自己训练出多个基学习器。然而，一方面由于训练数据不同，产生的分类器有着较大的差异。但是为了有更好的集成效果，单个分类器效果不能太差。基于此，可以采用互有交叉的的采样子集进行训练。

Bagging[63]是并行集成学习算法的代表，其训练过程基于自主采样法(bootstrap sampling)。Bagging的基本流程可以描述为：给定一个包含m个样本的数据集，随机选取一个样本放入采样集，再将该样本放回初始数据集，使其下次依旧可能被选中，经过m次随机采样就会生成一个包含m个样本的采样集。这样初始数据集中就会有一部分出现在采样集，另一部分没有出现。如此，可以产生N个包含m个训练样本的采样集，基于每个采样集可以训练出一个基学习器，最后将N个基学习器进行集合。最后，依据分类和回归任务，选择恰当的决策方式。

### 3.2.5 随机森林

随机森林(Random Forest, RF)[64]是bagging的一个扩展变体。RF在一树为基学习器构建集成学习系统的基础上，在决策树的训练过程中引入了随机属性选择。不同于传统决策树选择划分属性时是在当前节点的属性集合中选择一个最优属性，而对于RF中基学习器的每个节点，先从该节点的属性集合中随机选择k个属性子集，然后再从这个子集中选择一个最优属性用于划分。其中k值控制了随机性的引入程度，如果k等于训练集的属性数，那么基学习器的构建等同于传统决策树。

随机森林具有简单、容易实现、计算开销小以及强大的性能等优点，因此在现实生产环境中被广泛应用。和Bagging方法相比，随机森林仅仅做了一点改动，就实现了其基学习器来自样本扰动和属性扰动的多样性，这就使得其最后的集成泛化能力可以通过个体学习器之间的差异度增加而进一步提升。

## 3.3 实验算法

### 3.3.1 LightGBM

LightGBM由微软DMTK团对于2017年发布，其本质是对Gradient Boosting算法的一种改进。现有的GBDT实现算法如XGBoost和pGBRT相较于传统的GBDT有了极好的改进效果，但当特征维度极高且数据量很大时，pGBRT和XGBoost都不能提供较好的效率和伸缩性。为了解决此问题，DMTK团队提出GOSS(Gradient-based One-Side Sampling)和EFB(Exclusive Feature Bundling)两种算法，并将其用于GBDT的求解过程中。实现效果表明，LightGBM相比于传统的GBDT算法，在将训练速度提高20倍的同时也实现了更高的精度。

GOSS算法基本思想可以描述为：在GBDT算法中，如果数据样本没有初始权重。那么不同梯度的数据样本在计算信息增益的时候有不同的作用，梯度较大的数据样本有着更高的信息增益。因此，在对数据集进行下采样时，为了保证信息增益估计的准确性，应该保留具有较高梯度值的样本，而随机的去除有较低梯度的样本。基于此算法，如果下采样比例子固定，可以获取相较传统随机下采样更好的分类精度，尤其是当信息增益有着较大的范围时。

EFB算法基本思想可以描述为：在实际应用中，虽然众多数据集有着极高的特征维度，比如高光谱数据中的波段数目，然而特征空间普遍是稀疏的。基于此，可以考虑设计一种近似无损的压缩方法来减少有效的特征维度。具体而言，稀疏的特征空间中，许多特征是互斥的，比如这些特征会很少同时取非零值。因此，可以对这些互斥的算法进行相应捆绑，使之成为一种特征。基于此，EFB算法将优化捆绑特征问题转化为图着色问题（通过将特征设为图的顶点，如果两个特征不是互斥的，那么在这两种特征之间加一条边），最后经过贪心算法进行了求解。通过这种方式，提高了算法的执行速度。

LightGBM的具体实现中还使用到Histogram算法，通过把连续特征值的分为许多离散的小“桶”，进而在这些“桶”中寻找基分类器的分散点，以此减小了存储成本和计算成本。此外大多数决策树的分类算法是通过层次(level-wise)分裂节点而进行生长，如图3.2(a)所示，这样势必会造成较大的计算开销，而LightGBM使用基于叶子分裂(leaf-wise)节点进行生长，如图3.2(b)所示，当生成相同的叶子时，leaf-wise策略可以比level-wise减少更多的精度。但是，当数据较少时，leaf-wise可能会产生或拟合，但可以通过控制树的深度和叶子数目以避免过拟合的产生。



\

(a) Level-wise tree growth



(b) Leaf-wise three growth

图3.2 树的生长方式

### 3.3.2 空间聚类方式的空谱信息获取

传统像元级别的分类方法在不利用空间信息的情况下，单独处理各个像元。然而，空间邻近像元通常有着相似的光谱特性，将像元的光谱信息与空间上下文结合可以减少样本的不确定性。因此，空间上下文信息和光谱信息同等重要。

本章利用K-Means聚类算法，实现了一种轻量级的空谱信息提取方式，图3.3展示了本文空谱特征的生成方式。不同于以往每个像元的特征由其光谱值表示，我们提取一个的光谱立方块，立方体中蕴含的信息表示该立方块中中心像元的特征值。为了能均匀的表示每个像素的信息，一般取奇数，代表光谱维度。对于边缘和边角像元，可以通过依据其周围像元的信息对进行镜像填充操作，类似于卷积神经网络中的填充(padding)操作，但又有所不同。对于每一维度，可以获取N个的矩阵分别代表该维度各像元的信息，其中N为该维度像元的个数。最后通过将N个的矩阵分别向量化为的向量，并通过对这N个向量进行聚类，得到K个聚类簇，最后经过对各个簇和各像元的相关处理得到该维度相应空间信息。高光谱图像有着N个维度，通过将N个维度获取信息级联，则形成了整个图像的空-谱信息

每一个像元所属的矩阵块都被映射为一个的向量，通过求该向量和K个聚类中心的2范数可以将该像元映射到一个K维向量中，其具体细节如下式：

 (3-9)

其中代表像元到聚类中心的2范数，代表关于i的均值。需要注意的是，当K=1时，表示在求像元与其自身所在簇的相应关系，因此可以取其2范数的倒数来代表此种情况的信息。是一个很小的值，其存在是为了避免2-范数为0，其倒数不存在。

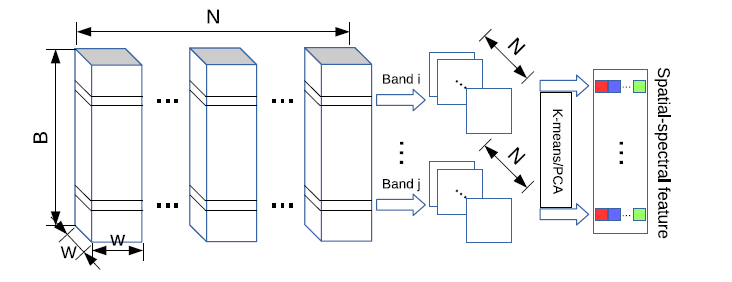
 经过以上的过程，针对每一维度，每个像元都保存了个分量，考虑到一共有N个维度，那么该像元就会获取维的空-谱信息。

图3.3 基于K-means的空-谱信息提取方式

### 3.3.2 基于LightGBM和K-Means的高光谱图像空-谱分类算法

本节选用LightGBM作为基础学习器，通过上文提出的基于K-means聚类算法的空-谱信息提取框架整合高光谱图像的空谱信息，实现了一种有着较高精度的分类算法——KC-LGM高光谱图像分类算法。其具体步骤如下：

(1) 通过R-PCA[65]操作对高光谱图像进行降维操作，减少计算量。

(2) 确定代表像元信息的矩阵块大小，确定镜像填充数量。

 (3-10)

(3) 依据值，对遥感图像每一维度信息进行填充操作，使每一维图像形状变为，其中代表原始图像的行数，代表原始图像的宽度，具体操作如图3.4所示。

(4) 为了产生较好的聚类效果，对每一维的图像中的每个信息矩阵块进行标准化操作：

 (3-10)

(5) 假设每一维有个像元，那么相应的就有个信息矩阵块，将个矩阵块分别转换为向量后，进行K-means聚类操作，生成个聚类簇。

(6) 通过计算每个像元信息矩阵到这个聚类中心向量的2范数，将像元矩阵 映射到一个维向量，中的个元素可以通过下式获得：

 (3-11)

(7) 对降维后的高光谱图像的每一维进行上述(2)-(6)操作，并将每一维度的信息进行级联，从而形成整个图像的空-谱信息。

(8) 联立空谱信息和样本标签并对数据集进行归一化操作，抛弃图像背景提取有用标签，并进行测试集合训练集的划分。

(9) 构建lightGBM学习器，将训练集送入该学习器。

以上就是该算法的详细步骤，其分类流程图如图3.4所示：

## 3.4 实验仿真结果与分析

**参数设置以及性能指标**

本章仿真实验是基于谷歌Colaboratory[66]实验平台，利用python语言进行编程仿真。为了验证本章所提出算法的有效性，本节选取PaviaU和Salinas Valley高光谱数据作为实验对象。PaviaU数据集包含9个地物类别，其地物参考图如图3.5(a)所示。Salinas Valley数据集包含16个地物类别，其地物参考图如图3.5(b)所示。

针对聚类方式的空-谱信息提取预处理，经过网格搜索法确定当聚类中心数K=5时，分类效果平均最好，因此实验中默认聚类中心为5。针对LightGBM算法，将迭代次数设置为1000。实验中采用OA和Kappa系数验证算法的效果。

(a) (b)

图3.5 地物分布图（a）PaviaU数据集（b）Salinas数据集

**实验I：**

本实验用来探究KC-LGM算法中中相关参数的设置对实验结果的影响，重点考察的参数是空-谱信息提取过程中，对原始数据集进行降维保留的主成分个数DIM和空间窗口大小FIL，这里DIM和FIL分别设为{10,20,30}和{3,5,7}；LightGBM算法中决策树的个最大深度max depth和算法学习速率learning rate，这两个数据分别设为{6,7,8}和{0.05,0.1,0.3,0.5}。对于两组高光谱数据，均随机选取10 %的数据作为训练集，其余90%的数据作为测试集。在探究DIM和FIL对分类的影响的实验，将max deapth和learning rate分别设为7和0.5；在探究max depth和learning rate对分类的影响的实验中，将DIM和FIL分别设为10和5。

表3.1和3.2分别给出两个数据在选取不同DIM和FIL时的分类表现，可以发现对于paviaU，当DIM为20时精度普遍不好，当DIM=30且FIL=7时分类效果最好，DIM=10时FIL=7时表现不好。这说明采用此种方法提出空间信息选取合适的参数至关重要，不合适的参数选择可能会导致样本特征出现大量冗杂信息，从而极大降低分类精度，这是选择不同的FIL和DIM会导致样本包含不同的特征数量，进而会带来产生冗余信息的可能。对于Salinas数据，可以发现在实验范围内，DIM和FIL值越高，分类效果越好。其中，当DIM = 30 , FIL = 7时算法的分类效果如图3.6(a)和3.7(b)所示。

表3.3和3.4给出了两个数据集在选取不同learning rate和max depth时的分类表现，通过上边实验可知，当DIM=10且FIL=5时，两数据集均有着极好的表现，因此在进行learning\_rate和max depth的实验时，在保证精度的前提下考虑到实验的运行速度，将

DIM和FIL分别设为5。对于paviaU数据集，可以看出在所选参数的前提下，分类的精度为一个量级，相对而言learning rate为0.3时分类效果比较好。对于Salinas数据，当learning rate为0.5时，如果盲目增加决策树的最大深度可能会导致精度急剧下降，这是由于对于特定数据集，过大的learning rate可能会使得进行梯度下降步伐过大，从而不能达到全局最优，甚至获得极差的分类表现。

(a) OA=98.78% (b) OA=99.6%

图3.6 分类效果图（a）PaviaU（b）Salinas Vallye

表3.1 针对PaviaU数据集不同主成分数和窗口大小对性能的比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 性能 | *DIM* | *FIL=3* | *FIL=5* | *FIL=7* |
| OA | 10 | 96.1 | 98.32 | 53.1 |
| 20 | 82.7 | 58.47 | 90.07 |
| 30 | 97.2 | 98.65 | 98．78 |
|  | 10 | 0.948 | 0.97 | 0.34 |
| Kappa | 20 | 0.77 | 0.41 | 0.87 |
|  | 30 | 0.96 | 0.98 | 0.983 |

表3.2 针对Salinas Vallye数据集不同主成分数和窗口大小对性能的比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 性能 | *DIM* | *FIL=3* | *FIL=5* | *FIL=7* |
| OA | 10 | 95.79 | 98.18 | 99.20 |
| 20 | 97.3 | 99.01 | 99.62 |
| 30 | 97.5 | 99.14 | 99.57 |
|  | 10 | 0.95 | 0.98 | 0.991 |
| Kappa | 20 | 0.972 | 0.991 | 0.993 |
|  | 30 | 0.973 | 0.992 | 0.995 |

表3.3 针对PaviaU数据集不同学习速率和决策树最大深度对性能的比较

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 性能 | *Depth* | *LR=0.05* | *LR=0.1* | *LR=0.3* | *LR=0.5* |
| OA | 6 | 98.27 | 98.48 | 98.49 | 98.36 |
| 7 | 98.44 | 98.45 | 98.46 | 98.40 |
| 8 | 98.42 | 98.49 | 98.50 | 98.37 |
|  | 6 | 0.977 | 0.979 | 0.980 | 0.978 |
| Kappa | 7 | 0.979 | 0.979 | 0.979 | 0.987 |
|  | 8 | 0.979 | 0.979 | 0.980 | 0.987 |

表3.4 针对Salinas Valley数据集不同学习速率和决策树最大深度对性能的比较

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 性能 | *Depth* | *LR=0.05* | *LR=0.1* | *LR=0.3* | *LR=0.5* |
| OA | 6 | 98.37 | 98.57 | 98.60 | 98.5 |
| 7 | 98.59 | 98.64 | 98.61 | 26.77 |
| 8 | 98.62 | 98.65 | 98.62 | 73.51 |
|  | 6 | 0.981 | 0.984 | 0.984 | 0.983 |
| Kappa | 7 | 0.984 | 0.985 | 0.984 | 0.110 |
|  | 8 | 0.984 | 0.985 | 0.985 | 0.705 |

**实验II：**

本实验的主要目的是为了验证提出的算法的分类效果高于传统分类算法。根据上一个实验，我们发现当DIM=10，FIL=5, learning rate=0.3, Depth=7时，KC-LGM算法在保证精度的前提下有着较快的执行速度，因此本节采用上述参数进行实验。首先，比较在不提取空间信息的情况下各基本分类器在两个数据集上的分类效果，选择的对比分类器分别为决策树、传统GBDT以及线性SVM，其分类效果图如图3.7和3.8所示。其次，比较通过本章提出方法提取空间信息后各分类器的执行表现，其分类效果图如图3.9和3.10所示。通过两组实验结果，可以看出本章算法是否利用空间信息效果都优于传统分类算法，

本章算法对两个数据集中各类别的分类结果如表3.5和3.6所

(a) KC-LGM OA=92.86% Kapp=0.904 (b) Linear SVM OA=89.88% Kappa =0.864

（c）DT OA=82.79% Kapp=0.771 (d) GBDT OA=90.95% Kappa =0.878

图3.7 针对PaviaU数据集针仅利用光谱信息不同算法的分类结果比较

 （a）KC-LGM OA=94.36% Kapp=0.937 (b) Linear SVM OA=93.02% Kappa=0.922

（c）DT OA=88.57% Kapp=0.873 (d) GBDT OA=93.29% Kappa=0.925

图3.8 针对Salinas数据集仅利用光谱数据不同算法的分类结果比较

表3.5 PaviaU数据集各地物分类结果

|  |  |  |  |
| --- | --- | --- | --- |
| 编号 | 地物类别 | 样本个数 | OA |
| 1 | Asphalt | 6641 | 0.99 |
| 2 | Meadows | 18649 | 1.00 |
| 3 | Gravel | 2099 | 0.90 |
| 4 | Trees | 3064 | 0.99 |
| 5 | Metal Sheets | 1345 | 1.00 |
| 6 | Soil | 5029 | 0.99 |
| 7 | Bitumen | 1330 | 0.94 |
| 8 | Bricks | 3682 | 0.95 |
| 9 | Shadows | 947 | 0.99 |

表2.3 萨利纳斯峡谷数据集合地物分类结果

|  |  |  |  |
| --- | --- | --- | --- |
| 编号 | 地物类别 | 样本个数 | OA |
| 1 | Brocoli\_green\_weeds\_1 | 2009 | 1.00 |
| 2 | Brocoli\_green\_weeds\_2 | 3726 | 0.99 |
| 3 | Fallow | 1976 | 1.00 |
| 4 | Fallow\_rough\_plow | 1394 | 0.97 |
| 5 | Fallow\_smooth | 2678 | 1.00 |
| 6 | Stubble | 3959 | 1.00 |
| 7 | Celery | 3579 | 1.00 |
| 8 | Crapes\_untrained | 11271 | 0.98 |
| 9 | Soli\_binyard\_develop | 6203 | 1.00 |
| 10 | Corn\_senesced\_green\_weeds | 3278 | 0.99 |
| 11 | Lettuce\_romaine\_4wk | 1068 | 1.00 |
| 12 | Lettuce\_romaine\_5wk | 1927 | 1.00 |
| 13 | Lettuce\_romaine\_6wk | 916 | 0.94 |
| 14 | Lettuce\_romaine\_7wk | 1070 | 0.99 |
| 15 | Vinyard\_untrained | 7268 | 0.95 |
| 16 | Vinyard\_vertical\_trellis | 1807 | 0.99 |



1. KC-LGM (b) SVM (c)DT (d) GBDT

OA =98.46 OA =98.46 OA = 93.85 OA = 96.84

Kappa=0.979 Kappa = 0.974 Kappa=0.91 Kappa = 0.958

图3.9 针对PaviaU数据集提取空间信息后各分类算法的分类结果比较



1. KC-LGM (b) SVM (c) DT (d) GBDT

OA =98.5 OA =96.95 OA =95.13 OA = 97.34

Kappa=0.983 Kappa = 0.966 Kappa=0.94 Kappa = 0.97

图3.10 针对Salinas数据集提取空间信息后各分类算法的分类结果比较

## 3.5 本章小结

本章提出了经过聚类预处理提取空间信息的基于LightGBM的高光谱图像分类算法，首先利用R-PCA降维算法对原始图像降维，然后对每一维数据进行聚类获取各维的中心向量，在通过求个区域向量到中心向量的2范数确定各个维度的空间信息，最后将各维度信息级联得到空-谱联合信息。将所得空-谱信息送入LightGBM分类器，最终得到分类结果。LighGBM在加快分类速度的同时，保证了强大的学习能力，保证了本章提出的高光谱分类算法的可行性。

本章进行了一系列实验，首先探究了相关实验参数对分类效果的影响。其次，在不进行提取空间信息的条件下，将LightGBM和其余传统经典分类算法进行对比实验，验证了LightGBM算法在高光谱数据分类中有着良好的表现。最后，利用本文提出的空-谱信息提取方式，将LightGBM与其余算法进行对比，验证了本章提出算法的有效性。然后，本文所提出的的空-谱信息提取方法需要消耗大量的内存，运算开销极大，必须对原始数据进行降维处理，这样势必会造成对光谱信息的浪费。因此，对于如何进行优化算法，减少对计算机资源的利用将是本课题的后续研究方向之一。

# 第4章 基于3D卷积神经网络的高光谱图像分类

## 4.1 引言

高光谱图像蕴含着丰富的光谱信息，大量的传统分类方法如K-NN、最小距离分类法以及逻辑回归等算法均被应用于高光谱图像分类中。近年来，许多有效的特征提取方式和高级分类算法被提出，如空-谱信息分类、局部Fisher判别分析等等。其中，基于SVM并联合空-谱信息的分类器被证明是兼顾稳定性与精度的高光谱图像分类算法。

传统的空-谱信息分类可以被分为两大类。第一种是分别利用光谱特性和空间上下文信息。空间依赖信息可以通过各种滤波器提前提取出，如形态轮廓[67]-69]、熵[70]、属性轮廓[71]、低秩表示[72][73]等等。然后将提取到的空间信息和光谱信息合并，依据具体情况选择降维或者将联合信息直接送入学习器进行训练。此外，还有学者利用空间依赖信息，通过一些正则化方法如马尔科夫随机场[74]和图割理论[75]来细化分类结果。第二类方法主要是通过融合空间信息和光谱细信息去制造联合特征。比如，可以通过对高光谱遥感数据应用一系列不同尺寸和频率的3D小波滤波器[76]和3DGabor滤波器[77]提取空-谱联合特征。同样，可以再利用降维技术在保留判别信息的同时提取到低维的空-谱特征。由于高光谱数据通常以3D数据立方形式呈现，这种方法可以产生大量包含空间、光谱、空-谱联合信息的立方块，这些信息对于提高分类精度有着重要的意义。然而，这类特征提取方法基于“浅”层学习提取出的手工制作特征，严重的依赖于领域内知识。手工特征在一些情况下可能并不能考虑到埋藏在数据中的真实细节，这样就会存在一些算法只适用于一部分数据的情况。

人工神经网络，如多层感知机、RBF神经网络在遥感数据分类中的应用也被众多学者研究。事实上，对于高光谱图像分类，SVM从效率和精度两个方面都优于传统神经网络。进年来，深度学习技术发展火热，越来越多学者将深度学习技术用于高光谱图像分类的研究中，且有众多基于深度学习的算法取得了优于支持向量机的效果。不同于传统“浅”层的手工制作特征，深度学习可以自动的从原始数据集学习到从低级到高级的结构化特征。利用这些学习到的特征进行分类，已经在许多机器视觉任务中取得了巨大的成功。陈等人利用栈式自动编码机(SAE)[78]和深度置信网络(DBF)[79]两种非监督深度学习算法提取空-谱特征并分类。这种方法虽然解决了提取深层特征的需求，但是却存在两个问题：没有充分利用样本标签信息；必须将输入图像样本展平到一维，因此损失了部分空间信息。赵等人[80]-83]采用卷积神经网络(CNN)进行高光谱图像分类，提取图像前几个主成分后采用2D-CNN获得空间信息。这种方法取得了优于SAE和DBF的方法，但是这类算法为了适应2D-CNN的可控范围避免生成过大的卷积核，对图像进行了极大的降维操作且分开提取空间和光谱特征，这种策略势必不能充分提取到空-谱特征。

鉴于上述内容，本文提出一种基于3D卷积神经网络的高光谱图像分类算法。解决了两个问题：在充分提取空-谱特征的同时，避免产生过长过深的网络结构，减轻了计算机的运算负载；相较于传统高光谱图像算法极大的提升了高光谱图像的分类精度。

## 4.2 深度学习基本理论

### 4.2.1 深度学习基本思想

深度学习最早起源于1957年Frank Rosenblatt提出的感知机[84](Percepton)。感知机是只有单层计算单元的神经网络，由线性单元和阈值单元构成。感知机模型如图4.1所示。

图4.1 感知机

其数学模型为：

 (4-1)

其中是阶跃函数：

 (4-2)

 (4-3)

感知机一种简单的二分类器，其分类边界条件为：

 (4-4)

感知机是整个神经网络的基础单元，神经元输出通过激励函数确定，神经元之间通过权值传递信息，权值通过误差进行调节。然而，单层感知机只针对线性问题有分类能力，对于非线性问题却没好的解决方案。

一条直线不能解决分类问题，那么弯曲的折线会不会有着较好的效果。基于此，多层感知机被提出，具体实现是通过在单层感知机的输入和输出层之间加入隐藏层。多层感知机模型如图4.2所示。



图4.2多层感知机

随着隐藏层数增多，多层感知机的的分类凸域可能成为任意形状，因此理论上可以解决任何复杂问题。然而，由于隐藏层的各节点不存在期望输出，导致不能通过感知机的学习规则来训练多层感知机，不能确定隐藏层的参数。

对于如何确定隐藏层参数，误差反向传播算法(Backpropagation, BP)[85]被提出。BP算法由两个过程组成：信号正向传播和误差反向传播。正向传播时，样本从输入层传入，经各隐层处理后到达输出层，如果实际输出和期望输出不符，那么进入反向传播阶段。反向传播时，将输出以某种形式通过隐层逐渐向输入层反向传递，同时将误差均摊分给各层所有单元而获得各单元的误差信号，并以此修正各单元权值。然而BP网络存在一定问题：利用局部梯度下降对权值进行调整会出现梯度弥散现象，其原因在于非凸目标的损失函数的求解可能会陷入局部最优而非全局最优。随着网络层数越深，反向传播的梯度的幅值急剧变小，造成整体损失函数相对最初几层的权值的倒数非常小。这样当使用梯度下降法时，最初几层的权值会变化异常缓慢，进而不能有效的学习样本特征。

深度学习便是为了解决神经网络以上缺点而提出，对于深度神经网络参数难以训练为最优的问题，可以通过逐层次的训练求解，即将本层训练好的结果作为下一层训练的初始化参数。此外，深度学习中用于解决梯度弥散问题的方法还包括：合理化的参数初始化、Relu激活函数的应用以及残差网络中提出的恒等映射(Identity)[86]等技术。深度学习是一系列在信息处理阶段利用非监督特征学习和模式分类功能，具有多层分层体系结构的机器学习技术[87]。其本质在于对数据进行分层特征表示，实现将低级特征逐步抽象为高级特征表示。

深度学习一般分为三类：(1) 生成型深度结构：生成型深度结构旨在模式识别分析过程中描述观察到的可见数据的高阶相关属性，或者描述可见数据和其相关类别的联合分布。生成型深度模型主要包括：自编码机、受限玻尔兹曼机、深度置信网络等等。(2) 判别型深度结构：此类结构主要通过描述可见数据类别的后验概率分布，从而为分类任务提供辨别力。此类深度模型主要的代表是卷积神经网络。(3)混合型深度结构：此类结构主要用老对数据进行判别分类，其包含了生成和判别两部分结构。此类深度模型的代表是通过深度置信网络进行预训练后的深度神经网络。

传统的学习器如SVM、LR、HMM等都属于浅层次的结构，其共同点是采用一层简单结构将原始输入信号或者特征转换到特定问题的特征空间中。其局限性在于对复杂函数的表示能力有限，针对复杂分类问题其泛华能力受到一定制约，难以解决一些更加复杂的自然信号处理问题，如人类语音和自然图像等问题。然而，深度学习以通过学习一种深层非线性网络结构表示输入数据特征，从而实现复杂函数的逼近，并且展现了强大的从少数样本集中学习数据集本质特征的能力。

### 4.2.2 深度学习框架

由于深度学习技术异常火爆，为了深度学习研究者的方便，各大公司、研究院和高校发布其深度框架，极大地简化了深度学习网络的搭建和系统的参数调优。

Theano[88]由加拿大蒙特利尔大学Lisa Lab团队于2018年开发，是一个高性能的符号计算和深度学习库。Theano诞生时间最早，因此可以算成是此类库的始祖之一。应用Theano可以将用户定义的各类型的计算编译为高效的底层代码，并链接CUDA[89]等加速库。然而，Theano存在这些问题：调试功能不能强大，调试信息难以读懂；在CPU上执行效率低下；编译复杂模型非常耗时。因此，在经过10年的发展后，Theano团队于2017年9月29日宣布停止对Theano的维护。

Caffe[90]是由UC-Berkeley贾扬清博士于2013年发布，目前由伯克利视觉中心维护，贾扬清博士也是另一个深度学习框架Tensorflow的作者之一。Caffe的核心单元是Layer，每个神经网络中每一模块都是一个Layer。Layer接收数据后，经过内部运算产生输出数据。通过拼接各Layer，那么就能构建一个神经网络系统。Caffe是第一个主流的工业级框架，复杂的网络计算、梯度求解以及科研领域内新算法一般都能快速在Caffe内找到实现方案。然而，Caffe对递归型网络和自然语言建模能力比较差。因此，Caffe2[91]在后续被提出。相较于Caffe，Caffe2将其基本单元变为Operator，对于适当数量和类型的输入参数，每个Operator都包括所需的计算逻辑，解决了网络定义的问题。

Tensorflow[92]是由Google AI团队研发，是目前流行最为广泛且功能最为强大的深度学习框架。Tensorflow基于数据流图计算，且包含内置一些上层组件，使得搭建网络变得容易。Tensorflow最主要的特点是其灵活的移植性，可将同一份代码不经修改就部署到包含任意数量CPU和GPU的个人电脑、服务器以及移动终端中。此外相较于Theano，Tnesorflow有着极快的编译速度。因次，Tensorflow是当前工业界应用最为广泛的深度学习框架。

Pytorch[93]是Facebook FAIR实验室开元的深度学习框架，PyTorch基于原生的Torch，为其增加了Python API。Pytorch利用Python语言编程，使用动态图定义神经网络，支持变量自动求导，对神经网络层和损失函数优化等操作进行了高层封装。相较于Caffe和Tensorflow，Pytorch使得网络结构定义变得极为简单。Pytorch目前热度不减，成为最主流的深度学习框架之一。

CNTK[94]是微软研究院(MSR)开源的深度学习框架，其目前在语音识别领域应用极为广泛。CNTK通过一个有向图将神经网络描述为一系列的运算操作，有向图中的子节点代表输入或者网络参数，其余节点代表各种矩阵运算。CNKI的设计以性能为导向，在单CPU、单GPU、多CPU、多GPU以及GPU集群上都有极为优越的性能。然而，CNTK目前不支持ARM架构，因此在移动设备上不能应用。

Keras[95]由Google深度学习专家François Chollet开发，是一个崇尚极简和高度模块化的神经网络库。Keras的运行后台可以选择Tensorflow或者Theano，使得操作Tensorflow或者Theano变得非常简单。其旨在快速搭建出网络结构，最快得到实验结果。因此，Keras非常适用于科学研究。

此外还有Amazon开源的Mxnet、Deeplearning4J、PaddlePaddle等优秀的机器学习框架。可见，当前深度学习研究已经达到空前的盛况。

## 4.3 实验算法

卷积网络(convolutional network)是由LeCun[96]于1988年提出，又被称为卷积神经网络(convolutional neural network, CNN)，是一种专门用来处理具有类似网络结构数据的神经网络[97]。当前，卷积神经网络在机器视觉领域取得了巨大的成功，尤其是在图像识别领域，是深度学习领域最为重要的网络结构之一。

### 4.3.1 卷积神经网络原理

相较于传统神经网络，卷积神经网络依旧是层级的网络结构。然而其各层的功能和形式发生了较大的变化，是传统神经网络的一种改进。卷积神经网络主要由数据输入层(Input layer)、卷积运算层(Convolutional layer)、激活函数层(Activation layer)、池化层(Pool layer)和全连接层(FC layer)等构造，其网络结果如图4.3所示。



图4.3 卷积神经网络结构

数据输入层是卷积网络的初始层，主要用于对原始数据进行预处理，其中主要包括：去均值、归一化、降维三个过程。去均值是将输入数据每一维度的中心置0，其目的在于将样本中心拉回坐标原点位置。归一化是机器学习中最常用的操作之一，主要是将每一维的幅度归一化到同样的范围，从而减少各维数据取值范围不同的差异而带来的干扰。

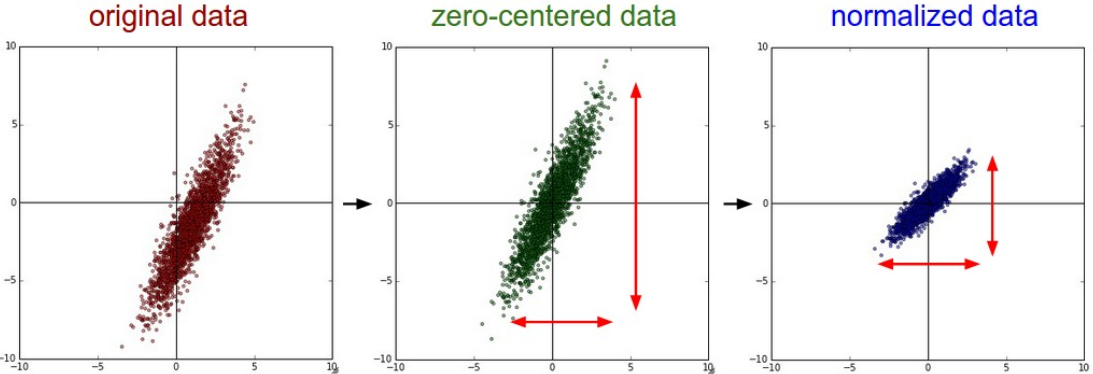
降维过程一般利用PCA提出相关主成分。去均值和归一化的效果如图4.4所示。

图4.4 去均值和归一化操作示意图

卷积运算层是第二层，也是卷积神经网络最为重要的一层，该层主要用来提取数据信息，学习神经网络的权值参数。才卷积层有着两个关键的操作：局部关联、窗口滑动。

局部关联指将每个神经元看成一个滤波器，窗口活动是指滤波器不断滑动，从而不断对局部数据进行计算，提取相关信息。卷积滤波器根据原始数据的通道数，确定其层数。根据滤波后是否保持原始图像大小，确定是否进行”0”填充操作。根据步长操作，确定滤波器每一步的移动步伐。在卷积层，各滤波器的权值是固定的，每一个滤波器只关注一个的特征，由一组固定权重的滤波器和不同图像窗口做卷积运算，这就是卷积神经网络的权值共享机制。对于图像数据来说，这些特征包括垂直边缘、水平边缘、颜色和纹理等信息，每各卷积层的所有滤波器加起来就相当于整个图像的特征提取器集合。权值共享机制极大的减少了需要学习参数的数量，也是卷积神经网络最为关键的一点。

激活函数层直接跟在卷积层后，用于对卷积层输出结果做非线性映射，这是为了保证各卷积层学习到的特征不存在线性关系。CNN中最常用的激活函数是ReLU(The Rectified Linear Unit)[98]，其特点是收敛快，求梯度简单，ReLU的图像如图4.5所示。

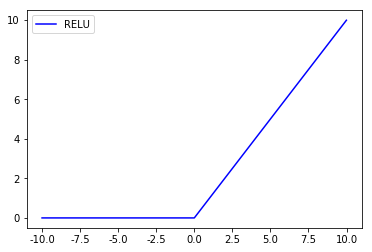


图4.5 ReLU 线性修正单元。

此外，常用的激励还是还包括较早提出的Sigmod以及Tanh和Leaky ReLU和Maxout。

池化层位于两个卷积层中间，主要用于对数据进行一定程度的压缩，减少过拟合。对于输入是图像而言，池化层的作用就是用来压缩图像。具体而言，池化层可以让输出图像保持特征不变性，并对图像进行特征降维。特征不变性相当于图像处理中的特征尺度不变性，对于一张被缩小的图像，依旧可以识别其主体信息，经过压缩去剩下特征不变的信息是最能表达图像特征的。特征降维指对于含有极多的信息和特征的图像，有些信息对于图像处理没有太多用途，可以将此类冗余信息剔除。

全连接层位于卷积神经网络最后阶段，两层网络之间所有神经元都相互连接，相当于传统的神经网络。在网络的最尾部是用于分类的函数，对于二分类是Logistic函数，对于多分类则是Softmax函数。Softmax函数在处理多分类问题中有着重要的作用，Softmax函数定义如下：

 (4-5)

其中，指分类器前置单元的输出。表示类索引，总类别个数为C。代表当前元素指数与所有元素指数和的比值。因此，Softmax函数将多分类输出的连续数值转化为相对概率，更容易理解和比较。

卷积神经网络共享卷积核，对高维数据处理方便；无需手动选取特征，训练好权重就可以自动获取较好的分类结果。然而，卷积网络是一个“黑盒”，我们并不知道其提取的是什么模型。典型的CNN网络主要包括：LeNet、AlexNet、深层LeNet、ZF Net、GoogLeNet以及VGGNet等等。总之，卷积神经网络本质是一种输入到输出的映射，可以学习大量输入到输出的映射关系，然而这一过程并不需要任何输入输出之间精确的数学表达式，只要设计特定网络结构训练好网络，就可以得到输入输出之间精确的映射能力。

### 4.3.2 3D-卷积神经网络

为了利用卷积神经网络进行各种类型的任务，根据传统2D卷积网络衍生出多种类型的卷积神经网络。 根据卷积运算方向和输出形状的区别，卷积神经网络可以分为1D-CNN、2D-CNN和3D-CNN。

 1D-CNN仅从一个方向去进行卷积操作，假设其输出数据形状为，滤波器形状为，通过填充和1D卷积操作，输出数据为，1D-CNN输入输出对应关系如图4.6(a)所示。此外，当输入数据为二维时，保证滤波器形状为，输出数据形状为一维，也属于1D卷积，如图4.6(b)所示。因此，当仅从一个方向进行卷积运算且输出形状为一维，那么就就是一维卷积网络。1D-CNN主要用于处理序列数据，如自然语言处理。

1. (b)

图4.6 1D-CNN

 2D-CNN也是应用最为广泛的卷积神经网络，当前主要被用于机器视觉领域。2D-CNN是指从两个维度进行卷及操作的卷积网络。对于输入形状为的二维数据，卷积核形状为，那么经过填充和二维卷积操作，输出数据形状为，其输入输出对应关系如图4.7(a)所示。对于输入是3维的数据，如RGB图像而言，假设其形状为，经过卷积核为的滤波器处理，输出形为的二维图像，如图4.7(b)所示。后者也是最为常见了2D-CNN形式，经典卷积神经网络如LeNet、VGG等等都属于此种。

 (a)

(b)

图4.7 2D-CNN

 3D-CNN是最晚被提出的卷积神经网络，当前主要用于处理视频、动画等数据，主要用来采集并处理时空特征。3D-CNN从三个维度进行卷积操作，对于输入为的三维数据，采用卷积核为的滤波器在三个方向进行卷积，可以获得形为的输出数据，如图4.8所示。2D-CNN最大的不同在于，3D-CNN从三个方向进行卷积，且得到的输出数据也是三维数据。利用3D-CNN对高光谱遥感数据进行处理，不用对图像进行降维操作，对光谱维也进行学习，因此通过卷积操作同时学习了高光谱图像空-谱特征。

图4.8 3D-CNN

## 4.4 基于3D-卷积神经网络的高光谱遥感图像分类

高光谱遥感图像是一个三维的数据立方体，有着较高的光谱维度。利用2D-CNN进行分类时，为了避免形成一个很大的卷积核从而产生大量待求权值，必须要对原始图像进行降维处理，这样就会导致初始光谱信息的丢失。因此，本章利用3D-CNN从高光谱图像长、宽和高光谱三个维度进行卷积提取信息，可以较好地利用光谱信息。

在前向传播阶段，本节提出的基于3D卷积神经网络主要包含空-谱信息提取阶段、卷积池化阶段、全连接学习阶段、最终分类阶段四个部分，网络结构如图4.9所示。

1. 空-谱信息提取阶段：利用3D卷积核对数据进行三维卷积操作，并采用ReLU对卷积后的结果进行处理。
2. 卷积池化阶段：利用3D或者1D卷积核，将卷积步长调整为合适的长度，从而实现对数据的池化运算，这使得神经网络在通过卷积提取数据的同时进行了池化操作，可以降低网络的过拟合。
3. 全连接阶段将：将提取到的三维信息展开为一维，将之前阶段采集的信息进行压缩处理。
4. 最终分类阶段：利用Softmax函数，将网络最后的输出数据映射为几个不同大小分数，根据分数大小从而判断输入数据的归属类。



图4.9 本章3D-CNN高光谱图像分类框架

同之前用于高光谱图像分类的3D-CNN相比，本章设计的网络的结构中没有使用Dropout函数并通过在各个维度设置合适步长的卷积层起到池化层的作用，这样使得网络能保留更多的原始数据，具备更好的分类能力。此外，通过设计较少的卷积池化层的滤波器数量，可以构建更深的神经网络，从而能学习更“深”层次的特征。

在网络反向传播阶段，通过梯度下降法求解使得损失函数最小的网络权值。其中多类交叉熵代价函数(Categorical Crossentrop)作为Softmax函数的损失函数，如式(4-5)所示：

 (4-6)

其中n为输出类别数目，y为期望输出，a为神经元实际输出。选择Nesterov- Momentum

SGD优化算法[插入]作为神经网络的优化器，求解网络最优权值，其表达式如式(4-6)所示：

 (4-7)

其中为网络权值，为学习速率，为损失函数针对的梯度，C为损失函数。采用添加Nesterov动量的SGD算法，可以更新梯度时顺应损失函数的梯度来调整速度，并对SGD进行加速，从而较快到达收敛。在更新权值参数时，为了克服批梯度下降(Batch Gradient Descient, BGD)计算开销大速度慢和SGD收敛的震荡问题，将采用小批量梯度下降(Mini-Batch Gradient Descient, MBGD)代替Nesterov Momentum SGD 算法中的 SGD。批量梯度下降指的是，把数据分为几个批，按照批更新参数。那么一个批中的数据就会共同决定本次梯度的方向，下降方向不容易跑偏，并且减少了随机性。另一方面，由于一个批中得数据比整个数据集小了很多，因此计算消耗也会降低。

## 4.5 仿真结果与分析

本章实验基于谷歌Colabotary平台，选择Kears搭建神经网络框架，利用Python编写实验代码。为了验证本章提出的3D-CNN高光谱分类框架有着强大的分类性能，选取分类难度较高的Indian Pines高光谱数据和KSC(肯尼迪航空中心)高光谱数据作为实验数。其中Indian Pines数据集包含16个地物，其地物参考图如图4.10(a)所示。KSC数据集包含13个地物类别，其地物参考图如图4.10(b)所示。本实验选择10%的样本作为训练集，10%的样本作为验证集，其余样本作为测试集。

本节设立两个分实验去验证本章算法的优劣。通过第一个实验探究网络结构和超参数对分类结果的影响，从而确定最优的分类方案；在第二个实验中，将本章节算法和传统高光谱分类算法和以及其余CNN结构进行对比，探究本章算法的优劣。

（a） （b）

图4.10分布图（a）Indian pines数据集（b）KSC数据集

**实验I：**

基于上述理论描述，本节构建出三种形式的网络结构，用于寻找最优的网络结果。三种网络分别是：4个卷积层和2个全连接层的A网络、包含6个卷积层和2全连接层的B网络结构、包含8个卷积层的2个全连接层C网络结构，如图4.10所示。对于C层网络，通过设置多个仅包含少数滤波器的池化层，将网络进行序列化拉伸，使得网络可以提取到更“深”层次的信息。在网络中，对于卷积层滤波器数量C和卷积池化层滤波器的数量P1分别利用网格搜索法进行确定，先在大范围内粗略确定，最后在精细范围内确定，其取值范围均为[5-40]。对于C网络中个卷积池化层的个数P2，其取值范围为[2-5]。实验中，用mini-batch的方式训练数据，每次batch大小为16，对于两组数据均进行100次迭代训练。

表4.1和4.2给出了A网络和B网络在两组高光谱数据上的分类效果，可以看出当空域窗口选为，A模型对两组数据的分类效果最好。表4.3给出了C层网络在两组高光谱数据上的分类表现，通过对比可以发现当空域窗口为时，B模型效果好，当空域窗口为时，C模型略好。总体而言，模型A分类效果最好，这是因为深度学习需要大量的训练样本，本实验只选择总体样本的10%作为训练样本，用小数量的样本训练复杂的网络，极有可能产生欠拟合。



(a) A网络



(b) B网络



(c) C网络

图4.10 本节实验三种网络结构

图4.11给出了模型A在选择的空域窗口时的分类效果图，Indiana Pines数据集的OA和Kappa分别达到了94.8%和0.932，KSC数据集的OA和Kappa分别达到了96.7%和0.964。图4.12和4.13分别是利用两数据集训练网络时训练集精度、训练集损失、验证集精度和验证集精度随训练次数的变化曲线，可以出处随着训练次数不断增加，验证集精度逐渐和测试集精度逐渐接靠近，验证集损失不断减小。最终，选择验证集精度最好的模型最为分类模型。

表4.1 A网络模型针对两组数据的分类表现

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 空域窗口 | 卷积层数量 | 卷积池化层数量 | 全连接层数量 | IndianPines  OA Kappa 权值数 | KSC  OA Kappa 权值数 |
| 2 2 2 91.7% 0.905 107817 94.7% 0.941 89168 | | | | | |
| 2 2 2 94.8% 0.941 108297 96.7% 0.964 89648 | | | | | |

表4.2 B网络模型针对两组数据的分类表现

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 空域窗口 | 卷积层数量 | 卷积池化层数量 | 全连接层数量 | IndianPines  OA Kappa 权值数 | KSC  OA Kappa 权值数 |
| 3 3 2 91.3% 0.909 95411 94.03% 0.933 84608 | | | | | |
| 3 3 2 93.7% 0.923 95891 95.51% 0.950 85088 | | | | | |

表4.3 C网络模型针对两组数据的分类表现

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 空域窗口 | 卷积层数量 | 卷积池化层数量 | 全连接层数量 | IndianPines  OA Kappa 权值数 | KSC  OA Kappa 权值数 |
| 4 4 2 88.3% 0.86.6 322477 93.7% 0.931 85387 | | | | | |
| 4 4 2 93.4% 0.925 322957 96.6% 0.962 85867 | | | | | |

****

（a）OA = 94.1% Kappa =0.932 （b）OA=96.7% Kappa=0.964

图4.11 分类效果图（a）Indian pines数据集（b）KSC数据集

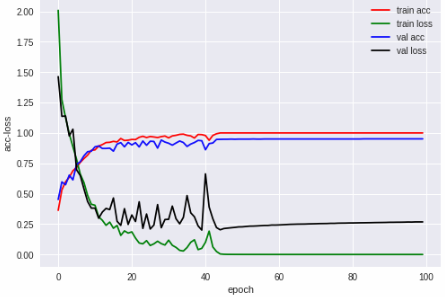


图4.12 Indian Pines ACC-LOSS曲线

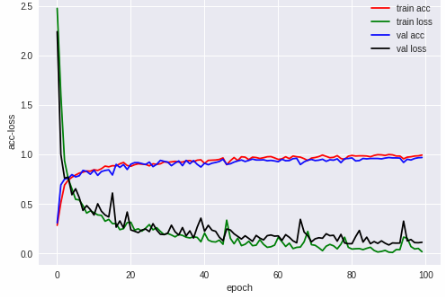
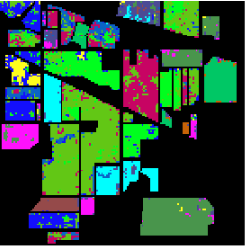


图4.13 KSC ACC-LOSS曲线

**实验II：**

本实验用于将本章提出方法和传统高光谱图像分类算法进行比较，证明本章提出的分类算法分类效果较好。根据上一个实验，我们发现当选择空域窗口大小为时A型网络的分类效果最好。因此，本节选择该模型的分类效果作为比较基准，对比分类算法分别为KC-LGM、DP-CNN[weili]、RPCNN[laowai]。对于KC-LGM，IndianPines留所有波段的信息，KSC保留前30主成分波段，空域窗口为，K-Means取5个聚类中心。对于DP-CNN分别构建2,3,4层网络进行测试，选择精度最好的4层网络模型。对于RPCA-CNN选择作者文中4层网络模型，分别提取前20、30、40层主成分，实验发现提取前30主成分时效果最好。

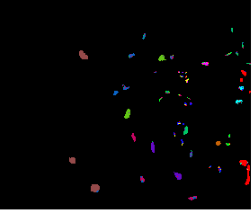
**** 图4.14和图4.15分别给出了四种分类算法在Indian pines和KSC分类效果图，表4.4和4.5分别给出了四种算法对两组实验数据各类地物的分类精度。可以明显的看出本章算法分类效果最好，RPCA-CNN效果分类最差，上章算法KC-LGM对于KSC数据集分类效果欠佳，对Indian pines分类效果较好。因此，利用3D-CNN对高光谱遥感图像分类可以解决传统分类算法的局限性。

 （a）本章算法OA = 94.1% Kappa =0.932 （b）KC-LGM OA = 92.7% Kappa =0.917

（c）DPCNN OA = 80.6% Kappa =0.779 （d）RPCNN OA = 77.2% Kappa =0.736

图4.14 Indian pines分类效果图

（a）本章算法OA=96.7% Kappa=0.964 （b）KC-LGM OA = 83.3% Kappa =0.814

****

（c）DPCNN OA=88.0% Kappa=0.866 （d）RPCNN OA = 65.1% Kappa =0.61

图4.15 KSC分类效果图

表 4.4 Indian Pines数据集地物分类OA

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 编号 | 地物类别 | 样本个数 | 本章算法 | KC-LGM | DPCNN | RPCNN |
| 1 | Alfalfa | 54 | 0.35 | 0.85 | 0.43 | 0.00 |
| 2 | Corn-notill | 1434 | 0.96 | 0.95 | 0.81 | 0.73 |
| 3 | Corn-min | 834 | 0.93 | 0.95 | 0.69 | 0.80 |
| 4 | Corn | 234 | 0.88 | 0.91 | 0.68 | 0.63 |
| 5 | Grass/Pasture | 497 | 0.96 | 0.94 | 0.88 | 0.82 |
| 6 | Grass/Trees | 747 | 0.99 | 0.95 | 0.96 | 0.91 |
| 7 | Grass/pasture-mowed | 26 | 0.88 | 0.79 | 0.46 | 0.96 |
| 8 | Hay-windrowed | 489 | 0.97 | 1.00 | 0.97 | 0.94 |
| 9 | Oats | 20 | 0.58 | 0.40 | 0.75 | 0.05 |
| 10 | Soybeans-notill | 968 | 0.94 | 0.92 | 0.73 | 0.76 |
| 11 | Soybeans-min | 2468 | 0.95 | 0.92 | 0.83 | 0.84 |
| 12 | Soybean-clean | 614 | 0.94 | 0.70 | 0.77 | 0.05 |
| 13 | Wheat | 212 | 0.99 | 0.99 | 0.95 | 0.94 |
| 14 | Woods | 1294 | 0.98 | 0.98 | 0.95 | 0.94 |
| 15 | Bldg-Grass-Tree-Drive | 380 | 0.90 | 0.96 | 0.61 | 0.67 |
| 16 | Stone-steeltowers | 95 | 0.97 | 0.99 | 0.90 | 0.00 |
|  |  |  |  |  |  |  |

表4.5 KSC数据集地物分类OA

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 编号 | 样本个数 | 本章算法 | KC-LGM | DPCNN | RPCNN |
| 1 | 761 | 0.99 | 0.96 | 0.96 | 0.94 |
| 2 | 243 | 0.94 | 0.61 | 0.88 | 0.53 |
| 3 | 256 | 0.89 | 0.75 | 0.94 | 0.49 |
| 4 | 252 | 0.96 | 0.33 | 0.05 | 0.03 |
| 5 | 161 | 0.86 | 0.61 | 0.71 | 0.55 |
| 6 | 229 | 0.90 | 0.70 | 0.49 | 0.00 |
| 7 | 105 | 0.99 | 0.98 | 0.92 | 0.30 |
| 8 | 431 | 0.99 | 0.80 | 0.88 | 0.50 |
| 9 | 520 | 0.96 | 0.90 | 0.99 | 0.83 |
| 10 | 404 | 0.97 | 0.82 | 0.94 | 0.50 |
| 11 | 419 | 0.98 | 0.98 | 0.98 | 0.91 |
| 12 | 503 | 1.00 | 0.73 | 0.96 | 0.57 |
| 13 | 927 | 1.00 | 0.98 | 1.00 | 0.88 |
|  |  |  |  |  |  |

## 4.6 本章小节

传统空-谱信息提取方式只能提取“浅”层的手工特征，这种特征提取方式严重地依赖于传领域内的知识，且针对不同的高光谱数据手工特征的提取方式可能有所不同。基于此，本文提出一种基于3D-CNN框架的高光谱图像分类算法。该算法相比于传统空-谱提取算法，可以自动生成数据自身“深层”的特征；相比于采用2D-CNN的高光谱图像算法，从对空间和光谱两个维度同时进行卷积，更好的利用了高光谱数据的光谱信息。通过两组实验进行了验证，首先通过第一组实验比较多种网络结构下本章算法的执行效果，验证了3D-CNN算法可以极大地提高高光谱图像分类精度；通过第二组实验将本章提出算法与传统分类算法进行比较，验证了本章算法有着更好的分类能力。然而，本章算法没有考虑利用相关无监督信息，可能会损失一些分类精度，因此本课题后续将着重探究3D-CNN和生成对抗网络(GAN)相结合形成的高光谱图像分类算法。

# 结 论

随着遥感科学的不断进步，高光谱遥感技术逐渐被应用到众多领域，研究高光谱影像的特征提取和分类技术有着重要的实际价值。本文针对高光谱数据的特点以及传统分类方法的不足，深入研究了高光谱数据的分类问题，总结概括如下：

(1) 对课题的研究背景进行了研究，了解高光谱分类的技术的发展历史。充分认识了高光谱遥感技术的成像特点以及高光谱数据空间高分辨率、波段数多和信息量大所带来的信息处理上的难题。

(2) 研究了高光谱影像分类算法理论基础。首先概述了高光谱影像的分类流程，分别介绍了有监督、无监督、半监督算法在高光谱影像分类中的应用。其次，阐述了高光谱数据分类结果的评价指标，并详细介绍了本课题中用到的高光谱数据集。最后，利用传统RBF-SVM和线性SVM对Indian Pines进行了分类实验。

(3) 提出一种基于树形集成学习的高光谱影像空-谱分类算法。首先简要概述了集成学习的基本理论，分别介绍了Boosting、Boosting Tree、GBDT、Bagging、随机森林等算法。其次，详细介绍了LightGBM算法的核心思想，并提出一种结合K-means聚类算法的空-谱特征提取方式，将两者结合形成一种新式高光谱影像分类算法——KC-LGM算法。最后，利用该算法对两组数据高光谱数据进行试验，验证了KC-LGM对于高光谱影像有着极好的分类效果。

(4) 提出一种基于3D卷积神经网络的高光谱影像分类算法。首先介绍了深度学习的基本概念、原理和实验框架。其次，介绍了卷积神经网络的原理，并详细阐述了1D-CNN、2D-CNN和3D-CNN的概念。考虑到传统空-谱特征提取方式依赖于领域内已有知识，需要人工合成特征、所得特征处于较“浅”层次等缺点，提出一种新型基于3D-CNN的高光谱影像分类架构。该分类框架有以下几个优点：通过设立合理的步长使得卷积层在提取信息的同时起到了池化层避免过拟合的作用；构建3D卷积滤波器，同时从三个方向对高光谱立方进行信息采集，更好地联立了空-谱信息；系统自发地提取了高光谱数据的“深”层特征。最后，通过两个实验验证了本章算法相对于传统分类算法、1D-CNN和2D-CNN有着更好地分类效果。

本文提出的两种算法均有着极好的分类表现，然而KC-LGM算法在进行分类过程中，需要占用大量的内存，运算开销极大，因此只能通过PCA减少原始数据量，这样必然会丢失一部分空间信息。因此针对KC-LGM算法，后序将着重研究计算机并行处理技术和高光谱数据的存储方式，以便优化算法，从而在充分利用光谱信息的同时加快运算速度。

对于3D-CNN算法，本文没有利用深度学习中的非监督方法进行辅助，没有更好地利用到数据信息。因此，后序将着重研究生成对抗网络(GAN)在高光谱数据中的应用，并考虑将GAN结构和3D-CNN结合，以便更好地利用高光谱数据中蕴藏的信息，提高分类精度。

# 参考文献

孙家抦. 遥感原理与应用[M]. 武汉: 武汉大学出版社, 2003:92.

钟燕飞. 遥感影像处理中的人工免疫系统理论及其应用方法研究[D]. 武汉大学博士学位论文, 2007.

常祎. 兼顾效率和精度的高光谱遥感影像分类技术研究[D]. 哈尔滨工程大学硕士学位论文, 2017.

郝思媛. 空谱协作的高光谱图像分类方法研究[D]. 哈尔滨工程大学博士学位论文, 2015.

David C. Plaut, Geoffrey E. Hinton. Learning sets of filters using back-propagation[J]. Computer Speech & Language, 1987, 2(1):35-61.

G. Hinton, S. Osindero, and Y. Teh, “A fast learning algorithm for deep belief nets,” Neural Computation, vol. 18, no. 7, pp. 1527–1554, 2006.

Krizhevsky A, Sutskever I,Hinton G E. Imagenet classification with deep convolutional neuralnetworks[C]. Advances in neural information processing systems. 2012:1097-1105.

He K, Zhang X, Ren S, etal. Deep residual learning for image recognition[J]. arXiv preprintarXiv:1512.03385, 2015.

李航. 统计学习方法[M]. 清华大学出版社, 2012.

Valiant L G. A theory of the learnable[J]. Communications of Acm, 1984, 27(11):1134-1142.

Schapire R E. A Brief Introduction to Boosting[C]. Sixteenth International Joint Conference on Artificial Intelligence. Morgan Kaufmann Publishers Inc. 1999:1401-1406.

Freund, Yoav, Schapire, Robert E. A decision-theoretic generalization of on-line learning and an application to boosting[C]. European Conference on Computational Learning Theory. Springer, Berlin, Heidelberg, 1995:23-37.

Breiman L. Bagging predictors[J]. Machine Learning, 1996, 24(2):123-140.

Breiman L. Random Forests[J]. Machine Learning, 2001, 45(1):5-32.

Chen T, Guestrin C. XGBoost:A Scalable Tree Boosting System[C]. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016:785-794.

Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree[C].Advances in Neural Information Processing Systems. 2017: 3146-3154.

Mishra D R. A Comparative Evaluation of ISODATA and Spectral Angle Mapping for the Detection of Saltcedar Using Airborne Hyperspectral Imagery[J]. Geocarto International, 2006, 21(2):59-66.

Subhash H M, Leahy M J. Microcirculation imaging based on full-range high-speed spectral domain correlation mapping optical coherence tomography[J]. Journal of Biomedical Optics, 2014, 19(2):21103.

Zhang B, Li S, Jia X, et al. Adaptive Markov Random Field Approach for Classification of Hyperspectral Imagery[J]. IEEE Geoscience & Remote Sensing Letters, 2011, 8(5):973-977.

Benediktsson J A, Pesaresi M, Amason K. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations[J]. Geoscience & Remote Sensing IEEE Transactions on, 2003, 41(9):1940-1949.

Chen Y, Nasrabadi N M, Tran T D. Hyperspectral Image Classification Using Dictionary-Based Sparse Representation[J]. IEEE Transactions on Geoscience & Remote Sensing, 2011, 49(10):3973-3985.

Tarabalka Y, Chanussot J, Benediktsson J A. Segmentation and classification of hyperspectral images using watershed transformation[J]. Pattern Recognition, 2010, 43(7):2367-2379.

Demir B, Erturk S. Hyperspectral Image Classification Using Relevance Vector Machines[J]. IEEE Geoscience & Remote Sensing Letters, 2007, 4(4):586-590.

Roscher R, Waske B, Forstner W. Incremental Import Vector Machines for Classifying Hyperspectral Data[J]. IEEE Transactions on Geoscience & Remote Sensing, 2012, 50(9):3463-3473.

Du P, Tan K, Xing X. A novel binary tree support vector machine for hyperspectral remote sensing image classification[J]. Optics Communications, 2012, 285(13-14):3054-3060.

Zhang H, Shi W, Liu K. Fuzzy-Topology-Integrated Support Vector Machine for Remotely Sensed Image Classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2012, 50(3):850-862.

Li C H, Kuo B C, Lin C T, et al. A Spatial–Contextual Support Vector Machine for Remotely Sensed Image Classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2012, 50(3):784-799.

Gao L, Li J, Khodadadzadeh M, et al. Subspace-Based Support Vector Machines for Hyperspectral Image Classification[J]. IEEE Geoscience & Remote Sensing Letters, 2015, 12(2):349-353.

Zhang C, Wang T, Atkinson P M, et al. A novel multi-parameter support vector machine for image classification[J]. International Journal of Remote Sensing, 2015, 36(7):1890-1906.

Xia J, Chanussot J, Du P, et al. Rotation-based support vector machine ensemble in classification of hyperspectral data with limited training samples[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(3): 1519-1531.

Li Y, Zhang H, Shen Q. Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network[J]. Remote Sensing, 2017, 9(1): 67.

Li W, Wu G, Zhang F, et al. Hyperspectral image classification using deep pixel-pair features[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(2): 844-853.

Hu W, Huang Y, Wei L, et al. Deep convolutional neural networks for hyperspectral image classification[J]. Journal of Sensors, 2015, 2015.

Zhang M, Li W, Du Q. Diverse Region-Based CNN for Hyperspectral Image Classification[J]. IEEE Transactions on Image Processing, 2018, 27(6): 2623-2634.

Lee H, Kwon H. Going deeper with contextual CNN for hyperspectral image classification[J]. IEEE Transactions on Image Processing, 2017, 26(10): 4843-4855.

Makantasis K, Karantzalos K, Doulamis A, et al. Deep supervised learning for hyperspectral data classification through convolutional neural networks[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, 2015: 4959-4962.

马晓瑞. 基于深度学习的高光谱影像分类方法研究[D]. 大连理工大学博士学位论文, 2017.

王立国, 赵春晖. 高光谱图像处理技术[M]. 国防工业出版社, 2013.

黄立贤, 沈志学. 高光谱遥感图像的监督分类[J]. 地理空间信息, 2011(5):81-83.

Uss M L, Vozel B, Chehdi K, et al. Maximum likelihood estimation of spatially correlated signal-dependent noise in hyperspectral images[J]. Optical Engineering, 2012, 51(11): 111712.

Yang H, Du Q, Chen G. Unsupervised hyperspectral band selection using graphics processing units[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2011, 4(3): 660-668.

Zhong Y, Zhang L, Huang B, et al. An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(2): 420-431.

Wagstaff K, Cardie C, Rogers S, et al. Constrained k-means clustering with background knowledge[C]//ICML. 2001, 1: 577-584.

Kanungo T, Mount D M, Netanyahu N S, et al. An efficient k-means clustering algorithm: Analysis and implementation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002 (7): 881-892.

Camps-Valls G, Marsheva T V B, Zhou D. Semi-supervised graph-based hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2007, 45(10): 3044-3054.

Chapelle O, Scholkopf B, Zien A. Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews][J]. IEEE Transactions on Neural Networks, 2009, 20(3): 542-542.

Böhning D. Multinomial logistic regression algorithm[J]. Annals of the institute of Statistical Mathematics, 1992, 44(1): 197-200.

Kwak C, Clayton-Matthews A. Multinomial logistic regression[J]. Nursing research, 2002, 51(6): 404-410.

童庆禧, 张兵, 郑兰芬. 高光谱遥感: 原理, 技术与应用[M]. 高等教育出版社, 2006.

Congalton R G. A review of assessing the accuracy of classifications of remotely sensed data[J]. Remote sensing of environment, 1991, 37(1): 35-46.

Stehman S. Estimating the kappa coefficient and its variance under stratified random sampling[J]. Photogrammetric Engineering and Remote Sensing, 1996, 62(4): 401-407.

Green R O, Eastwood M L, Sarture C M, et al. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)[J]. Remote sensing of environment, 1998, 65(3): 227-248.

Landgrebe D, Biehl K. AVIRIS NW Indiana’s Indian Pines 1992 data set[J]. 1992.

Tarabalka Y, Fauvel M, Chanussot J, et al. SVM and MRF-based method for accurate classification of hyperspectral images[J]. IEEE Geoscience and Remote Sensing Letters, 2010, 7(4): 736-740.

Plaza A, Martinez P, Plaza J, et al. Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations[J]. IEEE Transactions on Geoscience and remote sensing, 2005, 43(3): 466-479.

Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural processing letters, 1999, 9(3): 293-300.

周志华. 机器学习[M]. Qing hua da xue chu ban she, 2016.

Banfield R E, Hall L O, Bowyer K W, et al. A comparison of decision tree ensemble creation techniques[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007 (1): 173-180.

Freund Y, Schapire R E. Experiments with a new boosting algorithm[C]//Icml. 1996, 96: 148-156.

Hastie T, Rosset S, Zhu J, et al. Multi-class adaboost[J]. Statistics and its Interface, 2009, 2(3): 349-360.

Carreras X, Marquez L. Boosting trees for anti-spam email filtering[J]. arXiv preprint cs/0109015, 2001.

Friedman J H. Stochastic gradient boosting[J]. Computational Statistics & Data Analysis, 2002, 38(4): 367-378.

Quinlan J R. Bagging, boosting, and C4. 5[C]//AAAI/IAAI, Vol. 1. 1996: 725-730.

Liaw A, Wiener M. Classification and regression by randomForest[J]. R news, 2002, 2(3): 18-22.

Bingham E, Mannila H. Random projection in dimensionality reduction: applications to image and text data[C]//Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2001: 245-250.

Carneiro T, Da Nóbrega R V M, Nepomuceno T, et al. Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications[J]. IEEE Access, 2018, 6: 61677-61685.

Plaza A, Martinez P, Perez R, et al. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles[J]. Pattern Recognition, 2004, 37(6): 1097-1116.

Benediktsson J A, Palmason J A, Sveinsson J R. Classification of hyperspectral data from urban areas based on extended morphological profiles[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(3): 480-491.

Ghamisi P, Dalla Mura M, Benediktsson J A. A survey on spectral–spatial classification techniques based on attribute profiles[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(5): 2335-2353.

Tuia D, Volpi M, Dalla Mura M, et al. Automatic feature learning for spatio-spectral image classification with sparse SVM[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(10): 6062-6074.

Dalla Mura M, Villa A, Benediktsson J A, et al. Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(3): 542-546.

Jia S, Zhang X, Li Q. Spectral–Spatial Hyperspectral Image Classification Using $\ell\_ {1/2} $ Regularized Low-Rank Representation and Sparse Representation-Based Graph Cuts[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(6): 2473-2484.

Zhang X, Xu C, Li M, et al. Sparse and low-rank coupling image segmentation model via nonconvex regularization[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2015, 29(02): 1555004.

Zhang B, Li S, Jia X, et al. Adaptive Markov random field approach for classification of hyperspectral imagery[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(5): 973-977.

Tarabalka Y, Rana A. Graph-cut-based model for spectral-spatial classification of hyperspectral images[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International. IEEE, 2014: 3418-3421.

Qian Y, Ye M, Zhou J. Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(4): 2276-2291.

Shen L, Jia S. Three-dimensional Gabor wavelets for pixel-based hyperspectral imagery classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 49(12): 5039-5046.

Zhang L, Zhang L, Du B. Deep learning for remote sensing data: A technical tutorial on the state of the art[J]. IEEE Geoscience and Remote Sensing Magazine, 2016, 4(2): 22-40.

Chen Y, Zhao X, Jia X. 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.

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.

Yue J, Zhao W, Mao S, et al. Spectral–spatial classification of hyperspectral images using deep convolutional neural networks[J]. Remote Sensing Letters, 2015, 6(6): 468-477.

Makantasis K, Karantzalos K, Doulamis A, et al. Deep supervised learning for hyperspectral data classification through convolutional neural networks[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International. IEEE, 2015: 4959-4962.

Liang H, Li Q. Hyperspectral imagery classification using sparse representations of convolutional neural network features[J]. Remote Sensing, 2016, 8(2): 99.

Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. Psychological review, 1958, 65(6): 386.

Hagan M T, Menhaj M B. Training feedforward networks with the Marquardt algorithm[J]. IEEE transactions on Neural Networks, 1994, 5(6): 989-993.

He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436.

Bastien F, Lamblin P, Pascanu R, et al. Theano: new features and speed improvements[J]. arXiv preprint arXiv:1211.5590, 2012.

Yang C T, Huang C L, Lin C F. Hybrid CUDA, OpenMP, and MPI parallel programming on multicore GPU clusters[J]. Computer Physics Communications, 2011, 182(1): 266-269.

Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014: 675-678.

Jia Y. Caffe2[J]. 2017.

Abadi M, Barham P, Chen J, et al. Tensorflow: a system for large-scale machine learning[C]//OSDI. 2016, 16: 265-283.

Paszke A, Gross S, Chintala S, et al. PyTorch[J]. 2017.

Seide F, Agarwal A. CNTK: Microsoft's open-source deep-learning toolkit[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 2135-2135.

Chollet F. Keras[J]. 2015.

LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series[J]. The handbook of brain theory and neural networks, 1995, 3361(10): 1995.

Goodfellow I, Bengio Y, Courville A, et al. Deep learning[M]. Cambridge: MIT press, 2016.

Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.

# 攻读硕士学位期间发表的论文和取得的科研成果

王立国, 马赫男. 基于改进K\_Medoids算法的高光谱图像聚类[J]. 哈尔滨工程大学学报, 2018, 39(12).(已录用)

# 致 谢

转眼间学生生涯已接近尾声，在论文撰写阶段十分感谢老师、实验室同仁以及室友的大力支持，临表涕零，不知所言。

谢谢大家！