

学号 2014301200176

密级 _____

武汉大学本科毕业论文

基于深度学习的 遥感影像分类方法

院（系）名称：电子信息学院

专业名称：电子信息工程

学生姓名：张凌霄

指导教师：卢其楷 讲师

二〇一八年五月

郑重声明

本人呈交的学位论文，是在导师的指导下，独立进行研究工作所取得的成果，所有数据、图片资料真实可靠。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确的方式标明。本学位论文的知识产权归属于培养单位。

本人签名：_____ 日期：_____

摘 要

遥感技术已成为人类研究地球环境和资源的主要工具。作为遥感技术中的重要信息表现形式，遥感影像可以记录各种地物电磁波的大小，反映地物的光谱、空间、纹理等特征。面对海量遥感影像数据，快速识别和获取不同专题信息通常直接取决于遥感影像的内容解译。而作为遥感影像解译中的重要环节，遥感影像分类结果的准确性直接影响了图像解译质量的高低。本文基于深度学习方法对遥感影像分类问题进行研究，研究的具体内容包含以下三部分：

(1) 面向遥感影像场景分类任务，本文介绍了多种深度神经网络的分类方法，并在 UMCL 数据集上进行实验并比较了各种网络的优劣。实验结果表明，使用 GoogleNet 网络所获得的分类准确率最高。随后，采用基于 ImageNet 数据集的微训练模型可将分类准确率提高 9.05%。

(2) 面向高光谱遥感影像分类任务，本文介绍了一种光谱-信息融合分类方法，并在 Pavia 数据集和 Indian Pines 数据集上验证了该方法的可行性。实验结果表明，该方法可以很好地保持影像地物对象的形状和细节信息，同时消除了错分噪点的影响，能够实现良好的分类结果。

(3) 面向大尺度高空间分辨率遥感影像分类任务，本文介绍了基于影像瓦片的场景分类方法，并利用 GF2 卫星高空间分辨率遥感影像验证了该方法的可行性。同时比较了该方法与基于人工判别的影像解译方法的优劣。

关键词：深度神经网络；场景分类；高光谱遥感影像；大尺度高空间分辨率影像

ABSTRACT

Remote sensing technology is a powerful technical way and an important source of information to study the environment and the geology of the Earth. Remote sensing images can objectively and truly record and reflect the intensity of electromagnetic radiation of surface objects, which are significant manifestations for information obtained by remote sensing detection. And remote sensing image classification plays an significant role in remote sensing image processing.

This paper is carried out on the basis of deep learning methods for remote sensing image classification. The specific research content can be divided into the following three parts:

- (1) For aerial scene classification, this paper introduces the classification methods based various deep neural networks, and compares the advantages and disadvantages of various networks on UMCL dataset. The experimental results illustrate that the classification accuracy based on GoogleNet is the highest. In order to improve the accuracy, we also use Fine-Tuning to test on UMCL dataset.
- (2) For hyperspectral remote sensing image classification, this paper introduces joint spectral-spatial classification and validates the feasibility of the method on Pavia dataset and Indian Pines dataset. The experimental results show that the method retains the shape and detail of some objects, while simultaneously eliminating noisy scattered points of misclassification.
- (3) For large-scale and high-spatial resolution remote sensing image classification, this paper introduces two classification methods: Scene Classification based on Image Tiles and Geographic Object-Based Image Analysis (GEOBIA) based on Multiresolution Segmentation algorithm. By testing on a large-scale and high-spatial resolution remote sensing image over Guangzhou obtained by GF2 satellite, PMS1 sensor on January 23, 2015, the experimental results demonstrate that the former has the advantage in classification speed while the latter has a higher classification accuracy.

Keywords: deep neural network; scene classification; hyperspectral remote sensing image; large-scale and high-spatial resolution image

目 录

1 緒論	1
1.1 研究背景与意义	1
1.2 国内外研究现状	2
1.3 论文研究内容	3
1.4 论文组织结构	4
2 深度学习概述	5
2.1 神经网络算法	5
2.2 深度神经网络模型	5
2.2.1 LeNet	5
2.2.2 AlexNet	6
2.2.3 VGGNet	7
2.2.4 GoogleNet	8
2.2.5 ResNet	10
2.2.6 DenseNet	11
2.2.7 CapsNet	11
2.3 本章小结	13
3 遥感影像场景分类	14
3.1 基于神经网络的遥感影像场景分类	14
3.1.1 深度学习方法	14
3.1.2 基于 HOG 特征的支持向量机	14
3.1.3 视觉词袋模型	15
3.1.4 深度信念网络	17
3.2 基于微训练模型的遥感影像场景分类	18
3.2.1 模型预训练	18
3.2.2 特征融合框架	19
3.3 实验与分析	20
3.3.1 实验数据	20
3.3.2 实验配置	21
3.3.3 精度评价指标	21

3.3.4	实验结果与分析	22
3.4	本章小结	27
4	高光谱遥感影像分类	28
4.1	模型整体架构	28
4.2	光谱信息	29
4.2.1	光谱向量提取	29
4.2.2	R-PCA 降维	29
4.3	空间信息	29
4.3.1	邻域空间信息提取	29
4.3.2	卷积神经网络提取特征	29
4.4	空间-光谱信息融合分类	30
4.5	实验与分析	30
4.5.1	实验数据	30
4.5.2	实验模型	33
4.5.3	实验结果与分析	33
4.6	本章小结	36
5	大尺度高空间分辨率遥感影像分类	37
5.1	基于影像瓦片的场景分类	37
5.2	基于人工判别的影像解译	37
5.2.1	遥感影像面向对象分析技术	37
5.2.2	遥感影像分割与多尺度分割算法	38
5.3	实验与分析	39
5.3.1	实验数据	39
5.3.2	实验结果与分析	40
5.4	本章小结	43
6	总结与展望	44
	参考文献	45
	致谢	50

1 绪论

1.1 研究背景与意义

随着对地观测技术的发展，各种资源环境监测卫星的发射与运行为地表动态监测提供了多平台、多光谱、多时相、大范围的信息^[1]。遥感技术已成为人类研究地球环境和资源的主要工具，并在政府部门、科研单位和公司得到广泛的应用。

作为遥感技术中的重要信息表现形式，遥感影像可以记录各种地物电磁波的大小，反映地物的光谱、空间、纹理等特征。作为遥感影像解译的主要组成部分，遥感影像分类是根据各类地物的光谱信息和空间信息，按照像元相似的大小，归类相似的像元，分离不相似的像元，并给每一个像元赋类别值，确定遥感图像与实际地物类别对应关系的过程^[2]。从单波段遥感影像到多光谱影像，再到高光谱影像，遥感影像分类研究不断扩展和深入^{[3][4]}。遥感影像分类方式主要可以分为统计决策法模式识别和句法模式识别。统计决策法模式识别指的是在某些一定量度或观测基础上，依据对研究对象的统计分析，把待识别模式划分到各自的模式类中的过程。主要方法包括：费歇准则法、最小距离法、最小均方误差法、梯度法、等混合距离法、逐步聚类法等。句法模式识别则需要了解图像结构信息，从而对其进行分类，主要包括人工神经网络方法，模糊数学方法，决策树方法，专家系统方法等^[5]。

近年来，深度学习算法在图像处理和计算机视觉领域显现出优良的性能。深度学习算法具有数据利用率高，特征提取整合能力强，便于进行系统性的性能评估等优势^[6]，同时深层模型具有更强的建模和推广能力^[7]。利用深度学习算法解决遥感影像分类问题，有利于提升对于影像的光谱和空间信息的综合利用率，学习到对象更深刻本质的特征^[8]。与此同时，获得高鲁棒性、强判别能力的遥感影像特征，提高深度神经网络对影像特征的利用率也是遥感影像分类问题进一步研究的热点。

1.2 国内外研究现状

随着遥感技术的发展，众多卫星提供了丰富的对地观测数据。为了从遥感数据中提取到感兴趣的信息，影像分类技术得到了研究者的广泛关注。影像分类基于样本提供的类别先验信息，训练分类器对未标记样本的类别进行预测。其中，神经网络、决策树、光谱角分类、增强分类、多尺度分类（MFM-5-尺度）、马尔可夫随机场、线性约束判别分析、随机森林等众多分类方法在遥感影像分类中得到应用^[5]。随着遥感影像空间分辨率的不断提升，更多的地物空间细节被获取。传统的基于光谱的分类方式难以挖掘地物的空间分布关系，不利于高分辨率遥感影像的解译。因此，光谱-空间联合分类的方式得到了极大的发展。为了描述影像的空间信息，研究者提出了大量的特征提取算法，并将提取到的空间特征与光谱特征相叠加，形成更可靠的地物特征描述方式。另一方面，面向对象的方法也得到了重视，该方法使用分割算法将影像上具有相似特征且空间相邻的像素进行聚集形成对象块，并以此为基本处理单位进行影像解译。

近年来，基于光谱特征，采用形态学特征、小波变换、灰度纹理、指数特征等方法提取空间特征补充光谱信息，形成面向对象的分类方法，结合深度学习进行遥感影像分类成为当前的主要趋势之一。此方法可以克服仅仅采用光谱信息分类结果中的噪点，降低分类结果的不确定性；面向对象的分类方法可以有效处理景观异质性的问题，提升分类精度；利用深度学习算法相对于 SVM、KNN 等分类方法，能够更加准确地挖掘遥感影像的光谱-空间分布规律。未来，增加遥感影像特征，提高深度神经网络分类精度，增强分类器鲁棒性和容错性，提高网络与地学知识的融合性，提高算法运算速度，减少人工设置，使得分类器更自动化、智能化是遥感影像分类问题的发展趋势。国内众多大学和研究院，包括武汉大学、哈尔滨工业大学、电子科技大学、清华大学、北京交通大学、中科院遥感与数字地球研究所等，都在遥感影像分类问题方面做了大量的研究，极大地促进了遥感影像分类研究的发展。表 1.1 展示了近几年的相关研究工作。

表 1.1 深度学习算法在遥感影像分类中运用

时间	事件
2013 年	Chen 等人首次提出采用堆栈自编码器（SAE）的深度模型来提取高光谱图像的特征，并对这些特征采用逻辑回归的方式来分类 ^[9] 。
2014 年	美国 ORNL (Oak Ridge National Laboratory) 研究员 Cheriyadat 在处理高分辨率场景分类问题时首次使用了非监督特征学习算法，为后续的研究提供了新的研究思路 ^[11] 。
2015 年	巴西三星研究院 (SAMSUNG Research Institute) Penatti 在处理高分辨率遥感场景时首次使用了深度卷积神经网络，大大提高了场景分类的精度 ^[12] 。同年，Hu 等人在对高光谱遥感图像进行分类时采用了卷积神经网络模型 (CNN) ^[13]
2016 年	刘大伟等通过非下采样轮廓波变换计算影像的纹理特征，利用深度学习的常用模型—深度信念网络(DBN)对高分辨率遥感影像进行了基于光谱-纹理特征的分类 ^[3] 。
2017 年	袁林等利用自编码神经网络对高光谱数据进行光谱维的降维，然后将卷积神经网络作为分类器 ^[14] 。

1.3 论文研究内容

论文的研究内容包括：

- (1) 整理国内外关于深度学习的发展历程和最新的研究成果，对遥感影像分类的基本原理和算法进行简要介绍；研究卷积神经网络搭建和结构，描述几种常见的深度学习网络的模型；
- (2) 针对遥感影像分类的不同应用（影像场景分类、高光谱影像分类、大尺度高分辨率遥感影像分类），分别运用深度学习的方法实现遥感影像分类任务，并评估分类结果；
- (3) 运用 BOW、SVM、Geobia 等方法进行对比实验，评价各种方法结果。

图 1.1 显示了论文研究主要内容。

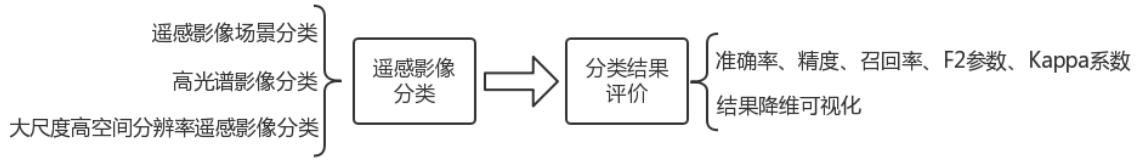


图 1.1 论文研究主要内容

1.4 论文组织结构

本论文共分为六章，各个章节的内容概述如下：

第一章介绍了论文选题的研究背景与意义，分析了该研究方向上国内外研究现状，并列出了本论文的主要研究内容和组织结构。

第二章主要介绍了神经网络算法，并具体介绍了深度学习的常用模型。

第三章面向遥感影像场景分类任务，给出了基于神经网络的分类方法和基于微训练模型的分类方法，并通过实验分析了各方法的优劣。

第四章面向高光谱遥感影像分类任务，介绍了一种光谱-空间信息融合分类方法，并通过实验验证了该方法的可行性。

第五章面向大尺度高空间分辨率遥感影像分类任务，介绍了基于影像瓦片的场景分类方法，同时比较了该方法与基于人工判别的影像解译方法的优劣。

第六章对本文工作进行了总结，并在此基础上进行了展望。

2 深度学习概述

2.1 神经网络算法

神经网络算法是基于仿生学，模仿人类神经结构而设计的。神经网络算法通过机器感知并解译数据，对输入信息进行标记或聚类。由于存在梯度消失的现象，深度神经网络一般比浅层网络更难训练，但是训练好的深度网络比浅层网络具有更强大的数据表征能力。在图像分类中，卷积神经网络于 2012 年运用到百万量级的 ImageNet 数据集合上，图像分类准确率大幅超过传统的方法，之后在视觉领域不断取得突破。图 2.1 显示了 2010-2017 年视觉领域竞赛 ImageNet Large Scale Visual Recognition Competition (ILSVRC) 图像分类任务中最优模型的 Top5 分类误差。

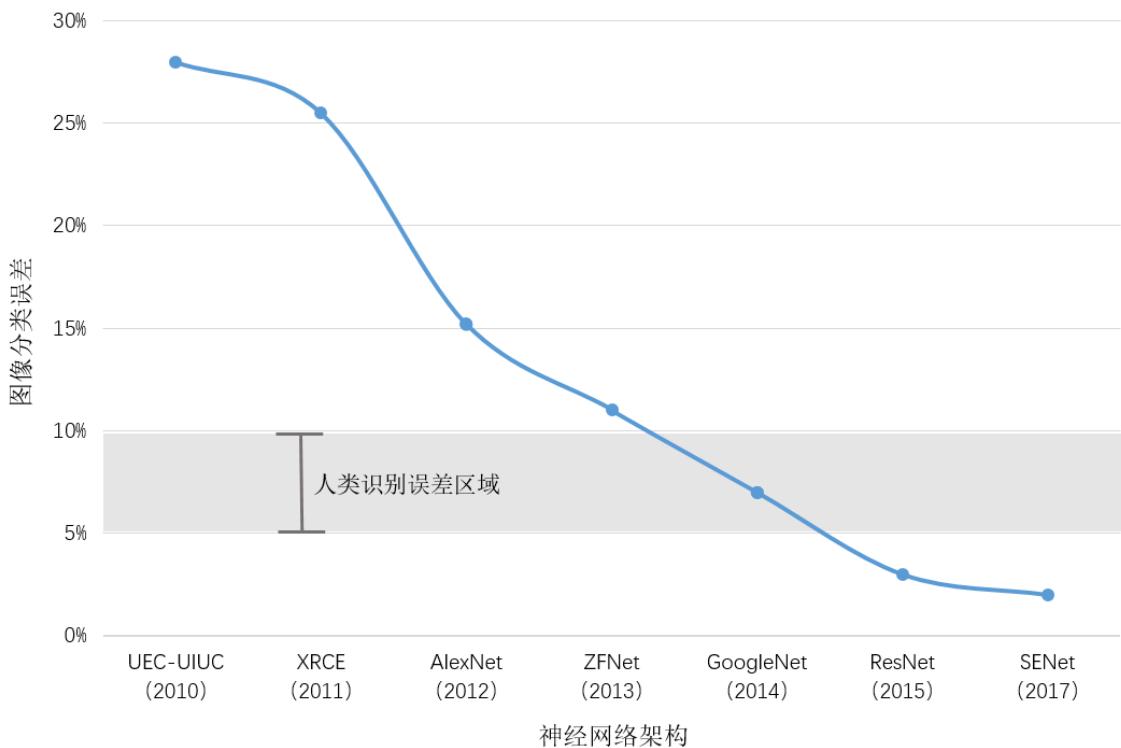


图 2.1 2010-2017 年 ILSVRC 图像分类任务最优模型的 Top5 分类误差

从上图可以看出，2015 年之后，深度学习分类器在 ILSVRC 图像分类任务中的分类误差已经低于人类识别的误差，即深度学习可以取得比人类更好的判别效果。

2.2 深度神经网络模型

2.2.1 LeNet

LeNet^{[15][16]}是最早的卷积神经网络结构，它开创了卷积神经网络的基本结构（如

图 2.2)。



图 2.2 卷积神经网络的基本结构

其中，卷积层（Convolutional Layers）对图像局部进行感知，并对图像中的每一小块进行分析，得出抽象程度更高的特征。通常情况下，通过卷积层处理的神经元结点矩阵的深度会增加。图 2.3 展示了卷积核或滤波器（filter）将当前层级上的一个子结点张量（ $32 \times 32 \times 3$ ）转化为下一层神经网络上的一个长和宽都为 1，深度为 5 的结点矩阵。这五个卷积核有不同的权重，但每一个卷积层使用一个卷积核的权重是一样的（共享权重）。

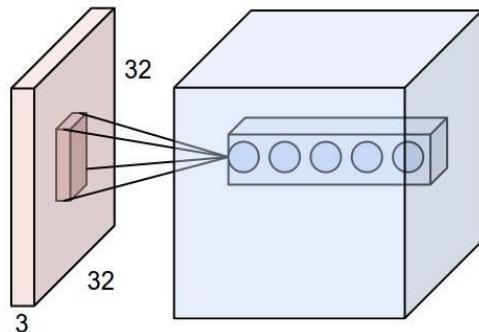


图 2.3 卷积核示意图

池化层（Pooling layers）通常接在卷积层后，它的主要作用是简化卷积层输出的信息^[17]。常见的池化有最大值池化(max-pooling)、最小值池化(min-pooling)等。

LeNet 是一个 7 层的网络，深度为 5，包含 2 个卷积层，2 个池化层和 2 个全连接层和 1 个高斯连接层，具体结构如图 2.4。

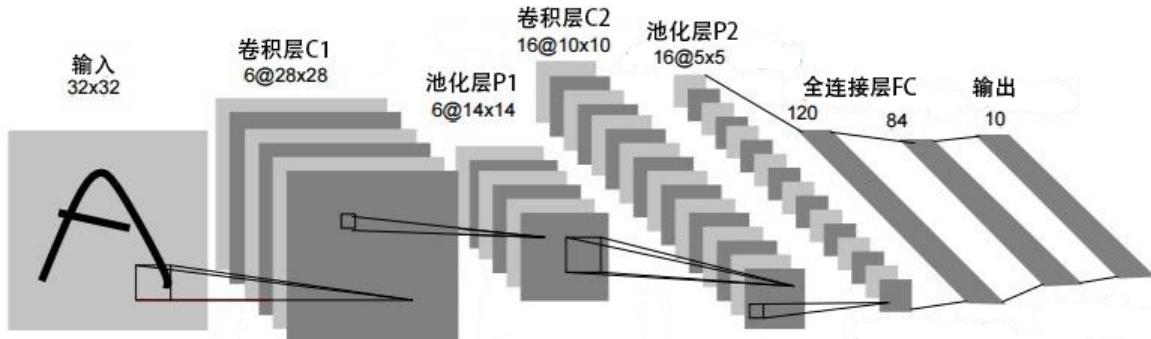


图 2.4 LeNet 结构示意图

2.2.2 AlexNet

由 Alex Krizhevsky 提出的 AlexNet^[18]是具有多个隐藏层的深度卷积神经网络，

具体包含 1 个输入层、5 个卷积层、3 个池化层和 3 个全连接层，1 个具有 1000 个输出类别标签的输出层，结构如图 2.5。

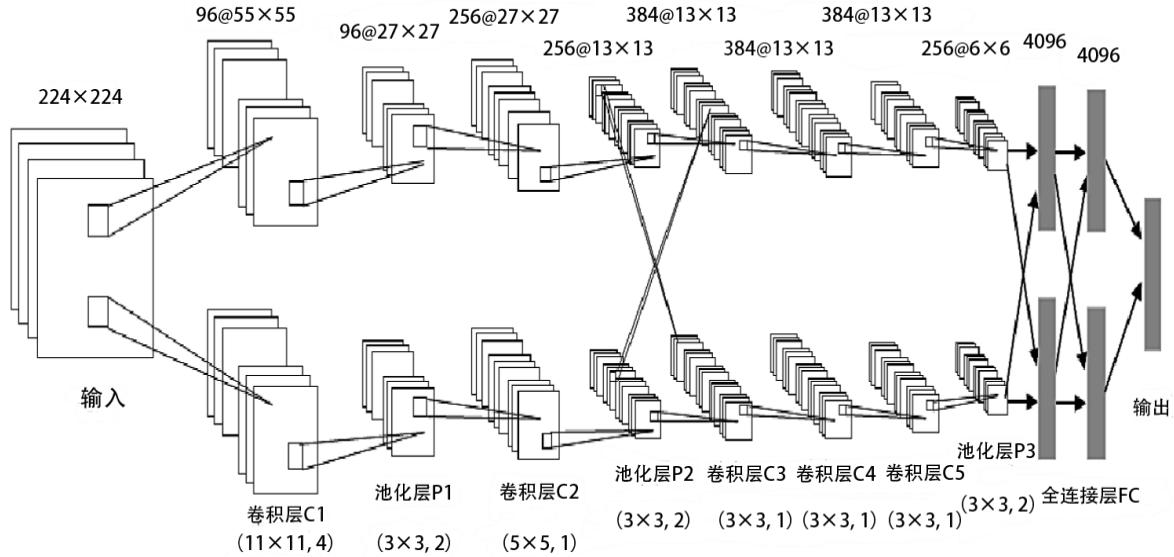


图 2.5 AlexNet 结构示意图

AlexNet 的主要改进包括：

- 采用线性整流单元（Rectified Linear Units, ReLU）取代 Sigmoid 函数作为网络中卷积层的激活函数，增加了模型的非线性，加快了模型的训练速度，解决了 Sigmoid 函数在网络较深时出现的梯度弥散问题；
- 全连接层采用的 Dropout 策略使隐含层节点有一定几率失活，这样可以提高所学习特征的鲁棒性，避免网络过拟合问题；
- 使用最大池化方法（max-pooling），在避免平均池化导致的模糊化问题的同时，提升特征的丰富性；
- 采用了 LRN（Local Response Normalization）结构，该结构可以创建局部神经元的竞争机制，刺激反馈较大的神经元，抑制其他反馈较小的神经元，从而提升模型的泛化能力。

2.2.3 VGGNet

VGGNet^[19]由 Oxford 大学 Visual Geometry Group 提出，旨在研究深度对卷积网络的影响。VGGNet 使用简单的 3×3 卷积，不断重复卷积层（中间有池化），最后经过全连接、池化、Softmax，得到输出类别概率。VGGNet 有 6 种不同类型配置，配置名称依次 A-E 为，深度从 11（8 个卷积核 3 个全连接）到 19（16 个卷积核 3 个全连接）；每个卷积层的深度范围为 64-512。具体配置如图 2.6。

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

图2.6 VGGNet配置

VGGNet 的主要特点包括：

- 使用了更小的卷积核 (3×3) 堆叠表示更大的卷积核 (5×5 或 7×7) 的视野，减少了模型的参数数量；使用了更深的网络，增加了模型的特征学习能力；
- 在 VGGNet 的 C 配置中，引入视野为 1×1 的卷积核，在不影响输入输出维度的情况下，引入非线性变换，增加了网络的数据表征能力；
- 采用 Multi-Scale 的方法进行训练和预测，增加了训练的数据量，防止模型过拟合，提升预测准确率。

2.2.4 GoogleNet

尽管神经网络的表达能力与特征学习能力不断提高，但是仍存在以下问题：参数空间大，容易过拟合，且训练数据集有限；网络结构复杂，计算资源不足，导致难以应用；深层次网络结构容易出现梯度弥散，模型性能下降。基于此，GoogleNet 对传统卷积层进行了修改，根据 NIN (Network-in-Network) 模型提出了一种能够高效表达特征的稀疏性结构，称为 Inception 结构，该结构可增加网络深度和宽度，提升深度神经网络性能。NIN 的结构如图 2.7。

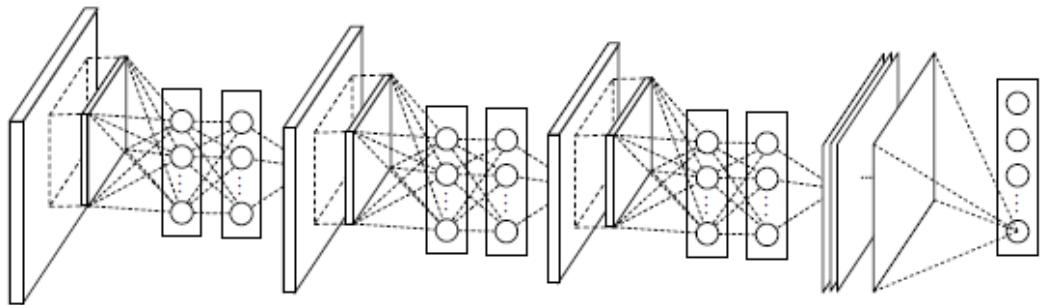


图 2.7 NIN 结构

Inception 结构共有四个版本，下面分别对其进行介绍：

(1) Inception v1 网络^[20]

将 1×1 , 3×3 , 5×5 的卷积层和 3×3 的最大池化层堆叠在一起（如图 2.8），一方面增加了网络的宽度，另一方面增加了网络对尺度的普适性。通过降维的方法，模型在不损失特征表达能力的前提下，减小了卷积核的数目，这使得模型复杂度大大降低。

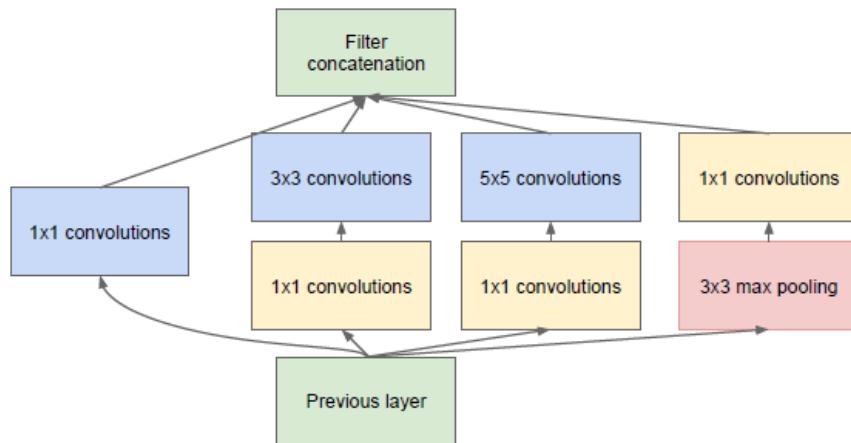


图 2.8 网络堆叠结构

(2) Inception v2 网络^[21]

一方面，加入 Batch Normalization (BN 层)，使每一层的输出都规范化到一个 $N(0, 1)$ 的高斯分布，减少了内部神经节的数据分布变化；另外一方面，使用 2 个 3×3 的卷积层替代 Inception v1 模块中的 5×5 的卷积层，降低参数数量，提升计算速度。

(3) Inception v3 网络^[22]

添加了分解卷积层 (Factorization) (如图 2.9)，将视野 7×7 和 3×3 的卷积层分别分解成两个一维的卷积层 ($1 \times 7, 7 \times 1$) ($1 \times 3, 3 \times 1$)，既加速了计算，又加深了网络深度，提高了网络的非线性。

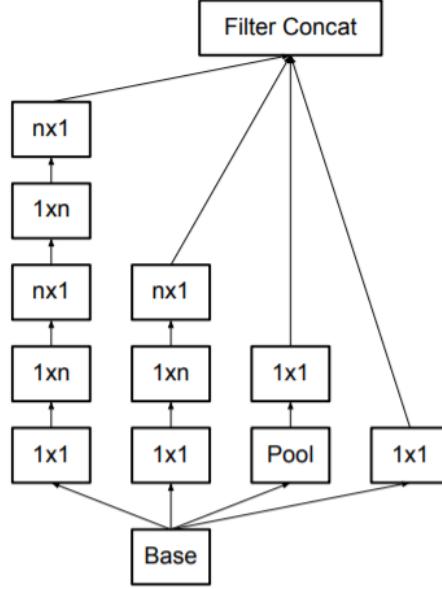


图 2.9 分解卷积层结构

(4) Inception v4^[23]

将 Inception 模块与残差连接(Residual Connection)结合, 得到 Inception-ResNet 网络, 提高了训练效率, 提升了模型性能。

2.2.5 ResNet

ResNet^[24]使用了跨连接（Shortcut Connection）的方法，引入了残差结构（Residual Net），如图 2.10。

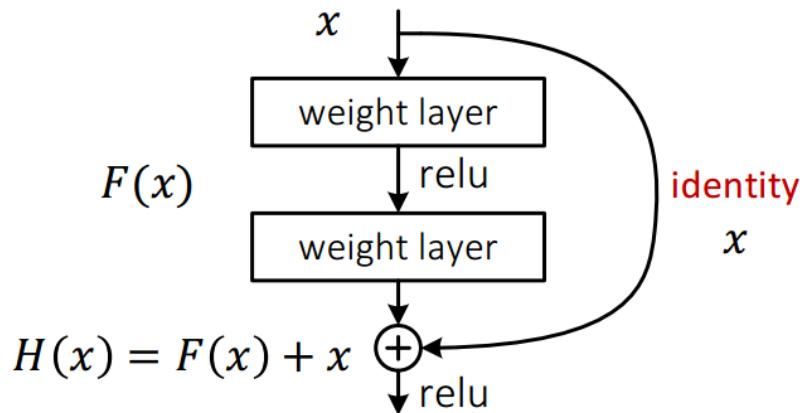


图 2.10 残差结构

其中, x 是某层网络的输入, $H(x)$ 是残差网络中某层原始的期望映射输出, $F(x)$ 是网络优化目标。该方法令模型的优化目标从“拟合输出 $H(x)$ ”变为“输出和输入的差($H(x) - x$)”, 这使得深层网络优化变得更容易, 能够有效地解决深层网络梯度消失的现象。

2.2.6 DenseNet

为解决网络结构加深导致的梯度消失现象，块正则化（Batch Normalization）通过规范化输入数据改变数据分布，在前传过程中消除梯度弥散。ResNet 的跨连接（Shortcut Connection）能更好地把梯度传到较浅的层次中，在后传过程中消除梯度弥散。基于前向过程和后传过程中消除梯度弥散的方法，Gao Huang^[25]等提出了一种新的网络结构，称为 Dense Block（如图 2.11）。该结构以“Batch Normalization + ReLU_激活函数 + 3×3 卷积核”作为组成函数，使得当前层与后面层的紧密连接（Dense Connectivity），使每一层的信息得到充分利用，从而在缓解梯度消失的同时，极大地减少参数。具体技术要点包括：

- 网络中的每一层直接与其前面层相连，实现特征的重复利用；
- 采用更窄的网络，只学习少量的特征图，达到降低冗余性的目的。

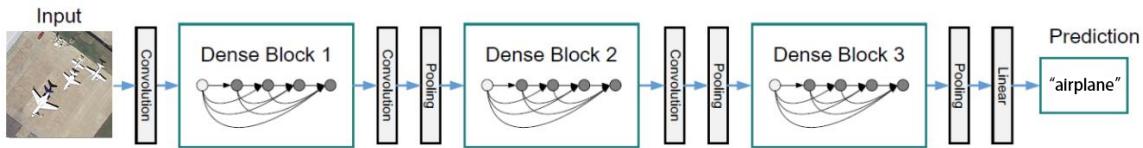


图 2.11 含有 Dense Blocks 的 DenseNet

2.2.7 CapsNet

卷积神经网络（CNN）中卷积层学习检测图像特征（较浅的层学习检测诸如边缘和颜色渐变等简单特征，较浅的层将简单特征组合成复杂特征），最大池化层可增加上层网络神经元的“视野”，减小模型参数。但是，由于卷积神经网络所组成高层特征的低层特征之间并不存在位姿（平移和旋转）关系，因此在网络的学习过程中，图像的空间关系特征往往丢失了。并且，卷积神经网络也容易混淆不同方向的图像，需要通过数据扩增将所有可能角度的图片纳入训练。

针对 CNN 以上的缺点，Geoffrey E. Hinton^[26]等人提出了一种新型的网络结构，称为胶囊网络（Capsule Network，CapsNet）。Capsule 是一组以向量表示特定实体类型的神经元，其输出向量长度表征实体存在的概率，向量方向表示实体的某些图形属性。前一层的 Capsule 通过变换矩阵对更高层的 Capsule 的实例化参数进行预测，从而激活更高层的 Capsule 单元。

Capsule 结构中使用“Squashing”非线性函数作为激活函数，在保留输入向量的方向的同时，将输入向量的长度压缩到区间 [0,1] 内。以下是该非线性函数的表达式：

$$V_j = \frac{\|s_j\|^2}{1+\|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (2.1)$$

其中, V_j 为 Capsule j 的输出向量, s_j 为 Capsule j 的输入向量。输入向量 s_j 的计算如下:

$$s_j = \sum_i c_{ij} \hat{u}_{j|i} \quad (2.2)$$

$$\hat{u}_{j|i} = W_{ij} u_i \quad (2.3)$$

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (2.4)$$

其中, c_{ij} 是动态路由算法 (Dynamic Routing algorithm) 中的连接参数, $\hat{u}_{j|i}$ 是先前层的预测向量, u_i 是先前层的输出向量, W_{ij} 是权重矩阵, b_{ij} 是初始未归一化概率。Capsule 层级结构如图 2.12。

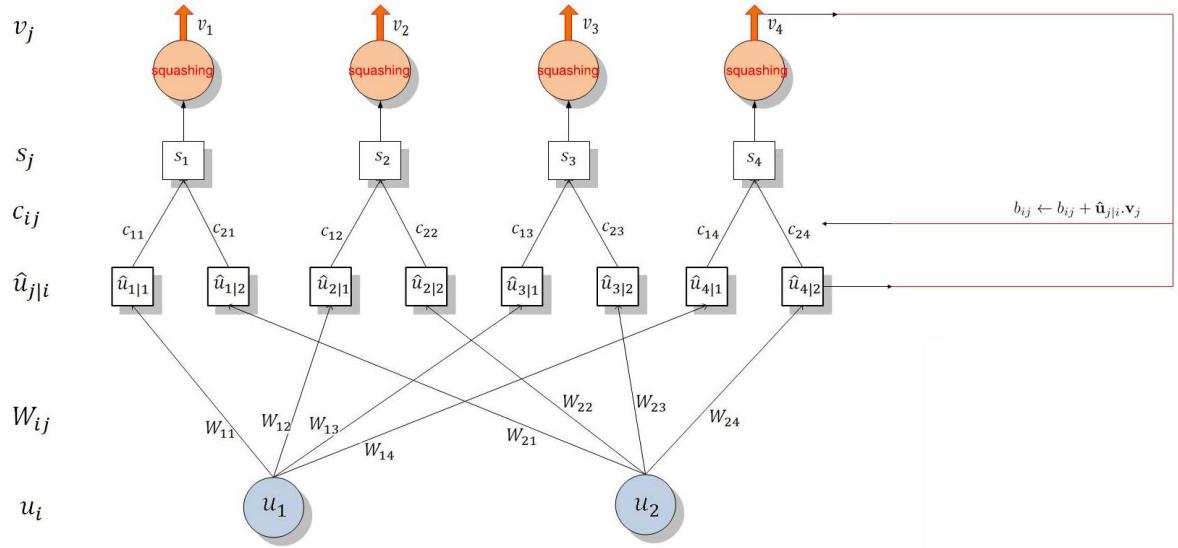


图 2.12 Capsule 层级结构图

CapsNet 由 1 个使用 ReLu 作为激活函数的传统卷积层, 1 个输出向量长度为 8、Capsule 单元个数为 32 的胶囊卷积层(PrimaryCaps), 1 个输出向量长度 16、Capsule 单元个数为 10 的胶囊卷积层(DigitCaps)构成, 并采用 Margin loss 作为损失函数。具体结构如图 2.13。

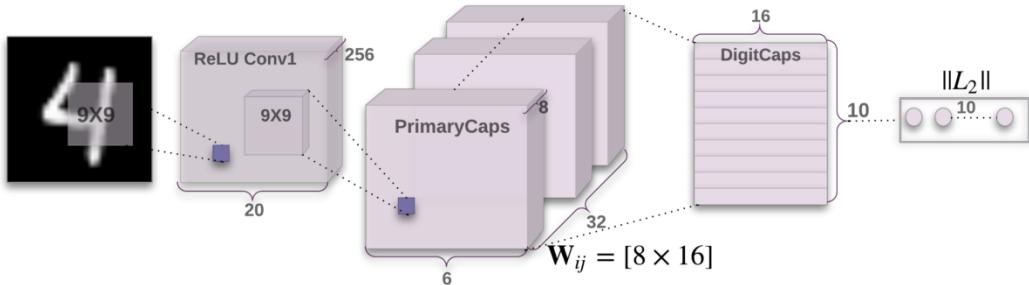


图 2.13 CapsNet 结构图

同时，CapsNet 可利用三层全连接层（FC）对 DigitCaps 层进行解码操作，生成图像。具体结构如图 2.14。

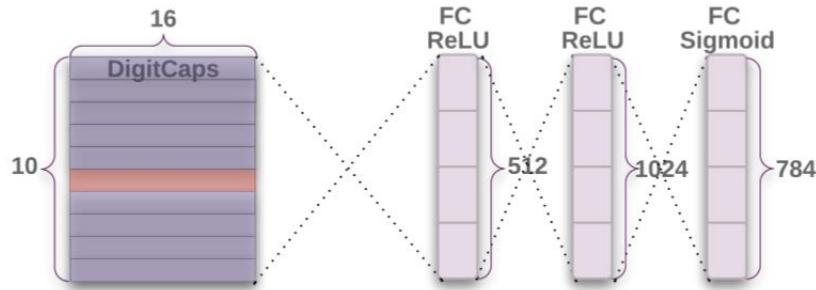


图 2.14 CapsNet 解码结构

表 2.1 具体显示了胶囊结构（Capsule）和传统的神经节（tradition neuron）结构的在输入（Input）、参数计算（Operation）、输出（Output）、激活函数（Activation）等方面的不同。

表 2.1 胶囊结构与传统的神经节对比

胶囊结构		传统神经节
Input from low-level neuron/capsule	vector(u_i)	scalar(x_i)
Operation	Affine Transformation	$\hat{u}_{j i} = W_{ij} u_i$
	Weighting	$s_j = \sum_i c_{ij} \hat{u}_{j i}$
	Sum	$a_j = \sum_{i=1}^3 W_i x_i + b$
	Non-linearity activation fun	$v_j = \frac{\ s_j\ ^2}{1 + \ s_j\ ^2} s_j$
output	vector(v_j)	scalar(h)
$\begin{array}{l} u_1 \xrightarrow{w_{1j}} \hat{u}_1 \\ u_2 \xrightarrow{w_{2j}} \hat{u}_2 \\ u_3 \xrightarrow{w_{3j}} \hat{u}_3 \\ \quad +1 \end{array}$		$\begin{array}{l} x_1 \xrightarrow{w_1} \\ x_2 \xrightarrow{w_2} \\ x_3 \xrightarrow{w_3} \\ \quad b \\ \quad +1 \end{array}$
$\begin{array}{l} \hat{u}_1 \xrightarrow{c_1} \\ \hat{u}_2 \xrightarrow{c_2} \\ \hat{u}_3 \xrightarrow{c_3} \\ \quad b \end{array}$		$\begin{array}{l} \Sigma \quad squash(\cdot) \rightarrow v_j \\ \Sigma \quad f(\cdot) \rightarrow h_{w,b}(x) \end{array}$

2.3 本章小结

本章先是简要介绍了神经网络算法，然后具体介绍了多个深度神经网络模型，包括 6 种卷积神经网络架构和 1 种胶囊神经网络架构。

3 遥感影像场景分类

3.1 基于神经网络的遥感影像场景分类

3.1.1 深度学习方法

深度学习方法具有强大的数据学习能力和优良的分类准确率，场景分类是遥感图像解译的一个重要环节，将深度学习方法应用于遥感影像场景分类流程如图 3.1。具体说明如下：

- (1) 搭建深度学习网络框架；
- (2) 将影像场景数据集进行随机打乱，并按照 9:1 的比例随机划分训练集和测试集；
- (3) 利用所构建的数据集，配置合适的超参数（优化器、学习率、迭代次数等）完成对网络的训练；
- (4) 对分类结果进行评估，计算分类准确率（accuracy）、精度（precision）、召回率（recall）、F1 参数、Kappa 系数等指标。

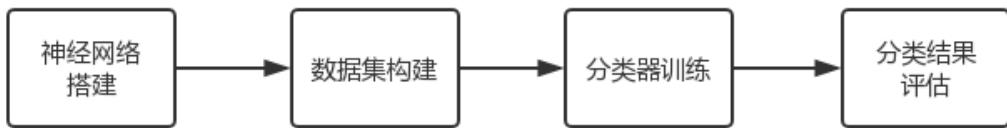


图 3.1 深度学习方法流程图

3.1.2 基于 HOG 特征的支持向量机

本小节介绍了一种结合方向梯度直方图（HOG）和支持向量机（SVM）的场景分类方法，该方法先对影像场景进行方向梯度直方图特征提取，再针对所提取的特征以支持向量机作为分类器进行遥感影像场景分类（如图 3.2）。以下具体介绍 HOG 特征和 SVM 分类器。



图 3.2 基于 HOG 特征的支持向量机分类流程图

3.1.2.1 方向梯度直方图

方向梯度直方图^[27]（Histogram of Oriented Gradient, HOG）核心思想是所检测的局部物体外形能够被梯度或边缘方向的分布所描述，能较好地捕捉局部形状信息，对几何和光学变化都有很好的不变性。HOG 主要分为以下步骤：

- 颜色空间归一化：将影像转化为灰度图，并对整幅图像进行归一化，从而减少了光照因素的影响。
- 计算梯度：根据图像横坐标和纵坐标方向的梯度，计算图像每个像素位置的梯度方向值。梯度计算可以更好地获取对象轮廓信息，并且进一步弱化光照影响。
- 方向梯度直方图：将图像分成若干个单元（cell），对单元内每一个像素在其梯度方向进行加权投影，进而得到该单元的梯度直方图。
- 组成区域：将几个单元组成一个区域（block），一个区域内所有单元的特征描述子（descriptor）串联起来得到该区域的 HOG 特征描述子。
- 组成特征：叠加图像内的所有区域的 HOG 特征描述子，获得该图像整体的 HOG 特征描述子。

3.1.2.2 支持向量机

支持向量机（Support Vector Machine, SVM）是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上的算法^[28]。基于少量样本信息，支持向量机可在模型的复杂性和学习能力之间寻求最佳折衷，以期获得最好的泛化能力^[29]。该算法在小样本、非线性及高维模式识别任务中优势明显。

3.1.3 视觉词袋模型

3.1.3.1 视觉词袋模型

视觉词袋模型（Bag-Of-visual-Words, BOW）模型^{[30][31]}是一种经典的文档建模方法，该方法最开始面向纹理分析问题。模型具体操作流程如下：首先，获取定量局部特征（如 SIFT、HOG 特征）；其次，利用聚类算法（如，K-means）选择主要特征，建立视觉词频；再者，利用视觉词频矩阵表达影像；最后，图像词向量可作为分类器输入进行图像分类。由于视觉词袋模型没有利用图像中的空间信息，因此视觉词袋模型对图像的局部细节特征有较强的依赖。视觉词袋模型实现流程如图 3.3。



图 3.3 视觉词袋模型流程图

以机场、高速公路、储油罐三张图片为例，具体介绍视觉词袋模型实现步骤

[32][33]:

- 把图像分割多块 (Patch), 并对每块的中心点计算 Sift 特征, 每个 Sift 向量为 128 维。(如图 3.4)



图 3.4 机场、高速公路、储油罐三张图片 SIFT 提取结果示意图

- 利用 K-means 算法, 将这些块的中心点的 SIFT 特征聚成为 K 类, 用这 K 个聚类中心 (Enters) 来构造单词表。图 3.5 显示了三张图片全部的特征, 图 3.6 显示了三张图片特征聚类结果。

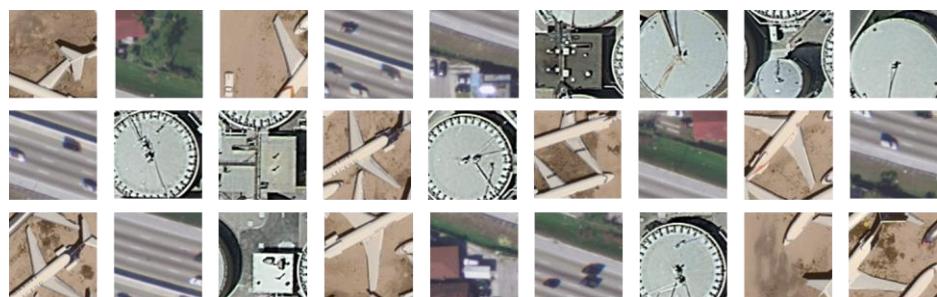


图 3.5 人脸、自行车、吉他三张图片全部的特征

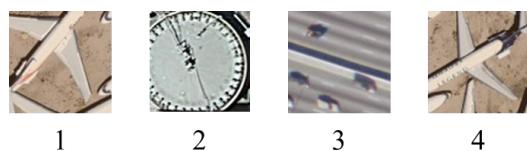


图 3.6 K-means 特征聚类结果 (K=4)

- 利用单词表表示每幅图像, 则每幅图像被表示成了一个与词序列相对应的词频向量 (K 维)。图 3.7 显示了三张图像词频表示结果。

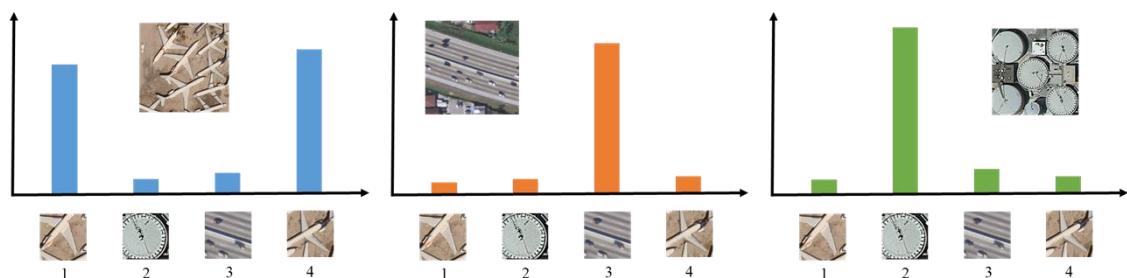


图 3.7 三张图像词频表示结果

- 以图像词向量作为分类器(本文采用 SVM 作为分类器)输入进行图片分类。

3.1.3.2 K 均值算法

K 均值算法^{[34][35]} (K-means) 是利用近邻信息来标注类别的一种聚类算法。该方法的核心思路为：指定 K 个初始质心(Initial Centroids)作为聚类的“簇”(Cluster)；计算剩余数据点到每个初始质心的距离，并把数据点和距离其最近质心标记为同一类别；在每一个“簇”内部，计算各“簇”数据点的均值向量，并将其设置为新的聚类质心；重复迭代直至新的质心与原质心之间距离小于指定阈值或算法达到最大迭代轮次。通常情况，算法以平方误差和（Sum of the Squared Error, SSE）作为聚类的目标函数，最优的聚类结果应使得 SSE 达到最小值。

3.1.2.3 尺度不变特征变换

尺度不变特征变换^[36] (Scale-Invariant Feature Transform, SIFT) 是一种图像局部特征描述子，该算法具有尺度不变性，主要分为以下步骤：

- 检测尺度空间的极值：构建尺度空间，通过高斯微分函数检测极值点，识别尺度和旋转鲁棒性较强点。
- 定位特征点：通过尺度空间 DoG 函数进行曲线拟合以确定位置尺度，根据尺度稳定程度过滤特征点，精确定位关键点。
- 赋值特征方向：计算图像局部梯度方向，为关键特征点分配一个或多个方向值。
- 生成特征描述子：以特征点为中心取 16×16 的邻域作为采样窗口，将采样点与特征点的相对方向通过高斯加权后归入方向直方图（8 个区间），最后获得 128 维的特征描述子。

3.1.4 深度信念网络

深度信念网络^{[37][38][39]} (Deep Belief Network, DBN) 是一种概率生成模型神经网络，通过建立一个观察数据和标签之间的联合分布，对 $P(\text{Observation}|\text{Label})$ 和 $P(\text{Label}|\text{Observation})$ 都做评估。DBN 由多个受限波尔兹曼机 (Restricted Boltzmann Machines, RBM) 以串联的方式堆叠而形成的^{[40][41]}。训练时通过自低到高逐层训练 RBM，将模型参数初始化为较优值，再使用全局反向传播算法对网络进行微调，使得模型收敛到接近最优值的局部最优点。RBM 可以利用对比散度 (Contrastive Divergence, CD)^[42] 等算法进行快速训练，该算法可将整体模型训练等效为多个 RBM 的训练，降低了直接训练模型的计算量，提升了模型推广能力。图 3.8 展示了 DBN 的结构。

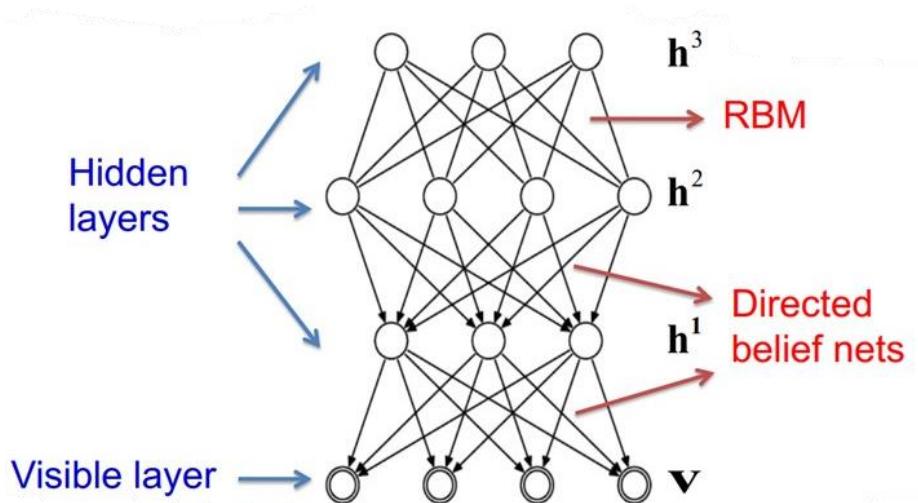


图 3.8 DBN 结构示意图

3.2 基于微训练模型的遥感影像场景分类

在神经网络中，充分利用数据的模型往往可以取得更好的效果，提取自预训练 CNN 的高层特征也利于提高模型的准确率。本节将介绍一种基于 ImageNet 数据集的微训练模型（Fine-Tuning）解决数据集有限的问题，提高模型的准确率，并减小模型训练所用时间。模型架构如图 3.9。

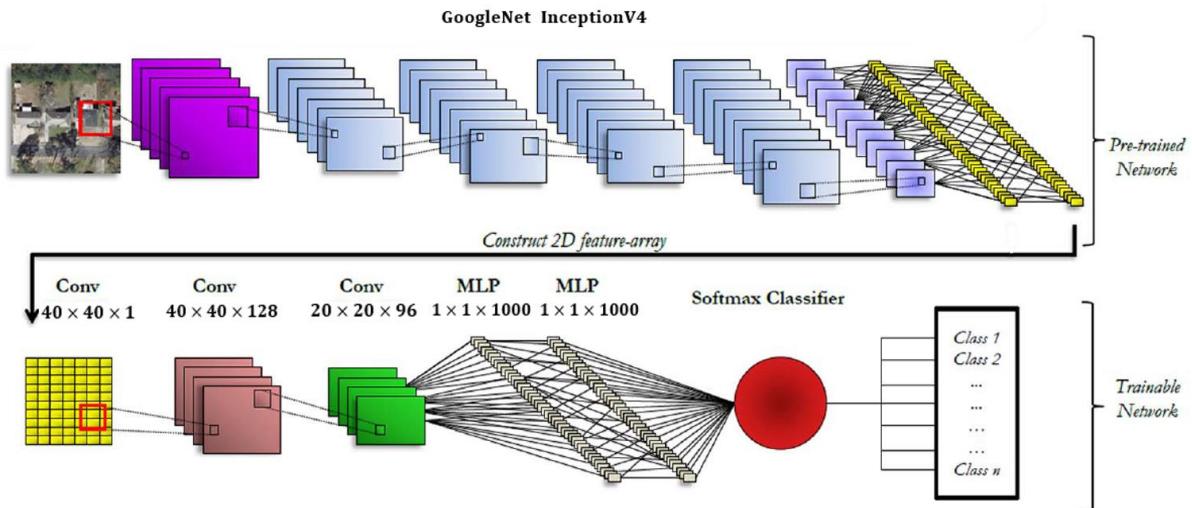


图 3.9 基于微训练模型的遥感影像场景分类模型架构

3.2.1 模型预训练

3.2.1.1 ImageNet 数据集

ImageNet 数据集是与视觉分类任务有关预训练模型所用的主要数据集，数据集中图像可在数据库中通过搜索引擎进行检索。ImageNet 由约 1500 万张已标记的高分辨率图像组成，大致分为 22000 个类别。在 ImageNet 框架下，每年举办视觉

识别竞赛（ImageNet Large-Scale Visual Recognition Challenge, ILSVRC），参与者可以使用这些数据的一个子集训练他们所选的分类算法。该比赛中使用的训练集的包括约 120 万张图片，同时参与者利用训练好的模型对 1000 个不同类别的测试集图片进行分类。

3.2.1.2 遥感影像分类的预训练网络

J. Donahue^[43]等人表明根据 ImageNet 数据集得到的深度预训练模型中的特征信息可以在很好地应用在不同分类任务中，可以取得比基本的分类算法更好的效果。Dimitrios Marmanis^[44]认为深度网络对于特征的表达在不同分类场景下也是可迁移的。

在遥感领域，标记数据往往是稀缺的，这需要投入大量时间建立地面实况标准。因此，可以利用已训练好的网络进行特征表达，实现遥感影像分类。由于大型预训练模型往往包含了图片共性的低维特征（边缘、角点），因此可在最小限度调整预训练模型的参数情况下，将预训练模型应用于不同的视觉分类任务。

本章采用 GoogleNet Inception V4 作为预训练网络（Pre-trained Network），如图 3.9。GoogleNet Inception V4 输入图像大小为 $299 \times 299 \times 3$ ，这里将 UCML 数据集从 $256 \times 256 \times 3$ 上采样为 $299 \times 299 \times 3$ 。

3.2.2 特征融合框架

本小节介绍了一种从预训练模型中提取特征，并进行二次训练的方法。

从 GoogleNet V4 中提取最后一层全连接层（ 1×1536 ）作为特征向量，并通过补零将其重塑为二维矩阵（ 40×40 ），如图 3.9。主要有以下要点：

- 预训练卷积神经网络最后一层的特性向量对分类准确度的影响最大^[43]。因此，从预训练模型的最后一层提取特性向量。
- 预训练模型中所提取的特征由一维向量重塑为二维矩阵，可减小第二个 CNN 网络的参数，并且特征维度的增加使得特征之间的空间关联度增加。

可训练网络（Trainable Network）对于提高模型判别力，学习高度抽象特征的复杂关系，正确地将数据分成互斥类至关重要。尽管其他分类算法也可以提供准确的结果，但卷积神经网络对于二维输入数据有着更显著的空间适应性。因此，这里可训练模型采用卷积神经网络架构（Trainable CNN），对得到的二维矩阵进行训练，具体包含 2 个卷积层，2 个最大值池化层，2 个全连接的层，以及顶部的 Softmax 分类器（具体参数请参阅图 3.9）。

3.3 实验与分析

3.3.1 实验数据

3.3.1.1 UCML 数据集

UCML 数据集 (UC Merced Land Use Dataset, UCML) 是从美国各个城市地区的大型遥感影像中手动提取的，共包含 21 种场景类别。每一种类别都包含 100 幅大小为 256×256 像素，空间分辨率为 1 inch/pixel 的影像。在 UCML 数据集中，不同类别的纹理模式与目标对象可能相同（如，稠密住宅与中等住宅），这使得 UCML 数据集的分类难度较大。图 3.10 展示了 UCML 数据集中部分类别的示例影像。

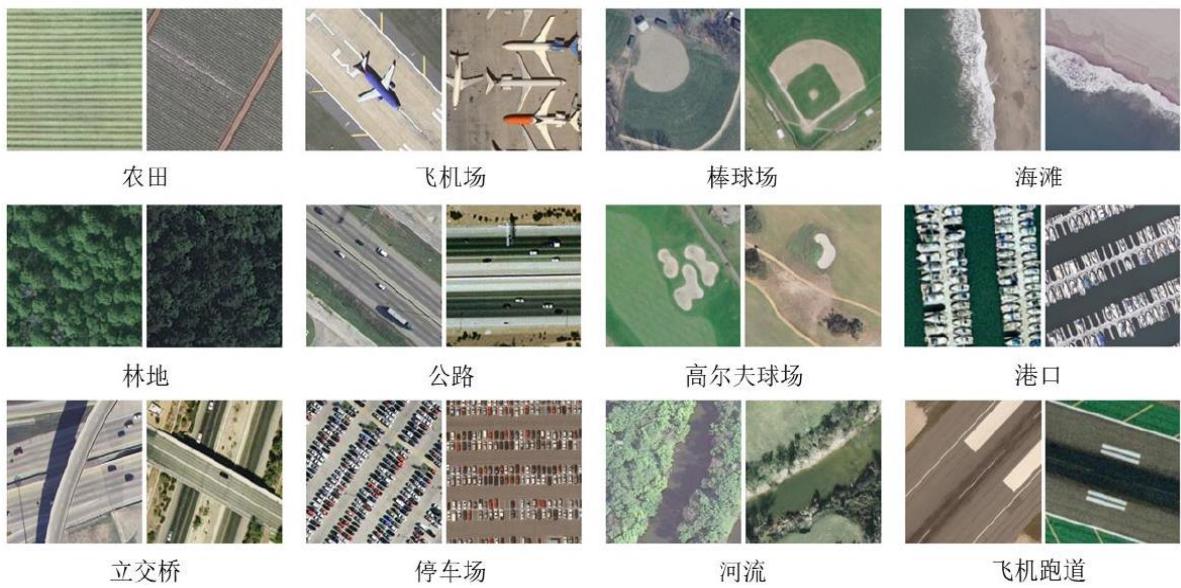


图 3.10 UCML 数据集示例

3.1.1.2 数据预处理

所有图片灰度值采用概率归一化方法（从 $[0,255]$ 归一化为 $[0,1]$ ），这样可以提升模型训练收敛速度，并使效果更为平稳。同时采用数据扩增方法，利用随机旋转变换、平移变换、翻转变换、尺度变换、剪切变换等方法，基于原始图片生成更多训练样本，扩充数据集的范围，减小模型过拟合风险，提升模型泛化能力。

3.3.2 实验配置

表 3.1 实验硬件配置

硬件名称	指标
显卡	NVIDIA GeForce GTX 850M
内存	12G
CPU	Intel core i7 2.50GHZ

3.3.3 精度评价指标

精度（precision）反映了被分类器判定的正例中真正的正例样本的比重；

召回率（recall）反映了被正确判定的正例占总的正例的比重；

准确率（accuracy）反映了分类器统对整个样本的判定能力；

F1 参数（F1-measure）是精度和准确率的调和平均数。

计算公式分别如下：

$$Precision = \frac{tp}{tp+fp} \quad (3.1)$$

$$Recall = \frac{tp}{tp+fn} \quad (3.2)$$

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (3.3)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (3.4)$$

在二分类问题中，

tp 为真正例（True Positives）：预测模型为 1，真实样本也为 1。

tn 为真负例（True Negatives）：预测模型为 0，真实样本也为 0。

fp 为假正例（False Positives）：预测模型为 1，但真实样本为 0。

fn 为假负例：（False Negatives）：预测模型为 0，但真实样本为 1。

多分类问题可以看作多个二分类问题的集合（0 代表不是某类别，1 代表是某类别），多个二分类指标可合并形成多分类评价指标。主要有三种合并方式：按全局方式计算指标（micro）；按标签类别计算指标，并求平均值（macro）；按标签类别依次计算指标，根据标签类别的样本数量计算权重，按“指标×权重”方式求平均值（weighted）。在本实验中均采用 weighted 的合并方式。

Kappa 系数用于一致性检验和分类精度衡量，其计算公式如下：

$$Kappa = \frac{p_o - p_c}{1 - p_c} \quad (3.5)$$

$$p_0 = \frac{\sum_{i=1}^r x_{ii}}{N} \quad (3.6)$$

$$p_c = \frac{\sum_{i=1}^r (x_{ii} \times x_{+i})}{N^2} \quad (3.7)$$

其中, p_0 称为观测精确性或一致性单元的比例(混淆矩阵对角线上元素的和),
 p_c 称为偶然一致单元的比例。

Kappa 系数的取值范围为-1~1, 但通常 Kappa 系数在 0~1 间。根据一致性高低, Kappa 系数具体又可分为五种: 0.81~1 表示几乎完全一致性(almost perfect), 0.61~0.80 表示高度一致性(substantial), 0.41~0.60 表示中等一致性(moderate), 0.21~0.40 表示一般一致性(fair), 0.0~0.20 表示极低一致性(slight)。

3.3.4 实验结果与分析

3.3.4.1 深度神经网络

表 3.2 算法分类结果

算法		训练集		测试集			
名称	参数个数	Accuracy(%)	Accuracy(%)	Precision(%)	Recall(%)	F1(%)	Kappa
LeNet	23,870,317	100.00	58.57	58.57	58.57	59.18	0.5630
AlexNet	58,367,381	99.68	63.81	63.81	63.81	63.20	0.6182
VGG-VD 16	165,705,539	100.00	75.24	75.24	75.24	74.92	0.7393
GoogleNet V2	6,838,725	99.37	82.86	82.86	82.86	82.65	0.8195
ResNet34	22,697,237	99.79	67.14	67.14	67.14	66.41	0.6548
DenseNet (13-8)	39,581	82.17	70.90	70.90	70.90	68.14	0.6548
CapsNet	9,927,696	50.11	49.05	49.05	49.05	47.18	0.4627
HOG+SVM (Linear)			36.19	36.19	36.19	35.30	0.3305
Linear			41.90	41.90	41.90	39.69	0.3899
BOW (K=50)	Rbf		49.05	49.05	49.05	48.95	0.4639
Sigmoid			19.52	19.52	19.52	17.87	0.1609
Poly			39.52	39.52	39.52	40.27	0.3637
DBN		15.71	15.71	15.71	15.71	13.42	0.1099

各算法分类结果如表 3.2 所示, 表格主要展示了 LeNet、AlexNet、VGG-VD16、GoogleNet V2、ResNet34、DenseNet(13-8)、CapsNet 七种深度学习网络的精度、召回率、准确率、F1 参数、Kappa 系数等指标。其次, 比较了 HOG+SVM、BOW、DBN 算法在这些指标方面与深度学习网络的差异。实验中, 将 UCML 数据集分为训练集和测试集, 训练集和测试集的样本数量之比为 9:1。实验结果显示, 在遥感影像场景数据集上, 深度学习方法均取得比传统算法更好的分类结果。GoogleNet V2 模型在 UCML 数据集上获得了最好的结果表现, 其准确率为 82.86%, Kappa 为 0.8195; DBN 模型则表现最差, 其准确率仅有 15.71%, Kappa 为 0.1099。值得注意的是, 深度学习网络在训练集获得的准确率往往大于测试集准确率, 其中 LeNet 的训练集准确率设置达到 100%, 而其测试集准确率只有 58.57%, 可见 LeNet 网络存在严重的过拟合现象。此外, 网络还存在着由于数据量小导致特征学习不充分等问题。

图 3.11—图 3.17 分别显示了 LeNet、AlexNet、VGG-VD、GoogleNet V2、ResNet34、DenseNet(13-8)、CapsNet 模型的准确率 (accuracy) 与损失值 (loss) 随迭代次数 (epoch) 变化情况。实验中, 批处理数 (batch) 设置为 32, 最大迭代次数 (max epoch) 设置为 100。从图 3.11 可以看出模型的测试集损失值在 40 次迭代之后逐渐上升, 因此验证了 LeNet 网络存在严重过拟合现象。图 3.12 中准确率与损失值波动较大, 可能由于 batch 值过小导致。图 3.13 中在 epoch=60 处左右存在一处明显尖峰, 可能是由于鞍点或极小值点导致。图 3.14、图 3.15 中, 模型拟合速度明显加快, 说明 Inception 结构、Shortcut Connection 结构有利于加快模型训练速度。图 3.16 中, 模型逐渐趋于拟合, 说明 Dense Block 结构有利于训练后期梯度消失的情况, 并且减小了 DenseNet 中参数。图 3.17 中训练集和测试集的准确率与损失值基本相同, 说明网络基本不存在过拟合现象, 但是 CapsNet 准确率在七种网络中最低, 说明 CapsNet 对遥感影像场景的描述能力仍然有待提高。

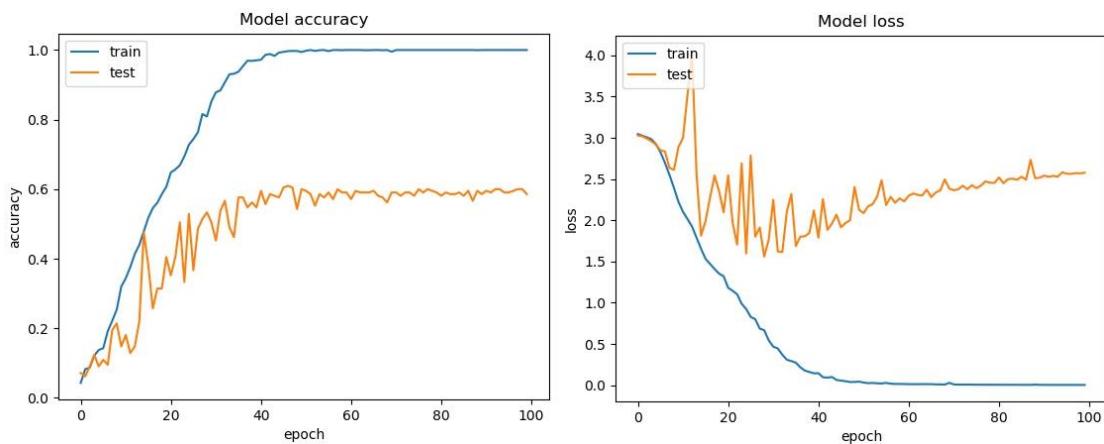


图 3.11 LeNet 模型准确率 (accuracy) 与损失值 (loss)

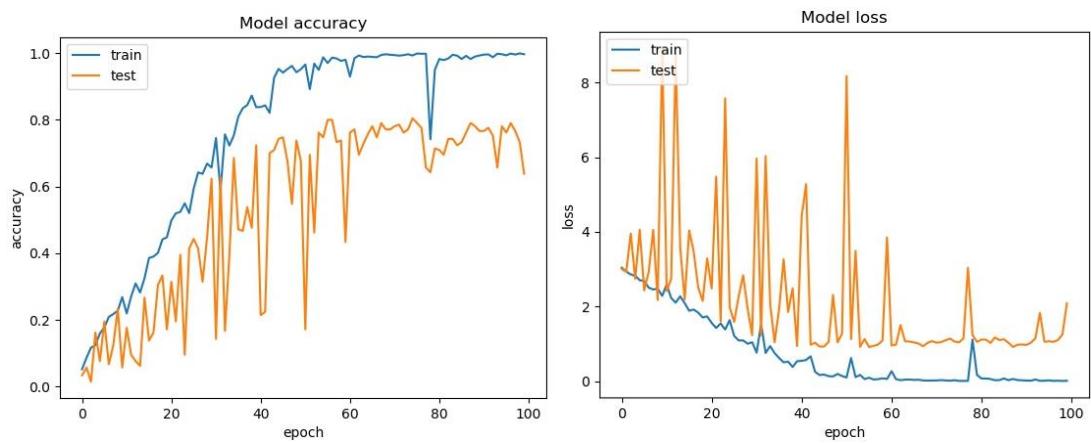


图 3.12 AlexNet 模型准确率 (accuracy) 与损失值 (loss)

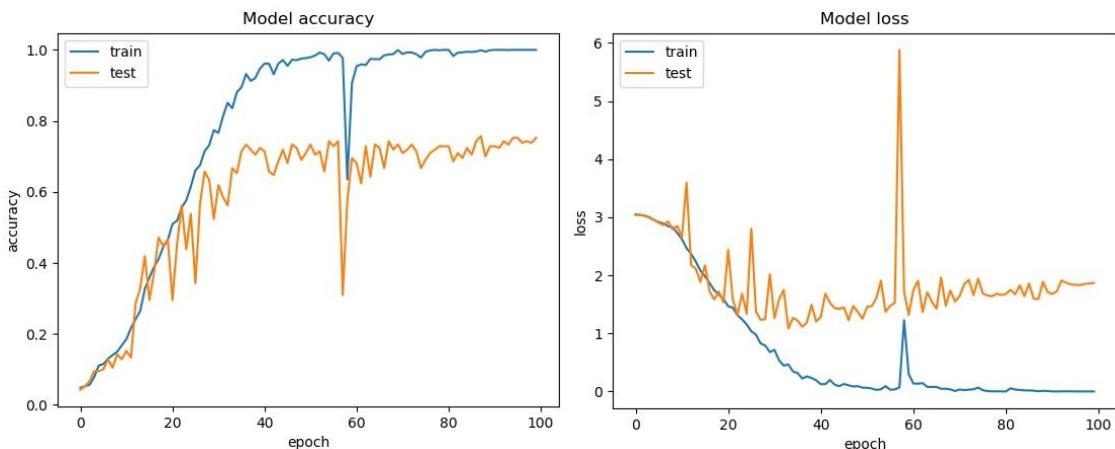


图 3.13 VGG-VD16 模型准确率 (accuracy) 与损失值 (loss)

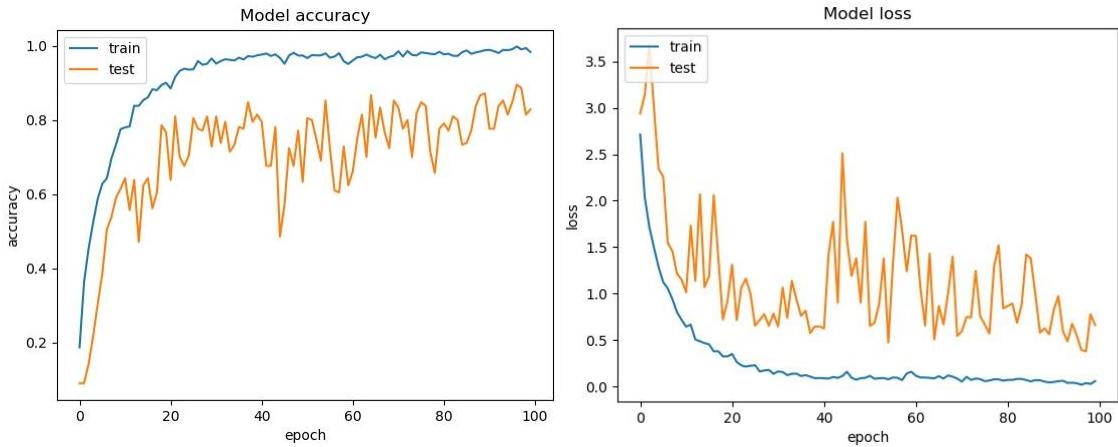


图 3.14 GoogleNet V2 模型准确率 (accuracy) 与损失值 (loss)

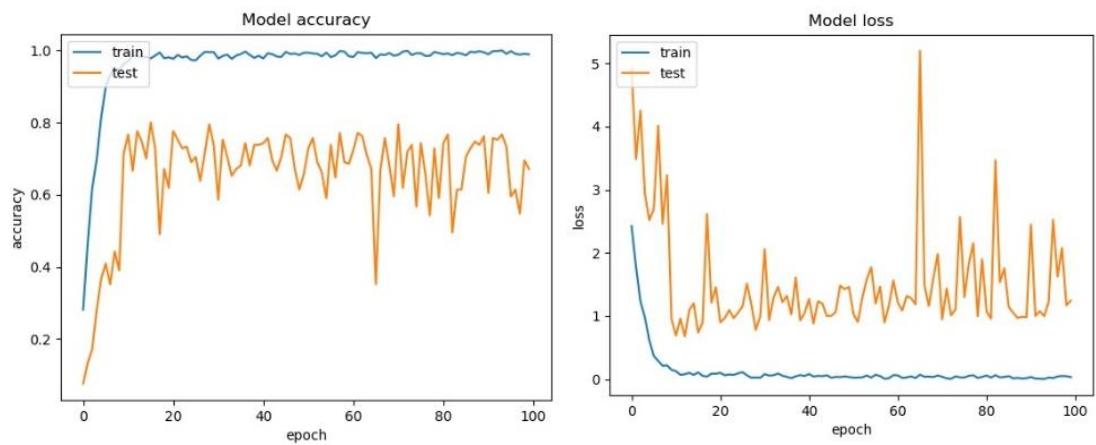


图 3.15 ResNet34 模型准确率 (accuracy) 与损失值 (loss)

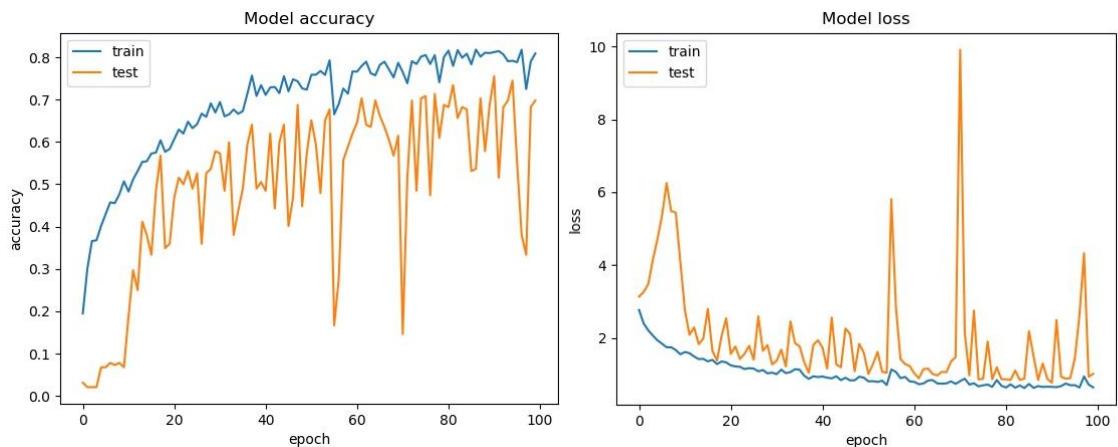


图 3.16 DenseNet(13-8)模型准确率 (accuracy) 与损失值 (loss)

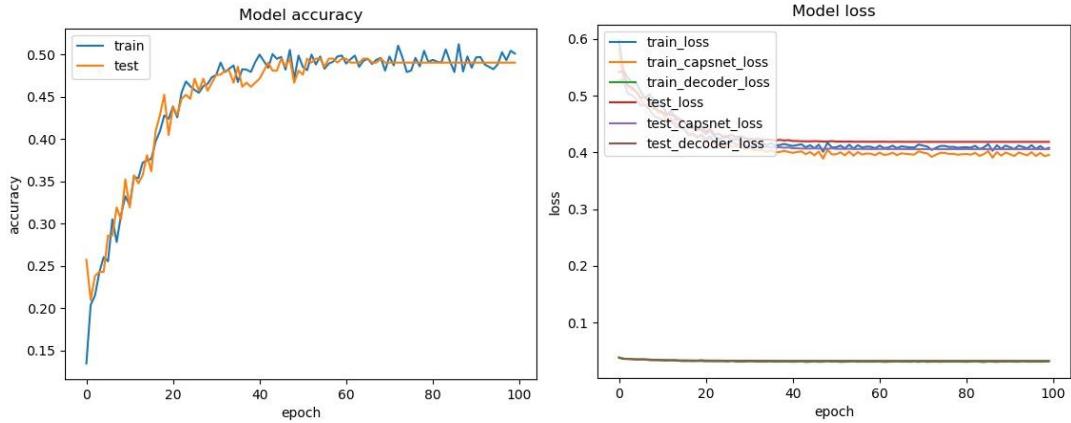


图 3.17 CapsNet 模型准确率 (accuracy) 与损失值 (loss)

3.3.4.2 微训练模型

这里利用预训练模型，采用 CNN 结构提取其中特征，并基于 UCML 数据集训练这些特征，得出分类结果（如表 3.3）。其中，数据集与测试集比例为 9:1。

表 3.3 微训练模型的场景分类结果

网络名称	训练集		测试集			
	Accuracy(%)	Accuracy(%)	Precision(%)	Recall(%)	F1(%)	Kappa
GoogleNet	100.00	91.91	91.91	91.91	91.86	0.9148
V4 + CNN						

从表 3.3 可以看出，微训练模型的测试集准确率可以达到 91.91%，相比较直接使用数据集进行训练获得的最优结果（82.86%）提高了 9.05%。说明从预训练网络中提取特征有利于解决原始数据集数据量小的问题，更容易提高网络分类表现。微训练模型的网络准确率（accuracy）与损失值（loss）如图 3.18 所示，图 3.19 显示了 t-SNE 算法^[45]对图像特征降维后的结果。

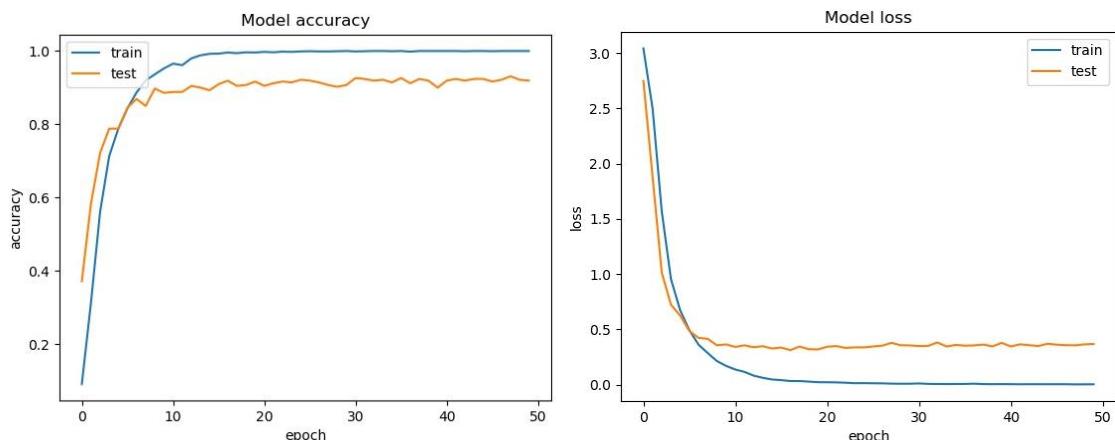


图 3.18 微训练模型的准确率 (accuracy) 与损失值 (loss)

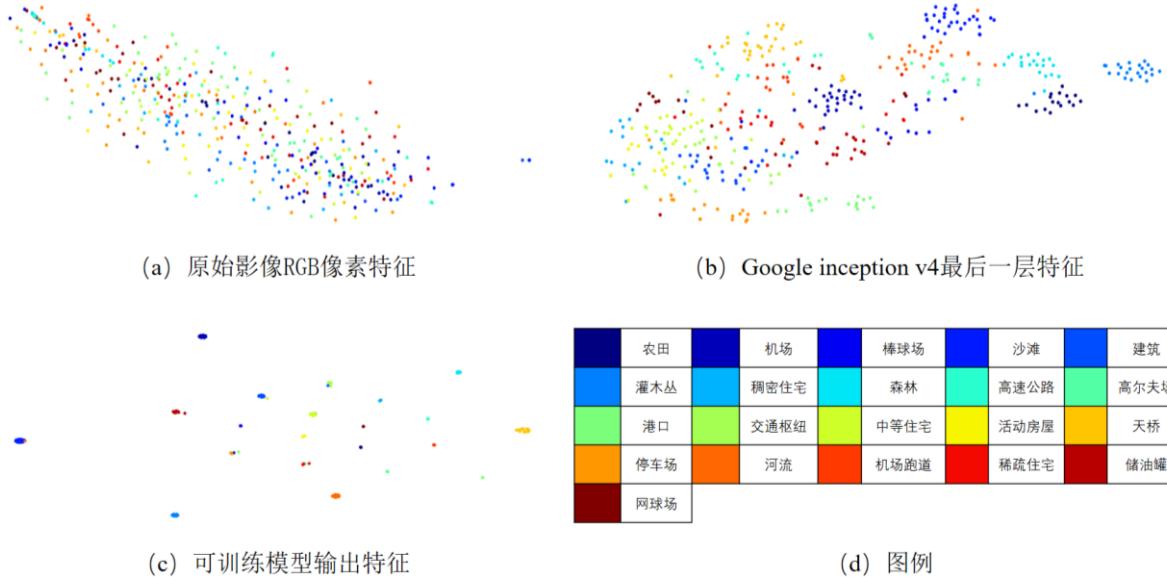


图 3.19 t-SNE 算法^[45]对图像特征降维后的二维散点图

3.4 本章小结

面向遥感场景分类任务，本章给出了两种分类方案：一种是基于神经网络的分类方案，采用遥感影像训练集直接训练网络，并利用测试集检测分类结果；另一种方案先利用预训练模型提取遥感影像数据集特征，再搭建卷积神经网络对所提取特征进行训练。最后，采用 UCML 数据集对两种方案进行测试，同时与传统分类方案进行比较，并分析了实验结果。

4 高光谱遥感影像分类

高光谱遥感影像是一种数据立方体，通过结合空间和光谱信息，实现了图谱合一^[46]。高光谱遥感可以在空间和光谱维度上实时得到连续数据，这令高光谱数据成为监测地球表面有价值的工具^{[47][48]}，并在农业^[49]，矿物学^[50]，监测^[51]，物理学^[52]，天文学^[53]，化学成像^[54]和环境科学等方面取得广泛的应用。

基于高光谱的分类任务主要有如下难点：（1）因为数据包含大量频谱信道，往往导致数据维度爆炸；（2）已标记的训练样本数量有限；（3）由于存在光照、传感器角度、大气、散射条件等差异，同一类别的对象光谱特性空间差异性大^[55]。传统方法把高光谱中的像元作为独立的光谱曲线进行分类，忽视了图像中的空间信息；利用 PCA、ICA、NWE 等图像降维算法，则容易丢失高光谱图像中光谱维度细节信息，降维后的图像与多光谱无异。

本章介绍一种基于深度学习通过联合光谱-空间信息对高光谱影像数据进行分类的方法。

4.1 模型整体架构

本节中简要描述方法的整体架构。首先，提取某位置点的光谱；其次，利用主成分分析的（PCA）降维方法，减少原始数据维度^[56]；再者，提取位置点的邻域像素，并利用卷积神经网络对空间信息进行特征提取；最后，将位置点的光谱信息和经过特征提取后的空间信息融合，并输入多层感知器（Multi-layer Perceptron, MLP）进行分类。整体架构如图 4.1 所示。

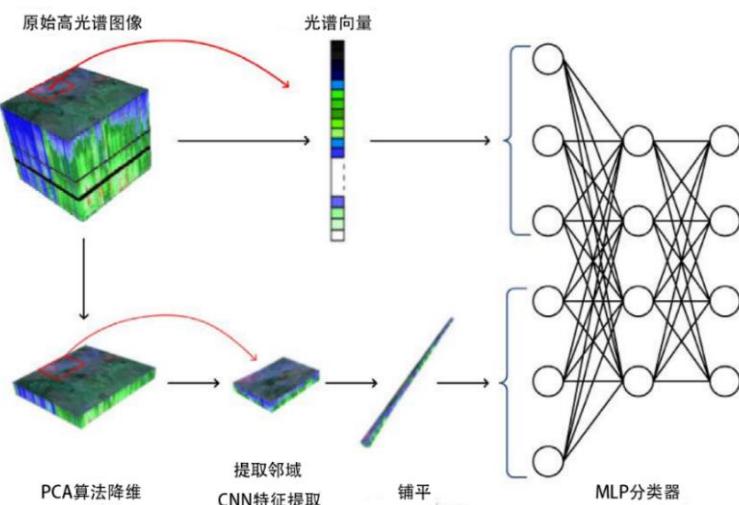


图 4.1 空间-光谱信息混合模型整体架构

4.2 光谱信息

4.2.1 光谱向量提取

对于表示为三维张量尺寸 $h \times w \times c$ 的高光谱图像（ h 和 w 对应于图像的高度和宽度， c 对应其通道），设 $\mathbf{p}_{x,y}$ 表示空间中在 (x, y) 处的光谱向量， $l_{x,y}$ 表示该点的标记（ground truth）。依次提取每一点的光谱向量 $\mathbf{p}_{x,y}$ 作为 MLP 的部分输入。

4.2.2 R-PCA 降维

训练卷积神经网络需要对输入数据的每一个光谱维度进行卷积运算，而对于频谱维度达到数百量级的高光谱遥感图像，网络计算负担过重。通过对属于同一类别的像素光谱进行统计分析，可以发现光谱响应的方差非常小。考虑到提取光谱向量时保留了像素点的全部光谱信息，因此使用 R-PCA 降维方式对高光谱进行降维，再提取像素点邻域信息。在实验过程中，使用前 10 光谱维度可以至少保留 99.9% 的初始信息。这种方式可以使原始数据输入维度减少 10-15 倍。

4.3 空间信息

4.3.1 邻域空间信息提取

采用矩形窗 ($S \times S \times Cr$) 对当前像素点的邻域进行提取，其中 S 是矩形窗的长宽， Cr 是经过降维后的光谱维度数量。

当参数 S 为大于 5 时，分类准确性没有进一步提高；当 S 的值超过 13，分类精度下降^[57]。因此，考虑到计算时间，实验中设置 S 值为 5。

4.3.2 卷积神经网络提取特征

卷积神经网络（CNN）用于提取空间信息特征，具体结构如图 4.2。

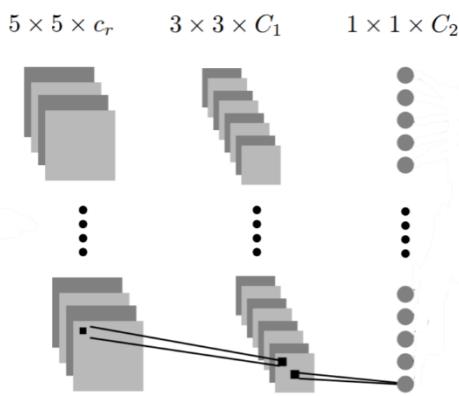


图 4.2 提取像素点空间信息网络架构

其中，第一层是深度为 $C1$ ($3 \times cr$)，尺寸为 3×3 的卷积层，第二层是深度为 $C2$ ($3 \times C1$)，尺寸为 3×3 的卷积层。因为这里的像素点邻域信息是整体遥感图片的部分，没有考虑任何的平移和尺度不变性，所以两层卷积层之间不设置最大值池化层，而采用两层卷积层直接连接的结构。空间数据经过模型训练后形成特性向量 $\mathbf{q}_{x,y}$ 。

4.4 空间-光谱信息融合分类

对于像素点 (x, y) ，将所提取光谱向量 $\mathbf{p}_{x,y}$ 和经过卷积网络训练所得到空间信息特征向量 $\mathbf{q}_{x,y}$ 铺平后融合，形成一个由光谱和空间组成的混合特征向量 $\mathbf{w}_{x,y}$ ，将混合特征向量作为 MLP 分类器输入向量（如图 4.1），标签为 $l_{x,y}$ 。模型根据数据集 $D = \{(\mathbf{w}_{x,y}, l_{x,y})\}$ 进行训练。

4.5 实验与分析

4.5.1 实验数据

4.5.1.1 Pavia 数据集

Pavia 数据集是由 ROSIS 传感器在意大利北部帕维亚飞行中所获得的，包括 Pavia Center 和 Pavia University 两个场景（分别如图 4.3、图 4.4 所示），几何分辨率是 1.3 米。Pavia Center 由 1096×1096 像素和 102 个光谱带组成，Pavia University 由 610×610 像素和 103 个光谱带组成。两幅图像中均包含 9 类地物，类别信息及样本数量如表 4.1 和表 4.2。



图 4.3 Pavia Center 数据集



图 4.4 Pavia University 数据集

表 4.1 Pavia Center 场景类别及样本数量

序号	类别	数量
0	背景	635488
1	水	65971
2	树	7598
3	柏油	3090
4	自阻塞砖	2685
5	沥青	6584
6	瓷砖	9248
7	阴影	7287
8	草地	42826
9	裸地	2863

表 4.2 Pavia University 场景类别及样本数量

序号	类别	数量
0	背景	164624
1	柏油	6631
2	草地	18649
3	碎石	2099
4	树	3064
5	彩绘金属板	1345
6	裸地	5029
7	沥青	1330
8	自阻塞砖	3682
9	阴影	947

4.5.1.2 Indian Pines 数据集

该场景由 AVIRIS 传感器在印第安纳州西北部的印度松树测试场地上采集，由 145×145 像素和波长范围为 $0.4\text{-}2.5\mu\text{m}$ 的 224 个光谱反射带组成（如图 4.5）。Indian Pines 数据集中三分之二的部分是农业用地，三分之一的部分是森林或其他天然植

被。除此之外，Indian Pines 数据集还包括两条双车道高速公路，一条铁路线，以及一些低密度住房和小道。地面实况包括 16 个类别，具体信息如表 4.3。

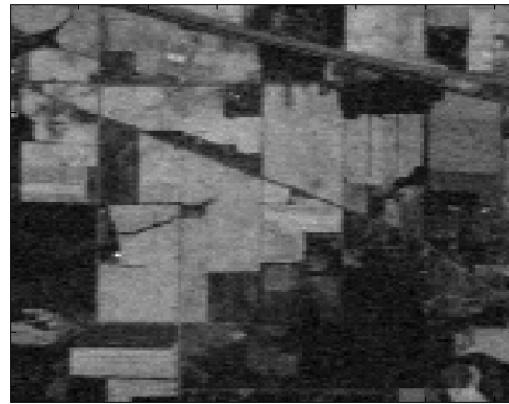


图 4.5 Indian Pines 数据集

表 4.3 Indian Pines 场景类别及样本数量

序号	类别	数量
0	Background	10776
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

4.5.2 实验模型

以 Pavia Center 数据集为例，实验总体模型如图 4.6。输入包括 $5 \times 5 \times 10$ 的空间信息和 102 维的光谱信息，输出为 10 类（9 类对象+1 类背景）。设置 MLP 中间层节点个数为 100，并添加 Dropout 层防止过拟合。训练集和测试集比例为 4:1。

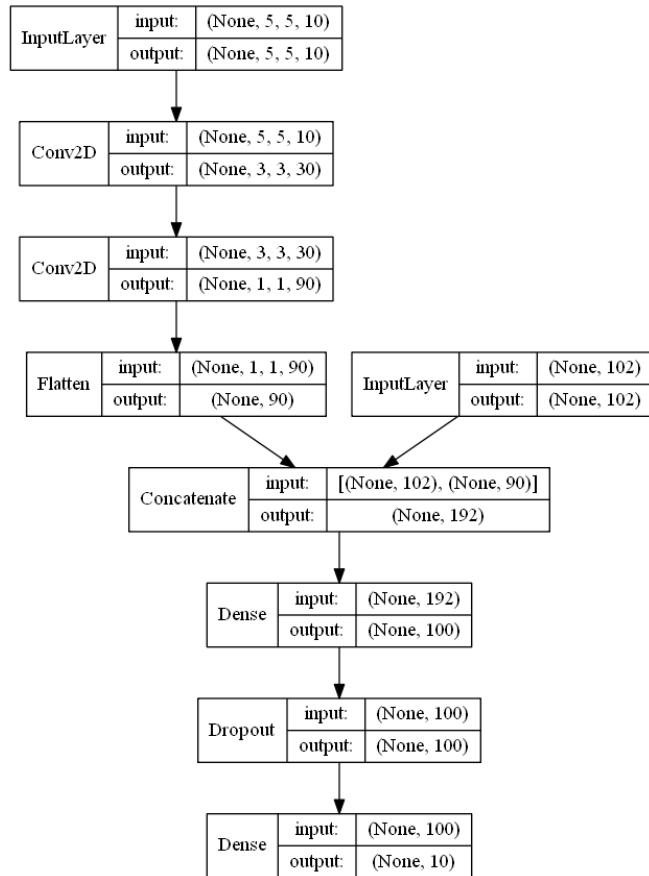


图 4.6 实验总体实验模型

4.5.3 实验结果与分析

分类结果如表 4.4 所示。可以看出，各数据测试集准确率均达到 94% 以上，分类效果良好。图 4.7—图 4.8 分别显示了基于 Pavia Center、Pavia University 和 Indian Pines 数据集的准确率与损失值。

表 4.4 空间-光谱信息混合模型分类结果

数据集	主成分	准确率 (%)
Pavia Center	10 (99.9%)	94.6 (95.2)
Pavia University	10 (99.9%)	96.1 (97.9)
Indian Pines	10 (96.9%)	94.0 (99.1)

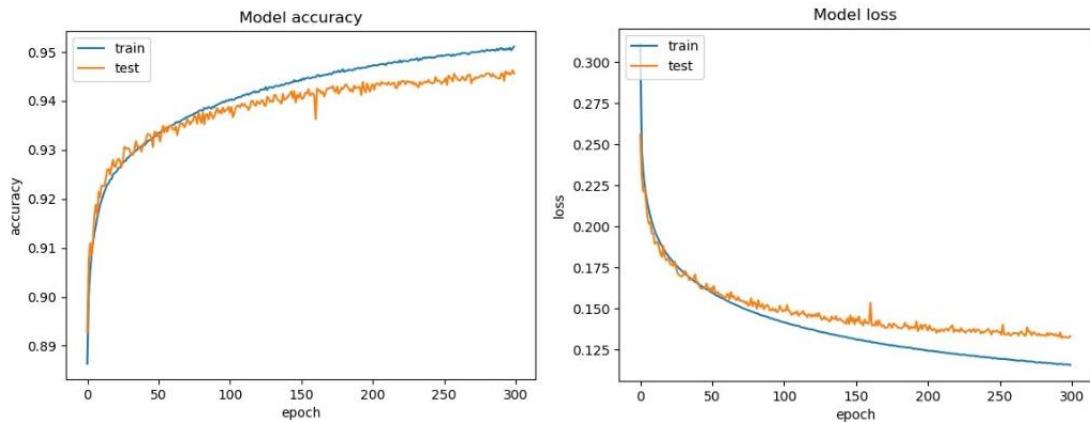


图 4.7 基于 Pavia Center 数据准确率 (accuracy) 与损失值 (loss)

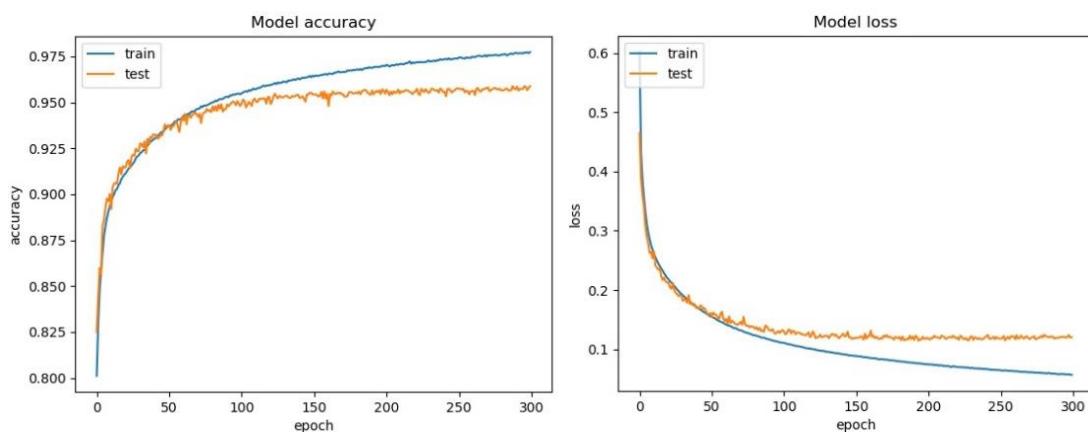


图 4.8 基于 Pavia University 数据准确率 (accuracy) 与损失值 (loss)

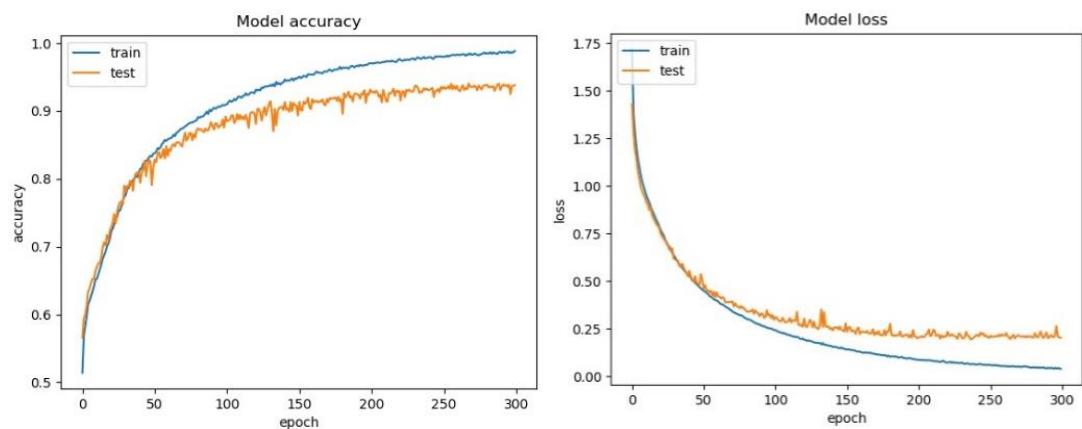


图 4.9 基于 Indian Pines 数据准确率 (accuracy) 与损失值 (loss)

图 4.10—图 4.12 分别显示了在 Pavia Center、Pavia University 和 Indian Pines 数据集的分类结果与测试数据。从图中可以看出，空间光谱融合分类模型可以很好地对地物对象进行分类，即保留了地物对象的形状和细节（如图 4.10 中的自阻塞砖，图 4.11 中的彩绘金属板均取得良好的分类效果），同时减少散乱的噪点。但

与 Pavia 数据集相比, Indian Pines 数据集分类误差较高, 可能由于空间信息所选邻域较小, 导致大块地块的某些细节丢失。

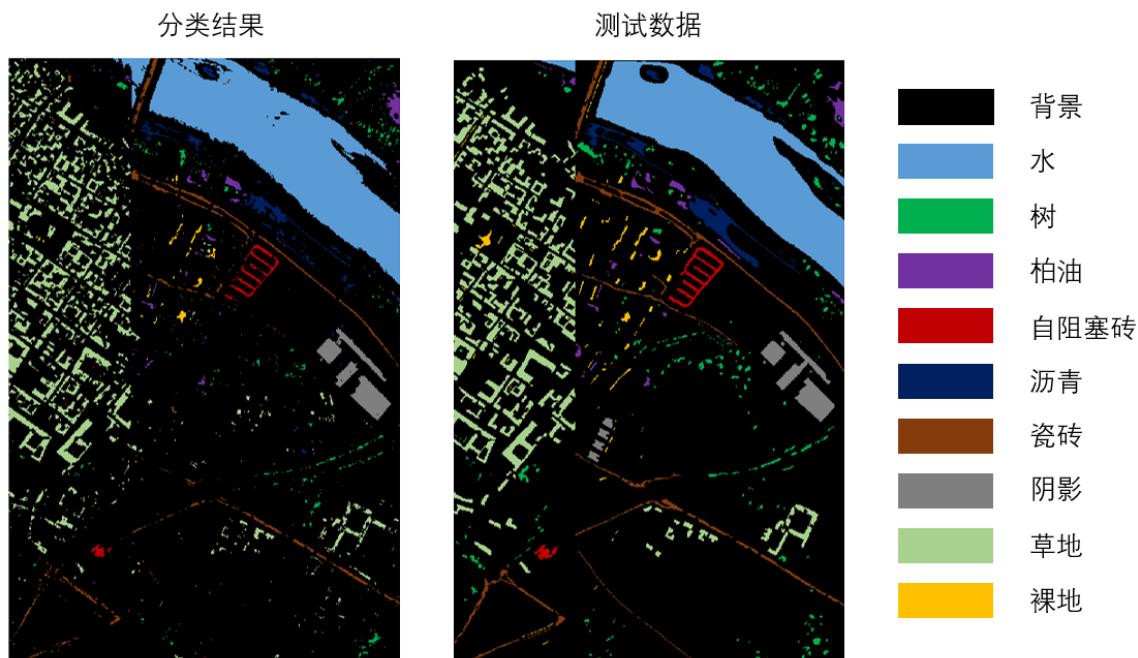


图 4.10 Pavia Center 数据集分类结果 (左) 与测试数据 (右)

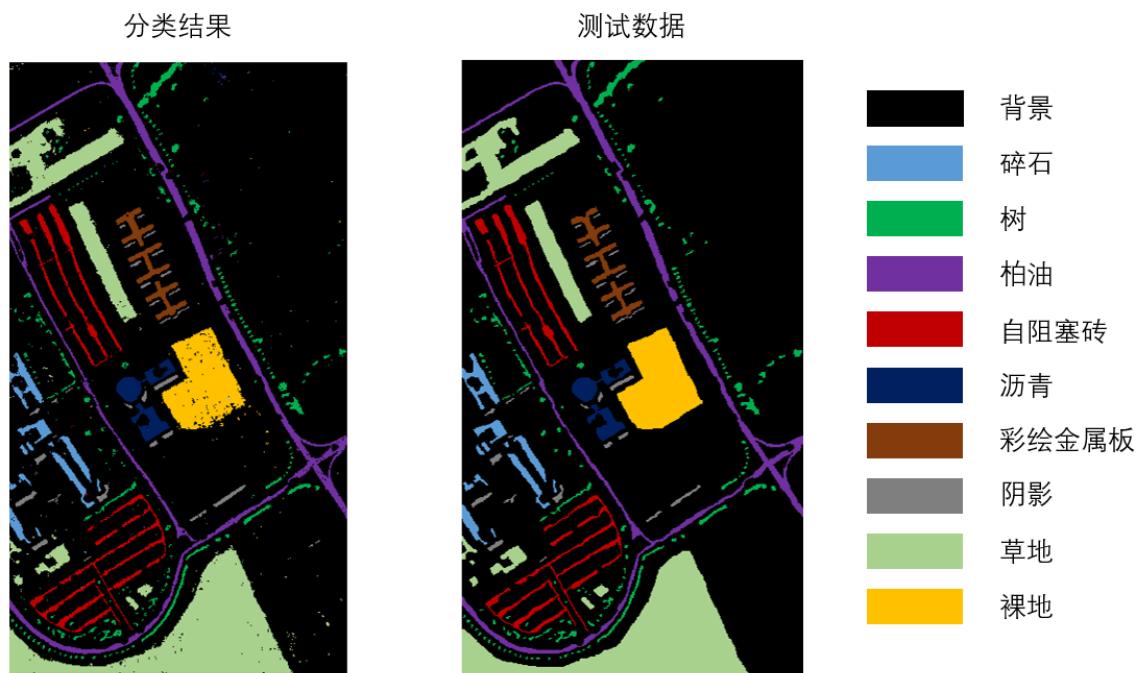


图 4.11 Pavia University 数据集分类结果 (左) 与测试数据 (右)

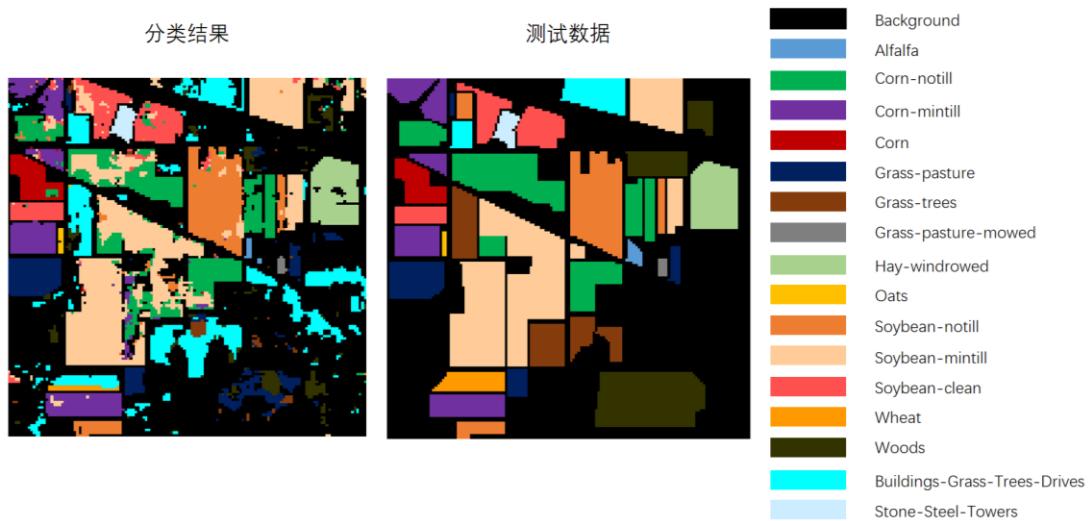


图 4.12 Indian Pines 数据集分类结果（左）与测试数据（右）

4.6 本章小结

面向高光谱遥感影像分类任务，本章给出一种空间-光谱信息融合分类方案。首先，介绍了模型整体架构；然后，分别详细说明了像素点光谱信息、空间信息的提取步骤与光谱-空间信息融合方式；最后，采用三个不同区域的数据集进行了实验，并对实验结果进行了分析。

5 大尺度高空间分辨率遥感影像分类

5.1 基于影像瓦片的场景分类

本节主要介绍一种基于影像瓦片的场景分类方法，对大尺度高空间分辨率遥感影像数据进行分类。

高空间分辨率遥感影像具有地物几何与属性细节信息丰富、目视效果直观等特点^[58]。直接获得的高空间分辨率遥感影像通常是大尺度的，而大尺度影像往往因为数据量大而难以直接处理。影像瓦片则可将大尺度遥感影像分为多个局部场景，从而便于影像分析和解译。本节将整张高空间分辨率遥感影像以（299, 299）正方形窗口进行分片，形成影像瓦片。分片后的影像瓦片作为单个场景，输入到后续预训练网络进行场景分类。

预训练网络采用 GoogleNet V4+自定义 CNN 架构。其中，GoogleNet V4 采用基于 ImageNet 数据集所训练的权重，自定义 CNN 采用基于 UCML 数据集所训练的权重。网络输出为 UCML 数据集的场景标签（21 种），包括：农田、机场、棒球场、沙滩、建筑、灌木丛、稠密住宅、森林、高速公路、高尔夫场、港口、交通枢纽、中等住宅、活动房屋、天桥、停车场、河流、机场跑道、稀疏住宅、储油罐、网球场。

5.2 基于人工判别的影像解译

本节采用面向对象分析技术中的多尺度分割算法（eCognition®软件）对大尺度高空间分辨率遥感影像进行影像分割，进而对影像类别进行人工判别。

5.2.1 遥感影像面向对象分析技术

遥感影像面向对象分析技术，又称基于地理对象的影像分析技术（Geographic Object-Based Image Analysis, GEOBIA）是一种新型的影像分析技术。与传统影像处理技术相比，该技术主要有以下优势^[59]：

- 特征提取时只考虑边界内的像素，排除了图斑外像素对分类结果的干扰，有效降低影像误判率，同时提升了分类速度。
- 以图斑为统计单位，可以获得更丰富的地类描述特征（如，对象的灰度直方图，类间信息）。
- 以同质对象作为分析单元，减少了像元间光谱异质性，使得影像的纹理、

结构等信息得到更加充分的利用。

- 能有效地解决基于像素分类方法的椒盐现象。
- 有利于多源数据的融合，引导 GIS 和 RS 的集成。

本节中主要使用 eCognition®软件对大尺度高空间分辨率遥感影像进行面向对象分析。eCognition®是一款常用的面向对象遥感信息提取商用软件，由德国 Definiens Imaging 公司开发。

5.2.2 遥感影像分割与多尺度分割算法

影像分割技术是遥感影像面向对象分析中的一项主要技术。常用的遥感影像分割算法^[60]如图 5.1。

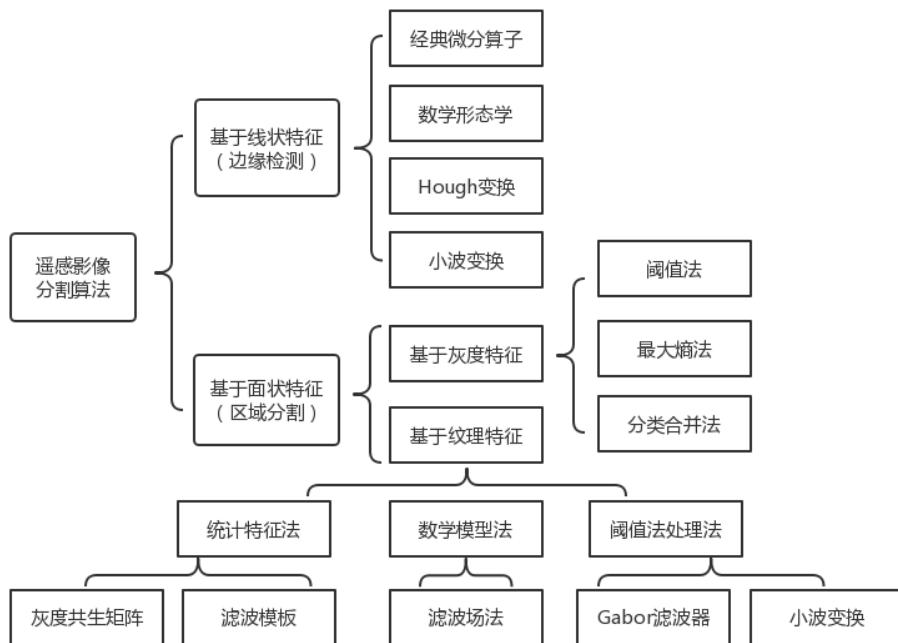


图 5.1 遥感影像常用分割算法分析框图

多尺度分割算法（Multiresolution Segmentation）是 eCognition®软件中的核心算法。该算法将遥感影像分割成一系列的影像对象，提取影像对象的特征，并以影像对象为最小处理单元进行分类，从而实现基于影像对象单元的分类任务^[58]。

多尺度分割算法的参数包括层权重（Layer Weight, LW）、尺度（Scale, Sc）、形状（Shape, Sh）、紧凑度（Compactness, Co）。其中，Scale 值决定了所分割结果中对象的规模，Scale 值越大，所得到的对象尺度也越大；Shape 值决定了影像形状和颜色对于分割结果的影响，Shape 值越大，则影像形状对于分割结果的影响越大，影像颜色对于分割结果的影响越小；Compactness 值定义紧凑度标准的权重，Compactness 值越大，则分割结果更紧凑，反之则分割结果更光滑。图 5.2 显示了

eCogniton®软件多尺度分割算法界面。

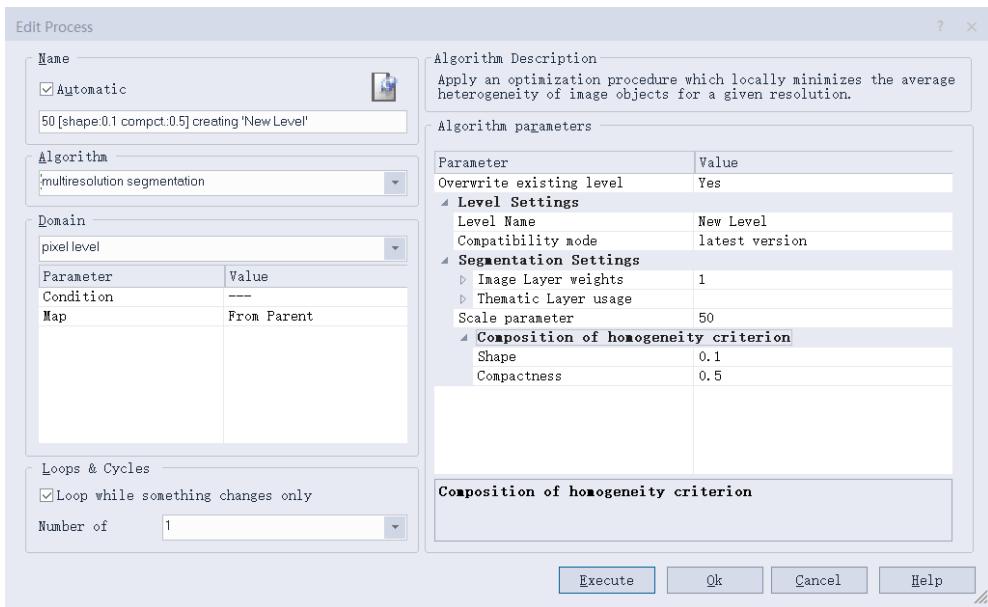


图 5.2 eCogniton®软件的多尺度分割算法界面

5.3 实验与分析

5.3.1 实验数据

实验采用 GF2 卫星 PMS1 传感器在 2015 年 1 月 23 日获取广州上空遥感影像（如图 5.3）进行测试。影像大小为 29200×27619 像素，空间分辨率为 4 米。

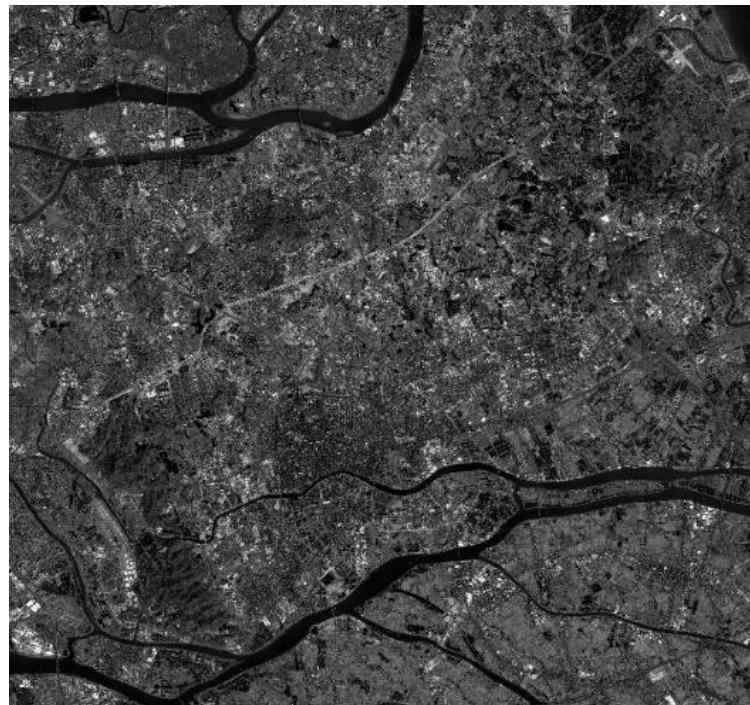


图 5.3 广州上空遥感影像数据

5.3.2 实验结果与分析

5.3.2.1 基于影像瓦片的场景分类

表 5.1 显示了基于影像瓦片的场景分类结果类别个数。测试大尺度高分辨率遥感影像被分成 8736 个正方形瓦片 (299×299 , 忽略部分边缘信息), 分类结果中河流场景最多 (4120 个), 机场、灌木丛、高尔夫场、稠密住宅、交通枢纽、中等住宅、活动房屋、天桥、停车场、机场跑道场景最少 (0 个)。

表 5.1 基于影像瓦片的场景分类结果类别个数

类别	个数
农田	1688
机场	0
棒球场	79
沙滩	148
建筑	12
灌木丛	0
稠密住宅	0
森林	1
高速公路	178
高尔夫场	0
港口	536
交通枢纽	0
中等住宅	0
活动房屋	0
天桥	0
停车场	0
河流	4120
机场跑道	0
稀疏住宅	3
储油罐	1931
网球场	40
总数	8736

图 5.4 显示了基于影像瓦片的场景分类的部分结果。从图中可以看出，大多场景均取得正确的分类结果，但依然存在少数场景（如储油罐）得到错误的分类结果。由于该方法依赖于预训练模型的类别，因此涉及到预训练模型中未包含的场景，模型往往无法得到正确的分类结果。同时瓦片大小的设置也影响分类结果，对于如储油罐、沙滩等相对较小的对象，大尺度的瓦片往往使得场景包含过多冗余信息，导致分类结果不正确。并且，场景的重叠问题也有待解决。

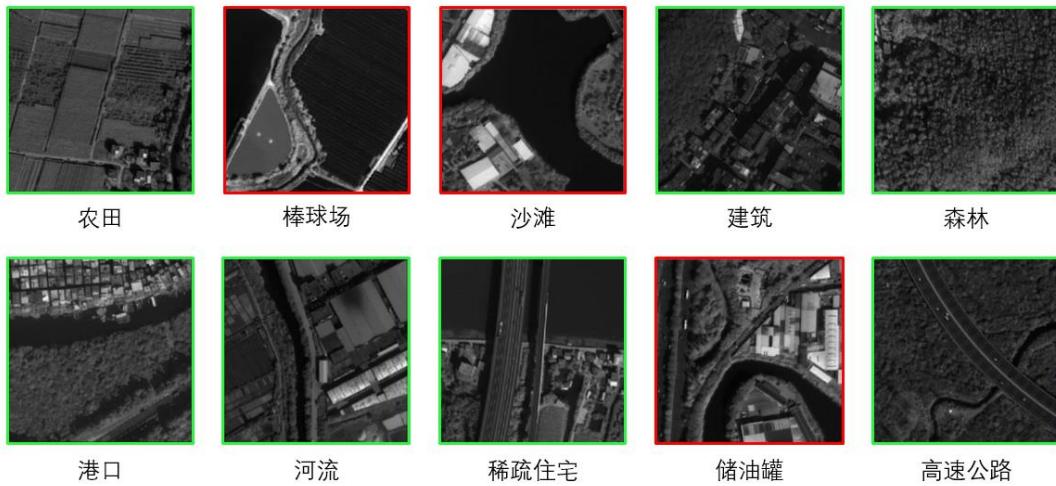


图 5.4 基于影像瓦片的场景分类结果类别示例（绿色为正确，红色为错误）

5.3.2.2 基于人工判别的影像解译

本实验中采用 $LW=1$, $Sc=80$, $Sh=0.8$, $Co=0.5$ 的参数设置。

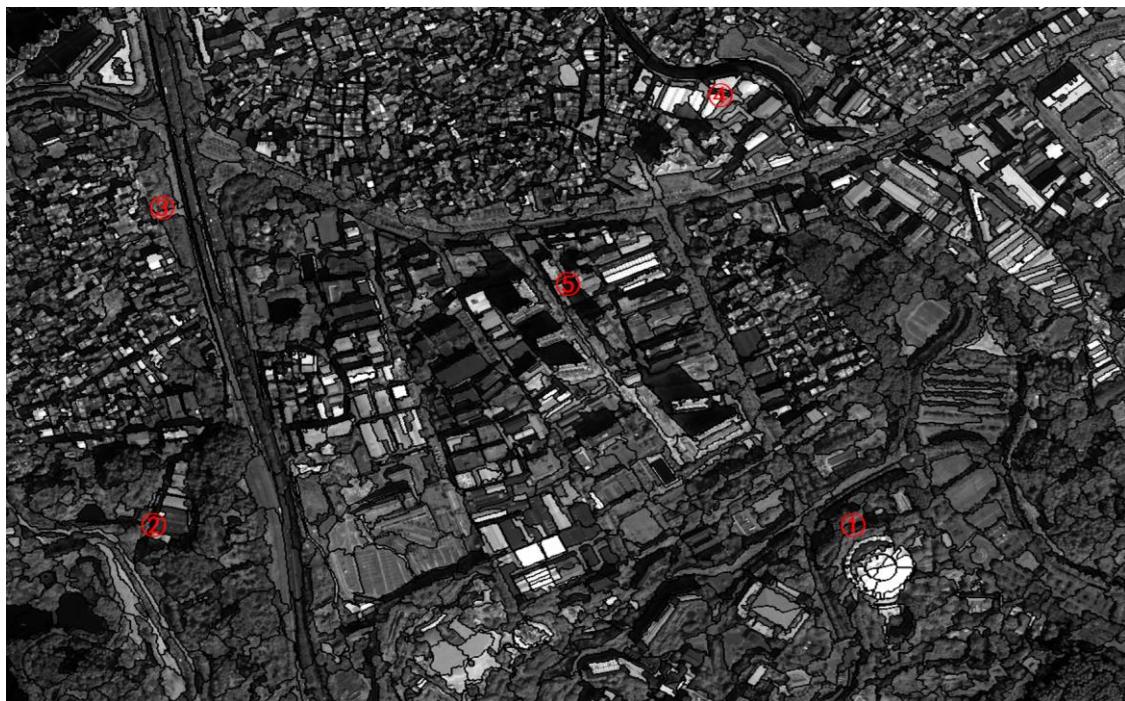


图 5.5 多尺度分割算法处理结果

图 5.5 显示了数据影像经过 eCognition® 软件的多尺度分割算法处理后结果示意图，这里分别挑选图中①—⑤号场景对象进行分析。放大图中①—⑤处影像场景如图 5.6 所示。

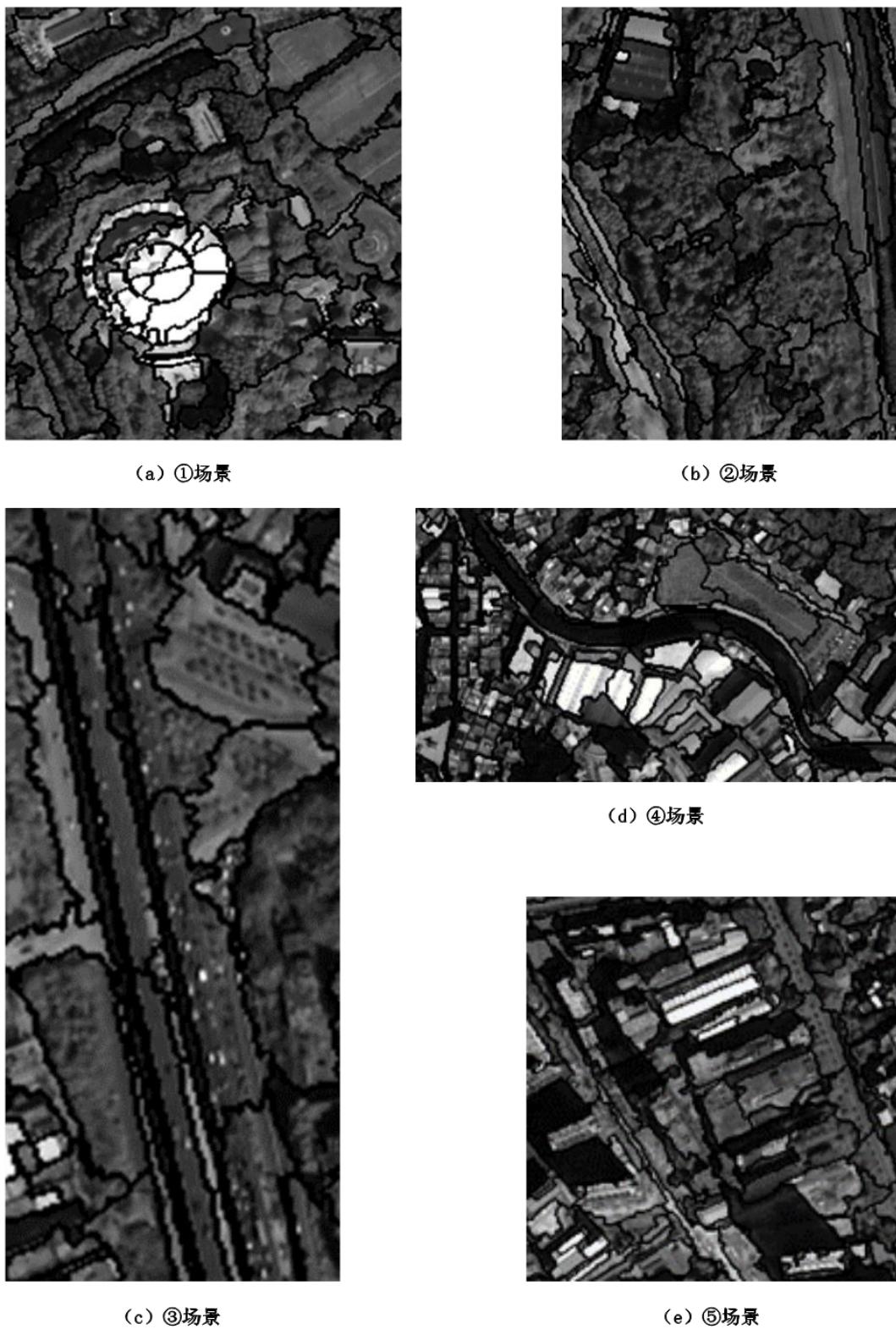


图 5.6 ①—⑤处影像场景放大示意图

对分割后的遥感影像进行人工影像判别，我们可以得出以下结果：①场景为

储油罐，②场景为森林，③场景为公路，④场景为河流，⑤场景为住宅。

该方法这使得场景规模不受窗口大小限制，场景标签也不局限于训练样本的类别。但由于场景类别确定受人为因素影响较大，参数选择对于分割结果影响较大，因此处理大规模数据时受到一定限制。

5.4 本章小结

面向大尺度高空间分辨率遥感影像，本章给出了基于影像瓦片的场景分类方案，并利用 GF2 卫星 PMS1 传感器在 2015 年 1 月 23 日获取广州上空遥感影像对该方案进行测试。同时，比较了该方法与基于人工判别的影像解译方法的优劣。

在处理实际问题大规模高分辨率遥感影像过程中，可将基于影像瓦片的场景分类方法和基于人工判别的影像解译方法结合使用，以场景分类结果作为遥感对象分析的重要参考，从而得到更准确的遥感对象分类，并且提高分类速度。

6 总结与展望

本文主要介绍了深度学习在遥感影像分类任务中的应用。具体介绍了神经网络算法和深度神经网络模型，并针对三种不同的应用场景分别使用深度学习方法进行影像分类处理。

面向遥感影像场景分类任务，本文分别采用基于神经网络的分类方法和基于微训练模型的分类方法。实验结果表明，深度神经网络的分类准确率比传统方法的分类准确率更高，基于微训练模型的分类方法可大幅提高深度神经网络的分类准确率。

面向高光谱遥感影像分类任务，本文采用了空间-光谱信息融合分类方法。实验结果表明，该方法可以很好地保持影像地物对象的形状和细节信息，分类效果良好。

面向大尺度高空间分辨率遥感影像，本文分别采用基于影像瓦片的场景分类方法和基于人工判别的影像解译方法。从实验结果可以看出，基于影像瓦片的场景分类方法在分类速度上更有优势，而基于人工判别的影像解译分类方法则有更高的分类准确率。实际问题中，可将两种方法结合使用。

针对遥感影像分类亟待解决的问题和未来前景，有以下设想：

- (1) 针对地表真实数据（Ground Truth）缺乏的问题，拟采用无监督学习对遥感影像进行分类；
- (2) 在遥感影像分类中，拟添加地理先验信息作为辅助提高分类精度；
- (3) 在高光谱遥感影像分类中，拟研究光谱信息与空间信息的权重自动确定方法；
- (4) 在大尺度高空间分辨率遥感影像分类中，拟采用目标检测算法，检测部分目标及所在区域，实现目标与场景类别的自动解译。

参考文献

- [1] 王贝贝, 崔巧云, 张栋. 探析测绘工作中遥感技术的应用[J]. 科技创新, 2014 年 12 月下:1005.
- [2] 孔祥慧. 基于知识的遥感信息分类方法实验研究[D]. 昆明: 昆明理工大学, 2009.
- [3] 刘大伟, 韩玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. 光学报, 2016, 36(4).
- [4] 高常鑫, 桑农. 基于深度学习的高分辨率遥感影像目标检测[J]. 测绘通报, 2014 (S1): 108-111.
- [5] 钱茹茹. 遥感影像分类方法比较研究[D]. 西安: 长安大学, 2007.
- [6] 深度学习的优越性 [EB/OL] , 雷锋网 , <https://www.leiphone.com/news/201702/Zt2HTvL4psFsDoWh.html>, 2017
- [7] 孙志远, 鲁成祥, 史忠植等. 深度学习研究与进展[J]. 计算机科学, 2016 (2016 年 02): 1-8.
- [8] 王巧玉. 基于深度学习的高光谱遥感图像分类[D]. 泉州: 华侨大学, 2016.
- [9] Lin Z, Chen Y, Zhao X, et al. Spectral-spatial classification of hyperspectral image using autoencoders[C]. In: Information, Communications and Signal Processing (ICICS) 2013 9th International Conference on. IEEE, 2013: 1-5.
- [10] 李晓旭. 基于概率主题模型的图像分类和标注的研究[D]. 北京: 北京邮电大学, 2012.
- [11] M. Lienou, H. Maitre, and M. Datcu. Semantic annotation of satellite images using latent dirichlet allocation[J]. IEEE Geosci. Remote Sens. Lett., vol. 7, no. 1, pp. 28–32, 2010.
- [12] O. A. Penatti, K. Nogueira, and J. A. dos Santos, “Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?” in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 44–51.
- [13] Hu W, Huang Y, Wei L, et al. Deep convolutional neural networks for hyperspectral image classification[J]. Journal of Sensors, 2015.
- [14] 袁林, 胡少兴, 张爱武, 柴沙陀, 王兴. 基于深度学习的高光谱图像分类方法[J]. 人工智能与机器人研究, 2017, 6(1): 31-39.

- [15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278 – 2324, 1998.
- [16] LeNet-5. convolutional neural networks[EB/OL].
<http://yann.lecun.com/exdb/lenet/>
- [17] Ioffe S., Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift [J]. *Journal of Machine Learning Research*, 2015, 37 448-456.
- [18] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks[J]. *Advances in neural information processing systems*, 2012.
- [19] Simonyan, Karen, and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition[DB/OL]. arXiv preprint arXiv:1409.1556 (2014).
<https://arxiv.org/pdf/1409.1556/>
- [20] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich. Going Deeper with Convolutions[DB/OL]. <https://arxiv.org/abs/1409.4842>
- [21] Sergey Ioffe, Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[DB/OL].
<https://arxiv.org/abs/1502.03167>
- [22] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna. Rethinking the Inception Architecture for Computer Vision[DB/OL].
<https://arxiv.org/abs/1512.00567>
- [23] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[DB/OL].
<https://arxiv.org/abs/1602.07261>
- [24] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition[DB/OL]. <https://arxiv.org/pdf/1512.03385.pdf>
- [25] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger. Densely Connected Convolutional Networks[DB/OL]. <https://arxiv.org/abs/1608.06993>
- [26] Sara Sabour, Nicholas Frosst, Geoffrey E. Hinton. Dynamic Routing Between

- Capsules. <https://arxiv.org/pdf/1710.09829v1.pdf>
- [27] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, 2005, 1: 886-893.
- [28] 李航. 统计学习方法[M]. 北京: 清华大学出版, 2012-3.
- [29] Pang-Ning Tan, Michael Steinbach, Vipin Kumar. Introduction to Data Mining[EB/OL]. <https://www-users.cs.umn.edu/~kumar001/dmbook/index.php>
- [30] Y. Yang and S. Newsam. Bag-of-visual-words and spatial extensions for land-use classification[R]. Proc. of the 18th Sigspatial International Conference on Advances in Geographic Information Systems, 2010: 270 – 279.
- [31] Csurka, G., C. R. Dance, L. Fan, J. Willamowski, C. Bray. Visual Categorization with Bags of Keypoints[R]. Workshop on Statistical Learning in Computer Vision. ECCV 1 (1-22), 1-2.
- [32] Mathwork. Image Classification with Bag of Visual Words[EB/OL]. <http://cn.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html>
- [33] Bag of Words model 应用于图像 [EB/OL]. <https://blog.csdn.net/happyer88/article/details/45769413>
- [34] Pang-Ning Tan, Michael Steinbach, Vipin Kumar. Introduction to Data Mining[EB/OL]. <https://www-users.cs.umn.edu/~kumar001/dmbook/index.php>
- [35] Xindong Wu, Vipin Kumar. The Top Ten Algorithms in Data Mining[J]. Knowledge and Information Systems, January 2008, Volume 14, Issue 1: 1-37
- [36] David G. Lowe. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 60, 2 (2004): 91-110.
- [37] 吕启, 窦勇, 牛新, 徐佳庆, 夏飞. 基于 DBN 模型的遥感图像分类[J]. 计算机研究与发展, 2014, 51(9): 1911-1918.
- [38] Hinton, G. E., Osindero, S. and Teh, Y. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18: 1527-1554.
- [39] Hinton, G. E. and Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks[J]. Science, 28 July 2006, Vol. 313. no. 5786: 504 - 507.

- [40] Smolensky P. Information processing in dynamical systems: Foundations of harmony theory[M]. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1986: 194-281.
- [41] Freund Y, Haussler n Unsupervised learning of distributions of binary vectors using two layer networks[R]. Santa Cruz: Computer Research Laboratory, University of Callfornia, 1994.
- [42] Hinton G E Training products of experts by minimizing contrastive divergence[J]. Neural Computation, 2002, 14(8): 1771—1800.
- [43] J. Donahue. Decaf: A deep convolutional activation feature for generic visual recognition[DB/OL]. Unpublished paper, 2013. <http://arxiv.org/abs/1310.1531>.
- [44] Dimitrios Marmanis, Mihai Datcu, Thomas Esch, and Uwe Stilla. Deep Learing Earth Observation Classification Using ImageNet Pretrained Networks[J]. IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 1, January, 2016:105-109.
- [45] Maaten, L., & Hinton, G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008.
- [46] 刘代志, 黄世奇, 王艺婷, 刘志刚, 王百合. 高光谱图像处理与应用[M]. 第一版. 北京: 科学出版社, 2016.
- [47] D. Landgrebe. Hyperspectral image data analysis[J]. IEEE Signal Process. Mag., vol. 19, no. 1, Jan. 2002:17-28.
- [48] J. A. Richards, Remote Sensing Digital Image Analysis: An Introduction. New York, NY, USA: Springer, 2013.
- [49] F. M. Lacar, M. M. Lewis, and I. T. Grierson. Use of hyperspectral imagery for mapping grape varieties in the Barossa Valley, South Australia[J]. Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Sydney, Australia, 2001, vol. 6: 2875 – 2877.
- [50] F. V. D. Meer. Analysis of spectral absorption features in hyperspectral imagery[J]. Int. J. Appl. Earth Observ. Geoinf., vol. 5, no. 1, Jan. 2004: 55 – 68.
- [51] P. W. Yuen and M. Richardson. An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition[J]. Imaging Sci. J., vol. 58, no. 5, May 2010: 241 – 253.
- [52] R. F. Egerton, Electron Energy-Loss Spectroscopy in the Electron Microscope[R].

New York, NY, USA: Plenum, 1996.

- [53] E. K. Hege. Hyperspectral imaging for astronomy and space surveillance[C]. Proc. SPIE' s 48th Annu. Meet. Opt. Sci. Technol., San Diego, CA, USA, 2004: 380 – 391.
- [54] A. A. Gowen. Hyperspectral imaging—An emerging process analytical tool for food quality and safety control[J]. Trends Food Sci. Technol., vol. 18, no. 12, Dec. 2007: 590 – 598.
- [55] G. Camps-Valls and L. Bruzzone. Kernel-based methods for hyperspectral image classification[J]. IEEE Trans. Geosci. Remote Sens., vol. 43, no. 6, Jun. 2005: 1351-1362.
- [56] Yushi Chen, Zhouhan Lin, Xing Zhao, Gang Wang, Yanfeng Gu. Deep Learning-Based Classification Of Hyperspectral Data[J]. IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 7, No. 6, June 2014.
- [57] Konstantinos Makantasis, Konstantinos Karantzalos, Anastasios Doulamis, Nikolaos Doulamis. Deep Supervised Learning For Hyperspectral Data Classification Through Convolutional Neural Networks[DB/OL].
http://www.evacuate.eu/media/24602/makantasis_igarss_camera_ready.pdf
- [58] 刘建华, 杜明义. 高空间分辨率遥感影像地学计算[M]. 第一版. 北京: 科学出版社, 2017.
- [59] 顾海燕、李海涛、闫利. 高分辨率遥感影像面向对象分析技术[M], 北京: 科学出版社, 2016.
- [60] 明东萍, 骆剑承, 周成虎. 高分辨率遥感影像特征分割及算法评价分析[J]. 地球信息科学, 2006, 8(1): 103-109.

致谢

感谢卢其楷老师在我论文撰写过程中的细心指导与协助。在论文完成期间，卢老师始终鼓励我进行学术创新和积极尝试，每次遇到问题时，卢老师也总是很细心的给予我指导，并毫无保留的给我传授经验。

感谢电子信息学院和大学间教育培养我的各位老师，是你们的辛勤培育让我逐渐了解并学习了广博的知识，同时发现了自我。

感谢卓工班与电信四班一起奋斗了四年的同学们，是优秀的你们激励着我不断向你们的方向迈进！

感谢我的家人，他们永远是我坚强的后盾，是他们的支持与鼓励成就了今天的我。

感谢不分昼夜工作的我的电脑的显卡与 CPU，是你们每天深夜的微光为本文照亮了光明的方向。

感谢武大美丽的校园，感谢三月的珞樱，六月的桂香，十月的枫叶，腊月的白雪。感谢武大给予我精彩的四年，那是一段无与伦比的青春。

本文中，参考引用了一些其他学者的研究成果，对此表示感谢！

时光荏苒，我们即将离开校园，向着未知的前方前进。回望武汉大学的四年，我学到最多的也许不是具体某一方面的知识，而是一种精神，一种理性思考的精神，一种严谨的实证精神；也是一种能力，一种自我决定能力，一种自我辨别能力；还是一种品质，是自强、弘毅、求是、拓新。

风荡千度，经年难堪不数。只愿归来时，身处沼泽，依然故我。