实训 5-3 遥感图像监督分类

1、实训目的

巩固所学的监督分类的基本原理和方法。
熟悉 Envi 软件进行监督分类的操作流程。
掌握利用监督分类进行土地类型类型图生成的方法。

2、实训内容

使用 Envi 软件进行监督分类。

任选一个广东省的县级行政区,利用监督分类生成该区域的土地利用类型图。

3、实训步骤



遥感影像的监督分类一般包括以下 6个步骤,如下图所示:

2.1 类别定义/特征判别

根据分类目的、影像数据自身的特征和分类区收集的信息确定分类系统;对影像进行特征判断,评价图像质量,决定是否需要进行影像增强等预处理。这个过程主要是一个目视查看的过程,为后面样本的选择打下基础。

启动 ENVI5.3, 打开待分类数据: can_tmr.img。以 R:TM Band 5, G: TM Band 4, B: TM Band 3 波段组合显示。

通过目视可分辨六类地物:林地、耕地、裸地、居民地、道路、河流、湖泊坑塘、其他 7 类。

2.2 样本选择

(1)在图层管理器 Layer Manager 中, can_tmr.img 图层上右键,选择"New Region Of Interest",打开 Region of Interest (ROI) Tool 面板,下面学习利用选择样本。

1) 在 Region of Interest (ROI) Tool 面板上,设置以下参数:

🕸 🕸 ú	🛊 🎯			
OI Name:	林地			
Geometry	Pixel	Grow	Threshold	

2)默认 ROIs 绘制类型为多边形,在影像上辨别林地区域并单击鼠标左键开始绘制多边 形样本,一个多边形绘制结束后,双击鼠标左键或者点击鼠标右键,选择 Complete and Accept Polygon,完成一个多边形样本的选择;



3)同样方法,在图像别的区域绘制其他样本,样本尽量均匀分布在整个图像上;

4) 这样就为林地选好了训练样本。

注:1、如果要对某个样本进行编辑,可将鼠标移到样本上点击右键,选择 Edit record 是修改样本,点击 Delete record 是删除样本。



2、一个样本 ROI 里面可以包含 n 个多边形或者其他形状的记录(record)。

3、如果不小心关闭了 Region of Interest (ROI) Tool 面板,可在图层管理器 Layer Manager 上的某一类样本(感兴趣区)双击鼠标。



(2)在图像上右键选择 New ROI,或者在 Region of Interest (ROI) Tool 面板上,选择

路、河流、湖泊坑塘、其他5类选择样本;

(3) 如下图为选好好的样本。



(4)计算样本的可分离性。在 Region of Interest (ROI) Tool 面板上,选择 Option>Compute ROI Separability,在 Choose ROIs 面板,将几类样本都打勾,点击 OK;

	ŧ
□□ <mark>例</mark> 裸地	
Junkov of items .	1+-4-7
wumber of items s	serected.

(5) 表示各个样本类型之间的可分离性,用 Jeffries-Matusita, Transformed Divergence 参数表示,这两个参数的值在 0~2.0 之间,大于 1.9 说明样本之间可分离性好,属于合格样本;小于 1.8,需要编辑样本或者重新选择样本;小于 1,考虑将两类样本合成一类样本。

```
- 0
                                                                                                                 23
ROI Separability Report
File
        耕地: (1.99863817 2.00000000)
                                                                                                                     ×
        裸地: (1.67222938 1.83445737)
裸地
        河流: (2.00000000 2.0000000)
林地: (1.99999947 2.00000000)
       ///1.11/1/1/2182 1.999999997)
湖泊坑塘: (2.00000000 2.0000000)
耕地: (2.00000000 2.0000000)
道路: (1.67222938 1.83445737)
Pair Separation (least to most);
道路 and 裸地 - 1.67222938
道路 and 裸地 - 1.67222938
居民地 and 道路 - 1.95926097
林地 and 耕地 - 1.97537540
林地 and 道路 - 1.99735213
耕地 and 道路 - 1.99863817
林地 and 湖泊坑塘 - 1.99911401
居民地 and 裸地 - 1.99912182
河流 and 湖泊坑塘 - 1.99946638
林地 and 湖泊坑塘 - 1.99979204
民地 and 湖泊坑塘 - 1.99979204
RE地 and 湖泊坑塘 – 1.99980599
湖泊坑塘 and 道路 – 1.99999903
林地 and 裸地 – 1.99999947
河流 and 道路 – 1.99999984
河流 and 居民地 - 2.00000000
耕地 and 裸地 - 2.00000000
                                                                                                                     Ξ
湖泊坑塘 and 裸地 - 2.00000000
湖泊坑塘 and 耕地 - 2.00000000
 ÷
```

注: 1、在图层管理器 Layer Manager 中,可以选择需要修改的训练样本。

2、在 Region of Interest (ROI) Tool 面板上,选择 Options > Merge (Union/Intersection) ROIs, 在 Merge ROIs 面板中,选择需要合并的类别,勾选 Delete Input ROIs。

刘 Merge ROIs	×
Select ROIs to Merge	
│ ────────────────────────────────────	
│ ──□ 22 居民地	
四月 建油	
Hard Contract Description of the second	
Number of items selected: 2	
Select All Items Liear All	Items
Merge Method: 💿 Union 🛛 🔘 In	tersection
V Delete Innut ROIs	
☑ Delete Input ROIs	

(6) 在图层管理器中,选择 Region of interest , 点击右键, save as, 保存为.xml 格式的样本文件。

C) R	egion of Interest (ROI) To	pol X	Select ROIs for Output Select ROIs for Output 「夏」「「」」 「夏」」 「夏」」 「」」 「」」 「」」 「」
File	Options Help		
	Open		┃
	Import Vector		
	Save		
	Save As	LEZUOLO	
0	Revert		
4	Export +	hap	Number of items selected: 7
	Close		Select All Items Clear All Items
H	4 1 + +i	×	Enter Output File [.xml]
			D:\训练区.xml
Are	a		OK Cancel

注: 1、早期版本的感兴趣文件格式为.roi,新版本的为.xml,新版本完全兼容.roi文件,在 Region of Interest (ROI) Tool 面板上,选择 File>Open 打开.xml 或.roi 文件。

2、新版本的.xml 样本文件(感兴趣区文件)可以通过,File>Export>Export to Classic 菜

单保存为.roi 文件。

2.3 分类器选择

根据分类的复杂度、精度需求等确定哪一种分类器。目前 ENVI 的监督分类可分为基于 传统统计分析学的,包括平行六面体、最小距离、马氏距离、最大似然,基于神经网络的, 基于模式识别,包括支持向量机、模糊分类等,针对高光谱有波谱角(SAM),光谱信息散 度,二进制编码。下面是几种分类器的简单描述。

(1) 平行六面体(Parallelepiped)

根据训练样本的亮度值形成一个n维的平行六面体数据空间,其他像元的光谱值如果落 在平行六面体任何一个训练样本所对应的区域,就被划分其对应的类别中。

(2)最小距离(Minimum Distance)

利用训练样本数据计算出每一类的均值向量和标准差向量,然后以均值向量作为该类在 特征空间中的中心位置,计算输入图像中每个像元到各类中心的距离,到哪一类中心的距离 最小,该像元就归入到哪一类。

(3) 马氏距离(Mahalanobis Distance)

计算输入图像到各训练样本的协方差距离(一种有效的计算两个未知样本集的相似度的 方法),最终技术协方差距离最小的,即为此类别。

(4)最大似然(Maximum Likelihood)

假设每一个波段的每一类统计都呈正态分布,计算给定像元属于某一训练样本的似然 度,像元最终被归并到似然度最大的一类当中。

(5) 神经网络(Neural Net)

指用计算机模拟人脑的结构,用许多小的处理单元模拟生物的神经元,用算法实现人脑 的识别、记忆、思考过程。

(6) 支持向量机(Support Vector Machine)

支持向量机分类(Support Vector Machine 或 SVM)是一种建立在统计学习理论(Statistical Learning Theory 或 SLT)基础上的机器学习方法。SVM 可以自动寻找那些对分类有较大区分能力的支持向量,由此构造出分类器,可以将类与类之间的间隔最大化,因而有较好的推广性和较高的分类准确率。

(7) 波谱角(Spectral Angle Mapper)

它是在 N 维空间将像元与参照波谱进行匹配,通过计算波谱间的相似度,之后对波谱 之间相似度进行角度的对比,较小的角度表示更大的相似度。

•••••

2.4 影像分类

基于传统统计分析的分类方法参数设置比较简单,在 Toolbox/Classification/Supervised Classification 能找到相应的分类方法。这里选择最大似然分类方法。选择待分类影像,点击 OK,按照默认设置参数输出分类结果。

244 S	Adaptive Coherence Estimator Classific:
200 () 200 () 200 ()	Ringerre Contraction Classification
200 () 200 ()	binary Encouring crassification
200	Constrained Energy Minimization Classi
	Mahalanobis Distance Classification
2.00 C	Maximum Likelihood Classification
2.00 Que 0 2.00 Que 0	Minimum Distance Classification
2.00 Qua 2.00 Qua 2.00 Qua	Neural Net Classification
244 C	Orthogonal Subspace Projection Classifi
7.00 C	Parallelepiped Classification
244 (244 (244 (Spectral Angle Mapper Classification
244 C	Spectral Information Divergence Classi
	Support Vector Machine Classification

Select Input File: data1.tif	File Information: File: E:\01学校\1上课\13遥感\test01\datal Dims: 3871 x 2458 x 7 [BSQ] Size: [Byte] 66,742,461 bytes. File Type : TIFF Sensor Type: Unknown Byte Order : Host (Intel)
	Projection : UTM, Zone 49 North Pixel : 30 Meters Datum : WGS-84 Wavelength : None Upper Left Corner: 1,1 Description: GEO-TIFF File Imported into ENVI [Wed Jun 14 21:46:10 2017]
Spatial Subset Full Scene Spectral Subset 7/7 Bands	

Select Classes from Regions: 可流 妹地 居氏地 明白坑塘 挂地 圓路	Output Result to File Memory Enter Output Class Filename Choose Compress D:\maxlikehood
Aumber of items selected: 7 Select All Items Clear All Items Set Probability Threshold None © Single Value © Multiple Values	Output Rule Images ? No
Probability Threshold 0.95 Data Scale Factor 255.00	



2.5 分类后处理

包括更改类别颜色、分类后统计、小斑块处理、栅矢转换等,这部分专门有一节课讲解。在此不做叙述。

2.6 精度验证

对分类结果进行评价,确定分类的精度和可靠性。有两种方式用于精度验证:一是混淆

矩阵,二是 ROC 曲线,比较常用的为混淆矩阵,ROC 曲线可以用图形的方式表达分类精度,比较抽象。

真实参考源可以使用两种方式:一是标准的分类图,二是选择的感兴趣区(验证样本区)。 两种方式的选择都可以通过主菜单->Classification->Post Classification->Confusion Matrix 或者 ROC Curves 来选择。

真实的感兴趣区验证样本的选择可以是在高分辨率影像上选择,也可以是野外实地调查 获取,原则是获取的类别参考源的真实性。由于没有更高分辨率的数据源,本例中就把原分 类的 TM 影像当作是高分辨率影像,在上面进行目视解译得到真实参考源。

(1) 在 Data Manager 中,分类样本上右键选择 Close,将分类样本从软件中移除

(2) 直接利用 ROI 工具,跟分类样本选择的方法一样,即重复第二步,在 TM 图上选择 6 类验证样本。

注:可直接 File>open,打开 can_tm-验证样本.roi。



(3)在 Toolbox 中,选择/Classification/Post Classification/Confusion Matrix Using Ground Truth ROIs,选择分类结果,软件会根据分类代码自动匹配,如不正确可以手动更改。点击 OK 后选择报表的表示方法(像素和百分比),点击 OK,就可以得到精度报表。

Select Input File:	File Information:
maxLikehood datal.tif	File: D:\maxlikehood Dims: 3871 x 2458 x 1 [BSQ] Size: [Byte] 9,514,918 bytes. File Type : ENVI Classification Sensor Type: Unknown Byte Order : Host (Intel) Projection : UTM, Zone 49 North Pixel : 30 Meters Datum : WGS-84 Wavelength : None Upper Left Corner: 1,1 Description: Maximum Likelihood Classification Result [Wed Jun 14 22:18:26 2017]
OK Cancel Previous	en v

Select Ground Truth ROI		Select Classification Imag	e
耕地 河流 居民地 道路 湖泊坑塘	* III *	Unclassified	
Ground Truth ROI			
Classification Class	į.		
Add Combination			
Matched Classes			
居民地 <-> 居民地 湖泊坑塘 <-> 湖泊坑塘 耕地 <-> 耕地 道路 <-> 道路 裸地 <-> 裸地	ŧ		4 III 4



下面对混淆矩阵中的几项评价指标进行说明:

(1) 总体分类精度

等于被正确分类的像元总和除以总像元数。被正确分类的像元数目沿着混淆矩阵的对角 线分布,总像元数等于所有真实参考源的像元总数,如本次精度分类精度表中的 Overall Accuracy = (1849/2346) 78.8150%。

(2) Kappa 系数

它是通过把所有真实参考的像元总数(N)乘以混淆矩阵对角线(X_{KK})的和,再减去某 一类中真实参考像元数与该类中被分类像元总数之积之后,再除以像元总数的平方减去某一 类中真实参考像元总数与该类中被分类像元总数之积对所有类别求和的结果。



(3) 错分误差

指被分为用户感兴趣的类,而实际属于另一类的像元,它显示在混淆矩阵里面。本例中,

林地有 419 个真实参考像元,其中正确分类 265,12 个是其他类别错分为林地(混淆矩阵中 林地一行其他类的总和),那么其错分误差为 12/419=2.9%。

(4)漏分误差

指本身属于地表真实分类,当没有被分类器分到相应类别中的像元数。如在本例中的耕地类, 有真实参考像元 465 个,其中 462 个正确分类,其余 3 个被错分为其余类(混淆矩阵中耕地 类中一列里其他类的总和),漏分误差为 3/465=0.6%

(5)制图精度

是指分类器将整个影像的像元正确分为 A 类的像元数(对角线值)与 A 类真实参考总数(混淆矩阵中 A 类列的总和)的比率。如本例中林地有 419 个真实参考像元,其中 265 个正确分类,因此林地的制图精度是 265/419=63.25%。

(6) 用户精度

是指正确分到 A 类的像元总数(对角线值)与分类器将整个影像的像元分为 A 类的像 元总数(混淆矩阵中 A 类行的总和)比率。如本例中林地有 265 个正确分类,总共划分为 林地的有 277,所以林地的用户精度是 265/277=95.67%。

注:监督分类中的样本选择和分类器的选择比较关键。在样本选择时,为了更加清楚的 查看地物类型,可以适当的对图像做一些增强处理,如主成分分析、最小噪声变换、波段组 合等操作,便于样本的选择;分类器的选择需要根据数据源和影像的质量来选择,比如支持 向量机对高分辨率、四个波段的影像效果比较好。

3、实训要求

(1)每人以自己的数据为基础,设置至少5类地物,分别选择不同的分类法,进行非监督分类,并且比较在相同的训练样本条件下的分类精度。

(2)以文字+截图的形式,简要撰写实训报告,包括主要步骤和结果,问题与解决办法, 意见和建议等。

4、提交结果

(1) 最终分类结果:不同的分类方法各1份, tiff 格式。

(2) 实训报告: 1 份, doc 格式。