

遥感影像分类是研究青海湖区域土地利用与覆盖分类的一项非常重要的支撑技术. 几十年来, 相关研究人员与机构不断提出各种遥感影像分类算法, 极大的推进了遥感数据处理与分析技术的进步.

上世纪 80 年代, 研究人员主要是利用统计模式识别的方法进行遥感影像的分类研究. 90 年代起, 出现了大量的机器学习遥感影像分类方法, 机器学习分类算法在遥感领域的应用极大的推动了遥感影像分类技术的发展. 2000 年陈华^[1]把 K-means 聚类方法应用于多光谱图像分类, 重点讨论了算法的技术过程. 尹世源^[2]用最大似然分类算法对山东丁字湾遥感影像进行分类, 并得到了较好的分类效果. 杨育桢^[3]基于 TM 影像的波段值与 NDVI(归一化植被指数)、DEM(数字高程模型) 坡度分析使用决策树分类算法进行地表覆盖分类的分析, 取得了比最大似然分类算法更好的分类效果. 惠文华^[4]提出一种基于 SVM、光谱特征和纹理特征相结合的遥感图像分类方法, 取得了明显优于最大似然分类的结果. 刘志刚^[5]探讨了一种基于支持向量机的遥感影像半监督分类新方法.

2006 年, 多伦多大学的 Geoff Hinton 提出深度学习的概念, 现在深度学习已经成为了机器学习、人工智能领域的研究热点, 尤其是在 2012 年的 ImageNet LSVRC-2010 比赛中, 由 Hinton 教授及其学生提出的 AlexNet 模型^[6]取得了惊人的效果, 将 to-5 的错误率由 25% 降到了 17%, 由此, 深度学习引起了学术界和工业界的广泛关注. 卷积神经网络 (Convolutional Neural Networks, CNN) 是深度学习网络的一种, AlexNet 网络模型就是 CNN 模型的一种. 曹林林^[7]将卷积神经网络应用到高分辨率遥感影像的分类中, 取得了比 SVM 更好的分类效果, 证明了卷积神经网络在高分辨率遥感影像分类中的可行性及精度优势.

针对当前的研究现状, 本文借鉴 GoogLeNet Inception 结构, 设计并提出了一种卷积神经网络模型用于 30 米分辨率 LandSat 8 OLI 青海湖区域遥感影像的特征提取与分类, 并设计实验分析了生成样本的邻域窗口尺寸对分类结果的影响, 以及本文模型与最大似然分类和 SVM 分类器对分类结果的对比. 实验结果表明窗口尺寸为 9×9 时, CNN 的总体分类效果最好, 且优于最大似然分类和 SVM 分类效果. 证明了卷积神经网络模型在 30 米分辨率 LandSat 8 OLI 青海湖区域遥感影像分类中特征提取与分类的有效性.

1 相关工作

许多研究者对青海湖区域的土地覆盖变化展开过研究, 李小雁^[8]以 1977~2004 年 4 期的遥感影像数据为基础, 采用土地利用转移矩阵、动态度、利用程度等方法研究了青海湖流域 1977 年以来的土地利用变化特征及其空间分异规律. 祁佳丽^[9]选用 2009~2011 年的遥感影像数据, 采用土地利用转移矩阵研究分析方法对青海湖流域的土地利用类型进行了分析. 刘娟^[10]以青海湖的 TM 数据和地形数据为主要数据源, 在 GeoEye-1 高分辨率影像和土壤图的帮助下, 采用最大似然的方法探讨了遥感技术在青海湖流域土壤分类中的可行性.

卷积神经网络在高分遥感影像分类的应用比较成熟, 许多研究者提出了不同的数据处理方法. 邢晨^[11]考虑到提取空间信息的窗口大小可能会包含进来与中心像素点不同类别的干扰像素, 其尝试了一维卷积的方式, 即将一个空间邻域的数据串联起来的光谱数据作为一个通道的图像, 然后将 3×3 窗口尺寸的空间的像素数作为通道数, 其使用的 INP 数据有 200 个波段, 就得到 9 通道 1×200 的图像作为卷积神经网络的输入, 这样得到的图像宽度就大大加宽. 宋欣益^[12]由原始高分数据得到某一个像元在所有波段的光谱数值, 然后转为二维矩阵, 并灰度化得到二维图像输入卷积神经网络.

适用于高分图像的处理方法并不适用于中分遥感影像, 因为这些处理方式得益于高分遥感影像具有非常多的波段数据, 而 LandSat8 遥感影像只有 11 个波段的数据, 即使全部利用也只能得到 1×11 的图像.

付秀丽^[13]在处理中分 LandSat5 遥感影像时, 按照 32×32 大小的尺寸对 TM 影像进行裁剪, 完成对数据的预处理及样本生成. 但是对于 30 分辨率的遥感影像来说, 32×32 的窗口实在太大, 必然会导致线条特征提取不明显, 分类边界不够细腻, 影响图像分类效果. 张伟^[14]在进行 16 米空间分辨率多光谱影像的特征提取时, 使用了动态窗口上采样的方法做样本生成, 其使用预训练的深度卷积神经网络 AlexNet 模型进行特征提取, 并用 SVM 进行分类, 取得了优于基于光谱+纹理特征的分类结果. 但是, 上采样的方法会额外增加计算量, 此外, 该方法也会给原始数据带来冗余信息, 扰乱图像原始的纹理结构, 在一定程度上会影响分类效果.

本文首次将卷积神经网络应用于青海湖区域遥感影像分类,同时为了保持图像原始的纹理结构,保证CNN提取到最准确的特征表示,本文不采用上采样的方法做样本生成,而是保持原始窗口大小的图像样本。同时为了能更好的提取到小窗口图像的特征表示,本文借鉴GoogLeNet Inception结构设计并提出了自己的卷积神经网络模型。

表1 LandSat8 波段说明

Band1	Band2	Band3	Band4	Band5	Band6	Band7	Band8	Band9	Band10	Band11
海岸波段	蓝波段	绿波段	红波段	近红外波段	短波红外 1	短波红外 2	全色波段	短波红外波段	热红外 1	热红外 2

表1中波段1主要应用于海岸带观测;波段9又称卷云波段,包含水汽强吸收特征,一般用于云层检测;波段8、10、11的空间分辨率分别为15 m、100 m、100 m,一般不参与分类。对于一些传统模型,通常会选择波段信息量丰富、波段相关性小、地物光谱差异大、可分性好的最佳波段组合来进行特征图像解译和特征提取,但是考虑到深度学习网络具有极其强大的学习及特征表达的能力,本文将把剩余的波段2~7全部利用起来,作为输入样本的6个通道。本文实验利用了其中的6个波段,对TIF波段文件的读取和处理使用的是Python GDAL库以及Python Numpy库。LandSat8数据是16位的,最大值为65 535,本文实验对每个波段的像元值做了层内归一化,归一化到[0,1]。

根据实地考察及相关资料参考,本文把青海湖区域的地物覆盖分为7类:草地、荒地、高寒草甸、农田、人工用地、沙地和水体。CNN训练是一种有监督训练,需要大量有标注的训练样本。本文利用Google Earth圈定了7类样本数据,然后将标记导出并导入到ArcGis中。

不同类别的样本数量有较大差异。其中沙地、水体、高寒草甸样本数量较大,而人工用地等样本数量较少。为了避免样本不平衡带来的模型预测偏移的情况,本文对每类样本随机选取了1.5万个样本,此举也使得样本分布更为分散,减少样本重叠的情况。

2.2 CNN 模型

图1即为GoogLeNet Inception结构的v1版本,该结构将1×1、3×3、5×5的卷积操作与3×3的池化操作stack在一起,一方面增加了网络的宽度,另一方面也有利于网络充分提取多尺度感受野的特征。

2 实验设计

2.1 实验数据

本文使用的实验数据是成像于2016年10月17日的30米分辨率LandSat8 OLI青海湖区域遥感影像,云层覆盖率低于1%。LandSat8共有11个波段的光谱数据,每个波段的说明见表1。

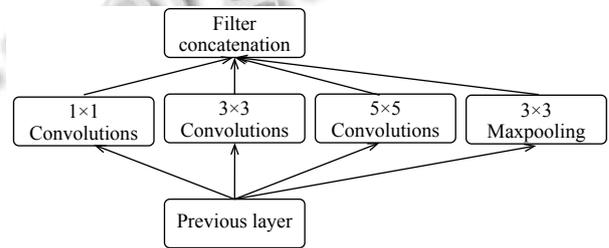


图1 GoogLeNet Inception v1

本文借鉴GoogLeNet Inception结构设计并提出的CNN网络模型见图2。

CNN能在图像处理领域取得极大成功的一个特性就是感受野,感受野是指CNN结构中某个特征映射到输入空间的区域大小。本文使用1×1、3×3、5×5不同大小的卷积核来提取不同大小的感受野特征,充分挖掘中心像元与邻域空间的特征联系,提取到相关的特征。本文设计的CNN模型包含两个类Inception结构,每个类Inception之后有一个Batch Normalization层,BN是将该层的输出归一化为均值为0,方差为1,BN操作可以有效的加快模型训练速度,提高模型精度。两个类Inception之后是一个3×3的卷积层和一个512维的全连接层,最后是输出层。

3 实验结果分析

3.1 邻域窗口大小ws对分类结果的影响

如2.1小节所述,为了选择最优的邻域窗口尺寸,本文选择不同尺寸的ws做了实验,选取的ws的尺寸为5、7、9、11、13。不同窗口尺寸的分类结果见图3。

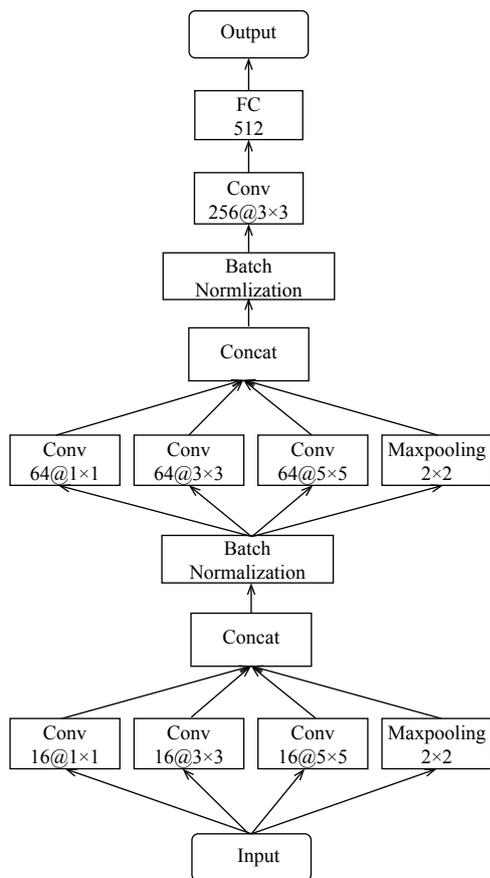


图2 本文设计并提出的 CNN 模型结构

从图3可以看到, ws 较小时分类图中出现了较多的“椒盐噪声”式的细碎图斑, 分类区块多且不连续, 观察 ws 等于 5 时可以发现, 模型较好的识别出了道路等细长的人工用地, 这是因为窗口较小时, 模型能较好的提取和表征狭小地物的特征, 但同时对于区域面积较大的其他 6 个分类, 小窗口能够提取的特征较少, 就非常容易出现错分的情况, 可以看到在高寒草甸内部出现了很多错分的零散的人工用地. 随着窗口 ws 的增大, 整体分类效果也在慢慢变得合理. 但是当 ws 变大为

11 或 13 时, 可以发现出现了非常大的连片区域, 分类边界模糊, 也出现部分区域被整体错分的情况; 观察人工用地, 出现了人工用地被明显加粗加宽的现象. 这是因为太大的邻域窗口会包含其他的地物信息或冗余信息, 而不能使用模型提取出最能代表某一地物的特征, 尤其是在边界附近的像元, 邻域窗口越大, 样本所包含进去的冗余信息就越多, 就越不利于分类.

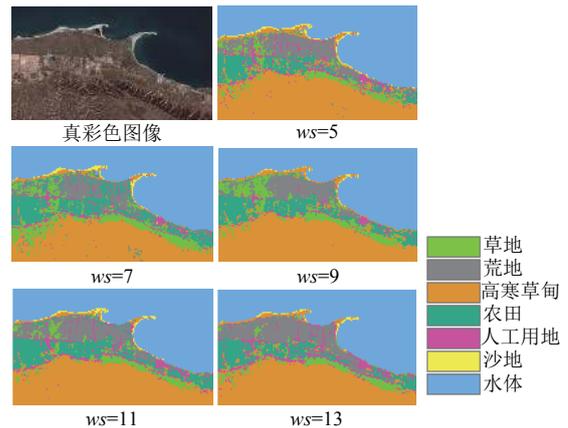


图3 不同邻域窗口尺寸下的 CNN 分类结果

经过对比不同窗口尺寸的分类结果与实地地物覆盖情况, 本文认为在邻域窗口 ws 等于 9 时, 分类效果达到最好, 分类结果与实地地物覆盖分类最为吻合.

3.2 最大似然分类和 SVM 与 CNN 实验结果对比

使用相同的训练像元, 本文又分别使用最大似然分类和 SVM 方法做了两组对比实验. 本文采用混淆矩阵进行精度对比, 结果如表 2、表 3 和表 4.

由表 2、表 3 和表 4 可知, 使用 CNN 模型计算得到的总体分类精度为 80.97%, Kappa 系数为 0.778, 使用最大似然分类计算得到的总体分类精度为 54.46%, Kappa 系数为 0.469. 由 SVM 方法得到的分类结果的总体分类精度 55.26%, Kappa 系数为 0.478.

表2 CNN 精度评价

地物类型	草地	荒地	高寒草甸	农田	人工用地	沙地	水体	总计	生产者精度 (%)
草地	401	1	88	4	6	0	0	500	80.2
荒地	4	285	8	152	50	1	0	500	57.0
高寒草甸	22	0	471	4	3	0	0	500	94.4
农田	0	98	1	380	21	1	0	500	76.0
人工用地	2	24	59	77	337	1	0	500	67.4
沙地	0	2	11	3	1	462	21	500	92.4
水体	0	0	0	1	0	1	498	500	99.6
总计	429	410	638	621	418	475	510	3500	
用户精度 (%)	93.47	69.51	73.82	61.19	82.67	97.26	95.88		

表3 最大似然分类精度评价

地物类型	草地	荒地	高寒草甸	农田	人工用地	沙地	水体	总计	生产者精度 (%)
草地	268	8	197	3	24	0	0	500	53.6
荒地	0	97	18	141	236	8	0	500	19.4
高寒草甸	85	19	307	4	83	2	0	500	61.4
农田	0	26	35	170	269	0	0	500	34.0
人工用地	1	19	39	26	403	12	0	500	80.6
沙地	0	1	1	0	29	469	0	500	93.8
水体	0	0	0	0	32	276	192	500	38.4
总计	354	170	597	344	1086	767	192	3500	
用户精度 (%)	75.7	57.1	51.4	49.4	37.1	61.1	100.0		

表4 SVM精度评价

地物类型	草地	荒地	高寒草甸	农田	人工用地	沙地	水体	总计	生产者精度 (%)
草地	446	1	1	51	0	1	0	500	89.2
荒地	175	35	10	207	54	19	0	500	7.0
高寒草甸	274	79	6	126	1	11	3	500	1.2
农田	67	28	1	388	14	2	0	500	77.6
人工用地	31	44	9	176	213	27	0	500	42.6
沙地	5	8	5	3	89	346	44	500	69.2
水体	0	0	0	0	0	0	500	500	100
总计	998	195	32	951	371	406	547	3500	
用户精度 (%)	44.69	17.95	2.00	40.80	57.41	85.22	91.41		

从以上结果可以看出,不论是总体的分类精度还是 Kappa 系数,本文设计的 CNN 模型都要高于最大似然分类和 SVM。

最大似然分类方法假设数据服从正态分布或联合正态分布,如果某一类的数据分布不服从正态分布,那么该方法就未必合适。我们观察表3中,水体的预测准确率只有 28.4%,而 SVM 和 CNN 关于水体的预测准确率都高达 99% 甚至 100%。这是因为,遥感影像中的水体的光谱数据一般只在一个非常小的范围内变化,其数据分布规律不符合正态分布,因此水体的预测准确率较低。此外,观察 SVM 的分类结果发现,荒地和高寒草甸的分类结果非常不理想,两者被更多的错误分类为草地和农田,本文使用的遥感影像成像时间为 10 月份,4000 多米海拔的青海湖区域已经处于深秋季节,所以这几类地物具有非常相似的景观特征,非常难以识别。而 CNN 的分类结果却好很多,这也说明了卷积神经网络具有非常强大的特征提取与表达的能力。

图4也直观的展现最大似然分类、SVM 分类与 CNN 分类效果的差异。

3.3 CNN 在青海湖区域遥感影像分类的应用

根据 2.3.1 及 2.3.2 两小节对实验过程及结果的描

述与对比,卷积神经网络在遥感影像分类中具有比最大似然分类和 SVM 更好的分类效果,且对于中等分辨率的遥感影像,选择邻域窗口为 9 时,分类效果最好。图5即为使用最优参数的卷积神经网络在 30 米中分辨率青海湖区域遥感影像的分类结果图。

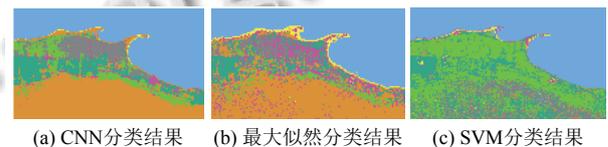


图4 CNN 与最大似然分类、SVM 分类结果对比

4 结论

鉴于卷积神经网络具有非常强大的特征学习和表达能力,本文提出并设计了一种卷积神经网络模型,并将其应用到中分辨率青海湖区域遥感影像的分类中。同时本文针对样本生成中邻域窗口尺寸对分类效果的影响,以及 CNN 与最大似然分类和 SVM 分类效果的优劣,分别作了对比实验。本文得出以下结论:

1) 与最大似然分类和 SVM 方法相比,卷积神经网络在遥感影像分类应用中具有更强的特征提取能力与更优的分类效果。

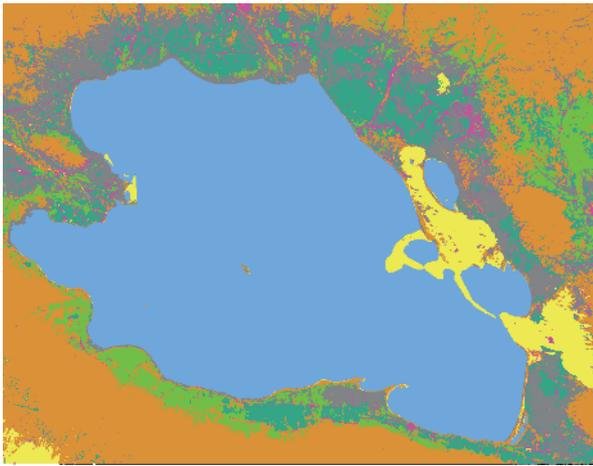


图5 青海湖区域分类结果

2) 针对30米分辨率的LandSat 8 OLI青海湖区域遥感影像,生成样本时所需的邻域窗口尺寸对最终的分分类结果有一定的影响。尺寸太小,会产生很多“椒盐噪声”式的错分点;尺寸太大,会导致部分整块区域的错分以及边界模糊。邻域窗口尺寸为 9×9 时取得最好的分类效果。

3) 本文方法也存在不足:本文在选取最优邻域窗口尺寸时,采用了与实际地物覆盖人工对比的方法,而没有找到一个可量化的衡量指标。此外,不论生成样本的窗口尺寸有多大,实际分类结果的边界都会存在边界模糊的情况。

在今后的工作中,我们将继续探索,力求找到能较好的解决3)中所提到问题的解决方案;我们还将将在样本生成方法、模型结构等方面继续研究,努力为青海湖区域遥感影像分类提供新的更科学的技术方法。

参考文献

1 陈华,陈书海,张平,等. K-means 算法在遥感分类中的应

用. 红外与激光工程, 2000, 29(2): 26–30. [doi: 10.3969/j.issn.1007-2276.2000.02.008]

2 尹世源. 计算机最大似然法分类在海岸调查中应用的探讨. 海岸工程, 1983, (1): 79–82.

3 杨育楨, 陈晓宁, 高永吉. 基于决策树方法的遥感分类研究. 测绘标准化, 2015, 31(2): 10–11.

4 惠文华. 基于支持向量机的遥感图像分类方法. 地球科学与环境学报, 2006, 28(2): 93–95. [doi: 10.3969/j.issn.1672-6561.2006.02.019]

5 刘志刚, 史文中, 李德仁, 等. 一种基于支撑向量机的遥感影像不完全监督分类新方法. 遥感学报, 2005, 9(4): 363–373.

6 krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 2012, 25(2): 1097–1105.

7 曹林林, 李海涛, 韩颜顺, 等. 卷积神经网络在高分遥感影像分类中的应用. 测绘科学, 2016, 41(9): 170–175.

8 李小雁, 许何也, 马育军, 等. 青海湖流域土地利用/覆被变化研究. 自然资源学报, 2008, 23(2): 285–296. [doi: 10.3321/j.issn:1000-3037.2008.02.013]

9 祁佳丽, 李志强, 李飞, 等. 青海湖流域土地利用/土地覆盖变化研究. 青海环境, 2013, 23(4): 197–202. [doi: 10.3969/j.issn.1007-2454.2013.04.011]

10 刘娟, 蔡演军, 王瑾. 青海湖流域土壤遥感分类. 国土资源遥感, 2014, 26(1): 57–62.

11 邢晨. 基于深度学习的高光谱遥感图像分类[硕士学位论文]. 北京: 中国地质大学, 2016.

12 宋欣益. 基于卷积神经网络的高光谱数据分类方法研究[硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2016.

13 付秀丽, 黎玲萍, 毛克彪, 等. 基于卷积神经网络模型的遥感图像分类. 高技术通讯, 2017, 27(3): 203–212. [doi: 10.3772/j.issn.1002-0470.2017.03.002]

14 张伟, 郑柯, 唐聘, 等. 深度卷积神经网络特征提取用于地表覆盖分类初探. 中国图象图形学报, 2017, 22(8): 1144–1153.