

引用格式: Li Hongda, Gao Xiaohong, Tang Min. Land Cover Classification for Different Spatial Resolution Images from CNN[J]. Remote Sensing Technology and Application, 2020, 35(4): 749-758. [李宏达, 高小红, 汤敏. 基于 CNN 的不同空间分辨率影像土地覆被分类研究[J]. 遥感技术与应用, 2020, 35(4): 749-758.]
doi: 10.11873/j.issn.1004-0323.2020.4.0749

基于 CNN 的不同空间分辨率影像土地覆被分类研究

李宏达, 高小红, 汤 敏

(青海师范大学地理科学学院, 青海省自然地理与环境过程重点实验室, 高原科学与可持续发展
研究院, 青藏高原地表过程与生态保育教育部重点实验室, 青海 西宁 810008)

摘要: 基于卷积神经网络(Convolutional Neural Networks, CNN)和 5 种不同空间分辨率的遥感影像, 对西宁市东部一区域开展土地覆被分类研究, 旨在探索 CNN 在不同空间分辨率下进行影像分类的差异性和对不同地物的提取能力。为提高样本的选择效率, 引入了窗口滑动方法进行辅助选择。研究表明 5 种不同空间分辨率影像的总体分类精度均达 89% 以上, Kappa 系数达 0.86 以上, 分类精度较高。在所涉及的分辨率尺度范围内, 空间分辨率越高, CNN 分类结果越精细, 并能保持较高的分类精度, 表明 CNN 更适合高空间分辨率影像分类; 但同时影像空间分辨率越高, 地物表现出较高的类内变异性和低类间差异性, 分类精度有降低的趋势。相比较而言, SPOT 6 影像的分类精度最高, 同时窗口滑动是一种有效的样本辅助选择方法。研究对今后同类工作具有一定的借鉴意义。

关 键 词: CNN; Landsat-8/Sentinel-2A/SPOT-6/GF-2 影像; 土地覆被分类

中图分类号: TP75 **文献标志码:** A **文章编号:** 1004-0323(2020)04-0749-10

1 引 言

传统参数化遥感数据分类方法(如最大似然法)计算简单、实现方便, 但因仅使用少量的影像特征进行分类, 同时假设了数据的参数化分布, 泛化能力差, 精度通常较低。非参数化机器学习方法, 如人工神经网络、支持向量机、决策树和随机森林等具有较强的学习能力及泛化能力, 在遥感影像分类中得到广泛应用。有研究表明^[1-3], 除决策树以外, 人工神经网络、支持向量机和随机森林在不同的分类场景下, 均有可能是最优的影像分类方法。但总体上它们属于浅层学习, 很难有效表达复杂的函数变化, 对样本缺少足够的适应性^[4]。同时, 这些机器学习算法也没有或未充分利用影像的上下文结构信息, 尤其对于高空间分辨率影像, 其波段数

量通常较少, 光谱信息有限, 表现出较高的类内变异性和类间低差异性, 仅用少量的波段信息去区分相似的地物无疑增加了分类的工作难度^[5]。

20 世纪 80 至 90 年代, 在人工神经网络的基础上, 兼顾图像上下文结构信息的卷积神经网络(Convolutional Neural Networks, CNN)开始出现, 2006 年以来随着计算机技术的发展和深度学习理论的提出, CNN 被广泛应用于计算机视觉、自然语言处理等方面, 并逐步应用到遥感影像分类领域。Masoud 等^[5]采用 7 种不同结构的 CNN 对 RapidEye 5 m 多光谱影像进行土地覆被分类, 总体分类精度最高为 96.17%, 同时所有 CNN 的分类结果均优于支持向量机和随机森林方法。孟祥锐等^[6]基于 CNN 对 GF-2 融合后的 1 m 多光谱数据进行地物覆

收稿日期: 2019-08-27; 修订日期: 2020-06-24

基金项目: 青海省科技厅自然科学基金项目(2016-ZJ-907)资助。

作者简介: 李宏达(1995—), 男, 湖北荆门人, 硕士研究生, 主要从事遥感应用与地理空间数据分析研究。E-mail: 2395789679@qq.com

通讯作者: 高小红(1963—), 女, 陕西白水人, 教授, 博士生导师, 主要从事遥感信息提取与土地覆被变化方面的研究。

E-mail: xiaohonggao226@163.com

被分类,取得了 87.58% 的总体分类精度和 0.86 的 Kappa 系数,验证了 CNN 对湿地群落精细划分的可行性。张伟等^[7]利用 CNN 对 GF-1 16 m 多光谱数据进行了特征提取,并以 SVM 作为分类器对北京密云水库周边地区开展土地覆被分类,得到了 97.99% 的总体分类精度和 0.98 的 Kappa 系数,表明 CNN 可以提取更精细、更准确的地表覆盖特征。李亚飞等^[4]基于 CNN 对 Landsat-8 30 m 的 OLI 影像开展土地覆被分类研究,取得了 97.83% 的总体分类精度和 0.97 的 Kappa 系数,在影像分辨率相对较低的情况下仍获得了较高的分类精度。目前,CNN 在土地覆被分类方面的研究更多倾向于 CNN 算法精度本身,而欠缺考虑影像空间分辨率对于 CNN 分类精度的影响。崔刚等^[8]基于深度学习对不同空间分辨率影像的冬小麦识别进行了研究,结果表明随着影像分辨率的提高,冬小麦的识别精度也越高。但该研究为单一地物的提取,影像空间分辨率提高导致地物光谱复杂程度增加对其精度的影响可能不大,这与多分类的情况存在较大差别。

CNN 在发展过程中也衍生出了一些新的深度学习方法,在影像分类方面较为出色的如全卷积神经网络(Fully Convolutional Network,FCN)。Maggiori 等^[9]基于空间分辨率为 1 m 的多光谱数据研究表明:相比于 CNN,利用 FCN 对建筑物进行提取会大幅度节省时间,同时提取精度也会有一定上升。杨瑞等^[10]基于 FCN 改进和扩展而来的 U-Net 对高分二号影像融合后进行了多要素分类,取得了优于支持向量机和面向对象方法的分类精度。同传统机器学习算法相比,CNN 不需要复杂的特征工程,拥有更好的泛化能力,分类精度更高;相比于 FCN,CNN 的计算量较大。但为获得更为普适性的规律,对其他衍生的深度学习方法也形成一定的参考价值,本文选用较为常规的 CNN 作为研究方法。

就目前文献检索结果来看,CNN 对影像进行分类时,影像的空间分辨率多集中在 4 m 以上,这与

CNN 对数据上下文结构信息的要求有很大关系。但通常来讲,影像的空间分辨率越高,数据获取的费用及对其进行处理的时间成本也愈高。为明确 CNN 对不同分辨率影像分类的差异性,选取 5 种不同空间分辨率的遥感影像数据,分辨率跨度为 4~30 m,探索不同空间分辨率对 CNN 分类精度的影响,为 CNN 方法下土地覆被分类影像的选择提供合理参考。

2 研究区及数据

2.1 研究区概况

西宁市地处湟水流域中游河谷盆地,面积为 497 km²,地理坐标为 36°12′27″~37°30′09″ N, 100°47′56″~101°56′49″ E,市区平均海拔 2 295 m,属于高原大陆性半干旱气候^[11],湟水河自西向东流经市区。研究区为西宁市东部的一区域(图 1),土地利用类型主要包括建设用地、林地、草地、耕地及水域,且分布无明显规律,能够充分验证 CNN 的学习能力。

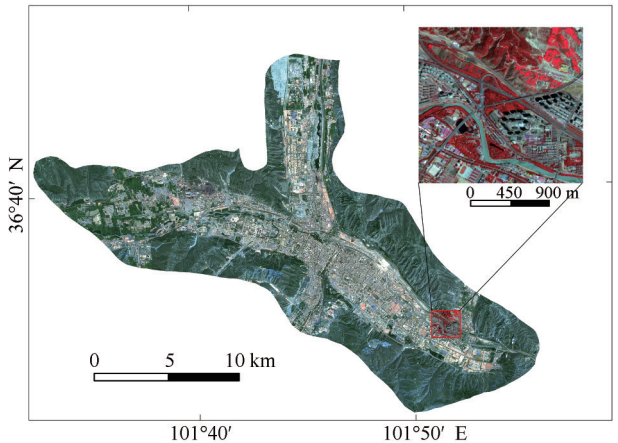


图 1 研究区位置图

Fig.1 Location of the study area

2.2 数据及预处理

研究采用的影像数据为 GF-2 4 m 多光谱数据(免费来源于高分辨率对地观测青海数据与应用中

表 1 卫星影像及其参数
Table 1 Satellite images and their parameters

卫星平台	所属国家	空间分辨率/m	空间参考	使用光谱波段	影像获取时间
GF-2	中国	4	WGS-1984	R、G、B、NIR	2015-07-28
SPOT-6	法国	6			2016-08-08
Sentinel-2A	欧空局	10			2016-07-27
Landsat-8	美国	15/30			2016-08-07

注: Sentinel-2A 影像的 10 m 空间分辨率波段为 b2(蓝)、b3(绿)、b4(红)和 b8(近红外); Landsat-8 的 15 m 影像是经全色波段 b8 与其多光谱数据融合后得到,分类时使用了 b2(蓝)、b3(绿)、b4(红)和 b5(近红外)4 个波段。

心)、SPOT-6 6 m 多光谱数据、Sentinel-2A 10 m 多光谱数据(免费来自于欧洲航天局 <https://scihub.copernicus.eu/dhus/#/home>)、Landsat-8 OLI 的 30 m 空间分辨率数据(免费来源于中国地理空间数据云 <http://www.gscloud.cn/sources/>)及其全色波段融合后的 15 m 空间分辨率数据。其中,GF-2 和 SPOT-6 的多光谱数据含有红、绿、蓝和近红外 4 个波段,Sentinel-2A 的多光谱数据共包含 13 个波段,红、绿、蓝和近红外 4 个波段为 10 m 空间分辨率,研究中仅采用了 Sentinel-2A 的 10 m 分辨率数据。为了保持数据波段的一致性,Landsat-8 两种空间分辨率的数据也仅采用了红、绿、蓝和近红外 4 个波段。

CNN 作为一个深度学习模型,其最初的目的就是避免复杂的特征工程^[12],因此 CNN 可以直接以图像的原始像素作为输入,不需要复杂的数据预处理工作。文中 4 种多光谱数据的地理参考均为 WGS-1984 椭球体,其中 Sentinel-2A 影像数据为 Level-1C 级数据,是经正射校正和亚像元级几何精校正后的正射影像产品^[13],所以仅对 SPOT-6 和 GF-2 影像进行了几何校正,并对所有数据进行了大气校正处理。

3 CNN 搭建与数据处理

3.1 CNN 简介

CNN 是多层前馈神经网络的一种变体,与传统

的神经网络一样,它也是由可学习的权重和偏置组成,含有较多的隐含层是其区别于人工神经网络的关键^[14]。相较于传统神经网络的全连接方式,CNN 最大的特点在于局部连接和权值共享。科学家通过研究发现猫的一个视觉神经元只处理一小块区域内的视觉图像,这个区域被称为感受野^[15]。具体的,局部连接和权值共享是指每个神经元不需要接收全部像元的信息,只需接受局部像元点的信息输入(感受野),然后将所有的局部信息综合起来便可得到整体信息。这一特点大大降低了训练的参数量及网络的复杂程度,所需训练的参数与影像无关,仅取决于卷积核的大小及数量,减小了训练的难度。同时,权值共享还赋予了 CNN 对图像平移的容忍性,提高了模型的泛化能力^[16]。一个完整的 CNN 一般由卷积层、池化层、激活函数及全连接层构成。

卷积层由多个卷积核构成,卷积核通俗地理解为一系列的权重阵列^[17]。卷积运算是指将卷积核的每个权重与相应的影像像元一一对应,分别求积后相加。一个卷积核负责提取图像的一种特征,为了从原始影像中提取到足够多的特征,一个卷积层往往含有多个卷积核。图 2 展示了一个卷积核对影像进行卷积运算得到一个特征图的过程,为了不改变卷积运算后图像的大小,在影像边缘部分采取补 0 的方式进行卷积运算。

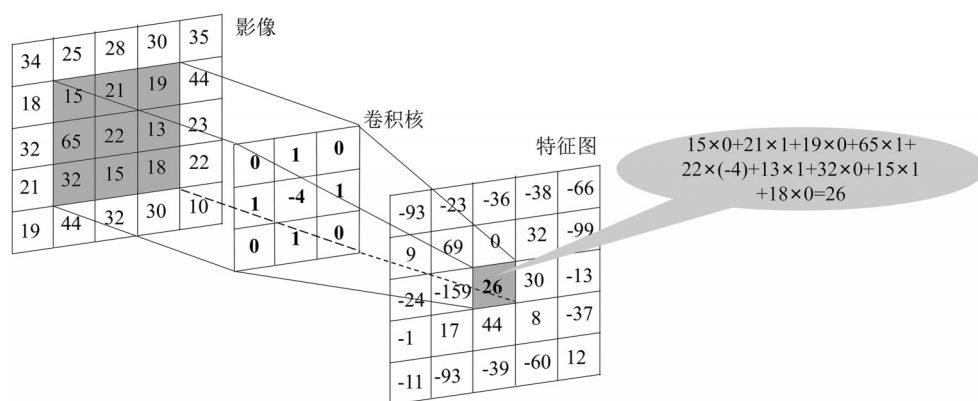


图 2 卷积运算示意图

Fig.2 The schematic diagram of convolution operation

池化层又称下采样层,一般连接在卷积层之后,它的主要作用是对数据进行降维处理。常用的池化有平均池化和最大池化,本文采用最大池化算法,减小计算量的同时保留数据最显著的特征,增强模型的泛化能力^[18]。采用的池化窗口大小为 2×2 像素,步长为 2,进行池化操作前卷积层输出的数

据大小为 5×5 像素,影像边缘部分采用补 0 的方式进行池化,经池化操作后数据大小为 3×3 像素(图 3)。

由于卷积为线性操作,无论堆叠多少层,其最终结果都是经线性计算得到。但现实世界中很多问题是非线性不可分的,为增强 CNN 的表达能

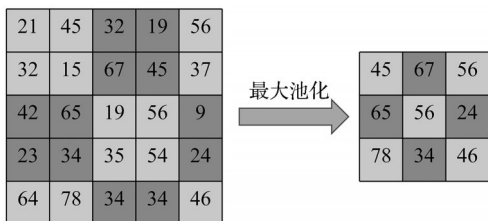


图3 最大池化示意图

Fig.3 Maximum pooling schematic diagram

往会在卷积层之后连接一个激活函数,它是一个非线性函数,用来学习影像的非线性特征,本文采用常见的 ReLu 函数^[19]作为激活函数,其函数图像见图4。

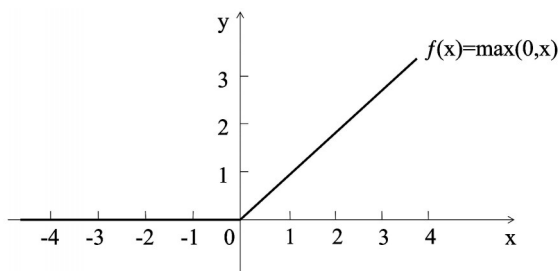


图4 ReLu 函数图像

Fig.4 Graph of ReLu function

经过若干次的卷积和池化操作后,可以得到一些更高抽象层次的特征图,将所有的二维特征图展开拼接为一个一维向量,作为全连接层的输入。全

连接层的每一个神经元都与前一层的所有像元相连接,一方面可以使用全连接层来扩充 CNN 的容量;另一方面,其主要作用是将学习到的抽象特征映射到样本标记空间,完成分类任务。全连接层常会产生过拟合现象,可以增加一个 dropout 层,在训练时随机抛弃一部分节点,减轻或者避免过拟合现象。

3.2 CNN 的搭建

为实现 CNN 分类结果的像素级定位,参考 FCN 经反卷积层实现上采样的思想,研究中根据周围像元确定待分类像元的上下文结构信息,采用三次卷积插值的方法对影像进行上采样后,使用 CNN 进行分类,对比发现该方法取得的分类结果优于通常采用的 patch-based^[9]方法。

CNN 的搭建基于 Google 旗下开源的深度学习框架 Tensorflow^[20]进行,文中搭建的 CNN 结构共 9 层,采用了 5×5 像素和 3×3 像素两种大小的卷积核,以便充分提取不同尺度的特征,每个卷积层后均跟有一个激活函数。由于输入图像较小,为了避免过多的信息损失,仅在第四层的卷积层和激活函数之后使用了一个大小为 2×2 像素,步长为 2 的最大池化层,并在第二个全连接层后使用了一个 dropout 层减轻过拟合现象,最后使用 softmax 分类器完成分类任务,网络详细结构如图 5 所示。

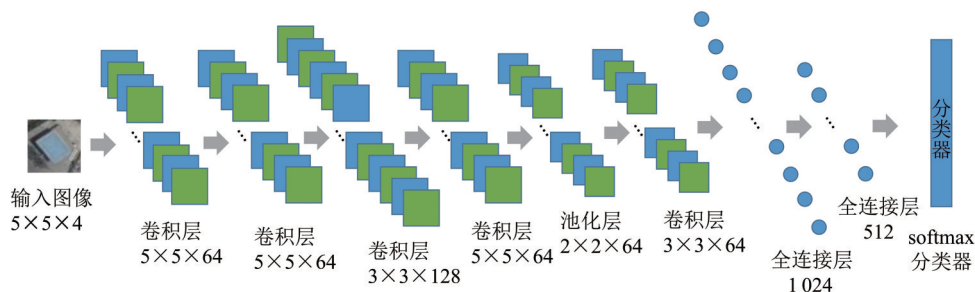


图5 卷积神经网络结构示意图

Fig.5 The schematic diagram of CNN structure

3.3 分类体系及训练样本选取

根据全国遥感监测土地利用覆盖分类体系及城市用地分类与规划建设用地标准,结合研究区实际情况,土地利用共分为 8 类:林地(包括灌木林和疏林地)、耕地、草地、河流、水库坑塘、富营养化水体、城镇建设用地及工业仓储用地。其中富营养化水体仅在 2016 年的影像(Landsat-8、Sentinel-2A 和 SPOT-6)上存在,该水体受藻类影响,整体呈现植被的光谱特征,由于其斑块面积较小,在 Landsat-8 30 m 空间分辨率影像上未对其进行单独分类。对

于影像中由高层建筑所产生的阴影,研究中简单地将其归为城镇建设用地。

训练样本是在参考高空间分辨率影像基础上,基于待分类影像进行选取,考虑到影像空间分辨率及获取时间不同,针对 5 种数据分别选取了不同的训练样本。CNN 要求输入数据大小相同,例如文中选用的为 5×5 像素,目前常见的选样方式为逐块勾选并标记,受样本大小及数量的影响,该方法的工作量较大。由于 CNN 具有旋转的不变性,当样本数量不足时,通常采用的方法是进行样本增强,即

将已有的样本旋转某个角度后作为新的样本,但这样难免会影响样本的丰富性和代表性。

为提高样本选择效率,采用了GDAL和C++相结合的辅助选样方法。首先对每类地物手动数字化选取一定数量较为纯净的像素块作为样本候选区,并尽量保证其在影像上均匀分布。然后使用一个 5×5 的窗口对影像进行遍历操作,窗口每前进一步,根据坐标判断该窗口是否落在数字化的某个多边形内,若是便将其裁剪并标记作为一个样本,否则继续遍历操作。使用固定大小的窗口对影像进行遍历,可以有效避免人工选择时的遗漏,既严格保证了样本大小,也能够较为快速地选出足够数量的样本。

由于研究区内不同的地物类型的面积差异显著,手动数字化的样本数量没有固定大小,主要根据地物的具体分布情况确定。朱秀芳等^[21]对训练样本与分类方法分类精度响应关系的研究表明,对于同一种分类方法,当训练样本数量达到一定程度时,其分类精度是相对稳定的。考虑到本文的研究区域较小,为了使最终的训练样本数量满足统计要求,减少其对研究结果的影响,手动数字化过程中尽可能勾选较多的样本候选区。在这种情况下,对于空间分辨率较高的影像或斑块面积较大的地物,窗口滑动方法得到的训练样本数量较大,可以采取等间距取样的方法从已选出的样本中抽取一部分作为训练时实际使用的样本,或者在进行窗口滑动时设置合适的步长,以保证样本在空间上分布的相对均匀性。

将窗口滑动方法对每种地物选出的最大样本数量设定为300个,其中80%的样本作为用于训练,剩余的20%用作交叉验证,对模型的精度进行评估。对于空间分辨率较低的影像或斑块面积较小的地物,选出的样本数量可能达不到300个,这种情况下将所有选出的样本均予采用,并按同样的比例划分训练样本和测试样本。

3.4 CNN分类参数设置

CNN的各项参数设置对训练及最终的分类结果有较大影响,前人的各项研究中参数设置均有较大差异^[22],根据研究区的实际情况结合多次实验对比,研究中各项参数设置如下:batchsize设置为15,学习率设置为0.000 1,采用Adam优化器进行优化训练。为了防止训练过程中产生过拟合现象,drop-out层的keep_prob在训练时设置为0.6,分类时设

置为1。

受影像接收时刻与分辨率不同的影响,地物产生的阴影在GF-2影像中颜色较深,与正常水体的光谱较为类似,CNN无法直接将两者区分开,在CNN分类结束后计算归一化水体指数^[23],阴影区域均为负值,选择0为阈值,将两者区分开,分类完成后对分类结果进行了碎斑去除处理。

4 结果和分析

4.1 分类结果

5种不同空间分辨率数据的CNN分类结果如图6所示,影像分辨率越高,分类结果对地物细节的展现越好。由于研究区内地物斑块整体偏小,导致影像分辨率为30 m时,分类结果中较大的斑块连成片,边界粗糙,过小的斑块无法有效区分,仅少量的小斑块零星分布,加之不同地物间的光谱信息混合严重,出现了一些错分现象,分类结果较为杂乱;影像分辨率上升到15 m后,小的地物斑块增多,但地物分布细节仍未能得到有效展现;影像分辨率为10 m时,CNN分类结果对地物细节的反映出现了较大的提升,呈现较为明显的相间分布,同时地物边缘多为锯齿状;在SPOT-6影像中,大部分地物斑块在分类结果中均得到了很好的展现,边缘清晰平滑,不再有明显的锯齿状;影像空间分辨率进一步提高到4 m后,一方面地物细节表现更加丰富,另一方面分类结果也显得更为破碎。

4.2 精度评价

分类结果精度评价主要通过总体分类精度、生产者精度、用户精度以及Kappa系数进行评价。验证样本是在野外调查和参考高空间分辨率影像的基础上,基于待分类影像手动数字化选取的感兴趣区,选取验证样本时尽可能覆盖到整个区域。考虑到影像的空间分辨率有较大差异,分别对5种数据选取了不同的验证样本,按空间分辨率从低到高,验证样本所包含的像元数量也依次递增,分别为828、2 636、4 886、10 490和23 287个。

5种不同空间分辨率下,CNN的总体分类精度均高于89.00%(表2),尤其是分辨率高于15 m以后,总体分类精度均高于90.00%,Kappa系数高于0.90,说明CNN对于不同空间分辨率的影像均有着较强的区分能力,能够获得较高的总体分类精度。

伴随着影像空间分辨率从低到高,不同地物的分类精度逐渐趋于稳定,生产者精度和用户精度方

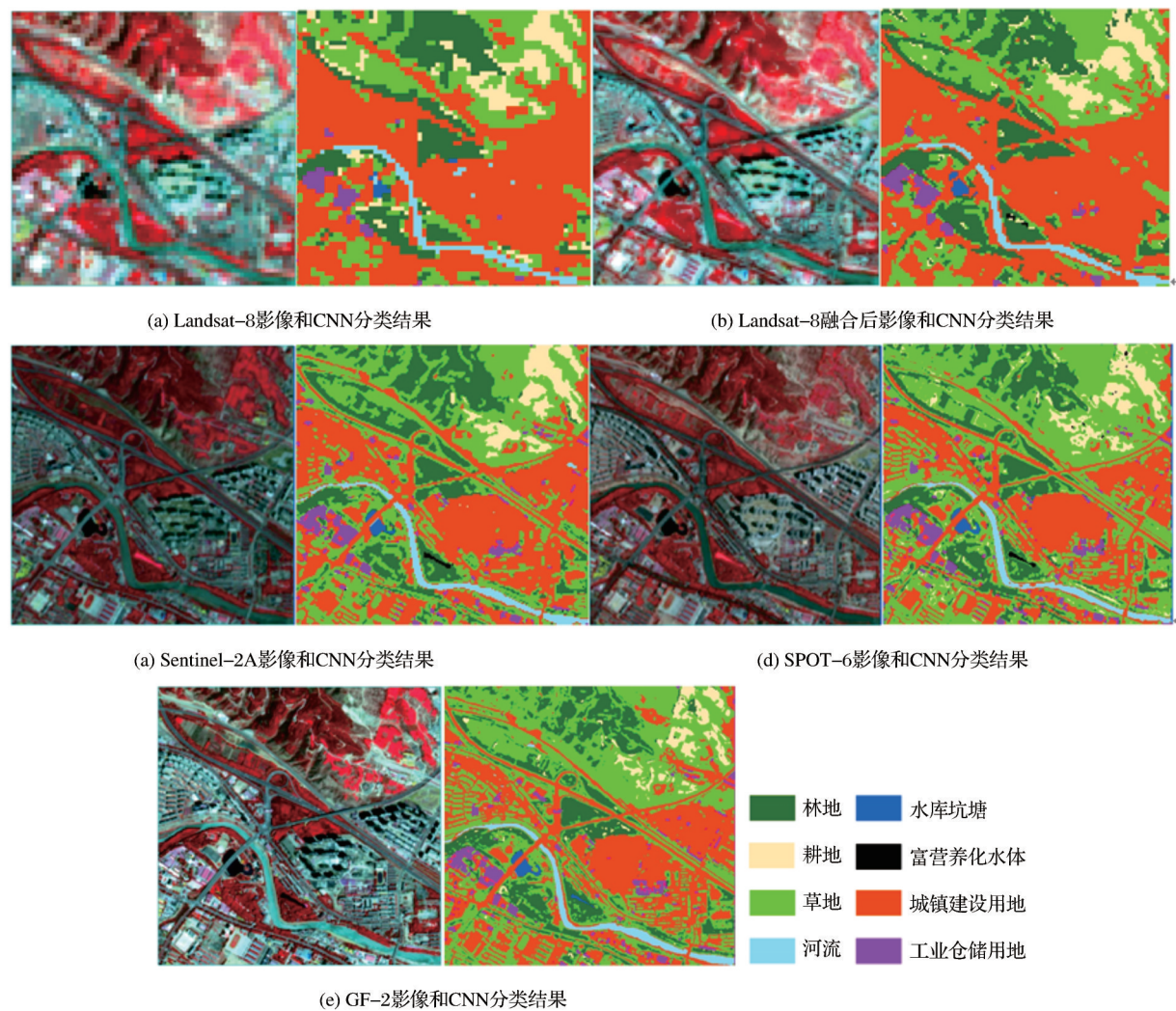


图 6 标准假彩色影像与 CNN 分类结果
Fig.6 Standard false color images and CNN classification results

表 2 分类精度评价

Table2 Classification accuracies evaluation

	Landsat-8		Lanssat-8(融合后)		Sentinel-2A		SPOT-6		GF-2	
	生产者精度	用户精度	生产者精度	用户精度	生产者精度	用户精度	生产者精度	用户精度	生产者精度	用户精度
	/ %	/ %	/ %	/ %	/ %	/ %	/ %	/ %	/ %	/ %
林地	88.48	94.12	96.37	96.54	93.12	97.73	93.55	98.84	93.77	97.95
耕地	89.29	78.13	93.42	87.98	94.83	98.21	93.68	91.05	96.60	91.65
草地	78.43	87.59	90.08	87.57	86.82	78.93	94.24	80.60	93.17	75.79
河流	76.56	96.08	84.11	99.45	89.94	96.40	97.07	100.00	90.72	98.55
水库坑塘	85.71	100.00	85.19	95.83	91.03	100.00	92.86	100.00	88.69	99.62
富营养化水体	——	——	57.14	100.00	96.88	100.00	98.53	91.78	——	——
工业仓储用地	87.10	93.10	94.92	100.00	99.60	93.98	99.01	94.17	96.89	84.52
城镇建设用地	98.87	87.63	94.22	91.47	93.56	93.61	91.46	97.83	90.95	95.62
总体精度/ %	89.01		92.64		92.31		93.65		92.46	
Kappa 系数	0.86		0.91		0.90		0.92		0.90	

差的平均值整体上呈现先降低后升高的趋势,且 SPOT-6 的平均方差最低,图 7 以生产者精度为例,展示了这一趋势,不同地物分类精度中,极差最小

的为 SPOT-6 影像。图 7 也反映出大部分地物随着影像分辨率的提高,生产者精度首先呈现较大幅度上升,但当影像分辨率达到一定程度时,这种上升

趋势开始趋于稳定(如林地),甚至开始呈现出下降的趋势(如工业仓储用地及河流),这说明CNN对多光谱影像进行土地覆被分类存在一个最佳的空间分辨率。伴随着影像分辨率的提高,地物间的规律性也会逐渐减弱,出现较高的类内变异性,相对而言,不同地物类型间的差异性也就随之减小,可能会导致CNN的错分误分比例上升。

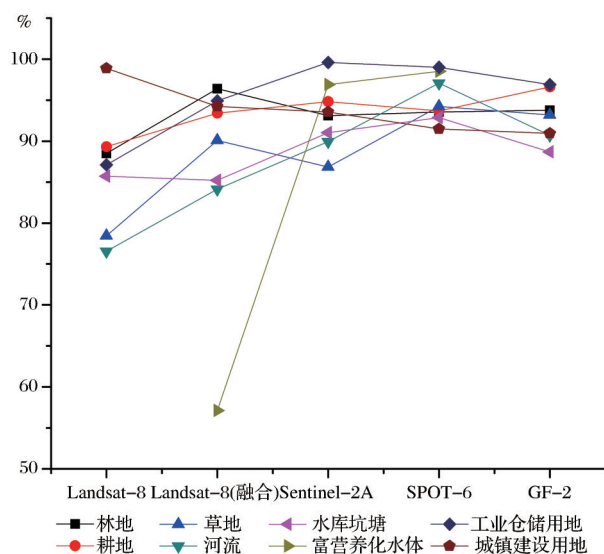


图7 不同空间分辨率影像下地物生产者精度

Fig.7 Producer accuracies of features under different spatial resolution images

总体来看,SPOT-6的总体分类精度优于GF-2的4 m多光谱数据,且其不同地物精度的方差更小,分类结果更稳定,同时也较好地反映了地物细节,可以认为在所涉及的分分辨率范围内,CNN对SPOT-6的6 m多光谱数据分类效果最佳。该结论与其他学者^[24-26]的研究成果相吻合,明冬萍等^[24-25]研究表明,土地覆被分类的最佳空间分辨率为3~7 m,杨闫君等^[26]研究表明多光谱影像土地覆被分类的最佳分辨率为5 m,均与本文的结果较为接近。

由图7中反映的信息来看,影像空间分辨率的提升在一定范围内对CNN提取面积较小或宽度较窄的地物提高分类精度有很大帮助,如富营养化水体本身的斑块面积较小,在Landsat 30 m的影像上未对其进行单独分类,与全色波段融合为15 m后仍未能取得较好的分类精度,但当影像分辨率升高到10 m后,其生产者精度出现大幅提高。受此类影响较为明显的还有同样面积较小的工业仓储用地及宽度较窄的河流。与其他地物不同,随着影像分辨率的提高,城镇建设用地的生产者精度一直呈现下降趋势,这是由于其本身的组成成分较为复杂,随

着影像分辨率的提高,类内的规律性进一步减弱,增加了正确分类的难度。

4.3 结果分析

结合精度评价发现当影像空间分辨率为15 m时,CNN取得了92.64%的总体分类精度,甚至略高于Sentinel-2A的10 m多光谱数据,但就分类结果中对地物细节的表达来看,Landsat融合后的15 m数据与后者仍有较大差距。在15~10 m空间分辨率范围,CNN开始区分出丰富的地物细节,Sentinel-2A的10 m多光谱数据因为处在该区间,一些较小的地物虽然能够被CNN识别,但由于特征不足够明显,一些地物的分类精度出现了明显下降(如草地、建筑物之间的小块草地开始被识别且精度不高),同时部分地物受影像空间分辨率升高带来的正向影响,分类精度有所提高,总体来看Sentinel-2A影像的分类精度波动幅度较大(图7)。

综合5种不同空间分辨率影像的分类结果,CNN对光谱特征较为单一的水库坑塘分类效果较差,这与传统的机器学习分类方法有较大差别。研究区内水库坑塘的CNN分类结果生产者精度偏低,结合分类混淆矩阵发现水库坑塘比较容易被错分为林地和城镇建设用地。由于CNN主要依靠图像的上下文结构信息进行分类,水库坑塘这类质地较为均一的地物,邻域中像元的光谱差异很小,但靠近边缘部分的像元容易受到岸边其他地物(主要为林地和城镇建设用地)的光谱影响,同时CNN模型使用了最大池化算法和Relu函数,二者均偏向于学习像元值较大的像素特征,容易对光谱吸收特征明显的水体形成抑制作用^[7],因此在水库坑塘的边缘部分会出现一定的错分现象。

CNN基于影像的上下文结构信息进行分类,一方面能够有效提高像元识别的准确率,另一方面对于边界部分的混合像元来讲,更容易受到其他地物的影响,造成CNN对地物边缘分类结果不准确。图8中将5种不同空间分辨率的影像进行分类细节展示,道路边缘部分的像元由于同时受到道路和林地的影响,被错分为了草地。同时由于顾及了影像上下文信息,CNN对较小的地物不敏感,在空间分辨率低于6 m的情况下,均未能提取出影像中的完整环路;当影像分辨率达到6 m时,CNN提取出了较为完整的环路轮廓;空间分辨率为4 m的GF-2影像上,CNN虽然完整提取出了该环路,但对于宽度为2~3个像元的道路,CNN仍未能完全识别。

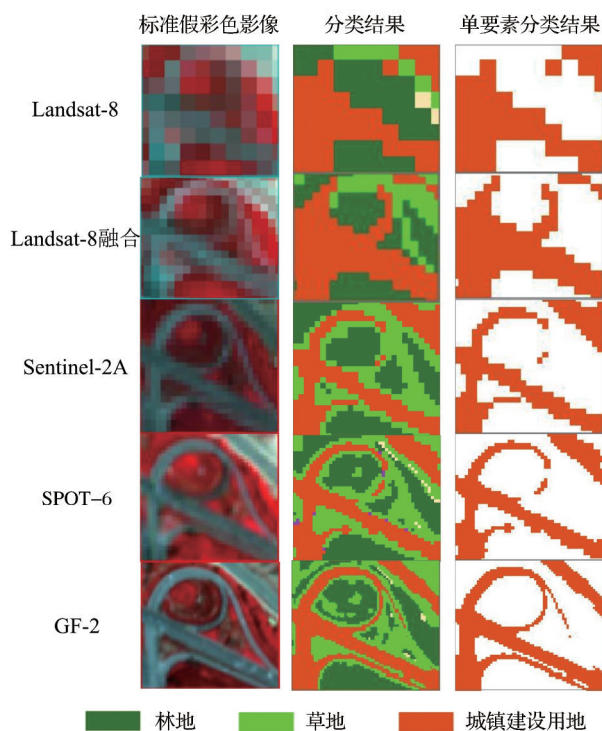


图 8 不同空间分辨率影像 CNN 分类细节对比

Fig.8 Comparison of CNN classification details for different spatial resolution images

5 结 语

研究针对 5 种不同空间分辨率的影像数据,设计并搭建了 CNN 模型分类框架,采用了窗口滑动辅助选样方法,能够在仅更改少量参数的情况下,为不同的遥感影像选出足量且满足大小要求的训练样本。就 5 种数据的分类精度来看,该方法可行的,为后期 CNN 在遥感影像分类中的样本选择提供了一种思路。

在不同空间分辨率下, CNN 均取得了高于 89.00% 的总体分类精度和 0.86 以上的 Kappa 系数,分类精度较高。但 CNN 对地物的边缘部分混合像元的分类效果较差。同时对斑块较小的地物不敏感,难以对其进行正确分类,若要提取地物的精细结构,有赖于影像空间分辨率的进一步提高。

在所涉及的空间分辨率范围(4~30 m)内,考虑到总体分类精度、对地物细节信息的反映及不同地物分类精度分布的集中程度等多方面因素,本文结果表明 CNN 对 SPOT-6 6 m 多光谱数据进行土地覆被分类的效果最佳。当影像的空间分辨率进一步提高时,地物呈现的规律性逐渐降低,表现出较高类内变异性 and 低类间差异性,异质性对 CNN 分类精度的影响超过了影像分辨率提高带来的正向

收益,增加了正确分类的困难程度,总体分类精度及 Kappa 系数反而出现了下降的趋势。

本文仍存在一些不足之处:一是针对采用的窗口滑动辅助选样方法没有量化其对分类结果的影响,仅通过最终的分类结果来间接验证了该方法的可行性;二是 CNN 分类结果对地物细节的反映程度未能使用一个具体的指标来衡量,仅通过直观感受对其进行了分析;三是在研究过程中未对地物的阴影进行单独处理,而是简单地将其归入了建设用地之中,可能由此产生一些错分误分现象,今后将考虑添加阴影去除的操作,以减小其对分类精度的影响。

参考文献 (References):

- [1] Tian Fuyou, Wu Bingfang, Zeng Hongwei, *et al.* Identifying Soybean Cropped Area with Sentinel-2 Data and Multi-layer Neural Network [J]. *Journal of Geoinformation Science*, 2019, 21(6): 918-927. [田富有, 吴炳方, 曾红伟, 等. 基于多层神经网络与 Sentinel-2 数据的大豆种植区识别方法[J]. *地球信息科学学报*, 2019, 21(6): 918-927.]
- [2] Thanh H N T, Doan T M, Tomppo E, *et al.* Land Use/Land Cover Mapping Using Multitemporal Sentinel-2 Imagery and Four Classification Methods—A Case Study from Dak Nong, Vietnam [J]. *Remote Sensing*, 2020, 12(9): 1367. doi: 10.3390/rs12091367.
- [3] Gu Xiaotian, Gao Xiaohong, Ma Huijuan, *et al.* Comparison of Machine Learning Methods for Land Use/Land Cover Classification in the Complicated Terrain Regions [J]. *Remote Sensing Technology and Application*, 2019, 34(1): 57-67. [谷晓天, 高小红, 马慧娟, 等. 复杂地形区土地利用/土地覆被分类机器学习方法比较研究[J]. *遥感技术与应用*, 2019, 34(1): 57-67.]
- [4] Li Yafei, Dong Hongbin. Classification of Remote-sensing Image based on Convolutional Neural Network [J]. *CAAI Transactions on Intelligent Systems*, 2018, 13(4): 550 - 556. [李亚飞, 董红斌. 基于卷积神经网络的遥感图像分类研究[J]. *智能系统学报*, 2018, 13(4): 550 - 556.]
- [5] Masoud M, Bahram S, Mohammad R, *et al.* Very Deep Convolutional Neural Networks for Complex Land Cover Mapping Using Multispectral Remote Sensing Imagery [J]. *Remote Sensing*, 2018, 10(7): 1119. doi: 10.3390/rs10071119.
- [6] Meng Xiangrui, Zhang Shuqing, Zang Shuying. Remote Sensing Classification of Wetland Communities based on Convolutional Neural Networks and High Resolution Images: A Case Study of the Honghe Wetland [J]. *Scientia Geographica Sinica*, 2018, 38(11): 1914-1923. [孟祥锐, 张树清, 臧淑英. 基于卷积神经网络和高分辨率影像的湿地群落遥感分类——以红河湿地为例[J]. *地理科学*, 2018, 38(11): 1914-1923.]
- [7] Zhang Wei, Zheng Ke, Tang Ping, *et al.* Land Cover Classifi-

- cation with Features Extracted by Deep Convolutional Neural Network[J]. Journal of Image and Graphics, 2017, 22(8): 1144-1153.[张伟,郑柯,唐娉,等.深度卷积神经网络特征提取用于地表覆盖分类初探[J].中国图象图形学报,2017,22(8):1144-1153.]
- [8] Cui Gang, Wu Jinsheng, Yu Zhen, *et al.* Comparison Analysis on Wheat Mapping Using Deep Learning Algorithm from Different Satellite Data Source[J]. Remote Sensing Technology and Application, 2019, 34(4): 748-755.[崔刚,吴金胜,于镇,等.深度学习对不同分辨率影像冬小麦识别的适用性研究[J].遥感技术与应用,2019,34(4):748-755.]
- [9] Maggiori E, Tarabalka Y, Charpiat G, *et al.* Fully Convolutional Neural Networks for Remote Sensing Image Classification[C]//2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2016.
- [10] Yang Rui, Qi Yuan, Su Yang. U-Net Neural Networks and Its Application in High Resolution Satellite Image Classification[J]. Remote Sensing Technology and Application, 2019, 34(6): 1-8.[杨瑞,祁元,苏阳.深度学习U-Net方法及其在高分辨卫星影像分类中的应用[J].遥感技术与应用,2019,34(6):1-8.]
- [11] Liu Xuemei, Gao Xiaohong, Jia Wei. Spatio Temporal Change of Urban Heat Island Effect in Xining from Landsat Image[J]. Journal of Earth Environment, 2019, 10(2): 166-176.[刘雪梅,高小红,贾伟.基于遥感影像的近30年西宁市热岛效应时空变化[J].地球环境学报,2019,10(2):166-176.]
- [12] Wang Can, Wu Xinhui, Li Lianqing, *et al.* Convolutional Neural Network in Prediction of Soil Moisture Content[J]. Spectroscopy and Spectral Analysis, 2018, 38(1): 36-41.[王璨,武新慧,李恋卿,等.卷积神经网络用于近红外光谱预测土壤含水率[J].光谱学与光谱分析,2018,38(1):36-41.]
- [13] Tian Ying, Chen Zhuoqi, Hui Fengming, *et al.* ESA Sentinel-2A/B Satellite: Characteristics and Applications[J]. Journal of Beijing Normal University (Natural Science Edition), 2019, 55(1): 57-65.[田颖,陈卓奇,惠凤鸣,等.欧空局哨兵卫星Sentinel-2A/B数据特征及应用前景分析[J].北京师范大学学报(自然科学版),2019,55(1):57-65.]
- [14] Dong Yunya, Zhang Qian. A Survey of Depth Semantic Feature Extraction of High-resolution Remote Sensing Images based on CNN[J]. Remote Sensing Technology and Application, 2019, 34(1): 1-11.[董蕴雅,张倩.基于CNN的高分遥感影像深度语义特征提取研究综述[J].遥感技术与应用,2019,34(1):1-11.]
- [15] Zhou Feiyan, Jin Linpeng, Dong Jun. Review of Convolutional Neural Network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.[周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.]
- [16] Huang H, Sun G Y, Zhang X M, *et al.* Combined Multiscale Segmentation Convolutional Neural Network for Rapid Damage Mapping from Postearthquake Very High-resolution Images[J]. Journal of Applied Remote Sensing, 2019, 13(2): 022007.doi:10.1117/1.JRS.13.022007.
- [17] Alshehhi R, Marpu P R, Woon W L, *et al.* Simultaneous Extraction of Roads and Buildings in Remote Sensing Imagery with Convolutional Neural Networks[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130: 139-149.
- [18] Scherer D, Müller A, Behnke S. Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition[C]//International Conference on Artificial Neural Networks. 2010.
- [19] Xu L, Choy C S, Li Y W. Deep Sparse Rectifier Neural Network for Speech Denoising[C]//2016 IEEE International Workshop on Acoustic Signal Enhancement (IWAENC). IEEE, 2016.
- [20] Shu Na, Liu Bo, Lin Weiwei, *et al.* Survey of Distributed Machine Learning Platforms and Algorithms[J]. Computer Science, 2019, 46(3): 9-18.[舒娜,刘波,林伟伟,等.分布式机器学习平台与算法综述[J].计算机科学,2019,46(3):9-18.]
- [21] Zhu Xiufang, Pan Yaozhong, Zhang Jinshui, *et al.* The Effects of Training Samples on the Wheat Planting Area Measurement Accuracy in TM Scale(1): The Accuracy Response of Different Classifiers to Training Samples[J]. Journal of Remote Sensing, 2007, 11(6): 826-837.[朱秀芳,潘耀忠,张锦水,等.训练样本对TM尺度小麦种植面积测量精度影响研究(1)——训练样本与分类方法间分类精度响应关系研究[J].遥感学报,2007,11(6):826-837.]
- [22] Nogueira K, Penatti Otávio A B, Santos J A D. Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification[J]. Pattern Recognition, 2017, 61: 539-556. doi:10.1016/j.patcog.2016.07.001.
- [23] McFEETERS S. K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features[J]. International Journal of Remote Sensing, 1996, 17(7): 1425-1432.
- [24] Ming Dongping, Wang Qun, Yang Jianyu. Spatial Scale of Remote Sensing Image and Selection of Optimal Spatial Resolution[J]. Journal of Remote Sensing, 2008, 12(4): 529-537.[明冬萍,王群,杨建宇.遥感影像空间尺度特性与最佳空间分辨率选择[J].遥感学报,2008,12(4):529-537.]
- [25] Chen J, Deng M, Xiao P F, *et al.* Optimal Spatial Scale Choosing for High Resolution Imagery based on Texture Features Frequency Analysis[J]. Journal of Remote Sensing, 2011, 15(3): 492-511.
- [26] Yang Yanjun, Tian Qingjiu, Zhan Yulin, *et al.* Effects of Spatial Resolution and Texture Features on Multi-spectral Remote Sensing Classification[J]. Journal of Geo-information Science, 2018, 20(1): 99-107.[杨闫君,田庆久,占玉林,等.空间分辨率与纹理特征对多光谱遥感分类的影响[J].地球信息科学学报,2018,20(1):99-107.]

Land Cover Classification for Different Spatial Resolution Images from CNN

Li Hongda, Gao Xiaohong, Tang Min

(School of Geographical Sciences, Qinghai Normal University, Qinghai Province Key Laboratory of Physical Geography and Environmental Process, Academy of Plateau Science and Sustainability, MOE Key Laboratory of Tibetan Plateau Land Surface Processes and Ecological Conservation, Xi'ning 810008, China)

Abstract: Based on convolutional neural networks and five different spatial resolution remote sensing images, the land use/land cover classification study was carried out on a small area in the eastern part of Xining City, aiming at exploring the differences of image classification by CNN with different spatial resolutions and CNN's ability to extract different features. In order to improve the selection efficiency of the samples, a window sliding method was introduced to assist the samples selection. The research shows that the overall classification accuracy of the five different spatial resolution images is above 89%, the Kappa coefficient is above 0.86. The result further shows that within the resolution scale the higher the resolution, the performance of the CNN classification results for the details is better, and can maintain high classification accuracy, indicating that CNN is more suitable for high spatial resolution images; at the same time, the image spatial resolution is too high, the ground objects exhibit high intra-class variability and low inter-class variability, the classification accuracy tends to decrease. In comparison, CNN has the best classification effect on SPOT 6 images in this study, and window sliding is an effective sample-assisted selection method. This research has certain reference significance for similar research in the future.

Key words: Convolutional Neural Network; Landsat-8/Sentinel-2A/SPOT-6/GF-2 images; Land cover classification