doi: 10.6046/gtzyyg.2018.01.04
引用格式: 金永涛,杨秀峰,高涛,等. 基于面向对象与深度学习的典型地物提取[J]. 国土资源遥感,2018,30(1):22-29. (Jin Y T, Yang X F, Gao T, et al. The typical object extraction method based on object-oriented and deep learning[J]. Remote Sensing for Land and Resources, 2018, 30(1):22-29.)

基于面向对象与深度学习的典型地物提取

金永涛[1,3], 杨秀峰[1,3], 高 涛[2], 郭会敏[2], 刘世盟[1]
(1. 北华航天工业学院,廊坊 065000; 2. 航天恒星科技有限公司,北京 100086;
3. 河北省航天遥感信息处理与应用协同创新中心,廊坊 065000)

摘要: 针对遥感图像地物分割问题面向对象方法可以将不同地物分割到不同的对象之中,在很大程度上解决了农作物、林地、水体、道路、建筑物等典型地物的混分问题,但面向对象方法对于形状、纹理等特征描述仍不够全面,信息量还不足以支撑完整的地物分类、识别。提出一种将面向对象与深度学习相结合的新方法,选用卷积神经网络 Caffe 框架,对训练样本数据进行深度学习,掌握不同对象的纹理等特性,形成深度学习模型,反过来指导对象分类。实验表明,新方法可以有效解决典型地物分不准的问题。

关键词: 面向对象; 深度学习; 卷积神经网络; 目标识别
中图法分类号: TP 79 **文献标志码**: A **文章编号**: 1001-070X(2018)01-0022-08

0 引言

遥感图像解译有两大难点,即不同地物难以分割且地物类型难以精准识别。目前,遥感影像的分类方法按其分类基本单元的不同可分为基于像元的分类方法和面向对象的分类方法 2 大类[1]。基于像元的分类方法以影像单一像元为基本单元,主要利用影像的光谱特征进行分类[2],分类结果容易出现椒盐现象,而且由于影像存在"同物异谱"或"异物同谱"现象,易造成地物类别的错分和漏分,导致分类结果精度较低。面向对象的分类方法是利用"同质均一"的多个像元为基础分类对象,利用影像的光谱和空间纹理特征,突破了基于像元分类方法的限制[3],很大程度上解决了农作物、林地、水体、道路及建筑物等典型地物区分不开的问题。

深度学习是计算机科学机器学习领域中一个新的研究方向。其概念源于人工神经网络的研究,含多隐层的多层感知器就是一种深度学习结构[4-7]。Lecun 等在 1998 年提出的卷积神经网络是第一个真正的多层结构学习算法,它利用空间相对关系减少参数数目以提高样本训练性能[8-9]。深度学习可通过学习一种深层非线性网络结构,实现复杂函数的逼近,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[10-13]。

本文结合面向对象与深度学习的技术特点提出了一种新的典型地物信息提取方法。首先,采用面向对象方法提取典型地物对象特征; 然后,选用卷积神经网络 Caffe 框架对分割后的不同尺度对象进行深度学习,获取不同对象的形状和纹理特征,用以指导对象分类; 最后,有效解决典型地物信息提取分不准的问题。

1 研究区概况及数据源

本研究选择廊坊市永清县龙虎庄乡为试验区。该区位于 E116°22′~116°36′, N39°11′~39°23′之间,主要土地利用类型为农作物、林地、水体、道路和建筑物等。选取了 2016 年 8 月 27 日 2 景国产高分二号(GF-2)卫星晴空影像数据,首先,对数据进行了几何精校正处理,将各波段数据进行影像的自动配准和镶嵌; 然后,对全色影像和多光谱(红、绿、蓝波段)影像采用 ENVI Gram-Schmiddt Spectral Sharpening 方法进行了融合,使多光谱影像具备了

收稿日期: 2017-06-03; **修订日期**: 2017-07-27
基金项目: 国家高分辨率对地观测系统重大专项"基于高分数据的京津冀一体化协同发展区域监测技术"(编号: 67-Y20A07-9002-16/17)、河北省科技计划"基于北斗系统和高分遥感数据的县域集成信息服务平台关键技术研究"(编号: 16210310D)和河北省军民结合产业发展专项资金"协同创新中心—基于国产高分数据的县域农业遥感监测与分析"(编号: JMJHZX-2016-01)共同资助。
第一作者: 金永涛(1978-),男,副教授,主要从事遥感技术与应用、计算机网络与信息安全方面的研究。Email: jsj_jin@126.com。

丰富的纹理特征,最后通过矢量裁剪出试验区域,结果如图 1 所示。

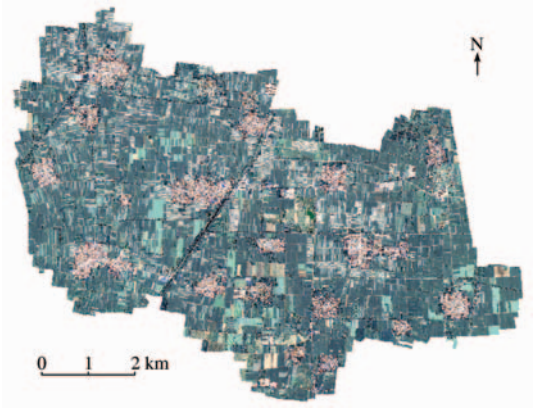


图 1　选取的实验样本区域 GF－2 影像

Fig. 1　GF－2 image of the selected sample area

2　研究方法

2.1　技术流程

传统的面向对象方法根据不同类对象在特征上表现出来的差异建立模糊规则。对象特征包括颜色、光谱特征(如亮度值、归一化水体指数(normalized difference water index, NDWI)和归一化植被指数(normalized difference vegetation index, NDVI))、形状纹理(如边界指数、紧致度、长宽比)等,但对于特征描述得不够全面、准确。深度学习方法可以通过样本训练,掌握不同地物的形状、纹理、背景等特性,但也需要大量的标签,人工标识工作量大。表 1 对 2 种方法的优缺点进行了对比分析。典型地物目标识别技术流程见图 2。

表 1　面向对象与深度学习方法优缺点对比分析

Tab. 1　Advantages and disadvantages of object－oriented and in－depth learning methods

分类方法	优点	缺点	两者结合点
面向对象方法	将不同的地物分割到不同的对象之中,对分类起了至关重要的作用,可以建立各种对象特征规则集来提取地物	对于形状、纹理特征描述得不够全面、准确,信息量还不足以支撑地物分类和识别	通过面向对象构建特征规则集可以进行前期的地物提取,并由地物对象构建深度学习所需要的数据训练集;由于深度学习往往都是通过 RGB 图像进行训练,没充分考虑遥感图像多波段的特性,结果会有错误和偏差,因此,可以在面向对象通过地物特征集中进行修正
深度学习方法	可以掌握不同地物的形状、纹理、背景等特性,比较准确地区分不同地物,并按照训练结果来分割地物	一方面对于图像分类模型需要大量的数据来训练,特别对于图像分割模型来说,需要大量的标签来标识不同地物,人工标识的话工作量太大;另一方面,深度学习的结果往往是栅格图像,最终结果很难修正	

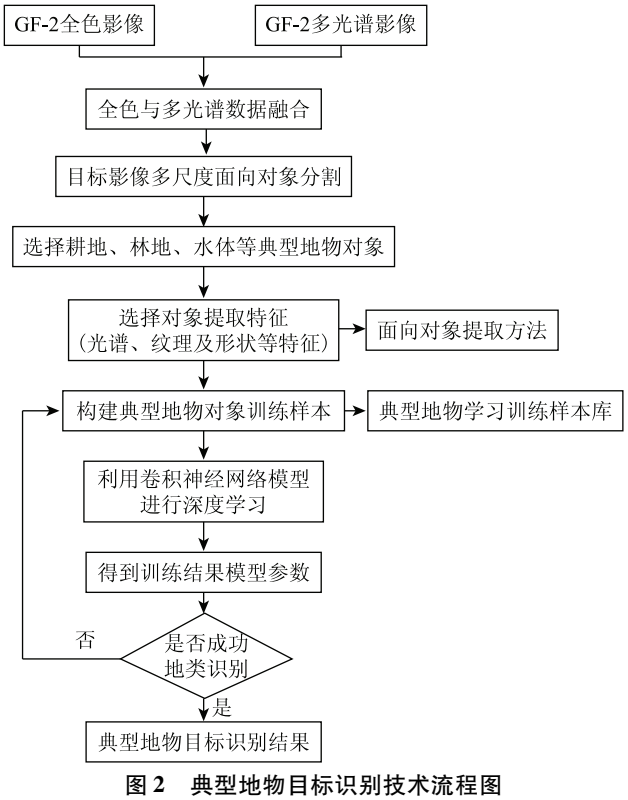


图 2　典型地物目标识别技术流程图

Fig. 2　The flow chart of typical objects recognition technology

在表 1 基础上,对 GF－2 影像全色波段和多光谱波段进行融合处理后,首先采用多尺度分割算法对试验区影像进行图像分割;然后在面向对象分割结果的基础上,构建耕地、林地、水体、道路、建筑物等典型地物对象特征规则集,对分割对象进行分类提取,得到训练样本,形成典型地物学习样本库;之后根据训练样本进行深度学习,得到训练结果,并进一步对训练样本进行分类;最后对提取的典型地物目标结果进行精度评价与分析。典型地物目标识别技术流程图如图 2 所示。

2.2　多尺度面向对象分割算法

多尺度面向对象分割算法将图像看作是一个由区域和区域之间的拓扑关系组成的一张区域邻接图[14],根据指定尺度进行分割,采用从单像元大小的区域开始,相邻影像区域两两合并增长的方法,设定阈值控制合并区域,保证生成高度同质性(或异质性最小)的影像分割区域(影像对象)[15],从而适于最佳分离和表示地物目标。

算法的基本思想是:从单一像元开始,与邻近像元分别进行差异性度量计算,降低异质性,完成一轮合并后,以上一轮生成的对象为基本单元,继续与

相邻对象分别进行计算，直到在规定的尺度上已经不能再进行任何对象的合并为止。这种异质性是由对象之间的光谱和几何形状差异决定的。初始时，每个像素作为一个区域，差异性度量准则(f)的计算公式为

$$f = w_1 x + (1 - w_1) y \quad , \tag{1}$$

式中：w_1 为权值，$0 \leqslant w_1 \leqslant 1$；$x$ 为光谱异质性；y 为形状异质性；x 与 y 的计算采用

$$x = \sum_{i=1}^{n} p_i \sigma_i \quad , \tag{2}$$

$$y = w_2 u + (1 - w_2) v \quad , \tag{3}$$

式中：p_i 为第 i 影像层的权值；σ_i 为第 i 影像层光谱值的标准差；u 为影像区域整体紧密度；v 为影像区域边界平滑度；w_2 为权值，$0 \leqslant w_2 \leqslant 1$。$u$，$v$ 的计算式为

$$u = E / \sqrt{N} \quad , \tag{4}$$

$$v = E / L \quad , \tag{5}$$

式中：E 为影像区域实际的边界长度；N 为影像区域的像元总数；L 为包含影像区域范围的矩形边界总长度。

在影像区域的合并过程中，从区域邻接图中的每个区域入手，寻找满足局部最优合并条件的区域对，将这 2 个区域合并，并更新与原来 2 个区域相连的所有区域的特征值及其与新区域的合并代价。当合并相邻的 2 个区域时，合并新生成的更大影像区域对象的异质性 f' 的计算式为

$$f' = w_1 x' + (1 - w_1) y' \quad , \tag{6}$$

式中：x'，y' 分别为合并新生成的更大影像区域的光谱异质性和形状异质性，即

$$x' = \sum_{i=1}^{n} p_i [N' \sigma_i' - (N_1 \sigma_{i1} + N_2 \sigma_{i2})] \quad , \tag{7}$$

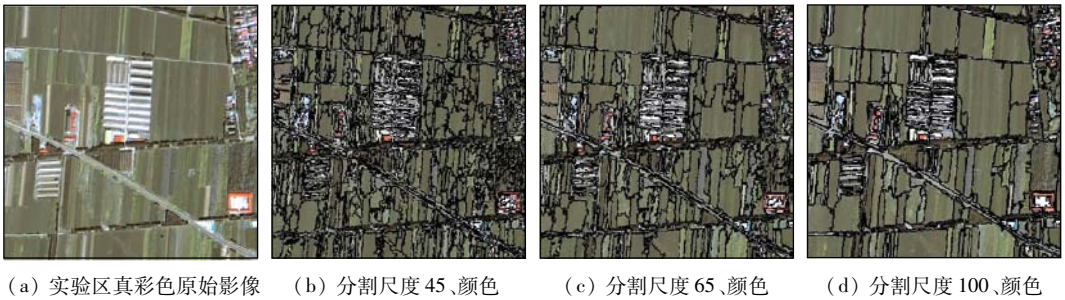
$$y' = w_2 u' + (1 - w_2) v' \quad , \tag{8}$$

式中：N' 和 σ_i' 分别为合并新生成的更大影像区域的像元总数和其所在 i 影像层光谱值的标准差；N_1 和 σ_{i1} 分别为合并前相邻影像区域 1 的像元总数和其所在 i 影像层光谱值的标准差；N_2 和 σ_{i2} 分别为合并前相邻影像区域 2 的像元总数和其所在 i 影像层光谱值的标准差。u' 和 v' 的计算公式分别为

$$u' = N' \frac{E'}{\sqrt{N'}} - \left(N_1 \frac{E_1}{\sqrt{N_1}} + N_2 \frac{E_2}{\sqrt{N_2}} \right) \quad , \tag{9}$$

$$v' = N' \frac{N'}{L'} - \left(N_1 \frac{E_1}{L_1} + N_2 \frac{E_2}{L_2} \right) \quad , \tag{10}$$

式中：E' 和 L' 分别为合并新生成的更大影像区域的实际边界长度和包含该新生成影像区域范围的矩形边界总长度；E_1 和 L_1 分别为合并前的相邻影像区域 1 的实际边界长度和包含该影像区域范围的矩形边界总长度；E_2 和 L_2 分别为合并前的相邻影像区域 2 的实际边界长度和包含该影像区域范围的矩形边界总长度。不同尺度的地类分割的结果如图 3 所示。



（a）实验区真彩色原始影像　（b）分割尺度 45、颜色 0.7、平滑度 0.5　（c）分割尺度 65、颜色 0.7、平滑度 0.5　（d）分割尺度 100、颜色 0.7、平滑度 0.5

图 3　多尺度分割结果示意图

Fig. 3　Schematic diagram of multiscale segmentation

2.3　基于多尺度规则集的样本库构建

结合试验区地表下垫面类型，通过对象的特征信息与地物之间的对应关系，建立样本分类选取的层次结构，即利用多尺度的分层分割，不同的地物采取不同的尺度进行分割；然后，根据对象的光谱特征、几何纹理特征以及拓扑结构特征设置分类规则。建筑物、道路、水体和植被（包括农作物、林地）大类在大尺度分割层（分割尺度 100、颜色 0.7、平滑度 0.5）中，分别利用亮度值、NDWI 和 NDVI 作为判断依据进行初步提取；再在植被大类的影像对象上选取适合其子类的分割尺度（分割尺度 65、颜色 0.7、平滑度 0.5）来进行分割，考虑耕地的形状比林地规则，综合两者对象的边界指数、紧致度、长宽比等形状指数进行区分，具体的对象规则集如表 2 所示。

表 2　对象规则集

Tab. 2　The rule set of objects

序号	层次	类别	判断依据	规则	备注
1	大尺度分割层(分割尺度 100,颜色 0.7,平滑度 0.5)	建筑物	亮度值	亮度 > 398	建筑物光谱较为高亮,方差较大
2		道路	宽度和长宽比	宽度 < 17.8,长宽比 > 12.5	呈条带状,长宽比较大
3		水体	NDWI	$NDWI > 0.5$	NDWI 用来提取影像中的水体信息,效果较好,$NDWI = (B2 - B4)/(B4 + B2)$①
4		植被	NDVI	$NDVI > 0$	NDVI 是植被生长状态及植被覆盖度的最佳指示因子。$NDVI = (B4 - B3)/(B4 + B3)$
5	小尺度分割层(分割尺度 65,颜色 0.7,平滑度 0.5)	农作物	提取于植被,形状规则	边界指数 > 1.5,同质性 < 12	考虑内部父对象与子对象关系、对象的边界指数、紧致度、同质性等
6		林地	提取于植被,颜色更绿	$NDVI > 0.45$,同质性 > 32	

①式中 $B2$,$B3$,$B4$ 分别为 2,3,4 波段亮度值。

基于对象判定规则,通过程序自动追踪各个类别下的样本边界,建立了包括耕地、林地、水体、道路及建筑物等主要典型地物信息遥感影像特征集,获取训练样本集。如表 3 所示。

表 3　典型地类训练样本

Tab. 3　The training sample of typical objects

序号	样本名称	样本	样本数
1	建筑物		612
2	水体		202
3	道路		645
4	农作物		700
5	林地		687

2.4　基于深度学习的典型地物自动识别

利用实验区各个类别的样本图片作为训练数据,将深度学习与面向对象的方法相结合,选用卷积神经网络模型,对训练样本数据进行深度学习,自动获取样本特征,利用面向对象分割结果实现典型地物的自动识别分类。其中,卷积神经网络的结构设计是问题关键,本文的深度学习框架采用 Caffe 框架,用到的卷积神经网络结构如图 4 所示。

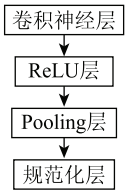


图 4 卷积神经网络结构

Fig. 4 Convolution neutral network structure

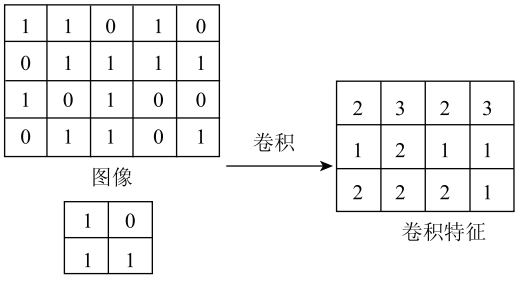


图 5 图像卷积过程示意图

Fig. 5 Imagine convolution process diagram

2.4.1 卷积神经层

卷积神经层是对图像的每个像素点进行卷积运算,卷积核作为训练参数,经过几次处理之后,能够提取出图像的"特征值"。卷积神经层中,卷积核越大,对图像"抽象"的效果越好,但需要训练的参数就越多; 卷积核越小,越能够精细地处理图像,但要达到同样的"抽象"效果,需要更多的层数[16]。图像卷积过程如图 5 所示。在经典的神经网络结构中,采用 11 ×11 的卷积核。为达到预期结果,需引入大量参数,在不考虑偏置的情况下,11 ×11 的卷积核参数是 3 ×3 卷积核的(11 ×11 +1)/(3 ×3 +1) = 12.2 倍,且每个输出点对应 11 ×11 次乘法和 11 × 11 次加法,计算量大,从而导致算法性能的降低,因此,本文采用 3 ×3 和 5 ×5 较小的卷积核。

2.4.2 ReLU 层

深度卷积网络一般都需要大量的数据进行训练,如果使用传统的激活函数(如 sigmoid 函数和 tanh 函数),数据计算量大,几乎不能够配合卷积神经层完成训练,因此,本文采用了线性激活函数(Rectified Linear Units, ReLU),不仅减少了训练时间,而且提高了算法性能。在 ReLU 中,对于给定的一个输入值 x,如果 $x > 0$,ReLU 层的输出为 x; 如果 $x < 0$,ReLU 层的输出为 0。sigmoid 函数、tanh 函数、ReLU 函数公式分别为:

$$\text{sigmoid 函数: } f(x) = 1/[1+\exp(-1)]\,g'(x) = [1-g(x)]g(x) \text{。} \tag{11}$$

$$\text{tanh 函数: } f(x) = \sinh(x)/\cosh(x) = [\exp(x) - \exp(-x)]/[\exp(x) + \exp(-x)] \text{。} \tag{12}$$

$$\text{Rectified 函数: } f(x) = \max(0,x) \text{。} \tag{13}$$

2.4.3 池化层

输入图像经过卷积神经层和 ReLU 处理之后,图像中的每个像素点都包含了相邻区域的信息,造成了信息冗余,继续计算不仅会降低算法性能,还会破坏算法的平移不变性。为了提升算法的性能和鲁棒性,需要对图像进行二次采样[17],在深度卷积网络中,这种操作称为池化(pooling),即将图像分成一小块一小块的区域,对每个区域计算出一个值,然后将计算出的值依次排列,输出为新的图像。如果划分的区域之间互不重叠,其算法称为非重叠型池化,否则称为重叠性池化。对每个区域计算输出的方法也分为 2 种: 求平均值(mean pooling)或者取最大值(max pooling)。本文采用可重叠的、取最大值的池化运算,可以在一定程度上降低过度拟合。池化计算过程如图 6 所示。

2.4.4 规范化神经层

为了让图像更加具有对比性,需要设计规范化神经层。规范化神经层的作用类似于对图像进行"增加对比度"的操作。本文选用局部响应归一化(local response normalization, LRN)算法,选择通道内空间区域归一化,局部区域在空间上扩展,将每个输入值都除以

$$(1 + (\alpha/n)\sum\nolimits_i x_i^2)^{\beta} \text{,} \tag{14}$$

式中,α 为缩放因子,默认值为 1; β 为指数项,默认值为 5; n 为局部尺寸大小,默认值为 5。如此完成"临近抑制"操作后,可有效提高主体部分与其他部分的区分度。

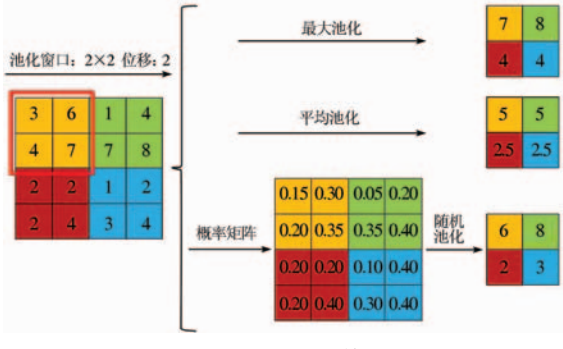


图 6 池化运算

Fig. 6 Pooling operation

3 结果与分析

根据龙虎庄乡试验区下垫面类型特征,将影像分类为建筑物、道路、水体、林地、农作物及其他 6 种类别。基于面向对象和结合深度学习模型的面向对

象分类方法的典型地区提取结果如图 7 和 8 所示。

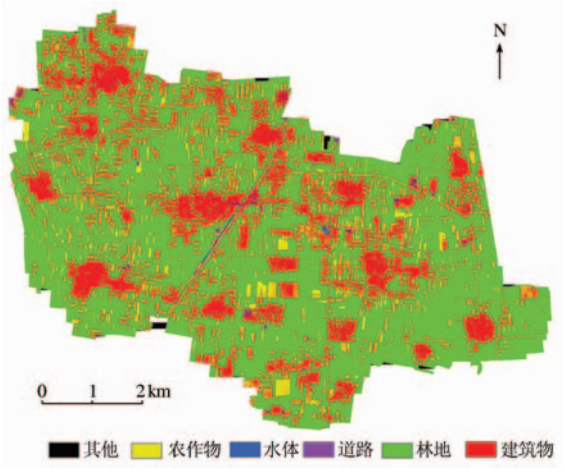


图 7　基于面向对象分类方法的典型龙虎庄乡地物提取结果图
Fig. 7　Typical class selection result of Longhuzhuang township by Object – Oriented method

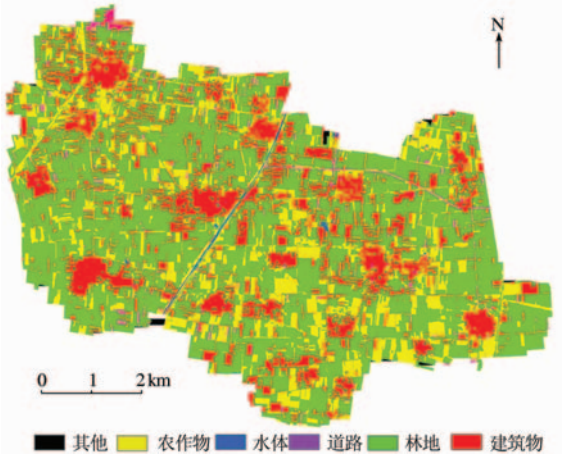


图 8　结合深度学习模型的龙虎庄乡典型地物提取结果图
Fig. 8　Typical class selection result of Longhuzhuang township by deep learning method

本文采用随机抽样方法进行精度评价[18],在龙虎庄乡试验区 GF – 2 卫星融合影像中随机选择了 300 个样本点像元,以目视解译结果为标准进行评价。随机点分布效果如图9所示。然后,通过混淆

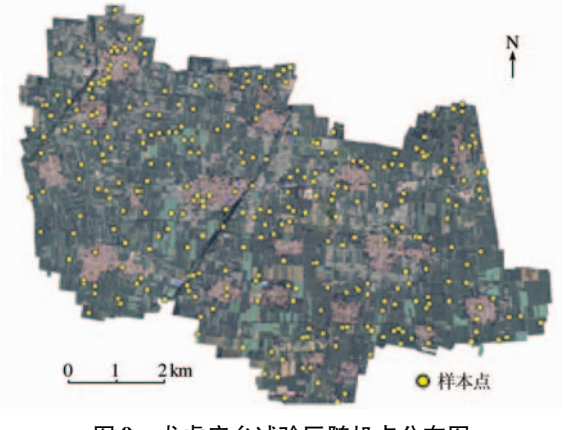


图 9　龙虎庄乡试验区随机点分布图
Fig. 9　Random point distribution of Longhuzhuang township

矩阵计算出用户精度、制图精度、错分误差、漏分误差、总体精度、Kappa 系数等精度指标,对分类结果进行精度评价,计算结果如表 4—表 7 所示。

表 4　基于面向对象分类方法的典型地物分类混淆矩阵
Tab. 4　Confusion matrix of typical classification by object – oriented method　（个）

地物类别	建筑物	农作物	林地	水体	道路	其他	分类总数
建筑物	32	1	3	1	4	1	42
农作物	2	65	22	0	0	0	89
林地	5	20	105	1	1	0	132
水体	1	0	1	12	1	0	15
道路	5	1	1	2	9	0	18
其他	0	0	0	0	1	3	4
实际总数	45	87	132	16	16	4	300

表 5　基于面向对象分类方法的地物分类错分误分、漏分误分、制图精度与用户精度
Tab. 5　Commission and omission error, production and user precision by object – oriented method　（%）

地物类别	错分误差	漏分误差	制图精度	用户精度
建筑物	23.81	28.89	71.11	76.19
农作物	26.97	25.29	74.71	73.03
林地	20.45	20.45	79.55	79.55
水体	20.00	25.00	75.00	80.00
道路	50.00	43.75	56.25	50.00
其他	25.00	25.00	75.00	75.00

表 6　结合深度学习模型分类方法的典型地物分类混淆矩阵
Tab. 6　Confusion matrix of typical classification by deep learning method　（个）

地物类别	建筑物	农作物	林地	水体	道路	其他	分类总数
建筑物	38	0	2	0	2	0	42
农作物	0	69	5	1	1	0	76
林地	5	18	125	0	1	1	150
水体	1	0	0	13	0	0	14
道路	1	0	0	1	12	0	14
其他	0	0	0	1	0	3	4
实际总数	45	87	132	16	16	4	300

表 7　结合深度学习模型分类方法的地物分类错分误分、漏分误分、制图精度与用户精度
Tab. 7　Commission and omission error, production and user precision by deep learning method　（%）

地物类别	错分误差	漏分误差	制图精度	用户精度
建筑物	9.52	15.56	84.44	90.48
农作物	9.21	20.69	79.31	90.79
林地	16.67	5.30	94.70	83.33
水体	7.14	18.75	81.25	92.86
道路	14.29	25.00	75.00	85.71
其他	25.00	25.00	75.00	75.00

根据 2 种分类结果的混淆矩阵计算得到,基于面向对象分类方法的分类总精度为 75.33%,Kappa 系数为 0.96;结合深度学习模型面向对象分类方法

的分类总精度为 84.40%，Kappa 系数为 0.97，分类结果优于前者，主要是新提出的分类方法考虑利用了对象之间的相邻关系模型，可以很好地切割地物边界线，较好地划分了农作物和林地，建筑物和道路等易混淆地区，优于基于单一对象特征的分类方法。

4 结论

针对农作物、林地、水体、道路、建筑物等典型地物信息提取，提出了一种将面向对象与深度学习相结合的新方法。新方法将图像多尺度分割基础上构建通用规则集获取的典型地物样本特征为输入，然后通过深度学习方法进一步进行样本训练和特征提取，最后再将学习后的样本特征运用于分割结果完成遥感影像的信息提取。选取廊坊市永清县龙虎庄乡为试验区进行实验和精度验证，精度为 84.40%。实验结果表明，新方法将面向对象与深度学习进行了合理、高效的结合，有效解决了典型地物分不准的问题，分类精度高于面向对象分类方法。但是新方法在通用规则集构建和深度学习结构设计方面尚不完善，有待进一步改进。

参考文献(References)：

[1] 郑 毅，武法东，刘艳芳. 一种面向对象分类的特征分析方法[J]. 地理与地理信息科学，2010，26(2)：19－22.
Zheng Y，Wu F D，Liu Y F. A feature analysis approach for object－oriented classification[J]. Geography and Geo－Information Science，2010，26(2)：19－22.

[2] 赵英时. 遥感应用分析原理与方法[M]. 北京：科学出版社，2003.
Zhao Y S. Theory and Methods of Remote Sensing Application Analysis[M]. Beijing：Science Press，2003.

[3] 黄慧萍. 面向对象影像分析中的尺度问题研究[D]. 北京：中国科学院研究生院遥感应用研究所，2003.
Huang H P. Scale Issues in Object－oriented Image Analysis[D]. Beijing：Institute of Remote Sensing Applications Chinese Academy of Sciences，2003.

[4] Bengio Y. Learning deep architectures for AI[J]. Foundations and Trends® in Machine Learning，2009，2(1)：1－127.

[5] Deng L. An overview of deep－structured learning for information processing[C]//Proceedings of Asian－Pacific Signal & Information Processing Annual Summit & Conference (APSIPA－ASC). Xi'an：Chinese Information Processing Society of China，2011：301－313.

[6] Arel I，Rose D C，Karnowski T P. Deep machine learning－a new frontier in artificial intelligence research[research frontier][J]. IEEE Computational Intelligence Magazine，2010，5(4)：13－18.

[7] Do V H，Xiao X，Chng E S. Comparison and combination of multilayer perceptrons and deep belief networks in hybrid automatic speech recognition systems[C]//Proceedings of Asian－Pacific Signal & Information Processing Annual Summit & Conference (APSIPA－ASC). Xi'an：Chinese Information Processing Society of China，2011.

[8] Farabet C，Couprie C，Najman L，et al. Learning hierarchical features for scene labeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2013，35(8)：1915－1929.

[9] Kavukcuoglu K，Sermanet P，Boureau Y L，et al. Learning convolutional feature hierarchies for visual recognition[C]//Proceedings of the 23rd International Conference on Neural Information Processing Systems. Vancouver，British Columbia，Canada：ACM，2010：1090－1098.

[10] Charalampous K，Kostavelis I，Amanatiadis A，et al. Sparse deep－learning algorithm for recognition and categorisation[J]. Electronics Letters，2012，48(20)：1265－1266.

[11] Olshausen B A，Field D J. Emergence of simple－cell receptive field properties by learning a sparse code for natural images[J]. Nature，1996，381(6583)：607－609.

[12] Olshausen B A，Field D J. Sparse coding with an overcomplete basis set：A strategy employed by V1？[J]. Vision Research，1997，37(23)：3311－3325.

[13] Goodfellow I，Courville A，Bengio Y. Large－scale feature learning with spike－and－slab sparse coding[J]. arXiv preprint arXiv：1206.6407，2012.

[14] 王卫红，何 敏. 面向对象土地利用信息提取的多尺度分割[J]. 测绘科学，2011，36(4)：160－161.
Wang W H，He M. Multi－scale segmentation in land－use information extraction based on object－oriented method[J]. Science of Surveying and Mapping，2011，36(4)：160－161.

[15] Woodcock C E，Strahler A H. The factor of scale in remote sensing[J]. Remote Sensing of Environment，1987，21(3)：311－332.

[16] Alain G，Bengio Y，Rifai S. Regularized auto－encoders estimate local statistics[J]. arXiv：1211.4246，2012：1－17.

[17] Rifai S，Bengio Y，Dauphin Y，et al. A generative process for sampling contractive auto－encoders[J]. arXiv：1206.6434，2012：2.

[18] Zhan Q M，Molenaar M，Tempfli K，et al. Quality assessment for geo－spatial objects derived from remotely sensed data[J]. International Journal of Remote Sensing，2005，26(14)：2953－2974.

[19] 宫 鹏，黎 夏，徐 冰. 高分辨率影像解译理论与应用方法中的一些研究问题[J]. 遥感学报，2006，10(1)：1－5.
Gong P，Li X，Xu B. Interpretation theory and application method development for information extraction from high resolution remotely sensed data[J]. Journal of Remote Sensing，2006，10(1)：1－5.

[20] 曹 雪，柯长青. 基于对象级的高分辨率遥感影像分类研究[J]. 遥感信息，2006(5)：27－30，51.
Cao X，Ke C Q. Classification of high－resolution remote sensing images using object－oriented method[J]. Remote Sensing Information，2006(5)：27－30，51.

The typical object extraction method based on object – oriented and deep learning

JIN Yongtao[1,3], YANG Xiufeng[1,3], GAO Tao[2], GUO Huimin[2], LIU Shimeng[1]

(1. *North China Institute of Aerospace Engineering*, *Langfang* 065000, *China*; 2. *Space Star Technology Co.*, *LTD*, *Beijing* 100086, *China*; 3. *Collaborative Innovation Center of Aerospace Remote Sensing Information Processing and Application of Hebei Province*, *Langfang* 065000, *China*)

Abstract: The object – oriented method solves the problem of segmentation of objects, divides different features into different objects and to a great extent separates the cultivated land, forest land, water, roads, buildings and other typical objects which are inseparable; nevertheless, the object oriented method for features such as shape, texture description is not comprehensive, the amount of information is not enough to support the whole classification and recognition. In this paper, a new method of combining object – oriented and deep learning is proposed, in which the Caffe framework of convolution neural network is used to study the training sample data in depth and, by grasping the texture of different objects and forming deep learning model, guides the classification of objects. The experiment shows that the new method can effectively solve the problem of the low classification accuracy.

Keywords: object – oriented; deep learning; convolution neural network; object recognition

(责任编辑: 李 瑜)