

基于深度学习的高分辨率遥感影像分类研究

刘大伟^{1,2} 韩 玲¹ 韩晓勇¹

¹长安大学地测学院, 陕西 西安 710054

²武警工程大学信息工程系, 陕西 西安 710086

摘要 针对高空间分辨率遥感影像的分类问题,提出了基于深度学习的分类方法。该方法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,并与基于单源光谱信息的DBN分类方法、支持向量机(SVM)分类方法、传统神经网络(NN)分类方法进行了比较分析。研究表明:相对于单源光谱信息,利用影像的光谱-纹理特征能够有效提高高分辨率遥感影像的分类精度;相对于SVM、NN等分类方法,DBN能够更加准确地挖掘高分辨率遥感影像的空间分布规律,提高分类的准确度。

关键词 遥感;深度学习;深度信念网络;高空间分辨率;遥感影像分类;非下采样轮廓波变换;纹理

中图分类号 TP751.1 文献标识码 A

doi: 10.3788/AOS201636.0428001

High Spatial Resolution Remote Sensing Image Classification Based on Deep Learning

Liu Dawei^{1,2} Han Ling¹ Han Xiaoyong¹

¹School of Geology Engineering and Geomatics, Chang'an University, Xi'an, Shaanxi 710054, China

²Department of Information Engineering, Armed Police Engineering University, Xi'an, Shaanxi 710086, China

Abstract A classification method based on deep learning is proposed for the classification of high spatial resolution remote sensing images. The texture features of the images are calculated through nonsubsampling contourlet transform, the deep learning common model—deep belief networks (DBN) are used to classify the high spatial resolution remote sensing images based on spectral and texture features. The proposed method is compared with the DBN classification method based on single spectral information, the support vector machine (SVM) method and the traditional neural network (NN) classification method. Experimental results show that comparing with the single spectral information, the use of spectral and texture information can effectively improve the classification accuracy of high spatial resolution remote sensing images, and comparing with methods of SVM and NN, the DBN method can accurately explore the distribution law of the high spatial resolution remote sensing images and improve the accuracy of classification.

Key words remote sensing; deep learning; deep belief networks; high spatial resolution; remote sensing image classification; nonsubsampling contourlet transform; texture

OCIS codes 280.4750; 100.2960; 100.5010; 110.6980

1 引言

随着遥感技术的不断进步,可以获得的遥感影像的空间分辨率也越来越高。在高分辨率的影像上,地物的光谱特征更加丰富,同类地物内的光谱差异增大,类间的光谱差异减少,同物异谱及同谱异物现象更加

收稿日期: 2015-09-21; 收到修改稿日期: 2015-12-04

基金项目: 国家自然科学基金(41171224, 41301386)、中央高校基本科研业务费专项(310826161009)

作者简介: 刘大伟(1980—),男,博士研究生,主要从事遥感图像处理方面的研究。E-mail: wjmicheal@163.com。

导师简介: 韩 玲(1964—),女,教授,博士生导师,主要从事遥感图像处理方面的研究。

E-mail: hanling@chd.edu.cn(通信联系人)

普遍。影像中大量细节的出现和地物光谱特征的复杂化导致了基于光谱统计特征进行分类的传统方法如极大似然法、最小距离法、K-均值聚类法等分类准确性的降低^[1]。基于此,研究者将机器学习算法如神经网络(NN)、支持向量机(SVM)等应用于高分辨率影像的分类,并且在分类过程中加入影像的纹理、结构等特征^[2-4],研究表明,相对于传统的统计方法,基于机器学习算法的影像多元特征分类能够获得更优的结果。然而无论是SVM还是NN都属于浅层学习算法^[5],由于计算单元有限,浅层学习的网络很难有效地表达复杂函数,所以随着样本数量的增大以及样本多样性增强,浅层模型也逐渐不能适应复杂的样本。而由多层非线性映射层组成的深度学习网络拥有强大的函数表达能力,在复杂分类上具有很好的效果和效率^[6]。

深度学习可以看作是神经网络的发展,它模拟大脑的学习过程,通过对输入数据从低层到高层渐进地进行特征提取,最终形成适合模式分类的较理想特征,从而提升分类的准确性。2006年, Hinton等^[7]采用深度学习模型实现数据的降维和分类,并指出深层的神经网络结构可以学习到对象更深刻更本质的特征,分类性能强。目前,深度学习已经在语音识别、图像识别、信息检索等领域取得了成功^[8-10],这些应用也证明了深度学习是一种行之有效的分类识别工具。然而,将深度学习应用于高分辨率遥感影像解译的研究工作相对较少,对此,本文提出了一种基于深度学习的高分辨率遥感影像分类方法,将深度学习的常用模型—深度信念网络(DBN)应用于高分辨率影像的分类,在分类时综合利用了影像的光谱和纹理特征。实验采用 Resurs DK1 影像进行了验证,实验结果表明,与浅层学习算法相比基于DBN的方法可以获得更好的分类效果。

2 深度信念网络

深度信念网络由多层受限波尔兹曼机(RBM)和一层反向传播(BP)网络组成,其结构如图1所示。DBN的训练过程包括预训练和微调两个过程。其中预训练采用非监督方式通过无标签样本来进行,训练时采用贪婪算法逐层进行优化,每层RBM的参数单独进行调整,训练完一层后,将该层的输出看作下一层的输入,继续训练下一RBM。预训练完毕后,采用有监督学习方式对最后一层的BP网络进行训练,并将误差逐层向后传播,对整个DBN网络的权值进行微调,微调时采用反向传播方法,这时可以将DBN看作是一个普通的神经网络,但与传统神经网络不同的是各层的权重都是事先训练好而非随机初始化,因此克服了传统神经网络训练时间长和容易陷入局部最优解的缺点。

RBM是一个无向图模型,结构如图2所示,它由一个可见层和一个隐含层组成,每层有若干个节点,层内节点无连接,不同层之间的节点全连接。

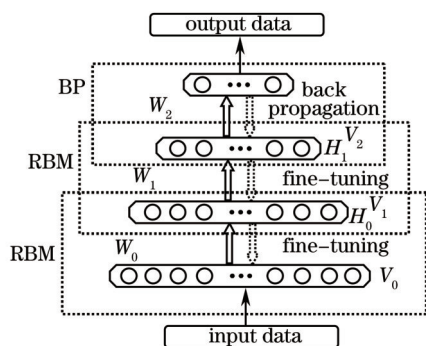


图1 DBN示意图

Fig.1 Schematic diagram of DBN

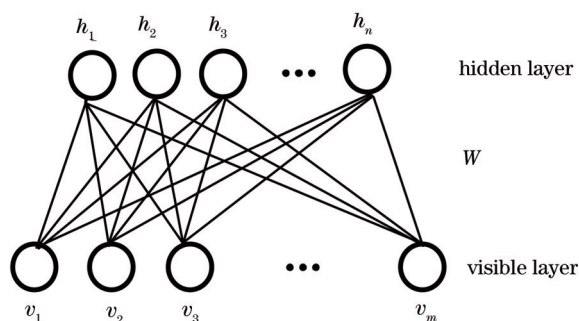


图2 RBM示意图

Fig.2 Schematic diagram of RBM

图2中, m 跟 n 分别表示可见层和隐含层中的神经元数目, $\mathbf{v} = (v_1, v_2, \dots, v_m)^T$ 表示可见层的状态向量, v_i 表示可见层中第 i 个神经元的状态, $\mathbf{h} = (h_1, h_2, \dots, h_n)^T$ 表示隐含层的状态向量, h_j 表示隐含层第 j 个神经元的状态, $\mathbf{a} = (a_1, a_2, \dots, a_m)^T$ 表示可见层的偏置向量, a_i 表示可见层第 i 个神经元的偏置, $\mathbf{b} = (b_1, b_2, \dots, b_n)^T$ 表示隐含层的偏置向量, b_j 表示隐含层第 j 个神经元的偏置, $\mathbf{W} = (w_{ij}) \in R^{m \times n}$ 表示隐含层和可见层之间的权值矩阵, w_{ij} 表示可见层第 i 个神经元与隐含层第 j 个神经元的连接权重, 记 $\boldsymbol{\theta} = (\mathbf{W}, \mathbf{a}, \mathbf{b})$ 表示RBM的参数。RBM是一个基于能量的模型,对于一组给定状态 (\mathbf{v}, \mathbf{h}) , RBM所具备的能量定义为

$$E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta}) = -\sum_{i=1}^m a_i v_i - \sum_{j=1}^n b_j h_j - \sum_{i=1}^m \sum_{j=1}^n w_{ij} v_i h_j . \quad (1)$$

利用(1)式,可以得到状态 (\mathbf{v}, \mathbf{h}) 的联合概率分布

$$P(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta}) = \frac{1}{Z_{\boldsymbol{\theta}}} \exp[-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})] , \quad (2)$$

式中 $Z_{\boldsymbol{\theta}} = \sum_{\mathbf{v}, \mathbf{h}} \exp[-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})]$, 对于观测数据 \mathbf{v} , 其概率分布 $P(\mathbf{v}|\boldsymbol{\theta})$, 它对应 $P(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})$ 的边缘分布, 表示为

$$P(\mathbf{v}|\boldsymbol{\theta}) = \sum_{\mathbf{h}} P(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta}) = \frac{1}{Z_{\boldsymbol{\theta}}} \sum_{\mathbf{h}} \exp[-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})] , \quad (3)$$

对于RBM, 当给定可见层上所有神经元的状态时, 隐含层各神经元激活状态之间是条件独立的, 隐含层第 k 个神经元被激活的概率为

$$P(h_k = 1|\mathbf{v}, \boldsymbol{\theta}) = f_{\text{sigmoid}}\left(b_k + \sum_{i=1}^m w_{ik} v_i\right) , \quad (4)$$

当给定隐含层上所有神经元的状态时, 可见层各神经元的激活状态也是条件独立的, 可见层第 k 个神经元被激活的概率为

$$P(v_k = 1|\mathbf{h}, \boldsymbol{\theta}) = f_{\text{sigmoid}}\left(a_k + \sum_{j=1}^n w_{kj} h_j\right) , \quad (5)$$

训练RBM就是通过调整参数 $\boldsymbol{\theta}$, 以拟合给定的训练数据。参数 $\boldsymbol{\theta}$ 可以通过最大化RBM在训练集上(假设包含 T 个样本)的对数似然函数得到, 即

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\text{arg max}} L(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}}{\text{arg max}} \sum_{t=1}^T \ln P(\mathbf{v}^{(t)}|\boldsymbol{\theta}) , \quad (6)$$

式中 $\mathbf{v}^{(t)}$ 表示训练集第 t 个样本。通常通过梯度上升法来求 $L(\boldsymbol{\theta})$ 的最大值, 其关键是计算梯度 $\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, 通过推导可得

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^T \left[-\sum_{\mathbf{h}} P(\mathbf{h}|\mathbf{v}^{(t)}, \boldsymbol{\theta}) \frac{\partial E(\mathbf{v}^{(t)}, \mathbf{h}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \sum_{\mathbf{v}, \mathbf{h}} P(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta}) \frac{\partial E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] , \quad (7)$$

式中 $P(\mathbf{h}|\mathbf{v}^{(t)}, \boldsymbol{\theta})$ 表示可见层单元限定为已知的训练样本 $\mathbf{v}^{(t)}$ 时隐含层的概率, 该值比较容易计算。由于归一化因子 $Z_{\boldsymbol{\theta}}$ 的存在, $P(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})$ 很难直接计算, 将式中 $\boldsymbol{\theta}$ 分别取 w_{ij}, a_i, b_j 进行计算时, 对每个样本需要求目标分布 \mathbf{v} , 而对于 \mathbf{v} 通常可以通过Gibbs采样来进行近似计算, 然而Gibbs采样需要足够次数的转移才能保证采集到的样本复合目标分布, 这种方法从效率上并不可取。采用Hinton^[11]提出的对比散度(CD)算法可以通过很少次数的采样就可以得到目标分布 \mathbf{v} , 并对参数进行更新, 其更新的规则为

$$\Delta w_{ij} = \varepsilon \left(\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}} \right) , \quad (8)$$

$$\Delta a_i = \varepsilon \left(\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}} \right) , \quad (9)$$

$$\Delta b_j = \varepsilon \left(\langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{recon}} \right) , \quad (10)$$

式中 ε 为学习率, $\langle \cdot \rangle_{\text{data}}$ 表示训练数据集定义分布上的数学期望, $\langle \cdot \rangle_{\text{recon}}$ 表示通过CD算法得到的目标分布的期望。

3 光谱及纹理特征提取

由于高分辨率影像地物特征的复杂化, 仅利用影像的光谱特征进行分类时结果往往并不理想, 因此需要利用其多元化特征, 通过影像变换获取潜在信息来补充光谱特征空间, 以提高影像的模式可分性, 在影像分类时结合光谱和纹理信息能够得到更有效的结果^[12]。对此, 综合考虑了地物的光谱和纹理特征, 构建了能够表达地物多元信息的光谱—纹理特征向量, 将其作为DBN分类器的输入进行分类。

对于每一个待分类像素点, 其光谱特征通过像素点本身及其周围邻域像素所组成的矩形区域来描述。

假设邻域窗口大小为 L , 则对每个像素点, 可以构建维数为 $L*L$ 的光谱特征向量。采用区域构成的向量来表示单个像素光谱特征可以一定程度上弱化同类地物的光谱差异, 同时消除图像斑点噪声的影响。

通过非下采样轮廓波变换(NSCT)来进行影像纹理特征的提取。轮廓波变换(CT)是由 Do 和 Vetterli 提出的一种图像稀疏表示方法^[13], 它可以对图像的高频信息进行多个方向的分解, CT 不仅具有小波变换的多尺度、时频局部特性, 而且具有很好的方向性和各向异性, 克服了小波变换方向选择性较差的不足。然而, CT 在对图像进行分解和重构时需要进行上下采样, 这使得 CT 不具备平移不变性。针对这一不足, Cunha 等^[14-15]又提出了 CT 的改进算法—NSCT, NSCT 不仅具有 CT 的多尺度、多方向等优良特性, 而且还具备完全平移不变性, 具有优良细节捕捉能力。

NSCT 变换包括非下采样金字塔(NSP)分解和非下采样方向滤波器组分解(NSDFB)两部分。其变换流程为: 通过 NSP 分解把源图像分解为一个低通子图像和带通子图像, 对带通子图像采用扇形方向滤波器组进行分解, 对低通子图像重复上述分解过程就可以得到图像的多尺度 NSCT 特征。图像经 k 级 NSP 分解后可以得到 1 个低通图像和 k 个尺度不同的带通图像, 而某一尺度上的带通图像再进行 l 级方向分解, 可以得到 2^l 个与源图像具有相同尺寸的方向子带图像, 从而实现频域中精确的方向分解。目前, NSCT 已经在图像去噪、图像融合、纹理图像的检索与分类等领域取得较好的效果^[16-21], 黄昕^[22]将 NSCT 用于高分辨率遥感图像纹理特征的提取, 得到了比灰度共生矩阵、小波变换等更好的效果, 充分说明了 NSCT 能够有效地反映图像的纹理和细节变化。

采用由 NSCT 变换后的各带通子图像系数构成的特征向量来表示源图像的纹理特征, 其计算方法为:

假设对源图像 I 进行了 k 级 NSCT 变换, 产生的带通子图像分别为 I_1, I_2, \dots, I_k , 某一子图像 $I_m (1 \leq m \leq k)$ 的方向分解数为 d_m , 则 $I_m (1 \leq m \leq k)$ 产生的方向子图像分别表示为 $I_m^1, I_m^2, \dots, I_m^{d_m} (1 \leq m \leq k)$, 对于源图像 I , 其 (i, j) 处的纹理特征向量定义为

$$\mathbf{F}_{ij} = \left[I_1^1(i, j), I_1^2(i, j), \dots, I_1^{d_1}(i, j), I_2^1(i, j), I_2^2(i, j), \dots, I_2^{d_2}(i, j), \dots, I_k^1(i, j), I_k^2(i, j), \dots, I_k^{d_k}(i, j) \right], \quad (11)$$

式中 $I_m^n(i, j)$ 表示方向子图像 I_m^n 在 (i, j) 处的系数值, 向量 \mathbf{F}_{ij} 的维数为 $\sum_{m=1}^k 2^{d_m}$ 。将每个待分像素点的光谱特征向量和纹理特征向量首尾相连, 得到的维数为 $L*L + \sum_{m=1}^k 2^{d_m}$ 的特征向量, 就是该像素点的光谱-纹理特征向量。

4 基于 DBN 的高分辨率遥感影像分类方法

采用 DBN 对高分辨率遥感影像进行分类的步骤为: 1) 数据预处理, 包括图像增强及图像配准; 2) 图像纹理特征计算及光谱-纹理特征构建; 3) 训练样本选取及 DBN 网络参数的设置; 4) 基于 DBN 的网络分类及精度评定。其中步骤 3) 中 DBN 网络参数的设置主要包括 RBM 参数设置及 DBN 网络层数的选择。

4.1 RBM 参数设置

4.1.1 RBM 训练参数设置

RBM 的训练参数主要包括学习率 ε , 权重和偏置的初始值以及训练样本的小批量设置。当 ε 设置较大时, 算法收敛速度较快, 但可能引起系统不稳定, ε 较小时, 虽可避免不稳定的情况出现, 但收敛速度较慢, 通常为保持算法的稳定而选择较小 ε 。权重矩阵 \mathbf{W} 各元素一般选择符合正态分布的较小随机数, 隐含层偏置 \mathbf{b} 各元素初始化为 0, 可见层偏置 \mathbf{a} 的元素进行初始化:

$$a_i = \log \frac{p_i}{1-p_i}, \quad (12)$$

式中 p_i 表示训练样本中第 i 个特征处于激活状态的样本所占的比例。此外为保证训练的效率, 充分利用图形处理器中矩阵运算的优势, 通常将整个样本集分为包含小批量数据的多个样本集逐个进行训练。

4.1.2 RBM 隐含层节点数设置

通过计算 RBM 的重构误差来选择隐含层的节点数。重构误差是以训练数据作为初始状态, 根据 RBM 的分布进行一次 Gibbs 采样后获得的样本与原数据的差异(一般用一范数或二范数来评估)。其计算方法为:

对于样本集中的每一训练样本 $v^{(t)}$, $t \in \{1, 2, \dots, T\}$, 通过 $P(\cdot|v^{(t)})$ 对隐含层进行采样, 得到隐含层状态 h , 通过 $P(\cdot|h)$ 对可见层进行采样后得到 v , 计算累计误差 $E_{\pi} = E_{\pi} + \|v - v^{(t)}\|$ 。对所有样本计算完毕后得到的 E_{π} 就是 RBM 的重构误差。重构误差能够在一定程度上反映 RBM 对训练数据的似然度, 虽然并不完全可靠^[23], 但是实现方法简单、高效, 在实践中比较有用。

在设置隐含层节点数时, 需确定节点数的大致范围, 通过计算该范围内不同节点数所对应 RBM 的重构误差, 取最小重构误差所对应的数值作为 RBM 的隐含层节点数。当确定了第 1 层 RBM 隐含层的节点数后, 将其作为下一层 RBM 的输入层, 按同样方法计算该层 RBM 的重构误差, 从而确定该 RBM 隐含层即 DBN 第 2 隐含层的节点数, 重复上述过程可以确定后续隐含层的节点数。

4.2 DBN 网络深度

在实际应用中 DBN 网络深度对分类效果影响明显, 更高的网络深度可以发掘出更为抽象的特征表示, 提高分类性能^[24], 但是过多的层数不仅使训练耗时增大, 训练效率降低, 还可能降低网络泛化性能, 出现过拟合的现象^[25]。合适的网络深度往往与具体的应用和数据集有关, 通过实验的方法, 针对不同的网络深度, 以计算分类后的总体精度(OA)和 Kappa 系数来确定最终的深度。

5 实验与分析

5.1 实验数据及样本选取

实验数据为美国佐治亚州亚特兰大市北部一住宅小区的绿、红、近红外三波段 Resurs DK1 影像, 覆盖范围为: $84^{\circ}32'28'' \sim 84^{\circ}32'43''W$, $34^{\circ}6'39'' \sim 34^{\circ}6'49''N$, 影像的空间分辨率为 1.25 m, 大小为 671 行 499 列, 成像时间为 2009 年 11 月, 图 3 是经过 ERDAS 预处理之后, 三波段彩色合成图。该区域地物类型主要包括房屋、道路、树木、小路、空地、草地。

在进行样本选取时, 考虑的地物的光谱和纹理特征, 对每一类地物选取约 1000 个像素点, 样本总数占实验数据量的 1.82%, 对于每一类样本选取 2/3 作为无标签数据用于无监督训练, 1/3 作为有标签数据进行网络的有监督微调。



图 3 Resurs DK1 高分辨率影像

Fig.3 High resolution image of Resurs DK1

实验中, 进行光谱特征计算的窗口设置为 $L=3$, NSCT 变换的尺度为 3, 每一级方向数分别为 4、4、3, 得到的纹理特征向量维数为 $16+16+8=40$, 实验分别计算了三个波段的光谱-纹理特征向量, 将 3 个特征向量组合成为一个特征向量作为分类的输入, 输入向量维数为 $3 \times (3 \times 3 + 40) = 147$ 。

5.2 实验参数设置

实验中学习率 ε 的初始值设定为 0.01, 训练时, 通过重构误差来动态调整 ε , 若重构误差下降则 ε 不变, 重构误差上升或不变则将 ε 除以 2。 W 各元素来自正态分布 $N(0, 0.01)$ 的随机数。隐含层偏置 b 各元素初始化为 0, 可见层偏置 a 采用(12)式得到。将样本集设置为包含 100 个样本的多个小样本集进行分批训练。对于 RBM 隐含层节点参数, 由于输入、输出层节点数分别为 147 和 6, 因此, 实验在 6~147 范围内以 5 为步长进行取值, 计算不同节点数对应的重构误差, 当节点数为 56 时重构误差最小, 从而确定第 1 隐含层节点数为 56,

后续隐含层节点数采用同样方法来进行确定。

在确定 DBN 网络深度时,实验将隐含层的数目分别设置为 1~4,通过重构误差确定每层的节点数及其他网络参数后,训练网络并对实验影像进行分类,计算分类后的精度系数,如图 4 所示,可知当隐含层数为 3 时 OA 和 Kappa 系数最大。因此实验最终确定的 DBN 隐含层数为 3。

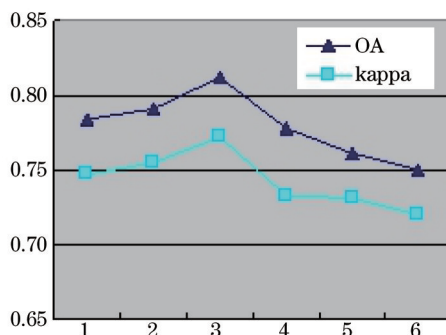


图 4 不同 DBN 层数的分类精度

Fig.4 Classification accuracy of different DBN layers

5.3 实验结果及分析

实验分类结果如图 5(a)所示。进行精度分析时,均匀选取各类地物共 32709 个已知像素点作为测试样本,如图 5(b)所示,根据测试样本计算得到分类混淆矩阵及各种地物的分类精度如表 1 所示。其中,房屋的用户精度为 73.44%,这是由于部分房屋与道路、树木的光谱特性相似,所以被错分为道路和树木,影响其精度。树木的用户精度为 82.13%,虽然树木与道路和部分房屋的光谱特征相似,但是加入 NSCT 多尺度纹理特征后,使其与另外两种地物得到了较好地区分。道路也得到了较好地识别,用户精度为 81.85%。

表 1 基于 DBN 分类方法的混淆矩阵

Table 1 Confusion matrix of classification based on DBN

Types	Roofs	Roads	Trees	Lawns	Trails	Clearings	User's accuracy /%
Roofs	3598	647	542	0	78	34	73.44
Roads	423	4935	174	0	435	62	81.85
Trees	117	83	6020	681	78	351	82.13
Lawns	0	0	567	2637	0	121	79.30
Trails	6	343	0	0	4246	427	84.55
Clearings	0	23	321	413	229	5128	83.87
Procuder's accuracy /%	86.82	81.82	78.96	70.68	83.81	83.75	

OA:81.2% Kappa coefficient:0.772

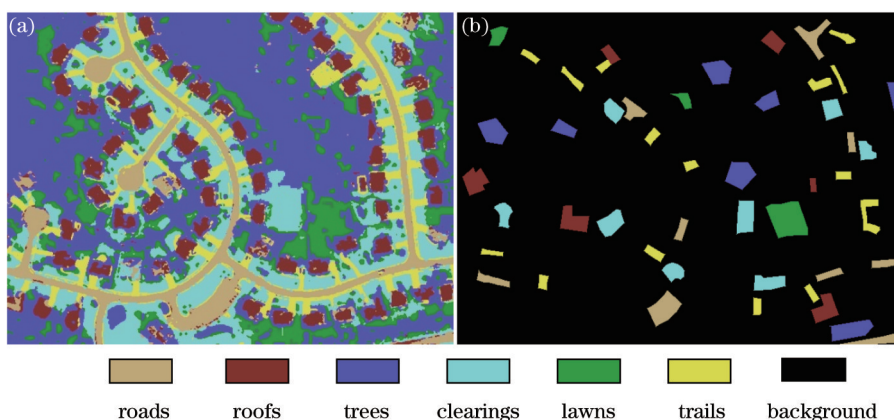


图 5 (a)DBN 方法的分类效果和(b)测试样本

Fig.5 (a) Classification results of DBN and (b) test samples

草地虽然与部分树木光谱特征相似,且存在一定混合分布,但通过纹理特征了得到了较好地区分。小路和空地由于相对独特的光谱、纹理特征也得到了较好地识别。从整体上看,树木成连续面状分布,道路成线状分布,房屋等呈块状分布,整体分类效果较好。

5.4 与其他方法对比

实验将本文方法与基于光谱单源信息进行分类的DBN方法(DBNO)及采用光谱-纹理特征的NN方法和SVM方法进行了对比,分类效果如图6所示。其中,DBNO方法的光谱特征设置与本文实验相同,隐含层数为1层,隐含层节点数为14,其他训练参数设置与本文方法一致。SVM方法的多类问题采用一对一方法,核函数采用径向基函数,采用基于交叉验证的网格搜寻法来确定惩罚因子 C 和核函数参数 γ 的取值。NN采用常见的BP神经网络,设置1个隐含层。4种方法的分类精度和Kappa系数对比如表2所示。

表2 四种方法的分类精度和用时对比

Table 2 Comparison of classification accuracy and execution time for the four methods

Method	DBN	DBNO	NN	SVM
OA / %	81.2	72.1	79.4	80.1
Kappa coefficient	0.772	0.675	0.748	0.756
Execution time/s	47	43	131	38

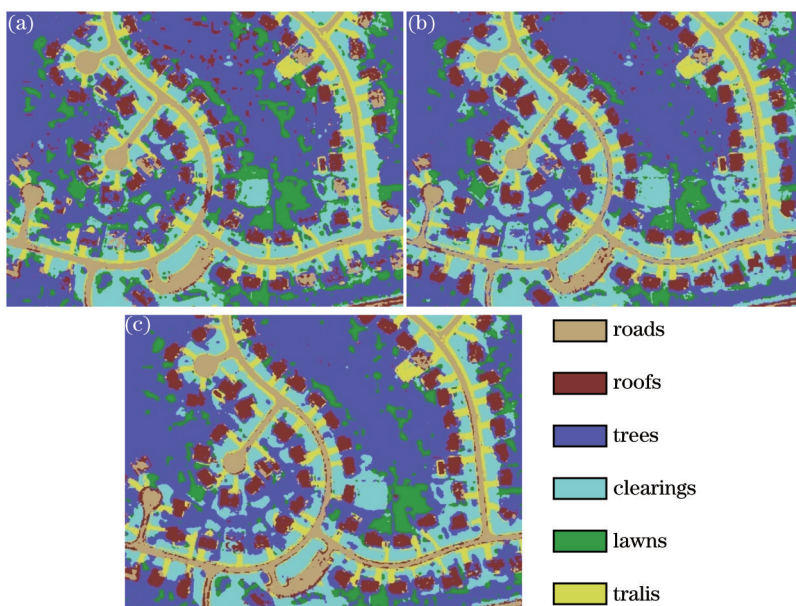


图6 其他方法的分类效果。(a) DBNO; (b) NN; (c) SVM

Fig.6 Classification results of other methods. (a) DBNO; (b) NN; (c) SVM

可以看出,在4种方法中,DBN方法的分类效果最优,DBNO分类效果最差,这说明纹理特征在高分辨率遥感影像分类中的重要性。相对于NN和SVM两种分类器,DBN取得了更好的分类效果,其原因在于,作为一种深层的网络结构,DBN可以获得对象更本质的特征,适应更复杂的样本。此外,DBN通过无监督学习得到的网络权值接近全局最优,而不像普通神经网络那样给权值赋以随机数,DBN结合了无监督学习和有监督学习各自的优点,能更好地发掘遥感数据的分布规律,提高分类精度。除进行精度对比外,实验还对4种方法的执行时间进行了比较,软件实现平台为Matlab,硬件配置为主频2.0 GHz,内存4 GB。可以发现,DBN法虽然用时要略高于SVM,但是大大低于NN方法,这是因为DBN在训练时逐层进行优化,层数跟训练用时的增加是线性关系,而每层在训练时通过CD算法可以极大降低时间复杂度,因此用时要低于传统的NN方法。

6 结 论

利用深度学习的代表性模型—DBN对高分辨率遥感影像进行了基于光谱-纹理特征的分类研究。通过与SVM和NN等传统机器分类方法进行比较,说明深度学习网络能够更好提取地物的本质特征,更准确地挖

掘高分辨率遥感影像的空间分布规律,从而提升分类的准确性。此外,通过与单源信息的DBN分类方法进行比较还表明,综合利用影像的光谱纹理特征能够有效提高高分辨率遥感影像的分类精度。

通过研究可以发现深度学习在遥感影像分类领域的潜力,但是作为一种新兴的技术,仍有很多工作需要研究,主要表现在网络结构的选择仍然没有完善的理论依据,很多时候通过经验或实验方法来选取最优参数,因此影响了分类的效率。下一步,将以如何更加快速准确确定深度学习的网络参数作为研究的方向。

参 考 文 献

- 1 Luo B, Zhang L. Robust autodual morphological profiles for the classification of high-resolution satellite images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(2): 1451-1462.
- 2 Pacifici F, Chini M, Emery W J. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification[J]. Remote Sensing of Environment, 2009, 113(6): 1276-1292.
- 3 Huang X, Zhang L. An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(1): 257-272.
- 4 Liu C, Hong L, Chen J, *et al.* Fusion of pixel-based and multi-scale region-based features for the classification of high-resolution remote sensing image[J]. Journal of Remote Sensing, 2015, 19(2): 228-239.
- 5 Bengio Y. Learning deep architectures for AI[J]. Foundations and Trends in Machine Learning, 2009, 2(1): 1-127.
- 6 Bottou L. Large-scale kernel machines[M]. Cambridge: MIT Press, 2007: 321-359.
- 7 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- 8 Yu D, Deng L. Deep learning and its applications to signal and information processing[J]. IEEE Signal Processing Magazine, 2011, 28(1): 145-154.
- 9 Jones N. The learning machines[J]. Nature, 2014, 505(7842): 146-148.
- 10 Yu Kai, Jia Lei, Chen Yuqiang, *et al.* Deep learning: Yesterday, today and tomorrow[J]. Journal of Computer Research and Development, 2013, 50(9): 1799-1804.
余 凯, 贾 磊, 陈雨强, 等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展, 2013, 50(9): 1799-1804.
- 11 Hinton G. Training products of experts by minimizing contrastive divergence[J]. Neural Computation, 2002, 14(8): 1771-1800.
- 12 Jin Jing, Zou Zhengrong, Tao Chao. Compressed texture based high resolution remote sensing image classification[J]. Acta Geodaetica et Cartographica Sinica, 2014, 43(5): 493-499.
金 晶, 邹峥嵘, 陶 超. 高分辨率遥感影像的压缩纹理元分类[J]. 测绘学报, 2014, 43(5): 493-499.
- 13 Do M N, Vetterli M. The Contourlet transform: An efficient directional multiresolution image representation[J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091-2106.
- 14 Cunha A L, Zhou J, Do M N. The nonsubsampling contourlet transform: Theory, design, and applications[J]. Image Processing, IEEE Transactions on, 2006, 15(10): 3089-3101.
- 15 Zhou J, Cunha A L, Do M N. Nonsubsampling contourlet transform: Construction and application in enhancement[C]. Image Processing, 2005. ICIP 2005. IEEE International Conference on, IEEE, 2005, 1: 469-472.
- 16 Wu Xiaoyue, Guo Baolong, Tang Lu, *et al.* A new adaptive image denoising method based on the nonsubsampling Contourlet transform algorithm[J]. Acta Optica Sinica, 2009, 29(8): 2147-2152.
武晓玥, 郭宝龙, 唐 璐, 等. 一种新的基于非下采样Contourlet变换的自适应图像去噪算法[J]. 光学学报, 2009, 29(8): 2147-2152.
- 17 Zhou Jinghong, Zhou Cui, Zhu Jianjun, *et al.* A method of super-resolution reconstruction for remote sensing image based on non-subsampling Contourlet transform[J]. Acta Optica Sinica, 2015, 35(1): 0110001.
周靖鸿, 周 璇, 朱建军, 等. 基于非下采样轮廓波变换遥感影像超分辨率重建方法[J]. 光学学报, 2015, 35(1): 0110001.
- 18 He Yongcong, Liu Wenbo, Zhang Gong, *et al.* Rotation-invariant texture image retrieval algorithm based on non-subsampling Contourlet transform[J]. Journal of Image and Graphics, 2011, 16(1): 79-83.
何永从, 刘文波, 张 弓, 等. 基于NSCT的旋转不变纹理图像检索算法[J]. 中国图象图形学报, 2011, 16(1): 79-83.
- 19 Wu Yiquan, Yin Jun, Zhu Li, *et al.* Method for improving reconstructed image quality of digital hologram based on SRAD and NSCT[J]. Chinese J Lasers, 2014, 41(2): 0209024.
吴一全, 殷 骏, 朱 丽, 等. 基于SRAD和NSCT的数字全息再现像像质改善方法[J]. 中国激光, 2014, 41(2): 0209024.
- 20 Zhou Liangji, Li Qingwu, Huo Guanying, *et al.* Fase image recognition method based on the NSCT and bionic pattern[J]. Laser & Optoelectronics Progress, 2015, 52(3): 031001.

- 周亮基, 李庆武, 霍冠英, 等. 基于 NSCT 和仿生模式的人脸图像识别方法[J]. 激光与光电子学进展, 2015, 52(3): 031001.
- 21 Chen Musheng, Cai Zhishan. Study on fusion of visual and infrared images based on NSCT[J]. Laser & Optoelectronics Progress, 2015, 52(6): 061002.
- 陈木生, 蔡植善. 基于 NSCT 的红外与可见光图像融合方法研究[J]. 激光与光电子学进展, 2015, 52(6): 061002.
- 22 Huang Xin. Multiscale texture and shape feature extraction and object-oriented classification for very high resolution remotely sensed imagery[D]. Wuhan: Wuhan University, 2009: 71-76.
- 黄 昕. 高分辨率遥感影像多尺度纹理、形状特征提取与面向对象分类研究[D]. 武汉: 武汉大学, 2009: 71-76.
- 23 Hinton G. A practical guide to training restricted Boltzmann machines[J]. Momentum, 2010, 9(1): 926.
- 24 Roux N L, Bengio Y. Representational power of restricted Boltzmann machines and deep belief networks[J]. Neural Computation, 2008, 20(6): 1631-1649.
- 25 Larochelle H, Bengio Y, Louradour J, *et al.*. Exploring strategies for training deep neural networks[J]. Journal of Machine Learning Research, 2009, 10(1): 1-40.

栏目编辑: 苏 岑