基于深度协同稀疏编码网络的海洋浮筏 SAR 图像目标识别

耿 杰 [1] 范剑超 [1,2] 初佳兰 [2] 王洪玉 [1]

摘 要 浮筏养殖广泛存在于我国近海海域, 可见光遥感图像无法完全准确地获取养殖目标, 而基于主动成像的合成孔径雷达 (Synthetic aperture radar, SAR) 遥感图像能够得到养殖目标, 因此采用 SAR 图像进行海洋浮筏养殖目标识别. 然而, 海洋遥感 SAR 图像包含大量相干斑噪声, 并且 SAR 图像特征单一, 使得目标识别难度较大. 为解决这些问题, 提出一种深度协同稀疏编码网络 (Deep collaborative sparse coding network, DCSCN) 进行海洋浮筏识别. 本文方法对预处理后的图像先提取纹理特征和轮廓特征, 再进行超像素分割并将同一个超像素块特征组输入该网络进行协同表示, 最后得到有效特征并分类识别. 通过人工 SAR 图像和北戴河海域浮筏养殖 SAR 图像的实验验证所提模型的有效性. 该网络不仅具有优异的特征表示能力, 能够获得更适合分类器的特征, 而且通过近邻协同约束, 有效抑制相干斑噪声影响, 所以提高了 SAR 图像目标识别精度.

关键词 合成孔径雷达, 深度学习, 稀疏自动编码器, 浮筏养殖, 目标识别

引用格式 耿杰, 范剑超, 初佳兰, 王洪玉. 基于深度协同稀疏编码网络的海洋浮筏 SAR 图像目标识别. 自动化学报, 2016, **42**(4): 593−604

DOI 10.16383/j.aas.2016.c150425

Research on Marine Floating Raft Aquaculture SAR Image Target Recognition Based on Deep Collaborative Sparse Coding Network

GENG Jie[1] FAN Jian-Chao[1,2] CHU Jia-Lan[2] WANG Hong-Yu[1]

Abstract Floating raft aquaculture is widely distributed in the offshore ocean of China. Since raft information cannot be obtained accurately in the visible remote sensing image, active imaging images acquired from synthetic aperture radar (SAR) are applied. However, oceanic SAR images are seriously contaminated by speckle noise, and effective features of SAR images are deficient, which make recognition difficult. In order to overcome these problems, a deep collaborative sparse coding network (DCSCN) is proposed to extract features and conduct recognition automatically. The proposed method extracts texture features and contour features from the pre-processed image firstly. Then, it segments the image into patches and learns features of each patch collaboratively through the DCSCN network. The optimized features are used for recognition finally. Experiments on the artificial SAR image and the images of Beidaihe demonstrate that the proposed DCSCN network can accurately obtain the area of floating raft aquaculture. Since the network can learn discriminative features and integrate the correlated neighbor pixels, the DCSCN network improves the recognition accuracy and has better performance in overcoming the contamination of speckle noise.

Key words Synthetic aperture radar (SAR), deep learning, sparse auto-encoders, floating raft aquaculture, target recognition

Citation Geng Jie, Fan Jian-Chao, Chu Jia-Lan, Wang Hong-Yu. Research on marine floating raft aquaculture SAR image target recognition based on deep collaborative sparse coding network. *Acta Automatica Sinica*, 2016, **42**(4): 593−604

收稿日期 2015-07-06 录用日期 2016-01-15
Manuscript received July 6, 2015; accepted January 15, 2016
国家自然科学基金 (61273307, 61301130), 中国博士后面上基金 (2014M551082), 北戴河邻近海域典型生态灾害与污染监控海洋公益专项 (201305003), 海域使用动态监测和污染监测研究专项资助
Supported by National Natural Science Foundation of China (61273307, 61301130), the China Postdoctoral Science Foundation (2014M551082), the Public Welfare Project of Beidaihe Marine Ecological Disasters and Pollution Monitoring (201305003) and the Research of Dynamic Monitoring and Pollution Monitoring of Sea Area
本文责任编委 刘跃虎
Recommended by Associate Editor LIU Yue-Hu
1. 大连理工大学电子信息与电气工程学部 大连 116024 2. 国家海洋环境监测中心 大连 116023
1. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024 2. National Marine Environment Monitoring Center, Dalian 116023

合成孔径雷达 (Synthetic aperture radar, SAR) 是一种主动成像传感器, 具有全天候、全天时、高分辨率的数据获取能力[1−2]. SAR 图像目标识别是 SAR 图像解译的重要内容, 在军事侦察、资源探测、地理测绘、海域监测等领域有着广泛的应用. 随着 SAR 成像技术的发展, 分辨率的不断提高, SAR 图像目标识别成为海量 SAR 数据处理的主要手段. 典型的 SAR 图像目标识别模型分为检测 (Detection)、鉴别 (Discrimination) 和分类 (Classification) 三个阶段[2]. 海洋浮筏养殖是海域监测的重要组成部分, 自动识别浮筏信息对实现大面积海

域动态使用监测有着重要意义[3−4]. 由于海洋散射信号较弱, 光谱成像卫星数据难以获取浮筏信息, 而 SAR 是主动式微波成像, 利于获取浮筏信息, 在海域监测中具有明显优势[5]. 因此, 利用 SAR 图像进行海洋浮筏养殖自动识别, 不仅可以大范围、宏观调查海域使用情况, 而且为海洋遥感业务化动态监测提供了有力手段. 但是, 由于海洋周期性波浪的影响, 后向散射系数相干叠加使得成像存在严重的相干斑噪声; 同时由于海洋每个区域的流、浪、涌、内波都不一样, 导致后向散射特性非常不规律, 同一目标在不同海域特征存在明显差异[6]. 这些问题增加了海洋 SAR 图像目标识别难度, 有效的特征表示是解决问题的关键.

在 SAR 图像目标识别问题中, 纹理特征被广泛用于 SAR 图像的特征提取. 纹理特征是图像灰度或色彩在空间上的重复或者变化的特征, 提取纹理特征的方法可以分为三类: 基于统计的方法、基于模型的方法和基于变换域的方法[7]. 在基于统计的方法中, 灰度共生矩阵 (Gray-level co-occurrence matrix, GLCM) 统计图像上保持某距离的两个像素的灰度状况, 主要包括均值、方差、对比度、熵、自相关等元素. 在基于模型的方法中, 马尔科夫随机场模型是最常用的分析模型. 在基于变换域的方法中, 小波变换得到变换域各个方向的系数能反映目标的后向散射特性.

为了更加有效地分析理解图像, 多尺度几何分析 (Multi-scale geometric analysis, MGA) 方法陆续被提出, 它是一种多方向、多尺度和局部性的图像稀疏表示方法, 被称为后小波分析. 由于小波分析的方向受限、频率分辨率低, 而 MGA 可以利用图像的几何正则性, 能够保持丰富的细节信息. 目前常用的 MGA 方法有 Gabor 变换、曲波变换、轮廓波变换、剪切波变换等. Gabor 变换是具有高斯函数的傅里叶变换, 二维 Gabor 函数相当于一个高斯函数调制的复正弦函数, 可以提取图像不同空间位置、空间频率的局部结构信息[8]. 由于二维 Gabor 变换在表征图像纹理变化方面的优势, 已经被普遍用于人脸识别等模式识别问题中. 为了更好地逼近图像的轮廓信息, Do 等提出了轮廓波变换[9], 通过多尺度分解和方向分解, 可以有效地表示图像轮廓. 然而, 对图像轮廓波变换时, 要进行降采样和上采样, 这样就不具备平移不变性, 而平移不变性在边缘保持、目标检测等方面起着重要作用. 为了解决这一问题, Da Cunha 等提出一种非下采样轮廓波变换 (Nonsubsampled contourlet transform, NSCT)[10], 具有平移不变性, 可以很好地保持图像的细节和边缘信息. 当前, NSCT 已经被广泛用于图像融合问题中[11]. 考虑到二维 Gabor 变换和 NSCT 在图像特征表示方面的优势, 本研究利用这两种方法分别提取 SAR 图像的纹理特征和轮廓特征.

在目标识别问题中, 有效的特征表示对于识别准确度起着关键作用. 但是在许多研究中, 特征提取和表示都需要大量的经验知识和人为地参数设置, 这个过程比较复杂而且效率较低, 尤其是当图像分辨率较高、尺寸较大时, 更增加了特征表示的难度. 近年来, 深度学习神经网络被用于特征自主学习和表示, 在智能识别任务中表现出极其优异的性能[12−15]. 深度学习网络具有类似于人脑感知视觉的层次化结构, 不过分依赖研究人员的经验知识, 能逐层抽象学习数据的特征, 从而发现数据的内在结构信息并改善分类识别的效果. 研究表明, 深度网络可以很好的挖掘图像的低层特征 (如边界) 和中层特征 (如形状), 在图像去噪、图像分类识别等问题中取得较好的效果. 在分类识别问题中, 利用深度网络得到的特征表示, 不同类别的区分度更大, 更适合于监督分类. 常用的深度网络结构有深度置信网络 (Deep belief networks, DBN)[12]、卷积神经网络 (Convolutional neural networks, CNN)[16] 和堆叠自动编码器 (Stacked auto-encoders, SAE)[17].

本文提出一种深度协同稀疏编码网络 (Deep collaborative sparse coding network, DCSCN) 来进行 SAR 图像目标识别, 并用在海洋浮筏养殖监测任务中. 为了从 SAR 图像获得更优异的特征, 提高目标与背景特征差异性, DCSCN 利用了自动编码器在特征表示中的优势, 获得更适合于分类器的特征. 为了克服海洋 SAR 图像严重的相干斑噪声, 对超像素分割的近邻像素特征利用 DCSCN 进行协同表示, 使得同个超像素块特征更相近并平滑掉相干斑噪声, 减小噪声对目标识别的影响.

1 深度学习原理

传统的神经网络都是随机初始化模型的参数, 然后通过反向传播来优化参数. 然而深度网络如果利用这种方式, 很容易陷入局部极值或者造成梯度弥散, 使得网络参数无法最优. 为了解决这一问题, 提出了逐层学习参数的预训练, 再对预训练后的模型进行反向传播来优化参数的方法[18]. 常用的预训练单元有限制玻尔兹曼机 (Restricted Boltzmann machine, RBM) 和自动编码器 (Auto-encoders), 多层预训练单元堆叠起来构成深度网络.

1.1 自动编码器

一个自动编码器由输入层、隐含层和输出层构成, 训练学习的目标是使得网络的输出尽量逼近输入[19]. 自动编码器训练过程由编码过程和解码过程组成, 编码过程将输入样本进行线性映射和非线性映射变换得到隐含层表示. 设样本集为

$\boldsymbol{X}=\{\boldsymbol{x}_i\}_{i=1}^{N}$, 输入样本 $\boldsymbol{x}_i$ 对应的隐含层表示为

$$\boldsymbol{h}_i=f\left(\boldsymbol{x}_i\right)=\text{sigmoid}\left(\boldsymbol{W}_1\boldsymbol{x}_i+\boldsymbol{b}_1\right) \tag{1}$$

其中, $\boldsymbol{W}_1$ 和 $\boldsymbol{b}_1$ 分别代表输入层与隐含层之间的权重和偏置, sigmoid$(\cdot)$ 表示隐含层的激励函数. 解码过程就是把编码后的表示重新投影到原信号空间, 得到解码信号 $\hat{\boldsymbol{x}}_i$ 可表示为

$$\hat{\boldsymbol{x}}_i=g\left(\boldsymbol{x}_i\right)=\text{sigmoid}\left(\boldsymbol{W}_2\boldsymbol{h}_i+\boldsymbol{b}_2\right) \tag{2}$$

其中, $\boldsymbol{W}_2$ 和 $\boldsymbol{b}_2$ 分别表示隐含层与输出层之间的权重和偏置, sigmoid$(\cdot)$ 表示输出层的激励函数. 自动编码器的目标是让解码输出尽量逼近编码之前的输入, 网络的参数通过重构误差最小化来优化, 目标函数如下

$$J\left(\boldsymbol{W}_1,\boldsymbol{W}_2,\boldsymbol{b}_1,\boldsymbol{b}_2\right)=\arg\min_{\boldsymbol{W}_1,\,\boldsymbol{W}_2,\,\boldsymbol{b}_1,\,\boldsymbol{b}_2}\sum_{i=1}^{N}\left\|\boldsymbol{x}_i-\hat{\boldsymbol{x}}_i\right\|_2^2 \tag{3}$$

训练好一层自动编码器后, 将其隐含层的激活值作为下一层自动编码器的输入. 这样将多层自动编码器堆叠起来构成了堆叠自动编码器.

1.2　堆叠自动编码器

堆叠自动编码器是一种典型的深度神经网络, 被广泛用于特征学习与表示. 该网络先逐层贪婪学习来确定参数, 再从最顶层反向传播来微调 (Fine-tune) 整个网络的参数. 如果最顶层不加入标签信息, 则这个学习过程为无监督的特征学习过程; 如果最顶层加入样本的标签信息来反向微调整个网络的参数, 则这个学习过程是监督学习过程. 图 1 给出了堆叠自动编码器结构图.

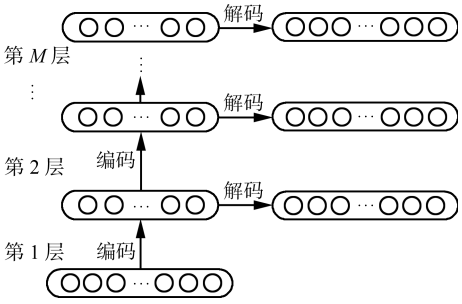


图 1　堆叠自动编码器结构图

Fig. 1　The structure of SAE

由图 1 所示, 该堆叠自动编码器共 M 层, 最高层自动编码器的隐含层就是网络的输出. 通常在堆叠自动编码器后连接 Softmax 分类器进行分类识别. 在有监督学习网络中, 样本的标签也是在 Softmax 分类器加入的, 整个网络的参数通过后向传播来整体调整. 网络整体参数微调的目标函数为

$$J\left(\boldsymbol{W}_k\left|_{k=1}^{M}\right.,\boldsymbol{b}_k\left|_{k=1}^{M}\right.\right)=\arg\min_{\boldsymbol{W}_k,\,\boldsymbol{b}_k}\sum_{i=1}^{N}\left\|\boldsymbol{y}_i-g_M\left(f_M\left(\cdots f_1\left(\boldsymbol{x}_i\right)\right)\right)\right\|_2^2 \tag{4}$$

其中, $\boldsymbol{y}_i$ 表示样本 $\boldsymbol{x}_i$ 对应的标签, $\boldsymbol{W}_k$ 和 $\boldsymbol{b}_k$ 分别表示整个网络各层的权重和偏置. 加入标签信息来优化整个网络, 可以使得网络的特征表示更适合于分类器, 即不同类别的样本特征差异更大.

2　SAR 图像预处理与特征提取

2.1　预处理

卫星拍摄的 SAR 图像需要先进行几何校正, 来解决由于传感器以及地球自转等因素造成的图像相对于地面目标发生的几何畸变. 由于原始的 SAR 图像对比度较差, 需采用灰度拉伸来提高图像灰度级的动态范围. 为了增大目标与背景间的对比度, 利用非线性函数对图像进行灰度拉伸处理. 为了抑制相干斑噪声对目标识别的影响, 采用增强 Lee 滤波算法进行去噪.

2.2　纹理特征提取

二维 Gabor 滤波器可以提取图像多尺度、多方向的纹理特征. 二维 Gabor 变换是具有高斯函数形式的傅氏变换, 由一个高斯函数乘以正弦函数, 正弦函数的频率决定特征的尺度, 相位决定特征的方向, 具体公式如下[20]

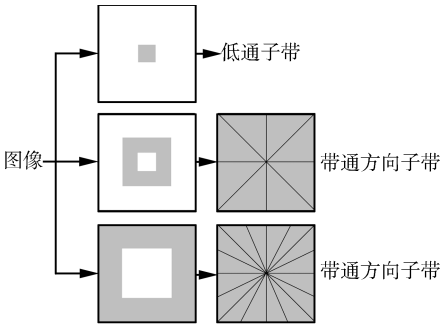
$$G\left(x_0,y_0,\theta,\omega_0\right)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x_0^2+y_0^2}{2\sigma^2}\right)\times\left[\exp\left(j\omega_0x_0\right)-\exp\left(-\frac{\omega_0^2\sigma^2}{2}\right)\right] \tag{5}$$

其中, $x_0=x\cos\theta+y\sin\theta$, $y_0=-x\sin\theta+y\cos\theta$, x 和 y 代表像素点的位置, ω_0 表示滤波器的中心频率, θ 表示滤波器的方向, σ 表示高斯函数的标准差. $\exp\left(j\omega_0x_0\right)$ 代表交流分量, $\exp\left(-\omega_0^2\sigma^2/2\right)$ 代表直流分量. 当直流分量趋于 0 时, Gabor 滤波器对光照强度不敏感. θ 和 σ 不同取值得到一组滤波器, 将滤波器组与图像进行卷积得到与图像相同大小的分解图像, 这些图像代表纹理特征.

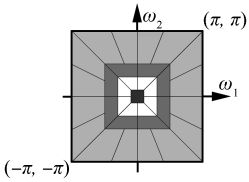
2.3　轮廓特征提取

非下采样轮廓波变换 (NSCT)[10] 可以将图像分解为不同频率和不同方向的子带图像, 各个子带图像表示了图像在各个频率和各个方向上的系数, 可以看作原图的一张特征图, 即轮廓特征图. NSCT 交替利用非下采样的塔型滤波器组和非下采样的方

向滤波器组, 来得到多方向、多尺度的二维分解图像. 图 2 给出了具体分解的过程.



(a) NSCT结构图
(a) Constructure of NSCT



(b) 经过滤波器组成的频域分解
(b) Frequency domain decomposition with filter bank

图 2 非下采样轮廓波变换
Fig. 2 Nonsubsampled contourlet transform

如图 2 所示, NSCT 先利用非下采样的塔型滤波器组将图像分解为低通、带通和高通子带图像, 然后利用非下采样方向滤波器组对子带图像进行方向滤波, 得到各个频域方向的子带图像. 具体过程如下: NSCT 的非下采样塔型滤波器由多个双通道非下采样滤波器级联构成. 设定分解的尺度为 j, 则在尺度 j 上的理想低通滤波器 $H_0(z)$ 包含的区域为 $\left[-\frac{\pi}{2^j}, \frac{\pi}{2^j}\right]^2$, 理想高通滤波器 $H_1(z)$ 包含的区域为 $\left[-\frac{\pi}{2^{j-1}}, \frac{\pi}{2^{j-1}}\right]^2 \backslash \left[-\frac{\pi}{2^j}, \frac{\pi}{2^j}\right]^2$. 由于每次分解中都只对上次分解结果进行上采样, 而不使用下采样, 所以保证算法的平移不变性. 经过 k 级非下采样塔型滤波, 可以得到 $k+1$ 个与原图像相同尺寸大小的子带图像. NSCT 所采用的方向滤波器组也是一组二通道非下采样滤波器, 其理想频域区域是扇形形状, 对扇形滤波器采用不同的采样矩阵进行上采样, 再对上一级方向分解的子带进行滤波, 得到频域的方向分解. 当方向滤波器组数目较多时, 得到的子带图像代表各个方向的分解图像, 具有优异的轮廓和边界保持效果. 如果对某尺度下子带图像做 t 级方向分解, 可以获得 2^t 个与原图像相同尺寸大小的方向子带图像. 因此, 图像经过 K 级 NSCT 后可得到 $\sum_{k=1}^{K} 2^{t_k} + 1$ 个与原图像尺寸大小相同的子带图像, 其中 t_k 是尺度 k 下的方向分解级数[21−22]. 综上描述, NSCT 分解得到的子带图像具有平移不变性、良好的轮廓和边界保持效果, 这些子带图像可以用来作为目标识别的特征图.

3 基于深度协同稀疏编码网络的 SAR 图像目标识别

海洋 SAR 图像存在严重的相干斑噪声, 使得图像像素的灰度值不完全反映地物的后向散射特性, 基于单个像素特征进行分类识别往往达不到很好的效果. 考虑到图像的统计特性, 近邻像素点的特性比较接近, 可以联合起来进行协同表示. 为了增加不同类别特征的差异性, 并且抑制相干斑噪声的影响, 本研究提出一个深度协同稀疏编码网络 (DCSCN), 采用超像素分割的方法将图像进行分割, 然后对每个超像素块内像素点对应的特征联合优化, 使得同个超像素块的像素特征更趋于相似并平滑掉相干斑噪声. 为了获得更丰富的特征, 在 DCSCN 网络前先进行预处理, 再通过二维 Gabor 变换和 NSCT 来提取纹理特征和轮廓特征. 图 3 给出了整个算法的流程图.

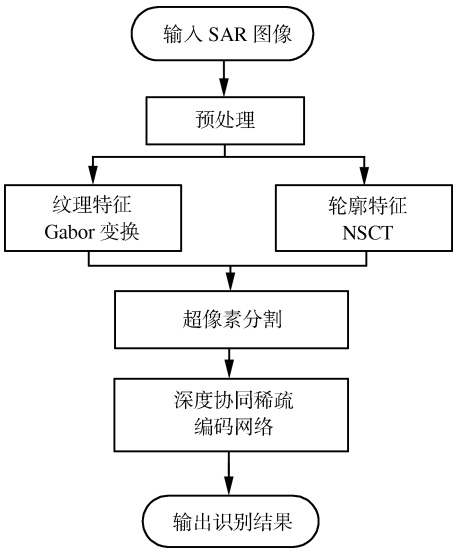


图 3 基于深层协同稀疏编码网络算法流程图
Fig. 3 The flow chart of the proposed algorithm

如图 3 所示, 本文算法先对输入 SAR 图像进行几何校正、灰度拉伸和增强 Lee 滤波去噪的预处理过程, 再利用二维 Gabor 滤波器和 NSCT 滤波器提取图像的特征, 并进行超像素分割, 然后将各个超像素块特征组输入 DCSCN 网络进行特征优化和分类识别, 最后输出识别结果.

本研究中使用了一种基于几何流的水平集方法 (TurborPixels)[23], 将 SAR 图像分割成 M 个超像素块 $\{S_1, S_2, \cdots, S_M\}$. 把每个超像素块特征组输入到 DCSCN 网络进行协同表示和分类识别. 下面详细描述了 DCSCN 网络结构和网络训练过程.

3.1 DCSCN 网络结构

根据第 1 节描述, 堆叠自动编码器可以学习到图像更高层次的特征, 使得不同类别样本特征区分度增大. 为了让隐含层单元激励的平均输出较小, 对每个自动编码器进行稀疏约束, 可以避免参数取值过大, 防止过学习. 稀疏自动编码器的目标是使得解码输出尽量接近输入, 本文为了有效抑制相干斑噪声, 希望各个像素特征通过编码器的解码输出与其相应的超像素块特征组的平均特征接近, 提出了协同稀疏编码器来对每个超像素块的特征组联合训练学习. 将多个协同稀疏编码器堆叠起来构成了深度协同稀疏编码网络 (DCSCN), 图 4 给出了网络结构图. 图中网络共包含 L 层, 第 1 层到第 $L-1$ 层网络为协同稀疏编码器, 第 L 层网络为 Softmax 分类器. 为了进行协同表示, 各层协同稀疏编码器的输入对输出进行约束, 即每个特征的输出都要与输入的这组特征的平均值尽量接近. 因此, 在前向逐层训练过程中, 每层参数都是将特征组的特征依次输入网络后确定的. 完成逐层贪心训练后, 整个网络的参数通过后向反馈整体微调.

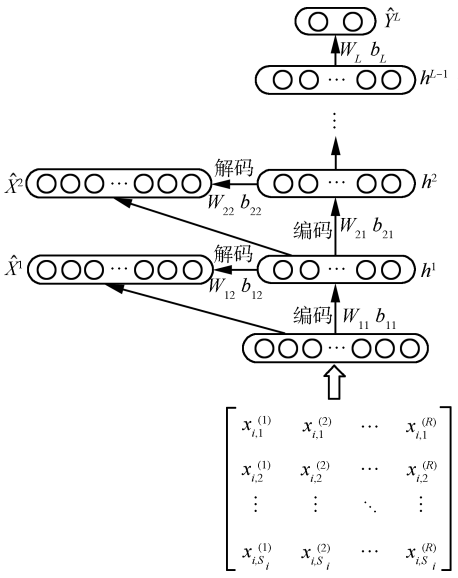


图 4　深度协同稀疏编码网络结构图

Fig. 4　The structure of deep collaborative sparse coding network

如图 4 所示, $\boldsymbol{X}_i = [\boldsymbol{x}_{i,1}\ \boldsymbol{x}_{i,2}\ \cdots\ \boldsymbol{x}_{i,S_i}]$ 表示输入到 DCSCN 网络的第 i 个像素块的 S_i 个像素点的特征组, 其中 $\boldsymbol{x}_{i,1} = [x_{i,1}^{(1)}\ \ x_{i,1}^{(2)}\ \ \cdots\ \ x_{i,1}^{(R)}]^{\mathrm{T}}$ 表示 $\boldsymbol{x}_{i,1}$ 对应的 R 维特征. 各层协同稀疏编码器的编码、解码过程如下: 设第 k 层协同稀疏编码器的输入样本特征为 $\boldsymbol{x}_{i,t}^k$, 进行线性映射得到 $\boldsymbol{z}_{i,t}^{k1} = \boldsymbol{W}_{k1}\boldsymbol{x}_{i,t}^k + \boldsymbol{b}_{k1}$, 通过隐含层激励元得到编码结果 $\boldsymbol{h}_{i,t}^k = \mathrm{sigmoid}\left(\boldsymbol{z}_i^{k1}\right)$. 同样, 解码过程也是先进行线性映射得到 $\boldsymbol{z}_{i,t}^{k2} = \boldsymbol{W}_{k2}\boldsymbol{h}_{i,t}^k + \boldsymbol{b}_{k2}$, 再通过输出层激励元得到解码输出 $\hat{\boldsymbol{x}}_{i,t}^k = \mathrm{sigmoid}\left(\boldsymbol{z}_{i,t}^{k2}\right)$. 为了让输入样本特征 $\boldsymbol{x}_{i,t}^k$ 逼近整个超像素块的平均特征, 网络中解码层的期望输出为 $\bar{\boldsymbol{x}}_i^k = \sum_{t=1}^{S_i}\boldsymbol{x}_{i,t}^k/S_i$, 表示第 i 个超像素块的特征平均值. 网络参数 $\boldsymbol{W}$ 和 $\boldsymbol{b}$ 通过训练学习来确定.

3.2 DCSCN 网络训练

DCSCN 网络需要先确定各层参数, 然后再进行分类识别. 网络各层参数的优化分为预训练 (Pre-training) 和整体微调 (Fine-tune) 两个过程. 预训练是逐层优化各层网络的参数. 整体微调是从顶层引入标签信息来优化整个网络参数, 可以使得网络输出的特征更适合于分类器.

3.2.1 DCSCN 网络预训练

预训练过程是将训练样本的特征输入到网络, 逐层训练得到各层网络的参数. 对于第 k 层网络, 参数优化的目标函数为

$$\begin{aligned} J\left(\boldsymbol{W}_{k1},\boldsymbol{W}_{k2},\boldsymbol{b}_{k1},\boldsymbol{b}_{k2}\right) = \\ \arg_{\boldsymbol{W}_{k1},\,\boldsymbol{W}_{k2},\,\boldsymbol{b}_{k1},\,\boldsymbol{b}_{k2}}\min \frac{1}{2N}\sum_{i=1}^{M}\sum_{t=1}^{S_i}\left\|\bar{\boldsymbol{x}}_i^k - \hat{\boldsymbol{x}}_{i,t}^k\right\|_2^2 + \\ \frac{\lambda}{2}\left(\|\boldsymbol{W}_{k1}\|_2^2 + \|\boldsymbol{W}_{k2}\|_2^2\right) + \\ \beta\sum_{j=1}^{l_k} KL\left(\rho\,\|\hat{\rho}_j\right) \end{aligned} \tag{6}$$

其中, $N = \sum_{i=1}^{M} S_i$ 代表样本总数目, λ 是权重衰减系数. $\|\boldsymbol{W}_{k1}\|_2^2 + \|\boldsymbol{W}_{k2}\|_2^2$ 是权重衰减项, 保证权重值不会过大. β 表示稀疏性惩罚因子权重, l_k 代表第 k 层协同稀疏编码器的隐含层节点个数. $KL(\rho\,\|\hat{\rho}_j)$ 是稀疏性惩罚项, 通过相对熵 (KL 散度, KL divergence) 计算, 公式如下

$$\begin{aligned} KL\left(\rho\parallel\hat{\rho}_j\right) &= \rho\log\frac{\rho}{\hat{\rho}_j} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_j}, \\ \hat{\rho}_j &= \frac{1}{M}\sum_{i=1}^{M} f_{k,j}\left(\bar{\boldsymbol{x}}_i^k\right) \end{aligned} \tag{7}$$

其中, ρ 是稀疏性参数. $\hat{\rho}_j$ 表示第 j 个隐含单元在所有超像素块特征组编码后的平均激励值, 其中 $f_{k,j}\left(\bar{\boldsymbol{x}}_i^k\right) = \mathrm{sigmoid}\left(\boldsymbol{W}_{k1}\bar{\boldsymbol{x}}_i^k + \boldsymbol{b}_{k1}\right)$. 可知, $\hat{\rho}_j$ 的大小由网络权重和偏置决定[24]. ρ 是人为设定的接近 0 的数值, 目的是让 $\hat{\rho}_j$ 的值接近 0, 这样隐含层的输出值接近 0, 使得隐含层具有稀疏性. 目标方程 (6) 利用此约束项求解网络权重和偏置, 目的是使得网络的隐含层输出具有稀疏性.

根据网络的目标函数, 通过梯度下降法来确定网络的参数. 对于第 k 层协同稀疏编码器, 设定

$$\boldsymbol{\delta}_{i,t}^{k2} = -\left(\bar{\boldsymbol{x}}_i^k - \hat{\boldsymbol{x}}_{i,t}^k\right) f'\left(\boldsymbol{z}_{i,t}^{k2}\right) \tag{8}$$

$$\boldsymbol{\delta}_{i,t}^{k1} = \left(\boldsymbol{W}_{k2}\boldsymbol{\delta}^{k2} + \beta\left(-\frac{\rho}{\hat{\rho}_j} + \frac{1-\rho}{1-\hat{\rho}_j}\right)\right) f'\left(\boldsymbol{z}_{i,t}^{k1}\right) \tag{9}$$

其中, $\beta\left(-\frac{\rho}{\hat{\rho}_j} + \frac{1-\rho}{1-\hat{\rho}_j}\right)$ 表示稀疏项偏导, $f'(\cdot)$ 表示 sigmoid 函数的导函数. 这样, 目标函数的偏导数如下

$$\begin{aligned}
&\frac{\partial}{\partial \boldsymbol{W}_{k1}} J\left(\boldsymbol{W}_{k1}, \boldsymbol{W}_{k2}, \boldsymbol{b}_{k1}, \boldsymbol{b}_{k2}\right) = \bar{\boldsymbol{x}}_i^k \boldsymbol{\delta}^{k1} \\
&\frac{\partial}{\partial \boldsymbol{W}_{k2}} J\left(\boldsymbol{W}_{k1}, \boldsymbol{W}_{k2}, \boldsymbol{b}_{k1}, \boldsymbol{b}_{k2}\right) = \boldsymbol{h}_i^k \boldsymbol{\delta}^{k2} \\
&\frac{\partial}{\partial \boldsymbol{b}_{k1}} J\left(\boldsymbol{W}_{k1}, \boldsymbol{W}_{k2}, \boldsymbol{b}_{k1}, \boldsymbol{b}_{k2}\right) = \boldsymbol{\delta}^{k1} \\
&\frac{\partial}{\partial \boldsymbol{b}_{k2}} J\left(\boldsymbol{W}_{k1}, \boldsymbol{W}_{k2}, \boldsymbol{b}_{k1}, \boldsymbol{b}_{k2}\right) = \boldsymbol{\delta}^{k2}
\end{aligned} \tag{10}$$

根据以上分析, 下面给出了利用梯度下降法求第 k 层网络参数的具体过程:

步骤 1. 设定初始值 $\Delta \boldsymbol{W}_{k1} = 0$, $\Delta \boldsymbol{W}_{k2} = 0$, $\Delta \boldsymbol{b}_{k1} = 0$, $\Delta \boldsymbol{b}_{k2} = 0$.

步骤 2. 对于所有超像素块 $i = 1, 2, \cdots, M$, 计算 $\Delta \boldsymbol{W}$ 和 $\Delta \boldsymbol{b}$.

$$\Delta \boldsymbol{W}_{k1} := \Delta \boldsymbol{W}_{k1} + \frac{\partial J\left(\boldsymbol{W}_{k1}, \boldsymbol{W}_{k2}, \boldsymbol{b}_{k1}, \boldsymbol{b}_{k2}\right)}{\partial \boldsymbol{W}_{k1}} = \Delta \boldsymbol{W}_{k1} + \bar{\boldsymbol{x}}_i^k \boldsymbol{\delta}^{k1} \tag{11}$$

$$\Delta \boldsymbol{b}_{k1} := \Delta \boldsymbol{b}_{k1} + \frac{\partial J\left(\boldsymbol{W}_{k1}, \boldsymbol{W}_{k2}, \boldsymbol{b}_{k1}, \boldsymbol{b}_{k2}\right)}{\partial \boldsymbol{b}_{k1}} = \Delta \boldsymbol{b}_{k1} + \boldsymbol{\delta}^{k1} \tag{12}$$

同理,

$$\Delta \boldsymbol{W}_{k2} := \Delta \boldsymbol{W}_{k2} + \boldsymbol{h}_i^k \boldsymbol{\delta}^{k2} \tag{13}$$

$$\Delta \boldsymbol{b}_{k2} := \Delta \boldsymbol{b}_{k2} + \boldsymbol{\delta}^{k2} \tag{14}$$

步骤 3. 更新权重 $\boldsymbol{W}$ 和偏置 $\boldsymbol{b}$.

$$\boldsymbol{W}_{k1} := \boldsymbol{W}_{k1} - \alpha\left(\frac{1}{M}\Delta \boldsymbol{W}_{k1} + \lambda \boldsymbol{W}_{k1}\right) \tag{15}$$

$$\boldsymbol{W}_{k2} := \boldsymbol{W}_{k2} - \alpha\left(\frac{1}{M}\Delta \boldsymbol{W}_{k2} + \lambda \boldsymbol{W}_{k2}\right) \tag{16}$$

$$\boldsymbol{b}_{k1} := \boldsymbol{b}_{k1} - \alpha\left(\frac{1}{M}\Delta \boldsymbol{b}_{k1}\right) \tag{17}$$

$$\boldsymbol{b}_{k2} := \boldsymbol{b}_{k2} - \alpha\left(\frac{1}{M}\Delta \boldsymbol{b}_{k2}\right) \tag{18}$$

其中, α 代表学习率. 权重 $\boldsymbol{W}$ 和偏置 $\boldsymbol{b}$ 迭代更新停止条件是目标函数取值不再变小或者达到最大迭代步数.

3.2.2 DCSCN 网络整体微调

通过预训练逐层学习各层参数, 再从顶层加入训练样本标签信息来优化整个网络参数, 可以使得网络输出的特征更适合于分类器, 即不同类别的样本特征差异更大. 整体微调网络参数采用了误差反向传播梯度下降算法, 设网络层数为 L, 则目标函数表示如下

$$J'\left(\boldsymbol{W}_L, \boldsymbol{W}_{k1}\big|_{k=1}^{L-1}, \boldsymbol{b}_{k1}\big|_{k=1}^{L-1}\right) = \arg_{\boldsymbol{W}_L,\, \boldsymbol{W}_{k1},\, \boldsymbol{b}_{k1}} \min \frac{1}{2N}\sum_{i=1}^{M}\sum_{t=1}^{S_i}\left\|\boldsymbol{y}_{i,t} - g_L\left(f_L\left(\boldsymbol{h}_{i,t}^{L-1}\right)\right)\right\|_2^2 \tag{19}$$

其中, $\boldsymbol{h}_{i,t}^{L-1} = f_{L-1}\left(f_{L-2}\left(\cdots f_1\left(\boldsymbol{x}_{i,t}\right)\right)\right)$ 表示第 $L-1$ 层网络隐含层激励值, $\boldsymbol{y}_{i,t}$ 表示样本 $\boldsymbol{x}_{i,t}$ 对应的标签, $\boldsymbol{W}_L$ 代表网络最后一层的权重, $\boldsymbol{W}_{k1}$ 和 $\boldsymbol{b}_{k1}$ 分别表示网络其他层的权重和偏置. 利用 BP 算法优化过程如下:

步骤 1. 设权重系数的修改量为 $\Delta \boldsymbol{W}$, 考虑到学习速率 η, 有 $\Delta \boldsymbol{W} = -\eta \frac{\partial J'}{\partial \boldsymbol{W}}$.

步骤 2. 求解 $\frac{\partial J'}{\partial \boldsymbol{W}}$.

$$\frac{\partial J'}{\partial \boldsymbol{W}_L} = \frac{\partial J'}{\partial \boldsymbol{z}^L}\frac{\partial \boldsymbol{z}^L}{\partial \boldsymbol{W}_L} = \frac{\partial J'}{\partial \boldsymbol{z}^L}\boldsymbol{h}^{L-1} \tag{20}$$

$$\frac{\partial J'}{\partial \boldsymbol{W}_{k1}} = \frac{\partial J'}{\partial \boldsymbol{z}^{k1}}\frac{\partial \boldsymbol{z}^{k1}}{\partial \boldsymbol{W}_{k1}} = \frac{\partial J'}{\partial \boldsymbol{z}^{k1}}\boldsymbol{h}^{k-1} \tag{21}$$

设 $\boldsymbol{d}^k = \frac{\partial J'}{\partial \boldsymbol{z}^{k1}}$, 则有

$$\boldsymbol{d}^k = \frac{\partial J'}{\partial \boldsymbol{h}^k}\frac{\partial \boldsymbol{h}^k}{\partial \boldsymbol{z}^{k1}} = \frac{\partial J'}{\partial \boldsymbol{h}^k}\boldsymbol{h}^k\left(1 - \boldsymbol{h}^k\right) \tag{22}$$

当 $k = L$ 时, 上式可以化简为

$$\boldsymbol{d}^L = \left(\boldsymbol{h}^L - \boldsymbol{Y}\right)\boldsymbol{h}^L\left(1 - \boldsymbol{h}^L\right) \tag{23}$$

当 $k < L$ 时, 式 (22) 可以化简为

$$\boldsymbol{d}^k = \boldsymbol{W}_{k1}^{\mathrm{T}}\boldsymbol{d}^{k+1}\boldsymbol{h}^k\left(1 - \boldsymbol{h}^k\right) \tag{24}$$

步骤 3. 按照负梯度方向修改权系数 $\boldsymbol{W}$ 和偏置系数 $\boldsymbol{b}$.

$$\Delta \boldsymbol{W}_{k1} = -\eta\frac{\partial J'}{\partial \boldsymbol{W}_{k1}} = -\eta \boldsymbol{d}^k \boldsymbol{h}^{k-1} = -\eta \boldsymbol{W}_{k1}^{\mathrm{T}}\boldsymbol{d}^{k+1}\boldsymbol{h}^k\left(1 - \boldsymbol{h}^k\right)\boldsymbol{h}^{k-1} \tag{25}$$

$$\Delta \boldsymbol{W}_L = -\eta\frac{\partial J'}{\partial \boldsymbol{W}_L} = -\eta \boldsymbol{d}^L \boldsymbol{h}^{L-1} = -\eta\left(\boldsymbol{h}^L - \boldsymbol{Y}\right)\boldsymbol{h}^L\left(1 - \boldsymbol{h}^L\right)\boldsymbol{h}^{L-1} \tag{26}$$

同理, 可得偏置系数的修改量

$$\Delta \boldsymbol{b}_{k1} = -\eta \frac{\partial J'}{\partial \boldsymbol{b}_{k1}} = -\eta \sum_{j=1}^{R} \boldsymbol{d}^k \tag{27}$$

由于 $\boldsymbol{W} := \boldsymbol{W} + \Delta \boldsymbol{W}$, 可得权重系数和偏置系数更新公式如下

$$\boldsymbol{W}_{k1} := \boldsymbol{W}_{k1} + \Delta \boldsymbol{W}_{k1} = \boldsymbol{W}_{k1} - \eta \boldsymbol{d}^k \boldsymbol{h}^{k-1} \tag{28}$$

$$\boldsymbol{W}_{L} := \boldsymbol{W}_{L} + \Delta \boldsymbol{W}_{L} = \boldsymbol{W}_{L} - \eta \boldsymbol{d}^L \boldsymbol{h}^{L-1} \tag{29}$$

$$\boldsymbol{b}_{k1} := \boldsymbol{b}_{k1} + \Delta \boldsymbol{b}_{k1} = \boldsymbol{b}_{k1} - \eta \sum_{j=1}^{R} \boldsymbol{d}^k \tag{30}$$

同理, 权重 $\boldsymbol{W}_{k1}$ 和偏置 $\boldsymbol{b}_{k1}$ 迭代更新停止条件是目标函数取值不再变小或者达到最大迭代步数.

3.2.3　分类识别

训练好整个网络的参数后, 最后一层 Softmax 分类器可以用来对样本进行识别. 如果样本 $\boldsymbol{x}_{i,t}$ 是目标, 则对应的标签为 +1; 如果样本不是目标, 则对应的标签为 −1. 样本 $\boldsymbol{x}_{i,t}$ 是否为目标的概率计算如下

$$\boldsymbol{h}_{i,t}^{L} = \begin{bmatrix} p\left(\boldsymbol{y}_{i,t} = +1 \,|\boldsymbol{x}_{i,t}; \boldsymbol{W}_L\right) \\ p\left(\boldsymbol{y}_{i,t} = -1 \,|\boldsymbol{x}_{i,t}; \boldsymbol{W}_L\right) \end{bmatrix} = \frac{1}{\sum\limits_{c=-1}^{+1} \mathrm{e}^{\boldsymbol{W}_c^{\mathrm{T}} \boldsymbol{h}_{i,t}^{L-1}}} \begin{bmatrix} \mathrm{e}^{\boldsymbol{W}_{+1}^{\mathrm{T}} \boldsymbol{h}_{i,t}^{L-1}} \\ \mathrm{e}^{\boldsymbol{W}_{-1}^{\mathrm{T}} \boldsymbol{h}_{i,t}^{L-1}} \end{bmatrix} \tag{31}$$

其中, $\boldsymbol{W}_L$ 代表网络最后一层的权重, $\boldsymbol{h}_{i,t}^{L-1}$ 代表样本 $\boldsymbol{x}_{i,t}$ 在第 $L-1$ 层编码后的特征. 最后, 样本 $\boldsymbol{x}_{i,t}$ 的类别判决如下

$$Class\left(\boldsymbol{x}_{i,t}\right) = \arg \min_{c=-1,+1} p\left(\boldsymbol{y}_{i,t} = c \,|\boldsymbol{x}_{i,t}; \boldsymbol{W}_L\right) \tag{32}$$

4　实验内容和结果分析

本文采用人工合成 SAR 图像和真实 SAR 图像 — 北戴河海域 Radarsat-2 数据进行实验. 本文与其他几种常用的识别方法进行比较, 分别是支持向量机 (Support vector machine, SVM)[25]、联合稀疏表示分类算法 (Simultaneous orthogonal matching pursuit, SOMP)[26]、未做改进的堆叠自动编码器 (SAE)[24] 和稀疏编码–池化分类算法 (Lasso-Pooling)[27]. Lasso-Pooling 算法, 先采用 Lasso 算法对每个像素点稀疏编码得到稀疏向量, 然后对超像素分割后同一像素块的各个像素点稀疏向量采用最大值池化 (Max pooling) 得到一个稀疏向量, 最后利用联合稀疏表示分类器进行分类. 为了比较各个算法的识别效果, 采用总分类精度 (Overall accuracy, OA) 和 Kappa 系数 (Kappa coefficient, κ) 作为识别精度的评价指标. 总分类精度反映了整体图像识别的误差性, 计算公式如下

$$OA = \frac{\sum\limits_{i=1}^{C} M_{ii}}{N} \tag{33}$$

Kappa 系数反映了样本的识别结果与真实标记类型相一致的概率, 计算公式如下

$$\kappa = \frac{N\left(\sum\limits_{i=1}^{C} M_{ii}\right) - \sum\limits_{i=1}^{C}\left(\sum\limits_{j=1}^{C} M_{ij} \sum\limits_{j=1}^{C} M_{ji}\right)}{N^2 - \sum\limits_{i=1}^{C}\left(\sum\limits_{j=1}^{C} M_{ij} \sum\limits_{j=1}^{C} M_{ji}\right)} \tag{34}$$

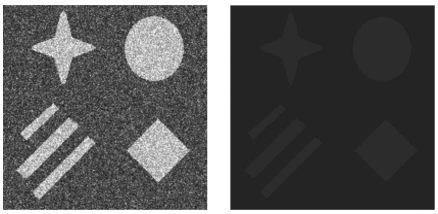
其中, M 表示误差混淆矩阵 (Confusion matrix), 是一个 $C \times C$ 的矩阵, C 是类别数量, M_{ji} 表示类别 j 被误分类为类别 i 的样本数量. N 表示总的样本数量.

本文实验还比较了各个算法的计算效率, 实验所用的计算平台为 PC (Intel Core i7 4.0 GHz CPU, 32 GB RAM), 操作系统为 Windows 7, 所有算法采用 Matlab 编写.

4.1　仿真实验

4.1.1　人工 SAR 图像

为了验证 DCSCN 网络在目标识别上的可行性和有效性, 设计人工 SAR 图像进行实验. 该图像中包含了多个不同形状的目标, 图像大小为 200 × 200. 通过对该图像的目标识别来说明 DCSCN 网络的普适性. 图 5 (a) 和图 5 (b) 分别给出了人工 SAR 图像及其相应的目标真值图.



(a) 原始图像 (a) Original image　　(b) 真值图 (b) Ground–truth image

图 5　人工 SAR 图像数据

Fig. 5　The artificial SAR data

4.1.2　网络参数设置

根据第 3 节介绍, 本文分别利用二维 Gabor 变换和 NSCT 来提取图像纹理特征和轮廓特征, 初始特征会对最后的识别精度有一定影响. 根据经验, 采用 5 个尺度、8 个方向的二维 Gabor 滤波器组来提取图像纹理特征, 利用 3 级 NSCT 分解得到 15 个

子带图像作为轮廓特征, 将这些特征堆叠起来构成 55 维的特征. 然后, 采用 TurborPixels 将图像分割为 1 000 个超像素块, 将各个小块特征组输入到 DCSCN 网络中进行特征学习和分类识别.

DCSCN 网络的结构和参数对识别精度有很大的影响. 考虑到算法的复杂度, 设计三层结构的 DCSCN, 前两层进行特征稀疏编码, 最高层进行分类识别. 根据实验经验, 网络的权重衰减系数 λ 设置为 0.005, 稀疏性惩罚因子 β 设置为 0.1, 稀疏性参数 ρ 设置为 0.2. 各层隐含单元数目对分类识别的精度有明显影响, 因此这里详细描述了隐含单元数目的确定过程. 由于最高层用于识别, 最高层隐含单元数等于类别数, 所以只需要分析前两层隐含单元数目对识别结果的影响. 实验中, 随机选取 20 % 的超像素块作为训练集, 其他超像素块用来验证算法的识别精度. 通过改变其中一个参数的同时固定另一个参数, 来讨论隐含单元数目对结果的影响.

图 6 (a) 和图 6 (b) 分别给出了总分类精度和 Kappa 系数随着两层隐含单元数目变化的变化情况. 实验中, 第一层隐含单元数目在 20 到 150 之间变化, 第二层隐含单元数目在 20 到 100 之间变化. 实验结果表明, 当前两层的隐含单元数目分别为 100 和 40 时, 网络的识别精度最高. 当第二层隐含单元数目一定, 第一层隐含单元数目取值在 90 到 110 之间时, 识别的精度优于其他情况的结果. 即第一层隐含单元数目比输入样本特征维度要高, 其原因是输入的样本初始特征是非线性的, 当隐含单元数目比输入单元数目大时, 网络第一层相当于对特征进行高维空间映射, 使得特征更加线性化、更利于分类识别. 当第一层隐含单元数目一定, 第二层隐含单元数目取值在 30 到 40 之间时, 识别的精度优于其他情况的结果. 网络第二层的作用相当于对特征进行降维, 降维处理不能降得太低或者太高. 这是因为降维维度过低会压缩掉有用信息, 而降维维度较高时训练分类器容易发生过拟合现象.

4.1.3 对比实验结果与分析

为了证明 DCSCN 在 SAR 图像目标识别中的效果, 本文与 SVM 算法、SOMP 算法、Lasso-Pooling 算法及未做改进的 SAE 网络进行对比. 对比算法采用的样本特征与输入到 DCSCN 网络的特征一致, 都是由二维 Gabor 滤波器和 NSCT 滤波器得到的 55 维特征. 根据前述实验分析, DCSCN 网络为三层, 前两层的隐含单元数目分别为 100 和 40. SVM 采用了高斯径向基作为核函数, 参数通过十次交叉验证得到. SOMP 算法、Lasso-Pooling 算法中的主要参数为稀疏度和重建误差极值, 分别设定为 10 和 10^{-3}. SAE 网络采用了与本文提出的 DCSCN 网络相同的结构, 都设计了三层结构的网络, 前两层的隐含单元数目分别为 100 和 40, 最高层进行分类识别. 通过五次实验来分析各个算法的识别效果, 每次实验随机选取 20 % 的样本作为训练集, 其他样本作为测试集, 计算五次实验结果的均值和标准差. 表 1 给出了各种算法识别精度和算法效率的结果对比.

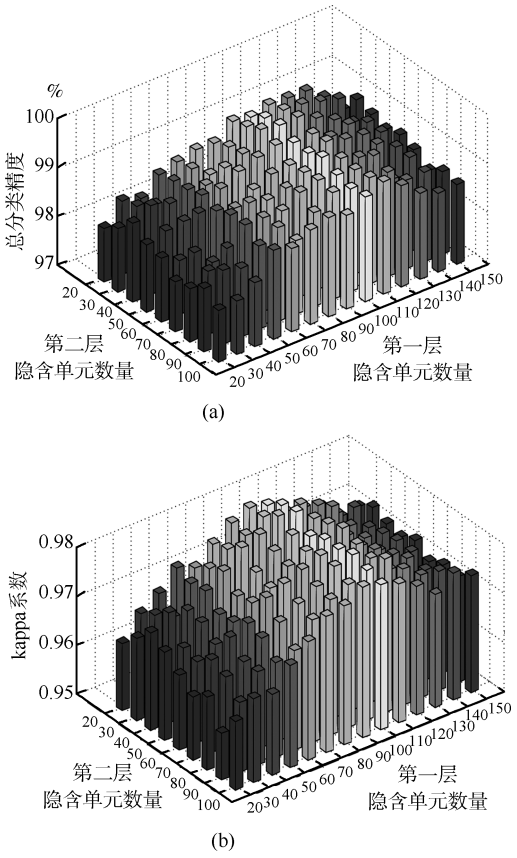


图 6 DCSCN 网络参数设置

Fig. 6 Parameter setting of the DCSCN

根据表 1 结果, DCSCN 识别精度要明显高于其他四种算法. 对比 SAE 和 DCSCN 实验结果, 基于 SAE 提出的 DCSCN 网络可以提高目标识别精度, 说明本文对 DCSCN 网络的设计是有效的. DCSCN 网络考虑了近邻像素的协同性, 通过联合优化得到近邻像素特征表示, 能平滑掉相干斑噪声, 从而有效降低相干斑噪声对识别结果的影响. 由于 SVM 算法、SOMP 算法、Lasso-Pooling 算法和 DCSCN 网络输入的初始特征都是由二维 Gabor 滤波器和 NSCT 滤波器分别提取的纹理特征和轮廓特征, 而 DCSCN 网络识别精度得到明显提高, 这说明经过 DCSCN 网络优化后的不同类别特征的区分度更大, 更适合于分类器进行识别. 对比各个算法的计算耗时发现, DCSCN 网络计算效率比 SVM 算法略差, 但要优于其他方法. 这是由于 SOMP 算法和 Lasso-Pooling 算法的计算复杂度随着样本数成高次方增长, 而 DCSCN 网络计算复杂度随迭代次数、样本数和网络节点数线性变化. 因此, 基于深度学习神经

表 1　各个算法在人工 SAR 图像上识别结果对比

Table 1　Recognition performance comparison of different algorithms on the artificial SAR image

算法	OA (%)	κ	计算效率 (s)
SVM[25]	83.574 ± 0.132	0.585 ± 0.006	**166.623 ± 1.854**
SOMP[26]	92.803 ± 0.588	0.775 ± 0.019	568.204 ± 3.376
SAE[24]	95.261 ± 0.375	0.858 ± 0.011	312.044 ± 2.975
Lasso-Pooling[27]	97.972 ± 0.183	0.938 ± 0.005	757.605 ± 6.334
DCSCN	**99.457 ± 0.079**	**0.983 ± 0.002**	192.630 ± 4.481

网络的 DCSCN 网络计算复杂度低于基于稀疏编码的 SOMP 算法及 Lasso-Pooling 算法.

图 7 显示了五种算法的目标识别结果图. 对比五幅图可知, DCSCN 网络表现出更好的识别结果, 误分类点比其他方法更少. 同时发现, SOMP 算法和 Lasso-Pooling 算法结果图的孤立误分类点也很少, 这是因为 SOMP 算法、Lasso-Pooling 算法和本文 DCSCN 网络都考虑了近邻像素的协同性, 而通过空间近邻协同可以平滑掉相干斑噪声, 所以这三种算法对相干斑噪声抑制作用明显. 由于 DCSCN 算法有优异的特征表示能力, 所以其识别精度明显高于 SOMP 算法和 Lasso-Pooling 算法. 此外, SVM 算法和 SAE 算法结果的孤立误分类点很多, 原因是这两种算法是基于像素的分类识别, 未考虑空间协同相关性, 相关斑噪声对识别产生很大的影响.

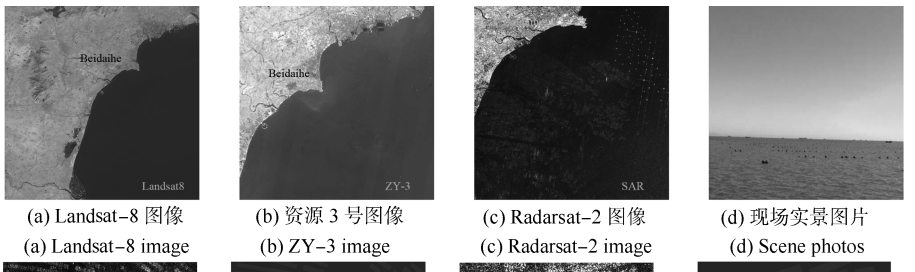
4.2　浮筏识别实验

为了验证 DCSCN 网络在海洋浮筏养殖目标识别的有效性, 选取北戴河海域作为研究区域. 图 8 (a) 和图 8 (b) 是北戴河海域的两个不同传感器的多光谱图像, 分别是 Landsat-8 卫星图像和资源三号卫星图像. 图 8 (c) 为 Radarsat-2 卫星的 SAR 图像, 极化方式为 HH, 分辨率为 5 m. 比较这三幅图可以发现, 多光谱图像不能有效地获取北戴河海域的浮筏目标, 而在 SAR 图像中有清晰的浮筏目标. 这是

(a) SVM　(b) SOMP　(c) SAE　(d) Lasso–Pooling　(e) DCSCN

图 7　人工 SAR 图像目标识别结果

Fig. 7　Target recognition results of the artificial SAR image



(a) Landsat–8 图像　(b) 资源 3 号图像　(c) Radarsat–2 图像　(d) 现场实景图片

(a) Landsat–8 image　(b) ZY–3 image　(c) Radarsat–2 image　(d) Scene photos

(e) 图像 1　(f) 真值图 1　(g) 图像 2　(h) 真值图 2

(e) First original image　(f) First ground–truth image　(g) Second original image　(h) Second ground–truth image

图 8　北戴河区域 SAR 数据

Fig. 8　The original SAR data of Beidaihe area

由于海洋散射信号较弱, 光谱成像卫星数据难以获取浮筏信息, 而 SAR 是主动式微波成像, 利于获取海洋信息. 图 8 (d) 为现场测绘时拍摄的实景图片. 从图 8 (c) 选取两块海域的图像来验证本文提出算法的效果, 选取的图像 1 和图像 2 分别如图 8 (e) 和图 8 (g), 图像大小分别为 $1\,275\times1\,140$ 和 600×600. 根据现场测绘的数据绘制出相应的真值图如图 8 (f) 和 8 (h).

在 SAR 图像 1 上进行实验, 先做几何校正、灰度拉伸和增强 Lee 滤波去噪, 再采用二维 Gabor 变换和 NSCT 提取图像纹理特征和轮廓特征. 与仿真实验相同, 采用 3 级 NSCT 分解和 5 个尺度、8 个方向的二维 Gabor 滤波器得到 55 维的特征. 采用 TurborPixels 将图像分割为 40 000 个超像素块, 将各个小块输入到 DCSCN 网络中进行特征学习并分类识别. DCSCN 网络共三层, 前两层的隐含层单元数目分别设置为 100 和 40. 网络其他参数的设置与仿真实验相同. 同样, 为说明本文算法在海洋浮筏养殖目标识别的优越性能, 本文与 SVM、SOMP、Lasso-Pooling 和 SAE 进行了比较. 表 2 给出了各个算法识别精度和算法效率的对比, 这里每个算法效率的统计都是指整个识别过程, 即从输入图像到最后得出分类识别结果. 图 9 给出了各个算法目标识别结果图.

根据表 2 结果, DCSCN 网络的总分类精度和 Kappa 系数都明显好于其他四种算法. 结果表明, 本文 DCSCN 网络具有优异的特征表示能力, 可以得到区分性更大、更适合分类器的特征. 同时, DCSCN 通过空间近邻协同约束, 有效地抑制 SAR 图像相干斑噪声的影响, 所以在分类精度上明显优于现有的三种方法. 此外, DCSCN 的算法效率更优, 比其他四种算法耗时都少. 这是由于 DCSCN 网络计算复杂度随迭代次数、样本数和网络节点数线性变化, 而 SVM 算法、SOMP 算法和 Lasso-Pooling 算法的计算复杂度随着样本数成高次方增长, 所以 DCSCN 网络计算复杂度较低. 在图像尺寸较大、样本较多的时候, DCSCN 能够有效地降低计算量, 使得计算效率优于其他算法.

对比图 9 各个算法的结果图, DCSCN 识别效果要明显优于其他四种算法的结果. 分析发现, SVM 和 SAE 识别结果图存在较多误分类噪点, 而本文算法、SOMP 算法和 Lasso-Pooling 算法误分类噪点较少. 这是由于这三种算法都采用了近邻协同约束, 可以有效抑制 SAR 图像相干斑噪声. 同时, 本文算法的误分类噪点更少, 说明本文算法在抑制相干斑噪声方面的效果更好. 本文算法要优于 SAE 算法结果, 说明稀疏约束和协同约束能使得深度网络学到区分性更大的特征. 此外, 在实际 SAR 图像上的实验结果与人工 SAR 图像的实验结果得到了一致的结论, 充分证明了算法的有效性和适应性.

由于海洋不同区域的后向散射特性存在差异, 为了说明本文 DCSCN 网络对不同海域浮筏的识别能力, 对 SAR 图像 2 进行相同的实验. 与 SAR 图像 1 上实验一样, 先进行几何校正、灰度拉伸和增强 Lee 滤波去噪, 再采用 5 个尺度、8 个方向的二维 Gabor 滤波器和 3 级 NSCT 滤波器提取出 55 维的特征, 然后采用 TurborPixels 将图像分割为 10 000 个超像素块, 最后将各个小块输入到三层结构的 DCSCN 网络中进行特征学习和分类识别. 表 3 给出了各个算法识别精度和算法效率的对比, 图 10 给出了各个算法目标识别结果图.

分析图 10 和表 3 结果, DCSCN 网络的识别精

表 2 各个算法在 SAR 图像 1 浮筏识别结果对比

Table 2 Recognition performance comparison of different algorithms on the first SAR image

算法	OA (%)	κ	计算效率 (s)
SVM[25]	79.804 ± 0.375	0.653 ± 0.004	528.170 ± 4.266
SOMP[26]	86.243 ± 0.615	0.708 ± 0.007	12 411.229 ± 62.799
SAE[24]	81.960 ± 0.392	0.686 ± 0.005	7 963.060 ± 29.862
Lasso-Pooling[27]	86.925 ± 0.073	0.692 ± 0.006	15 796.110 ± 79.618
DCSCN	**89.046 ± 0.433**	**0.759 ± 0.005**	**458.050 ± 2.145**

表 3 各个算法在 SAR 图像 2 浮筏识别结果对比

Table 3 Recognition performance comparison of different algorithms on the second SAR image

算法	OA (%)	κ	计算效率 (s)
SVM[25]	87.714 ± 1.026	0.693 ± 0.007	794.139 ± 7.148
SOMP[26]	93.261 ± 0.068	0.821 ± 0.021	6 432.545 ± 39.582
SAE[24]	93.179 ± 0.240	0.816 ± 0.003	1 245.995 ± 16.942
Lasso-Pooling[27]	93.614 ± 0.175	0.832 ± 0.042	8 381.801 ± 44.702
DCSCN	**98.762 ± 0.172**	**0.966 ± 0.003**	**268.498 ± 1.633**

图 9　SAR 图像 1 浮筏识别结果

Fig. 9　Target recognition results of the first SAR image

图 10　SAR 图像 2 浮筏识别结果

Fig. 10　Target recognition results of the second SAR image

度和算法效率都优于其他四种算法, 这一结论和前面分析一致. 这充分说明, 不同海域即使海洋后向散射特性有差别的情况下, DCSCN 网络都能达到较好的浮筏养殖识别结果.

5　结论

本文针对海洋遥感 SAR 图像严重的相干斑噪声污染和 SAR 图像特征单一的问题, 提出一种深度协同稀疏编码网络, 进行 SAR 图像特征表示和目标识别. 通过人工 SAR 图像和北戴河两个海域浮筏养殖 SAR 图像的实验, 证明本文所提网络的有效性和适用性. 结果表明, 深度协同稀疏编码网络不仅可以有效优化近邻像素特征组来提高识别精度, 而且能有效抑制 SAR 图像相干斑噪声的影响. 此外, 深度协同稀疏编码网络计算效率较高, 在图像尺寸较大时表现出更好的效率.

References

1　Yu Hang, Jiao Li-Cheng, Liu Fang. Context based unsupervised hierarchical iterative algorithm for SAR segmentation. *Acta Automatica Sinica*, 2014, **40**(1): 100−116
(余航, 焦李成, 刘芳. 基于上下文分析的无监督分层迭代算法用于 SAR 图像分割. 自动化学报, 2014, **40**(1): 100−116)

2　Zhao Ming-Bo, He Jun, Fu Qiang. Survey on fast CFAR detection algorithms for SAR image targets. *Acta Automatica Sinica*, 2012, **38**(12): 1885−1895
(赵明波, 何峻, 付强. SAR 图像 CFAR 检测的快速算法综述. 自动化学报, 2012, **38**(12): 1885−1895)

3　Chu Jia-Lan, Zhao Dong-Zhi, Zhang Feng-Shou. Wakame raft interpretation method of remote sensing based on association rules. *Remote Sensing Technology and Application*, 2012, **27**(6): 941−946
(初佳兰, 赵冬至, 张丰收. 基于关联规则的裙带菜筏式养殖遥感识别方法. 遥感技术与应用, 2012, **27**(6): 941−946)

4　Fan Jian-Chao, Zhang Feng-Shou, Zhao Dong-Zhi, Wen Shi-Yong, Wei Bao-Quan. Floating raft aquaculture extraction based on high resolution satellite remote sensing SAR images. In: Proceedings of the 2nd Symposium on Disaster Risk Analysis and Management in Chinese Littoral Regions. 2014. 59−63
(范剑超, 张丰收, 赵冬至, 文世勇, 卫宝泉. 基于高分辨率卫星遥感 SAR 图像的海洋浮筏养殖信息提取. 见: 第二届中国沿海地区灾害风险分析与管理学术研讨会. 2014. 59−63)

5　Qazi W A, Emery W J, Fox-Kemper B. Computing ocean surface currents over the coastal california current system using 30-min-lag sequential SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, **52**(12): 7559−7580

6　Pan De-Lu, Lin Ming-Shen, Mao Zhi-Hua. *Microwave Remote Sensing and Application of Ocean*. Beijing: Ocean Press, 2013.
(潘德炉, 林明森, 毛志华. 海洋微波遥感与应用. 北京: 海洋出版社, 2013.)

7　Collins M J, Allan J M. Modeling and simulation of SAR image texture. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, **47**(10): 3530−3546

8　Liu C J, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, 2002, **11**(4): 467−476

9　Do M N, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 2005, **14**(12): 2091−2106

10　Da Cunha A L, Zhou J P, Do M N. The nonsubsampled contourlet transform: theory, design, and applications. *IEEE Transactions on Image Processing*, 2006, **15**(10): 3089−3101

11　Yang X H, Jiao L C. Fusion algorithm for remote sensing images based on nonsubsampled contourlet transform. *Acta Automatica Sinica*, 2008, **34**(3): 274−281

12 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527−1554

13 Salakhutdinov R, Tenenbaum J B, Torralba A. Learning with hierarchical-deep models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(8): 1958−1971

14 Goh H, Thome N, Cord M, Lim J H. Learning deep hierarchical visual feature coding. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(12): 2212−2225

15 Zheng Yin, Chen Quan-Qi, Zhang Yu-Jin. Deep learning and its new progress in object and behavior recognition. *Journal of Image and Graphics*, 2014, **19**(2): 175−184
(郑胤, 陈权崎, 章毓晋. 深度学习及其在目标和行为识别中的新进展. 中国图象图形学报, 2014, **19**(2): 175−184)

16 Chen X Y, Xiang S M, Liu C L, Pan C H. Vehicle detection in satellite images by hybrid deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 2014, **11**(10): 1797−1801

17 Chen Y S, Lin Z H, Zhao X, Wang G, Gu Y F. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, **7**(6): 2094−2107

18 Feng Bo, Chen Bo, Wang Peng-Hui, Liu Hong-Wei. Feature extraction method for radar high resolution range profile targets based on robust deep networks. *Journal of Electronics & Information Technology*, 2014, **36**(12): 2949−2955
(冯博, 陈渤, 王鹏辉, 刘宏伟. 基于稳健深层网络的雷达高分辨距离像目标特征提取算法. 电子与信息学报, 2014, **36**(12): 2949−2955)

19 Li Shuai, Xu Yue-Lei, Ma Shi-Ping, Ni Jia-Cheng, Wang Kun. SAR target recognition using wavelet transform and deep sparse autoencoders. *Video Engineering*, 2014, **38**(13): 31−35
(李帅, 许悦雷, 马时平, 倪嘉成, 王坤. 基于小波变换和深层稀疏编码的 SAR 目标识别. 电视技术, 2014, **38**(13): 31−35)

20 Rai P, Khanna P. A gender classification system robust to occlusion using gabor features based $2D^2$PCA. *Journal of Visual Communication and Image Representation*, 2014, **25**(5): 1118−1129

21 Zhang Qiang, Guo Bao-Long. Fusion of multi-sensor images based on the nonsubsampled contourlet transform. *Acta Automatica Sinica*, 2008, **34**(2): 135−141
(张强, 郭宝龙. 基于非采样 Contourlet 变换多传感器图像融合算法. 自动化学报, 2008, **34**(2): 135−141)

22 Yan C M, Guo B L, Yi M. Fast algorithm for nonsubsampled contourlet transform. *Acta Automatica Sinica*, 2014, **40**(4): 757−762

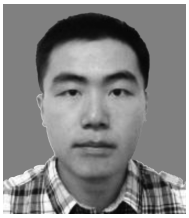
23 Levinshtein A, Stere A, Kutulakos K N, Fleet D J, Dickinson S J, Siddiqi K. TurboPixels: fast superpixels using geometric flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(12): 2290−2297

24 Ng A. Sparse autoencoder [Online], available: http://web.stanford.edu/class/cs294a/sae/, November 8, 2015.

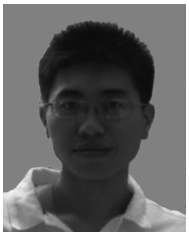
25 Chang C C, Lin C J. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2011, **2**(3): Article No. 27

26 Chen Y, Nasrabadi N M, Tran T D. Simultaneous joint sparsity model for target detection in hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 2011, **8**(4): 676−680

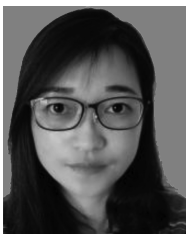
27 Mairal J, Bach F, Ponce J. Sparse modeling for image and vision processing. *Foundations Trend in Computer Graphics Vision*, 2014, **8**(2−3): 85−283



耿 杰 大连理工大学电子信息与电气工程学部博士研究生. 主要研究方向为 SAR 图像处理, 模式识别.
E-mail: gengjie@mail.dlut.edu.cn
(**GENG Jie** Ph. D. candidate at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His research interest covers SAR image processing and pattern recognition.)



范剑超 国家海洋环境监测中心, 大连理工大学电子信息与电气工程学部副研究员. 主要研究方向为神经网络, 模式识别和遥感图像处理. 本文通信作者.
E-mail: fjchaonmemc@163.com
(**FAN Jian-Chao** Associate professor at National Marine Environment Monitoring Center and Dalian University of Technology. His research interest covers neural network, pattern recognition, and remote sensing image processing. Corresponding author of this paper.)



初佳兰 国家海洋环境监测中心工程师. 主要研究方向为海域动态卫星遥感应用.
E-mail: jlchu@nmemc.org.cn
(**CHU Jia-Lan** Engineer at National Marine Environment Monitoring Center. Her research interest covers sea dynamic surveillance remote sensing application.)

王洪玉 大连理工大学电子信息与电气工程学部教授. 主要研究方向为图像处理, 模式识别和无线传感器网络.
E-mail: whyu@dlut.edu.cn
(**WANG Hong-Yu** Professor at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His research interest covers image processing, pattern recognition, and wireless sensor networks.)