

doi: 10.3969/j.issn.0490-6756.2020.01.014

多层局部感知卷积神经网络的高光谱图像分类

池涛^{1,2}, 王洋¹, 陈明¹

(1. 上海海洋大学 农业部渔业信息重点实验室, 上海 201306;

2. 喀什大学计算机科学技术学院, 喀什 944104)

摘要: 针对高光谱图像分类中光谱特征的高度非线性问题, 提出一种基于多层感知器卷积层和批标准化层的改进卷积神经网络模型, 提高模型在光谱域处理的非线性特征提取能力. 该算法通过构建七层网络结构, 实现多层局部感知结构, 逐个像素对光谱信息开展分析, 区分不同目标物的光谱信息, 将全光谱段集合作为输入, 舍去空间信息, 利用动量梯度下降训练算法对多层局部感知卷积神经网络训练, 实现对不同目标物体光谱特征的提取与分类. 实验中, 采用两组高光谱遥感影像进行对比分析, 以 Pavia University 数据集为例, 在 3 600 个训练样本情况下, 测试集为 1 800 个样本, 本文方法正确率为 90.23%, LeNet-5 正确率为 87.94%, Linear-SVM 正确率为 90.00%; 在 21 000 个训练样本情况下, 测试集为全部样本, 本文方法正确率为 97.23%, LeNet-5 正确率为 96.64%, Linear-SVM 正确率为 92.40%. 实验结果表明, 在训练集较小的情况下, 本文方法优于传统神经网络, 能有效提取数据特征, 并且在精度上和计算成本上略优于在小样本分类中具有高效和鲁棒性良好的 SVM 算法. 在大规模训练集时, 本文方法表现出良好的学习能力, 在分类精度上优于 LeNet-5. 本文提出的多层局部感知网络结构增强了对非线性特征的学习能力, 无论训练集规模大小, 都比传统的 SVM 和一般的深度学习网络更能有效的利用高光谱图像中的逐像素点的光谱域信息, 能有效提高分类精度.

关键词: 高光谱图像; 卷积神经网络; 支持向量机; 分类; 非线性特征; 多层局部感知

中图分类号: P237

文献标识码: A

文章编号: 0490-6756(2020)01-0103-10

Hyperspectral image classification of multi-layer local perceptual convolutional neural networks

CHI Tao^{1,2}, WANG Yang¹, CHEN Ming¹

(1. Ministry of Agriculture Key Laboratory of Fishery Information, Shanghai Ocean University, Shanghai 201306, China;

2. Department of Computer Science and Technology, Kashgar University, Kashgar 944104, China)

Abstract: To solve the problems on high nonlinearity in spectral features in hyperspectral image classification, a classification algorithm, based on multi-layer perceptron convolutional layer and batch normalization layer improved convolutional neural network in spectral domain processing, is proposed to improve the nonlinear feature extraction ability. By constructing a seven-layer network structure, the algorithm implements a multi-layer local sensing structure, analyzes the spectral information pixel by pixel, distinguishes the spectral information of different targets, takes the full spectrum segment set as input, discards the spatial information, and uses the momentum gradient descent training. The algorithm trains

收稿日期: 2019-01-05

基金项目: 国家自然科学基金(61561027); 上海市自然科学基金(16ZR1415100)

作者简介: 池涛(1976-), 男, 博士后, 教授, 主要研究领域为农业信息化、物联网. E-mail: tchi@shou.edu.cn

通讯作者: 王洋. E-mail: 1034394549@qq.com

multi-layer local perceptual convolutional neural networks to realize the extraction and classification of spectral features of different target objects. In the experiment, two sets of hyperspectral remote sensing images are used for comparative analysis. Taking the Pavia University data set as an example, in the case of 3 600 training samples, the test set is 1 800 samples, the accuracy of the proposed method is 90.23% and the accuracy of the LeNet-5 and Linear-SVM method are 87.94% and 90.00% respectively. In the case of 21 000 training samples, the test set is all samples, the accuracy is 97.23%, 96.64% and 92.40% respectively for the proposed method, LeNet-5 and Linear-SVM method. The experimental results show that the proposed method is superior to the traditional neural network in the case of small training set, which can effectively extract the data features, and is superior to SVM algorithm for the small sample classification in terms of accuracy and computational cost. In the large-scale training set, this method shows good learning ability and is superior to LeNet-5 in classification accuracy. The multi-layer local perceptual network structure proposed in this paper enhances the learning ability of nonlinear features, it can utilize hyperspectral images much more effectively than traditional SVM and general deep learning networks, both in small sample classification and large sample classification. The spectral domain information of the pixel-by-pixel point can effectively improve the classification accuracy.

Keywords: Hyperspectral image; Convolutional neural network; Support vector machine; Classification; Nonlinear feature; Multi-layer local perception

1 引言

高光谱影像 HSI(hyperspectral image)^[1]含有丰富的光谱信息和空间信息在空间观测和目标识别中有着重要地位. 它广泛使用在现代军事、精准农业以及环境监测等诸多领域.

高光谱图像分类^[2]是高光谱影像处理和应用的的重要组成部分. 高光谱图像包含了上百个光谱通道,并且具有高维特性、波段间高度相关性以及光谱混合等特性. 这使得高光谱图像分类成为一个巨大的挑战. 由于光谱信号的高维度、不确定性、信息冗余以及表面覆盖的异构性和同质性,使得高光谱数据结构具有高度非线性的特征,这导致基于统计模式识别的分类模型难以直接分类和识别高光谱数据. 同时,样本数量有限并且质量好坏不一,从而导致分类器模型的参数难以估计或者估计不准确. 在这些情况下,高光谱图像的准确分类需要建立复杂的数学模型,从而可以真实的反映数据的内在本质. 模型训练过程需要繁琐的预处理和后处理. 通过研究和实践表明,传统的遥感影像技术在分析高光谱图像时无法解决高维数据小样本识别、高光谱图像分类精度等问题,目前机器学习、计算机视觉以及模式识别的理论和方法成为高光谱图像分类的重要技术手段.

为了实现高效、高精度的高光谱图像分类,曾在高光谱图像上采用 K 邻近算法、主成分分析

(principal components analysis, PCA)、支持向量机 SVM(support vector machine)等方法. 随着特征提取和分类方法改进,提出了光谱空间分类法、局部 Fisher 判别法、U-Net 卷积神经网络分割图像法^[2]等多种方法,取得了较好的结果. 其中,支持向量机 SVM^[3]被认为是高效和鲁棒性良好的方法,适用于小规模训练样本. 它是一种监督学习模型,通过非线性映射将样本空间映射到一个高维甚至无穷维的特征空间(Hilbert 空间)中. 将原始样本空间中的非线性可分问题变成比原始样本空间更高维度空间中的线性可分问题,并在高光谱图像分类中有着良好的效率和正确率.

以卷积神经网络 CNN(convolutional neural network)^[4]为代表的深度学习技术在图像分类和模式识别方面具有良好的性能. 随着神经网络的发展,CNN 越来越多的被应用于高光谱数据的分类. 例如,多层感知器 MLP(multilayer perception)^[5]和径向基函数 RBF(radial basis function)^[6]的应用.

CNN 方法在机器视觉方面分类效果优于传统的 SVM 分类器. 实际上,在高光谱图像分类任务中,SVM 在分类精度、时间复杂度和空间复杂度方面优于传统的神经网络的,但不能因此放弃神经网络这样强大的工具. 在本文实验过程中,当训练数据足够大时,本文提出的改进型卷积神经网络的精度可以优于 SVM.

文献[7]提出的卷积神经网络, 只有 Alexnet 网络参数的 10% 大小, 在计算机视觉上的分类精度和算法效率上优于 Alexnet 网络. 文献[1]中将 CNN 直接应用于高光谱分类中, 本文通过实现该论文中方法作为改进前 CNN 网络. 本文通过文献[7]中提高卷积神经网络的非线性特征学习能力的改进策略应用对高光谱影像分类的 CNN 方法改进, 改进后的网络结构能够有效提高对高光谱影像非线性特征的学习能力, 有效提高分类精度[7].

2 多层局部感知卷积神经网络

针对高光谱图像每个像素的不同频带的特征具有非线性, 本文提出在传统 CNN 网络的基础上通过改变卷积层感知器和激活函数, 并且引入批标准化层, 实现多层局部感知网络结构, 增强其对非线性特征的学习能力.

2.1 增强非线性特征学习的改进方法

2.1.1 多层感知器卷积层 (Mlpconv layer) 经典卷积神经网络中的卷积层实际使用线性滤波器对图像进行内积运算, 与单层神经网络类似, 卷积滤波器实际就是一种广义线性模型 GLM (generalize linear model), GLM 的抽象能力相对较低. 例如 Lenet-5[8] 是一个经典的卷积神经网络结构, 它的卷积层为线性滤波器.

线性卷积层假设其上层一个输入为 x_m , 步长为 s , 卷积核大小为 k , 激活函数为 f , 输出为 n (指代由一个样本数据通过 n 个神经元输出 n 个数据, 在这里没有给定具体数值仅代表这个变量与图 1 中下标 n 对应, 以更好描述一个线性卷积层结构和运算过程, 在 2.2 节中给出模型的具体数值), i 代表 $0 \sim n$ 之间的第 i 个输出数据或第 i 个神经元; j 代表输出数据中第 j 个基本数据; 神经元激活函数为 f ; 权重矩阵 W ; 偏置为 b (每个神经元由激活函数、权重矩阵和偏置三部分组成); T 为转置, 该卷积层的过程如图 1 所示.

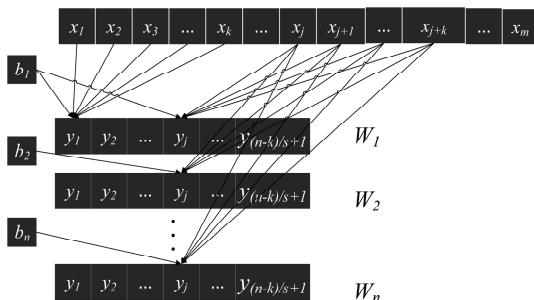
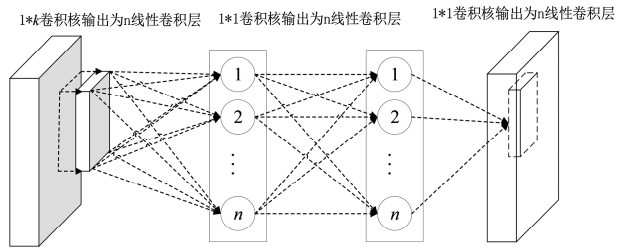


图 1 线性卷积层
Fig. 1 Linear convolution layer

则线性卷积层中第 i 个神经元的计算公式如下.

$$y_j = f(\omega_k^T x_{j,j+k} + b) \tag{1}$$

本文对卷积神经网络的卷积层进行了较大的改进, 使用 MLP (多层感知器) 作为卷积层模型, 从而提高非线性特征学习能力. 在实际使用过程中, 通过在经典卷积层后增加两层卷积核为 1×1 的卷积层实现. 由于采用了 Mlpconv 卷积层, 网络在前面卷积阶段特征提取效果提高, 极大地减少了参数和过拟合风险.



输出为 n 多层感知器卷积层

图 2 多层感知器卷积层
Fig. 2 Mlpconv layer

从图 2 可以看到, Mlpconv 卷积层网络结构就是在图 1 的基础上增加两层卷积核为 1×1 , 输出为 n , 步长为 1, 相同激活函数的卷积层. 简而言之, 多层感知器卷积层使用 Mlp 网络结构, 对局部感受野的神经元进行更复杂的运算操作, 可以对非线性特征更加敏感, 而线性卷积层, 局部感受野的运算仅仅只是一个单层的神经网络.

2.1.2 激活函数层 (activation function layer) 所谓激活函数 (activation function), 就是在人工神经网络的神经元上运行的函数, 负责将神经元的输入映射到输出端. 通过在激活函数给神经元引入了非线性因素, 从而使得神经网络可以更加逼近任何非线性函数, 这样的神经网络就可以被应用于更多的非线性模型中.

常用的 Relu 激活函数 (the rectified linear unit), 被使用于隐藏层神经元输出. 与线性激活函数相比, Relu 激活函数克服了梯度消失、训练速度等问题, 但在实验过程中, 大量局部神经元出现饱和和情况 (也称神经元死亡), 无法有效学习特征. 因而本文方法采取 Relu 激活函数 (the rectified linear unit) 的变种, 带泄露线性整流 (Leaky ReLU)[9]. 当输入 x 为负时, 带泄露线性整流函数 (Leaky ReLU) 的梯度为一个常数 $\lambda \in (0, 1)$, 而不是 0. 当输入为正时, 带泄露线性整流函数和 Relu 激活函数保持一致. 数学表达方式:

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ \lambda x, & \text{if } x \leq 0 \end{cases} \quad (2)$$

本文使用 MATLAB 神经网络工具箱函数提供的默认值 0.01 作为常数 λ 的取值,同时参考多数神经网络模型取较小的数值作为常数 λ 的取值就可以避免饱和现象. 通过实验测试,当常数 $\lambda \in (0.005, 0.1)$ 时既可以避免饱和现象,同时加速模型收敛. 当取值过大训练时间增加,当取值过小则无法避免饱和现象影响正确率.

2.1.3 批标准化层(batch normalization) 本文在每个卷积层中激活函数层之前增加批标准化层(batch normalization)^[10],实际过程表现为对卷积层的神经元批量标准化处理,对于在非线性函数映射后不断向取值区间饱和和区靠拢的输入分布强行映射到比较标准的正态分布,使非线性变换函数的输入落入对输入值比较敏感的取值区间,从而解决梯度消失问题. 设同一最小批次 B 有 m 个样本: $B = \{x_1, \dots, x_m\}$; 神经元激活函数为 f ; 权重矩阵 W ; 偏置为 b ; 左箭头符号为命题的“条件”运算,一般可以根据实际数据再做调整,在本文中该符号与等号含义相同. 本文通过实验测试最小批次样本数为 96 个,可以使多数样本数值调整到合理区间范围内,能有效提取特征,提高训练模型分类精度.

最小批次 B 均值如下式.

$$\mu_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad (3)$$

最小批次 B 方差如下式.

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (4)$$

标准化处理如下式.

$$\hat{x} \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (5)$$

卷积激活如下式

$$y_i = f(\omega \hat{x}_i + b) \equiv \text{BN}_{\omega, b}(x_i) \quad (6)$$

传统线性卷积层中神经元做卷积和激活操作,改进后卷积层对同一神经元同一批次做批标准化处理,然后在标准化后的样本上进行卷积和激活操作. 本文将批标准化层(batch normalization)引入卷积层中,可以增加训练速度,加快收敛过程,减小局部神经元死亡风险.

2.2 多层局部感知卷积神经网络分类器

2.2.1 多层局部感知 CNN 网络模型 卷积神经网络是一种人工神经网络. 它实际上是一种前馈式神经网络,一般包含卷积层(convolutional lay-

ers)、池化层(pooling layers)和全连接层(fully connect layers). 每个隐藏层神经元通过连接输入的一部分并非全连接,通过利用过滤器的局部敏感性实现模型对于目标不同空间域的相关性的学习. 经典的卷积神经网络结构为 LeNet-5,每个隐藏层连接一部分输入图像区域. 该网络结构受猫的脑部视觉皮层工作原理启发,通过模拟视觉皮层细胞对局部的视野非常敏感,敏感区域也是感受野,大量的感受野感受整个视野,同时这也会造成神经元过多,参数过大问题,这些感受野通过共享权重的方式解决该问题. 对于大多数二维图像识别的卷积神经网络结构,输入层之后是由卷积层和最大池化层交替叠加构成,接近输出层由全连接层构成. 典型的卷积网络结构如图 3 所示.

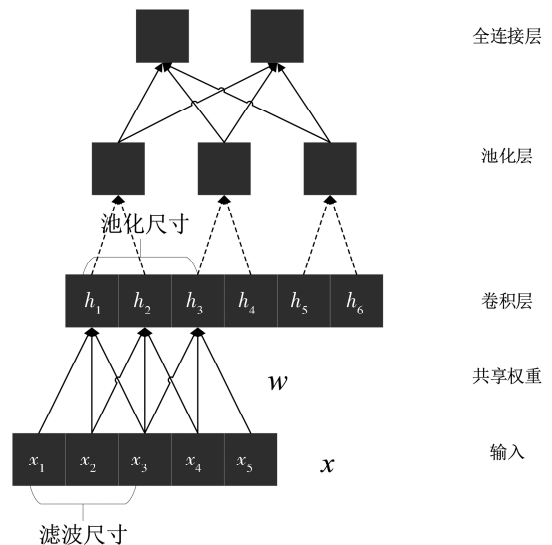


图 3 典型的卷积神经网络结构
Fig. 3 Typical convolutional neural network structure

在卷积神经网络中用多层感知器卷积层(Mlpcnv Layer)代替线性卷积层(linear convolution layer)提高卷积神经网络对于非线性特征的学习能力. 改进后卷积神经网络结构如图 4 所示.

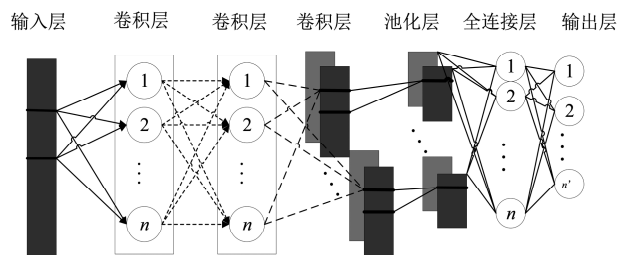


图 4 改进后卷积神经网络结构
Fig. 4 Improved convolutional neural network structure

2.2.2 构建增强非线性学习的 CNN 分类器 本文构建的卷积神经网络结构如图 5,其中共包含 7

层, 分别是输入层、卷积层 C_1 、卷积层 C_2 、卷积层 C_3 、最大池化层 M_4 、全连接层 FC_5 、输出层. 设输入尺寸为 $1 \times m_1 \times 1$, 其他各层输出尺寸 $1 \times m_2 \times 1$ 、 $1 \times m_3 \times 1$ 、 $1 \times m_4 \times 1$ 、 $1 \times m_5 \times 1$ 、 m_6 、 m_7 .

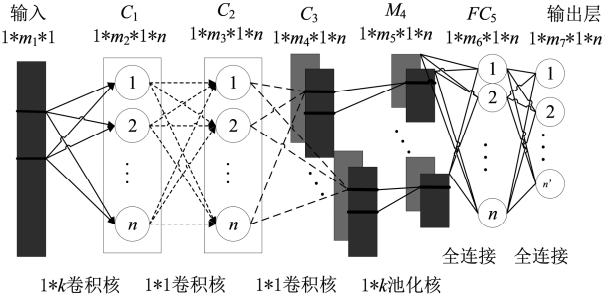


图 5 多层局部感知 CNN 分类器

Fig. 5 Multi-layer local perceptual CNN classifier

高光谱图像中的每一个像素样本是一个三维矩阵, 该矩阵的列数为 1, 行数 m_1 为高光谱图像频谱的个数 (以 Pavia University 数据为例, 该数据集频谱个数为 115, 但采用良好的 103 个频段, 所以行数为 103), 该数据同时是单通道所以通道数为 1. 这样构造方式将一个像素点作为一张图片识别. 因此输入层为 $(1 \times m_1 \times 1)$, m_1 为频谱的个数. 卷积层 C_1 是 n 个大小为 $(1 \times k)$ 的卷积核构成. 因此, 该卷积层输出 $1 \times m_2 \times 1 \times n$. 最后输出层 n' , 由识别物体种类决定; 详细网络结构及计算过程如表 1 和式 (7) 所示.

表 1 多层局部感知 CNN 分类器

Tab. 1 Multi-layer local perceptual CNN classifier

网络层	核大小	输出	步长
输入层		$1 \times m_1 \times 1$	
卷积层 C_1	$1 \times k$	$1 \times m_2 \times 1 \times n$	S_1
卷积层 C_2	1×1	$1 \times m_3 \times 1 \times n$	1
卷积层 C_3	1×1	$1 \times m_4 \times 1 \times n$	1
最大池化层 M_4	1×3	$1 \times m_5 \times 1 \times n$	3
全连接层 FC_5		$1 \times m_6 \times 1 \times n$	
输出层		$1 \times m_7 \times 1 \times n'$	

以下是各层尺寸计算公式.

$$\begin{cases} m_2 = (m_1 - k) / S_1 + 1 \\ m_3 = (m_2 - 1) / 1 + 1 = m_2 \\ m_4 = (m_3 - 1) / 1 + 1 = m_3 \\ m_5 = (m_4 - 3) / 3 + 1 \end{cases} \quad (7)$$

2.3 基于像素分类策略

本文提出的多层局部感知卷积神经网络结构有利于高效并且准确地提取非线性特征. 但对于高光谱图像分类, 面临另一个主要问题是如何建模

和调整输入层的数据格式.

在图 6~图 8 中所展示的光谱信息, 将 Pavia University 数据中每个像素点所包含的所有光谱通道融合成一条曲线反应在二维坐标值上. 横坐标表示 103 个谱段, 纵坐标代表每一个像素点该波段的幅值. 在图 6~图 8 中很难直观上判别不同地物的频谱特征差别, 但可以观察到每一种地物的光谱曲线明显不同于其他地物的光谱曲线.

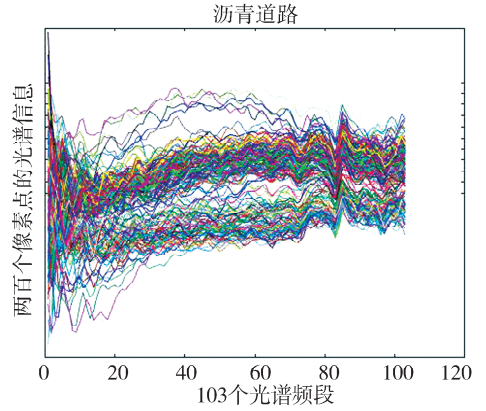


图 6 沥青道路像素的光谱信息

Fig. 6 Spectral information of asphalt road pixels

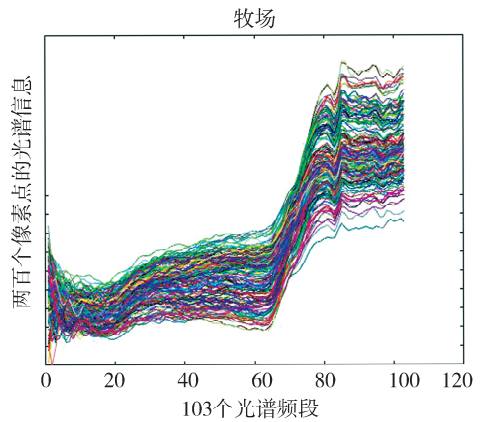


图 7 牧场像素的光谱信息

Fig. 7 Spectral information of pasture pixels

本文对高光谱图像数据的分类处理方式为硬分类^[11], 将高光谱图像中每个像素都赋予单个类别, 划分方式只要是通过像素的光谱特征和已知每个类别光谱特征的相似性来区分.

基于像素分类策略, 通过各种方法分析和提取不同目标物之间光谱域特征, 调整数据输入格式, 对不同方法进行测试.

以 Pavia University 数据集为例, 对于该图像 640×340 中的每个像素的 103 个谱段上的幅值组成 $1 \times 103 \times 1$ 的矩阵作为输入. 这种提取像素的

方式,不考虑像素间的空间关系,仅仅通过每个像素在不同波段的光谱信息特征提取用于分类。

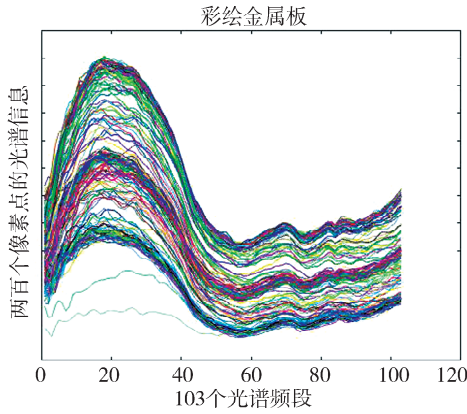


图 8 彩绘金属板像素的光谱信息
Fig. 8 Spectral information of painted metal plate pixels

2.4 基于改进 CNN 的分类及训练过程

2.4.1 正向传播 多层局部感知卷积网络结构包含 7 层,输入单元 n_1 构成输入层,输出单元 n_7 构成输出层,此外隐藏单元包含卷积层 C_1 、卷积层 C_2 、卷积层 C_3 、最大池化层 M_4 和全连接层 FC_5 。设第 i 层的输入为 x_i ,也是第 $(i-1)$ 层的输出,数学表示如下式。

$$x_{i+1} = f(W_i^T x_i + b_i) \quad (8)$$

式(8)中, W 是第 i 层的权重矩阵,对输入进行权重相乘; b 为第 i 层的偏置矩阵; f 为第 i 层的激活函数。卷积层 C_1 、卷积层 C_2 、卷积层 C_3 和全连接层 FC_5 适用。

2.4.2 动量梯度下降方式反向传播误差更新权重

本文使用梯度下降算法反向传播误差进行权重和偏置的更新。首先我们先定义一个代价函数(误差定义),如下。

$$C(w, b) = \frac{1}{2m} \sum_x \|y(x) - a\|^2 \quad (9)$$

式中, C 被称为二次代价函数,也称均分误差或者 MSE; m 为训练数据集的个数; a : 目标值(对应输入 x)。

综上,训练卷积神经网络的最终目标是求出最小化二次代价函数 $C(w, b)$ 的权重和偏置梯度下降算法对权重和偏置进行训练使得代价函数最小化。简化的更新公式如下所示。

$$\begin{cases} w_i \rightarrow w'_i = w_i - \eta \frac{\partial C}{\partial w_i} \\ b_i \rightarrow b'_i = b_i - \eta \frac{\partial C}{\partial b_i} \end{cases} \quad (10)$$

式中,右箭头符号为命题的“条件”运算; η 为学习率。

同过迭代上述更新权重和偏置的公式,对每一个输入和目标不断训练参数,从而使得最终代价函数最小化。一般学习率设置在 0.01 左右,学习率过大或者过小都会使得模型无法达到预想的精度,当加入 batch normalization 层后,可以增大学习率对精度影响变小,本文通过实验测试,将学习率调整为 0.035。

为进一步提高本文方法分类精度,引入加入动量的梯度下降算法,从而进一步提高训练出模型的精度。改进后的更新公式如下。

$$\begin{cases} w_i^{t+1} = w_i^t - \eta \frac{\partial C}{\partial w_i} + \gamma(w_i^t - w_i^{t-1}) \\ b_i^{t+1} = b_i^t - \eta \frac{\partial C}{\partial b_i} + \gamma(b_i^t - b_i^{t-1}) \end{cases} \quad (11)$$

式中, t 为更新的次数,公式指代当前更新的权重和偏置; γ 为动量,为考虑上次更新的权重、偏置与这次权重、偏置之间的因素对更新权重、偏置的影响。增加动量这个常数。本文采用默认值 0.9,通过增加动量优化后,进一步提高了训练模型的分类精度。

3 实验

本文算法在 MATLAB2018b 上运行,PC 机配置 Intel(R) Core(TM) 2 Duo CPU E7500, 2.93 GHz 处理器,ATI Radeon HD 3400 Series 显卡。

3.1 数据集

本文实验选用的常用的高光谱数据集为 Pavia University 数据和 Salinas 数据集^[12]。通过这种小训练样本的方式,检测该方法是否能够提取不同目标物像素点光谱域的有效特征。在输入前数据做 z-score 标准化,提高其他算法的计算速度和精度,同时加快神经网络梯度训练时候的收敛速度。

3.1.1 Pavia University 数据 Pavia University 数据是由德国的机载反射光学光谱成像仪(reflective optics spectrographic imaging system, ROSIS-03)对帕维亚大学所拍摄的一部分高光谱数据。其中,光谱成像仪在 0.43~0.86 μm 波长范围内的 115 个波段连续成像,空间分辨率为 1.3 m。其中剔除了 12 个受噪声影响的波段,因此采用剩余 103 个光谱波段所呈的图像。该图像的尺寸为 610×340,包含 207 400 个像素,但是其中包含 164 624 个背景像素,因而只有 42 776 个包含地物的像素,

这些像素中共有 9 类地物, 包括树(tree)、沥青道路(asphalt)、牧场(meadows)、彩绘金属板(painted metal sheets)等, 详细数据信息见表 2.

表 2 Pavia University 高光谱数据集

Tab. 2 Pavia University hyperspectral dataset

类别	类名	每类样本数	训练集 1
1	Asphalt	6 631	400
2	Meadows	18 649	400
3	Gravel	2 099	400
4	Trees	3 064	400
5	Painted metal sheets	1 345	400
6	Bare Soil	5 029	400
7	Bitumen	1 330	400
8	Self-Blocking Bricks	3 682	400
9	Shadows	947	400

3.1.2 Salinas 数据 Salinas 数据是由 AVIRIS 成像光谱仪所拍摄的美国加利福尼亚州的 Salinas 山谷图像. 该图像的空间分辨率为 3.7 m. 该图像原本有 224 个波段, 因为噪声剔除了第 108~112, 154~167, 还有不能被水反射的第 224 个波段后剩下的 204 个波段的图像. 该图像的尺寸为 512×217, 因此包含 111 104 个像素, 其中包含 56 975 个背景像素, 可用于分类的像素有 54 129 个, 这些像素总共分为 16 类, 包括休耕地 (Fallow)、芹菜 (Celery) 等, 详细数据见表 3.

3.2 实验结果与分析

本文为测试各个模型的时间与空间复杂度、过拟合程度和精度三个方面, 从而设计了以下实验.

(1) 实验中所用到的训练集的选择方式如下.

训练集 1 在每一类目标物当中随机挑选 400 个像素点作为训练集.

训练集 2 在每一类目标物当中随机挑选一半像素作为训练集.

(2) 实验中所用到的测试集的选择方式如下.

测试集 1 在训练集 1 随机挑选完后剩余像素中的每一类目标物当中随机挑选 200 个像素点作为测试集.

测试集 2 所有目标物像素点作为测试集.

本文提出的多层感知卷积神经网络、改进前原型和 Linear-SVM 三种方法形成对比, 本文方法与改进前原型的卷积核大小、参数设置、训练方法、训

练批次、批次大小等方面完全一致. Linear-SVM 和其他方法未进行参数优化. Linear-SVM 和 RBF-SVM 使用的是 MATLAB 中统计和机器学习工具箱实现.

卷积神经网络都采用动量梯度下降算法, 图 9 展示本文方法在训练集 1 上的训练过程, 其中包含正确率和损失率.

表 3 Salinas 高光谱数据集

Tab. 3 Salinas hyperspectral dataset

类别	类名	每类样本数	训练集 1
1	Brocoli_green_weeds_1	2 009	400
2	Brocoli_green_weeds_22	3 726	400
3	Fallow	1 976	400
4	Fallow_rough_plow	1 394	400
5	Fallow_smooth	2 678	400
6	Stubble	3 959	400
7	Celery	3 979	400
8	Grapes_untrained	11 271	400
9	Soil_vinyard_develop	6 203	400
10	Corn_senesced_green_weeds	3 278	400
11	Lettuce_romaine_4wk	1 068	400
12	Lettuce_romaine_5wk	1 927	400
13	Lettuce_romaine_6wk	916	400
14	Lettuce_romaine_7wk	1 070	400
15	Vinyard_untrained	7 268	400
16	Vinyard_vertical_trellis	1 807	400

表 4 表明, 在小样本分类中, 本文方法相对于改进前 CNN 在精度上有着明显提高, 同时略微优于传统高光谱分类的 Linear-SVM 方法. 测试结果说明, 在小规模训练样本情况下, 本文改进策略有效提高了卷积神经网络高光谱分类的精度. 同时说明本文改进策略提高了模型对非线性特征的提取能力.

表 4 不同数据集的三种算法对比结果测试结果(训练集 1 和测试集 1)

Tab. 4 Three algorithm comparison results test results of different data sets

数据集 1	改进前 CNN 精度/%	Linear-SVM 精度/%	本文方法 精度/%
Pavia University	86.32	90.00	90.23
Salinas	87.21	91.00	91.52

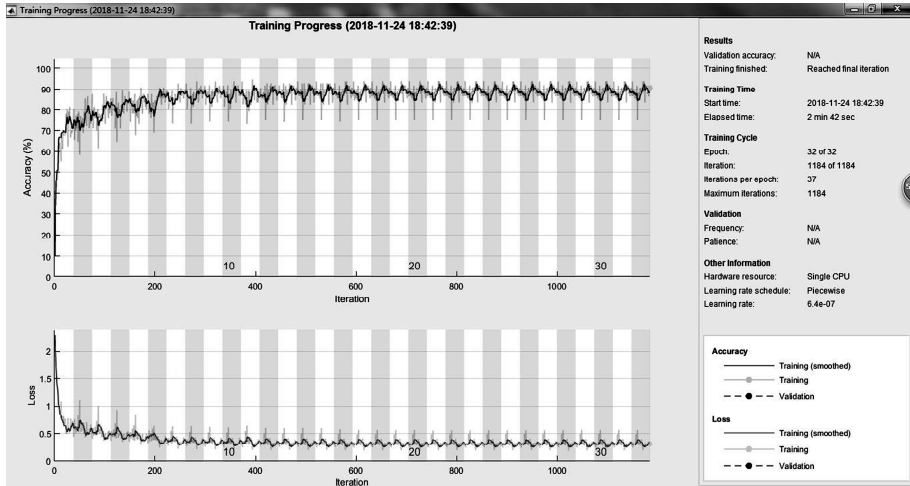


图 9 本文方法训练过程
Fig. 9 Method training process

表 5 对于训练集 1 和测试集 1 各个方法的性能对比 (Pavia University 数据集)

Tab. 5 Performance comparison of each method of training set 1 and test set 1

方法	训练时间 /s	测试时间 /s	精度 /%
本文方法	132.00	0.31	90.23
改进前 CNN	86.00	0.28	86.34
LeNet-5	61.00	0.29	87.94
KNN($k=7$ 最优)	0	31.42	83.86
Linear-SVM	11.15	0.48	90.00
RBF-SVM	4.06	3.25	82.39

在表 5 中,本文方法相对于改进前 CNN 训练时间增加,有更多的网络参数需要计算. 在仅增加 0.03 s 情况下提高了模型分类精度. 同时在分类时间几乎相等情况下,在精度上击败了经典的卷积神经网络结构 LeNet-5. 相对于传统的 SVM 在小样本分类中不仅分类速度上和分类精度上都有明显改进.

表 6 对于训练集 1 和测试集 2 各个方法的性能对比 (Pavia University 数据集)

Tab. 6 Performance comparison of each method of training set 1 and test set 2

方法	训练时间/s	测试时间/s	精度/%
本文方法	180.06	8.46	86.52
改进前 CNN	87.56	6.52	80.66
LeNet-5	65.00	4.02	79.49
Linear-SVM	7.92	5.87	87.74

将测试集 1 替换成测试集 2 得出的表 6 实验结果. 在表 6 中,所有模型分类精度都有下降,

出现不同程度的过拟合现象. 本文方法,随着测试集增大,测试时间相对于其他模型增长稍大,实际使用过程中依然可以接受. 传统 SVM 依然能保持良好高效的性能.

表 7 量化过拟合现象,过拟合程度数值由第二列减第三列的差值然后与第三列的比值得出. 在小样本训练集情况下,传统 SVM 能有效提取数据特征,本文方法相对于其他传统卷积神经网络大大减小了过拟合程度,多层局部感知结构能有效提高对非线性特征的提取能力.

表 7 对于训练集 1 各个方法的过拟合程度

Tab. 7 Over-fitting degree for each method of training set 1

方法	测试集 1 精度 /%	测试集 2 精度 /%	过拟合程度 /%
本文方法	90.23	86.52	4.25
改进前 CNN	86.34	80.66	7.04
LeNet-5	87.94	79.49	10.63
Linear-SVM	90.00	87.74	2.57

表 8 对于训练集 2 和测试集 2 各个方法的性能对比 (Pavia University 数据集)

Tab. 8 Performance comparison of each method of training set 2 and test set 2

方法	训练时间/s	测试时间/s	精度/%
本文方法	843.85	8.29	97.23
改进前 CNN	755.45	7.12	94.56
LeNet-5	743.00	4.26	96.64
Linear-SVM	452.94	5.87	92.40

训练 2 测试得到表 8 实验结果,通过增大训练集卷积神经网络可以达到更高的分类精度,多层局部感知神经网络相对于其他神经网络在高光谱图

像分类上有着更高的精度,有着更好的非线性特征提取能力. SVM 依然有着高效性能和分类精度,但无法提取更多的特征.

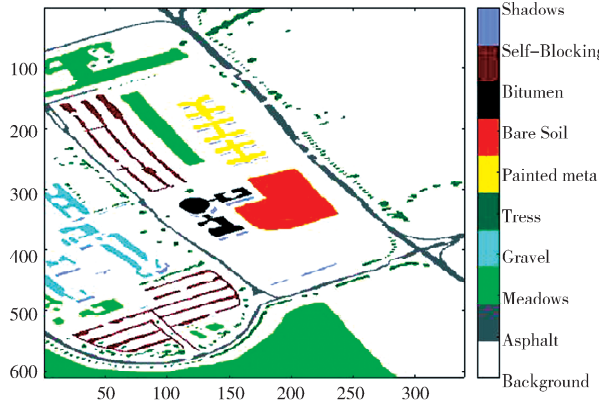


图 10 Pavia University 数据集目标可视化
Fig. 10 Pavia University dataset target visualization

综合上述实验结果说明,SVM 在小规模训练样本上的计算成本和训练时间都是有着较高优势,但是随着训练集增大,精度无法做到最优,并且需要参数优化,才能达到较好的精度. 改进前 CNN 方法和 LeNet-5 在小规模训练样本上,不能达到良好的分类精度. 在较大规模训练样本情况下,神经网络有着更高的分类精度. 随着训练集增大,可以学习到更多的有效特征. 本文提出多层局部感知神经网络能更加有效提取样本光谱特征,同时也随着训练集增大,精度能有效提高. 相比传统的 LeNet-5 深度学习网络和 SVM,本文提出的方法无论在小样本分类和大样本分类都能有效提取特征,相对于改进前的 CNN 分类器和 LeNet-5 在精度上有着明显提升^[13].

图 11~图 13 展示各种方法在训练集 2 训练下的测试集 2 的测试可视化结果.

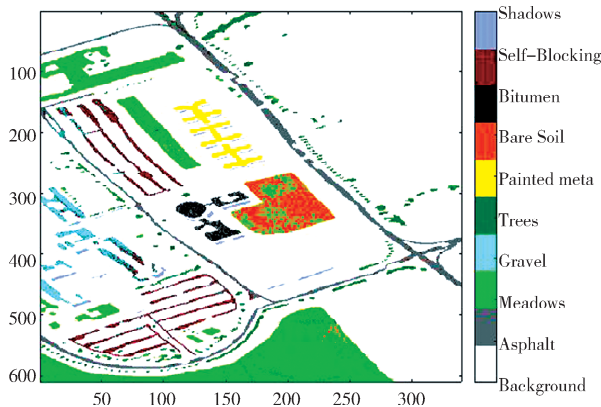


图 11 Linear-SVM 预测结果 92.40% 精确度
Fig. 11 Linear-SVM prediction results 92.40% accuracy

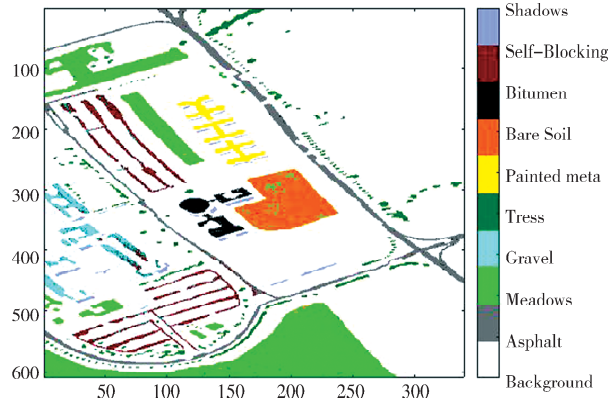


图 12 LeNet-5 预测结果 96.64% 精确度
Fig. 12 LeNet-5 forecast results 96.64% accuracy

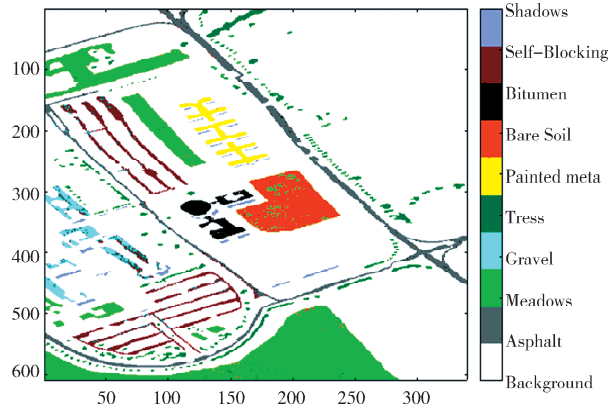


图 13 本文方法预测结果 97.23% 精确度
Fig. 13 This method predicts 97.23% accuracy

4 结 论

本文在 Pavia University 数据和 Salinas 数据集上进行多种方法测试,SVM 在小样本训练中有着有效提取特征的能力,本文通过对传统卷积神经网络进行改进确实有效提高了在小样本和大样本训练时的非线性特征学习能力,都达到最佳的精度. 在较大规模训练样本时,卷积神经网络有着良好的进一步学习能力. LeNet-5 网络结构在大训练样本时表现出高效的分类精度,表明深度学习网络在高光谱图像分类上有着良好的潜力^[14].

下一步,将 U-net 网络和 NIN 网络结构应用于高光谱分类中,进一步提高分类精度同时避免过拟合,在考虑新的分类策略,通过软分类方式考虑像素空间域和光谱域相关性特征,更进一步提高分类精度^[15].

参考文献:

- [1] 赵漫丹, 任治全, 吴高昌, 等. 利用卷积神经网络的高光谱图像分类[J]. 测绘科学技术学报, 2017, 34: 501.
- [2] 杜培军, 夏俊士, 薛朝辉, 等. 高光谱遥感影像分类研究进展[J]. 遥感学报, 2016, 20: 236.
- [3] Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: A review[J]. ISPRS J Photogramm, 2011, 66: 247.
- [4] 刘健, 袁谦, 吴广, 等. 卷积神经网络综述[J]. 计算机时代, 2018, 2018: 19.
- [5] Atkinson P M, Tatnall A R L. Introduction neural networks in remote sensing [J]. Int J Remote Sens, 1997, 18: 699.
- [6] Bruzzone L, Prieto D F. A technique for the selection of kernel-function parameters in RBF neural networks for classification of remote-sensing images [J]. IEEE T Geosci Remote, 1999, 37: 1179.
- [7] Lin M, Chen Q, Yan S. Network in network [J]. arXiv preprint arXiv:1312.4400, 2013.
- [8] Shi B, Bai X, Yao C. Script identification in the wild via discriminative convolutional neural network [J]. Pattern Recogn, 2016, 52: 448.
- [9] He K, Zhang X, Ren S, *et al.* Delving deep into rectifiers: surpassing human-level performance on imagenet classification [C]//Proceedings of the IEEE international conference on computer vision. [S. l.]: IEEE, 2015.
- [10] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the International Conference on International Conference on Machine Learning. . [S. l.]: [s. n.], 2015.
- [11] 伍广明, 陈奇, Shibasaki R, 等. 基于 U 型卷积神经网络的航空影像建筑物检测[J]. 测绘学报, 2018, 47: 864.
- [12] Grupo deInteligencia Computacional. Hyperspectral remote sensing scenes [EB/OL]. [2017-10-10]. (2019-01-01). http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.
- [13] 张睿萍, 马宗梅. 基于 Hadoop 平台的大数据图像分类机制[J]. 吉林大学学报: 理学版, 2018, 56: 1206.
- [14] 但斌, 罗晓, 刘墨林. “互联网+”产品服务供应链的价值创造及其实现路径[J]. 重庆邮电大学学报: 社会科学版, 2019, 31: 83.
- [15] 凌语, 孙自强. 基于卷积神经网络的乳腺病理图像识别算法[J]. 江苏大学学报: 自然科学版, 2019, 40: 573.

引用本文格式:

中文: 池涛, 王洋, 陈明. 多层局部感知卷积神经网络的高光谱图像分类[J]. 四川大学学报: 自然科学版, 2020, 57: 103.

英文: Chi T, Wang Y, Chen M. Hyperspectral image classification of multi-layer local perceptual convolutional neural networks [J]. J Sichuan Univ: Nat Sci Ed, 2020, 57: 103.