文章编号: 2095-2163(2019)05-0050-04 中图分类号:TP391.41 文献标志码:A

一种改进的稀疏编码模型图像分类算法

刁 琦
(浙江东方职业技术学院, 浙江 温州 325000)

摘 要: 针对传统稀疏编码方法对图像分类，本文采用 SIFT 算法提取的图像特征，分类器主要通过构造多尺度小波核极限学习机进行图像分类。采用该方法对 wine 数据集与微软 Corel1K 图像库进行测试，实验表明，较单尺度小波核，多尺度小波核具有更好的分类效果。同时，多尺度小波核极限学习机的分类性能优于多尺度支持向量机、支持向量机与极限学习机的分类性能。

关键词: 稀疏编码; 图像分类; 极限学习机; 支持向量机

An improved image classification approach of sparse coding

DIAO Qi
(Zhejiang Dongfang Polytechnic, Wenzhou Zhejiang 325000, China)

【Abstract】 According to the traditional sparse coding method for image classification, the image features extracted by SIFT algorithm are used in this paper. The classifier mainly constructs multi-scale wavelet kernel limit learning machine for image classification. This method is used to test the wine dataset and the Microsoft Corel1K image library. Experiments show that the multi-scale wavelet kernel has better classification effect than the single-scale wavelet kernel. At the same time, the classification performance of multi-scale wavelet kernel limit learning machine is better than that of multi-scale support vector machine, support vector machine and extreme learning machine.

【Key words】 sparse coding; image classification; extreme learning machine; support vector machine

0 引 言

稀疏编码(Sparse Coding,SC)算法作为无监督的学习方法之一,其使用基向量的线性组合方式(完备基)来描述图像的输入信息,且能够提取较好的数据集特征,广泛应用于图像分类、图像检索、数据预测等问题。张立和等人[1]提出核拉普拉斯稀疏编码,在分类效果上优于拉普拉斯稀疏编码;谢成军等人[2]采用稀疏编码金字塔模型对农田害虫图像进行识别,较传统方法识别精度提高 14.1%;张勇等人[3]提出一种非负弹性网稀疏编码算法,较传统的稀疏编码算法的有效性更高,分类的效果更好;徐佳庆等人[4]提出基于二阶矩空谱联合稀疏编码的遥感图像分类方法,与支持向量机(Support Vector Machine,SVM)方法相比有较高的准确性。在分类方法上,SVM 是一种常用的分类器[5]。满瑞君等人[6]将多尺度小波支持向量机(Multi-scale wavelet SVM)应用于交通流数据的预测,指出该方法的预测准确性高于 BP 神经网络和高斯 SVM。随着极限学习机(Extreme Learning Machine,ELM)的提出,也被研究者广泛应用于图像分类问题[7]。王杰等人[8]将小波核极限学习机(Wavelet ELM)应用于 UCI 图像集分离,并指出 WELM 的分类性能优于 ELM。为此,本文将多尺度小波核函数(Multi-scale wavelet kernel)作为 ELM 的核函数,以在测试集中获得更好的分类效果。对此拟展开研究论述如下。

1 稀疏编码描述

SC 概念来源于仿生学,其作为一种特征的表示方法,目的在于求解基向量的系数,采用基向量的线性组合来表示输入特征。其数学公式即如式(1)所示:

$$X = \sum_{i=1}^{k} a_i \boldsymbol{\Phi}_i , \tag{1}$$

其中, $\boldsymbol{\Phi}_i$ 代表基向量,a_i 为基向量的系数。

类似于主成分分析(Principal Component Analysis,PCA)描述信息特征,采用 SC 也是为了使用较少的数据去构造一个新的数据。根据式(1)中的变量数量来看,需要求解基向量以及相应系数两个变量。由于基向量个数大于输入向量的维数,这

基金项目: 浙江省社科规划课题成果(19NDJC396YBM); 校级课题(DF2017YB08)。
作者简介: 刁 琦(1989-),男,硕士,助教,主要研究方向:计算机应用技术、智能计算及应用。
收稿日期: 2019-07-30

使得求解基向量的系数不再唯一取决于输入向量。为了能够求解 2 个变量，需要对基向量系数进行约束。那么，稀疏编码的目标函数表达式见式（2）：

$$\min_{a_i^{(j)},\Phi_i}\sum_{j=1}^{m}\left\|x^{(j)}-\sum_{i=1}^{k}a_i^{(j)}\boldsymbol{\Phi}_i\right\|^2+\lambda\sum_{i=1}^{k}S(a_i^{(j)}),\tag{2}$$

分析可知，式（2）中包含 2 部分，第一项可理解为构造项，第二项为系数的惩罚系数，λ为常量，用以平衡上述两项间的代价权重。由于系数 a 的无限减小与基值无限大使得第一项基本保持不变，而第二项会很小。为了获得最优解，需对基向量进行约束，最终的目标函数描述如式（3）所示：

$$\min_{a_i^{(j)},\Phi_i}\sum_{j=1}^{m}\left\|x^{(j)}-\sum_{i=1}^{k}a_i^{(j)}\boldsymbol{\Phi}_i\right\|^2+\lambda\sum_{i=1}^{k}S(a_i^{(j)})$$
$$\text{subject to }\ \|\boldsymbol{\Phi}_i\|^2\leqslant C,\ \forall i=1,\cdots,k.\tag{3}$$

2　本文算法

2.1　SIFT 特征

本文在特征提取上主要采用尺度不变特征变换（Scale Invariant Feature Transform，SIFT）算法，其本质是一种局部特征的提取方法，主要优势在于对图像的旋转、尺度缩放、亮度变化等保持不变形[9]。采用 SIFT 算法提取图像特征主要包括 4 个步骤，即：

（1）尺度空间的关键点检测。

（2）定位关键点。

（3）确定关键点位置。

（4）生成特征描述符。

限于篇幅，此处将不再详述 SIFT 算法的特征提取过程。

2.2　多尺度小波核 ELM 分类器

极限学习机的表达式如式（4）所示：

$$\begin{aligned}&\sum_{i=1}^{L}\beta_i G(a_i\cdot x_1+b_i)=t_1;\\&\qquad\vdots\\&\sum_{i=1}^{L}\beta_i G(a_i\cdot x_N+b_i)=t_N;\\&i=1,2,\cdots L.\end{aligned}\tag{4}$$

其中，$G(\cdot)$ 为激活函数；a_i 为输入权值；β_i 为输出权值；b_i 为第 i 个隐含层单元的偏置，其矩阵形式如式（5）所示：

$$\boldsymbol{H}(a_1,\cdots,a_L,b_1,\cdots,b_L,x_1,\cdots,x_N)=\begin{pmatrix}G(a_1\cdot x_1+b_1)\cdots G(a_L\cdot x_1+b_L)\\ \vdots \qquad \ddots \qquad \vdots\\ G(a_1\cdot x_N+b_1)\cdots G(a_L\cdot x_L+b_L)\end{pmatrix}_{N\times L},\tag{5}$$

其中，$\boldsymbol{H}$ 为隐含输出层。

较 SVM 学习机制有所不同，ELM 不仅需要考虑结构风险最小化，同时还需要考虑经验误差最小化。可构造出式（6）：

$$\min:\|\boldsymbol{\beta}\|^2\sum_{i=1}^{N}\|\boldsymbol{\beta}\cdot\boldsymbol{h}(x_i)+t_i\|,\tag{6}$$

为便于采用拉格朗日定理求解最值问题，式（6）可转换为式（7）的形式：

$$\begin{aligned}&\min:L=\frac{1}{2}\|\boldsymbol{\beta}\|^2+\frac{1}{2}C\sum_{i=1}^{M}\|\boldsymbol{\zeta}\|^2\\&\text{s.t. }\boldsymbol{h}(x_i)\boldsymbol{\beta}=\boldsymbol{t}_i^{\mathrm{T}}-\boldsymbol{\zeta}_i^{\mathrm{T}},i=1,\cdots,N.\end{aligned}\tag{7}$$

其中，ξ是训练样本 x_i 对应的网络输出值和实际值间的误差。式（7）又等同于式（8）：

$$\begin{aligned}&\min:L_{DELM}=\frac{1}{2}\|\boldsymbol{\beta}\|^2+\frac{1}{2}C\sum_{i=1}^{M}\|\boldsymbol{\zeta}_i\|^2-\\&\sum_{i=1}^{N}\sum_{j=1}^{m}\alpha_{i,j}(\boldsymbol{h}(x_i)\boldsymbol{\beta}_j-t_{i,j}+\boldsymbol{\zeta}_{i,j}),\end{aligned}\tag{8}$$

其中，a_i 和 u_i 为拉格朗日乘子，均大于零。β_j 为第 j 个输出节点的权值。

用式（8）分别对未知量求偏导，可得式（9）：

$$\frac{\partial L_{DELM}}{\partial\beta_j}=0\rightarrow\beta_j=\sum_{i=1}^{N}\alpha_i h(x_i)^{\mathrm{T}}\rightarrow\beta=\boldsymbol{H}^{\mathrm{T}}\alpha;\tag{a}$$

$$\frac{\partial L_{DELM}}{\partial\zeta_i}=0\rightarrow\alpha_i=C\zeta,i=1,2,\cdots,N;\tag{b}$$

$$\frac{\partial L_{DELM}}{\partial\alpha_i}=0\rightarrow h(x_i)\beta_i-t_i^{\mathrm{T}}+\zeta_i^{\mathrm{T}}=0,i=1,2,\cdots,N.\tag{c}$$

（9）

通过式（9a）、（9b）带入式（9c），进一步可得式（10）：

$$\begin{aligned}&\boldsymbol{h}(x_1)\boldsymbol{H}^{\mathrm{T}}C\boldsymbol{\zeta}_1-\boldsymbol{t}_1^{\mathrm{T}}+\boldsymbol{\zeta}_1^{\mathrm{T}}=0;\\&\qquad\vdots\\&\boldsymbol{h}(x_N)\boldsymbol{H}^{\mathrm{T}}C\boldsymbol{\zeta}_N-\boldsymbol{t}_N^{\mathrm{T}}+\boldsymbol{\zeta}_N^{\mathrm{T}}=0.\end{aligned}\tag{10}$$

为简化表达式，合并为式（11）：

$$\left(\frac{\boldsymbol{I}}{C}+\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\right)\alpha=\boldsymbol{T},\tag{11}$$

其中，$\boldsymbol{T}=\begin{bmatrix}t_1^{\mathrm{T}}\\ \vdots\\ t_N^{\mathrm{T}}\end{bmatrix}=\begin{bmatrix}t_{11}&\cdots&t_{1m}\\ \vdots&\vdots&\vdots\\ t_{N1}&\cdots&t_{Nm}\end{bmatrix}$，$\boldsymbol{H}=\begin{bmatrix}h(x_1)\\ \vdots\\ h(x_N)\end{bmatrix}$，

那么，输出权值表达式如式（12）所示：

$$\beta=\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\frac{\boldsymbol{I}}{C})^{-1}\boldsymbol{T},\tag{12}$$

极限学习机的逼近函数如式（13）所示：

$$f(\boldsymbol{x}_i)=h(\boldsymbol{x}_i)\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}+\frac{\boldsymbol{I}}{C})^{-1}\boldsymbol{T}.\tag{13}$$

其中，$\boldsymbol{H}=\begin{bmatrix} h(\boldsymbol{x}_1) \\ \vdots \\ h(\boldsymbol{x}_N) \end{bmatrix}_{N\times L}$ 。

隐层输出可视为样本的非线性映射，可以是线性核的形式，也可以采用 RBF 核的形式。于是可得式(14)：

$$\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}=\begin{bmatrix} h(\boldsymbol{x}_1) \\ \vdots \\ h(\boldsymbol{x}_N) \end{bmatrix}_{N\times L}^{\mathrm{T}} \cdot \begin{bmatrix} h(\boldsymbol{x}_1) \\ \vdots \\ h(\boldsymbol{x}_N) \end{bmatrix}_{N\times L} = \begin{bmatrix} h(\boldsymbol{x}_1)\cdot h(\boldsymbol{x}_1)\cdots h(\boldsymbol{x}_1)\cdot h(\boldsymbol{x}_N) \\ \vdots \quad \ddots \quad \vdots \\ h(\boldsymbol{x}_N)\cdot h(\boldsymbol{x}_1)\cdots h(\boldsymbol{x}_N)\cdot h(\boldsymbol{x}_N) \end{bmatrix}_{N\times N}, \tag{14}$$

根据核函数理论，通过构造一个隐式的映射来代替 $\boldsymbol{HH}^{\mathrm{T}}$，如式(15)所示：

$$\boldsymbol{HH}^{\mathrm{T}}(i,j)=K(\boldsymbol{x}_i,\boldsymbol{x}_j);$$

$$\boldsymbol{HH}^{\mathrm{T}}=\boldsymbol{\Omega}_{ELM}=\begin{bmatrix} K(\boldsymbol{x}_1,\boldsymbol{x}_1)\cdots K(\boldsymbol{x}_1,\boldsymbol{x}_j) \\ \vdots \quad \ddots \quad \vdots \\ K(\boldsymbol{x}_N,\boldsymbol{x}_1)\cdots K(\boldsymbol{x}_N,\boldsymbol{x}_N) \end{bmatrix};$$

$$h(\boldsymbol{x})\boldsymbol{H}^{\mathrm{T}}=\begin{bmatrix} K(\boldsymbol{x},\boldsymbol{x}_1) \\ \vdots \\ K(\boldsymbol{x},\boldsymbol{x}_N) \end{bmatrix}, \tag{15}$$

又因式(16)：

$$f(\boldsymbol{x})=h(\boldsymbol{x})\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{HH}^{\mathrm{T}}+\frac{\boldsymbol{I}}{C})^{-1}\boldsymbol{T}, \tag{16}$$

极限学习机的求解公式经变换后可写成式(17)：

$$f(\boldsymbol{x})=\begin{bmatrix} K(\boldsymbol{x},\boldsymbol{x}_1) \\ \vdots \\ K(\boldsymbol{x},\boldsymbol{x}_N) \end{bmatrix}^{\mathrm{T}}(\frac{\boldsymbol{I}}{C}+\boldsymbol{\Omega}_{ELM})^{-1}\boldsymbol{T}, \tag{17}$$

为了方便，本文直接将 SVM 的核函数应用在 ELM 中。已知一个母小波函数 $h(x)$，其伸缩因子和平移因子分别是 a 和 b，那么小波基函数可以表示为式(18)，即：

$$h_{a,b}(x)=\sqrt{|a|}h(\frac{x-b}{a}), \tag{18}$$

根据张量积理论，一个多维的小波函数可写成多个一维小波函数的张量积，由此得到：

$$h(\boldsymbol{x})=\prod_{i=1}^{n}h(x_i), \tag{19}$$

根据式(19)可以构造出平移不变的核函数，如式(20)所示：

$$K(x,x')=K(x+x')=\prod_{i=1}^{n}h\left(\frac{x_i-x'_i}{a}\right), \tag{20}$$

本文选用 Morlet 小波函数 $h(x)=\cos(1.75x)\exp(-x^2/2)$，那么小波核函数的表达式如式(21)所示：

$$waveletkernel(x,x')=\prod_{i=1}^{n}\left[\cos(1.75\times\left(\frac{x_i-x'_i}{a}\right)\exp(-\frac{(x_i+x'_i)^2}{2a^2})\right], \tag{21}$$

由于单个核函数不能在大规模图像分类中获得较好的分类效果，本文采用多个核函数叠加形式来构造混合的小波核函数，以提高分类效果。如式(22)所示：

$$K=a_1*K_1+\cdots+a_n*K_n, \tag{22}$$

同时，本文在小波核函数的基础上提出多尺度小波核，如式(23)所示：

$$K=\prod_{i}^{n}\left[\cos(1.75\times\left(\frac{x_i-x'_i}{a_1}\right)\exp\left(-\frac{(x_i-x'_i)^2}{2a_1^2}\right))\right]+\prod_{i}^{n}\left[\cos(1.75\times\left(\frac{x_i-x'_i}{a_2}\right)\exp\left(-\frac{(x_i-x'_i)^2}{2{a_2}^2}\right))\right]+\prod_{i}^{n}\left[\cos(1.75\times\left(\frac{x_i-x'_i}{a_N}\right)\exp\left(-\frac{(x_i-x'_i)^2}{2{a_N}^2}\right))\right] \tag{23}$$

最终可简写成式(24)：

$$K=\sum_{l}^{L}\prod_{i}^{n}\left[\cos(1.75\times\left(\frac{x_i-x'_i}{a_l}\right)\exp\left(-\frac{(x_i-x'_i)^2}{2a_l^2}\right))\right]. \tag{24}$$

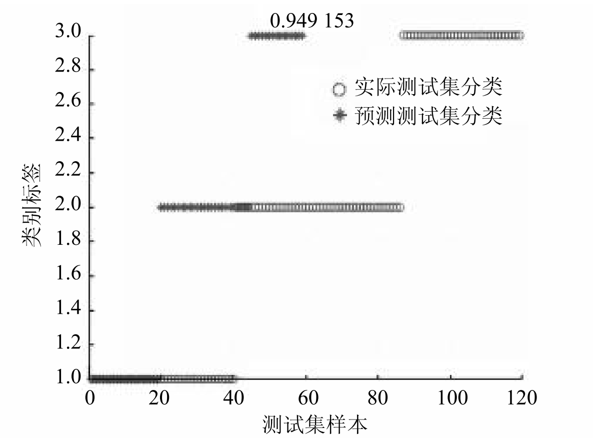
3 实验结果与分析

首先，本文对 Wine 数据集进行测试。Wine 数据集包含了 3 类葡萄酒数据，共计 178 个样本，每类葡萄酒有 13 个属性特征。第一类训练样本为 1 ~ 30，第二类训练样本为 60 ~ 95，第三类训练样本为 131 ~ 153，其余为测试样本。分别对比单尺度、二尺度以及三尺度的分类性能。采用不同尺度小波核的分类结果见表 1，采用文献[6]获得的分类精度为 93.09%，这也表明采用多尺度小波核 ELM 分类性能优于小波核 ELM 和小波核 SVM。分类的具体情况如图 1 所示。

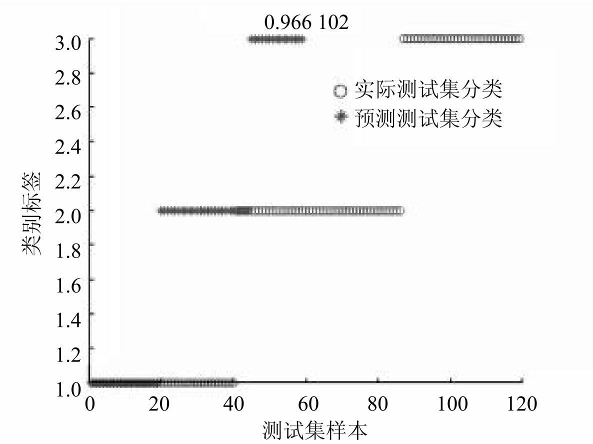
表 1 不同尺度小波核极限学习机的分类性能

Tab. 1 Classification performance of Wavelet kernel limit learning machines with different scales %

尺度	分类精度
单尺度小波核	94.91
二尺度小波核	96.61
三尺度小波核	96.61



(a) 单尺度
(a) Single scaling



(b) 多尺度
(b) Multi scaling

图 1　不同尺度小波核极限学习机的分类效果

Fig. 1　Classification effect of different scale wavelet kernel limit learning machines

其次，采用微软 Corelk 图像库进行测试。其中，Corel 图像库包含 10 类图像，每类图像为 100 张，共计 1 000 张图像。方法实现中主要包括 2 部分，分别是：对图像进行 sift 特征提取，图像分类。每类图像的分类精度见表 2。

表 2　Corelk 各类图像的分类精度

Tab. 2　Classification accuracy of Corelk images

Corelk	样本数量	分类精度
Africa	30	0.771 4
beaches	30	0.828 6
buildings	30	0.757 1
buses	30	1
dinosaurs	30	1
elephants	30	0.928 6
flowers	30	0.957 1
food	30	0.985 7
horses	30	0.975 1
mountains	30	0.890 0

采用不同分类器方法的对比情况如图 2 所示。从图 2 中可看出，本文的分类精度要高于文献[6]方法、SVM 和 ELM。对于 buses 与 dinosaurs 两个类别，4 种分类器的分类精度相同，主要由于这 2 类图片背景较为简单，对于复杂图像的分类，多尺度小波核 ELM 具有较好的分类效果。

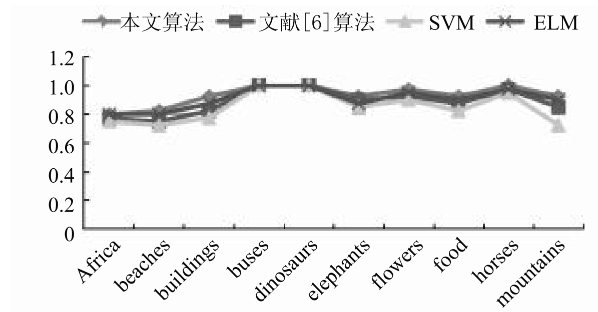


图 2　不同分类器的分类效果

Fig. 2　Classification effect of different classifiers

4　结束语

本次研究及实验表明，多尺度小波核极限学习机具有较高的分类性能。由于核函数在分类问题方面具有较强的映射能力和非线性分类能力，但对于解决实际复杂背景及大规模图像分类问题效果不佳，在实际应用中可采用多个核函数由组合形式来获得更好的结果。

参考文献

[1] 张立和，潘磊，刘涛，等. 基于核拉普拉斯稀疏编码的图像分类[J]. 大连理工大学学报，2015，55(2)：192-197.

[2] 谢成军，李瑞，董伟，等. 基于稀疏编码金字塔模型的农田害虫图像识别[J]. 农业工程学报，2016，32(17)：144-151.

[3] 张勇，张阳阳，程洪，等. 基于非负弹性网稀疏编码算法的图像分类方法[J]. 计算机工程，2017，43(7)：239-243，249.

[4] 徐佳庆，万文，吕启. 基于二阶矩稀疏编码的高光谱遥感图像分类[J]. 计算机科学，2018，45(9)：288-293.

[5] 吴园园，申立勇. 基于类重叠度欠采样的不平衡模糊多类支持向量机[J]. 中国科学院大学学报，2018，35(4)：536-543.

[6] 满瑞君，梁雪春. 基于多尺度小波支持向量机的交通流预测[J]. 计算机仿真，2013，30(11)：156-159.

[7] HUANG Guangbin，ZHU Qinyu，SIEW C K. Extreme learning machine：Theory and applications[J]. Neurocomputing，2006，70(1/3)：489-501.

[8] 王杰，郭晨龙. 小波核极限学习机分类器[J]. 微电子学与计算机，2013，30(10)：73-76，80.

[9] 李兵，刘磊，魏志强. 一种具有强实时性、强鲁棒性的图像匹配算法[J]. 软件学报，2014，25(7)：1583-1592.