

PDEs 对 RSVM 中模型参数的影响

江珊珊 范丽亚

(聊城大学 数学科学学院, 山东 聊城 252059)

摘 要 支持向量机(Support Vector Machines, SVMs)的分类结果严重受限于模型参数的选择. 很多学者都在研究调参问题,但目前还没有一个行之有效的办法,常用的办法是网格搜索,一个近似选择参数的方法. 以噪声(光暗,有遮挡)图像分类为背景,以正则化支持向量机(Regularized SVM, RSVM)为分类器,研究了偏微分方程组(Partial Differential Equations, PDEs)对 RSVM 模型中参数选择的影响. 实验结果表明通过 PDEs 的进化可以弱化参数的影响,甚至不需要考虑调参.

关键词 支持向量机;偏微分方程;参数选择;进化次数;图像分类

中图分类号 TP391

文献标识码 A

0 引言

近期,一档名为《机智过人》的栏目引起了人们的强烈兴趣,其中人工智能所呈现的多种多样的功能更是让人大开眼界. 实际上,随着计算机技术的不断发展,人工智能在我们生活中得到了广泛的应用,尤其是图像识别在地理勘测、文物修复、病理研究等领域发挥着愈来愈重要的作用. 图像识别的过程通常包含图像处理和图像分类两部分. 图像处理是指利用变换、编码压缩、增强复原、分割等方法提取或清洗图像特征,以期提高图像的分类准确性,如局部二值模式(Local Binary Pattern, LBP)^[1]、方向梯度直方图(Histogram of Oriented Gradient, HOG)^[2]、Haar 特征(Haar-like Features)^[3]、卷积神经网络(Convolutional Neural Networks, CNNs)^[4]、线性判别分析(Linear Discriminant Analysis, LDA)^[5]、主成分分析(Principal Component Analysis, PCA)^[6]等,但这些方法对 SVM 型分类器中的参数选择不起作用.

SVMs^[7]作为机器学习中的一种常用的分类工具,于 1995 年由 V. Vapnik 首次提出,之后得到了长足的发展^[8-14],被广泛应用于健康与疾病监测^[15,16]、情绪检测^[17,18]等诸多领域. 然而, SVM 型分类器的分类准确度严重依赖于模型参数的选择. 很多学者都在研究调参问题,但还没有一个行之有效的方法,常用的办法是网格搜索,一个近似选择参数的方法. 本文以噪声(光暗,有遮挡)图像分类为背景,以正则化支持向量机(Regularized SVM, RSVM)为分类器,研究偏微分方程组(Partial Differential Equations, PDEs)对 RSVM 模型中参数选择的影响.

在图像识别或图像分类任务中,所面临的图像大都是噪声图像,如光暗、遮挡、异常图像等. 为此需要面对三个方面的问题:一是图像的清洗,即图像的特征提取,特征越清晰分类效果越好;二是分类器的选择,本文选择正则化支持向量机(Regularized SVM, RSVM)作为分类器;三是分类器中模型参数的选择. 图像特征提取的方法有很多,如上面提到的 LDA, PCA, CNNs 等,但大多数方法只是起到了清洗图像的作用,对模型参数的选择没有影响. 本文希望找到一个既能清洗图像,又能弱化模型参数影响的图像特征提取方法.

Liu 等人于 2010 年针对图像恢复提出了一种基于最优控制的 PDEs 学习方法^[19],2013 年针对计算机视觉提出了一种基于最优控制的智能 PDEs 学习方法^[20],2016 年针对视觉分析提出了一种新的 PDEs 学习方法-learning to diffuse (LTD)^[21]. 2017 年, Fang 等人^[22]提出了一个基于偏微分方程组(Partial Differential Equations, PDEs)的特征学习方法,并应用于面部识别中. 2018 年,江等人^[23]提出了基于 PDEs 的图像特

收稿日期:2019-11-20

基金项目:国家自然科学基金项目(11801248);山东省自然科学基金项目(ZR2016AM24)资助

通讯作者:范丽亚,女,汉族,博士,教授,研究方向:机器学习理论与应用, E-mail: fanliya63@126.com.

征提取方法,大量的实验结果表明 PDEs 本身具有降噪(遮挡)功能、对光照具有鲁棒性且可提取图像的判别信息.本文在文献[22,23]的基础上,研究了 PDEs 对 RSVM 模型中参数选择的影响.

基于 Extended YaleB 数据集^[24]和 PIE 数据集^[25]上的实验结果表明,通过 PDEs 的进化可以弱化模型参数的影响,甚至不需要考虑模型参数的选择.

1 预备知识

本节简要回顾线性 RSVM 和基于 PDEs 的图像特征学习模型,具体内容见文献[7,19-20].

1.1 线性 RSVM

给定二分类数据集 $T = \{(\mathbf{x}_i, y_i)\}_{i=1}^m \in \mathbf{R}^d \times \{\pm 1\}$, 记 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_m] \in \mathbf{R}^{d \times m}$ 为样本矩阵, $\mathbf{y} = [y_1, \dots, y_m]^T \in \mathbf{R}^m$ 为类标签向量, $\mathbf{D} = \text{diag}(y_1, \dots, y_m)$ 为类标签矩阵, $\mathbf{e}_m = (1, \dots, 1)^T \in \mathbf{R}^m$ 为 m 维 1 向量, $\mathbf{I}_m \in \mathbf{R}^{m \times m}$ 为 m 阶单位矩阵.

线性 SVM 是通过构建下面的二次规划模型

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i, \\ \text{s. t.} \quad & y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, m, \end{aligned} \quad (1)$$

来寻找分类决策函数 $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$ (也称分类器), 其中 $\mathbf{w} \in \mathbf{R}^d$ 和 $b \in \mathbf{R}$ 分别是决策函数的法向量和阈值, $C > 0$ 是调节参数(也称模型参数), 主要是调节结构风险和惩罚项之间的权重, $\xi_i \geq 0$ 是松弛变量, $\langle \mathbf{w}, \mathbf{x}_i \rangle$ 表示 \mathbf{w} 与 \mathbf{x}_i 之间的内积, $\|\mathbf{w}\|$ 表示向量 \mathbf{w} 的欧式范数. 当 $C = 0$ 时, 模型(1)退化为硬间隔线性 SVM (hard-SVM).

引入 Lagrange 乘子向量 $\boldsymbol{\alpha} \in \mathbf{R}_+^m$, 并考虑模型(1)的 Lagrange 函数, 可得其 Wolfe 对偶模型

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D} \mathbf{X}^T \mathbf{X} \mathbf{D} \boldsymbol{\alpha} - \mathbf{e}_m^T \boldsymbol{\alpha}, \\ \text{s. t.} \quad & \mathbf{y}^T \boldsymbol{\alpha} = 0, 0 \leq \boldsymbol{\alpha} \leq C \mathbf{e}_m, \end{aligned} \quad (2)$$

通过求解模型(2)就可得到分类决策函数了. 需要注意的是, 模型(2)的解有可能不唯一.

为了得到对偶模型的唯一解, 同时也避免矩阵 $\mathbf{X}^T \mathbf{X}$ 的奇异性, 线性 RSVM 将模型(1)的目标函数中的间隔项 $\|\mathbf{w}\|^2/2$ 改进为 $(\|\mathbf{w}\|^2 + b^2)/2$, 从而得到下面的二次规划模型

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} \quad & \frac{1}{2} (\|\mathbf{w}\|^2 + b^2) + C \sum_{i=1}^m \xi_i, \\ \text{s. t.} \quad & y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, m, \end{aligned} \quad (3)$$

同理, 当 $C = 0$ 时, 模型(3)退化为硬间隔线性 RSVM (hard-RSVM). 类似于线性 SVM, 引入 Lagrange 乘子向量 $\boldsymbol{\alpha} \in \mathbf{R}_+^m$, 并考虑模型(3)的 Lagrange 函数, 可得其 Wolfe 对偶模型

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D} (\mathbf{X}^T \mathbf{X} + \mathbf{e}_m \mathbf{e}_m^T) \mathbf{D} \boldsymbol{\alpha} - \mathbf{e}_m^T \boldsymbol{\alpha}, \\ \text{s. t.} \quad & 0 \leq \boldsymbol{\alpha} \leq C \mathbf{e}_m. \end{aligned} \quad (4)$$

通过求解模型(4), 可得到分类决策函数, 具体算法如下.

算法 1(RSVM)

步 1 给定数据集 $T = \{(\mathbf{x}_i, y_i)\}_{i=1}^m \in \mathbf{R}^d \times \{\pm 1\}$, 选择适当的模型参数 $C > 0$.

步 2 求解模型(4), 得最优解 $\boldsymbol{\alpha}^* \in \mathbf{R}_+^m$.

步 3 计算 $\mathbf{w}^* = \mathbf{X} \mathbf{D} \boldsymbol{\alpha}^* = \sum_{i=1}^m \boldsymbol{\alpha}_i^* y_i \mathbf{x}_i$, $b^* = \mathbf{e}_m^T \mathbf{D} \boldsymbol{\alpha}^* = \sum_{i=1}^m \boldsymbol{\alpha}_i^* y_i$.

步 4 构造分类决策函数 $f(\mathbf{x}) = \langle \mathbf{w}^*, \mathbf{x} \rangle + b^*$.

步 5 对任一输入样本 $\tilde{\mathbf{x}} \in \mathbf{R}^d$, 其类标签为 $y_{\tilde{\mathbf{x}}} = \text{sign}(f(\tilde{\mathbf{x}})) = \text{sign}(\langle \mathbf{w}^*, \tilde{\mathbf{x}} \rangle + b^*)$.

需要说明的是, 算法 1 的分类准确性严重依赖于模型参数 $C > 0$ 的选择, 常用的方法是在区间 $[10^{-4}, 10^4]$ 内进行网格搜索, 选择一个尽可能合适的 C 值.

1.2 基于 PDEs 的图像特征学习模型

考虑具有 c 个类的噪声(光暗、有遮挡)图像数据集 $\{(\mathbf{I}_m, \mathbf{h}_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times d_2} \times \mathbf{R}^c$, 其中 \mathbf{I}_m 为原始输入图

像, \mathbf{h}_m 为 \mathbf{I}_m 所对应的标签向量, 若 \mathbf{I}_m 属于第 i 类, 则第 i 个分量 $h_{mi} = 1$; 否则 $h_{mi} = 0$. 基于 PDEs 的图像特征学习模型为

$$\begin{aligned} & \min_{\mathbf{w}, \{a_i(t)\}} \frac{1}{M} \|\mathbf{H} - \mathbf{W} \cdot \mathbf{U}\|_{t=T}^2 + \lambda \|\mathbf{W}\|_F^2, \\ & \text{s. t. } \frac{\partial \mathbf{u}_m}{\partial t} = \sum_{i=0}^5 a_i(t) g(\text{inv}_i(\mathbf{u}_m(t))), (x, y, t) \in \mathbf{Q}, \\ & \mathbf{u}_m(x, y, t) = 0, (x, y, t) \in \mathbf{\Gamma}, \mathbf{u}_m|_{t=0}(x, y, t) = \mathbf{I}_m, (x, y) \in \mathbf{\Omega}, \end{aligned} \quad (5)$$

其中 $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_M] \in \mathbf{R}^{c \times M}$, $(x, y) \in \mathbf{\Omega}$ 为空间变量, $t \in [0, T]$ 表示时间变量, $\mathbf{\Omega}$ 是 \mathbf{R}^2 上的一个开区间, $\partial\mathbf{\Omega}$ 是 $\mathbf{\Omega}$ 的边界, $\mathbf{Q} = \mathbf{\Omega} \times [0, T]$, $\mathbf{\Gamma} = \partial\mathbf{\Omega} \times [0, T]$, $\mathbf{U}|_{t=T} = [\text{vec}(\mathbf{u}_1|_{t=T}), \dots, \text{vec}(\mathbf{u}_M|_{t=T})] \in \mathbf{R}^{d_1 d_2 \times M}$, $\mathbf{u}|_{t=T}$ 表示当 $t = T$ 时产生的特征矩阵, $\text{vec}(\mathbf{u})$ 表示矩阵 \mathbf{u} 的向量化, $\lambda > 0$ 是调节参数, $\text{inv}(\mathbf{u})$ 表示基本微分不变量且 $\{\text{inv}_i(\mathbf{u}_m(x, y, t))\}_{i=0}^5$ 是直到二阶的基本微分不变量(具体见表 1, 其中 \mathbf{u} 表示进化后的特征矩阵, $\nabla \mathbf{u}$ 表示矩阵 \mathbf{u} 的梯度, \mathbf{H}_u 表示 \mathbf{u} 的 Hessian 阵, $\text{tr}(\mathbf{H}_u)$ 表示矩阵 \mathbf{H}_u 的迹, \mathbf{u}_x 表示 \mathbf{u} 对矩阵 \mathbf{x} 求偏导), 控制系数 $a_i(t)$ 与空间变量无关, 保证了图像的平移和旋转不变性. $\mathbf{W} \in \mathbf{R}^{c \times d_1 d_2}$ 表示权矩阵, $g(x)$ 为添加的非线性压缩映射, 能够保证 $\{g(\text{inv}_i(\mathbf{u}_m))\}_{i=0}^5$ 在灰度条件下仍保持不变.

表 1 直至 2 阶的基本微分不变量

阶数 i	基本微分不变量 $\text{inv}_i(\mathbf{u})$
0, 1, 2	$1, \mathbf{u}, \ \nabla \mathbf{u}\ ^2 = \mathbf{u}_x^2 + \mathbf{u}_y^2$
3	$\text{tr}(\mathbf{H}_u) = \mathbf{u}_{xx} + \mathbf{u}_{yy}$
4	$(\nabla \mathbf{u})^T \mathbf{H}_u \nabla \mathbf{u} = \mathbf{u}_x^2 \mathbf{u}_{xx} + 2\mathbf{u}_x \mathbf{u}_y \mathbf{u}_{xy} + \mathbf{u}_y^2 \mathbf{u}_{yy}$
5	$\text{tr}(\mathbf{H}_u^2) = \mathbf{u}_{xx}^2 + 2\mathbf{u}_{xy}^2 + \mathbf{u}_{yy}^2$

为了求解模型(5), 先要求解 PDEs.

$$\begin{cases} \frac{\partial \mathbf{u}_m}{\partial t} = \sum_{i=0}^5 a_i(t) g(\text{inv}_i(\mathbf{u}_m(t))), (x, y, t) \in \mathbf{Q} \times [0, T], \\ \mathbf{u}_m(x, y, t) = 0, (x, y, t) \in \mathbf{\Gamma} \times [0, T], \\ \mathbf{u}_m(x, y, 0) = \mathbf{I}_m, (x, y) \in \mathbf{\Omega}, \\ m = 1, \dots, M. \end{cases} \quad (6)$$

通过离散化时间变量 $t_i = i \cdot \Delta t$, $i = 0, \dots, N$, 并用前向差分离散时间偏导数, 用中心差分离散时间变量, 可将 PDEs(6) 转化为下面的迭代公式(具体推导见文献[20])

$$\begin{cases} \mathbf{u}_m^{n+1} = \mathbf{u}_m^n + \Delta t \sum_{i=0}^5 a_i^n g(\text{inv}_i(\mathbf{u}_m^n)) \in \mathbf{R}^{d_1 \times d_2}, \\ \mathbf{u}_m^0 = \mathbf{I}_m, \end{cases} \quad (7)$$

其中 $\mathbf{u}_m^n = \mathbf{u}(x, y, t_n)$, $a_i^n = a_i(t_n)$, $n = 0, 1, \dots, N-1$, $m = 1, \dots, M$. 将(7)式带入模型(5)中可得如下图像特征学习模型

$$\begin{aligned} & \min_{\mathbf{w}, \{a_i(t)\}} \frac{1}{M} \|\mathbf{H} - \mathbf{W} \cdot \mathbf{U}\|_{t=T}^2 + \lambda \|\mathbf{W}\|_F^2, \\ & \text{s. t. } \mathbf{u}_m^{n+1} = \mathbf{u}_m^n + \Delta t \sum_{i=0}^5 a_i^n g(\text{inv}_i(\mathbf{u}_m^n)) \in \mathbf{R}^{d_1 \times d_2}, \\ & \mathbf{u}_m^0 = \mathbf{I}_m, n = 0, 1, \dots, N-1, m = 1, \dots, M. \end{aligned} \quad (8)$$

2 基于 PDEs 特征学习的 RSVM

给定二分类噪声(光暗, 有遮挡)图像数据集 $\{(\mathbf{I}_m, \mathbf{h}_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times d_2} \times \{\pm 1\}$, 如果直接利用 RSVM 进行分类, 可通过下面的二次规划模型

$$\begin{aligned} & \min_{\mathbf{w}, b, \xi} \frac{1}{2} (\|\mathbf{w}\|^2 + b^2) + C \mathbf{e}_M^T \boldsymbol{\xi}, \\ & \text{s. t. } h_m (\langle \mathbf{w}, \mathbf{i}_m \rangle + b) \geq 1 - \xi_m, \xi_m \geq 0, m = 1, \dots, M, \end{aligned} \quad (9)$$

来寻找分类决策函数 $f(\mathbf{I}) = \langle \mathbf{w}, \mathbf{i} \rangle + b$, 其中 $\mathbf{w} \in \mathbf{R}^{d_1 d_2}$ 和 $b \in \mathbf{R}$ 分别为决策函数的法向量和阈值, $\mathbf{i}_m = \text{vec}(\mathbf{I}_m) \in \mathbf{R}^{d_1 d_2}$ 为原始图像的向量化. 然而, 由于噪声的存在将导致分类准确度下降. 为了解决这一问题, 需要对图像进行降噪处理. 一种方法是采用降噪处理方法对图像进行预处理, 然后放入模型(9)中学习分类器, 这是一种次优方法, 因为预处理好的图像并不一定适合模型(9). 另一种方法是将降噪处理与学习分类器放入同一个模型中联合学习. 本文将图像特征学习模型(8)与模型(9)融合, 得到最优化模型

$$\begin{aligned}
 & \min_{\mathbf{w}, b, \{a_i^n(t)\}, \xi} \frac{1}{2} (\|\mathbf{W}\|_F^2 + b^2) + C \mathbf{e}_M^T \boldsymbol{\xi}, \\
 & \text{s. t. } \mathbf{h}_m (\langle \mathbf{W}, \mathbf{u}_m |_{t=T} \rangle + b) \geq 1 - \xi_m, \\
 & \quad \xi_m \geq 0, \mathbf{u}_m^{n+1} = \mathbf{u}_m^n + \Delta t \sum_{i=0}^5 a_i^n g(\text{inv}_i(\mathbf{u}_m^n)) \in \mathbf{R}^{d_1 \times d_2}, \\
 & \quad \mathbf{u}_m^0 = \mathbf{I}_m, n=0, 1, \dots, N-1, m=1, \dots, M.
 \end{aligned} \tag{10}$$

这时的分类决策函数可表示为 $f(\mathbf{I}) = \langle \mathbf{W}, \mathbf{u} |_{t=T} \rangle + b$, 其中 $\mathbf{W} \in \mathbf{R}^{d_1 \times d_2}$ 表示权矩阵, $\mathbf{u} |_{t=T} \in \mathbf{R}^{d_1 \times d_2}$ 是图像 $\mathbf{I} \in \mathbf{R}^{d_1 \times d_2}$ 经过 PDEs 进化提取的特征矩阵, $\|\cdot\|_F$ 表示矩阵的 F -范数. 记 $\mathbf{w} = \text{vec}(\mathbf{W}) \in \mathbf{R}^{d_1 d_2}$, $\mathbf{D} = \text{diag}(\mathbf{h}_1, \dots, \mathbf{h}_M) \in \mathbf{R}^{M \times M}$, $\mathbf{U}^N = [\text{vec}(\mathbf{u}_1^N), \dots, \text{vec}(\mathbf{u}_M^N)] \in \mathbf{R}^{d_1 d_2 \times M}$, 则分类决策函数和模型(10)可分别表示为 $f(\mathbf{I}) = \langle \mathbf{w}, \text{vec}(\mathbf{u}_i^N) \rangle + b$ 和

$$\begin{aligned}
 & \min_{\mathbf{w}, b, a_i^n, \xi} \frac{1}{2} (\|\mathbf{w}\|_2^2 + b^2) + C \mathbf{e}_M^T \boldsymbol{\xi}, \\
 & \text{s. t. } \mathbf{D}((\mathbf{U}^N)^T \boldsymbol{\tau} \mathbf{w} + b \mathbf{e}_M) \geq \mathbf{e}_M - \boldsymbol{\xi}, \boldsymbol{\xi} \geq 0.
 \end{aligned} \tag{11}$$

下面用迭代法求解模型(11). 在第 k 步迭代中, 首先, 固定系数矩阵 $\mathbf{A} = \mathbf{A}^k \in \mathbf{R}^{N \times 6}$, 更新 (\mathbf{W}^k, b^k) , 其中 $\mathbf{A} = [a(n, i)]_{N \times 6}$, $a(n, i) = a_i^n$, $n=0, 1, \dots, N-1, i=0, 1, \dots, 5$.

利用(7)式计算 $(\mathbf{U}^N)^k$. 考虑模型(11)的 Lagrange 函数 $L(\mathbf{w}, b, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = (\|\mathbf{w}\|_2^2 + b^2)/2 + C \mathbf{e}_M^T \boldsymbol{\xi} - \boldsymbol{\alpha}^T (\mathbf{D}((\mathbf{U}^N)^k)^T \boldsymbol{\tau} \mathbf{w} + b \mathbf{D} \mathbf{e}_M - \mathbf{e}_M + \boldsymbol{\xi}) - \boldsymbol{\beta}^T \boldsymbol{\xi}$, 其中 $\mathbf{U}^N = (\mathbf{U}^N)^k$, 并令 $\partial L / \partial \mathbf{w} = \partial L / \partial b = \partial L / \partial \boldsymbol{\xi} = 0$, 得

$$\begin{cases} \mathbf{w} = \mathbf{U}^N \mathbf{D} \boldsymbol{\alpha}, b = \mathbf{e}_M^T \mathbf{D} \boldsymbol{\alpha}, \\ C \mathbf{e}_M - \boldsymbol{\alpha} - \boldsymbol{\beta} = 0 \Rightarrow 0 \leq \boldsymbol{\alpha} \leq C \mathbf{e}_M, \end{cases} \tag{12}$$

将(12)式带入 Lagrange 函数中, 有

$$\begin{aligned}
 L(\boldsymbol{\alpha}, \mathbf{U}^N) &= \frac{1}{2} (\|\mathbf{w}\|_2^2 + b^2) - \boldsymbol{\alpha}^T \mathbf{D}(\mathbf{U}^N)^T \mathbf{w} - b \mathbf{e}_M^T \mathbf{D} \boldsymbol{\alpha} + \mathbf{e}_M^T \boldsymbol{\alpha} \\
 &= -\frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D}(\mathbf{U}^N)^T \mathbf{U}^N \mathbf{D} \boldsymbol{\alpha} - \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D} \mathbf{e}_M \mathbf{e}_M^T \mathbf{D} \boldsymbol{\alpha} + \mathbf{e}_M^T \boldsymbol{\alpha} \\
 &= -\frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D}[(\mathbf{U}^N)^T \mathbf{U}^N + \mathbf{e}_M \mathbf{e}_M^T] \mathbf{D} \boldsymbol{\alpha} + \mathbf{e}_M^T \boldsymbol{\alpha},
 \end{aligned} \tag{13}$$

综合(12)-(13)式, 可得模型(11)的 Wolfe 对偶形式

$$\begin{aligned}
 & \min_{\boldsymbol{\alpha}} \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D}[(\mathbf{U}^N)^T \mathbf{U}^N + \mathbf{e}_M \mathbf{e}_M^T] \mathbf{D} \boldsymbol{\alpha} - \mathbf{e}_M^T \boldsymbol{\alpha} \\
 & \text{s. t. } 0 \leq \boldsymbol{\alpha} \leq C \mathbf{e}_M.
 \end{aligned} \tag{14}$$

求解模型(14), 得最优解 $\boldsymbol{\alpha}^k$. 将 $\boldsymbol{\alpha}^k$ 代入(12)式中, 得 $\mathbf{w}^k = \mathbf{U}^N \mathbf{D} \boldsymbol{\alpha}^k, b^k = \mathbf{e}_M^T \mathbf{D} \boldsymbol{\alpha}^k$, 其中 $\mathbf{U}^N = (\mathbf{U}^N)^k$.

其次, 固定 $\boldsymbol{\alpha} = \boldsymbol{\alpha}^k$, 更新 $\mathbf{A}^k \in \mathbf{R}^{N \times 6}$.

利用梯度下降法, 可得迭代公式

$$(a_i^n)^{k+1} = (a_i^n)^k - \eta \left(\frac{\partial L}{\partial a_i^n} \right)^k, i=0, \dots, 5, n=0, \dots, N-1, \tag{15}$$

其中 $\eta > 0$ 是迭代步长. 偏导数 $\partial L / \partial a_i^n$ 可由公式(7)推出(详细推导见文献[20])

$$\begin{aligned}
 \frac{\partial L}{\partial \mathbf{U}^N} &= -\mathbf{U}^N \mathbf{D} \boldsymbol{\alpha} \boldsymbol{\alpha}^T \mathbf{D}, \\
 \frac{\partial L}{\partial \mathbf{U}^{N-1}} &= -\mathbf{U}^N \mathbf{D} \boldsymbol{\alpha} \boldsymbol{\alpha}^T \mathbf{D} + \Delta t \sum_{i=0}^5 a_i^{N-1} \mathbf{C}_i^{N-1} \in \mathbf{R}^{d_1 d_2 \times M}, \\
 \frac{\partial L}{\partial \mathbf{U}^n} &= \frac{\partial L}{\partial \mathbf{U}^{n+1}} + \Delta t \sum_{i=0}^5 a_i^n \mathbf{C}_i^n \in \mathbf{R}^{d_1 d_2 \times M}, n=0, \dots, N-2, \\
 \frac{\partial L}{\partial a_i^{N-1}} &= -\Delta t \text{Tr}[(\mathbf{U}^N)^T \mathbf{B}_i^N \mathbf{D} \boldsymbol{\alpha} \boldsymbol{\alpha}^T \mathbf{D}] = -\Delta t \langle \mathbf{U}^N, \mathbf{B}_i^N \mathbf{D} \boldsymbol{\alpha} \boldsymbol{\alpha}^T \mathbf{D} \rangle, \\
 \frac{\partial L}{\partial a_i^n} &= \Delta t \text{Tr}[(\frac{\partial L}{\partial \mathbf{U}^{n+1}})^T, \mathbf{B}_i^n], i=0, \dots, 5, n=0, \dots, N-1,
 \end{aligned} \tag{16}$$

其中 $\mathbf{B}_i^n = [\text{vec}(g(\text{inv}_i(\mathbf{u}_1^n))), \dots, \text{vec}(g(\text{inv}_i(\mathbf{u}_M^n)))] \in \mathbf{R}^{d_1 d_2 \times M}, i=0, \dots, 5, n=0, \dots, N-2, \mathbf{C}_i^n = [\text{vec}(\mathbf{Z}(i, 1,$

$n)), \dots, \text{vec}(\mathbf{Z}(i, M, n))] \in \mathbf{R}^{d_1 d_2 \times M}, i=0, \dots, 5, n=0, \dots, N-1$, 且

$$\begin{cases} \mathbf{Z}(i, m, n)(p, q) = \frac{\partial \mathbf{L}}{\partial \mathbf{u}_m^{n+1}}(p+1, q) \cdot \frac{\partial g(\text{inv}_i(\mathbf{u}_m^n)(p+1, q))}{\partial \mathbf{u}_m^n(p, q)} + \frac{\partial \mathbf{L}}{\partial \mathbf{u}_m^{n+1}}(p-1, q) \cdot \frac{\partial g(\text{inv}_i(\mathbf{u}_m^n)(p-1, q))}{\partial \mathbf{u}_m^n(p, q)} \\ + \frac{\partial \mathbf{L}}{\partial \mathbf{u}_m^{n+1}}(p, q+1) \cdot \frac{\partial g(\text{inv}_i(\mathbf{u}_m^n)(p, q+1))}{\partial \mathbf{u}_m^n(p, q)} + \frac{\partial \mathbf{L}}{\partial \mathbf{u}_m^{n+1}}(p, q-1) \cdot \frac{\partial g(\text{inv}_i(\mathbf{u}_m^n)(p, q-1))}{\partial \mathbf{u}_m^n(p, q)} \\ + \frac{\partial \mathbf{L}}{\partial \mathbf{u}_m^{n+1}}(p, q) \cdot \frac{\partial g(\text{inv}_i(\mathbf{u}_m^n)(p, q))}{\partial \mathbf{u}_m^n(p, q)}, \\ \mathbf{Z}(i, m, n) = [\mathbf{Z}(i, m, n)(p, q)] \in \mathbf{R}^{d_1 \times d_2}. \end{cases}$$

下面给出具体算法.

算法2 (基于 PDEs 特征学习的 RSVM)

步1 给定二分类噪声(光暗,有遮挡)图像数据集 $\{(I_m, \mathbf{h}_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times d_2} \times \{\pm 1\}$, 置 $k=0, \epsilon > 0, \Delta t = T/N, N=5, k_{\max}=10, \eta=1, \rho=0.95$. 在区间 $[-1, 1]$ 内均匀采样, 得矩阵 $\mathbf{A}^k \in \mathbf{R}^{N \times 6}$.

步2 利用迭代公式(7) 计算 $(\mathbf{U}^N)^k = [\text{vec}((\mathbf{u}_1^N)^k), \dots, \text{vec}((\mathbf{u}_M^N)^k)] \in \mathbf{R}^{d_1 d_2 \times M}$, 其中 $\mathbf{A} = \mathbf{A}^k$.

步3 求解对偶模型(14), 得最优解 α^k . 计算 $\mathbf{w}^k = \mathbf{U}^N \mathbf{D} \alpha^k$ 和 $b^k = \mathbf{e}_M^T \mathbf{D} \alpha^k$, 其中 $\mathbf{U}^N = (\mathbf{U}^N)^k$.

步4 利用(16)式计算 $(\partial \mathbf{L} / \partial \mathbf{U}^N)^k, n=0, \dots, N$ 和 $(\partial \mathbf{L} / \partial a_i^n)^k, i=0, \dots, 5, n=0, \dots, N-1$, 其中 $\mathbf{A} = \mathbf{A}^k, \alpha = \alpha^k$.

步5 利用迭代公式(15) 更新 \mathbf{A}^k 至 \mathbf{A}^{k+1} .

步6 利用公式(7) 更新 $(\mathbf{U}^N)^k$ 至 $(\mathbf{U}^N)^{k+1}$, 其中 $\mathbf{A} = \mathbf{A}^{k+1}$.

步7 求解对偶模型(14), 得最优解 α^{k+1} . 计算 $\mathbf{w}^{k+1} = \mathbf{U}^N \mathbf{D} \alpha^{k+1}, b^{k+1} = \mathbf{e}_M^T \mathbf{D} \alpha^{k+1}$, 其中 $\mathbf{U}^N = (\mathbf{U}^N)^{k+1}$.

步8 若 $\|\mathbf{w}^{k+1} - \mathbf{w}^k\| < \epsilon, \|b^{k+1} - b^k\| < \epsilon$ 或 $k = k_{\max}$, 停止迭代, 置 $\mathbf{w}^* \leftarrow \mathbf{w}^{k+1}, b^* \leftarrow b^{k+1}$, 转步9; 否则, 置 $k \leftarrow k+1, \eta \leftarrow \rho \eta$, 转步4.

步9 构造分类决策函数 $f(I) = \langle \mathbf{w}^*, \text{vec}(\mathbf{u}_I^N) \rangle + b^*$.

步10 对任意输入的图像样本 $I \in \mathbf{R}^{d_1 \times d_2}$, 其类标签为 $\mathbf{h}_I = \text{sign}(f(I))$.

3 实验结果与分析

为了检验 PDEs 对 RSVM 中模型参数的影响, 本节选用 Extended YaleB 数据集的 7 对数据(见表 7) 和 PIE 数据集的 4 对数据(见表 8) 进行了实验. Extended YaleB 数据集包括 38 个人分别在 $12^\circ, 25^\circ, 50^\circ, 77^\circ, 90^\circ$ 的不同光照角度下的共 2414 个面部图像, 图像尺寸为 48×42 . PIE 数据集包含 68 个人, 每人 24 个不同角度光照和表情的面部图像, 图像尺寸为 64×64 .

3.1 平台的选择

对一对图像执行算法 2 可用两种方式: 一种是在不同平台下分别进行 PDEs 进化; 另一种是在同一平台下进行 PDEs 进化. 也就是说, 第一种方式是每个人训练自己的 PDEs 进化器, 第二种方式是两个人联合训练一个 PDEs 进化器, 如图 1 和图 2 所示

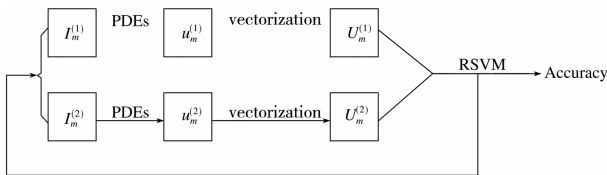


图 1 不同平台下进行 PDEs 进化

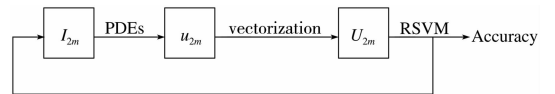


图 2 同一平台下进行 PDEs 进化

其中图 1 中的 $I_m^{(i)} = [\mathbf{x}_1^i, \dots, \mathbf{x}_m^i] \in \mathbf{R}^{d_1 \times m d_2} (i=1, 2)$ 表示第 i 人的图像集, $\mathbf{x}_j^i \in \mathbf{R}^{d_1 \times d_2} (j=1, \dots, m)$ 表示第 j 个图像, $\mathbf{u}_m^{(i)} \in \mathbf{R}^{d_1 \times m d_2}$ 表示经过 PDEs 进化后的特征矩阵, $U_m^{(i)} \in \mathbf{R}^{d_1 d_2 \times m}$ 为 $\mathbf{u}_m^{(i)}$ 的向量化. 图 2 中 $I_{2m} = [\mathbf{x}_1^{(1)}, \dots, \mathbf{x}_m^{(1)}, \mathbf{x}_1^{(2)}, \dots, \mathbf{x}_m^{(2)}] \in \mathbf{R}^{d_1 \times 2m d_2}$ 表示两个人的图像集, $\mathbf{x}_j^{(i)} (i=1, 2, j=1, \dots, m)$ 表示第 i 人的第 j 个图像, $\mathbf{u}_{2m} \in \mathbf{R}^{d_1 \times 2m d_2}$ 表示经过 PDEs 进化后的特征矩阵, $U_{2m} \in \mathbf{R}^{d_1 d_2 \times 2m}$ 为 \mathbf{u}_{2m} 的向量化.

为了说明两种进化方式的不同, 本小节选用 Extended YaleB 数据集中第 1 人和第 2 人各 10 张面部图像进行试验, 执行 10 次 PDEs 进化, 结果见图 3 至图 6, 其中图 3 和图 4 是同一平台下的进化结果, 图 5 和图 6 是

不同平台下的进化结果. 在图 3 至图 6 中,第一行表示 10 个原始图像,第二行至第十一行分别表示进化一次至进化十次的结果.

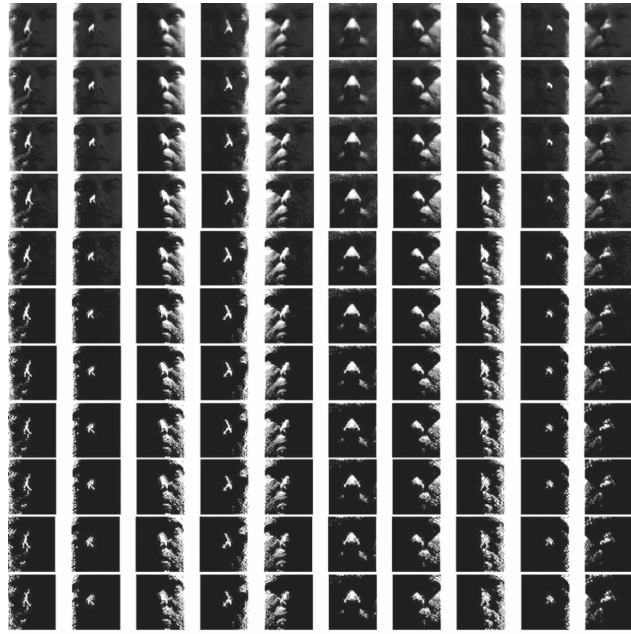


图 3 同一平台下第 1 人的 PDEs 进化结果

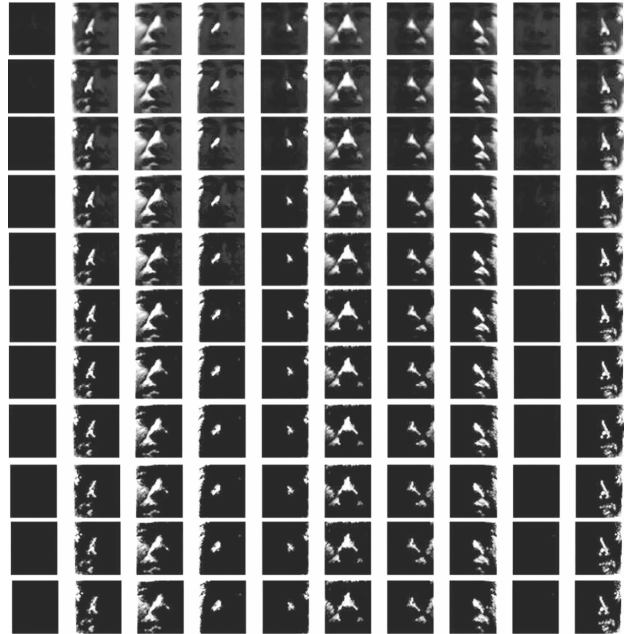


图 4 同一平台下第 2 人的 PDEs 进化结果

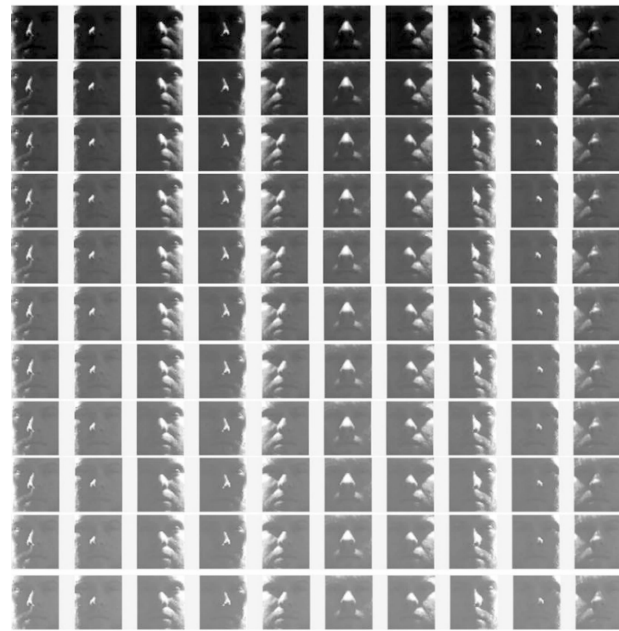


图 5 不同平台下第 1 人的 PDEs 进化结果

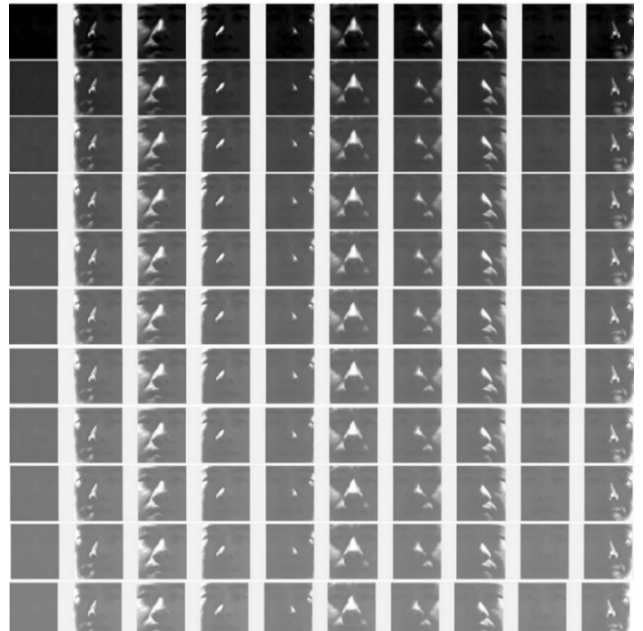


图 6 不同平台下第 2 人的 PDEs 进化结果

从图 3 和图 4 中可以看出,随着 PDEs 进化次数的增加,特征图像的像素差也随之逐渐增大,这有利于对图像的识别和分类.而在图 5 和图 6 中,随着进化次数的增加,特征图像整体有变亮的趋势,但像素差越来越小.因此,为了提高分类准确率,本文采用在同一平台下进行 PDEs 特征提取.

3.2 PDEs 对 RSVM 中模型参数的影响

为了检验 PDEs 对 RSVM 中模型参数的影响,本节随机选取 Extended YaleB 数据集的 7 对数据(见图 7),每人选取 30 张受光照影响较大的图像和 PIE 数据集的 4 对数据(见图 8),每人 24 张图像进行了实验.为了便于表述,图 7 中的 7 对数据依次标记为第 1 人至第 14 人,第一行从左至右依次为第 1 人至第 7 人,第二行依次为第 8 人至第 14 人;图 8 中的 4 对数据从左至右依次标记为第 1 人至第 8 人.对 Extended YaleB 数据集的实验采用五折交叉验证,对 PIE 数据集的实验采用四折交叉验证.且在 PDEs 进化过程中不使用压缩

函数,模型参数分别取为 $C=10^{-3}$ 、 10^{-2} 、 10^{-1} 、 10^0 、 10^1 、 10^2 . 实验结果见表 2 和表 3,其中 0 列表示未经过 PDEs 进化的结果,即直接利用 RSVM 进行分类的结果. 1 列至 10 列分别是进化 1 次至 10 次的结果,其中,为了便于比较,分类准确率降至 0.5 以下用“—”表示.



图 7 选用 Extended YaleB 数据集中的 7 对数据



图 8 选用 PIE 数据集中的 4 对数据

表 2 Extended YaleB 数据集的分类结果

第 1 人和第 2 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.5167	0.5167	0.5167	0.5167	0.5167	0.5167	0.5167	0.5167	0.5167	0.5167	0.5167
10^{-2}	0.9167	0.9833	1.0000	1.0000	1.0000	0.9167	0.8500	0.6667	0.6333	0.6167	0.7000
10^{-1}	0.9667	0.9667	0.9833	0.9667	0.9667	0.9167	0.8000	0.6667	0.6000	0.5500	—
1	0.9500	0.9667	0.9833	0.9667	0.9500	0.9167	0.7333	0.5000	0.5833	—	—
10^1	0.9500	0.9667	0.9833	0.9667	0.9500	0.9000	0.7667	0.5333	0.5333	—	—
10^2	0.9500	0.9667	0.9833	0.9667	0.9500	0.9000	0.7667	0.5333	0.5333	—	—

第 3 人和第 4 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.5667	0.5667	0.5667	0.5667	0.5667	0.5667	0.5667	0.5667	0.5667	0.5667	0.5667
10^{-2}	0.9500	1.0000	1.0000	1.0000	1.0000	0.9667	0.8833	0.8333	0.7833	0.7333	0.6833
10^{-1}	1.0000	1.0000	1.0000	1.0000	1.0000	0.9667	0.7667	0.7000	0.4167	—	—
1	1.0000	1.0000	1.0000	1.0000	0.9333	0.7833	0.5833	0.4833	—	—	—
10^1	1.0000	1.0000	1.0000	1.0000	0.9333	0.7833	0.5833	—	—	—	—
10^2	1.0000	1.0000	1.0000	1.0000	0.9333	0.7833	0.5833	—	—	—	—

第 5 人和第 6 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.5000	0.5000	0.5167	0.5167	0.5333	0.5333	0.5333	0.5500	0.5500	0.5500	0.5667
10^{-2}	1.0000	1.0000	1.0000	0.9833	0.9833	0.9667	0.9667	0.9333	0.8333	0.7833	0.7167
10^{-1}	1.0000	1.0000	0.9833	0.9833	0.9833	0.9500	0.9167	0.8167	0.7000	0.5333	—
1	1.0000	1.0000	0.9833	0.9833	0.9833	0.9500	0.8833	0.7667	0.6000	0.6333	0.5333
10^1	1.0000	1.0000	0.9833	0.9833	0.9833	0.9500	0.8833	0.7667	0.6000	0.6333	0.5333
10^2	1.0000	1.0000	0.9833	0.9833	0.9833	0.9500	0.8833	0.7667	0.6000	0.6333	0.5333

第 7 人和第 8 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.6333	0.6333	0.6333	0.6333	0.6333	0.6333	0.6333	0.6500	0.6500	0.6500	0.6500
10^{-2}	0.9500	0.9833	0.9833	0.9833	0.9500	0.8833	0.7500	0.7333	0.7500	0.6500	0.6333
10^{-1}	0.9833	0.9833	0.9667	0.9500	0.9500	0.8500	0.8167	0.6000	—	—	—
1	0.9833	0.9833	0.9833	0.9500	0.8833	0.8667	0.7333	0.5333	—	—	—
10^1	0.9833	0.9833	0.9500	0.9167	0.9167	0.8500	0.7333	0.5500	—	—	—
10^2	0.9833	0.9833	0.9500	0.9167	0.9167	0.8500	0.7333	0.5500	—	—	—

第 9 人和第 10 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.6000	0.6000	0.6000	0.6000	0.6000	0.6000	0.6167	0.6167	0.6167	0.6167	0.6167
10^{-2}	1.0000	1.0000	1.0000	0.9833	0.9667	0.8000	0.7667	0.7500	0.7000	0.6667	0.6333
10^{-1}	1.0000	1.0000	1.0000	0.9833	0.9833	0.9333	0.8500	0.7000	0.6500	0.5167	—
1	1.0000	1.0000	1.0000	0.9833	0.9833	0.9500	0.8167	0.6667	0.6000	0.5000	—
10^1	1.0000	1.0000	1.0000	0.9833	0.9833	0.9500	0.8167	0.6667	0.6000	0.5000	—
10^2	1.0000	1.0000	1.0000	0.9833	0.9833	0.9500	0.8167	0.6667	0.6000	0.5000	—

第 11 人和第 12 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
10^{-2}	0.9167	0.9833	1.0000	0.9833	0.9500	0.9000	0.8167	0.6833	0.5000	0.4833	0.4833
10^{-1}	0.9500	0.9667	0.9667	0.9000	0.8333	0.6833	0.5667	0.4333	0.5333	—	—
1	0.9333	0.9500	0.9167	0.8500	0.6167	0.5000	—	—	—	—	—
10^1	0.9333	0.9167	0.9167	0.8333	0.5667	0.5500	0.5167	—	—	—	—
10^2	0.9000	0.9333	0.9167	0.6500	0.4667	0.5667	—	—	—	—	—

第 13 人和第 14 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.4833	0.4833	0.4833	0.4833	0.4833	0.4833	0.4833	0.4833	0.4833	0.4833	0.4833
10^{-2}	0.8833	1.0000	1.0000	1.0000	1.0000	0.9500	0.7667	0.7167	0.8000	0.7833	0.7500
10^{-1}	1.0000	1.0000	1.0000	0.9833	0.9833	0.8500	0.7500	0.6833	0.6000	0.4167	—
1	1.0000	1.0000	1.0000	1.0000	0.9833	0.9667	0.8667	0.6667	—	—	—
10^1	1.0000	1.0000	1.0000	1.0000	0.9833	0.9667	0.8667	0.6667	—	—	—
10^2	1.0000	1.0000	1.0000	1.0000	0.9833	0.9667	0.8667	0.6667	—	—	—

表 3 PIE 数据集的分类结果

第 1 人和第 2 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.6875	0.6875	0.7083	0.7083	0.7292	0.7292	0.7500	0.7917	0.7917	0.8333	0.8542
10^{-2}	0.8958	0.8958	0.9375	0.9375	0.9375	0.8750	0.7083	0.5833	0.5625	0.5208	0.5417
10^1	0.8542	0.9167	0.9167	0.9167	0.9375	0.9167	0.7083	0.6250	—	—	—
10^0	0.8750	0.9167	0.9167	0.8958	0.9167	0.6250	0.6458	0.5000	—	—	—
10^1	0.8750	0.9167	0.9167	0.8958	0.9167	0.6250	0.6458	0.5000	—	—	—
10^2	0.8750	0.9167	0.9167	0.8958	0.9167	0.6250	0.6458	0.5000	—	—	—

第 3 人和第 4 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.5000	0.5000	0.5000	0.5000	0.5000	0.4792	0.4792	0.4792	0.4792	0.4792	0.5000
10^{-2}	0.9167	0.9167	0.9167	0.9167	0.8958	0.8333	0.7083	0.6458	0.7083	0.6667	0.6250
10^{-1}	0.9167	0.9167	0.9167	0.8958	0.8958	0.8750	0.7708	0.7083	0.6667	—	—
1	0.9167	0.9167	0.9167	0.8958	0.8958	0.8750	0.7500	0.6667	0.5208	0.5000	—
10^1	0.9167	0.9167	0.9167	0.8958	0.8958	0.8750	0.7500	0.6667	0.5208	0.5000	—
10^2	0.9167	0.9167	0.9167	0.8958	0.8958	0.8750	0.7500	0.6667	0.5208	0.5000	—

第 5 人和第 6 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
10^{-2}	0.8958	0.9167	0.9167	0.9167	0.9167	0.9375	0.8542	0.7500	0.6667	0.6458	0.6250
10^{-1}	0.8958	0.8958	0.9167	0.9167	0.9167	0.9375	0.8958	0.7917	0.7500	0.7083	0.5417
1	0.8958	0.8958	0.9167	0.9167	0.9167	0.9375	0.8958	0.7917	0.7500	0.7083	0.5417
10^1	0.8958	0.8958	0.9167	0.9167	0.9167	0.9375	0.8958	0.7917	0.7500	0.7083	0.5417
10^2	0.8958	0.8958	0.9167	0.9167	0.9167	0.9375	0.8958	0.7917	0.7500	0.7083	0.5417

第 7 人和第 8 人的分类结果

n/C	0	1	2	3	4	5	6	7	8	9	10
10^{-3}	0.6875	0.6875	0.6875	0.6875	0.7083	0.7083	0.7083	0.7083	0.7292	0.7292	0.7500
10^{-2}	0.9375	0.9375	0.9375	0.9375	0.8542	0.8333	0.7917	0.7917	0.7083	0.7083	0.7083
10^{-1}	0.9375	0.9375	0.9375	0.9375	0.8542	0.8333	0.7917	0.7917	0.6875	0.6875	0.5417
1	0.9375	0.9375	0.9375	0.9375	0.8542	0.8333	0.7917	0.7917	0.7083	0.6875	0.5208
10^1	0.9375	0.9375	0.9375	0.9375	0.8542	0.8333	0.7917	0.7917	0.7083	0.6875	0.5208
10^2	0.9375	0.9375	0.9375	0.9375	0.8542	0.8333	0.7917	0.7917	0.7083	0.6875	0.5208

从表 2 和表 3 中可以看出:(1) 随着 PDEs 的逐次进化,图像分类准确度也相应发生变化,总体上呈先升后降的趋势。(2) 参数 C 的取值越小,需要进化的次数越多.当 $C = 10^{-3}$ 时,进化 10 次以上最为合适,而当 $C = 10^{-2}, 10^{-1}$ 时,进化 4-5 次最为合适;当 $C = 10^0, 10^1, 10^2$ 时,进化 2-3 次即可。(3) 当 $C = 10^{-3}$ 时,RSVM 实际上等同于 hard-RSVM,故分类准确率明显下降。

综上所述,不论 C 取何值,都可通过适当次数的 PDEs 进化,得到较高的图像分类精度.因此,采取 PDEs 进行图像特征学习,可以弱化模型参数 C 的选择.换句话说,可以任意取一个 C 值,然后通过 PDEs 进化找到较高的分类精度,这回避了调参问题。

4 结论

本文以噪声(光暗,有遮挡)图像分类为背景,以 RSVM 为分类器,研究了 PDEs 对 RSVM 模型中参数选择的影响.从实验结果可以看出,PDEs 不仅具有降噪功能,还能弱化模型参数对分类结果的影响.如果参数取的小,则多进行几次 PDEs 学习;取的大,则少进行几次.由于 SVM 型分类器主要考虑的是结构风险极小化原则,而模型参数 C 的作用是调节结构风险和惩罚项之间的权重,故不可能取值太大.因此,通过 PDEs 的进化学习可以起到回避调参的作用.本文只是针对 RSVM 讨论了 PDEs 对模型参数的影响,能否在其他分类器中也有这样好的作用呢?这是我们下一步的研究工作。

参 考 文 献

- [1] Moscoso T E, Biasotti S. Description and retrieval of geometric patterns on surface meshes using an edge-based LBP approach[J]. Pattern Recognition, 2018, 82: 1-15.
- [2] Wei Y, Tian Q, Guo J H, et al. Multi-vehicle detection algorithm through combining Harr and HOG features[J]. Mathematics and Computers in Simulation, 2019, 155: 130-145.
- [3] Mohamed A, Issam A, Mohamed B, et al. Real-time detection of vehicles using the Haar-like features and artificial neuron networks[J]. Procedia Computer Science, 2015, 73: 24-31.
- [4] Li Y Z, Pang Y W, Wang J, et al. Patient-specific ECG classification by deeper CNN from generic to dedicated[J]. Neurocomputing, 2018, 314: 336-346.
- [5] Wang N F, Lao K Y, Zhang X H, et al. The recognition of grasping force using LDA[J]. Biomedical signal processing and control, 2019, 47: 393-400.
- [6] Li Y S, Qu J H, Dong W Q, et al. Hyperspectral pansharpening via improved PCA approach and optimal weighted fusion strategy[J]. Neurocomputing, 2018, 315: 371-380.
- [7] 邓乃扬, 田英杰著. 数据挖掘中的新方法—支持向量机[M]. 北京科学出版社, 2004.
- [8] Zhu G Y, Huang D, Zhang P, et al. ϵ -proximal support vector machine for binary classification and its application in vehicle recognition[J]. Neurocomputing, 2015, 161: 260-266.
- [9] Pei H M, Wang K N, Zhong P. Semi-supervised matrixized least squares support vector machine[J]. Applied Soft Computing, 2017, 61: 72-87.
- [10] 许颖春, 范丽亚. 基于 L_1 范数的非线性 TSVR[J]. 聊城大学学报(自然科学版), 2017, 30(3): 6-11.
- [11] 侯明, 张新新, 范丽亚. 四种基于支持向量机的多类分类器的性能比较[J]. 聊城大学学报(自然科学版), 2014, 27(3): 54-60.
- [12] Chen W J, Shao Y H, Li C N, et al. MLTSVM: A novel twin support vector machine to multi-label learning[J]. Pattern Recognition 2016, 52: 61-74.
- [13] 夏文静, 陈耿, 范丽亚. 八种最小二乘 SVM 型学习算法的优势比较[J]. 聊城大学学报(自然科学版), 2016, 29(2): 33-41.

- [14] Nasiri J A,Charkari N M,Jalili S. Least squares twin multi-class classification support vector machine[J]. Pattern Recognition,2015,48: 984-992.
- [15] Niell S,Jesus F,Diaz R,et al. Beehives biomonitor pesticides in agroecosystems;simple chemical and biological indicators evaluation using Support Vector Machines (SVM)[J]. Ecological Indicators,2018,91:149-154.
- [16] Tang X Y,Zeng W M,Shi Y H,et al. Brain activation detection by modified neighborhood one-class SVM on fMRI data[J]. Biomedical Signal Processing and Control,2018,39:448-458.
- [17] Rajisha T M,Sunija A P,Piyas K S. Performance analysis of Malayalam language speech emotion recognition system using ANN/SVM[J]. Procedia Technology,2016,24:1097-1104.
- [18] Tomar D,Agarwal S. A comparison on multi-class classification methods based on least squares twin support vector machine[J]. Knowledge-Based Systems,2015,81:131-147.
- [19] Liu R S,Lin Z C,Zhang W,et al. Learning PDEs for image restoration via optimal control[C]. //The 11th European Conference on Computer Vision,2010.
- [20] Liu R S,Lin Z C,Zhang W,et al. Toward designing intelligent PDEs for computer vision;An optimal control approach[J]. Image and vision computing,2013,31(1):43-56.
- [21] Liu R S,Zhong G Y,Cao J J,et al. Learning to diffuse:A new perspective to design PDEs for visual analysis[J]. IEEE transactions on pattern analysis and machine intelligence,2016,38(12):2457-2471.
- [22] Fang C,Zhao Z Y,Zhou P,et al. Feature learning via partial differential equation with applications to face recognition[J]. Pattern Recognition,2017,69:14-25.
- [23] 江珊珊,杨静,范丽亚. 基于 PDEs 的图像特征提取方法[J]. 山东大学学报(工学版),2018,48(4):27-35.
- [24] Georgiades A,Belhumeur P,Kriegman D. From few to many:illuminationcone models for face recognition under variable lighting and pose[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2001,23(6):643-660.
- [25] Saul L K,Weinberger K Q,Ham J H,et al. Spectral methods for dimensionality reduction[J]. Semisupervised Learning,2006:293-308.

Effects of PDEs for Modeling Parameters in RSVM

JIANG Shan-shan FAN Li-ya

(School of Mathematical Sciences,Liaocheng University,Liaocheng 252059,China)

Abstract The classification results of support vector machines (SVMs) were heavily limited by model parameters selection. Parameter selection problem was studied by many researchers, but at present, there is no effective method. The grid searching method was usually used, which is only an approximate method. This paper is devoted to research the influence of partial differential equations (PDEs) for parameters selection with regularized SVM (RSVM) and a noisy image (illumination, shade) classification task. Experiment results indicate that the evolution of PDEs can weaken the parameters influence, even no need to consider this problem.

Key words support vector machines; partial differential equations; parameters selection; evolutionary times; image classification