

## 特征空间中基于半遗传稀疏表示的图像识别

石林瑞, 黄祎婧, 符进武, 郭心悦, 范自柱

(华东交通大学理学院, 江西 南昌 330013)

**摘要:** 经典的稀疏表示分类 (SRC) 通常是基于求解  $L_1$  最小化问题的。SRC 在原始输入空间中求解  $L_0$  范数最小化问题, 无法很好地获取数据中的非线性信息。为了解决这一问题, 应用非线性映射将原始输入数据映射到一个新的高维特征空间, 并提出了一种新的基于  $L_0$  范数的表示方法。在所提方法中, 表示测试样本的字典包含两个部分: 第一部分固定在测试样本的近邻; 第二部分的训练样本通过半遗传算法 (SGA) 来选择, 利用表示误差确定第二部分的表示字典。在所提方法中, 如果训练样本和已确定的测试样本的近邻产生最小表示误差, 那么这些训练样本将被 SGA 确定为表示字典的第二部分。在一些常用的人脸数据集和一个手写体数据集上的实验表明, 所提方法能够获得更好的分类性能。

**关键词:** 稀疏表示; 图像识别; 特征空间;  $L_0$  范数; 遗传算法

**中图分类号:** TP391.4

**文献标识码:** A

**doi:** 10.11959/j.issn.2096-6652.202137

## Sparse representation for image recognition based on semi-genetic algorithm in feature space

SHI Linrui, HUANG Yijing, FU Jinwu, GUO Xinyue, FAN Zizhu

School of Science, East China Jiaotong University, Nanchang 330013, China

**Abstract:** The typical sparse representation for classification (SRC) is usually based on  $L_1$  minimization problem. Conceptually, SRC is essentially an  $L_0$  norm minimization problem solved in the original input space, which cannot capture well the nonlinear information within the data. In order to address this problem, a nonlinear mapping to map the original input data into a new high dimensional feature space was applied, and a new representation approach based on  $L_0$  norm was proposed. The representing dictionary used to represent the test sample contains two parts in the proposed approach. The first part is fixed to the neighbors of the test sample. The training samples of the second part is chosen by the variation of genetic algorithm (GA), i.e., the semi GA (SGA) algorithm, which exploits the representation error to determine the second part of the representing dictionary. In the approach, if the training samples combining the determined neighbors of the test sample yield the least representation error, these training samples are determined as the second part of the representing dictionary by SGA. Experiments on several popular face databases and one handwritten digit data set demonstrate that the proposed approach can achieve better classification performance.

**Key words:** sparse representation, image recognition, feature space,  $L_0$  norm, genetic algorithm

### 1 引言

基于  $L_1$  范数及其变体<sup>[1-4]</sup>的稀疏表示分类

(sparse representation for classification, SRC)<sup>[5-8]</sup>已被广泛应用于图像识别, 如人脸识别<sup>[9-10]</sup>、手写数字数据识别和其他应用<sup>[11-14]</sup>。从本质上讲, SRC

收稿日期: 2021-06-15; 修回日期: 2021-08-17

通信作者: 范自柱, zxfan3@163.com

基金项目: 国家自然科学基金资助项目 (No.61991401); 江西省自然科学基金资助项目 (No.20192ACBL20010)

**Foundation Items:** The National Natural Science Foundation of China (No.61991401), Jiangxi Provincial Natural Science Foundation (No.20192ACBL20010)

可以看作经典最近邻 (nearest neighbor, NN) 分类器的泛化。与使用一个子空间学习获得的特征提取器的经典 NN 分类器相比,如线性判别分析 (linear discriminative analysis, LDA)<sup>[15]</sup>, SRC 可以在分类精度方面获得更好的分类性能,特别是当它被应用于对有噪声或损坏的数据进行分类时。

SRC 算法的主要思想是对于每个测试样本或数据点, SRC 基于  $L_1$  范数最小化问题的解,使用训练样本来表示测试样本,然后利用表示结果将测试样本分类到表示误差最小的类<sup>[16]</sup>。需要注意的是,经典 SRC 算法的缺点之一是不能很好地应对训练样本数量相对较少的学习场景。基于  $L_1$  范数的经典 SRC 算法通常需要每类有足够的训练样本来表示测试样本<sup>[17]</sup>。因此,  $L_1$  范数最小化问题可以足够稀疏,从而获得理想的分类性能<sup>[18]</sup>。从本质上讲, SRC 算法是  $L_0$  范数最小化问题,可以产生最稀疏的解,这可能会带来更好的分类性能<sup>[5]</sup>。如果每一类的训练样本不足,基于  $L_1$  范数的经典 SRC 模型就不能保证获得稀疏解。因此,它可能无法达到理想的分类性能。在这种情况下,  $L_1$  范数最小化不再适用于求解 SRC 模型。如果采用  $L_0$  范数最小化来求解 SRC 模型,该模型仍然可以获得稀疏解,这在理论上比基于  $L_1$  范数的模型具有更好的分类性能。CF-SECL<sup>[19]</sup>算法通过稀疏编码器将测试样本映射到其所属的子空间中,然后将分类器应用到稀疏编码中,通过计算预测出测试样本的标签;基于深度稀疏表示的多模态信号融合分类<sup>[20]</sup>算法提出具有共享的全连接层的多模态编码器和解码器,对潜在空间特征进行训练,使其具有判别性,并适用于稀疏表示。基于核方法的稀疏表示解决了非线性分布问题,在图像分类任务中受到了广泛关注, K-LSDSR<sup>[21]</sup>算法充分研究了图像数据间的局域性构造问题,可有效地捕获嵌入判别信息,提高了分类准确率。

本文利用  $L_0$  范数最小化来建立一种新的 SRC 模型,以克服经典 SRC 模型的不足。经典的 SRC 模型可能存在另一个问题,即经典的 SRC 模型不能很好地处理不同类的样本分布在同一向量方向上的数据集<sup>[22]</sup>,即不同类别相同方向 (different class same direction, DCSD)。为了解决这个问题, Zhang L 等人<sup>[22]</sup>借用了基于核方法的思想,如支持向量机 (support vector machine, SVM)<sup>[23]</sup>。他们选择了一个合适的非线性映射,利用该映射将原始输入数据

映射到一个高维数据空间,即特征空间,这个特征空间被称为再现核 Hilbert 空间 (reproducing kernel Hilbert space, RKHS)<sup>[24]</sup>,然后在该特征空间中执行 SRC。这种方法被称为基于核的稀疏表示分类 (kernel based SRC, KSRC)<sup>[24-26]</sup>,理论上可以避免上述 DCSD 问题。此外,它可以很好地获取数据集的非线性信息,有助于对样本进行正确分类<sup>[26]</sup>。基于核的稀疏表示方法可被应用在许多领域中。例如, Song B 等人<sup>[27]</sup>利用核稀疏表示对遥感图像进行一类分类。Zhou Y 等人<sup>[28]</sup>提出了一种基于核稀疏表示 (kernel sparse representation, KSR) 的手势识别算法,该算法对人类手势较大的变化具有鲁棒性。Huang K K 等人<sup>[29]</sup>结合 SRC 和核判别分析 (kernel discriminant analysis, KDA) 算法开发了一种核扩展字典 (kernel extended dictionary, KED) 方法,该方法可以学习大量用于人脸识别的字典。Wang L 等人<sup>[30]</sup>采用核坐标下降法 (kernel coordinate descent, KCD) 求解 KSR,该方法可用于视觉跟踪任务。受核方法思想<sup>[31]</sup>的启发,本文提出的方法首先使用适当的非线性映射将原始输入数据转换到高维特征空间,然后利用  $L_0$  范数最小化实现 SRC 模型。因此,本文提出的方法可以有效地获取非线性信息,避免了之前工作<sup>[32]</sup>在理论上无法有效解决的 DCSD 问题。

$L_0$  范数最小化问题是 NP 难问题,没有直接的解析解。实际上,  $L_0$  范数最小化问题的解决方案是寻找最优训练样本,即表示字典<sup>[33]</sup>,通过字典来表示测试样本。使用梯度下降法等解析优化方法很难确定表示测试样本的表示字典<sup>[34]</sup>。原则上,通过一种随机优化方法可以有效地确定表示字典,即遗传算法 (genetic algorithm, GA)。遗传算法是解决 NP 难问题的一种常用方法,已广泛地应用于特征选择及其他应用<sup>[35-37]</sup>。传统的遗传算法没有利用先验信息,即样本或数据点之间的距离或相似性。将这些信息加入遗传算法中,可以在一定程度上提高其有效性和效率。为此,本文提出的方法的表示字典<sup>[38]</sup>分为两部分:第一部分使用先验信息 (如距离信息) 实现快速计算;第二部分是通过表征结果来确定的,借鉴了参考文献<sup>[39]</sup>的思想。

给定一个测试样本,它通常与其邻域共享相同的类标签,这些邻域在该测试样本的表示中起着关键作用。因此,它们自然构成了表示字典的第一部分<sup>[40]</sup>。欧氏距离是一种常用的、简单的确定邻域的

方法。本文应用欧氏距离来确定表示字典的第一部分。虽然欧氏距离可被简单度量,但不能保证使用这个距离就可以获得整个合适的表示字典。因此,需要在确定表示字典的第一部分后确定第二部分,故利用遗传算法选择一些不属于字典第一部分的训练样本。如果这些训练样本结合字典的第一部分表示在测试样本时产生最小的表示误差,则使用这些训练样本构建表示字典的第二部分。表示字典的一部分是固定的,另一部分是基于遗传算法确定的,因此将该方法中确定字典的过程称为半遗传算法(semi-genetic algorithm, SGA)。

综上,本文提出的方法包含以下步骤。第一步是指定一个非线性映射,将输入数据转换到高维特征空间;第二步是采用半遗传算法确定表示字典,然后用确定的字典来表示测试样本;最后,利用表示结果对测试样本进行分类。本文提出的方法被称为核遗传稀疏表示(kernel GA sparse representation, KGASR)算法,它有以下贡献。

首先,本文提出的方法是在特征空间中进行的,它可以有效地克服经典 SRC 方法不能很好地处理规模非常小的训练集,即训练样本数量非常少的问题。具体来说,KGASR 算法可被应用于经典 SRC 模型无法获得充分稀疏解的学习场景。此外,该方法可以有效地避免 DCSD 问题,而经典的 SRC 方法不能很好地解决 DCSD 问题。

其次,KGASR 算法是一种被普遍认为可以获取数据中非线性信息的核方法。与经典的 SRC 方法相比,其工作是寻找一个更强大的工具来寻找类之间的非线性边界。特别是,本文提出的方法非常适合高维、小规模的数据集。主要原因是,在本文提出的方法中,非线性数据被映射到特征空间中,在这个空间中类之间的边界趋向于线性。显然,这种边界很容易找到。

最后,笔者使用半遗传算法来确定表示字典。SGA 方案有两个优点,第一个优点是 SGA 使用欧氏距离来计算测试样本的邻域,该方案有效地利用了数据分布的先验信息来确定最适合表示测试样本的训练样本,进而有效地提高其表现性能;第二个优点是因为表示字典的第一部分是固定的,所以搜索空间显著缩小,从而提高了计算效率。

## 2 SRC 概述

经典 SRC 的基本思想最初来自信号处理领域。

2009 年, J. Wright 等人首次将 SRC 引入人脸识别领域,并引起了该领域的广泛关注<sup>[1,41]</sup>。下面简单介绍 SRC 的基本原理。

假设有  $N$  个训练样本  $X=[x_1, x_2, \dots, x_N] \in R^{d \times N}$  和一个测试样本  $y \in R^d$ 。经典的 SRC 模型可解决以下优化问题<sup>[1]</sup>。

$$\hat{\alpha}_1 = \arg \min \|\alpha\|_1, \text{ subject to } y = X\alpha \quad (1)$$

其中,  $\alpha=[\alpha_1, \alpha_2, \dots, \alpha_N]^T$ ,  $\alpha_i$  是第  $i$  个训练样本对应的表示系数,  $\|\alpha\|_1 = \sum_{i=1}^N |\alpha_i|$ 。当处理数据中存在小密度噪声时,式(1)可以修改为:

$$\hat{\alpha}_1 = \arg \min \|\alpha\|_1, \text{ subject to } \|y - X\alpha\|_2 \leq \delta \quad (2)$$

其中,  $\delta$  是一个小的正数。显然,上面的 SRC 模型是基于  $L_1$  范数最小化的。最小化问题通常可转为下面的最小化问题:

$$\hat{\alpha}_1 = \arg \min_{\alpha} (\|y - X\alpha\|_2^2 + \lambda \|\alpha\|_1) \quad (3)$$

其中,  $\lambda$  是平衡稀疏性和表示误差的正则化参数。当每类的训练样本足够多、解足够稀疏时,式(2)中的模型可以取得很好的学习效果。如果训练集规模较小,  $L_1$  范数最小化不能保证在此类数据集中表现良好。事实上, SRC 的目标是为测试样本找到最稀疏的表示。它本质上是基于  $L_0$  范数最小化的,其计算式如下:

$$\hat{\alpha}_1 = \arg \min \|\alpha\|_0, \text{ subject to } \|y - X\alpha\|_2 \leq \delta \quad (4)$$

其中,  $\|\alpha\|_0$  为系数向量中非零项的数量,即  $\alpha$  的  $L_0$  范数。理论上,当训练集规模较小时,式(4)可以很好地为测试样本产生最稀疏的表示。本文提出的 KGASR 算法基于式(4)对小规模高维训练集进行分类。

## 3 核遗传稀疏表示

### 3.1 核稀疏表示

核稀疏表示可以被看作一种基于核的方法,它采用非线性映射将原始输入数据映射到高维特征空间,其中内积可以通过核函数计算<sup>[42-43]</sup>。非线性映射是不需要明确表示的,在核方法中由特定核函数及其参数决定。核稀疏表示是在特征空间中进行的,本质上是在原始输入数据空间中实现的传统稀疏表示的基于核的非线性扩展。

首先采用非线性映射 ( $\phi: R^d \rightarrow F$ , 其中  $F$  是

特征空间) 将训练样本和测试样本映射到特征空间, 这样在这个特征空间中就获得了训练集和测试样本; 然后利用核函数计算样本向量  $\phi(X)=[\phi(x_1), \phi(x_2), \dots, \phi(x_N)] \in R^{D \times N}$  和测试样本  $\phi(y) \in R^D$  之间的内积, 这个方法被称为核技巧。核函数有 3 种基本类型: 线性核函数、多项式核函数和高斯核函数。使用高斯核函数在理论上可以将原始输入数据映射到非常高甚至无限维的特征空间, 并且其核参数易于调整, 因此本文采用目前广泛使用的高斯核函数, 其表达式如下:

$$k(x_i, x_j) = \exp \frac{-\|x_i - x_j\|^2}{2\sigma^2} \quad (5)$$

其中,  $\sigma$  是核函数的高斯核宽度参数。

理论上来说, 特征空间中的测试样本可以由映射的训练样本稀疏表示。与输入空间中的经典 SRC 算法一样, 特征空间中基于  $L_1$  范数的核稀疏表示模型定义为:

$$\hat{\beta}_1 = \arg \min \|\beta\|_1, \text{ subject to } \|\phi(y) - \phi(X)\beta\|_2 \leq \xi \quad (6)$$

其中,  $\xi$  是一个小的正数。同样, 当训练集规模较小且数据维数很高时, 基于  $L_1$  范数的 KSR 无法保证获得足够稀疏的解。因此, 可以利用  $L_0$  范数最小化来构建新的 KSR 模型, 如下所示:

$$\hat{\beta}_1 = \arg \min \|\beta\|_0, \text{ subject to } \|\phi(y) - \phi(X)\beta\|_2 \leq \xi \quad (7)$$

类似于在输入空间中基于  $L_0$  范数的经典 SRC 分类算法, 在理论上可获得非常理想的分类性能, 在特征空间中基于  $L_0$  范数的 KSR 算法也可以得到

最稀疏解。需要注意的是, 基于  $L_0$  范数的 KSR 模型也是一个 NP 难问题, 可以采用一种改进的遗传算法来解决它。后文将详细描述这种改进的遗传算法, 其被称为半遗传算法。图 1 给出了 SRC 方法与 KSRC 方法的基本原理。

### 3.2 本文提出的方法

#### 3.2.1 确定表示字典

本文提出的 SGA 在特征空间中进行。给定一个测试样本  $\phi(y)$ , 在特征空间中计算它与其他样本  $\phi(x)$  之间的欧氏距离, 计算式如下:

$$\text{ed} = \|\phi(y) - \phi(x)\|_2^2 = (\phi(y) - \phi(x))^\top (\phi(y) - \phi(x)) = \phi^\top(y)\phi(y) - \phi^\top(y)\phi(x) - \phi^\top(x)\phi(y) + \phi^\top(x)\phi(x) \quad (8)$$

通过使用核技巧, 有:

$$\text{ed} = k(y, y) - 2k(x, y) + k(x, x) \quad (9)$$

因此, 使用欧氏距离来确定近邻的数量 (如  $K$  个近邻), 并将其作为表示字典的第一部分, 表示为  $D_1 = \{\phi(x'_1), \phi(x'_2), \dots, \phi(x'_K)\}$ 。

在确定了测试样本的  $K$  个近邻后, 需要使用 SGA 确定整个表示字典。在 SGA 中, 表示字典的第一部分是固定的。SGA 将表示误差最小的样本作为表示字典的第二部分。在描述 SGA 过程之前, 先给出一些关于遗传算法的基本概念 (详情请见参考文献[44])。第一个概念是个体或状态, 通常表示为 0 和 1 的串。个体的每个元素都与一个训练样本相关联。许多个体生成种群, 这是遗传算法术语中的第二个基本概念。每个个体通常通过适应度函数进行评估。

在本文中, 每个个体都被视为一个向量, 其中每个元素都是 0 或 1。向量的长度等于训练样本的

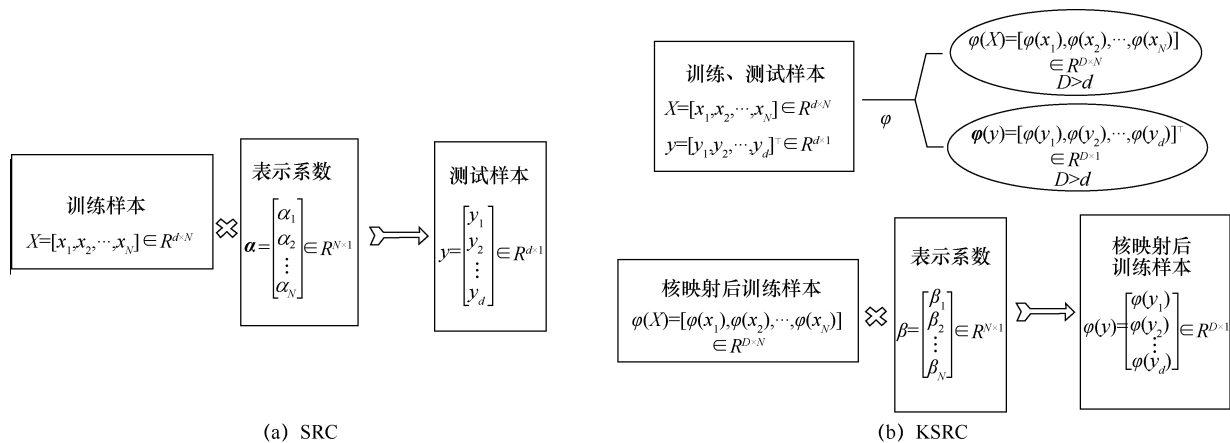


图1 SRC方法与KSRC方法的基本原理

数量。如果个体的元素为 1, 则其对应的训练样本属于表示字典。首先, 将与测试样本相关联的元素在个体中设置为 1, 并且它们在整个 SGA 过程中是固定的。这些元素产生了个体的第一部分。虽然有些训练样本没有被包含在表示字典的第一部分, 但它们在表示测试样本方面可能起着非常重要的作用。因此, 它们也应该被确定并放入表示字典中。换句话说, 它们对应的元素在个体的第二部分是 1。确定个体第二部分元素的第一步是将这些元素随机设置为 0 或 1; 第二步是根据适应度函数对个体的第二部分进行多代选择、交叉和变异操作。如果多个训练样本结合确定的近邻产生最小的表示误差, 即最小的适应度, 则它们的关联元素在个体的第二部分中为 1。

SGA 中的适应度函数被定义为使用个体对应的训练样本来表示测试样本产生的表示误差。假设个体中 1 的数量为  $M$ , 有  $M$  个训练样本 (包括测试样本的  $K$  个最近邻) 与这些 1 相关联。个体的适应度计算如下:

$$f = \left\| \phi(y) - \sum_{i=1}^M \beta_i \phi(x'_i) \right\|_2^2 \quad (10)$$

通过使用核技巧, 有:

$$f = k(y, y) - 2 \sum_{i=1}^M \beta_i k(x'_i, y) + \sum_{i=1}^M \sum_{j=1}^M \beta_i \beta_j k(x'_i, x'_j) \quad (11)$$

其中,  $\beta_i$  ( $i=1, 2, \dots, M$ ) 是使用训练样本  $I = [\phi(x'_1), \phi(x'_2), \dots, \phi(x'_M)]$  表示测试样本的系数。

### 3.2.2 表示和分类

确定测试样本  $\phi(y)$  的表示字典后, 使用该字典对  $\phi(y)$  进行表示和分类。假设已确定的表示字典来自  $t$  个类, 即  $C = \{c'_1, c'_2, \dots, c'_t\}$ 。分别利用表示字典中属于每个类的样本来表示测试样本, 并计算表示误差, 其计算式如下:

$$r_j(\phi(y)) = \left\| \phi(y) - \sum_{i=1}^{n_j} \beta_{c'_j, i}^* \phi(x_{c'_j, i}^*) \right\|_2^2 \quad (12)$$

同样, 使用核技巧, 可得

$$r_j(\phi(y)) = k(y, y) - 2 \sum_{i=1}^{n_j} \beta_{c'_j, i}^* k(x_{c'_j, i}^*, y) + \sum_{i=1}^{n_j} \sum_{j=1}^{n_j} \beta_{c'_j, i}^* \beta_{c'_j, j}^* k(x_{c'_j, i}^*, x_{c'_j, j}^*) \quad (13)$$

其中,  $n_j$  是字典中属于类  $c'_j$  ( $j=1, 2, \dots, t$ ) 的样本数,

$\phi(x_{c'_j, i}^*)$  是特征空间中类  $c'_j$  的第  $i$  个训练样本,  $\beta_{c'_j, i}^*$  是当应用表示字典表示测试样本  $\phi(y)$  时, 样本  $\phi(x_{c'_j, i}^*)$  的表示系数。本文使用了高斯核函数, 因此式 (13) 中的  $k(y, y)=1$ 。最后, 将测试样本分配到导致表示误差最小的类中:

$$\text{id}(y) = \arg \min_j r_j(\phi(y)) \quad (j=1, 2, \dots, t) \quad (14)$$

其中,  $\text{id}$  表示测试样本  $y$  的类别。

KGASR 算法如下。

#### 算法 1 核遗传算法稀疏表示 (KGASR)

输入 训练样本  $X = [x_1, x_2, \dots, x_N] \in R^{m \times N}$ , 测试样本  $y \in R^m$ , 进化代数  $P$ 、个体数量  $Q$ 。

步骤 1: 利用特征空间中的欧氏距离确定测试样本的  $K$  个近邻。

步骤 2: 生成包含  $Q$  个个体的种群, 其中与确定的近邻相关的元素为 1, 其他元素随机设置为 1 或 0。

步骤 3: For  $i = 1:Q$

- 1) 使用式 (12) 计算种群中每个个体的适应度。
- 2) 与确定的近邻相关联的元素在每个个体中始终为 1。对其他元素执行经典遗传算法的操作以产生新的种群。
- 3) 如果当前代个体的最小适应度没有减少, 则转到步骤 5。

步骤 4: 确定导致最小适应度的表示字典。

步骤 5: 使用式 (13) 和式 (14) 对测试样本进行分类。

输出 测试样本的标签。

### 3.3 KGASR 算法分析

需要注意的是, 如果使用传统的遗传算法来确定表示字典, 其搜索空间规模为  $2^N$ , 其中  $N$  是训练样本的总数。测试样本的  $K$  个确定的近邻对应的元素固定为 1, 因此提出的算法的搜索空间规模为  $2^{N-K}$ , KGASR 算法在理论上比传统的遗传算法更快。KGASR 算法的训练时间适中, 测试时间快。在 KGASR 算法中, 每次表示和分类一个测试样本。学习一个测试样本的过程与学习另一个测试样本的过程是独立的。也就是说, KGASR 算法的学习过程是并行的。如果利用并行程序来实现 KGASR 算法, 其计算效率将大大提高。

此外, KGASR 算法选择训练样本来生成具有非常强大的表示能力的表示字典来表示测试样本。

所选的表示字典包含测试样本的近邻。因此,确定的字典可以获取测试样本的局部结构。众所周知,局部性通常会导致稀疏性,这个优点有助于 KGASR 算法产生绝对稀疏的解决方案。也就是说,在表示测试样本时,在表示系数向量中得到了多个零。因为除表示字典外的训练样本的系数都是零,所以该方案可以在很大程度上减少噪声或那些表示能力较差的训练样本的影响。这是 KGASR 算法能够达到良好分类性能的原因之一。

近年来,深度学习在机器学习领域引起了很高的关注。深度学习可以在训练集规模非常大的情况下取得很好的分类性能。然而,笔者的工作侧重于小规模的数据集,并且数据维度非常高。如果在笔者的工作中使用深度学习,可能无法获得很好的分类性能。

KGASR 算法与相关深度学习方法有学习特征并通过学习到的特征对测试样本进行稀疏表示的共性<sup>[20]</sup>。深度学习方法可以在大规模的训练集下取得很好的分类性能,其将模型的上一层输出作为下一层的输入,实现了对输入数据信息的分级表达,最终达到特征学习的目的。KGASR 算法侧重于小规模且数据维度高的数据集,且应用非线性映射很好地捕捉了数据的非线性信息,有效地处理了不同类样本在同一向量方向上的数据分布问题。

## 4 实验

笔者在 4 个流行的数据集上进行了实验,以证明提出的方法可以在分类精度方面达到理想的性能。数据集包括 CMU-PIE、EYB (Extended YaleB)、AR 和 MNIST 数据集。前 3 个数据集是人脸图像数据集,第 4 个数据集是手写数字数据集。在本文的实验部分,笔者比较了一些经典的基于表示的分类方法,如协作表示分类 (collaborative representation for classification, CRC)<sup>[16]</sup>、SRC<sup>[1]</sup>、KSRC<sup>[22]</sup>、核坐标下降<sup>[30]</sup>、扩展 SRC (extended SRC, ESRC)<sup>[17]</sup>、两阶段测试样本稀疏表示 (two-phase test sample sparse representation, TPTSR)<sup>[39]</sup>和通过结构和非凸约束的核群稀疏表示方法 (KGSRSN)<sup>[45]</sup>。本文提出的算法可被看作字典学习方法的一种。因此,将提出的方法与一些典型的字典学习方法进行了比较,它们分别是核扩展字典 (KED) 算法<sup>[29]</sup>、局部约束和标签嵌入 (locality-constrained and label embedding, LCLE) 字典学习方法<sup>[46]</sup>、K 奇异值分解 (K-singular value decomposition, KSVD) 算法<sup>[47]</sup>

和鉴别 Fisher 嵌入字典学习算法 (DFEDL)<sup>[48]</sup>。

随机选择训练样本可以保证分类性能与训练集的任何特殊选择无关。与之前的许多工作相同,笔者随机选择了一部分样本进行训练,其余样本用于在每个数据集进行测试。在本文实验中,KSRC 算法利用了高斯核函数,其核函数参数  $\sigma$  (即高斯核宽度) 设置为  $r \times d_i$ , 其中,  $d_i$  表示原始输入空间中训练样本的平均欧氏距离,即  $d_i = \frac{1}{N^2} \sum_{i,j} (x_i - x_j)^2$ ,  $N$  表示

训练样本的数量。在所有实验中,KSRC 中的参数  $r$  设置为 0.001。在经典的 SRC 算法中,正则化参数设置为 0.01,在 ESRC 算法中也是如此。在 KGASR 算法中,表示字典中的样本数与训练样本数之比设为 0.1,进化代数  $P$  设为 50,每个种群中的个体数  $Q$  设为 20。

### 4.1 在 CMU-PIE 数据集上的实验

第一个实验是在包含 68 个人的图像的卡内基梅隆大学 (CMU) -PIE 人脸数据集上进行的。每个人都有在 13 种不同姿势和 43 种不同照明条件以及 4 种不同表情下获取的图像<sup>[49]</sup>,图 2 为该数据集的一些示例。这个实验选取前 30 个人的图像,所有图像都被裁剪并调整为  $32 \times 32$  像素的分辨率。对于每个人,随机选择  $M$  ( $M=5, 10$  和  $15$ ) 张图像进行训练,其余图像用于测试。

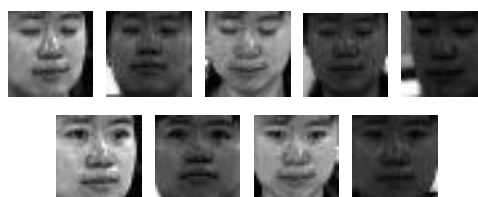


图2 CMU-PIE 人脸数据集的一些图像样本

为了节省计算成本,在所有算法开始之前使用主成分分析 (principal component analysis, PCA) 法将人脸图像数据降维至 100 维。在 KGASR 和 KCD 算法中,从  $\{0.1, 0.2, \dots, 1\}$  和  $\{1.1, 1.2, \dots, 10\}$  两个集中选择高斯核宽度参数  $r$ 。如果这两组中的一个值导致最佳分类性能,则将  $r$  设置为该值。本文的其他实验使用相同的方案来选择 KGASR 和 KCD 算法中的参数  $r$ 。 $r$  在 KGASR 和 KCD 中设置为 3。在 TPTSR 算法中,测试样本的邻居数设置为 35,随机运行每个算法 10 次。

表 1 给出了分类结果,包括每种算法的平均分类准确率和标准偏差。其中,  $M$  表示每个人的脸

图像数量。从表 1 可以观察到, KGASR 算法的分类精度明显优于其他所有比较算法。此外, 当训练集规模较小时, KGASR 算法的分类性能有较大的提升。

表 1 在 CMU-PIE 数据集上的分类精度

| 算法    | $M=5$       | $M=10$      | $M=15$      |
|-------|-------------|-------------|-------------|
| CRC   | 62.09%±1.76 | 75.75%±1.71 | 82.65%±1.36 |
| TPTSR | 59.43%±2.02 | 75.05%±1.59 | 82.91%±0.81 |
| SRC   | 61.78%±1.84 | 77.13%±1.77 | 84.84%±0.92 |
| KSRC  | 66.63%±1.61 | 80.43%±1.47 | 86.33%±0.61 |
| KCD   | 70.49%±1.98 | 82.54%±1.65 | 87.99%±1.04 |
| KED   | 53.96%±1.94 | 63.95%±1.35 | 69.82%±1.50 |
| KSVD  | 68.73%±1.95 | 78.67%±1.81 | 81.66%±1.34 |
| LCLE  | 70.25%±1.93 | 81.26%±0.89 | 84.35%±1.07 |
| ESRC  | 68.60%±2.08 | 82.45%±1.75 | 88.46%±0.93 |
| KGSRN | 46.55%±0.54 | 66.06%±0.52 | 77.20%±0.35 |
| DFEDL | 62.85%±0.71 | 78.58%±0.68 | 86.35%±0.39 |
| KGASR | 71.55%±1.65 | 85.25%±0.99 | 91.05%±0.22 |

#### 4.2 在 EYB 数据集上的实验

EYB 数据集包含 38 个受试者的 2 414 张正面图像。所有图像均是在不同照明条件下拍摄的<sup>[1]</sup>, 图 3 为该数据集的一些示例。所有图像被裁剪并调整为 40×35 像素的分辨率。对于每个受试者, 随机选择  $M$  ( $M=4, 5, 6, 7$ ) 张图像进行训练, 其余图像用于测试。

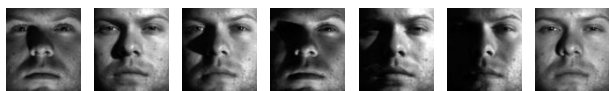


图 3 EYB 数据集的一些图像样本

使用 PCA 将数据维数降至 100 维。KGASR 算法的高斯核宽度设置为  $4 \times d_i$ , KCD 算法中的参数  $r$  设置为 1。在 TPTSR 算法中, 测试样本的近邻数设置为 50, 从而使该算法的分类效果最好, 随机运行每个算法 10 次。

表 2 给出了该数据集的分类结果, 从表 2 可以看出, 本文提出的算法性能优于其他所有比较算法。与 SRC 算法相比, KGASR 算法再次证明了当训练集规模较小时, 它可以实现更大的分类性能提升。当  $M=4$  时, KGASR 算法的分类准确率为 66.25%, 比 SRC 算法高 4.66%。而当  $M=7$  时, KGASR 算法的分类准确率比 SRC 算法高 1.27%。

实验表明, 本文提出的方法比 SRC 算法更适合处理小规模训练集。

表 2 在 EYB 数据集上的分类精度

| 算法    | $M=4$       | $M=5$       | $M=6$       | $M=7$       |
|-------|-------------|-------------|-------------|-------------|
| CRC   | 60.05%±1.22 | 64.48%±1.68 | 68.85%±1.3  | 71.59%±1.96 |
| TPTSR | 60.16%±1.18 | 65.22%±1.76 | 70.34%±1.49 | 74.47%±1.26 |
| SRC   | 61.59%±1.64 | 67.01%±2.01 | 71.92%±1.56 | 75.59%±1.45 |
| KSRC  | 62.25%±0.93 | 66.97%±1.8  | 70.81%±1.3  | 74.18%±1.37 |
| KCD   | 63.16%±1.26 | 67.64%±1.95 | 72.09%±1.49 | 75.42%±1.18 |
| KED   | 62.4%±3.4   | 67.35%±2.75 | 71.28%±2.41 | 74.9%±2.49  |
| KSVD  | 63.3%±1.39  | 67.65%±1.87 | 69.05%±1.23 | 71.55%±1.56 |
| LCLE  | 63.72%±1.17 | 67.41%±1.92 | 70.63%±0.98 | 73.21%±1.14 |
| ESRC  | 62.84%±1.47 | 67.50%±1.91 | 72.72%±1.24 | 76%±1.3     |
| KGSRN | 43.8%±0.35  | 49.96%±0.47 | 54.77%±0.57 | 60.51%±0.46 |
| DFEDL | 31.52%±0.31 | 61.59%±0.54 | 66.29%±0.46 | 70.34%±0.47 |
| KGASR | 66.25%±1.44 | 71.31%±1.69 | 75.7%±1.14  | 76.86%±1.03 |

#### 4.3 在 AR 数据集上的实验

AR 人脸数据集包含 126 个人 (126 个类别) 的图像, 每个人包含不同面部表情、光照条件和遮挡的人脸正面视图<sup>[50-51]</sup>, 图 4 为该数据集的一些示例。本文使用了 120 个人的人脸图像, 每个人有 26 张图像。所有图像首先被更改为灰色图像, 然后被裁剪并调整大小为 50×40 像素。对每个人选择的训练图像数量为 3 和 4, 它们构成了训练数据的两个子集。



图 4 AR 数据集的一些图像样本

与前两个实验相同, 在所有方法实现之前, 使用 PCA 将人脸图像数据维数降至 120 维。在 KGASR 算法中, 高斯核宽度设置为  $5 \times d_i$ , KCD 算法中的参数  $r$  设置为 1.5。对于 TPTSR 算法, 测试样本的近邻数设置为 30, 随机运行每个算法 10 次。

表 3 给出了该数据集的分类结果。从表 3 可以看出, KGASR 算法优于其他比较先进的分类算法。KGASR 算法在本实验中的分类性能与前两次实验一致。也就是说, 本文提出的方法在分类准确率方面达到了稳定且理想的分类性能。

#### 4.4 在 MNIST 数据集上的实验

第 4 个实验是在 MNIST 数据集上进行的, 该

数据集是一个手写数字数据集,共有 10 个类,即 0、1、2、 $\cdots$ 、9 每个数字产生一个类<sup>[52]</sup>。在每个类中,训练集包含 6 000 个图像样本,测试集包含 1 000 个图像样本。每张图片的像素大小为  $28 \times 28$ 。图 5 为该数据集的一些示例。在这个实验中,从每个类的训练集中随机选择样本,这些样本被用作训练样本。每个类别选择的训练样本数分别为 5、10 和 15。因此,获得了训练数据的 3 个子集。对于每个子集,从每个类的测试集中随机选择样本作为测试样本,为每个类别选择的测试样本数为 300。因此,生成了 3 个测试子集,分别与 3 个训练数据子集相关联。每个测试子集包含 3 000 个测试样本。

表 3 在 AR 数据集上的分类精度

| 算法    | $M=3$              | $M=4$              |
|-------|--------------------|--------------------|
| CRC   | $78.86\% \pm 1.74$ | $84.07\% \pm 0.97$ |
| TPTSR | $78.75\% \pm 1.38$ | $84.06\% \pm 0.88$ |
| SRC   | $78.39\% \pm 1.31$ | $84.19\% \pm 0.7$  |
| KSRC  | $79.4\% \pm 1.33$  | $84.69\% \pm 0.83$ |
| KCD   | $80.97\% \pm 1.52$ | $86.14\% \pm 0.75$ |
| KED   | $82.54\% \pm 1.19$ | $87.31\% \pm 0.91$ |
| KSVD  | $75.24\% \pm 1.77$ | $81.94\% \pm 0.91$ |
| LCLE  | $78.9\% \pm 1.54$  | $82.86\% \pm 0.96$ |
| ESRC  | $81.25\% \pm 1.22$ | $86.46\% \pm 1.23$ |
| KGSRN | $64.5\% \pm 0.47$  | $72.49\% \pm 0.35$ |
| DFEDL | $78.59\% \pm 0.59$ | $84.37\% \pm 0.27$ |
| KGASR | $82.99\% \pm 1.45$ | $87.47\% \pm 0.82$ |

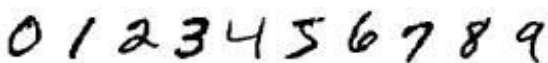


图 5 MNIST 数据集的一些图像样本

类似地,使用 PCA 将人脸图像数据维数减少到 40 维。KGASR 算法中的高斯核宽度设置为  $d_i$ 。在 TPTSR 算法中,测试样本的近邻数设置为 15。此外,随机运行每个算法 10 次。

图 6 给出了该数据集的分类结果。从图 6 可以看出,与 SRC、KSRC 和 ESRC 算法相比,KGASR 算法的分类精度高了近 10%。在该数据集中,KGSRN 算法的分类效果稍高于 KGASR 算法,但综合比较 4 个数据集的分类性能,KGASR 的分类性能更稳定、更鲁棒。

通过以上实验可知,本文提出的算法达到了理想的分类性能。其性能较好的主要原因是可以有效

地克服经典 SRC 和 KSRC 算法的缺点。具体来说,KGASR 算法可以有效地避免 DCSD 问题。此外,它使用半遗传算法确定合适的表示字典,从而很好地处理小规模和高维数据集,即欠采样数据集<sup>[17]</sup>,因此,在处理欠采样数据集时,ESRC 算法可能会被 KGASR 算法取代。

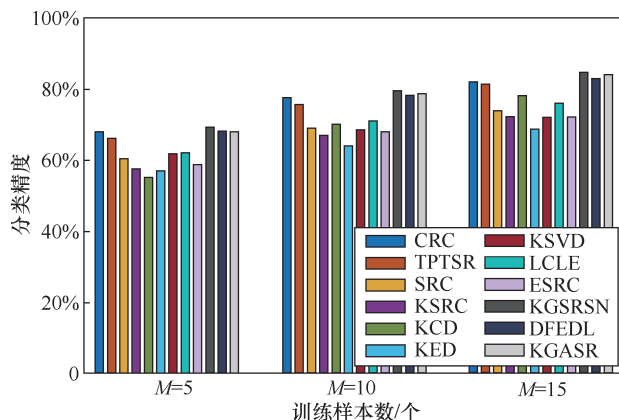


图 6 在 MNIST 数据集上的分类精度

#### 4.5 收敛性分析

KGASR 算法中表示字典的确定主要基于全局优化能力强的遗传算法。特征空间中的 SGA 是 KGASR 算法中的关键过程。实际上,SGA 是一种具有良好的收敛性的遗传算法,通常用于解决全局优化问题。接下来通过实验证明 SGA 在 4 个数据集上的收敛性。

在每个数据集中,随机选择一个测试样本,并根据式(12)中定义的适应度函数,使用 SGA 来确定表示字典。在 CMU-PIE、EYB 和 AR 数据集中,进化代数  $P$  设置为 5 000,个体数量  $Q$  设置为 5 000。在 MNIST 数据集中, $P$  设为 500, $Q$  设为 5 000。SGA 中的其他参数与前 4 次实验相同。图 7 显示了不同代在 4 个数据集上的适应度值。从图 7 可以看出,KGASR 算法实现了适应度函数的全局最小化。对于 CMU-PIE 数据集,在 1 200 代左右时适应度值下降到最小;对于 EYB 数据集,在 2 200 代左右时适应度值下降到最小;对于 AR 数据集,代数为 5 000 左右时降到最低;对于 MNIST 数据集,代数为 1 500 左右(MNIST 数据集中个体数  $Q$  为 500)时适应度值下降到最小。KGASR 算法在 4 个数据集上都是收敛的。因此,可以得出结论:KGASR 算法具有良好的收敛性。此外,该实验表明,如果设置更大的  $P$  和  $Q$ ,KGASR 算法可以获得更好的分类性能。

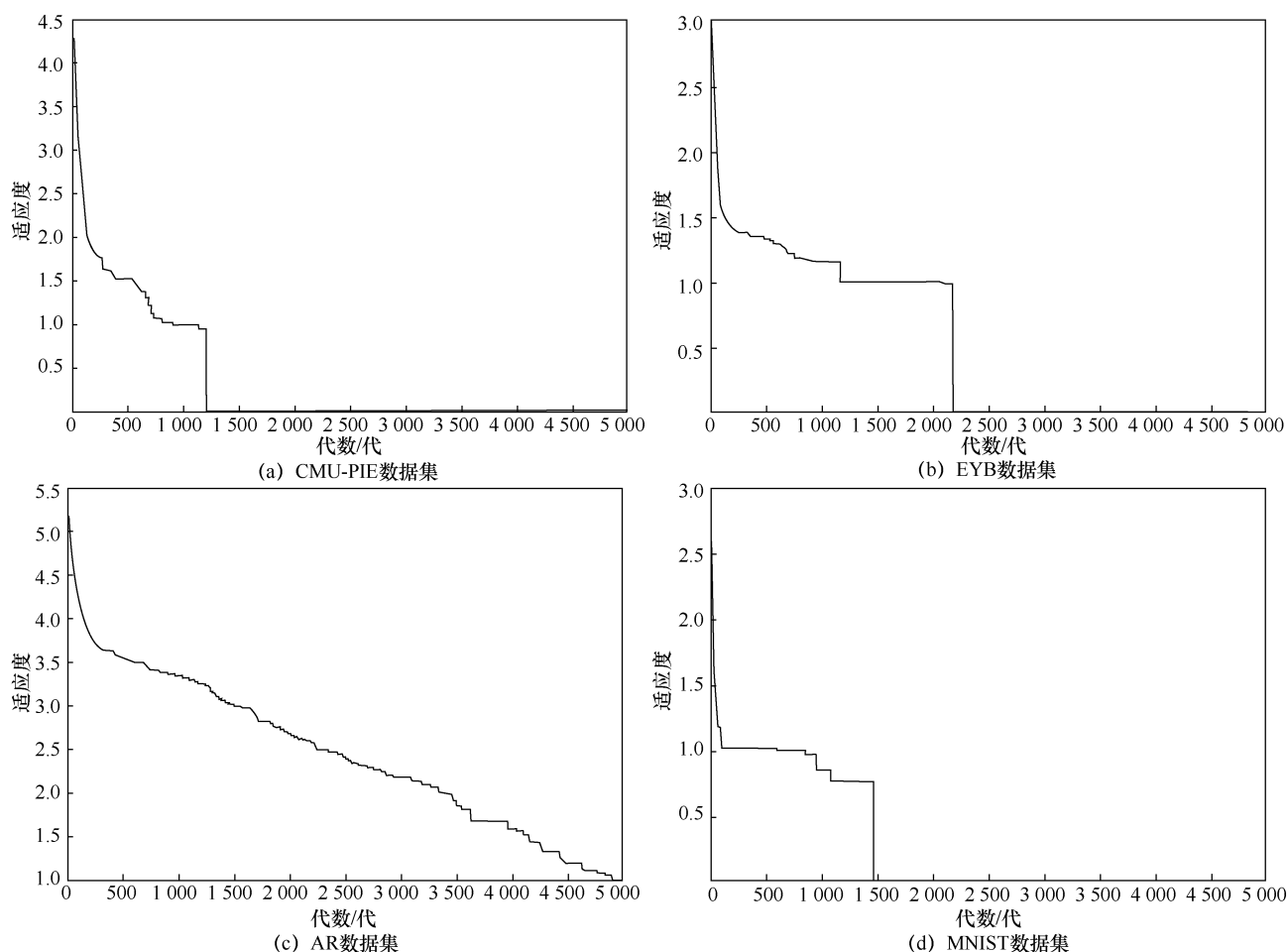


图7 不同代在CMU-PIE、EYB、AR和MNIST数据集上的适应度值

## 5 结束语

本文从一个新的角度研究了经典SRC算法。也就是说,经典SRC算法可以通过最小化特征空间中的 $L_0$ 范数来改进。为了有效地解决 $L_0$ 范数最小化问题,引入一种半遗传算法来确定一个具有强大表示和分类测试样本能力的表示字典。本文提出的KGASR算法在理论上克服了现有的基于表示分类的算法的不足,特别是基于 $L_1$ 范数的表示方法,如经典SRC、KSRC和ESRC算法。KGASR算法适用于处理小规模和高维数据集,还可以有效地避免DCSD问题。KGASR算法适用于许多应用,如图像分类和特征选择。实验证明,与比较算法相比,KGASR算法的分类精度最好。

## 参考文献:

- [1] SHEKHAR S, PATEL V M, NASRABADI N M, et al. Joint sparse representation for robust multimodal biometrics recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 36(1): 113-126.
- [2] YANG M, ZHANG L, FENG X, et al. Sparse representation based fisher discrimination dictionary learning for image classification[J]. International Journal of Computer Vision, 2014, 109(3): 209-232.
- [3] WANG J, LU C, WANG M, et al. Robust face recognition via adaptive sparse representation[J]. IEEE Transactions on Cybernetics, 2014, 44(12): 2368-2378.
- [4] YANG J, LUO L, QIAN J, et al. Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(1): 156-171.
- [5] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31(2): 210-227.
- [6] ZHANG Z, XU Y, YANG J, et al. A survey of sparse representation: algorithms and applications[J]. IEEE. Translations and Content Mining, 2015, 3: 490-530.
- [7] SHU T, ZHANG B, TANG Y Y. Sparse supervised representation-based classifier for uncontrolled and imbalanced classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 31(8): 2847-2856.
- [8] ABAVISANI M, PATEL V M. Deep multimodal sparse representation-based classification[C]//Proceedings of the IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2020:

- 773-777.
- [9] YANG J, CHU D, ZHANG L, et al. Sparse representation classifier steered discriminative projection with applications to face recognition[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, 24(7): 1023-1035.
  - [10] QIAN J, LUO L, YANG J, et al. Robust nuclear norm regularized regression for face recognition with occlusion[J]. *Pattern Recognition*, 2015, 48(10): 3145-3159.
  - [11] ZHANG J, ZHAO D, GAO W. Group-based sparse representation for image restoration[J]. *IEEE Transactions on Image Processing*, 2014, 23(8): 3336-3351.
  - [12] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861-2873.
  - [13] HONG Z, MEI X, PROKHOROV D, et al. Tracking via robust multi-task multi-view joint sparse representation[C]//*Proceedings of the IEEE International Conference on Computer Vision*. Piscataway: IEEE Press, 2013: 649-656.
  - [14] WANG S J, YANG J, SUN M F, et al. Sparse tensor discriminant color space for face verification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, 23(6): 876-888.
  - [15] BELHUMEUR P N, HESPAHIA J P, KRIEGMAN D J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711-720.
  - [16] ZHANG L, YANG M, FENG X. Sparse representation or collaborative representation: which helps face recognition?[C]//*Proceedings of the International Conference on Computer Vision*. Piscataway: IEEE Press, 2011: 471-478.
  - [17] DENG W, HU J, GUO J. Extended SRC: undersampled face recognition via intraclass variant dictionary[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(9): 1864-1870.
  - [18] SHEN C, CHEN L, DONG Y, et al. Sparse representation classification beyond  $L_1$  minimization and the subspace assumption[J]. *IEEE Transactions on Information Theory*, 2020, 66(8): 5061-5071.
  - [19] YANG C, WANG W, FENG X, et al. Classification-friendly sparse encoder and classifier learning[J]. *IEEE Access*, 2020, 8: 54494-54505.
  - [20] ABAVISANI M, PATEL V M. Deep multimodal sparse representation-based classification[C]//*Proceedings of the IEEE International Conference on Image Processing (ICIP)*. [S.l.:s.n.], 2020: 773-777.
  - [21] BENUWA B B, GHANSAH B, ANSAH E K. Kernel based locality-sensitive discriminative sparse representation for face recognition[J]. *Scientific African*, 2020, 7: e00249.
  - [22] ZHANG L, ZHOU W D, CHANG P C, et al. Kernel sparse representation-based classifier[J]. *IEEE Transactions on Signal Processing*, 2011, 60(4): 1684-1695.
  - [23] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2(3): 1-27.
  - [24] ZHANG G, SUN H, XIA G, et al. Multiple kernel sparse representation-based orthogonal discriminative projection and its cost-sensitive extension[J]. *IEEE Transactions on Image Processing*, 2016, 25(9): 4271-4285.
  - [25] JIAN M, JUNG C. Class-discriminative kernel sparse representation-based classification using multi-objective optimization[J]. *IEEE Transactions on Signal Processing*, 2013, 61(18): 4416-4427.
  - [26] GAO S, TSANG I W H, CHIA L T. Sparse representation with kernels[J]. *IEEE Transactions on Image Processing*, 2012, 22(2): 423-434.
  - [27] SONG B, LI P, LI J, et al. One-class classification of remote sensing images using kernel sparse representation[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(4): 1613-1623.
  - [28] ZHOU Y, LIU K, CARRILLO R E, et al. Kernel-based sparse representation for gesture recognition[J]. *Pattern Recognition*, 2013, 46(12): 3208-3222.
  - [29] HUANG K K, DAI D Q, REN C X, et al. Learning kernel extended dictionary for face recognition[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 28(5): 1082-1094.
  - [30] WANG L, YAN H, LV K, et al. Visual tracking via kernel sparse representation with multikernel fusion[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, 24(7): 1132-1141.
  - [31] LIU J, WU Z, XIAO L, et al. Learning multiple parameters for kernel collaborative representation classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(12): 5068-5078.
  - [32] FAN Z, NI M, ZHU Q, et al.  $L_0$ -norm sparse representation based on modified genetic algorithm for face recognition[J]. *Journal of Visual Communication and Image Representation*, 2015, 28: 15-20.
  - [33] LI Z, ZHANG Z, QIN J, et al. Discriminative fisher embedding dictionary learning algorithm for object recognition[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(3): 786-800.
  - [34] BOYD S, BOYD S P, VANDENBERGHE L. *Convex optimization*[M]. Cambridge: Cambridge University Press, 2004.
  - [35] ESTÉVEZ P A, TESMER M, PEREZ C A, et al. Normalized mutual information feature selection[J]. *IEEE Transactions on Neural Networks*, 2009, 20(2): 189-201.
  - [36] 杜宏庆, 陈德旺, 黄允许, 等. 基于改进遗传算法与支持度的模糊系统优化建模方法[J]. *智能科学与技术学报*, 2020, 2(2): 179-185.
  - [37] 辛峻峰, 张永波, 伯佳更, 等. 基于数据驱动的遗传算法的无人艇路径规划研究[J]. *智能科学与技术学报*, 2019, 1(2): 171-180.
  - [38] XIN J F, ZHANG Y B, BO J G, et al. Study on path planning of unmanned surface vessel based on data-driven genetic algorithm[J]. *Chinese Journal of Intelligent Science and Technology*, 2019, 1(2): 171-180.
  - [39] SUN Y, ZHANG Z, JIANG W, et al. Discriminative local sparse representation by robust adaptive dictionary pair learning[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(10): 179-185.

4303-4317.

- [39] XU Y, ZHANG D, YANG J, et al. A two-phase test sample sparse representation method for use with face recognition[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(9): 1255-1262.
- [40] HAN N, WU J, FANG X, et al. Projective double reconstructions based dictionary learning algorithm for cross-domain recognition[J]. IEEE Transactions on Image Processing, 2020, 29: 9220-9233.
- [41] YANG J, ZHANG L, XU Y, et al. Beyond sparsity: the role of  $L_1$ -optimizer in pattern classification[J]. Pattern Recognition, 2012, 45(3): 1104-1118.
- [42] CAI D, HE X, HAN J. Speed up kernel discriminant analysis[J]. The VLDB Journal, 2011, 20(1): 21-33.
- [43] XU Y, ZHANG D, JIN Z, et al. A fast kernel-based nonlinear discriminant analysis for multi-class problems[J]. Pattern Recognition, 2006, 39(6): 1026-1033.
- [44] RUSSELL S J, STUART J. Norvig[J]. Artificial Intelligence: A Modern Approach, 2003: 111-114.
- [45] ZHENG J, QIU H, SHENG W, et al. Kernel group sparse representation classifier via structural and non-convex constraints[J]. Neurocomputing, 2018, 296: 1-11.
- [46] LI Z, LAI Z, XU Y, et al. A locality-constrained and label embedding dictionary learning algorithm for image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 28(2): 278-293.
- [47] XU Y, LI Z, YANG J, et al. A survey of dictionary learning algorithms for face recognition[J]. IEEE Access, 2017, 5: 8502-8514.
- [48] LI Z, ZHANG Z, QIN J, et al. Discriminative Fisher embedding dictionary learning algorithm for object recognition[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(3): 786-800.
- [49] BEVERIDGE J R, DRAPER B A, CHANG J M, et al. Principal angles separate subject illumination spaces in YDB and CMU-PIE[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31(2): 351-363.
- [50] MARTINEZ A M, KAK A C. PCA versus LDA[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2): 228-233.
- [51] FAN Z, XU Y, ZUO W, et al. Modified principal component analysis: an integration of multiple similarity subspace models[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 25(8): 1538-1552.
- [52] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

# [作者简介]



**石林瑞** (1996-), 女, 华东交通大学理学院硕士生, 主要研究方向为字典学习、迁移学习、图像处理和模式识别。



**黄玮婧** (1999-), 女, 华东交通大学理学院硕士生, 主要研究方向为规则学习。



**符进武** (1998-), 男, 华东交通大学理学院硕士生, 主要研究方向为行人重识别。



**郭心悦** (1999-), 女, 华东交通大学理学院硕士生, 主要研究方向为目标跟踪。



**范自柱** (1975-), 男, 博士, 华东交通大学理学院教授、博士生导师, 江西省“井冈学者”特聘教授, 江西省“百人远航工程”入选者, 华东交通大学天佑人才计划第一层次(杰出人才)入选者, 华东交通大学数学学科计算机应用技术方向带头人, 主要研究方向为人工智能领域的模式识别与机器学习。