

工程与应用

## 基于SVM的非相关线性判别分析算法研究

张小丹, 吕建平

苏州大学 电子信息学院, 江苏 苏州 215021

收稿日期 2007-6-1 修回日期 2007-8-9 网络版发布日期 2008-1-21 接受日期

**摘要** 基于基因表达谱对组织样本进行分类, 在疾病诊断领域, 是个非常重要的研究课题。在基因表达数据中, 基因的数量(几千个)相对于数据样本(几十个)的个数通常比较多; 也就是说, 数据的维数相比于数据点的个数来说比较高(这个就是采样不足问题)。过高的维数(特征或基因数)将给分类问题带来极大的挑战。提出了结合非相关线性判别式分析方法(ULDA)和支持向量机(SVM)分类算法, 对结肠癌组织样本进行分类识别, 并同其他方法作了比较研究, 分类效果得到了提高; 结果表明了该方法的可行性和有效性。

**关键词** [非相关线性判别分析](#) [支持向量机](#) [基因表达谱](#) [分类](#)

分类号

## Research of uncorrelated linear discriminant analysis based on support vector machine

ZHANG Xiao-dan, LV Jian-ping

School of Electronics and Information Engineering of Soochow University, Suzhou, Jiangsu 215021, China

### Abstract

The classification of tissue samples based on gene expression data is an important problem in medical diagnosis of diseases. In gene expression data, the number of genes is usually very high (in the thousands) compared to the number of data samples (in the tens); that is, the data dimension is large compared to the number of data points (such is undersampled problem). Too high dimension (the number of features or genes) makes the task of classification quite challenging. This paper presents that ULDA and SVM are combined to classify colon tissue samples. Compared to other methods, the effect of classification is improved, the results prove the feasibility and effectiveness of this method.

**Key words** [Uncorrelated Linear Discriminant Analysis](#) [SVM](#) [gene expression profiling](#) [classification](#)

DOI:

通讯作者 张小丹 [xiaodan.albert@gmail.com](mailto:xiaodan.albert@gmail.com)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(683KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

#### 服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

#### 相关信息

▶ [本刊中 包含 “非相关线性判别分析” 的相关文章](#)

▶ 本文作者相关文章

· [张小丹](#)

· [吕建平](#)