

# 手写字体识别方法研究

信息科学与工程学院 计算机科学与工程系

计算机科学与技术专业学生 李嘉理

指导教师 高大启

本文以经典的手写体阿拉伯数字识别为应用背景，研究基于线性分类器的手写字体特征提取和分类方法，重点研究并完成的主要工作有以下几个方面：

1. 用主成份分析对数据集进行特征提取，用累计贡献率衡量降维效果，并分析维数和分类器正确率之间的关系。
2. 面向多类问题，比较一对一和一对多这两种分解方法对Fisher线性分类器、线性感知器、线性核支持向量机分类正确率和时间复杂度的影响。
3. 比较学习因子和最大迭代数对线性感知器分类结果的影响。
4. 比较Fisher线性分类器、线性感知器、线性核支持向量机对USPS、MNIST数据集的分类正确率和学习时间。

通过USPS和MNIST数据集的实验结果，我们得到如下结论：(A) Fisher线性分类器的学习速度最快；(B) 线性感知器的学习时间和分类精度均随学习因子和最大迭代数的变化而变化；(C)支持向量机的学习时间和样本数关系很大；(D)在实现的上述线性分类器中，线性核支持向量机的分类性能相对最好。