

# “手写识别” 实例介绍

ML19

---



礼欣

[www.python123.org](http://www.python123.org)

# 图像识别

- 图像识别 ( Image Recognition ) 是指利用计算机对图像进行处理、分析和理解，以识别各种不同模式的目标和对像的技术。
- 图像识别的发展经历了三个阶段：文字识别、数字图像处理与识别、物体识别。机器学习领域一般将此类识别问题转化为分类问题。

# 手写识别

- 手写识别是常见的图像识别任务。计算机通过手写体图片来识别出图片中的字，与印刷字体不同的是，不同人的手写体风格迥异，大小不一，造成了计算机对手写识别任务的一些困难。
- 数字手写体识别由于其有限的类别（0~9共10个数字）成为了相对简单的手写识别任务。DBRHD和MNIST是常用的两个数字手写识别数据集。



# MNIST数据集

MNIST的下载链接：<http://yann.lecun.com/exdb/mnist/>。

MNIST是一个包含数字0~9的手写体图片数据集，图片已归一化为以手写数字为中心的 $28 \times 28$ 规格的图片。MNIST由训练集与测试集两个部分组成，各部分规模如下：

- 训练集：60,000个手写体图片及对应标签
- 测试集：10,000个手写体图片及对应标签

# MNIST数据集

**MNIST数据集的手写数字样例：**

- MNIST数据集中的每一个图片由 $28 \times 28$ 个像素点组成
- 每个像素点的值区间为 $0 \sim 255$ ， $0$ 表示白色， $255$ 表示黑色。



图. 手写数字样例

# DBRHD数据集

DBRHD ( Pen-Based Recognition of Handwritten Digits Data Set ) 是UCI的机器学习中心提供的数字手写体数据库：<https://archive.ics.uci.edu/ml/datasets/Pen-Based+Recognition+of+Handwritten+Digits>。

DBRHD数据集包含大量的数字0~9的手写体图片，这些图片来源于44位不同的人的手写数字，图片已归一化为以手写数字为中心的32\*32规格的图片。DBRHD的训练集与测试集组成如下：

- 训练集：7,494个手写体图片及对应标签，来源于40位手写者
- 测试集：3,498个手写体图片及对应标签，来源于14位手写者

# DBRHD数据集

### DBRHD数据集特点：

- 去掉了图片颜色等复杂因素，将手写体数字图片转化为训练数据为大小 $32 \times 32$ 的文本矩阵
- 空白区域使用0代表，字迹区域使用1表示。

[illegible]

图. 手写数字样例

# “手写识别” 实例

已有许多模型在MNIST或DBRHD数据集上进行了实验，有些模型对数据集进行了偏斜矫正，甚至在数据集上进行了人为的扭曲、偏移、缩放及失真等操作以获取更加多样性的样本，使得模型更具有泛化性。

- 常用于数字手写体的分类器：

- |                    |             |
|--------------------|-------------|
| 1 ) 线性分类器          | 2 ) K最近邻分类器 |
| 3 ) Boosted Stumps | 4 ) 非线性分类器  |
| 5 ) SVM            | 6 ) 多层感知器   |
| 7 ) 卷积神经网络         |             |

- 后续任务：利用全连接的神经网络实现手写识别的任务