

# 模式识别大作业——花卉识别

王贵锦，李文涛，钟雷声，张晨爽

# 数据库： flower dataset



# Camellia 山茶花





# Water lily 水莲



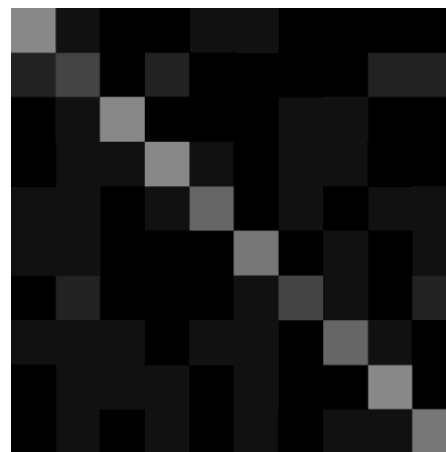
# rose 玫瑰



# 题目要求 (1/5)

## ■ PCA花卉识别。请采用PCA+KNN/Fisher分类方法完成下面的实验。图像特征可以采用灰度像素值、颜色直方图等

- ◆ 用每个品种的前60%做训练，后40%数据做测试。采用K近邻分类(必做：K=1, 选做：K=3, 5)，分析选取不同的主分量个数，对识别率和虚警率的影响。进行闭集测试评价该方法的性能
- ◆ 分析PCA降维对识别的影响
- ◆ 每个品种 
$$P_{i,j} = N_{i,j} / \sum_{j=1}^c N_{i,j}$$
  - ◆ 对角线为每个品种的正确识别率
  - ◆ 非对角线表示每行对应品种识别为其他品种的误识率
- ◆ 进行开集测试并分析(见题目要求3)



## 题目要求 (2/5)

- 使用其他特征和方法进行花卉识别，比如HOG特征/SIFT特征+SVM分类方法等。
  - ◆ 只需要实现一种算法即可，特征和分类器均可以使用开源代码。
  - ◆ 进行闭集测试并分析
  - ◆ 进行开集测试并分析（见题目要求3）
  - ◆ 同PCA+KNN一样，评价该方法的性能
  - ◆ 比较这两个方法的优缺点
    - ◆ 准确率
    - ◆ 速度
    - ◆ 复杂度等

# 题目要求 (3/5)

## ■ 开集测试

- ◆ 图像数据中还给出负样本(neg, 其他类型的花卉图片)。此时需要给出一个合理的拒识方式来判断某张图像是否属于训练的10个类别。请设计一个合理的拒识方式（对于K近邻来说，最简单的方式是对测试图像到训练图像的最近距离设定一个阈值），并对正样本负样本进行识别（11个类别，最后一个为neg类），观察阈值不同时对识别结果的影响。
- ◆ 注意：不应该使用给定的负样本既进行训练又进行测试，可以另外收集一些负样本来帮助确定阈值。



# 开集测试数据集

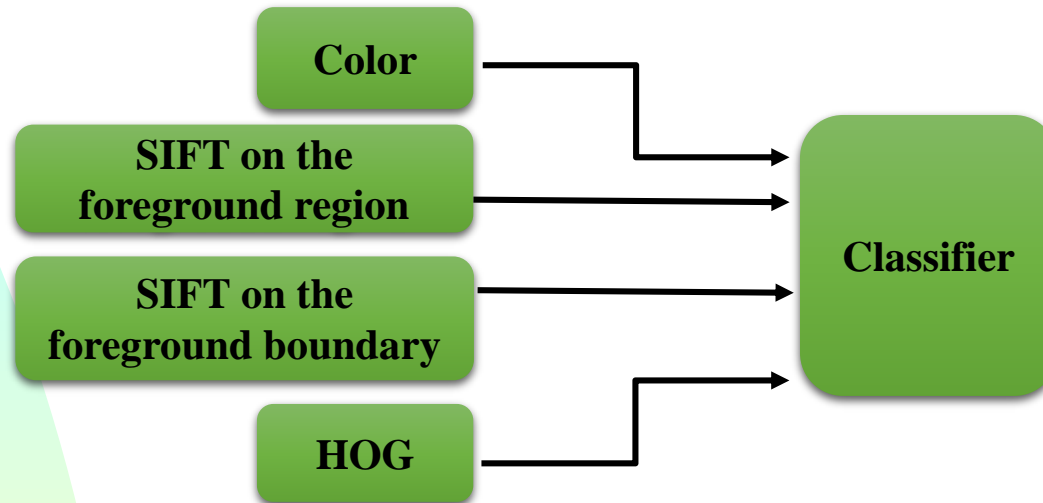


# 题目要求（4/5）

- （选做1）采用其他方法进行识别。
- 算法可以借鉴文献，也可借鉴给出的参考思路，鼓励同学们创新。
- 鼓励使用多种算法进行比较。
- 需要有可视化结果，如曲线、表格和结果示例。

# 参考思路1：多特征融合

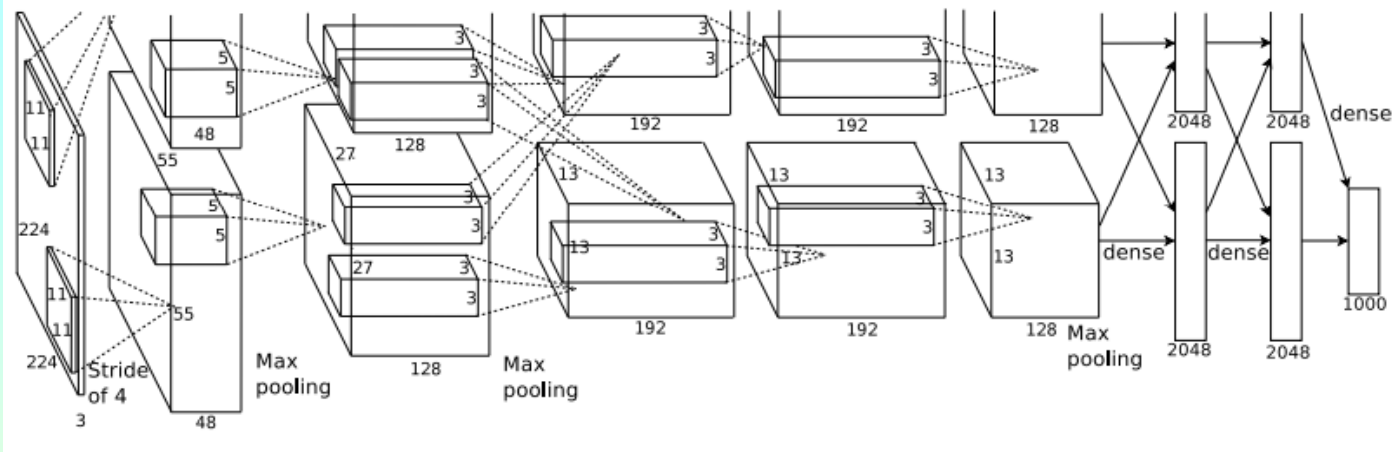
- 多特征融合：



- 参考文献： Nilsback, M-E. and Zisserman, A. Automated flower classification over a large number of classes. Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing (2008)

# 参考思路2: CNN + Transfer Learning

## ■ CNN example: AlexNet (2012)



- Transfer Learning/ Fine-tuning: 使用在大规模数据集（如 ImageNet）中训练好的网络模型参数，在较小规模数据集中进行参数微调。
- Caffe tutorial for fine-tuning :  
[http://caffe.berkeleyvision.org/gathered/examples/finetune\\_flickr\\_style.html](http://caffe.berkeleyvision.org/gathered/examples/finetune_flickr_style.html)



# 题目要求 (5/5)

- (选做2) 在给定的未分割的原始数据集(extend 文件夹)上的30种花卉图片做分类。
- 未分割的原始数据集：
  - ◆ 图像大小不同
  - ◆ 花个数不同
  - ◆ 背景不同
  - ◆ 花所处图像位置不同
  - ◆ 花所处区域面积不同

# 未分割的原始数据集



# 未分割的原始数据集



# 注意事项

- 提交一份报告（word或者pdf均可）；
- 提交源程序，并注明运行方法及运行环境（不要附在报告后面，提交单独的程序包），可以用matlab、C/C++或Python实现；
- 报告中要有理论分析，能体现出自己的思考；
- 鼓励同学们相互讨论，或者利用微信群开展讨论，但是报告和程序必须独立完成；如果发现程序或报告中有内容和表现形式基本相同、可判为抄袭的情况，抄袭者和被抄袭者的project成绩都为零分，所以请同学们要独立完成作业；
- 时间要求，见网络学堂作业提交的截止日期。



# 参考文献

- 模式识别通用方法：
- P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, July 1997. (Fisher)
- N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Computer Vision and Pattern Recognition (CVPR), 2005 IEEE Conference on, 2005, pp. 886–893. (HOG)
- Cortes, C.; Vapnik, V. (1995). "Support-vector networks". (SVM)
- Fei-Fei Li; Perona, P.(2005) "A Bayesian Hierarchical Model for Learning Natural Scene Categories". (BOW)
- Viola, Paul; Jones, Robert (2001). "*Rapid Object Detection Using a Boosted Cascade of Simple Features*". (Adaboost)
- Simonyan K, Zisserman A. "Very deep convolutional networks for large-scale image recognition " [J]. arXiv preprint arXiv:1409.1556, 2014. (CNN)
- Zeiler M D, Fergus R. "Visualizing and understanding convolutional networks[C]//European conference on computer vision. "Springer International Publishing, 2014: 818-833.(CNN)
- Chatfield K, Simonyan K, Vedaldi A, et al. "Return of the devil in the details: Delving deep into convolutional nets " [J]. arXiv preprint arXiv:1405.3531, 2014.(CNN)

# 参考文献

- 数据集相关文献：
- Nilsback, M-E. and Zisserman, A. A Visual Vocabulary for Flower Classification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2006) (Dataset)
- Nilsback, M-E. and Zisserman, A. Automated flower classification over a large number of classes. Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing (2008) (Dataset)
- Yuning CHAI, Victor LEMPITSKY, Andrew ZISSERMAN. BiCoS: A Bi-level Co-Segmentation Method for Image Classification. ICCV, 2011 (Segmentation)

# 大作业补充说明

- 对于c个类别（闭集测试中为10个类别），每个类别i(i=1,...,c)，有测试图片 $N_i$ 张（实验中为25张），测试中，属于类别j的被判为类别i的测试图片记为 $N_{i,j}$ 张，则：

- ◆ 每个类别的识别率：  $P_{TP,i} = N_{i,i} / N_i = N_{i,i} / 40$

- ◆ 整体识别率：  $P_{TP} = (\sum_{i=1}^c N_{i,i}) / (\sum_{i=1}^c N_i) = (\sum_{i=1}^c N_{i,i}) / (10 * 40)$

- ◆ 每个类别的虚警率：  $P_{FP,i} = \frac{\sum_{j=1, \dots, c, j \neq i} N_{i,j}}{\sum_{j=1, \dots, c, j \neq i} N_j} = (\sum_{j=1, \dots, c, j \neq i} N_{i,j}) / (9 * 40)$

- ◆ 整体的虚警率：

$$P_{FP,i} = \frac{\sum_{i=1, \dots, c, j=1, \dots, c, i \neq j} N_{i,j}}{(c-1) * \sum_{j=1, \dots, c} N_j} = (\sum_{i=1, \dots, c, j=1, \dots, c, i \neq j} N_{i,j}) / (9 * 10 * 40)$$

# 大作业补充说明

- 对于开集测试，则有c个类别+一个neg类别（c'=10+1个类别），识别率的计算方式与闭集测试一致，只统计c=10个类别的识别率；计算虚警率时，需考虑到neg测试样本判断为10个类别中的一种（判别为类别i的数量表示为 $N_{i,c'}$ ）所带来的影响，则计算式变为：

- ◆ 每个类别的虚警率：

$$P_{FP,i} = \frac{\sum_{j=1,\dots,c', j \neq i} N_{i,j}}{\sum_{j=1,\dots,c', j \neq i} N_j} = \frac{\sum_{j=1,\dots,c', j \neq i} N_{i,j}}{9*40+200} \quad (i=1,\dots,c)$$

- ◆ 整体的虚警率：

$$P_{FP,i} = \frac{\sum_{i=1,\dots,c, j=1,\dots,c', i \neq j} N_{i,j}}{(c-1) * \sum_{j=1,\dots,c} N_j + c * N_{c'}} = \left( \sum_{i=1,\dots,c, j=1,\dots,c', i \neq j} N_{i,j} \right) / (9 * 10 * 40 + 10 * 200)$$



# 相关的Matlab函数

- `imresize`: 放缩图像大小
- `reshape`: 改变矩阵形状, 可以将图像拉成一维向量
- `imadjust`: 调整图像的灰度直方图
- `pca`: 训练PCA
- `vlfeat`: 可以用于提取HOG特征的库(另有文档介绍如何使用)
- `svmtrain`, `svmclassify`: SVM训练、识别
- `fopen`, `fclose`, `fscanf`, `fprintf`: 文件打开、关闭、读、写

# 基于PCA的识别（续1）

- 计算训练数据的均值  $\mu = \sum_{i=1}^M x_i$
- 训练数据的协方差阵

$$C = \sum_{k=1}^M (x_k - \mu)(x_k - \mu)^T$$

记：  $A = [x_1 - \mu, x_2 - \mu, \dots, x_M - \mu]$  则  $C = AA^T$

# 基于PCA的识别（续2）

- 对协方差阵进行特征值分解，求出特征脸
  - ◆  $Cv_i = \lambda_i v_i$ ，其中  $\lambda_i \geq 0$ ， $v_i$  即为第  $i$  个特征脸
  - ◆ 一般的，图像的维数  $N$  要远大于训练的样本数  $M$ ， $C$  是一个非常大的矩阵。因此，转求下式的特征值分解：

$$A^T A u_i = \lambda u_i$$

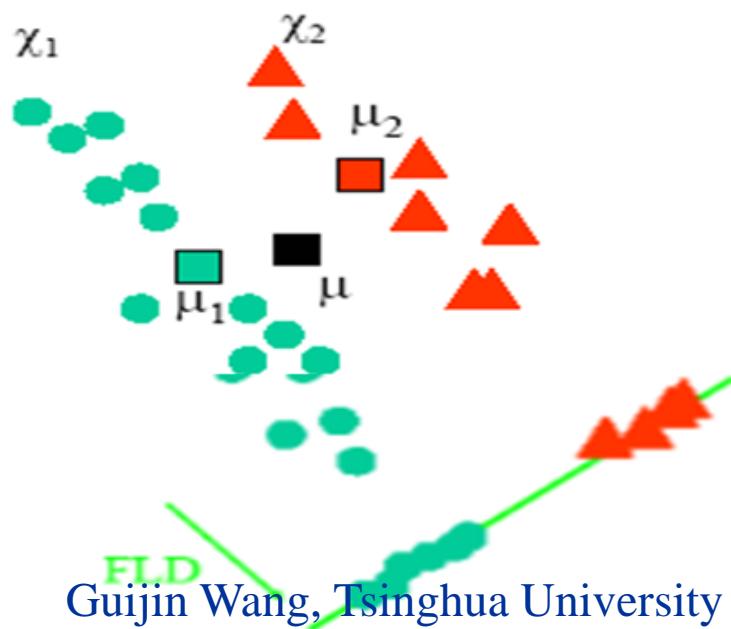
则有：  $AA^T A u_i = \lambda_i A u_i$ ，  $v_i = A u_i / \|A u_i\|$  (归一化)

- ◆ 选取最大的前  $K$  个特征值对应的特征向量来作为特征脸空间的基向量构成变换矩阵  $W = [v_1, \dots, v_K]$

# Fisherface: 基本思想 (续1)

- FLD 选择一种最优的投影变换, 满足:

$$W_{opt} = \operatorname{argmax}_W \frac{|W^T S_B W|}{|W^T S_W W|}$$





# Fisherface: 投影矩阵的计算

- 优化分析表明，满足上述最大化的 $W$ 是下述方程的解：

$$S_B W = S_W W \Lambda$$

- 进一步  $S_w$  假设是非奇异的。则可以通过求解下面的广义特征值问题来得到 $W$ ，

$$S_W^{-1} S_B W = W \Lambda$$

- 如果 $S_w$ 是奇异的，怎么办？

# Fisherface: 投影矩阵的计算 (续1)

- 实际上 $S_B W = S_W W \Lambda$ , 最多只有 $c-1$ 个特征值, 而 $S_W$ 的秩最多为 $M-c$
- 如何求出投影矩阵? 一个简单的策略是先做PCA, 然后再做FLD

$$W_{pca} = \operatorname{argmax}_W |W^T S_T W|, S_T = \sum_{k=1}^M (x_k - \mu)(x_k - \mu)^T$$

将原始特征降到 $M-c$ 维, 接下来做FLD, 降维到 $c-1$

$$W_{fld} = \operatorname{argmax}_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}$$

$$W_{opt}^T = W_{fld}^T W_{pca}^T$$