

模式识别

——猕猴桃膨大果识别

谭耀 SA17168009

目录

一.原理介绍:	3
1.图像预处理	4
1.1 概述	4
1.2 图像去背景及滤波	4
2. 颜色空间的选取	5
2.1 概述	5
2.2 颜色空间选取	5
3. 特征量的提取	7
3.1 形态学处理及阈值分割	7
3.2 特征提取	7
3.2.1 纹理特征	7
3.2.2 纹理特征提取	8
4. 分类识别	9
4.1 BP 神经网络	9
4.2 BP 神经网络的结构	10
4.3 BP 神经网络的学习算法	10
二. 程序设计说明:	10
1.图像处理	10
2.分类器设计:	11
3.GUI 实现	11
三. 实验结果:	12
1.图像处理结果分析:	13
2.分类器结果:	15
四. 结论.....	17

一.原理介绍：

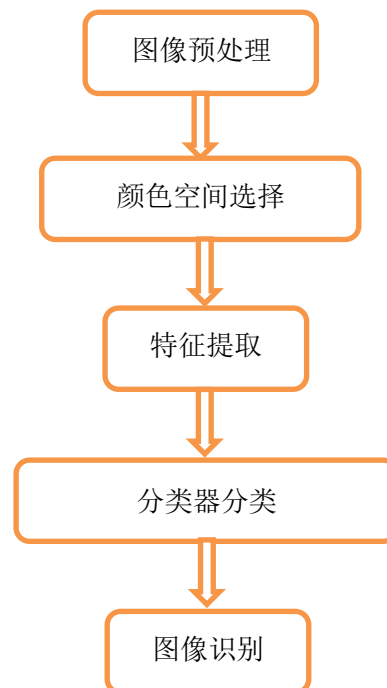
实验内容主要分三个部分。

第一部分为图像预处理，采用中值滤波对图像进行预处理，并去除背景减少干扰。

第二部分为颜色空间的选取及特征提取，通过对 RGB、Lab 等颜色空间进行分析，对比后采用选取了 Lab 颜色空间中的 a 分量进行开闭运算并且进行阈值分割以后，得到膨大果与正常果的图像特征，并且对特征进行了能量、熵、惯性矩、相关性、逆差矩等 5 种特征量的提取。以相同的程序进行批量特征与特征量的提取以后，得到所有样本的特征量。

第三部分为分类和识别，达到识别膨大果的目的。

具体流程：



1.图像预处理

1.1 概述

图像预处理的目的是提高图像质量,确保目标分割的精确度和保真度。因此,需要对图像进行预处理,达到滤除噪声,提高图像质量的目的。本课题的预处理部分主要包括有中值滤波、去背景。

1.2 图像去背景及滤波

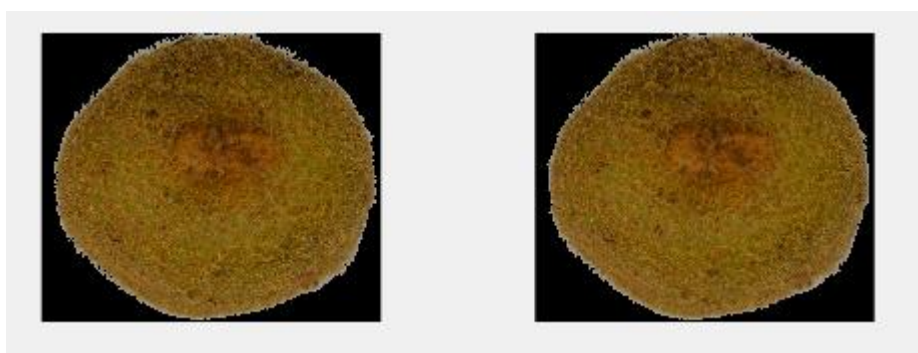
信号和噪声常常交织在一起,如果平滑不当就会使图像的很多细节变模糊,从而导致图像质量降低。因此最大限度地平滑掉图像的噪声同时又能尽量保持图像的细节,是研究图像平滑处理的主要问题。

用 MATLAB 进行图像处理时,背景对特征提取会造成很大的影响。一般背景噪点较多,必须除去。

去除背景并采用 3*3 中值滤波法滤波后的图像:



膨大果去背景及滤波结果图



正常果去背景及滤波结果图

2. 颜色空间的选取

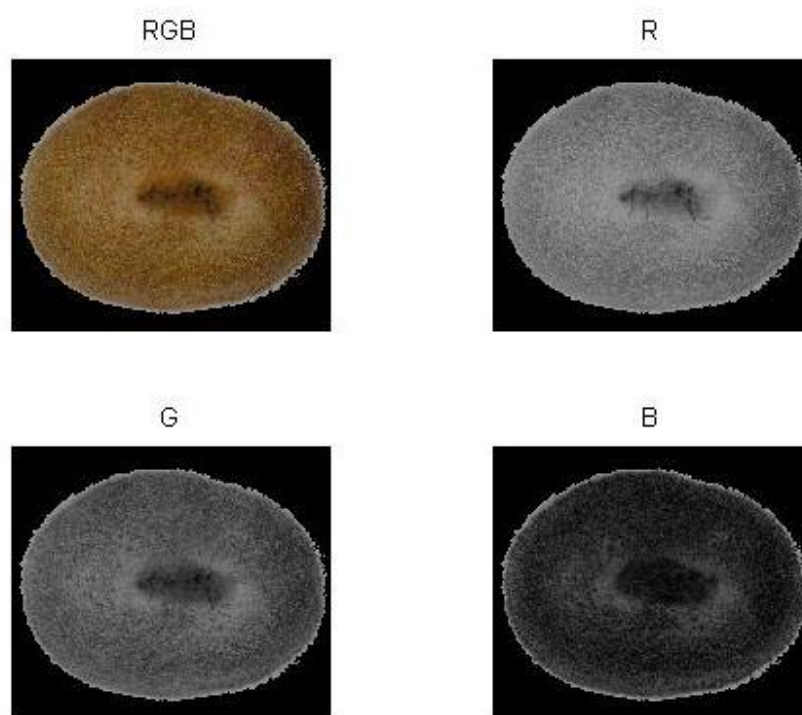
2.1 概述

颜色空间是对颜色的模型化描述, 是使用一组多个分量来描述颜色的数学模型, 如: RGB、Lab、HIS 等。不同颜色空间对我们所需要的膨大果与正常果特征呈现效果不同, 选取合适的颜色空间对后面的特征提取有重大意义, 所以需要多个颜色空间比较以后选择合适的一种。

2.2 颜色空间选取

RGB 颜色空间是最为常见的颜色空间。在 RGB 颜色空间中, 各种颜色的光都可以由红、绿和蓝 3 种基色加权混合而成。RGB 颜色空间在视觉上是非均匀的, 同时它还是一个与设备相关的、颜色描述不完全直观的空间。

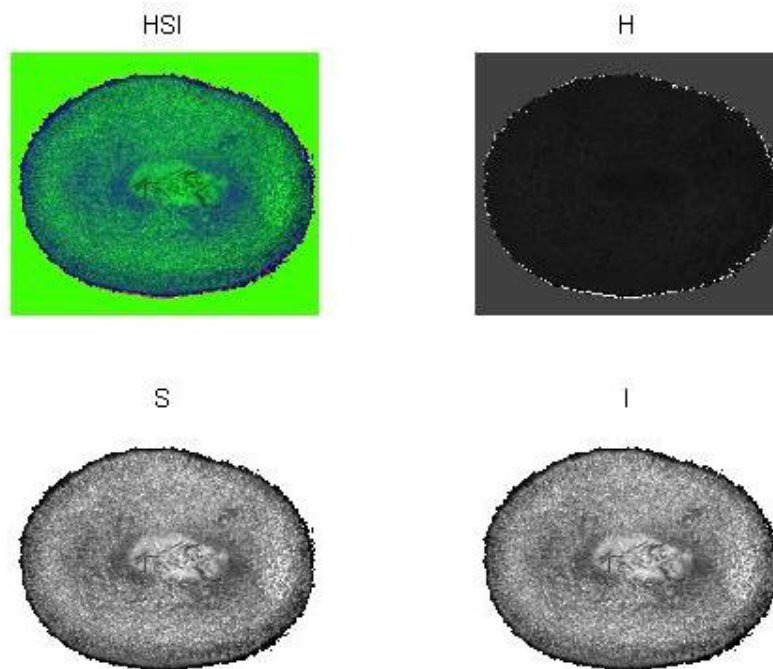
在 MATLAB 中猕猴桃膨大果的 RGB 空间各分量如图:



膨大果 RGB 空间各分量

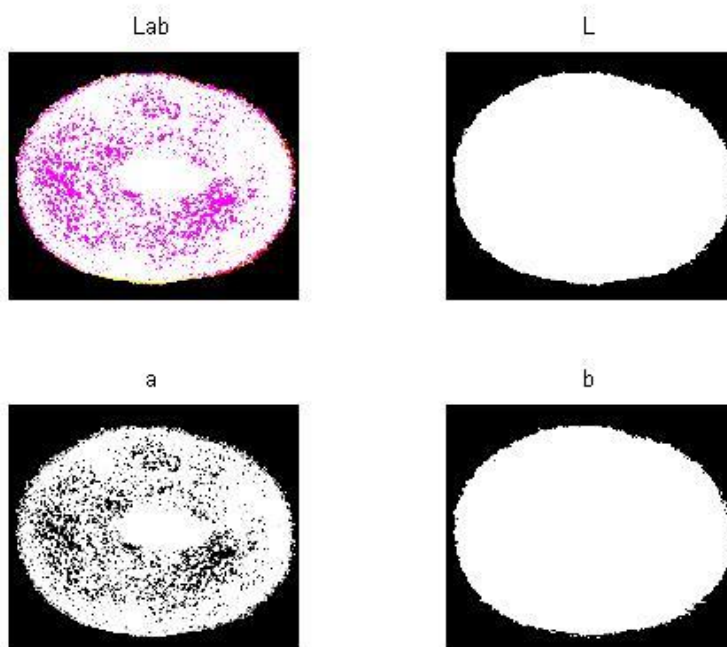
当人们观察某个彩色物体时, 通常用色调、饱和度和亮度来描述。色调是用来描述纯色的属性, 反过来, 饱和度也给出了一种纯色被白光稀释程度的度量。亮度是个较为主观的描述子, 它是不可预测的。亮度体现了无色的强度概念, 也是描述色彩感觉的关键参数。HSI 颜色模型就是这样一个用来反映色调、亮度以及饱和度的彩色模型。

在 MATLAB 中猕猴桃膨大果的 HSI 空间各分量如图:



膨大果 HSI 空间各分量

在 MATLAB 中猕猴桃膨大果的 Lab 空间各分量如图：



膨大果 Lab 空间各分量

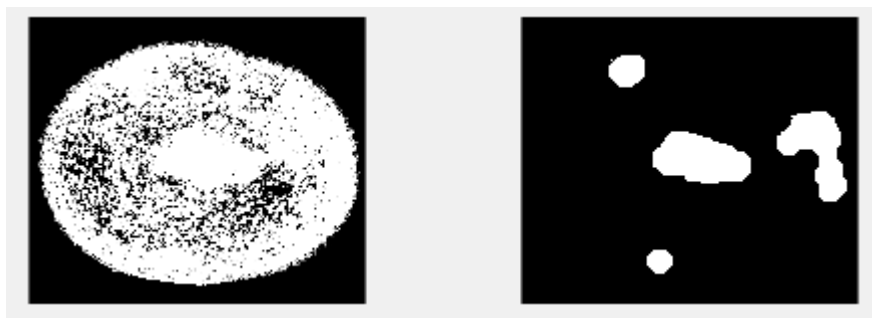
对比以上颜色空间各分量，发现 Lab 空间中的 a 分量特征部位明显，适用于特征提取。故单独提取猕猴桃膨大果和正常果的 a 分量进行特征量的提取。

3. 特征量的提取

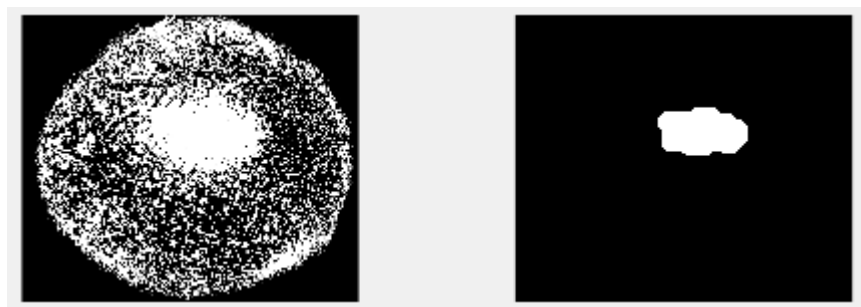
3.1 形态学处理及阈值分割

数学形态学是由形态学的代数运算子组成的，其基本运算有 4 个：膨胀、腐蚀、开启和闭合，它们在二值图像与灰度图像中各具特点。在这些基本运算的基础上还可继续推导、组合成各种数学形态学算法，利用它们进行图像的形状以及结构的分析与处理，主要包括图像分割、特征的提取等。

对猕猴桃膨大果和正常果 Lab 颜色空间中 a 分量进行开启运算和闭合运算后，画出灰度直方图，选取合适的阈值进行分割得到图像如图：



膨大果 Lab 空间 a 分量开闭运算、阈值分割结果



正常果 Lab 空间 a 分量开闭运算、阈值分割结果

3.2 特征提取

3.2.1 纹理特征

纹理特征是一种整体性的特征，作为图像识别领域里较为常用的特征，实际是利用计算机技术从图像中计算得出相关的数值进行描述这一特征。因对纹理的理解和认识不同，对其的定义也不同，纹理特征的提取方法也不同。纹理特征提取目的是使用一些数学参量描述图像中的纹理信息（图像大体走向、规则性等），作为图像分割的参数或是识别分类输入数据。

3.2.2 纹理特征提取

统计学方法是纹理特征提取的实际应用中使用最广泛的方法,而该方法中最经典的算法是灰度共生矩阵。它反映图像灰度情况、相同亮度的像素点在图像中的分布情况和出现频率。在确定相对距离的情况下,它使用相同亮度的像素对在图像不同位置上联合概率密度创建共生矩阵,提取共生矩阵的二阶统计特征来描述纹理特征。

灰度共生矩阵的定义是在确定移动距离(a, b)的情况下,图像中任一点(x, y)点与另一点(x+a, y+b),这两点的灰度值相同都为(i, j),令(x, y)在整幅图像中移动,得到不同的(i, j),统计(i, j)在整幅图像中出现的概率 P_{ij} ,由此就得到了灰度共生矩阵。灰度共生矩阵为对称矩阵,实质是两个像素对的联合直方图。灰度共生矩阵的数学表达式:

$$P(i, j, d, \theta) = \{(x, y), (x + Dx, y + Dy) | f(x, y) = i, f(x + Dx, y + Dy) = j\}$$

灰度共生矩阵反映了图像灰度的变化情况、相邻像素的相关性、同种亮度的两像素出现的频率,并不能直接用于描述纹理特征,需要从共生矩阵中计算出能量、熵、惯性、相关性等值,用这些特征值对整幅图像的纹理特征进行描述。灰度共生矩阵中可以计算出20个用于描述纹理的特征值,根据其各自的意义和实验效果,这20个参数中有重复描述纹理特征的特征值,故可选择常用的5个主要的特征值进行特征描述,即使用能量、熵、惯性、相关性和逆差矩5个特征表征纹理特征。

1. 能量(角二阶矩)

$$ASM = \sum_i \sum_j (p_{ij})^2$$

能量是矩阵中各元素的平方和,反映的是纹理粗细程度。实验证明,粗纹理的元素分布集中于主对角线,其局部灰度变化是均匀的,不存在大幅的灰度变化,能量值较大。而细纹理的元素值所差不大,均匀的分布在整個矩阵中,能量值相对较小。

2. 熵

$$BNT = -\sum_i \sum_j p_{ij} \lg(p_{ij})$$

熵值是图像所含信息量多少的度量。粗纹理的熵值较小,而细纹理变化细小且遍布图像中,所含图像的信息量就大,熵值也大。

3. 惯性(对比度)

$$CON = \sum_i \sum_j (i - j)^2 p_{ij}$$

惯性是反映图像的清晰程度和纹理的沟纹深浅的度量。惯性值越大, 纹理沟纹深, 图像清晰, 视觉效果好, 相邻像素对的灰度差距就大; 反之, 对比值小, 则沟纹浅, 效果模糊, 相邻像素对的灰度差距小。由于细纹理的数值不在主对角线附近, 惯性较大, 而粗纹理正好相反, 惯性较小。

4. 相关性

$$COR = [\sum_i \sum_j ij \times p_{ij} - u_x u_y] / [\sigma_x \sigma_y]$$

相关性反映了纹理在水平和垂直方向上的相似程度。当相关值就值较小时, 纹理在水平和垂直每个方向的相邻区域内走势方向不同; 相反, 相关值较大时, 说明纹理在相邻区域内方向相差不大。相关性可以表示纹理的方向。

膨大果与正常果相关性分布如图:

5. 逆差矩

$$IDM = \sum_i \sum_j p_{ij} / \{1 + (i - j)^2\}$$

反映图像纹理的同质性, 度量图像纹理局部变化的多少。其值大则说明图像纹理的不同区域间缺少变化, 局部非常均匀。

膨大果与正常果逆差矩分布如图:

4. 分类识别

4.1 BP 神经网络

BP 神经网络又称误差反传神经网络, 是最重要的神经网络模型之一, 它是在 BP 神经网络现有算法基础上提出来的, 是通过选出一组随机权值, 其建立线性方程组是通过将目标的输出作为方程代数和, 权值得以解出。BP 神经网络本质上是一种输入到输出的映射, 输入与输出之间的映射关系可以被学习, 而不需要任何输入与输出之间的精确数学表达式。只要用已知的模式对 BP 网络加以训练, 网络就具有输入与输出之间的映射能力, 因此被广泛应用于模式匹配、分类、识别和自动控制等应用领域。

4.2 BP 神经网络的结构

神经网络主要由以下三个神经元组成：输入层、隐含层以及输出层。同 层次内的神经元之间没有任何的连接，仅与相邻的层次的神经元之间有连接但无反馈。神经网络的学习包括两部分：首先是信号的正向传输，当样本数据从输入层输入网络，通过隐含层的处理后到达输出层，输出层的输出结果将与期望输出进行比较。如果两者的误差过大，则进入误差信号的反向传递，则是神经网络的第二个步骤。误差通过隐含层向输入层传播，通过原来的路径返回，分摊给每层的神经元，修正各输出值的权值。就这样不断的通过信号的正向传播和误差的反向传递，直到网络输出的误差达到期望值。

4.3 BP 神经网络的学习算法

整个神经网络的学习算法分为以下 7 个步骤：

- (1) 神经网络的参数设置，设定学习效率、训练函数、传输函数、期望误差和隐含层节点数等；
- (2) 从训练样本中随机输入一组样本值以及对应的期望输出值；
- (3) 通过神经网络正向信息传递，计算各神经元的输出；
- (4) 计算神经网络的实际输出与期望输出值之间的误差；
- (5) 判断误差是否达到期望误差，若达到即结束神经网络的学习；
- (6) 如果误差没有达到期望误差，则继续神经网络的学习，利用反向传播，逐层修正网络的连接权值；
- (7) 返回第 3 步，直至训练集的样本误差达到期望值。

二. 程序设计说明：

本次程序的语言为 python，主要分为三个模块的内容：图像处理，分类器设计，GUI 实现。

1.图像处理

图像处理采用的是 opencv 模块，数据临时存储和处理采用的是 panda。在程序文件中为 picprocess 类，其处理过程如下：

- (1) opencv 读取图像
- (2) 利用蓝色阈值差异，消除背景。
- (3) 对图像进行 3*3 的矩阵进行中值滤波
- (4) 将图像转换到 lab 色彩空间，并进行均值化操作。
- (5) 对图像进行先后进行模板为 (10*10) (50*50) 开闭操作。

(6) 将图像长宽压缩成为原图的 1/16，并计算相应的能量、熵、惯性、相关性和逆差矩

同时此类还可以对图像进行批处理，将数据写入 csv 文件中，可以用 excel 打开操作。

2.分类器设计：

分类器采用的是 BP 网络，采用的框架是 google 旗下的 tensorflow，也采用了 numpy 对数据进行处理。其类名为：NeuralBPNet

(1) 分类器结构

分类器采用两层结构，第一层为隐含层，输入为 5 个特征参数，输出为 7 个，采用的激活函数是 sigmoid。其表达式如下所示：

$$f(x) = \frac{1}{1 + e^{-x}}$$

第二层为输出层，输入为 7 输出为 1，采用的是的是线性激活函数

(2) 模型训练

训练样本采用的是：如果是非膨大果则预期输出为 1，如果是膨大果则输出为-1，，对于辨别是否为膨大果，则将数据输入，计算得到输出，若输出结果大于零，则判定为非膨大果，若小于等于零，则判定为膨大果。损失函数设定为最小均方根误差，对于网络的训练的方法采用的是梯度下降的方法，梯度的 $\beta = 0.001$ ，将所有的膨大果和非膨大果的数据分成按照 5: 3 的比例，随机分成训练数据和测试数据，设计循环的次数为 100000 次，则并将准确度最高的实验参数保存下来。文件默认保存在 ckpt 目录下。

(3) 模型加载与识别：

可以通过加载保存神经的参数文件夹来对文件进行加载，加载好后，输入图像的五个参量，可以对图像进行辨别，若正常果则输出为 1，若膨大果则输出为-1，其他情况输出为 0

3.GUI 实现

GUI 设计主要采用的是 PyQt5 模块，画图采用的是 matplotlib，其主要的类有：

MyMplCanvas：画模型训练过程中的损失精度图像

MainOp：另外一个线程，负责执行后台的图像处理，识别，模型的训练过程。

MainWin：主要负责界面的显示，与用户进行交互。

(1) 图形界面说明：

图形界面显示如下：



菜单栏中，可以选择即将进行的操作：

图像菜单：包含对图像的批处理和图像模式识别操作

模型菜单中有：模型加载和模型训练操作。

中间图片区域：显示正处理的图片或者模型训练的实时精度。

中间文本框：显示即将操作的内容和参数。

中间按钮：点击将执行中间文本框所将要进行的操作。

底部文本框：显示操作得到的结果或者实时的信息。

（2）操作说明：

加载训练好模型：打开模型菜单，点击载入，选择模型的文件夹，中间文本框显示相应内容，点击执行。

训练模型：打开模型菜单，点击训练，出现文件对话框，依次选择保存正常果数据和 非正常果的 csv 后缀的文件，中间文本框显示相应内容，点击执行。

识别图片：加载模型或者训练模型后，点击打开图片菜单，点击识别，通过文件框选择相应的图片，中间文本框显示相应内容，点击执行。

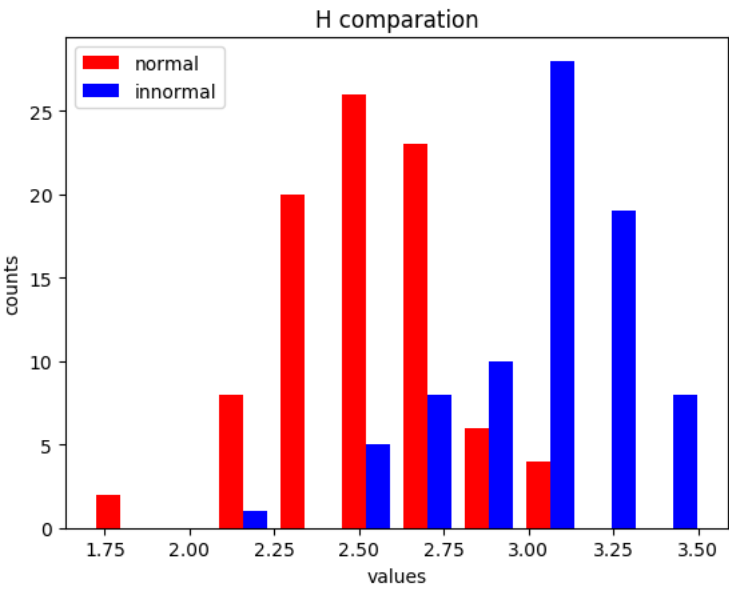
批处理文件：打开图片菜单，移到批处理子菜单，点击膨大果或者非膨大果菜单，进入文件夹选择菜单，点击选择包含有相应图片的文件夹。中间文本框显示相应内容，点击执行。训练后的数据，如果是膨大果将保存在 false.csv 如果是非膨大果则保存至 true.csv

三. 实验结果：

实验选用了 142 个猕猴桃样本，其中膨大果 79，正常果 89 作为样本集。

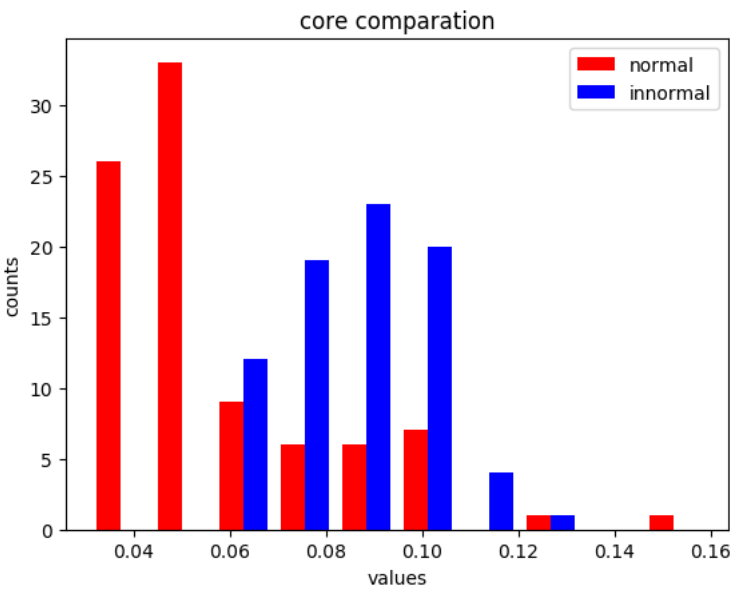
1.图像处理结果分析：

我们对正常果和膨大果的能量、熵、惯性、相关性和逆差矩进行统计，分别得到其不同取值区间的个数，可以观察其参数分布随属性不同的变化。



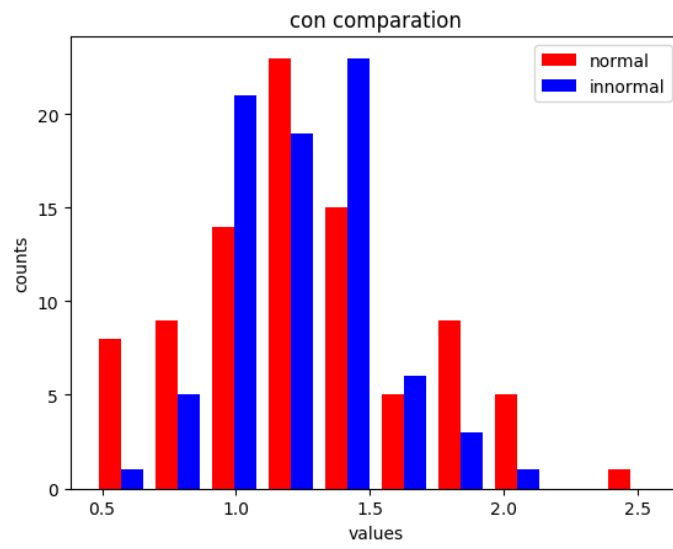
熵统计图像

可以看出正常果实的熵比较小



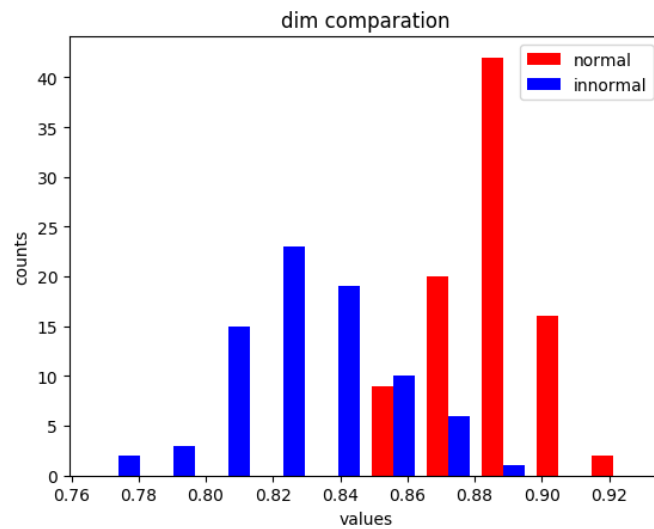
惯性值统计

可以看出正常果的惯性值偏大



相关性统计

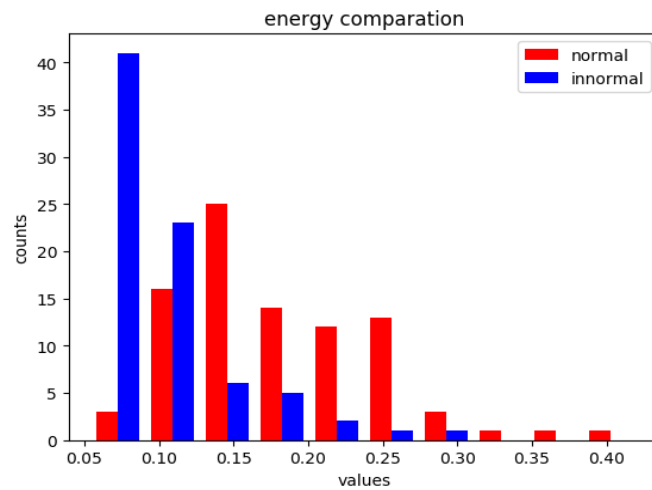
可以看出这两者的相关性统计分布区分度不是很大，但是正常果的相关性值略



大。

逆差矩统计

可以看出此参数，正常果和异常果区分明显。正常果的逆差矩比较大

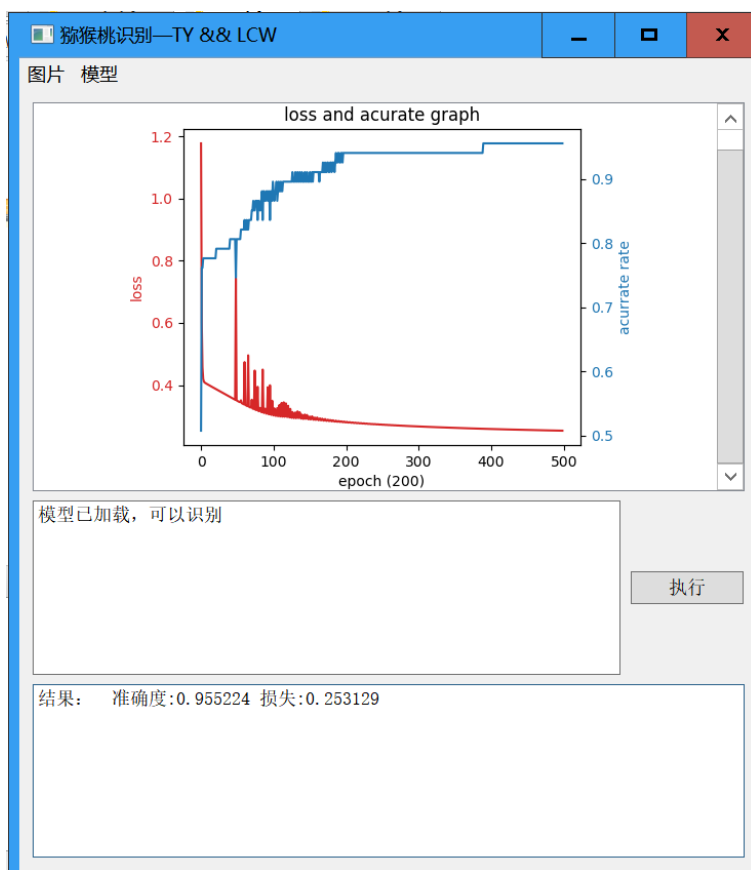


能量统计

可以看出正常果的能量略微比异常果大。

2.分类器结果：

利用上面计算得到的五个特征值，对神经网络进行训练。其训练结果如下图所示：



模型训练运行图

可以发现随着训练步数的增加，其 loss 不断降低，精度不断提高，最终达到稳定的状态。最终可以使得其准确率达到百分之 95.5，说明其准确率达到很高的程度，也尝试使用核方法和 svm 方法，但是最终的效果不如 bp 网络的感知器模型。

(1) 识别图像：

输入一张膨大果的图片对其进行识别，其结果如下：



可以对其进行正常识别。

四．结论

此次对猕猴桃图片进行图片进行处理，得到了相应的参数，并设计了含隐含层的感知器分类器对其进行识别，得到了较高的准确率，但是样本数比较少。图片处理和特征值的提取也存在改进的空间，需要进一步的改进。