

北京邮电大学

现代数字通信论文



题目：基于数据挖掘的神经网络算法研究

小组成员及分工：

姓名	学号	实验室	导师
黄锦雨	2015140011	主楼 1017	陈萍
系统整体设计，神经网络算法的设计与实现，报告撰写			
石志倩	2015140010	主楼 1017	陈萍
图像特征提取算法的设计与实现，报告撰写			
朱亚亭	2015140100	主楼 1017	李屹
系统级联与数据仿真,结果分析，报告撰写			

2016 年 1 月

基于数据挖掘的神经网络算法研究

摘 要

随着信息科学技术的发展，信息以爆炸式的速度增长，这些信息包含了很多有价值的隐藏信息。数据挖掘是指从大量的隐藏信息中通过特定算法来挖掘有用信息的过程。数据挖掘是一个跨学科的研究领域，吸收了从模式识别，统计抽样，人工智能，机器学习等学科的思想方法，近年来已成为热点研究领域。神经网络算法在数据挖掘领域被广泛使用，尤其是其强大的非线性映射能力得到了很广泛的应用，其在数据分类中有着特殊的作用。图像分类是数据挖掘的一个重要分支，在数码相机、摄像机移动等诸多终端流行的今天，数字图像的数据量急剧增加，人工分类既耗时又费力，急需有效的数字图像分类算法。

本文研究了基于神经网络的数字图像分类方法，首先提取图像的颜色特征，纹理特征，不变性特征，每张图片提取了一个 32 维的特征向量。设计了 BP 神经网络分类器，在不同参数设置下对 Corel image Datasets 数据集中的图像进行了分类，并分析了不同参数设置下分类的效果和算法的性能。

本文通过对数据分析和对算法的性能测试，通过对不同参数设置的分析，在几种不同的方案中找到了一种最优的分类方法。

关键词：图像分类 BP 神经网络 图像特征提取

Research on the neural network algorithm based on Data Mining

ABSTRACT

with the development of science and technology of information, the information is rising with a rapid speed. This information contains a lot of valuable hidden information. Data mining is a process of searching hidden information from massive data using special algorithms. Data mining is an interdisciplinary research field, which take advantage of ideas of statistics, estimation and hypothesis testing, artificial intelligence, pattern recognition and machine learning algorithm, modeling technology and learning theory. These years have seen that data mining has been a hot research field. The neural network algorithm has been widely used in the field of data mining, especially its powerful nonlinear mapping ability which has a special role in data classification. Image classification is an important branch in the field of data mining, and today with the popularity of digital camera, mobile terminal, the amount of digital image data increasing rapidly, manual classification is time-consuming and laborious, so that image classification algorithm is of urgent need.

In this paper, we study the digital image classification method based on neural network. First, we extract color feature, texture feature, invariant features, each image extracted a 32 dimensional feature vector. BP neural network classifier is designed, under different parameter settings on Corel image datasets data centralized image are classified, and to analyze the different parameter settings under the classification effect and the algorithm performance.

Based on data analysis and the algorithm performance testing, through the analysis of different parameter settings, an optimal classification method in several different scenarios was found.

KEY WORDS: Image Classification BP Neural Network Image feature extraction

目 录

第一章 绪论.....	- 1 -
1.1 研究背景.....	- 1 -
1.1.1 神经网络简介.....	- 1 -
1.1.2 图像分类简介.....	- 3 -
1.2 本文的主要工作及章节安排.....	- 4 -
1.2.1 主要工作.....	- 4 -
1.2.2 章节安排.....	- 4 -
第二章 神经网络原理.....	- 5 -
2.1 人工神经网络.....	- 5 -
2.1.1 人工神经网络简介.....	- 5 -
2.1.2 神经网络模型.....	- 5 -
2.2 BP 神经网络原理.....	- 6 -
2.2.1 BP 神经网络结构.....	- 6 -
2.2.2 BP 算法的学习算法.....	- 7 -
2.2.3 改进的 BP 神经网络.....	- 9 -
第三章 基于 BP 神经网络的图像分类算法.....	- 12 -
3.1 引言.....	- 12 -
3.2 算法描述.....	- 12 -
3.2.1 算法流程描述.....	- 12 -
3.2.2 系统实现框图.....	- 13 -
3.3 图像的特征提取.....	- 14 -
3.3.1 特征提取流程图.....	- 14 -
3.3.2 颜色特征的提取.....	- 15 -
3.3.3 纹理特征的提取.....	- 16 -
3.3.4 不变性特征的提取.....	- 19 -
3.3.5 数据归一化处理.....	- 20 -
3.4 分类算法的设计.....	- 21 -
3.4.1 训练数据集.....	- 21 -
3.4.2 BP 神经网络的设计.....	- 22 -
3.4.3 BP 神经网络的训练.....	- 25 -
第四章 算法仿真实现及结果分析.....	- 26 -
4.1 算法的仿真实现.....	- 26 -
4.1.1 BP 神经网络的仿真.....	- 26 -
4.1.2 实验数据与实验环境.....	- 26 -
4.1.3 不同参数下的结果分析.....	- 27 -
4.2 算法鲁棒性测试.....	- 37 -
4.2.1 测试方法.....	- 37 -
4.2.2 算法鲁棒性测试结果.....	- 37 -

4.3 本章小结	- 38 -
第五章 总结与展望	- 39 -
5.1 总结	- 39 -
参考文献	- 40 -

第一章 绪论

1.1 研究背景

1.1.1 神经网络简介

人工神经网络是由许多由可调节连接权重的神经元连接而成的，它具有分布式信息存储、自组织自适应能力、大规模并行处理能力好,有很强的学习能力^[1]，BP（Back Propagation）算法，也称误差反向传播算法，是人工神经网络中的一种监督学习算法。神经网络基本结构由非线性变化单元组成，具有非常强的非线性映射能力，在理论上神经网络算法可以拟合任何函数。神经网络的参数设置灵活性很大，网络的隐含层数、每一层的神经元数及网络的学习系数，传递函数等参数可根据不同的场景确定，在数据分类，信息处理与场景识别、自动化控制、故障检测处理等众多学科领域都有很大的潜在应用价值。

（1）人工神经网络的起源及发展：

人工神经元的起源源于脑神经元学说。19 世纪末，在生物、生理学领域，Waldeger 等人创建了神经元学说。研究者意识到复杂的神经系统是由不计其数的神经元连接组成。人脑包含超过百亿以上的神经元，每 mm^3 大约包括数万个神经元细胞，它们通过复杂的连接组成神经网络，庞大的神经网络控制着人的所有活动。神经元细胞的结构很特别，如图 1-1 具有许多的分枝，神经元细胞大致分为三个部分，轴突、细胞体和树突。细胞体延长后渐渐变细形成树突，树突的每个部分皆可与另外的神经细胞的轴突末梢互相“通信”，形成“突触”。在突触接触的地方两神经细胞并没有完全连接，神经末端只是传递信息功能的连接部位，突触可分为兴奋性与抑制性两种类型，它取决于神经细胞之间耦合的极性。每个神经细胞相互的连接极性和强度各不相同，且全部可以根据不同的需要进行调整、由于这种特点，大脑可以记忆和存储大量的信息。

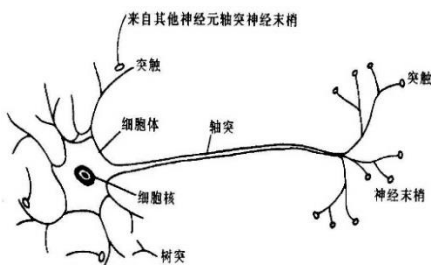


图 1-1 神经元图例

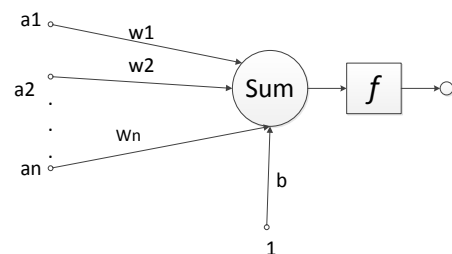


图 1-2 神经元模型

神经网络研究并不是一帆风顺的，其过程可谓是跌宕起伏，有高潮也有低谷。人工神经网络先驱者的研究可以追溯到 20 世纪 40 年代。图 1-2 是由心理学家 W·Mcculloch 和数理逻辑学家 W·Pitts 于 1943 年在分析、总结神经元基本特性的基础上首先提出神经元的数学基础模型。 $a_1 \sim a_n$ 为神经元输入特征向量的特征维； $w_1 \sim w_n$ 为神经元各个

突触的权值； b 为偏置量； f 为激活函数； t 为神经网络的输出。这个神经网络模型得到的广泛的应用，更是神经网络数学理论的基础模型，这个模型是神经网络研究过程中里程碑式的创举。所以，此两人可称为人工神经网络研究的奠基性人物。

50 多年来，神经网络研究工作按时间顺序历经了五个阶段萌芽期、第一次高潮期、低谷期、第二次高潮期、理性发展期与应用研究期。目前，神经网络已成为包括人工智能、信息科学、计算机科学、脑神经科学和智能控制等不同学科和领域的前沿的新兴交叉研究学科。

（2）人工神经网络的研究主要内容：

目前神经网络的研究主要集中在三个方面^[2]：理论基础研究、技术实现研究、应用研究。

a) 神经网络理论基础研究

神经网络理论基础研究重点在于找到恰当的神经网络算法模型和训练算法。其中，模型算法研究指的是构造恰当的理论神经元数学模型，并找到一种神经元之间有效的连接方式，并且确定其可以应用的场景；训练算法研究指的是在模型算法研究的架构上，找到合适的算法来修改权值和神经网络结构，与此同时必须达到训练样本的内在要求，更重要的要进一步提高训练速度。从理论上分析常用的神经网络设计是神经网络理论基础研究的另外一方面的重要内容。

b) 神经网络技术实现研究

神经网络技术实现的研究的侧重点在于讨论利用光学、光电、电子、生物等其他学科的技术实现神经网络的途径和方法，其中有利用传统计算机科学技术实现神经网络模型的计算机，及新型神经网络计算机结构体系的研究等。

c) 神经网络应用研究

确定如何有效利用神经网络合理解决实际中的问题是神经网络应用研究的重点。近些年，神经网络的应用研究已经取得了一定的成果，这得益于神经网络理论基础以及与神经网络相关技术取得的成效。近几年科研人员深入的研究，许多的领域中已经应用了神经网络。神经网络目前应用的主要学科领域有：专家系统、故障测试、模式识别、非线性系统辨识和控制、信号处理、知识处理、知识工程、机器人控制、物资调用、市场分析、智能接口和认知科学等等。

（3）BP 神经网络的研究现状

1986 年 Rumelhart 和 McClelland 共同发表了 BP 神经网络是一种多层前馈神经网络。当时提出的神经网络使用的是经典的 BP 训练算法，即梯度下降算法或者称负梯度算法，目标函数是平方误差。他们提出的方法为根据梯度下降的方向迭代修改网络阈值的和权值以使目标函数达到设定的最小值。BP 神经网络也存在一些问题，比如网络结构难以确定、收敛速度慢、泛化能力差、易得到局部极小值的缺陷，为了满足实际的需要，近些年来，很多学者研究出了很多提高性能的方法。

1.1.2 图像分类简介

目前, 互联网上的数字图像信息正在快速增长并在各个领域广泛应用, 数字图像成为了一种不可或缺的信息资源^[4]。数据统计结果表明, 人们获取信息主要是通过视觉, 有一般以上的信息获取是通过视觉的。因此图像分类问题越来越受到科研工作者的关注, 人工分类已经越来越费时费力, 所以图像分类算法是需要迫切研究的。由于图像信息更加复杂, 所以图像分类问题也是一个复杂的问题, 往往没有一个合理的算法对所有的图像都有效, 图像分类算法是研究者关注的焦点。

图像分类是指: 根据体现在图像中的信息不同特点, 对不同类别的事物来进行识别的图像处理方法, 是由计算机进行一个模式识别的过程^[5]。当前, 解决图像分类问题有两方面要解决的问题, 一个如何跨越语义鸿沟问题: 人工对图像分类时一般是基于图像内容和图像所代表的事物的语义而不是图像的底层特征。从图像的底层特征中获取图像的事物语义对人来说是很困难的, 人只能依据日常生活的一点一滴的经验和常识知识来判别图像的内容。通常来说计算机判别一张图像的类别时是依据图像的低层特征去判断的, 而人是基于图像的内容或者说视觉语义来判断类别的, 这就在人所理解的内容和计算所理解的内容之间产生了“语义鸿沟”^[5]。还有一个难题是图像的特征提取问题, 由于在照相的时候环境有很大的不同, 所以照相时的光线, 远近, 角度都会影响图像的效果, 而这些影响都会对图像的特征提取产生不同程度的影响, 这些影响也加大了跨越语义鸿沟的困难^[6]。

(1) 基于文本的图像分类技术

基于文本的图像分类技术首先是对图像内容进行人工识别, 然后标出表示其内容的文本, 再将文本作为图像分类的依据。由于基于内容的分类算法不是很成熟, 现阶段各大网站, 搜索引擎所使用的检索办法依然是基于文本的图像检索, 计算机会自动根据检索词查找匹配图像库中已经给出图像名标签的图像, 这种匹配是不涉及图像的底层信息的, 检索相对较简单, 但是前期需要繁杂的人工分类, 而且由于标签的不准确其准确性已经不能满足人们对于准确性的要求。一般在基于文本的图像分类初期, 都要建立数据, 数据库中包括描述图像内容的关键字, 分类或检索时用户输入关键字, 计算机到数据库中找到对应的关键字, 然后根据这些关键字对图像进行分类或检索。这种方法的一个缺点是, 不适合数据量很大数据集, 数据量大时分类会很有难度, 表现在第一, 人工标注关键字会很繁琐, 工作量巨大; 第二点是对复杂性很高的图像, 几个关键字不足以描述图像的内容, 关键字不准确会导致后面分类或检索结果不准确; 第三点是地球上语言多种多样, 不可能对每张图像都标记上所有的语言, 因此关键词的分类有一定的局限性。

(2) 基于内容的图像分类技术

基于内容的图像分类技术的核心是图像的底层可视特征, 其本质是一种近似匹配技术, 而且在最关键的特征提取步骤上, 特征提取的好坏直接影响到分类效果的好坏, 特

征提取的一个优点是提取特征的过程由计算机完成,可以一定程度上杜绝人为操作所产生的主观臆断,特征提取方法对每张图像都是“公平”的。图像特征提取和匹配这两个关键技术的解决是基于内容的图像分类技术的实现的关键。图像的特征提取分为两类,一是底层视觉特征的提取,如颜色、纹理、形状,不变性等特征;二是语义特征,它属于高层概念级别的翻译,语义特征一般要人的经验知识来解释^[7]。由于当前计算机视觉与模式识别的发展还存在着一定的局限,导致无法真正完全实现语义特征级别的图像分类技术。在当前,大部分研究人员所研究的算法大多还是基于图像的底层视觉特征。基于文字的图像分类过程不存在“语义鸿沟”的问题,描述图像的关键字基本就能够反映意图。而基于内容的图像分类存在着一个底层的视觉特征与上层语义理解间的差异,这就是著名的“语义鸿沟”(semantic gap)。

1.2 本文的主要工作及章节安排

1.2.1 主要工作

本文的主要工作是设计了一个基于BP神经网络的图像分类模型,主要包括对图像颜色,纹理,不变性三大特征的提取,神经网络的设计与实现,在Corel数据集上测试了算法的性能。本文通过对数据分析和对算法的性能测试,通过对不同参数设置的分析,在几种不同的方案中找到了一种最优的分类方法。

1.2.2 章节安排

本文一共分为五章,各章的主要内容简介如下:

第一章为绪论。该章主要介绍本文的选题背景与研究意义,并大致介绍了一下神经网络和图像特征提取研究的现状,并给出了全文的整体组织结构。

第二章介绍了神经网络原理的相关知识,介绍了神经网络的发展现状,以及神经网络的结构模型,神经网络的学习算法和训练算法的数学原理。

第三章是本文的核心部分,介绍了基于神经网络的图像分类算法,图像特征的提取方法,包括颜色,纹理和不变性特征,同时介绍了神经网络算法的设计训练和仿真,给出了系统的流程图和算法步骤。

第四章介绍了在不同参数设置下的算法性能分析,测试了算法的有效性,给出了一个基于本文分类算法的图像检索的应用。

第五章是本文的总结与展望。

第二章 神经网络原理

2.1 人工神经网络

2.1.1 人工神经网络简介

人工神经网络（Artificial Neural Network, ANN），通常简称神经网络，是一种在生物神经网络启示下建立的数据处理模型。神经网络由大量的人工神经元相互连接进行计算，根据外界信息改变自身的结构，主要通过调整神经元之间的权值来对输入的数据进行建模，最终具备解决实际问题的能力。

从某种意义上讲，人类自身就是一个训练好的模式识别系统。现在最好性能的模式识别算法也无法与人类相比，当人看到一个人脸时，能够在毫秒的时间内分辨出人物是谁，这对现阶段的计算机来说是无法完成的工作。在人类刚出生时，其神经元几乎不存储任何信息。在周围环境的刺激下，神经元之间的连接关系逐渐发生改变，人也就慢慢有了记忆，有了处理信息的能力。

人工神经网络就是模仿生物神经网络建立起来的，是对生物神经网络的抽象，不可能完全反应大脑的全部功能和特点。在人工神经网络中最重要的是神经元节点与权值。节点对应有向图的节点，权值表示节点间的相互连接强度。人工神经网络的特点在于其权值是可以调整的，当权值调整到恰到好处是，就能输出正确的结果。

2.1.2 神经网络模型

尽管神经网络的起源于生物神经网络有着密不可分的关系，但神经网络是对生物神经网络的抽象和简化。神经网络模型包含节点和连接权值，可以分为不同的种类。根据种类的不同可分为前向网络和反馈网络。单层感知器，线性网络是单层前向网络，多层感知器、径向网络都是多层前向网络。如图 2-1 是前向网络的结构。

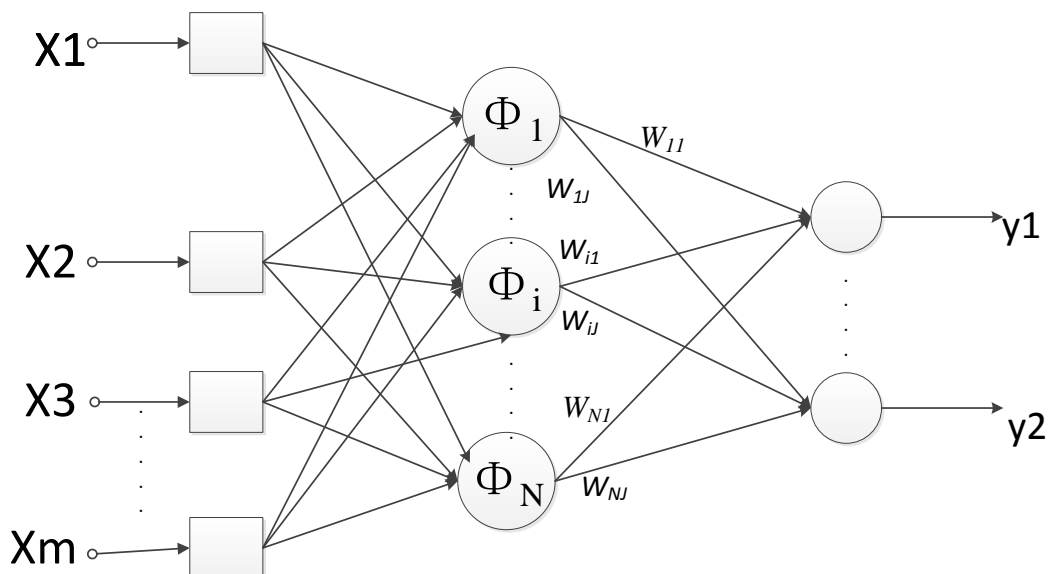


图 2-1 前向网络结构

根据学习方式不同可分为有监督学习网络和无监督学习网络。BP 网络，径向网络均是有监督网络，需要人为地给出已知目标的样本进行训练。大部分自组织网络是典型的无监督学习网络，只需要将待求的样本输入网络即可。

2.2 BP 神经网络原理

2.2.1 BP 神经网络结构

误差反向传播神经网络，简称BP (Back Propagation)网络，是一种有隐含层的多前馈神经网络，如果网络的输入节点数为M、输出节点数为L，则此神经元网络可看成是从M维欧氏空间到L维欧氏空间的映射^[1]。从本质上讲，BP算法把一组样本的输入输出问题，变换为一个非线性优化的问题。BP神经网络是目前神经网络中使用最为广泛的网络模型，它无论在网络理论还是网络性能方面都已经相当成熟，其突出的优点就是具有很强的非线性映射能力和柔性的网络结构。现实世界中的许多问题，如模式识别、图像处理、系统辨识、函数拟合、优化计算、最优预测和自适应控制等，都可以转换成这种神经元网络来处理。BP算法的神经元网络结构图和BP网络结构的图形公式法分别如图2-2，2-3。

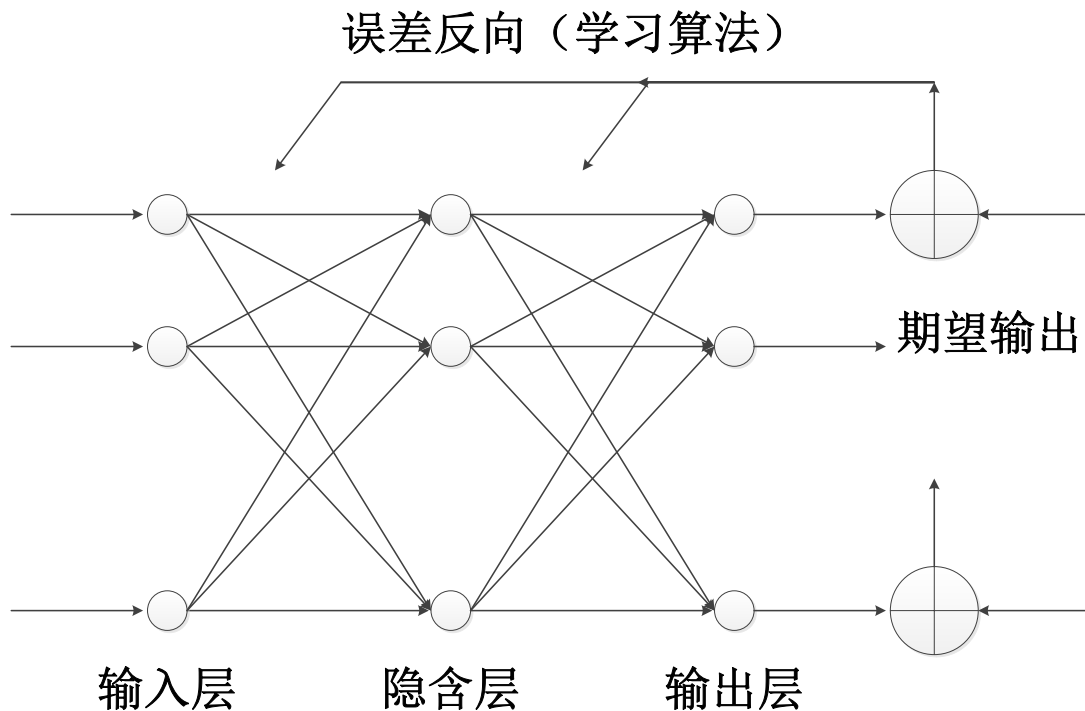


图 2-2 基于 BP 的神经网络模型

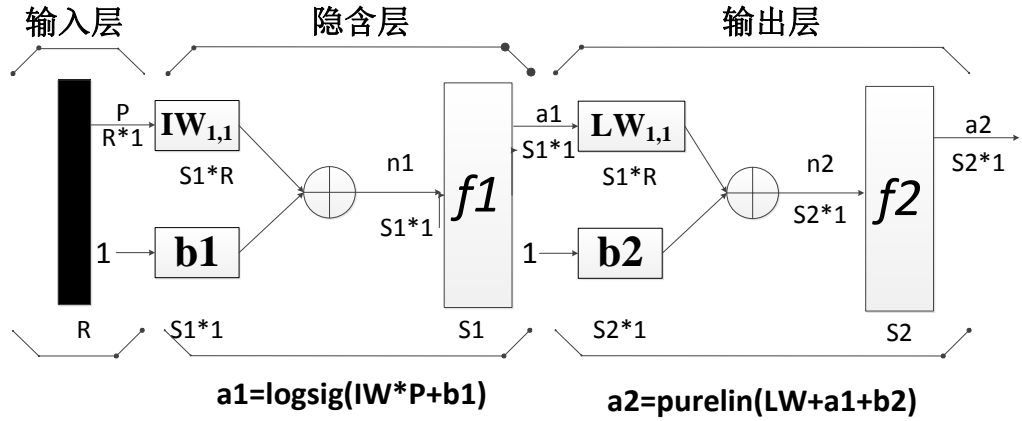


图 2-3 BP 网络的图形公式法

误差反向传播的BP算法(简称BP算法)，其基本思想是最小二乘算法。它采用梯度搜索技术，以期使网络的实际输出值与期望输出值的误差均方值最小。BP算法的学习过程由信号的正向传播与误差的反向传播两个过程组成。正向传播时，输入样本从输入层传入，经各隐含层逐层处理后，传向输出层。若输出层的实际输出与期望输出不符，则转入误差的反向传播阶段。误差反向传播是将输出误差以某种形式通过隐含层向输入层逐层反传，并将误差分摊给各层的所有单元，从而获得各层单元的误差信号，此误差信号即作为修正各单元权值的依据。这种信号正向传播与误差反向传播的各层权值调整过程，是周而复始地进行的。权值不断调整的过程，也就是网络的学习训练过程。此过程一直进行到网络输出的误差减少到可接受的程度，或进行到预先设定的学习次数为止^[22~25]。

目前，在人工神经网络的实际应用中，绝大部分的神经网络模型是采用BP网络和它的变化形式，它是前向网络的核心部分，并体现了人工神经网络精华的部分。

2.2.2 BP 算法的学习算法

设有 n 个神经元的任意多层BP网络，对于输入层神经元，其输入和输出相同，而隐含层神经元和输出层神经元的操作特性为^[9]：

$$net_{jk} = \sum_i w_{ji} o_{ik} \quad (2.1)$$

$$o_{jk} = f(net_{jk}) \quad (2.2)$$

其中， k 表示当前的输入样本， w_{ji} 为从神经元 i 到神经元 j 的连接权值， o_{ik} 为神经元的当前输入， o_{jk} 为其输出， f 是激励函数。BP网络通常使用的网络误差是均方差，其定义如下：

$$E_k = \frac{1}{2} \sum_j y_{jk} - o_{jk} \quad (2.3)$$

这里， E_k 是第 k 个输出响应矢量的误差； y_{jk} 是第 j 个输出神经元的期望值； o_{jk} 是第 j 个输出神经元的实际值。设 $E = \sum E_k$ 为整个训练集中所有样本产生的输出误差之和， $\Delta_k W_k$

为 w_{ji} 的调整值，并令

$$\Delta_k W_{ji} = \left\{ -\frac{\partial E}{\partial W_{ji}} \right\} \quad (2.4)$$

由式(2.4)有：

$$\frac{\partial net_{jk}}{\partial W_{ji}} = \frac{\partial}{\partial W_{ji}} \left\{ \sum_b W_{jb} O_{bk} \right\} \quad (2.5)$$

及

$$\frac{\partial E_k}{\partial W_{ji}} = \frac{\partial E_k}{\partial net_{jk}} * \frac{\partial net_{jk}}{\partial W_{ji}} \quad (2.6)$$

若定义误差符号 δ_{jk} 为：

$$\delta_{jk} = -\frac{\partial E_k}{\partial net_{jk}} \quad (2.7)$$

于是有：

$$\frac{\partial E_k}{\partial W_{ji}} = -\delta_{jk} O_{ik} \quad (2.8)$$

这就是说，要使 E 按梯度下降，就必须按下式进行权值调整：

$$\Delta_k W_{ji} = \eta \delta_{jk} O_{ik} \quad (2.9)$$

真正的梯度下降是沿着梯度确定的方向以无穷小步长进行的。很明显，这是不切实际的。因此定义 η 常数为学习速率，用它控制在权值空间中权值对应每步沿梯度方向变化的大小。其中心思想就是选择足够大的 η ，使得网络迅速收敛，而不会因调整过度而振荡。BP算法采用输出层产生的误差反向传播到输入层来计算靠。为了得到可用的微分方程，暂不考虑 δ_{jk} 的变化，运用偏微分公式：

$$\delta_{jk} = -\frac{\partial E_k}{\partial net_{jk}} = -\frac{\partial E_k}{\partial O_{jk}} * \frac{\partial net_{jk}}{\partial W_{ji}} \quad (2.10)$$

由式(2.2)有

$$\frac{\partial O_{jk}}{\partial net_{jk}} = f' net_{jk} \quad (2.11)$$

而对式(2.10)中第一项 $\frac{\partial E_k}{\partial O_{jk}}$ 必须分两种情况

(1) 若 j 为网络的输出神经元，由 E_k 的定义可知

$$\frac{\partial E_k}{\partial O_{jk}} = -y_{jk} - O_{jk} \quad (2.12)$$

这样有：

$$\delta_{jk} = y_{jk} - O_{jk} f' net_{jk} \quad (2.13)$$

(2) 若 j 是隐含层神经元，则有：

$$\sum_m \frac{\partial E_k}{\partial net_{mk}} \cdot \frac{\partial net_{mk}}{\partial O_{jk}} = \sum_m \frac{\partial E_k}{\partial net_{mk}} \cdot \frac{\partial}{\partial O_{jk}} \left\{ \sum_i W_{mi} O_{ik} \right\} = - \sum_m \delta_{mk} W_{mj} \quad (2.14)$$

将结果带入式 (2.10) 有:

$$\delta_{jk} = f' net_{jk} \sum_m \delta_{mk} W_{mj} \quad (2.15)$$

概括上述结果, 式 (2.10) 给出了关于如的微分方程, 它对隐含层和输出层都有效。式 (2.13) 和式 (2.15) 分别是对应输出层和隐含层权值如的表达式。为了得到适于数字计算机的微分方程, 现仅计算 $f'(net_{jk})$, 利用S型函数则得:

$$O_{jk} = f net_{jk} = \frac{1}{1 + \exp -net_{jk} + \theta} \quad (2.16)$$

易知:

$$f' net_{jk} = \frac{1}{1 + \exp -net_{jk} + \theta} \left\{ 1 - \frac{1}{1 + \exp -net_{jk} + \theta} \right\} \quad (2.17)$$

将 (2.16) (2.17) 代入 $f'(net_{jk})$ 表示为 O_{jk} 的函数:

$$f' net_{jk} = O_{jk} (1 - O_{jk}) \quad (2.18)$$

将式 (2.18) 代入式 (2.13), (2.15) 可得在计算机上用BP算法对网络训练所需的微方程^[10]。

基于以上原理, 加上BP网络的算法到现在已经得到了很成熟的发展, 有很多改进算法, 因此选用BP网络来作为图片的分类器是可行的。

2.2.3 改进的 BP 神经网络

标准的最速下降算法在实际应用中往往有收敛速度慢的缺点^[11]。针对标准 BP 算法的不足, 出现了几种标准 BP 算法的改进, 如动量算法, 牛顿法等。

1. 动量 BP 法

动量 BP 法是在标准的 BP 算法的权值更新阶段引入动量因子 $\alpha \in (0, 1)$, 使权值修订具有一定惯性:

$$\Delta \omega_{n-1} = -\eta (1 - \alpha) \nabla e_n + \alpha \Delta \omega_{n-1} \quad (2.19)$$

与标准的最速下降 BP 算法相比, 更新权值时, 上式多了一个因式 $\alpha \Delta \omega_{n-1}$ 。它表示, 本次权值的更新方向和幅度不但与本次计算所得的幅度有关, 还与上一次更新的方向和幅度有关。这一因式的加入, 使权值的更新具有一定的惯性, 且具有一定的抗震荡能力, 和加快收敛的能力。原理如下:

- (1) 如果前后两次计算所得的梯度方向相同, 则按标准 BP 法, 两次权值的更新方向相同。在上述公式中, 表示本次梯度反方向的 $-\eta(1 - \alpha)\nabla e(n)$ 项与上一次

的更新方向相加，得到的权值较大，可以加速收敛过程，不至于在梯度方向单一的位置停留过久。

- (2) 如果前后两次计算所得的梯度方向相反，则说明两个位置之间可能存在一个极小值。此时应减少权值修改量，防止产生能振荡。标准的最速下降法采用固定大小的学习率，无法根据情况调整学习率的值。在动量 BP 法中，由于本次梯度的反方向 $-\eta(1-\alpha)\nabla e(n)$ 与上次权值更新方向相反，其幅度会被 $\alpha\Delta\omega(n-1)$ 抵消一部分，得到一个较小的步长，更容易找到最小值点，而不会陷入来回振荡。具体应用时动量因子一般取 0.1~0.8。

2. 学习率可变的 BP 算法

在标准的最速下降 BP 算法中，学习率是一个常数，因此学习率的选择对于性能的影响巨大。如果学习率过小，则收敛速度慢；如果学习率过大，则容易出现振荡。对于不同的问题，只能通过经验来大致确定学习率。事实上，在训练的不同阶段，需要的学习率的值是不同的，如方向较为单一的区域，可选用较大的学习率，在“山谷”附近，应选择较小的学习率。如果能自适应的判别出不同的情况，调整学习率的值，将会提高算法的性能和稳定性。

学习率可变的可变的 BP 算法（Variable Learning Rate Back propagation VLBP）是通过观察误差的增减来判断的。当误差以减小的方向方式趋于目标时，说明修正方向是正确的，可以增加学习效率；当误差增加超过一定范围时，说明前一步修正进行的不正确，应减小步长，并撤销前一步修正过程。学习率的增减通过乘以一个增量/减量因子实现：

$$\eta_{n+1} = \begin{cases} K_{inc}\eta_n & e_{n+1} < e_n \\ K_{dec}\eta_n & e_{n+1} > e_n \end{cases} \quad (2.20)$$

3. 拟牛顿法

牛顿法是一种基于泰勒级数展开的快速优化算法。迭代公式如下：

$$\omega_{n+1} = \omega_n - H^{-1} \nabla g_n \quad (2.21)$$

H 为误差性能函数的 Hessian 矩阵，其中包含了误差函数的导数信息。例如，对于一个二元可微函数 $f(x, y)$ ，其 Hessian 矩阵为：

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \quad (2.22)$$

牛

顿法具有收敛快的优点，但需要计算误差性能的二阶导数，计算较为复杂，如果 Hessian 矩阵非正定，可能导致搜索方向不是函数下降方向。因此提出了改进算法，用一个不包含二阶导数的矩阵近似 Hessian 矩阵的逆矩阵，此方法为拟牛顿法。

拟牛顿法只需要知道目标函数的梯度，通过测量梯度的变化进行迭代，收敛速度大大优于最速下降法。

4. LM (Levenberg-Marquardt) 算法

LM 算法类似牛顿算法，都是为了在修正速率时避免计算 Hessian 矩阵而设计的。当误差函数具有平方和误差的形式时，Hessian 矩阵可近似表示为：

$$H = J^T J \quad (2.23)$$

梯度可表示为：

$$g = J^T e \quad (2.24)$$

J 是包含误差性能函数对网络权值一阶导数的雅可比矩阵。LM 算法根据下式修正网络权值：

$$\omega_{n+1} = \omega_n - [J^T J + \mu J]^{-1} J^T e \quad (2.25)$$

当 $\mu = 0$ 时，LM 算法退化为牛顿算法：当 μ 很大时，上式相当于步长较小的梯度下降算法，由于雅可比矩阵的计算比 Hessian 矩阵易于计算，因此速度非常快，但是 LM 算法在 MATLAB 仿真时是一个比较耗费内存的算法。

第三章 基于 BP 神经网络的图像分类算法

3.1 引言

对人类来说简要浏览一下图片就足以描述出看到的大量细节。然而，这种非凡的能力对视觉识别模型来说却是一项困难的工作^[25]。图像分类技术是根据图像的底层视觉特征将图像分门别类到预定义的类中，它是实现计算机自动获取图像语义的重要途径。它的主要流程包括图像预处理阶段、特征提取阶段、分类器设计与学习阶段。图像分类效果的好坏主要取决于两个因素，特征提取和分类器的设计。特征提取的好坏关键在于特征提取的算法设计，好的特征提取算法可以是分类达到很好的分类的效果，目前主要的特征提取包括：颜色特征，文理特征，不变性特征，这些特征的都从一定程度上反映了图像的内容信息，另外一个方面，分类器的设计与学习是图像分类技术的核心，它帮助人们在图像特征与关键字类别之间建立准确的映射关系，提出与人们认知一致图像语义信息。

3.2 算法描述

3.2.1 算法流程描述

Step1: 数据预处理：提取图像特征，产生对应标签

Step2: 初始化权系数及各个网络参数(学习速率，迭代次数，允许误差，隐含层节点数等)。文中网络参数的初始化见表 3-1

表 3-1 网络参数设置

学习率	允许误差	隐含层节点数	最大迭代次数
0.1	1E-4	14	10000

Step3: 输入样本， X^k ， $n=0$ 。

Step4: $Y_k(n)$ 为第 n 次迭代时网络的实际输出。由期望输出 T_k 和实际输出 $Y_k(n)$ 计算出误差 $E(n)$ ，判断其是否满足要求，若满足转至 Step6，不满足转至 Step5。

Step5: 判断 $n+1$ 是否大于最大迭代次数，若大于转至 Step6，若不大于，对输入样本 X^k ，反向计算每层的神经元的局部梯度 Dn 。

Step6: 判断是否学完所有的训练样本，是转至 Step7，否则转至 Step3。

Step7: 输入测试样本进行网络仿真，得到对应的测试样本的标签。

Step8: 将 Step7 中输出标签与已知标签比较，得出分类正确率。

3.2.2 系统实现框图

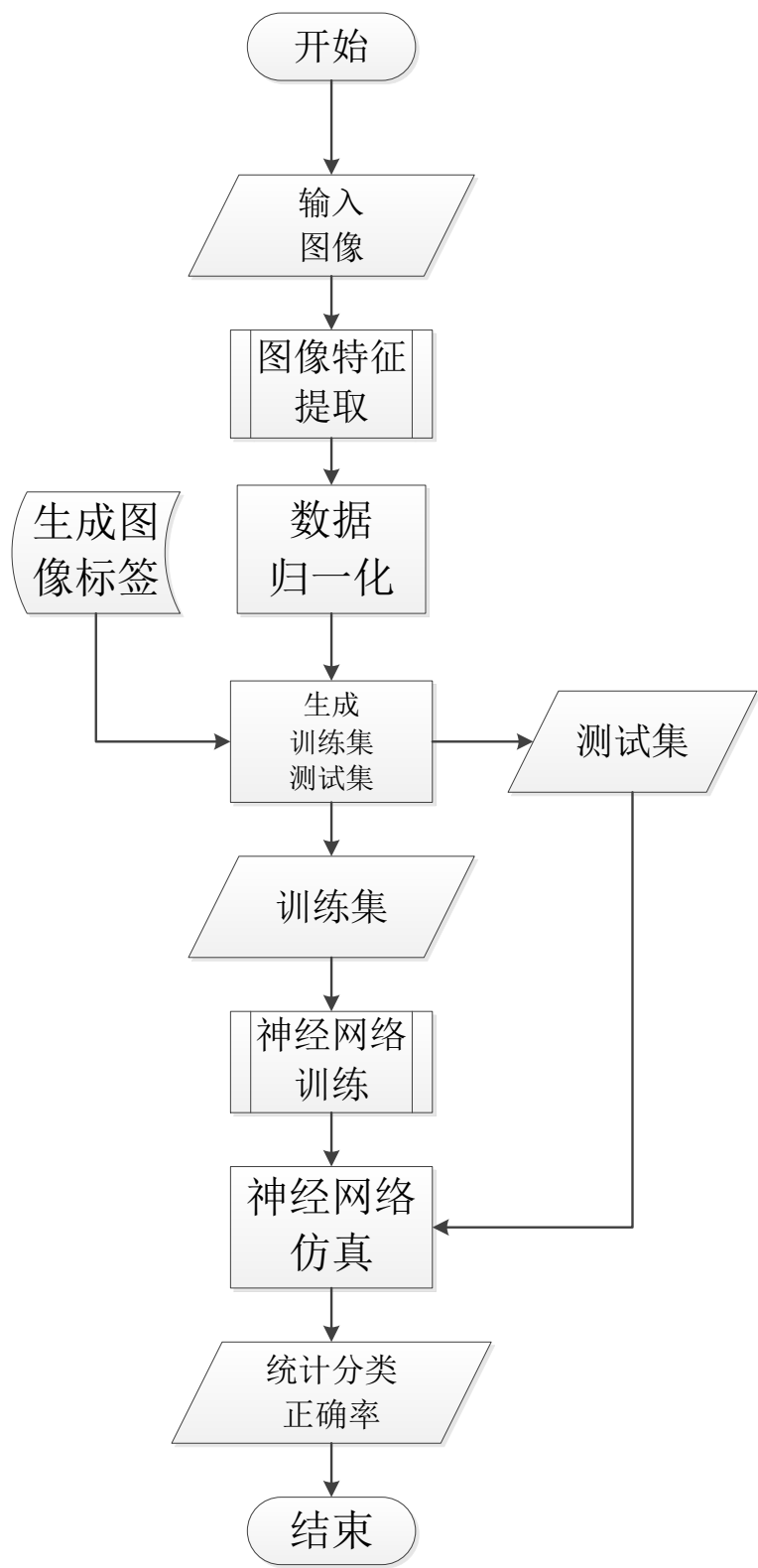


图 3-1 系统实现流程图

图 3-1 是系统实现流程图，首先对数据集图像进行特征提取，然后进行数据归一化并产生标签，将数据集和对应标签一起分成训练集和测试集，将训练在神经网络中训练，用训练好的网络对测试集进行测试，最后统计分析分类正确率。详细的各模块实现见以下各小节。

3.3 图像的特征提取

3.3.1 特征提取流程图

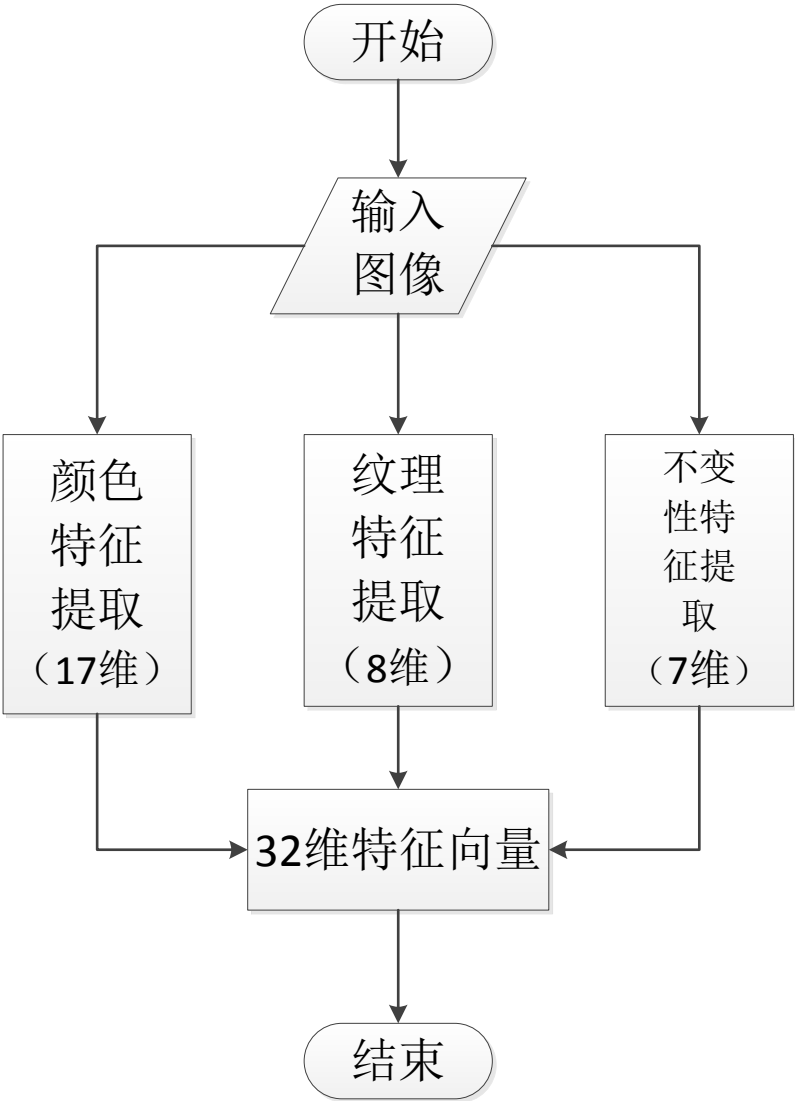


图 3-2 特征提取流程图

如图 3-2 是图像特征提取的颜色、纹理、不变性，每幅图像用一个 32 维的特征列向量表示，其中颜色特征 17 维，纹理特征 8 维，不变性特征 7 维，此 32 维特征向量是神经网络的输入。具体的特征提取原理见 3.3.2 节（颜色特征提取），3.2.3（纹理特征提取）和 3.3.4 节（不变性特征提取）。

3.3.2 颜色特征的提取

图像分类的一个关键在于特征提取算法。在一些算法中提取复杂性比较高的各种特征可能能够解决用于目标检测目的的问题，但是需要处理更多的数据，从而处理的效果成本很高。在基于内容的图像搜索的过程中，颜色特征是应用范围最广的一个特征，它有着简单并且计算量比较底的性质，颜色特征处理只用把数字图像所包含的像素值与其对应的数值相转换，所以具有低复杂度的颜色便成为了特征提取中一个较好的特征。在图像处理中，我们可以对特定像素的颜色以不同方法进行分析，提取出其颜色特征分量。

在众多的图像检索系统里，颜色直方图是其中应用极为广泛的颜色特征之一，其内容是图像里各种色彩的组成分布，也就是各种色彩出现在图像中的频率。最先提出通过颜色直方图来对图像进行特征提取的是Swain和Ballard^[26]，他们对色彩特征空间里的三个颜色分量进行分离以后，获得图像的颜色直方图，图像在经过了多次翻转、尺寸变换以及虚化等物理变换以后，其颜色直方图的变化并不明显，即其具有较好的稳定性，因此往往提取幅图像的颜色特征并利用颜色直方图来对比其全局差。颜色直方图里的数据一般通过统计得到，它用来表示图像中颜色的数量特征，但其只表明图像里某一颜色的出现频率却忽略了它的空间位置信息。因此，对于任意一副图像，均可以给出其相应的直方图，然而相同的直方图却有可能与多幅不同图像相对应。由于以上缺点本文并未采取此方法来提取图像的颜色特征，而是提取了图像的主要颜色，和各颜色分量的1, 2, 3阶矩，以下是本文所用颜色特征的提取方法：

(1) 图像主要颜色的提取

图像主要有8种颜色，黑色、白色、红色、黄色、绿色、青色、蓝色和紫色。每个像素点对应成这8种颜色中的一种。用 $C_1, C_2, C_3, \dots, C_8$ 分别代表8种颜色所对应像素点的个数。每个图像能得到一组8维的向量 $\{C_1/C, C_2/C, C_3/C, \dots, C_8/C\}$ (其中 C 为相应图像中总的像素点的个数)。由于一般认为RGB颜色空间与人眼的感知差异很大，HSV颜色空间能够比较好地模拟人的颜色感知，所以本文采用HSV颜色空间来提取其颜色特征。程序读入的图像是RGB图像，首先将每一个像素的 R, G, B ，转换成 H, S, V (其中 $H \in [0, 2\pi], S \in [0, 1], V \in [0, 1]$)，然后根据HSV颜色模型做如下的特殊处理，当 V 的平均值小于0.2时，颜色为黑色 C_1 ，当 S 的平均值小于0.15时，且 V 的平均值大于0.8时，对应颜色为白色 C_2 。在其他情况，按照公式(3.1)把色调 H 空间分成6份，分别代表红 C_3 ，黄 C_4 ，绿 C_5 ，青 C_6 ，蓝 C_7 ，紫色 C_8 。

$$H = \begin{cases} 0,22 & \& 330,360 & C_3 \\ 22,70 & & C_4 \\ 70,155 & & C_5 \\ 155,200 & & C_6 \\ 200,270 & & C_7 \\ 270,330 & & C_8 \end{cases} \quad (3.1)$$

对每张图像进行处理得到8维的列向量:

$$C_{color} = C1, C2, C3, \dots, C8^T \quad (3.2)$$

用此 8 维列向量表示每张图像的主要颜色特征。

(2) 颜色矩的提取

从图像中提取颜色特征时, 大多数方法都先要对图像进行量化处理。量化处理容易导致误检, 并且产生的图像特征维数较高, 不利于检索。Stricker和Orengo提出了颜色矩的方法^[27], 颜色矩是一种简单有效的颜色特征表示方法, 有一阶矩(均值, mean)、二阶矩(方差, variance)和三阶矩(斜度, skewness)等, 由于颜色信息主要分布于低阶矩中, 所以用一阶矩, 二阶矩和三阶矩足以表达图像的颜色分布, 颜色矩已证明可有效地表示图像中的颜色分布, 该方法的优点在于: 不需要颜色空间量化, 特征向量维数低, 适合应用于BP神经网络的图像分类中。三个颜色矩的数学定义如下:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N P_{i,j} \quad (3.3)$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N P_{i,j}^2 - \mu_i^2 \right)^{\frac{1}{2}} \quad (3.4)$$

$$S_i = \left(\frac{1}{N} \sum_{j=1}^N P_{i,j}^3 - \mu_i^3 \right)^{\frac{1}{3}} \quad (3.5)$$

其中, $P_{i,j}$ 表示彩色的第*i*个颜色通道分量中灰度为*j*像素出现的概率, *N*表示图像中的像素个数。图像的3个分量H, S, V图像的钱三阶颜色矩组成一个9维的特征向量, 即:

$$F_{color} = \mu_H, \delta_H, S_H, \mu_S, \delta_S, S_S, \mu_V, \delta_V, S_V^T \quad (3.6)$$

综合 (1) 和 (2) 每张图片的颜色特征提取到的特征向量为一个17维的向量:

$$Im_{color} = C_{color}, F_{color} \quad (3.7)$$

3.3.3 纹理特征的提取

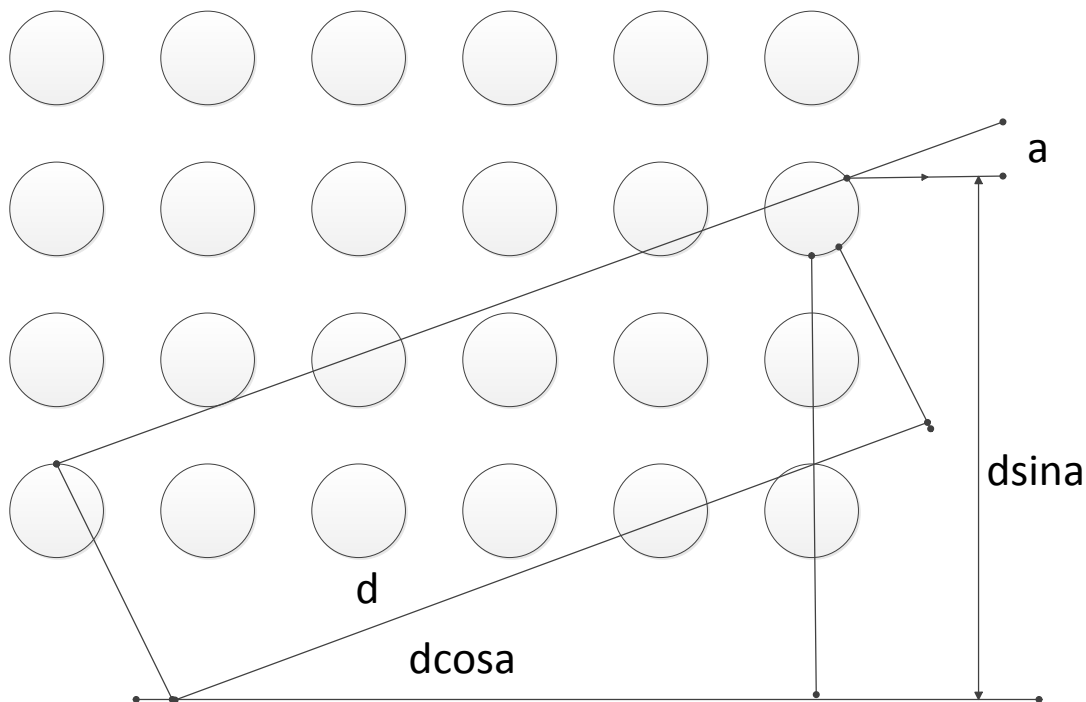
图像的纹理有时候被称为结构、影纹等。Pickett 认为纹理是保持一定的特征重复性并且间隔规律可以任意安排的空间结构; HawKins 认为纹理具有三大标志: 某种局部序列性不断重复、非随机排列和纹理区域内大致为均匀的统一体; LiWang 和 D. C. He 认为, 纹理是由纹理基元组成的, 纹理基元被认为是表现纹理特征的最小单元, 是一个像元在其周围 8 个方向上的特征的反应。纹理特征有时是明显的, 以某种基本图形在某一区域有规律地周期性出现; 而有时文理特征是不明显的、隐晦的, 具有不稳定性。

纹理作为一种区域特征, 是对图像各像元之间空间分布的一种描述。由于文理充分利用图像信息, 无论从理论上还是常识出发它都可以成为描述与识别图像的重要依据,

与其他的特征相比较，它能更好地兼顾图像宏观性质与细微结构两个方面，因此纹理成为目标识别需要提取的重要特征。

提取纹理特征的方法有很多，如基于局部统计的特征、基于随机场模型的特征、基于空间频率的特征、分型特征和基于灰度共生矩阵的特征。其中应用最广泛的是基于灰度共生矩阵的特征。本文所采用的纹理提取方法即为基于灰度共生矩阵的特征，下面介绍灰度共生矩阵的提取方法：

在图像中任取一点 (x, y) 以及偏离它的另一点 $(x + a, y + b)$ ，形成一个点对，设该点对的灰度值为 (i, j) ，即点 (x, y) 的灰度为 i ，点 $(x + a, y + b)$ 的灰度为 j 。固定点 a 与点 b ，令点 (x, y) 在整幅图上移动，则会得到各种 (i, j) 的值。设灰度值的级数为 L ，则点 i 与 j 的组合共有 L^2 种。在整幅图像中，统计出每一种 (i, j) 值出现的次数，再将它们归一化为出现的概率 P_{ij} ，则称方阵 $[P_{ij}]_{L \times L}$ 为灰度联合概率矩阵，也称灰度共生矩阵。灰度共生矩阵实际上就是两个像素点的联合直方图，其形成过程如图 3-1 所示：



图

3-3 灰度共生矩阵的形成过程

距离差分值 (a, b) 取不同的数值组合，可以得到沿一定方向（如： 0° 、 45° 、 90° 、 135° ）相隔一定距离 $d = \sqrt{a^2 + b^2}$ 的像元之间灰度联合矩阵。 a 和 b 的取值要根据纹理周期分布的特性来选择。当 a 和 b 取值较小时，对应于变化缓慢的纹理图像，其灰度联合概率矩阵对角线上的数值较大，倾向于做对角线分布；若纹理的变化较

快，则对角线上的数值较小，倾向于均匀分布。反映矩阵状况的参数有能量、相关、熵和反差。

1. 能量

灰度共生矩阵元素值的平方和，也称能量。它反映了图像灰度分布的均匀程度和纹理粗细度。能量的数学定义如下：

$$f_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{ij}^2 \quad (3.8)$$

当 P_{ij} 数值分布比较集中时，能量较大；当 P_{ij} 数值分布比较分散时，能量较小。

2. 相关

相关是度量空间灰度共生矩阵元素在行或列方向上的相似程度，因此，相关值大小反映了图像中局部灰度相关性。当矩阵元素值均匀相等时，相关值就越大；相反就越小。相关的数学定义如下：

$$f_2 = \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - \mu_x)(j - \mu_y) P_{ij} \quad (3.9)$$

式中， μ_x ， μ_y ， σ_x ， σ_y ，分别为： $m_x = \sum_{j=0}^{L-1} P_{ij}$ 与 $m_y = \sum_{i=0}^{L-1} P_{ij}$ 的均值和标准差。

$$\mu_x = \sum_{i=0}^{L-1} i \sum_{j=0}^{L-1} P_{ij} \quad (3.10)$$

$$\mu_y = \sum_{j=0}^{L-1} j \sum_{i=0}^{L-1} P_{ij} \quad (3.11)$$

$$\sigma_x^2 = \sum_{i=0}^{L-1} (i - \mu_x)^2 \sum_{j=0}^{L-1} P_{ij} \quad (3.12)$$

$$\sigma_y^2 = \sum_{j=0}^{L-1} (j - \mu_y)^2 \sum_{i=0}^{L-1} P_{ij} \quad (3.13)$$

3. 熵

熵是图像所具有的信息的度量，纹理信息也属于图像的信息，是一个随机性的度量，它表示了图像中纹理的非均匀或复杂程度。熵的数学表达式如下：

$$f_3 = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{ij} \log_2 P_{ij} \quad (3.14)$$

若 P_{ij} 为均匀分布则，熵较大；反之， P_{ij} 数值分布比较集中时，则熵较小。

4. 反差

反差又称主对角线惯性矩，其数学表达式如下：

$$f_4 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} |i-j|^k P_{ij} \quad (3.15)$$

对于粗纹理， P_{ij} 的数值集中于主对角线附近，此时 $|i-j|$ 的值较小，所以反差也越小。对于细纹理，反差较大。

本文采用的方法是在 $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ 情况下计算四个共生矩阵，然后求能量、熵、相关和反差的均值和标准差共 8 维作为最终的纹理特征：

$$Im_{Texture} = [T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8]^T \quad (3.16)$$

$T(1) \dots T(8)$ 分别为能量、熵、相关和反差的均值和标准差。

3.3.4 不变性特征的提取

图像分析中的一个关键问题就是如何获得一种有效的图像描述量，用一个很小的数据集就可以代表图像^[13]。不变矩（Invariant Moments）是一种高浓度的图像特征，具有平移，灰度，尺度，旋转不变性，因此矩和矩函数被广泛应用于图像的模式识别、图像分类、目标识别和场景分析中。

矩在统计学中表征随机量的分布，一幅灰度图像可以用二维灰度密度函数来表示，因此可以用矩描述灰度图像的特征。

一幅 $M \times N$ 的数字图像 $f(i, j)$ ，其 $p+q$ 阶几何矩 m_{pq} 和中心矩 μ_{pq} 为：

$$m_{pq} = \sum_{i=1}^M \sum_{j=1}^N i^p j^q f(i, j) \quad p, q = 0, 1, 2 \dots \quad (3.17)$$

$$\mu_{pq} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^p (j - \bar{j})^q f(i, j) \quad p, q = 0, 1, 2 \dots \quad (3.18)$$

式中： $\bar{i} = \frac{m_{10}}{m_{00}}$; $\bar{j} = \frac{m_{01}}{m_{00}}$ 。

若将 m_{00} 看作是图像的灰度质量，则 (\bar{i}, \bar{j}) 为图像灰度质心坐标，那么，中心矩 μ_{pq} 反映的是图像的灰度相对于其灰度质心的分布情况。可以用几何矩来表示中心矩，0~2 阶中心矩与几何矩的关系如下。

$$\mu_{00} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^0 (j - \bar{j})^0 f(i, j) = m_{00} \quad (3.19)$$

$$\mu_{10} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^1 (j - \bar{j})^0 f(i, j) = 0 \quad (3.20)$$

$$\mu_{01} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^0 (j - \bar{j})^1 f(i, j) = 0 \quad (3.21)$$

$$\mu_{11} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^1 (j - \bar{j})^1 f(i, j) = m_{11} - \bar{i}m_{10} \quad (3.22)$$

$$\mu_{20} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^2 (j - \bar{j})^0 f(i, j) = m_{20} - \bar{x}m_{10} \quad (3.23)$$

$$\mu_{02} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^0 (j - \bar{j})^2 f(i, j) = m_{02} - \bar{y}m_{01} \quad (3.24)$$

为了消除图像比例变化带来的影响，定义规格化中心矩如下：

$$\mu_{02} = \sum_{i=1}^M \sum_{j=1}^N (i - \bar{i})^0 (j - \bar{j})^2 f(i, j) = m_{02} - \bar{y}m_{01} \quad (3.25)$$

利用二阶和三阶规格中心矩可以导出下面七个不变矩组（ $\Phi_1 \sim \Phi_7$ ），它们对图像平移、旋转和比例变化保持不变：

$$\Phi_1 = \eta_{20} + \eta_{02} \quad (3.26)$$

$$\Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (3.27)$$

$$\Phi_3 = \eta_{20} - 3\eta_{12}^2 + 3\eta_{21} - \eta_{03}^2 \quad (3.28)$$

$$\Phi_4 = \eta_{30} + \eta_{12}^2 + \eta_{21} + \eta_{03}^2 \quad (3.29)$$

$$\begin{aligned} \Phi_5 = & \eta_{30} - 3\eta_{12} \eta_{30} + \eta_{12} \\ & \left[\eta_{30} + \eta_{12}^2 - 3\eta_{21} + \eta_{03}^2 + 3\eta_{21} - \eta_{03} \eta_{21} + \eta_{03} \right] \\ & \left[3\eta_{30} + \eta_{12}^2 - \eta_{21} + \eta_{03}^2 \right] \end{aligned} \quad (3.30)$$

$$\begin{aligned} \Phi_6 = & \eta_{20} - \eta_{02} \left[\eta_{30} + \eta_{12}^2 - \eta_{21} + \eta_{03}^2 \right] \\ & + 4\eta_{11} \eta_{30} + \eta_{12} \eta_{21} + \eta_{03} \end{aligned} \quad (3.31)$$

$$\begin{aligned} \Phi_7 = & 3\eta_{21} - \eta_{03} \eta_{30} + \eta_{12} \\ & \left[\eta_{30} + \eta_{12}^2 - 3\eta_{21} + \eta_{03}^2 \right] \\ & + 3\eta_{12} - \eta_{30} \eta_{21} + \eta_{03} \\ & \left[3\eta_{30} + \eta_{12}^2 - \eta_{21} + \eta_{03}^2 \right] \end{aligned} \quad (3.32)$$

本文采用以上方法求解图像的不变特征共 7 维列向量：

$$Im_{inv7} = \Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6, \Phi_7^T \quad (3.33)$$

综合以上 3.2.1、3.2.2、3.2.3 中颜色、纹理、不变性，每幅图像用一个 32 维的特征列向量表示，其中颜色特征 17 维，纹理特征 8 维，不变性特征 7 维，如式 3.35 所示：

$$Im = Im_{color}, Im_{texture}, Im_{inv7} \quad (3.34)$$

3.3.5 数据归一化处理

在训练神经网络前一般需要对数据进行预处理，一种重要的预处理手段是归一化处理。数据归一化，就是将数据映射到[0, 1]或[-1, 1]区间或更小的区间，比如

(0. 1, 0. 9) 。由于输入数据的单位不一样，有些数据的范围可能特别大，导致的结果是神经网络收敛慢、训练时间长，数据范围大的输入在模式分类中的作用可能会偏大，而数据范围小的输入作用就可能会偏小，由于神经网络输出层的激活函数的值域是有限制的，因此需要将网络训练的目标数据映射到激活函数的值域。例如神经网络的输出层若采用 S 形激活函数，由于 S 形函数的值域限制在(0, 1)，也就是说神经网络的输出只能限制在(0, 1)，所以训练数据的输出就要归一化到[0, 1]区间。基于以上原因在数据输入神经网络之前对其进行归一化是必要的。

一种简单而快速的归一化算法是线性转换算法。线性转换算法常见有两种形式：

(1)

$$y = \frac{x - \min}{\max - \min} \quad (3.35)$$

其中 x 待归一化的向量， \min 为 x 的最小值， \max 为 x 的最大值，输入向量为 x ，归一化后的输出向量为 y 。上式将数据归一化到[0, 1]区间，当激活函数采用 S 形函数时（值域为（0, 1））时这条式子适用。

(2)

$$y = 2 \frac{x - \min}{\max - \min} - 1 \quad (3.36)$$

这条公式将数据归一化到 [-1, 1] 区间。当激活函数采用双极 S 形函数（值域为(-1, 1)）时这条式子适用。

3.4 分类算法的设计

3.4.1 训练数据集

由于 BP 神经网络采用有监督的学习，因此用 BP 神经网络解决一个具体问题时，首先需要有一个训练数据集。本文采用的训练数据集是 Corel 数据集。Corel 数据集有 1k 图像，包括土著、沙滩、建筑、公交、山峰、大象、恐龙、花、马和食物，总共 10 个类别，每个类有 100 张图片。

首先对数据集进行分组，抽取每个类别的前 80 张图像作为训练数据集，后 20 张作为测试数据集，用 3.2 中的图像特征提取的方法分别对训练集和测试集进行图像的特征提取，同时生成图像对应的标签。最后生成四个文件：分别为 32×800 的矩阵， 4×800 的矩阵， 32×200 的矩阵， 4×200 的矩阵。 32×800 的矩阵每一列代表一张训练图像的特征向量，每一行代表图像的一个特征， 4×800 ，每一列代表训练图像的一个标签（四位二进制表示的 1 到 9）； 32×200 的矩阵每一列代表一张测试图像的特征向量，每一行代表图像的一个特征， 4×200 ，每一列代表测试图像的一个标签。

3.4.2 BP 神经网络的设计

在使用基于 BP 算法的人工神经网络时，必将会遇到选择基于 BP 算法人工神经网络的最佳结构问题，具体说就是，给定了某个问题，设计神经网络所需要考虑的问题有：网络层数的选择；每层神经元个数的选择；输出神经元的个数的确定；传递函数的选择，训练方法的选择等。

(1) 网络的层数选择

网络的性能好坏取决于神经网络中隐含层的层数及其神经元个数，这两个参量要在实际的测试中确定。BP 神经网络可以包含一到多个隐含层。不过理论上已经证明，单个隐含层的网络可以通过增加神经元节点数实现任意非线性映射。但是如果样本数过多，增加一个隐含层可以明显减小网络的规模。由于数据量较大，笔者电脑内存不是很充足，所以在实际的算法实现时为了减小网络的规模，防止内存溢出，本文采用的是三层的神经网络。

(2) 输入，输出层节点数

输入层节点数取决于输入向量的特征维数，由于本文采用的图像特征提取算法所提取的特征是 32 维，所以，输入层节点数是 32 维。

输出层节点数即为图像所对应的类别，由于数据设计时使用的是四位二进制表示 10 个类别，所以输出层节点数应为 4，每个节点对应四位二进制的 1 位。

(3) 隐含层节点数

隐含层节点数对 BP 神经网络的性能影响很大。一般较多的隐含层节点数可以带来更好的性能，但是可能导致训练时间过长。网络中的隐节点太少，将降低学习能力，甚至学习过程可能不收敛，不能记忆任何学习模式，训练后的网络不够强壮，不能识别未知的样本，容错性差；隐含节点太多，长时间不收敛，还会出现过拟合现象，学习后的网络对从前未见过的输入，泛化能力差^[14]。目前没有一个理想的解析式可以确定合理的神经元节点数，这也是 BP 神经网络的一个缺陷。

需要寻找一个最佳的隐层神经元个数。一般参照经验公式：

$n = \sqrt{n_i + n_o} + a$ 计算隐含层节点数，式中 n_i ， n_o ， n 分别代表输入层、输出层、隐含层神经元数， a 为 1~10 之间的常数^[15]，由公式可知，隐含层神经元数的取值范围为 6~16。为进一步确定隐含层节点数，这里首先令 $n=3$ ，并设定训练误差为 ±3%，取样本数据训练该网络，记录其达到误差精度要求时经过的进化代数，并用测试数据验证其预测准确性，然后采用同样的方法，令 n 分别取 7, 8, ..., 16，最后从中选择预测效果最好时候的 n 值。因此最终确定 $n=14$ 。

(4) 传递函数的选择

传递函数 $f(\cdot)$ ，又叫响应函数或者转移函数，其作用是激活神经元，使其对输入产生影响，将可能的无限域变换到指定的有限范围内输出。这类似于生物神经元所具有的非线性生物转移特性。

一般来说，神经元对来自外界刺激的响应函数有以下三种：

1) 阈值型：如图 3-2 (a) 所示：

$$f(x_j) = \begin{cases} 1 & x_j > 0 \\ 0 & x_j < 0 \end{cases} \quad (3.37)$$

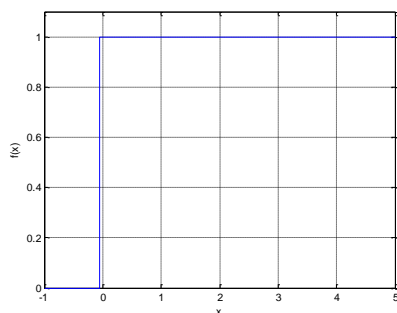
2) 线性：如图 3-2 (b) 所示：

$$f(x_j) = x_j \quad (3.38)$$

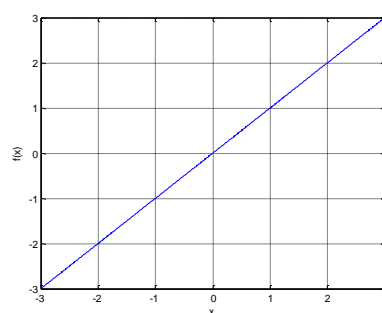
3) Sigmoid 函数：相应函数为光滑曲线函数，常用的有 logsig (3.39) 和 tansig (3.40)，如图 3-2 (c) 和 (d) 所示：

$$f(x_j) = \frac{1}{1 + \exp(-x_j)} \quad (3.39)$$

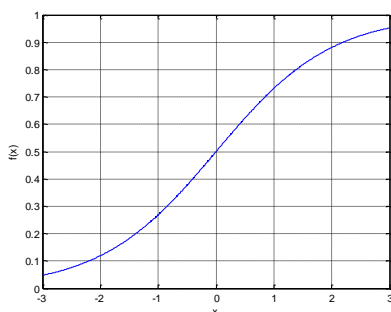
$$f(x_j) = \frac{2}{[1 + \exp(-2x_j)] - 1} \quad (3.40)$$



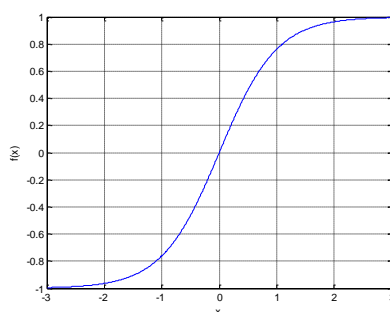
(a) 阶跃函数



(b) 线性函数



(c) logsig 函数



(d) tansig 函数

图 3-4 传递函数

本文采用的传输函数为，隐含层 logsig ，为了保证输出端输出的值在 (0, 1) 之间，输出层采用纯线性函数 purelin 。

(5) 初始权值的选择

初始权值的选择直接影响网络训练时间。如果每个节点的净输入均在零值附近，对单极性 Sigmoid 函数来说，那么不论输入如何，网络初始学习阶段的速度将是很快的。神经元要尽力避免工作在 Sigmoid 函数的饱和区。因为函数在该区的导数值很小，致使节权值在多隐节点的情况下，如果同一输入节点连至各隐节点的权值完

全相同的话，那么每个隐含层节点的输出也将完全相同。致使它们对在输出层造成的误差应承担的“责任”也完全相同。其结果是各权值的修改量也完全相同。这样，训练出的节点的权值总是相同的^[15]。

为此，节点的权值要有所不同。应采用随机数，且其值还要足够的小，以确保在任何输入模式作用下，任意隐含层节点的初始净输入都能接近零。

综上所述，初始权值按以下方式确定：

置隐含层节点的初始值均为分布在零附近的很小的随机值；

置输出层节点所连的权值数的一般为+1，另一半为-1。

网络节点的偏置统一设置为零。

(6) 训练算法的选择

BP 神经网络学习规则的指导思想是：对网络权值和阈值的修正要沿着表现函数下降最快的方向一负方向。

$$x_{k+1} = x_k - a_k g_k \quad (3.41)$$

其中 x_k 是当前的权值和阈值矩阵， g_k 是当前表现函数的梯度， a_k 是学习速率。

传统的 BP 算法本质上是最快速下降的，因而也具有最快速下降法的固有缺点，即学习效率低，收敛速度慢，容易陷入局部极小状态。即使是初始化较好，且采用最优学习率，仍不能完全解决这个问题。特别是当训练接近收敛点时，误差减小的速度将变得非常缓慢。

为此，出现了许多改进算法。BP 算法的改进主要有两种途径；一种是采用启发式学习算法，另一种则是采用更有效的优化算法。

改进的 BP 神经网络算法在第二章原理部分的 2.2.3 节已经详细介绍，这里不再赘述。本文选取的训练算法是 LM 算法，该算法的优点是训练所需时间较短。

(7) 训练参数的选择^{[16][17]}

BP 算法中几个常用的参数，包括学习率 η ，动量因子 α 及收敛误差界值 E_{min} 等，这些参数对训练速度的影响很关键。

学习率影响系统学习过程的稳定性。大的学习率可能使网络权值每一次的修正量过大，甚至会导致权值在修正过程中超出某个误差的极小值呈不规则跳跃而不收敛；但过小的学习率导致学习时间过长，不过能保证收敛于某个极小值。所以，一般倾向选取较小的学习率以保证学习过程的收敛性(稳定性)，通常在 0.01---0.8 之间。实践中，取该范围中较小的学习速率以保持系统稳定，通过观察误差下降曲线来判断，下降较快说明学习速率比较合适，若有较大振荡则说明学习速率偏大。增加动量项的目的是为了避免网络训练陷于较浅的局部极小点。理论上动量因子口取值大小与权值修正量的大小有关，但实际应用中一般取常量。通常在 0~1 之间，而且一般比学习率要大。

本研究对 η 和 α 的最终取值视实际情况而定。

在网络训练过程中应根据实际情况预先确定误差界值，误差界值的选择完全根据网络模型的收敛速度大小和具体样本的学习精度来确定。当 E_{min} 值选择较小时，学习效果好，但收敛速度慢，训练次数增加。如果 E_{rain} 值较大则相反。

3.4.3 BP 神经网络的训练

给定 p 组学习样本 $\{x_1, x_2, \dots, x_n\}^k$, $k = 1, 2, \dots, p$, 及与之相应的 p 个期望输出序列 $\{T_1, T_2, \dots, T_n\}^k$, $k = 1, 2, \dots, p$, 在对网络进行训练前, 先将连接权及阈值赋初值, 然后将第一组学习样本:

$$\{x_1, x_2, \dots, x_n\}^1$$

以及相应的期望输出序列:

$$\{T_1, T_2, \dots, T_n\}^1$$

输入网络, 计算网络输出:

$$\{y_1, y_2, \dots, y_n\}^1$$

然后, 将余下的 $p-1$ 组学习样本相应的学习样本及期望输出序列依次输入神经网络, 计算出相应的网络输出。

定义网络输出与期望输出的误差为:

$$E = \sum_{j=1}^p \sum_{i=1}^n T_{ij} - y_{ij}^2 \quad (3.42)$$

其中, P 为学习样本数。

训练的目就是要通过调节网络内部的连接权值和阈值, 使网络输出与期望输出的误差达到最小。

本文中学习样本 P 为 32×800 的矩阵, 每一列代表一个学习样本, $\{x_1, x_2, \dots, x_n\}^k$, 此处 $n=32$, $p=800$;

第四章 算法仿真实现及结果分析

神经网络仿真是算法测试的工程，本部分主要针对第三章的分类算法设计进行实现，测试不同参数对分类效果的影响，用与训练算法不同的数据集测试算法，检测算法的性能，此算法的优势在于对与训练集大小不同的图像也有很好的分类效果，既系统有很强的鲁棒性。

4.1 算法的仿真实现

4.1.1 BP 神经网络的仿真

神经网络的训练完成以后，即可进行网络仿真，也就是回忆操作。神经网络仿真，即对网络输入向量 X ，使网络输出系统输出系列 Y 的近似值序列 Y' 。从数学角度来看，神经网络是一组输入单元到一组输出单元的映射。若网络输入层有 n 个单元，输出层有 m 个单元，即有如下映射：

$$y = F(x) \quad (4.1)$$

式中： $\{x\}$ --输入向量；

$\{y\}$ --输出向量；

F --从输入层到输出层的映射。

对于输入样本 $\{x\}$ 和输出 $\{y\}$ ，可认为存在映射 G ，使得下式成立：

$$y_i = G(x_i) \quad i=1,2,\dots, m \quad (4.2)$$

利用神经网络求得映射 G ，使得在某种意义上 G 是 F 的最佳逼近。Hechi—Nielson 在 Kolmogorov 连续函数表示定理的基础上已经证明，只要采用有隐含层的网络，那么这种映射就是一定存在的。

本文的仿真输入是测试样本数据集即 32×200 的矩阵。

4.1.2 实验数据与实验环境

本实验的软件仿真环境为在 windows 7 下安装的 matlab 2012Rb。实验平台为 Intel Core i5 处理器，2GB 内存，32 位操作系统的旗舰版 win7，图像来源于 Corel 公司发布的权威图像数据库 Corel，图像中的 10 类，每类各 100 幅分辨率为 384×256 的彩色图片，格式为 JPEG，共 1000 幅图像。各类图像如图 4.1 示，采用本文中 3.2 节介绍的方法对图像进行处理，分别提取图像的颜色、纹理、不变性，每幅图像用一个 32 维的特征列向量表示，其中颜色特征 17 维，纹理特征 8 维，不变性特征 7 维，最后生成一个包含所有训练集的 .MAT 文件，包括图像的特征向量和对应标签。实验所用 BP 神经网络为三隐含层，五层神经网络，层神经元与输出神经元数目分别为 16 和 4。

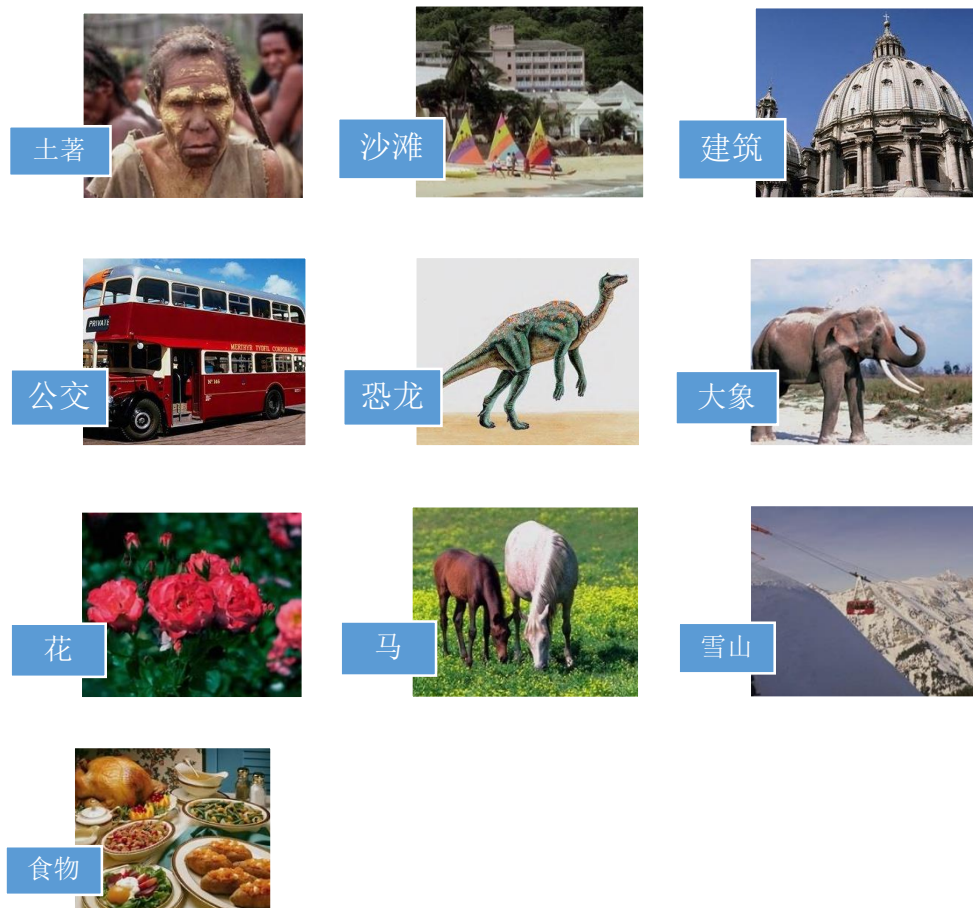


图 4-1 corel 数据集图像类别

4.1.3 不同参数下的结果分析

本节进行在不同参数设置下分类效果，均方误差和回归性能分析。

(一) 参数设置一：

BP神经网络参数设置最大迭代次数：100000，目标性能(均方误差性能)：0.0001，学习率：0.1，验证数据的最大允许错误样本数：10。训练函数为采用常规的最速下降算法训练算法。

(1) 分类效果：

在以上参数设置下训练用时16分58秒，迭代次数达到设置的允许最大迭代次数100000，测试集共200张图像，每个类别20张图像，总的分类正确率只有39%，各个类别的分类结果如图4.2。

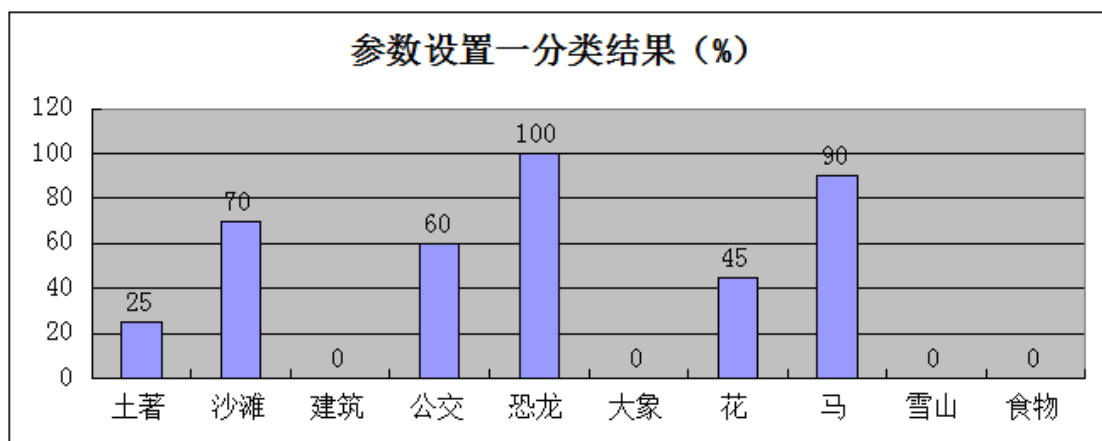


图4-2各个类别的分类结果 (%)

从图 4-2 中可以看出，在常规的最速下降法的训练函数下，分类效果很差。

(2) 均方误差性能：

均方误差性能函数 (MSE) 是衡量神经网络训练效果的一个参数，是误差的平方的算数平均，MSE 的越小越好，一般在神经网络训练时会设置一个初始的目标值，达到此初始的目标值则表示神经网络训练已经达到预期，则不再训练。均方误差的定义如下：

$$MSE = \frac{\sum_{i=1}^N e_i^2}{N} \quad (4.3)$$

其中 i 是样本编号， N 是样本总量， e_i 是第 i 个样本的误差。

神经网络在训练时，数据被自动分成 training set、validation set 及 test set 三部分，training set 是训练样本数据，validation set 是验证样本数据，test set 是测试样本数据，这样这三个数据集是没有重叠的。在训练时，用 training 训练，每训练一次，系统自动会将 validation set 中的样本数据输入神经网络进行验证，在 validation set 输入后会得出一个误差（不是网络的训练误差，而是验证样本数据输入后得到的输出误差，本文所用为均方误差），而此前对 validation set 会设置一个步数，本文设置的为 10 步，则系统判断这个误差是否在连续 10 次检验后不下降，如果不下降或者甚至上升，说明 training set 训练的误差已经不再减小，没有更好的效果了，这时再训练就没必要了，就停止训练，不然可能陷入过学习。validation set 的设置步数，作用就在这里。

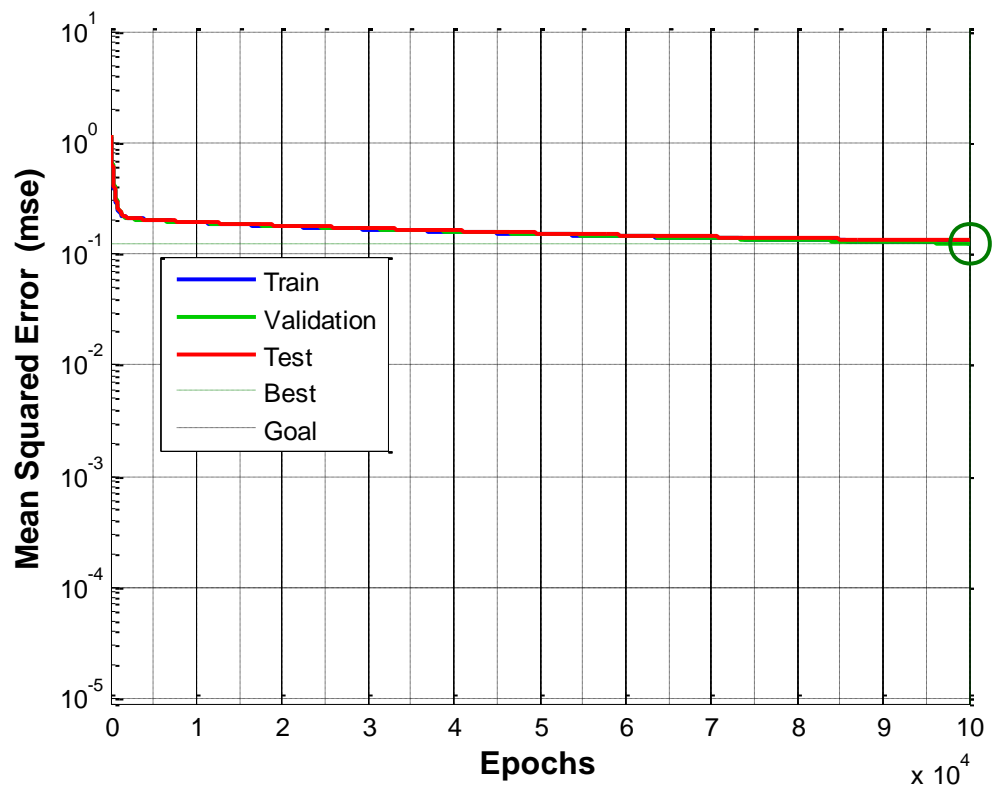


图 4-3 参数设置（一）条件下均方误差

从图 4-3 中可以看出，在常规的最速下降法的训练函数下，迭代 100000 次的情况下，均方误差性能只达到了 0.1 左右，远远大于设定的最大允许误差 0.0001。

(3) 回归性能 (Regression) 分析

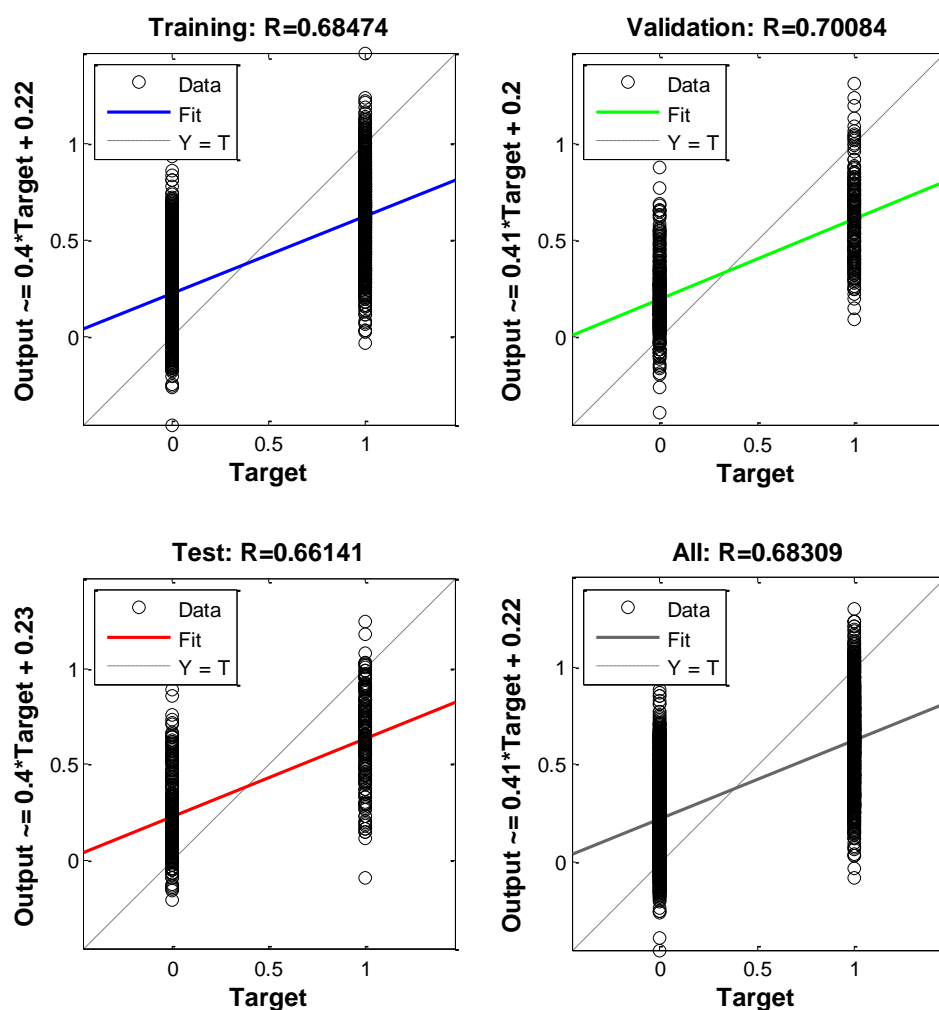


图 4-4 参数设置（一）条件下回归性能分析

从图4-4中可以看出training set、validation set 及test set中图像和对应标签的相关性只有68%左右，所以也能看出分类效果很差。

（二）参数设置二：

BP神经网络参数设置最大迭代次数：100000，目标性能(均方误差性能)：0.0001，学习率：0.1，验证数据的最大允许错误样本数：10。训练函数为采用改进的最速下降算法LM算法。

（1）分类效果：

在以上参数设置下训练用时45秒，测试集共200张图像，每个类别20张图像，总的分类正确率为72%，各个类别的分类结果如图4.5，分类结果的混淆矩阵如表4-1

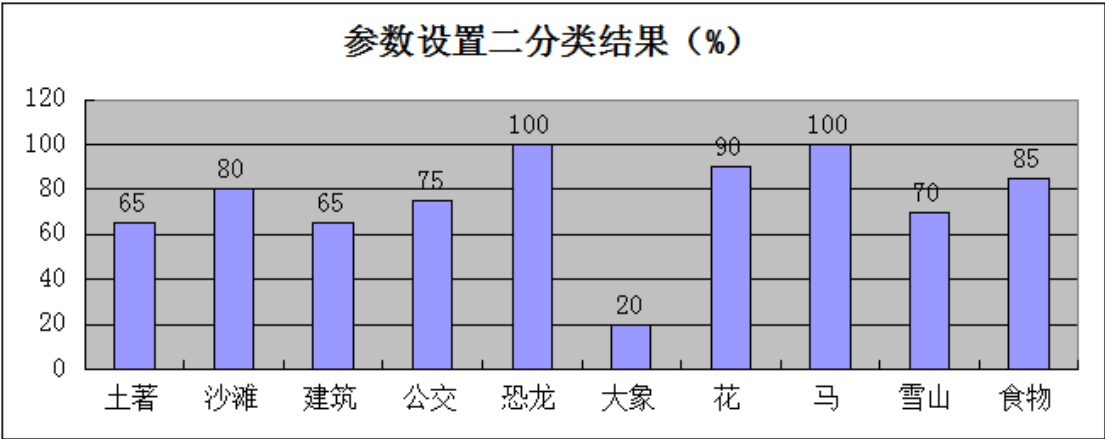


图4-5 各个类别的分类结果 (%)

从图4-5可以看出，分类正确率最高的是恐龙和马，分类正确率都为100%，分类最差的是大象，分类结果只有20%。

表4-1分类结果的混淆矩阵（分类正确数/百分比）

	土著	沙滩	建筑	公交	恐龙	大象	花	马	雪山	食物
土著	13/65	2/10	2/10		1/5				2/10	
沙滩		16/80		2/10		1/5			1/5	
建筑	3/15	3/15	13/65							1/5
公交			2/10	15/75	1/5					2/10
恐龙					20/100					
大象	2/10	1/5				4/20		12/60		1/5
花	1/5				1/5		18/90			
马								20/100		
雪山	3/15	3/15							14/70	
食物	1/5						2/10			17/85

表4-1是分类结果的混淆矩阵，混淆矩阵的每一行代表了测试数据的分类信息，每一行中的数值等于测试图像在相应类别中的数量。从中可以看出各个类别的分类正确数量和分类正确率，也可以看出各个类别的之间的影响。如表4-1中大象的分类正确率只有20%，而进步一观察发现，把“大象”分类成“马”的概率为60%，可见“马”这个类别对大象的影响很大，导致了大象分类的失败。分析其中的原因可能是“大象”类别和“马”类别的背景和前景都很相像（背景都是绿色为主，前景都是棕色为主，形状信息也差别不大）。针对数据集对分类效果影响的分析见（三）修改数据集后的分析。

(2) 均方误差性能

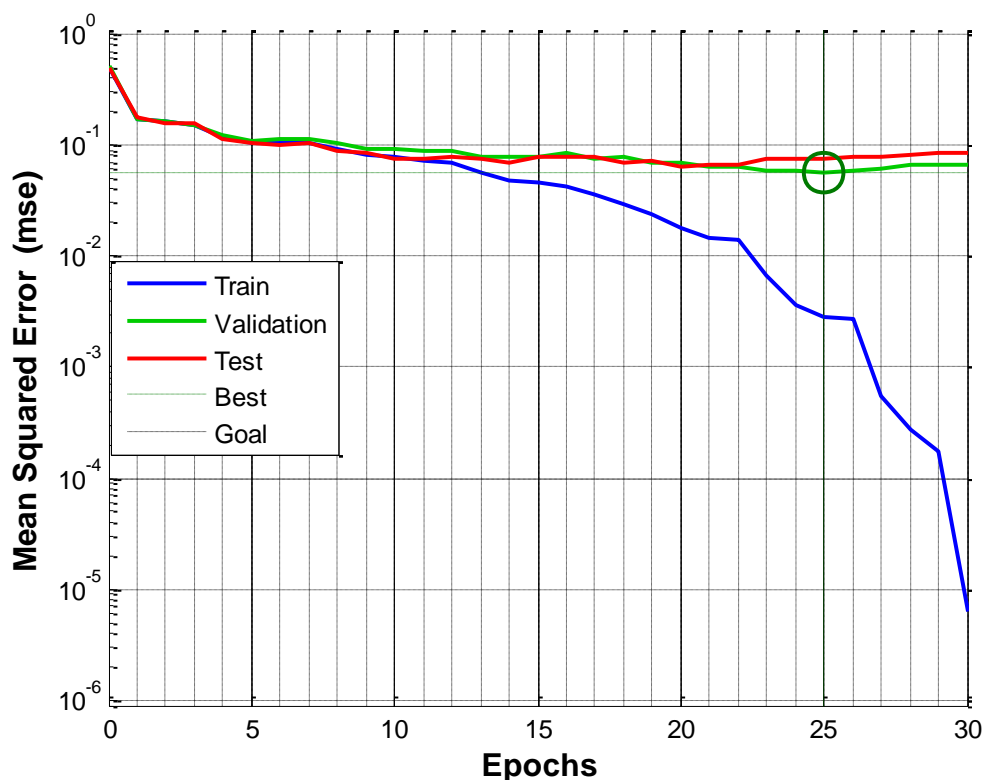


图 4-6 均方误差性能

如图 4-6 训练数据集在 30 步迭代达到标均方误差: $6.42e-5 < 0.0001$ (目标值), 在 30 步迭代后达到了预设的训练效果。

(3) 回归性能 (Regression) 分析

与均方误差性能一样, 神经网络在训练时把训练集分成 training set、validation set 及 test set 三部分, 图 4.7 分别画出了 training set、validation set、test set 和全部训练样本的回归性能。

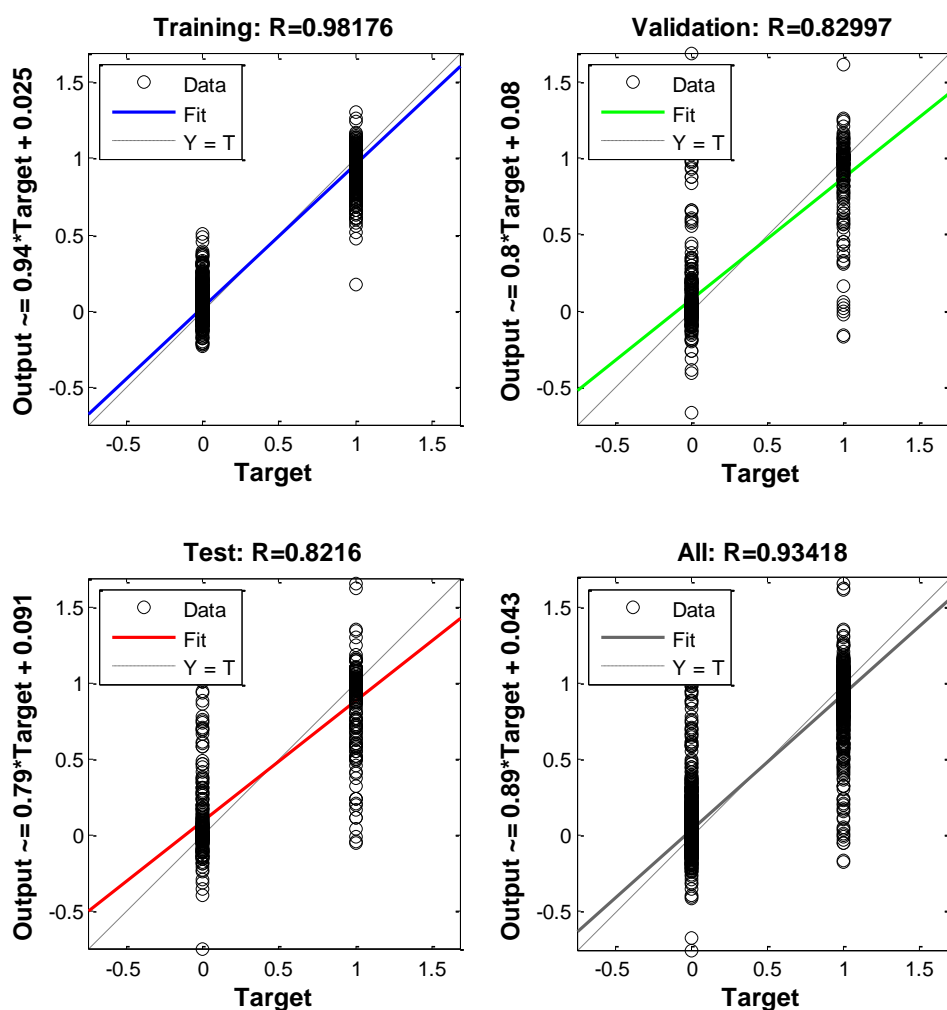


图 4-7 回归性能

从图 4-7 中可以看出，training set、validation set、test set 和全部训练样本中训练数据与各个类别的所对应的标签是正相关的，且其相关性都很高（越接近于 1 代表相关性越大），分别为：98.17%，82.99%，82.16%，93.41%。

（三）修改后的数据集测试分析

参数设置（二）中出现“大象”类别受“马”的影响较大，所以不能完全体现算法的性能，因此本文对 Corel 数据集做了修改，将“大象”类别改成“鸟”类别（图像来自百度）。

参数设置与参数设置（二）相同，即：BP神经网络参数设置最大迭代次数：100000，目标性能(均方误差性能)：0.0001，学习率：0.1，验证数据的最大允许错误样本数：10。训练函数采用改进的最速下降算法LM算法。

（1）分类效果：

在以上参数设置下训练用时1分01秒，测试集共200张图像，每个类别20张图像（“大象”换成“鸟”），总的分类正确率为83%，各个类别的分类结果如图4.8，分类结果的混淆矩阵如表4-2

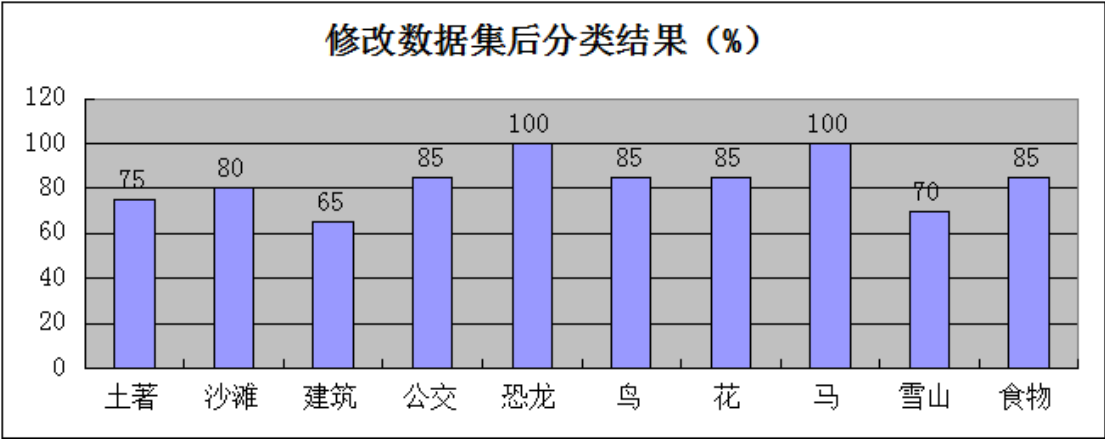


图4-8 各个类别的分类结果 (%)

从图 4-8 可以看出：在把“大象”类别改为“鸟”后算法性能有了很大的提高，从之前总体 72% 的正确率，上升至现在 83% 的总体正确率。

表4-2分类结果的混淆矩阵（分类正确数/百分比）

	土著	沙滩	建筑	公交	恐龙	鸟	花	马	雪山	食物
土著	15/75	1/5	3/15						1/5	
沙滩	1/5	16/80	1/5	1/5		1/5				1/5
建筑	2/10	2/10	13/65						3/15	
公交	1/5	1/5	1/5	17/85						
恐龙					20/100					
鸟	1/5				1/5	17/85	1/5			
花					1/5	1/5	17/85	1/5		
马								20/100		
雪山	3/15	1/5	1/5						14/70	1/5
食物	2/10								1/5	17/85

对比参数设置二中的分类效果的混淆矩阵表 4-1，可以看出，在改变数据集的“大象”类别后，其他类别的分类效果变化不大，而“大象”换成“鸟”后，“鸟”类别受其他类别的影响很小，分类正确率有很大的提高，总的分类正确率也有很大的提高。从中可以得出结论，此分类算法适合于色彩，对象不是很相近的图像，对北京色和前景中对象都很相像的图像分类效果略差。

(2) 均方误差性能

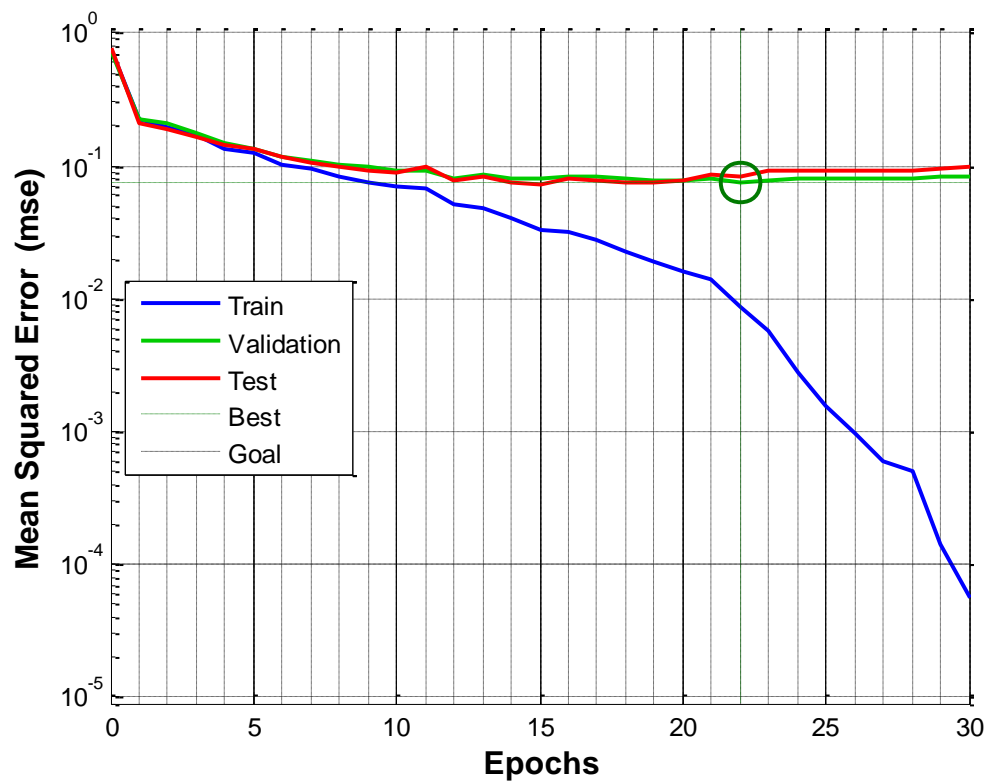


图 4-9 均方误差性能

如图 4-9 训练数据集在 30 步迭代达到均方误差： $5.69e-5 < 0.001$ (目标值)。在 30 步迭代后达到了预设的训练效果。

(3) 回归性能 (Regression) 分析

图 4-10 分别画出了 training set、validation set 、test set 和全部训练样本在改变训练数据集情况下的回归性能。

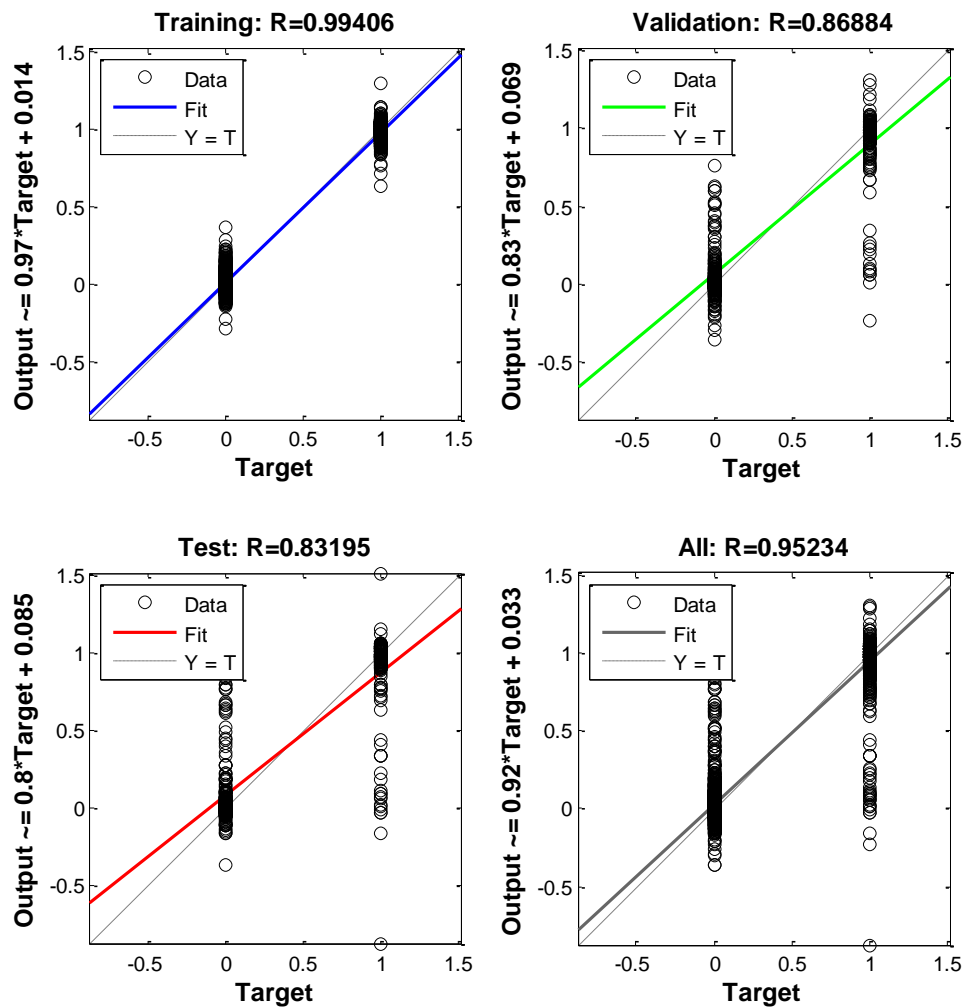


图 4-10 回归性能

从图 4-10 中可以看出, training set、validation set、test set 和全部训练样本中训练数据与各个类别所对应的标签是正相关的, 且其相关性都很高 (越接近于 1 代表相关性越大), 分别为: 99.41%, 86.88%, 83.19%, 95.23% 高于原始的 Corel 数据集情况下的相关性: 98.17%, 82.99%, 82.16%, 93.41%。。这也说明了, 训练数据与各个类别所对应的标签的相关性越高, 分类效果越好。

4.2 算法鲁棒性测试

4.2.1 测试方法

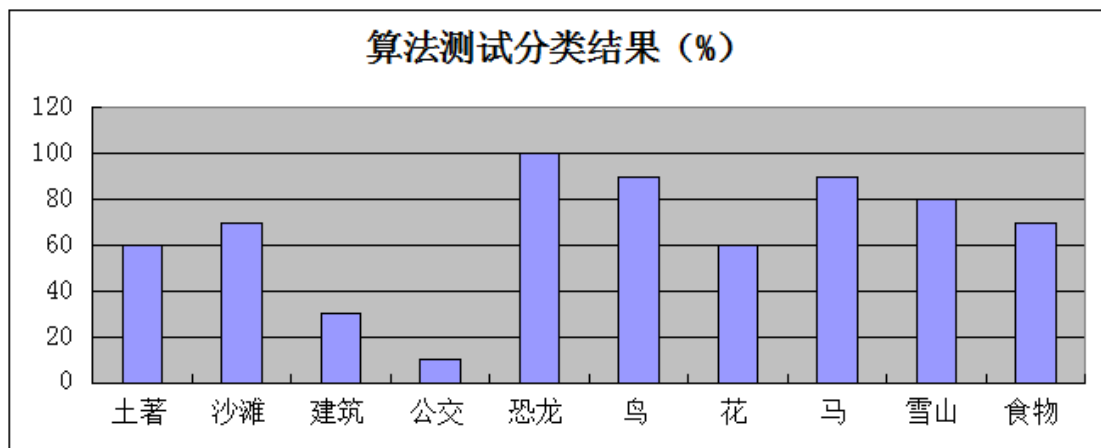
为了测试算法的性能，本文采用的方法是，从网上随机下载“土著、沙滩、建筑、公交、恐龙、鸟、花、马、雪山、食物”十类图像各十张，不限定图像的大小，用第三章 3.3 节的方法提取图像的特征，将提取特征后的数据集输入训练好的网络，得到测试数据集的标签，并统计分类正确率。

此测试数据集与原数据集相比，测试数据的选取更加随机，没有限定图像的大小，而原训练数据集和测试数据集的都是 384×256 像素的。用此测试数据集对分类算法进行测试，更能体现分类算法的鲁棒性。

4.2.2 算法鲁棒性测试结果

将提取的图像特征数据输入训练好的网络，仿真并统计分类结果。得到分类效果如表 4-6 所示。

表 4-6 算法测试的分类效果 (%)



从表 4-6 可以看出，“恐龙”的分类正确率是 100%，“鸟”和“马”的分类正确率是为 90%，“建筑”和“公交”的分类正确率较低，分别为 30% 和 10%。此结果如原测试集的分类结果比较可以看出，分类算法的鲁棒性较好，对多数类别的分类效果比较稳定。

4.3 本章小结

本章主要内容共三个小结，4.1 节是算法的分析，分别分析了在常规最速下降算法下的神经网络性能和 LM 算法下的神经网络的性能，常规的训练算法在进行 16 分钟运算迭代 100000 次的情况下并未完成训练，而 LM 算法仅用 45 秒就完成了训练，达到了预设的均方误差目标，说明 LM 算法比常规的最速下降算法更加的有效，分类效果更好，但有一个缺点：内存消耗大；之后又分析了在把大象类别用鸟来代替时，分类效果有很大的提高，从混淆矩阵中可以看出，各类别之间有一定的影响。

4.2 节对算法的鲁棒性进行了测试，用从网络随机下载的 100 张图像（每个类别 10 张）对分类算法进行了测试，结果显示此分类算法有不错的鲁棒性。

第五章 总结与展望

图像中包含着丰富多彩的信息，并不能简单地由底层的视觉特征来全面描述它。如何对图像进行高效准确的检索与分类仍然是一个需要处理的问题。本章主要对全文做一个总结性陈述，并对未来工作进行展望。

5.1 总结

随着图像数据的急剧增加，人工图像分类已经越来越费时费力，有效的图像分类算法是近些年来研究的热点，本文的研究也是基于这个大背景之下。

本文的主要研究内容如下：

本文研究了基于图像底层信息的分类方法，利用图像的颜色，纹理，不变性特征，利用 BP 神经网络做为分类器。设计了有效的分类算法，并在 corel 数据集以及改进的 corel 数据集上进行了测试，实验表明，该方法取得了较好的效果。

参考文献

- [1] 杨源杰. 黄道. 人工神经网络算法研究及应用. 华东理工大学学报(自然科学版). 2002 年 5 期
- [2] 魏海坤. 神经网络结构设计的理论与方法【M】. 北京: 国防工业出版社, 2005. 2.
- [3] 吕琼帅. BP 神经网络的优化与研究. 郑州大学. 2011. 5. 第 7 页
- [4] 罗易. 浅谈互联网对高校大学生的双重影响与对策分析[J]. 南方论刊, 2009, 06:104-105.
- [5] 温超, 耿国华. 基于内容图像检索中的“语义鸿沟”问题[J]. 西北大学学报(自然科学版), 2005, 05:46-50.
- [6] 钟洪, 夏利民. 基于本体的图像检索[J]. 计算机工程与应用, 2007, 17:37-40.
- [7] 曾奇森. 基于内容的图像检索相关技术研究. 南京理工大学, 2007
- [8] 韩力群. 人工神经网络理论、设计及应用. 北京: 化工工业出版社, 2002.
- [9] 张德丰. MATLAB 神经网络应用设计. 北京: 机械工程出版社, 2009. 1
- [10] 刘曙光. 前馈神经网络中的反向传播算法及其改进. 进展与展望, 计算机科学, 1996 年第 23 卷第 1 期
- [11] 陈明. MATLAB 神经网络原理与实例精解. 北京: 清华大学出版社. 2013. 3 163-165.
- [12] 席晓聪. 图像分类方法研究. 山东大学. 2013. 5
- [13] 赵小川. 何灏. 缪远诚. MATLAB 数字图像处理实战. 北京: 机械工程出版社. 2013. 6. 159~160.
- [14] 覃登攀. 基于遗传算法和人工神经网络相结合的南宁市空气质量预报研究. 广西: 广西大学, 2008.
- [15] 张立明. 人工神经网络的模型及其应用. 上海: 复旦大学出版社, 1993.
- [16] 魏海坤. 神经网络结构设计的理论与方法. 北京: 国防工业出版社【M】, 2005.
- [17] 吴简彤, 王建华. 神经网络技术及其应用. 哈尔滨: 哈尔滨工程大学出版社, 1997.
- [18] 孙兴华. 基于内容的图像检索研究. 南京理工大学. 2001. 10
- [19] Zhou ZH, Zhang ML. Neural networks for multi-instance learning. In: Shi ZZ, He Q, eds. Proceedings of the International Conference on Intelligent Information Technology. Beijing: People's Post and Telecommunications Publishing House, 2002, 455-459.
- [20] M. Minsky, S. Papert. Perceptrons [M]. In: Boston: MIT Press 1969.
- [21] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities, "in Proc. Nat. Academy Sci., vol. 79, 1982:2554-2558.
- [22] David M. Mark & Max J. Egeuhofer, Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subject Testing, Cartography and Geographic Information Systems, 1995, 21(4).
- [23] Ampazis, N.; Perantonis, S. J.; Leveuberg-Marquardt algorithm with adaptive momentum for the efficient training of feed forward networks; Neural Networks, 2000. UCNN2000,

Proceedings of the IEEE—D烈 S—ENNS International Joint Conferenceon, Volume: 1, 2000'1.

- [24]Mafia B. Diaz Aguado and C. Gonzblez Nieleza, Control and prevention of gas outbursts in coal mines, International Journal of Coal Geology, Volume 69, Issue 4, 1 March 2007, 253. 266.
- [25] Andrej Karpathy. Li Fei-Fei. Deep Visual-Semantic Alignments for Generating Image Descriptions.
- [26] Ramon J, Raedt LD. Multi-Instance neural networks[A]. Proceedings of ICML-2000 workshop on Attribute-Value and Relational Learning[C], 2000.
- [27] stricker M , Orengo M. similarity of col0r images[J]. Proc. SPIE St0rage and Retrieval f0r Image and Video D tabases, 1995, 2420: 381—392