

谨以此论文献给我的导师和亲人！

————— 朱亚菲

# 基于多特征多分类器组合的海洋浮游动物图像 分类研究

学位论文答辩日期: \_\_\_\_\_

指导教师签字: \_\_\_\_\_

答辩委员会成员签字: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## 独 创 声 明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的科研成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含未获得 \_\_\_\_\_（注：如没有其他需要特别声明的，本栏可空）或其他教育机构的学位或证书使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：                      签字日期：        年    月    日

---

## 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，并同意以下事项：

1. 学校有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。
2. 学校可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同时授权清华大学“中国学术期刊(光盘版)电子杂志社”用于出版和编入 CNKI《中国知识资源总库》，授权中国科学技术信息研究所将本学位论文收录到《中国学位论文全文数据库》。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：

签字日期：        年    月    日

导师签字：

签字日期：        年    月    日

# 基于多特征多分类器组合的海洋浮游动物图像 分类研究

## 摘 要

海洋覆盖了地球上约 70% 左右的面积，对人类有着重要而深远的影响，能够为人类创造巨大的经济、社会、环境效益。海洋生态系统是一个结构复杂的大系统，浮游动物是其中的重要一员，其生物量、种群结构以及群落多样性对海洋生态系统、海洋生物地球化学循环、海洋环境以及全球气候变化研究都起着重要的作用。

随着浮游生物光学成像系统的不断发展，海量的浮游动物图像资料涌现出来，为浮游动物的监测和分类带来了挑战。本论文主要针对浮游动物图像分类技术中的特征提取和分类器设计两个关键步骤进行了分析和研究，提出了一种基于多特征多分类器组合的浮游动物图像分类方法，主要研究工作如下：

1. 以混淆矩阵中的分类准确率指标作为评价函数，对由图像处理领域常用特征、各种分类竞赛中采用的经典特征、浮游动物形态特征等组成的特征集进行特征选择，最终挑选出适用于浮游动物图像三类特征，分别为 PkID 软件中的 22 个特征、局部二值模式特征、内距离形状上下文特征。
2. 对于挑选出的三类特征，分别对其采用不同的分类算法进行分类，将分类结果用交叉验证方法进行评价，根据评价结果挑选出适用于每类特征的分类算法，得到三个分类器。
3. 单个分类器对浮游动物图像分类可能存在一定的片面性，分类性能不能达到最佳，对此，我们提出了一种基于多特征多分类器组合的浮游动物图像分类方法，从而能够综合不同分类器的信息，提高分类准确率。

我们在包含 13 类浮游动物的数据集上进行了对比实验，证明了本文的算法与其它浮游动物分类方法相比在分类准确率上有明显的提高，对于浮游动物分类是有效和可行的。

**关键词：** 海洋浮游动物；图像分类；机器学习；多特征多分类器组合

# Marine zooplankton image classification based on combination of multiple features and multiple classifiers

## Abstract

The ocean occupies about 70% of the earth surface, producing significant economic, social and environmental benefits for human beings. Marine ecosystem is a complicated system, in which zooplankton is an important member. Marine zooplankton biomass, population structure and community diversity play an important role in marine ecosystems, marine biogeochemical cycles, marine environment and global climate change research.

With the development of marine plankton optical imaging system, a tremendous amount of zooplankton images sprang up, which challenges the monitoring and classification of zooplankton. The thesis mainly focuses on the feature extraction and classifier design steps in zooplankton image classification, presenting a new zooplankton image classification method based on combination of multiple features and multiple classifiers. Main tasks are as follows:

1. The classification accuracy indicator of confusion matrix is taken as the evaluation function for feature selection from the feature sets, which is composed of the most frequently used features in image processing fields, classic features in various classification competitions and zooplankton configuration features. We finally select three suitable types of features for zooplankton images: 22 features in PkID, local binary pattern and inner-distance shape context.
2. For each of the three types of features we choose, we apply different classification algorithm for it, and then use cross validation for evaluation. According to the evaluation results, we can select the most proper classification algorithm for each feature type and achieve three classifiers.
3. Single classifier for zooplankton classification has some certain one-sidedness, and the classification performance may not get the best. We propose a method based on combination of multiple features and multiple classifiers, thus can combine the information of different classifiers and improve the performance.

The performance comparison on zooplankton dataset with 13 categories validates that the proposed method has an obvious increasement on classification accuracy compared with other methods in zooplankton image identification. It is feasible and effective in zooplankton classification.

**Keywords:** marine zooplankton, image classification, machine learning, combination of multiple features and multiple classifiers

## 目 录

1 绪论 .....	1
1.1 课题的研究背景及意义 .....	1
1.2 浮游生物图像自动识别方法的国内外研究现状 .....	2
1.2.1 浮游生物成像系统的发展 .....	2
1.2.2 浮游生物图像自动识别技术的发展 .....	3
1.3 课题来源 .....	4
1.4 论文组织结构安排 .....	5
2 浮游动物图像分类基础 .....	6
2.1 机器学习分类技术 .....	6
2.2 数据集 .....	6
2.3 评价方法 .....	9
2.3.1 混淆矩阵简介 .....	9
2.3.2 混淆矩阵分类 .....	11
2.3.3 交叉验证 .....	11
2.4 ZooScan 系统介绍 .....	12
2.4.1 简介 .....	12
2.4.2 硬件部分 .....	12
2.4.3 软件部分 .....	13
3 浮游动物图像特征提取 .....	20
3.1 特征提取方法 .....	20
3.1.1 图像处理领域常用特征 .....	20
3.1.2 经典特征 .....	21
3.1.3 其它特征 .....	28
3.2 特征选择 .....	30
3.2.1 PkID 中的 22 个特征 .....	30
3.2.2 局部二值模式 .....	32
3.2.3 内距离形状上下文 .....	32

4 分类器设计 .....	33
4.1 常用的分类技术 .....	33
4.1.1 支持向量机方法 .....	33
4.1.2 随机森林方法 .....	34
4.1.3 极限学习机方法 .....	36
4.2 分类算法选择 .....	38
5 多特征多分类器组合 .....	44
5.1 基本原理 .....	44
5.2 图像预处理 .....	45
5.3 实验对比与分析 .....	45
6 总结与展望 .....	50
6.1 总结 .....	50
6.2 展望 .....	50
参考文献 .....	52
致 谢 .....	55
个人简历、在学期间发表的学术论文与研究成果 .....	56



# 1 绪论

## 1.1 课题的研究背景及意义

地球上海洋的面积为 3.6 亿平方千米, 约占地球表面总面积的 71%, 海洋是地球上综合生产力最大的生态系统, 对于经济的快速发展起到至关重要的作用。而在海洋当中, 海洋浮游生物是数量最多、规模最庞大的生物种类, 其主要包括浮游植物和浮游动物两大类。这些浮游生物, 绝大多数个体较小(一般从几微米到几毫米), 需要借助显微镜才能看清楚它们的身体构造, 并且它们缺乏发达的运动器官, 游动能力很弱。其中浮游植物属于初级生产者, 是食物链的第一个环节, 其数量变动对浮游植物群落结构和海洋生态环境的维持起着十分重要的作用。而浮游动物属于次级生产者, 它的作用主要表现为: 第一, 它相当于在浮游植物和其它捕食者之间搭起一座桥梁, 使能量可以顺利地由低向高传输, 所以其在维持整个海洋中生物(包括浮游植物和鱼类等)数量方面起着举足轻重的作用。第二, 浮游动物自身的尸体和摄食产生的粪便中含有有机碳元素, 并且以这种形式将 N、C、P 等元素传输到深海, 以丰富海洋的有机物元素。第三, 栖息环境的变换会很大程度上地影响浮游动物, 表现为群落的结构和功能参数上的差异, 而这些差异可以反应气候变化及海洋环境变化。第四, 某些种类的浮游动物在较深水层大量密集, 会形成深海散射层, 阻碍或干扰水下声纳的信号传播, 使声纳失效, 因而, 它们在国防军事中的作用也越来越受到人们的重视。

浮游生物监测的重点是调查其种类组成和数量分布, 其中最重要的监测指标是“优势种类判定”及“密度计算”。传统的浮游生物监测主要是由工作人员在显微镜下观测样本, 并对样本进行分类和计数。但是这种监测手段对浮游生物工作者的专业性要求较高, 且需要大量人力、财力的支撑, 识别速度较慢, 无法达到实时分析的需求。在过去的几十年间, 浮游生物成像系统得到了快速发展, 从而可以在更精细的时间、空间尺度上对浮游生物进行研究。与此同时, 为了有效解决由于浮游生物分类学人才稀缺而造成的种类鉴别滞后现象, 越来越多的科研工作者投入到利用图像处理和模式识别技术进行浮游动物的自动分类与计数这一研究中, 并且在目前阶段已经取得显著进展。

## 1.2 浮游生物图像自动识别方法的国内外研究现状

浮游生物图像自动识别技术的发展与成像系统的研究和改进密不可分，以下具体介绍这两类技术近年来在国内外的研究现状。

### 1.2.1 浮游生物成像系统的发展

传统的浮游生物图像通常是经过海洋调查采水取样，然后用显微镜对样本进行拍照而得。由于近几年摄像技术的不断提高和日益增长的海洋浮游生物实时监测的需求，使得新的浮游生物现场及光学成像系统不断涌现。

美国伍兹霍尔海洋研究所（WHOI）的科研团队研制的视频浮游生物录像机（Video Plankton Recorder, VPR）<sup>[1]</sup>可以近距离拍摄到浮游生物的影像，为研究浮游动物提供充足的样本；美国南佛罗里达大学研发的灰度图像颗粒探测系统（Shadowed Image Particle Platform and Evaluation Recorder, SIPPER）<sup>[2]</sup>采用高速线扫描相机可对水下浮游生物连续成像；法国研制的水下颗粒物和浮游动物图像原位采集系统（Underwater Video Profiler, UVP）<sup>[3]</sup>采用传统的照明设备和经电脑处理的光学技术，可实现浮游动物和颗粒物的剖面观测以及浮游动物图像和颗粒物图像的原位采集；由美国流体成像技术公司研发生产的流式细胞摄像系统（FlowCam）<sup>[4]</sup>能够从流体中快速检测出有机和无机悬浮体，具有连续成像拍照和流式计数功能。近年来，全息成像技术也逐步成为获取浮游生物图像的一种重要手段。英国数家科研机构联合开发一种新的全息成像样机（HOLOMAR）<sup>[5]</sup>，这款试验机是基于离轴技术，可以获得水下浮游生物图像，在水下  $10^5\text{cm}^3$  的容量内都能精确记录，并且最小分辨率能达到  $10\mu\text{m}$ 。美国约翰霍普金斯大学于 1999 年研发了一款基于水下同轴激光的全息成像系统，在水下  $732 - 1964\text{cm}^3$  的体积内都能进行勘察，用电荷耦合器件替代了在暗室中进行全息干板记录<sup>[6][7]</sup>。

国内，厦门大学的戴君伟等人<sup>[8]</sup>为采集海洋藻类生物图像研制了一种新的系统，该系统融合了流式细胞技术和常用的成像采集技术，能够快速获得图像数据。国家海洋技术中心于连生等人<sup>[9]</sup>于 2009 年研发的“水下全自动显微成像仪”，该仪器可以很好地对浮游生物进行拍照；浙江大学陈耀武、洪炎峰等<sup>[10]</sup>开发的“实时浮游生物图像目标智能识别系统”能够对海洋水体中的目标生物进行实时连续走航大面积探测和图像采集；中国海洋大学于新生、周章国等<sup>[11]</sup>开展了基于光学成像的浮游生物实时在线监测系统研究。综上，在对海洋浮游生物成像的研究中，不断涌现的采集系统保证了充足的研究资源，使基于图像处理的浮游生物研究更具有发展前景。

### 1.2.2 浮游生物图像自动识别技术的发展

海洋浮游生物的种类多种多样,对这么多的物种进行分类是一项非常繁重的任务。传统的浮游生物识别方法一般是人工地依据形态、尺寸、纹理、颜色等特征进行分类,再根据医学和化学上的一些知识进行进一步判定。但是这种人工识别方法存在一些明显的缺点:首先是对专业性要求较高,工作量大,效率低,浮游生物工作者在识别过程中易疲劳,从而难以保证海洋浮游生物监测的实时性和准确性。近年来发展了很多用来鉴定浮游生物的新技术,例如光谱法、流式细胞术等,但这些方法大多过程繁琐,严重依赖于浮游生物的生理状态,同时仪器昂贵,费用消耗较多。而目前利用数字图像处理技术进行浮游生物自动分类的方法,与其他方法相比,研究成本较低,可以实现精确的种属识别,因而越来越受到重视。

在对海洋浮游生物进行监测的技术中,数字图像处理这一方法的使用可以追溯到1970年,当时计算机处理能力有限,加上光学成像传感器的灵敏度和分辨率较低,对图像只能进行一些基本的信息获取。例如,利用模式识别方法对显微镜下的浮游动物进行简单的计数、尺寸度量和分类<sup>[12]</sup>。近几年,图像处理技术快速发展,计算机性能不断提升,利用图像处理技术进行浮游生物的分类识别也取得了长足的进展。

二十世纪末期,欧洲自动硅藻识别和分类项目建立了自己的硅藻图像数据库并实现了硅藻的自动判识,其使用的方法包括数字图像处理和模式识别。基于ADIAC项目,硅藻显微图像分析取得一系列进展<sup>[13][14][15]</sup>。此外,Embleton等<sup>[16]</sup>结合计算机图像分析利用多层感知器(MLP)神经网络方法实现4种浮游植物自动计数。Culverhouse等<sup>[17]</sup>基于人工神经网络进行6种甲藻分类。Blaschko等<sup>[18]</sup>采用图像分割、特征提取和模式识别方法实现12类浮游生物自动识别。Davis和Hu<sup>[19]</sup>利用形状和纹理特征采用神经网络和支撑向量机进行6类浮游生物自动识别及丰度估计。Sosik和Olson<sup>[20]</sup>通过边缘检测等图像处理后提取尺寸、形状、纹理等特征再采用支撑向量机完成22类浮游植物自动分类和丰度估计。Hense等<sup>[21]</sup>结合荧光图像进行浮游生物结构分析(PLASA)。Verikas等<sup>[22]</sup>结合相位一致性圆形目标检测、随机优化目标轮廓确定以及支撑向量机和随机森林分类方法实现微小原甲藻细胞目标的检测和识别。腰鞭毛虫分类项目中,研究者们利用形状和纹理特征,采用人工神经网络分类算法对不同种类的腰鞭毛虫进行了分类<sup>[23]</sup>。法国国家科学院的研究者们对浮游动物进行了位置特征、尺寸特征、灰度特征、形状特征、生物统计特征以及其它自定义特征的提取,并提供了7种常见的分类算法供选用,实现了对浮游动物的分类识别,识别准确率达到80%左右。Ellen等

人<sup>[24]</sup>对已有的机器学习方法进行了研究,能够对浮游动物分类问题的算法选择、性能调整等提供指导,从而最大限度地提高分类准确率。

与国外相比,国内在利用数字图像处理技术来实现浮游生物的分类识别上起步较晚,因而不管在在研究时间还是研究深度上都是有很大的距离的,并且大多数项目的研究对象都是针对浮游植物(如藻类)的。王明时等人<sup>[25]</sup>首先使用阈值法和形态学法对已有的赤潮生物显微图像进行处理,在图像中得到单个的浮游生物个体,再对这些个体的灰度和轮廓特征进行提取,最终完成分类识别。汪振兴等人<sup>[26]</sup>在对藻类图像的分类识别当中,提取了纹理和形状两种类型的特征,并采用了人工神经网络分类算法训练出分类器,实现了对赤潮藻的自动识别。骆巧琦等人首先对硅藻样本进行了采集,然后通过双轮廓叠加法对采集到的图像进行了分割,再提取了几何描述特征和形状描述特征,经过人工神经网络算法进行训练,最终得到精确的分类器<sup>[27]</sup>。魏雅娟等提取了基本的形态特征和纹理特征,采用支持向量机作为分类算法,实现了浮游动物图像的自动识别<sup>[28]</sup>。

综上所述,目前国内外基于图像技术的浮游生物分析主要在于浮游生物的目标识别和计数,且主要呈现出以下特点:

1. 就研究对象而言,国内外关于浮游动物分类识别的研究明显少于对浮游植物(如藻类)的识别研究。
2. 就研究范围而言,大多集中在某一个种类或者少量几个类别。
3. 就目标表示的特征提取而言,大多数研究方法选择外部形状特征和纹理特征作为浮游生物特征提取和描述,而多数浮游生物纹理特征并不明显,难以作为分类特征,且对特征没有进行有效研究与分析。
4. 就分类技术而言,逐渐趋于使用机器学习算法作为分类器,但一般使用单一的分类器得到分类结果,而没有考虑不同分类器分类结果的融合。

### 1.3 课题来源

课题来源:国家自然科学基金青年科学基金项目“基于视觉注意结合生物形态特征的海洋浮游植物显微图像分析”(批准号:61301240)、国家自然科学基金项目“基于生物形态特征的中国海常见有害赤潮藻显微图像识别”(批准号:61271406)和中央高校基本科研业务费“海洋浮游动物原位探测与分析系统”(批准号:201562023)。

## 1.4 论文组织结构安排

本文总结了国际上现有的浮游动物图像分类方法，对其中基于机器学习的浮游动物图像分类算法进行了深入研究，并提出了基于多特征多分类器组合的浮游动物分类算法。本文的主要安排如下：

第一章为绪论部分，主要对浮游动物图像分类研究的背景、意义及国内外研究现状进行介绍，并对全文的主要安排进行了说明。

第二章介绍浮游动物图像分类的一些基础知识。详细介绍了用来进行分类的数据集以及利用机器学习方法进行浮游动物图像分类的过程，对浮游动物图像分类准确率最高的 ZooScan 系统进行了大致的描述，最后，说明了如何评价一个分类器性能的好坏。

第三章将图像处理和机器视觉领域一些常用的特征进行整理和总结，通过一定的评判标准进行特征选择，最终选取三类特征作为浮游动物图像的特征提取。

第四章简述了机器学习领域常用的分类技术，以分类准确率作为评判，设计适合各类特征的最优分类器。

第五章针对单一分类器存在分类片面性，提出了多特征多分类器组合的方法，详细介绍了算法的基本原理，并通过在包含 13 类浮游动物的数据集上进行实验证明了我们的方法的有效性。

第六章对全文的工作进行了总结，指出工作中存在的问题并对以后的工作提出了展望。

## 2 浮游动物图像分类基础

### 2.1 机器学习分类技术

利用机器学习方法对浮游动物图像进行分类的主要模块是图像预处理、特征提取、分类，主要步骤为：

(1) 将浮游动物图像样本分为训练样本和测试样本两类，并建立实验所需的训练样本和测试样本数据；(2) 分别对训练样本和测试样本进行预处理；(3) 对训练样本进行特征提取，从中选择合适的特征信息构成特征向量；(4) 设计分类算法，通过支持向量机、随机森林、神经网络等方法设计分类器；(5) 将训练样本图像的特征输入到分类器，训练出分类模型；(6) 提取测试样本特征，并输入到分类模型中，输出测试样本的分类结果。具体见图 2-1。

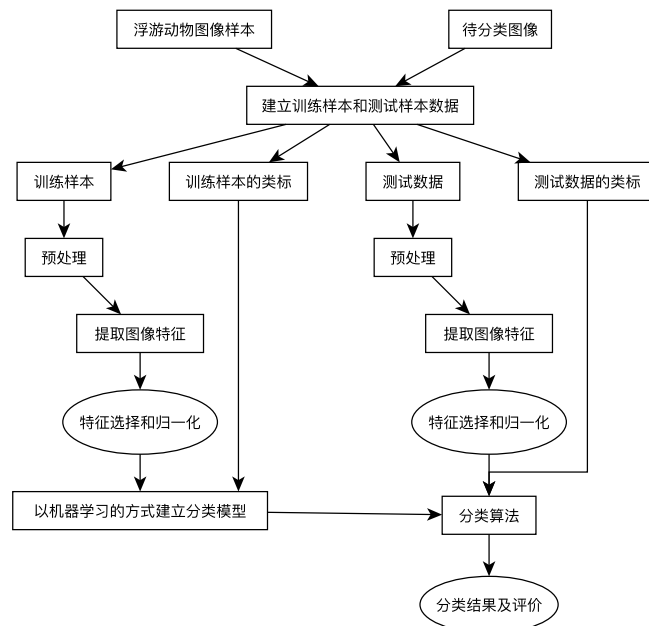


图 2-1 利用机器学习方法实现浮游动物图像分类的基本流程

### 2.2 数据集

自 1949 年起，加利福尼亚海洋渔业研究合作组织（California Cooperative Oceanic Fishery Investigation, CalCOFI）就开始对包括浮游生物在内的海洋生物进行采样<sup>[29]</sup>，根据标准邦戈网络协议采集海洋中的浮游生物样本，然后立即保存起来。部分被保存的 CalCOFI 样本被连接到加利福尼亚对当前生态系统进行长期生

态研究的网站上，并用 ZooScan 进行了扫描<sup>[30]</sup>。得到的灰度图像质量较好，在对比度、噪声等方面都得到了精确控制。整个数据集包含的种类如表 2-1 所示，其中每种类别图像所占比例见图 2-2，以下对该数据集中的每一类别分别进行了详细说明和描述。

表 2-1 浮游动物数据集包含种类

Appendicularia	尾海鞘纲
Bubble	气泡
Chaetognatha	毛颚动物门
Cladocera Penilia	尖头溞属
Copepoda	桡脚类
Decapoda	十足目
Doliolida	海樽目
Egg	卵
Fiber	纤维
Gelatinous	明胶
Multiple	多个生物
Pteropoda	翼足目
Nonbio	非生物

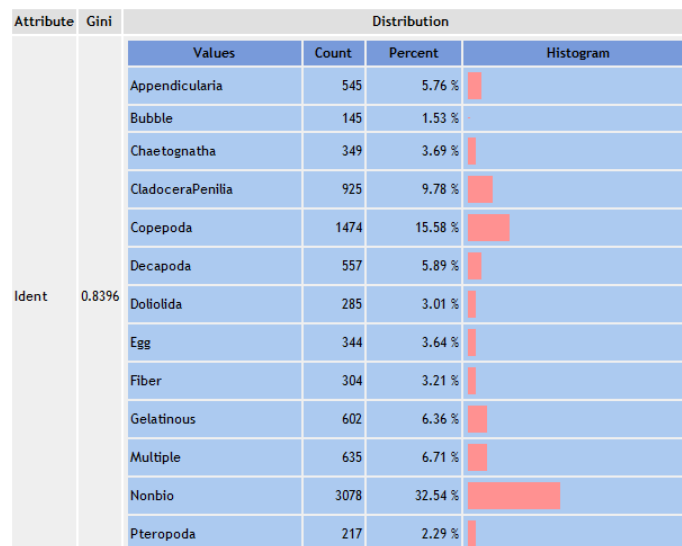


图 2-2 数据集中每种类别的数目及所占比例

**Appendicularia**（尾海鞘纲）属于脊索动物门，体型像蝌蚪，身体分为躯干和尾两部分。躯干较大且灰度较深，呈不太规则的椭圆形。尾部比躯干要长，长

度大多小于 5mm。尾部大致呈现两种形状：一种细长弯曲，另一种较粗（粗细甚至于头部差不多），呈现柳叶状。尾部的灰度相比于躯干较浅，轮廓不太清晰。

**Bubble（气泡）** 非生物，圆形。气泡四周灰度深，中间灰度很浅，呈亮白色。

**Chaetognatha（毛颚动物门）** 身体修长，可以明显看出身体分为头、躯干和尾 3 部分。头部略膨大，躯干与头连接处稍微缢缩为颈部，躯干较粗，轮廓清晰。尾部慢慢变窄，末端尖细。头部两侧各着生一系列 4 ~ 13 根镰刀状、几丁质的颚毛。大小：最大成年成年个体长 105mm，该类成年个体的长度一般都大于 5mm。

**Cladocera Penilia（Penilia，尖头蚤属）** 属于节肢动物门，鳃足纲，枝角目，俗称水跳蚤。身体短小，有两条长长的触角。该类浮游动物身体中轴线的地方灰度较深（感觉类似人体的脊柱），这个颜色较深的中轴线上还有一条条纹理线连向边缘（就像人体脊柱上连着的骨骼）。由于运动，扫描得到的图像中浮游动物的中轴线并不是都在其身体中间。大小约为 1mm 左右。

**Copepoda（桡脚类）** 属于节肢动物门，颚足纲，桡足类属于其下的一个亚纲。体形像泪珠，有大的触角。分为前体部和后体部，前体部较为宽大，后体部较为短小。前体部前体部由头和胸部组成，头部有两对触角，胸部有颚足、五对胸足。后体部无附肢，由 3—5 节组成。最末的腹节称尾节，末端具 1 对尾叉，尾叉的末端有 5 根不等长的刚毛，常呈羽状。

**Decapoda（十足目）** 属于节肢动物门，软甲纲。又分为两类：Lucifer hansenii 和 Crab larvae。体躯延长呈虾形（腹部发达）或缩短扁圆呈蟹形（腹部化）。

**Doliolida（海樽目）** 属于脊索动物门，樽海鞘纲。体型一般呈桶状，体壁最外是被囊层，其内层是外套膜。被囊层下有 8 ~ 9 条肌带环绕着体躯。由于该类浮游动物比较透明，因此在图像中灰度较浅，并且其桶状轮廓也不完整了，但最明显的是能看到大概 7、8 条环状的肌肉带，有的图像中还能看到内部器官。

**Egg** 很多种类生物的卵。形状大致都呈圆形。有的卵整体灰度都很深；有的卵中间有一块灰度较深的区域，四周灰度较浅（结构像细胞）。由于是不同动物的卵，因此其灰度特征差异较大。

**Fiber（纤维）** 非生物。弯曲的线状，有的纤维有分叉和交叉。该类图像中噪声较多，纤维的边缘也不是很规则。

**Gelatinous（明胶）** 该类包括很多不同种类的胶状生物，它们体内都含有很高的水分。包括 Aglaura（属于刺胞动物门）、Medusa（水母，属于刺胞动物门，水螅纲）、Siphonophora（管水母，属于刺胞动物门，水螅纲，管水母目）、



**Radiolaria**（放射虫，属于原声动物门，辐足纲）和 **Salps**（樽海鞘，属于脊索动物门 **Chordata**，樽海鞘纲 **Thaliacea**，纽鳃樽目 **Salpida**）。其中水母大部分都有三个主要部位：圆伞状或是钟状（寺院里面敲得那种钟）的身体、触器和口腕。由于该类呈胶状，因此该类物体灰度整体较浅，边缘也不是十分清晰。大部分是水母，有小部分的樽海鞘（与海樽目形态很相似），小部分的放射虫。

其中水母也包括很多类，形态大致呈现以下几种：

- 一些水母身体呈现类似钟状（这里呈现钟状有长有短，有粗有细，还有的会发生一点弯曲），灰度较浅，内部有一块颜色较深的椭圆形区域。
- 一些水母也呈钟状，但内部没有颜色较深的椭圆形区域，整个身体灰度均匀。
- 还有的个头稍微偏小，形状有的类似圆形、像半个胶囊（应该是由于拍摄原因，有的拍到顶部，有的拍到侧面），体内有颜色较深的一个大点和几个小点。（可能是灯塔水母）

放射虫：形状近似圆形（但由于整体灰度较浅，形状保存是完整），中间有一块灰度较深的区域，四周灰度较浅，可以看到淡淡的细纹从中心连接到边界。

**Multiple**（多个生物） 由于浮游动物的重叠，导致分割过程中多个浮游动物被分割到一张图像上。

**Pteropoda**（翼足目） 属于软体动物门，腹足纲。该类浮游动物灰度较深，形状总体都呈现一头宽一头窄。形状总体呈现三类：有的呈现象牙状，有点弯曲；有的较粗短，像一顶尖的小帽子；有的呈现细长的三角形状。

**Nonbio** 非生物的集合。

## 2.3 评价方法

### 2.3.1 混淆矩阵简介

在机器学习中，混淆矩阵 (CM) 是一种比较简单的对分类模型的性能进行评价的评估准则，主要用于比较分类结果和实际测得值。混淆矩阵是一个  $n$  行  $n$  列的矩阵， $n$  代表类别的数量。矩阵的每一列代表预测的每一类的数量，每一行代表实际的每一类的数量。对角线上表示分类正确的每一类的数量。例如：有 180 个样本数据，这些数据实际分为 3 类，每类 60 个，分类结束后得到的混淆矩阵如表 2-2 所示。第一行说明类 1 的 60 个样本有 53 个样本分类正确，5 个错分为类 2，2 个错分为类 3。第一列说明类 1 有 53 个样本分类正确，类 2 的 2 个样本被错分为类 1，类 3 没有样本被错分为类 1。对角线上的数据表示，类 1、2、3 分别有

53、55、59 个样本被分类正确。

表 2-2 混淆矩阵例子

	类 1	类 2	类 3
类 1	53	5	2
类 2	2	55	3
类 3	0	1	59

根据混淆矩阵可以导出以下几个参数：

- true positives (TP): 正样本被识别出的数量
- true negatives (TN): 负样本被识别出的数量
- false positives (FP): 负样本被错误分为正样本的数量
- false negatives (FN): 正样本被错误分为负样本的数量
- Accuracy: 准确率, 针对分类器的整个预测情况。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Error rate: 误分类率, 针对分类器的整体预测情况。

$$Errorrate = \frac{FP + FN}{TP + TN + FP + FN}$$

- The true positive rate (TPR) : 召回率, 正样本被识别出的概率。(文中和 PkID 中使用的的评价参数)

$$TPR = \frac{TP}{TP + FN}$$

- The false positive rate (FPR): 虚警率, 负样本被错误分为正样本的概率。

$$FPR = \frac{FP}{FP + TN}$$

- The true negative rate (TNR): 负样本被识别出的概率

$$TNR = \frac{TN}{FP + TN}$$

- The false negative rate (FNR) : 漏警率, 正样本被错误分为负样本的概率。

$$FNR = \frac{FN}{FN + TP}$$

- False discovery rate (FDR): 1-precision。(文中和 PkID 中使用的的评价参数)

$$FDR = \frac{FP}{FP + TP}$$

- Positive predictive value (PPV)

$$PPV = \frac{TP}{TP + FP}$$

- Negative predictive value (NPV)

$$NPV = \frac{TN}{TN + FN}$$

### 2.3.2 混淆矩阵分类

根据训练集和测试集的不同选取方式，混淆矩阵可以通过以下两种方式得到：

1. **Re-substitution** 混淆矩阵当采用的测试集和训练集为同一个数据集时，得到的混淆矩阵叫做 **Re-substitution** 混淆矩阵。在这个过程中，用产生的分类器对测试集进行分类时，得到的分类结果错误较少甚至可能没有错误，在进行评价时就会低估分类器的错误率。
2. **交叉验证混淆矩阵**采用交叉验证（交叉验证介绍见 2.3.3）的方法得到的混淆矩阵就叫做交叉验证混淆矩阵。并且这里采用的训练集和测试集之间在样本上是没有重合的。首先将一个数据集分成  $K$  个相等的子集，用其中  $K - 1$  个子集来训练产生分类器，用剩下的 1 个子集来进行测试，重复进行  $K$  次来构建混淆矩阵。

### 2.3.3 交叉验证

交叉验证是用来评价分类器的分类性能的一种方法，其基本步骤是是对原始图像集进行分组，然后选择没有交叉的两部分，一部分用作训练，一部分用作测试，用得到的混淆矩阵来评价分类器的性能好坏。

在机器学习的常见算法中，首要的步骤通常是将数据集分为训练集和测试集两个子集，其中训练集用来建立模型，而测试集则用来评价该模型对未知样本进行预测时的准确度，即泛化能力。如何将完整的数据集分为训练集跟测试集，必须遵守以下几点：

1. 在训练分类器阶段用到的是训练集，不能使用测试集，训练完成后在测试集上测试分类效果，将两者在分类器建立前后做到严格分离。
2. 一般来说，训练集中图像数量要多于测试集，以保证训练出的分类器有效。
3. 对于训练集和测试集，都必须做到对原始数据中每一类都均匀采样，不能过分偏向其中某几类，这样会导致训练出的分类器没有将所有类别的特点都考虑进去。

常见的交叉验证形式有 **Holdout** 验证、 $K$  折交叉验证和留一验证三种，其中用的比较多是  $K$  折交叉验证。首先将原始的整个图像集平均分成  $K$  份，一般是随

机挑选分配。每一次训练分别将其中一份作为训练集，另外的  $K - 1$  份作为测试集，得到一个混淆矩阵，由于有  $K$  份，所以会进行  $K$  次训练和测试，得到  $K$  个混淆矩阵，将这  $K$  个混淆矩阵累加起来，就能得到  $K$  折交叉验证的混淆矩阵。 $K$  的取值通常大于等于 2。在将图像集分成子集时由于是随机分配的，带有一定的随机性，所以在实际应用中，为了避免随机性对分类器效果的影响，我们通常会随机分配  $n$  次，将每一次得到的  $K$  折交叉验证的混淆矩阵进行叠加，得到最终的混淆矩阵。

本文中对分类结果采用  $K$  折交叉验证方式得到的混淆矩阵进行评价，其中  $K$  取值为 2， $n$  取值为 5，对混淆矩阵计算 Recall 和 1-Precision 值，得到每一类的分类准确率和错误率，再对所有类别求平均，得到对整个数据集进行分类的分类准确率和错误率。

## 2.4 ZooScan 系统介绍

利用机器学习方法进行浮游动物图像分类的方法有很多，目前取得最好的分类效果的是法国国家科学研究院研制的 ZooScan 系统<sup>[30]</sup>，受到该系统的启发，我们对其中的浮游动物图像特征提取和分类器设计两个关键环节进行了更深入的研究，从而对其进行了有效的改进，使分类准确率进一步提高。下面具体介绍 ZooScan 系统的主要组成部分。

### 2.4.1 简介

ZooScan Integrated System 是由法国国家科学研究院 Villefranche 海洋实验室的 Gorsky 等人发明、Hydroptic 公司生产的浮游动物样品图像扫描和处理系统，能够快速实现对液体中的浮游动物样品的扫描、计数、种类鉴定和生物量的测定，目前在国内和国外都影响很广。

ZooScan 系统是由 ZooScan、ZooProcess 和 Plankton Identifier (PkID) 三个部分组成的，其中前者是系统的硬件部分，用来对液体中的浮游动物样品进行扫描，得到高分辨率的数字图像。后两者是软件部分，先对扫描得到的图像进行一些标准化处理，对个体的一些形态、几何、灰度等参数进行提取，然后利用一些基本的机器学习算法对特征进行训练，得到分类器，最终实现对浮游动物图像的分类识别。

### 2.4.2 硬件部分

ZooScan 系统的扫描设备分为了上下两个部分（图 2-3），为了能够扫描带液体的样本，对设备的防水性有较高的要求，ZooScan 扫描系统在这方面作了很好



图 2-3 ZooScan 主扫描区

地处理，其内部隔绝水的性能相当好。扫描设备的上面部分提供了两种功能，一是照明，二是进行光密度的检测。下面部分是扫描区域，即液体样本的放置地方。由于扫描时光是从空气进去水中，再从水中进入玻璃，要分两次完成在两种不同介质之间的穿透，因而在成像效果上会相应差一些，表现为图像分辨率的降低。针对浮游动物样本的不同尺寸，ZooScan 系统提供了 1200dpi 和 2400dpi 两种分辨率以及 11cm × 24cm 和 15 × 24cm 两种大小的扫描框以供选用。

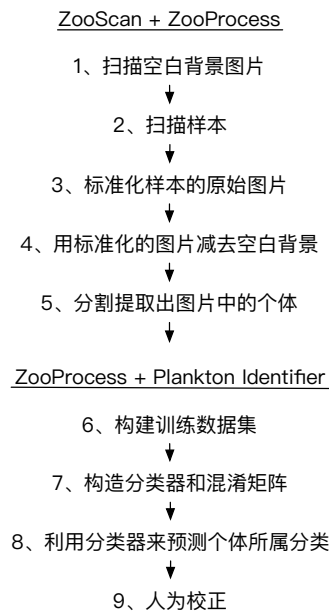


图 2-4 利用机器学习方法实现浮游动物图像分类的基本流程

### 2.4.3 软件部分

利用 ZooScan 系统对浮游动物进行扫描和分析的基本步骤如图 2-4所示。最初的步骤是使用 ZooProcess 软件完成：(i) 扫描空白背景图片，(ii) 扫描样品以获得高质量的原始图片和元数据信息，(iii) 标准化原始图片，(iv) 去掉扫描框

的边缘，以空白背景校正图片，(v) 分割提取并测量图片中的个体。随后的分析步骤是利用 ZooProcess 结合 PkID 完成：(vi) 构建训练数据集，(vii) 构造分类器和混淆矩阵，(ix) 利用分类器来预测个体所属分类，(x) 人为校正。下面会详细说明这些步骤。

### 1. 扫描空白背景图片

在扫描空白背景图片前需要打开 ZooProcess 软件创建一个新的项目，设置扫描参数，有 1200dpi 和 2400dpi 两种分辨率可供选择。然后输入样本的元数据信息。

背景图像是一张空白图像，在与样品相同环境下（自来水或是过滤海水），先扫描背景再扫描样品。并且最好是在每个扫描任务开始时，都扫描一次背景图像。首先用清水清洁和冲 ZooScan 托盘和表面玻璃，时不时地检查并清除在玻璃和扫描框上的污点。然后倒一些清水（保持在室温）没过托盘，它可以防止扫描框刮擦托盘。再放置扫描框（15cm × 24cm），这取决于之前 ZooProcess 中的“创建项目”中的选项。在预扫描和实际扫描之间，等待 30 秒。

### 2. 扫描样品以获得高质量的原始图片和元数据信息

首先需要准备好样品，存储几升清水，保持在室温，用来为 ZooScan 注水。然后用筛子（网格间隙为 100μm）过滤掉防腐剂和海水中物质。样品通过间隙为 1mm 和 200μm 的两种网格，将浮游生物分成不同体型的两部分。分别将分开后的样品，添加标签 d1 和 d2，用于扫描后的数据处理过程中。使上述两个分开后的样品中，保持只有 10001500 个浮游生物。

在扫描托盘中加入量水，放置扫描框（15cm × 24cm），调整扫描框放置的位置（位置在扫描托盘中有标注）。倒入样品，加入清水，直到没过扫描框的台阶。将个体较大的浮游生物放在扫描框的中心，用木棒分离粘连浮游生物，避免浮游生物贴靠扫描框边缘。对于漂浮在水面的浮游生物，轻轻用木棒将浮游生物压入水中。如果存在无法没入水中的浮游动物，且数量不多，则将它们移出。检查托盘是否有气泡，顶盖玻璃的表面是否冷凝。加载 ZooProcess，选择项目，点击扫描样品，选择两部分样品中的一个样品（d1 和 d2）。扫描完成后，输入相关元数据，用于将来的使用 and 比较。

### 3. 标准化原始图片

对扫描图片进行标准化可以允许使用不同参数得到的扫描图片之间进行转化和运算。对扫描图片和背景图片都要进行标准化，然后再进行相减操作。其中灰度和尺寸大小是浮游动物自动识别过程的两个最重要的变量，16 位的原始图片可以通过计算黑点（Black point, Bp）和白点（White point, Wp）

转化为 8 位图片。

$$Wp = Mg \cdot 1.15 \quad (2-1)$$

$$Bp = \frac{Mg}{1.15 \cdot \log(OD)} \quad (2-2)$$

其中,  $Mg$  (Median grey level) 表示图片的平均灰度值,  $OD$  表示光密度值。

4. 去掉扫描框的边缘, 以空白背景校正图片

对标准化后的原始图片和空白图片进行相减。

5. 分割提取并测量图片中的个体

默认的分割灰度值为 243, 等周长直径 (Equivalent Circular Diameter, ECD) 大于  $0.3mm$  的个体会被检测出来并进行下一步处理。

6. 相关参数提取

ZooProcess 中对个体提取的参数一共有 67 个: Area, Mean, StdDev, Mode, Min, Max, X, Y, XM, YM, Perim., BX, BY, Width, Height, Major, Minor, Angle, Circ., Feret, IntDen, Median, Skew, Kurt, %Area, XStart, YStart, Area\_exc, Fractal, Skelarea, Slope, Histcum1, Histcum2, Histcum3, XMg5, YMg5, Compentropy, Compmean, Compslope, CompM1, CompM2, CompM3, Symetrieih, Symetrieiv, Tag, ESD, Elongation, Range, MeanPos, CentroidsD, CV, SR, PerimAreaexc, FeretAreaexc, PerimFeret, PerimMaj, Circexc, CDexc, Nb1, Nb2, Nb3, Symetrieihc, Symetrieivc, Convperim, Convarea, Fcons, ThickR。大致可以分为位置特征、尺寸特征、灰度特征、形状特征、生物统计特征、自定义特征六大类。

(a) 位置特征

**BX、BY** 能够包围物体, 且平行于图像两条边的最小外界矩形的左上角顶点的 X 坐标和 Y 坐标。

**Height、Width** 能够包围物体, 且平行于图像两条边的最小外界矩形的高和宽。

**XStart、YStart** 图像最左上角像素点的 X 坐标和 Y 坐标。

**XM、YM** 物体灰度重心的 X 坐标和 Y 坐标。

**XMg5、YMg5** gamma 值为 51 时的物体灰度重心的 X 坐标和 Y 坐标 (gamma 值表示图像输出值与输入值关系的斜率)。

**X、Y** 物体重心点的 X 坐标和 Y 坐标。

**Angle** 浮游动物主轴与图片 x 轴形成的夹角, 在图片切割后旋转图片测量相关参数使用。

这类特征反映的是浮游动物在图像中的位置信息, 浮游动物特征与位

置信息无关, 因此它们不适合作为特征直接用于分类 (会降低分类的准确率), 而是用来计算其他特征 (尺寸特征、灰度特征和形状特征)。

(b) 尺寸特征

**Area** 包含物体的最小矩形面积。

**Perim** 周长, 物体最外层边缘的长度。

**Major** 物体的最佳拟合椭圆的长轴。

**Minor** 物体的最佳拟合椭圆的短轴。

**Feret Maximum feret diameter** (最大费雷特征), 沿物体边缘任意两个点的最长距离。

**Area\_exc** 去掉物体空洞后的表面积, 空洞是指灰度值与背景相同的部分。

**%area** 物体表面积中空洞所占的百分比, 即背景所占的比例。

这类特征表示了图像中目标的大小尺寸。它的根据是同类浮游动物的表面积、周长等尺寸特征应该是大致相同的。但是这些特征还存在着问题: 1、同类浮游动物在不同时期 (如幼年和成年) 的个体大小尺寸是不同的。2、拍摄照片的方位不同 (比如正面和侧面), 得到的尺寸特征也是不同的。

(c) 灰度特征

**Min、Max** 物体内部所有像素点的最小灰度值和最大灰度值 (0 = black, 255 = white)。

**IntDen (Integrated density)** 总密度, 物体内部像素点的灰度值的总和 ( $IntDen = Area * Mean$ )。

**Slope** 归一化的灰度累计直方图的斜率。

**Histcum1** 灰度累计直方图的值为 25% 时所对应的灰度值。

**Histcum2** 灰度累计直方图的值为 50% 时所对应的灰度值。

**Histcum3** 灰度累计直方图的值为 75% 时所对应的灰度值。

**CentroidsD**  $\sqrt{(XM - X)^2 + (YM - Y)^2}$  目标物体重心和灰度重心之间的距离。

**Fcons** 灰度对比度。

**Nb1** 图像在用阈值 Histcum1 二值化后剩余对象的数量。

**Nb2** 图像在用阈值 Histcum2 二值化后剩余对象的数量。

**Nb3** 图像在用阈值 Histcum3 二值化后剩余对象的数量。

**ThickR** 物体最大厚度和平均厚度 (不包括最大厚度) 的比值。

利用灰度特征进行浮游动物图像分类的依据是同类浮游动物的灰度特



征（灰度的范围和整体灰度变换趋势）应该是相似的。但观察图像发现并不是所有同类浮游动物的灰度都是相似的，例如 Gelatinous 类中有的个体灰度跨度较小，整体灰度都较浅，而有的个体灰度跨度较大。同时由于拍摄时光线的原因，也会造成同类浮游动物中个体灰度的深浅不一。

(d) 形状特征

**Fractal** 物体边界的分形维数，表明物体边界的不规则程度。

**Skelarea** 骨架像素的表面积（在二值图像中，不断地从物体边缘处减去像素点直到仅剩一个像素的宽度，最后所得图形的像素点数）。

**Symetrieih** 关于水平轴的对称性。

**Symetrieiv** 关于垂直轴的对称性。

**Symetrieihc** 图像在用阈值 Histcum1 二值化后物体的水平对称性。

**Symetrieivc** 图像在用阈值 Histcum1 二值化后物体的垂直对称性。

**Circ**  $Circularity = (4 * \pi * Area) / Perim^2$  圆形度，表征物体接近圆的程度，值等于 1 时，说明物体为正圆形，值越接近 0，物体体形越长。

**ESD**  $2 \times \sqrt{\frac{Area}{\pi}}$  相应球形直径（也称为等效球直径），是指一不规则外形物体，其体积相同球体的直径。

**Elongation**  $\frac{Major}{Minor}$  延伸率，最佳拟合椭圆的长轴和短轴之比。

**Circexc**  $\frac{4 \times \pi Area_{exc}}{Perim^2}$  去掉目标内部空洞的圆形度。

**Convperim** 包围物体凸包的周长。

**Convarea** 包围物体凸包的面积。

这类特征描述的是浮游动物的形状特征，依据是不同种类浮游动物之间形状不同。但在本文所采用的数据集中存在的问题是有不同种类的浮游动物形状相似，例如 Appendicularia 和 Chaetognatha，Bubble 和 Egg，也有同种浮游动物形状不同，例如 Decapoda、Gelatinous。

(e) 生物统计特征

**Mean** 物体内的平均灰度值，物体中所有像素点的灰度值的总和除以总的像素个数。

**Range**  $Max - Min$  极差，灰度的范围。

**CV**  $100 \times \frac{StdDev}{Mean}$  变异系数（也称离散系数或相对偏差），是灰度标准偏差与平均值之比，用百分数表示。

**SR**  $100 \times \frac{StdDev}{Max - Min}$  灰度标准差比上极差。

**Skew** 灰度直方图的偏度，衡量灰度分布的不对称性。正态分布的偏度为 0，概率密度函数两侧的尾部长度互相对称。偏度为负表示概率密度函数左侧的尾部相对于右侧要长，大部分的值位于平均值的右侧。偏度为正表示概率密度函数右侧的尾部比左侧的长，即处于平均值左侧的值要比右侧多。

**Kurt** 峰度，描述灰度直方图的陡缓程度。

**Mean\_exc** 物体内部去掉空洞后的平均灰度值 ( $Mean\_exc = IntDen/Area\_exc$ )。

**Median** 物体内部像素的灰度值的中值。

**StdDev** 物体内部像素的灰度值的标准差。

**Mode** 表示灰度的众数。

(f) 自定义特征

**MeanPos**  $\frac{Mean - Max}{Max - Min}$

**PerimAreaexc**  $\frac{Perim}{\sqrt{Area\_exc}}$

**FeretAreaexc**  $\frac{Feret}{\sqrt{Area\_exc}}$

**PerimFeret**  $\frac{Perim}{Feret}$

**PerimMaj**  $\frac{Perim}{Major}$

**CDexc**  $\frac{\sqrt{(XM - X)^2 + (YM - Y)^2}}{\sqrt{Area\_exc}}$

**ESD**  $2 \times \sqrt{\frac{Area}{\pi}}$

**Elongation**  $\frac{Major}{Minor}$

**Range**  $Max - Min$

**CentroidisD**  $\sqrt{(XM - X)^2 + (YM - Y)^2}$

**CV**  $100 \times \frac{StdDev}{Mean}$

**SR**  $100 \times \frac{StdDev}{Max - Min}$

$$\mathbf{Circexc} = \frac{4 \times \pi Area\_exc}{Perim^2}$$

### 3 浮游动物图像特征提取

#### 3.1 特征提取方法

图像特征提取是进行图像分析与图像识别的前提，可以通过这种方式对图像数据进行抽象和描述，给定一幅图像，我们只能得到一个  $M \times N$  的数据矩阵，而总结不出任何有效的信息，所以我们有必要对这个数据矩阵进行一定的处理，从而提取出有用的信息，方便我们作出对图像的描述。

在本节中，我们主要整理了图像处理领域的常用特征和网上分类、识别竞赛排名较为靠前的算法中用到的一些经典特征。

##### 3.1.1 图像处理领域常用特征

1. 边界的周长：轮廓边界的周长，对轮廓边缘上的像素点进行统计。
2. 边界的曲率：表示目标物体边缘轮廓曲线的弯曲程度。
3. 面积：描述区域大小的特征，对区域内总像素点的数目进行统计。
4. 宽度和高度：最小外接矩形的宽度和高度。
5. 矩形度：表示目标物体的整个边缘轮廓在其最小外接矩形中的充满程度，当目标的轮廓形状越接近矩形时，矩形度的值越接近 1。

$$R = \frac{A}{WH} \quad (3-1)$$

其中，A 为目标的面积，W、H 分别为最小外接矩形的宽度和高度。

6. 体态比：目标最小外接矩形的长与宽的比值。

$$C = \frac{W}{H} \quad (3-2)$$

7. 圆形性：用目标区域的所有边界点定义的特征向量。

$$C_I = \frac{\mu_R}{\sigma_R} \quad (3-3)$$

$\mu_R$  为区域重心到边界点的平均距离， $\sigma_R$  为从区域重心到边界点的距离的平均方差。

8. 偏心率：在一定程度上反映了物体轮廓形状与理想圆环的偏离。定义为焦点间距离  $p$  与长轴  $q$  的比值。

$$E = \frac{p}{q} \quad (3-4)$$

9. 凸率：表征物体边界轮廓的不规则程度，为目标区域面积与目标区域凸包面积两者的比值。

$$C_R = \frac{A}{\sum_{x=1}^M \sum_{y=1}^N k(x, y)} \quad (3-5)$$

分母为凸包区域的面积。

10. 密集度：表示目标物体围成的区域的密集程度。

$$C = \frac{A}{P^2} \quad (3-6)$$

$A$  表示物体区域的面积， $L$  表示物体轮廓的周长。圆形是所有图形中密集程度最高的，因此密集度的最大值为  $1/(4\pi)$ 。

11. 球状性：内切圆的直径与外接圆的直径之比。

$$S = \frac{r_i}{r_c} \quad (3-7)$$

12. 伸长度：周长与目标区域最小外接矩形面积之比。

$$P = \frac{L}{WH} \quad (3-8)$$

### 3.1.2 经典特征

目前每年都会举办的视觉方面比较著名的挑战赛有 PASCAL VOC 挑战赛、ImageNet 大规模视觉识别挑战赛和户外脸部检测竞赛（Labeled Faces in the Wild, LFW）。整理这些竞赛中取得较好效果的一些算法中用到的特征如下：

#### 1. SIFT 特征

SIFT（Scale invariant feature transform，尺度不变特征变换）是一种检测局部特征的算法，该算法通过求得一幅图像中的特征点及有关尺度和方向的描述子得到特征并进行图像特征点匹配，它由 David Lowe 在 1999 年<sup>[31]</sup>所发表，2004 年<sup>[32]</sup>完善总结。

算法主要步骤如下：

- (a) 构建尺度空间，检测极值点，获得尺度不变性。

利用不同参数的高斯微分函数可以在不同尺度下对要处理的图像进行滤波，将滤波后的图像与原图像进行对比，其中有明显变化的像素就是我们所感兴趣的特征点。

对原始图像用不同参数的高斯微分函数进行滤波所得到的一组图像空间表示如下：

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3-9)$$

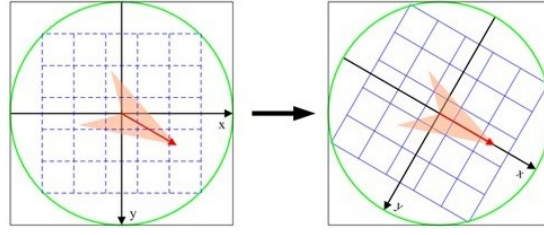


图 3-1 将坐标移至关键点主方向

其中  $I(x, y)$  是原始图像,  $\sigma$  大小决定图像的平滑程度,  $G(x, y, \sigma)$  是尺度可变高斯函数:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (3-10)$$

为了使特征点的检测更为稳定, 我们可以在高斯尺度空间上对图像进行滤波处理, 也就是用不同参数的高斯差分核与图像进行卷积。

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3-11)$$

(b) 确定特征点所在的位置。

(c) 指定每个特征点的方向参数。

以特征点为中心, 其邻域内像素的梯度方向分布是具有一定特点的, 可以帮助我们为每个特征点设置方向参数, 使得对特征点的描述子在不同方向上都能保持不变, 方向参数的求取是通过计算每个极值点的梯度得到的。在某点  $(x, y)$  处的梯度大小和方向的计算公式如下:

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2} \quad (3-12)$$

$$\theta(x, y) = \alpha \tan 2 \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (3-13)$$

(d) 求取特征点描述子。

首先确定计算描述子所需的区域, 即确定特征点邻域范围, 区域的半径的计算公式如下:

$$radius = \frac{3\theta_{oct} \times \sqrt{2} \times (d+1)}{2} \quad (3-14)$$

然后, 将坐标移至关键点主方向, 如图 3-1。

旋转后邻域内采样点的新坐标为:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \sigma & -\sin \sigma \\ \sin \sigma & \cos \sigma \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3-15)$$

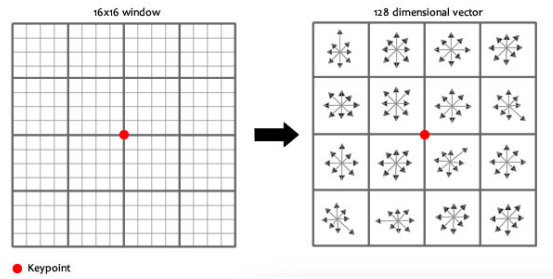


图 3-2 描述子用 128 维向量表征

下面就要对每个特征点求取方向直方图了，以特征点为原点，求其邻域内像素点的梯度方向直方图。梯度方向直方图是在  $0 \sim 360$  度的范围内，通常平均分为 8 等份，直方图中每一份是  $45$  度的范围。以特征点为中心，其周围  $16 \times 16$  的邻域内，计算每一个像素的梯度，看其梯度方向在哪一个度数范围内，直方图中就给相应的那一份加上一个权重，这里的权重就是每个像素梯度的大小。将  $16 \times 16$  的邻域再分为  $4 \times 4$  个区域，每个区域中有 16 个像素点，计算每个区域的梯度方向直方图，如图 3-2 所示。这样对每个特征点就得到一个  $4 \times 4 \times 8 = 128$  维的特征描述子。

再以特征点为原点，求其邻域内像素点的梯度方向直方图。梯度方向直方图是在  $0 \sim 360$  度的范围内，通常平均分为 8 等份，直方图中每一份是  $45$  度的范围。假设邻域范围为  $16 \times 16$ ，计算其中每一个像素的梯度，看其梯度方向在哪一个度数范围内，直方图中就给相应的那一份加上一个权重，这里的权重就是每个像素梯度的大小与高斯权重的乘积。当邻域范围为  $16 \times 16$  时，我们通常将其又分为  $4 \times 4$  个区域，计算每一个区域中的 8 个方向的梯度方向直方图，即可形成一个种子点，一共可以生成 16 个种子点。

描述子向量元素门限化及门限化后的描述子向量规范化。

根据特征点的尺度对特征描述向量进行排序，生成最终的 SIFT 特征向量。

## 2. 方向梯度直方图特征

方向梯度直方图<sup>[33]</sup> (Histogram of Oriented Gradients, HOG) 特征是由 Navneet Dalal 等人于 2005 年提出的一种特征描述子，它是一种局部特征，主要通过统计图像上局部区域内的梯度方向来构成直方图，最后将所有区域内的直方图串联而成。方向梯度直方图特征结合支持向量机分类算法生成分类器在图像识别方面取得了很好的效果，特别是在行人检测的应用中更是起到了重要作用。对方向梯度直方图进行求解的具体步骤如图 3-3。

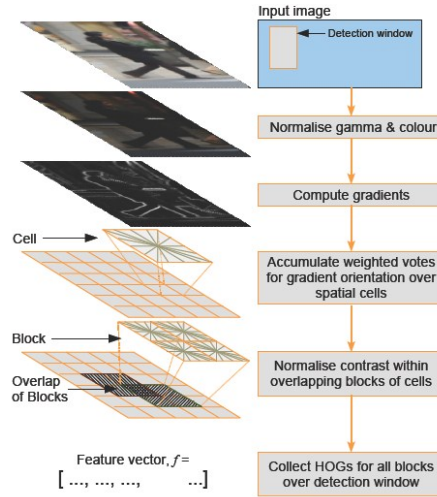


图 3-3 HOG 特征

- (a) 首先对图像进行标准化或归一化操作。由于拍照光线不好、角度问题以及噪声等的影响，所得到的图像会有对比度低、局部昏暗等问题，对于随后的特征提取会有一定的干扰。通常采用 **Gamma** 校正法来对颜色空间进行标准化，**Gamma** 压缩公式为：

$$I(x, y) = I(x, y)^{gamma} \quad (3-16)$$

- (b) 利用一阶微分计算每一个像素的梯度（包括大小和方向），梯度方向同时也表示了目标物体的轮廓边界方向。

采用模板  $[-1 \ 0 \ 1]$  计算水平和垂直方向的梯度：

$$G_h(x, y) = f(x + 1, y) - f(x - 1, y) \quad \forall x, y \quad (3-17)$$

$$G_v(x, y) = f(x, y + 1) - f(x, y - 1) \quad \forall x, y \quad (3-18)$$

梯度值和梯度方向的计算如下：

$$M(x, y) = \sqrt{G_h(x, y)^2 + G_v(x, y)^2} \quad (3-19)$$

$$\theta(x, y) = \arctan \frac{G_h(x, y)}{G_v(x, y)} \quad (3-20)$$

- (c) 将图像划分成小的单元（cell）（例如  $6 \times 6$  像素/单元）。
- (d) 统计每个单元内所有像素点的梯度直方图，将其作为每个单元的描述子。首先将梯度方向分为  $m$  份（共  $360$  度），如果分为了  $16$  份（如图 3-4），则每一份中是  $22.5$  度。然后看每个像素点的梯度方向是落在了哪一份中，对应直方图中的那一份就加上一个权值，这里的权值就是



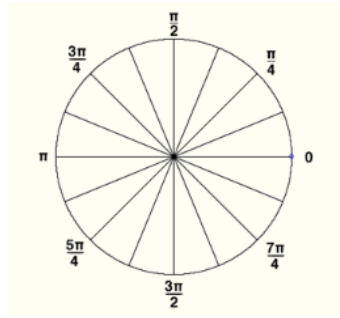


图 3-4 梯度方向 bin

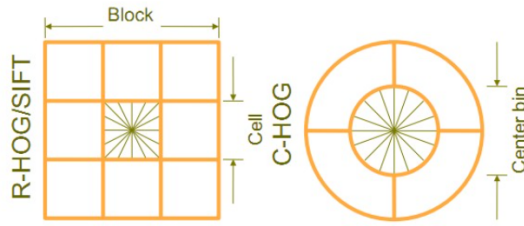


图 3-5 矩形区间和环形区间

像素点的梯度大小。

- (e) 相邻的几个单元可以组成一个更大的块 (block) (例如  $3 \times 3$  个 cell/block), 再将一个块内所有单元的梯度直方图串联起来, 最后得到该块的方向梯度直方图特征描述子。在将单元组合成块时, 块与块之间是可以有重合的, 这样, 每个单元的梯度直方图都会被多次使用, 以串联生成块的梯度直方图特征。其中块既可以按矩形划分 (R-HOG), 也可以按环形划分 (C-HOG), 具体见图 3-5。
- (f) 将图像内的所有块的梯度方向直方图特征串联起来就可以得到整幅图像的梯度方向直方图特征描述子了, 这个就是最终的可供分类使用的特征向量。该向量的维数为  $m \times n \times \alpha$ , 其中  $m$  表示每个单元中方向 bin 的数目,  $n$ 、 $\alpha$  分别表示块的个数以及一个块中的单元数目。

### 3. SURF

SURF 与 SIFT 算法相似, 其优点是运算简单, 效率高, 缺点是不够稳定, 并且检测出的特征点数目少一些。

算法步骤如下:

- (a) 构建 Hessian 矩阵。采用的是 Hessian 矩阵行列式近似值图像。
- (b) 构造多尺度空间。

构造对尺度空间的目的是实现尺度不变性, SURF 算法则用高斯滤波与 Hessian 矩阵结合近似实现多尺度空间, 计算复杂度比 SIFT 低。

- (c) 根据非极大值抑制初步确定特征点。

(d) 主方向确定

(e) 构造特征点算术描述子，SURF 算法是选取一个  $20S$ （单位）的区域，分成  $4 \times 4$  分，每一份中有  $5 \times 5S$ ，统计一份中的  $\sum d_x$ ， $\sum d_y$ ， $\sum |d_x|$ ， $\sum |d_y|$ ，这样就得到  $4 \times 4 \times 4 = 64$  的向量描述子。SIFT 是 128 个描述子，比 SURF 复杂一点。

#### 4. 局部二值模式特征

局部二值模式（Local Binary Pattern, LBP）<sup>[34]</sup> 是一种用来表征图像局部纹理特征的描述子，具有旋转不变性和灰度不变性等优点。

对图像中的某个像素点计算其局部二值模式值的步骤如下：以该像素点为中心，将其灰度值与其周围  $3 \times 3$  窗口的邻域内的 8 个像素点进行大小比较，如果邻域内像素的灰度值大于中心点的灰度值，则该邻域内像素点的值被置为 1，否则置为 0。经过这一阈值处理后，对每一个像素点，就得到了其邻域 8 个像素点的二进制表示（方向是从左上方的像素点开始往右顺时针走），再求出该二进制数的十进制表示，这就是该中心像素点上的局部二值模式值，如图 3-6。

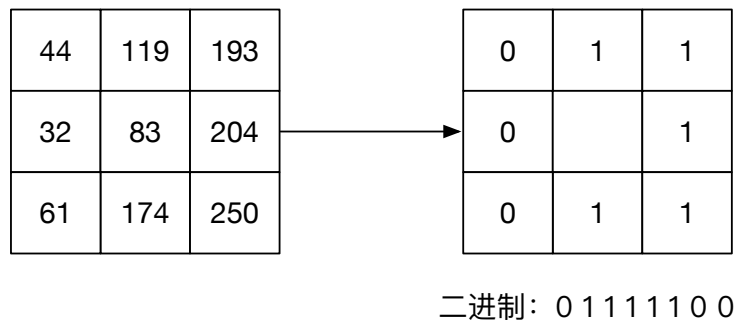


图 3-6 LBP 特征

#### 5. FV(Fisher vector)

Fisher vector 本质上是用似然函数的梯度 vector 来表达一幅图像。主要思路是用生成式模型 (GMM) 对样本输入进行建模，进而得到样本的一种表示 (fisher vector)，再将这种表示 (fisher vector) 输入判别式分类器 (SVM) 得到图像分类结果。fisher vector 是 fisher kernel 中对样本特征的一种表示，它把一幅图片表示成一个向量。在高斯分布的基础上再找到变化的方向，可以更加准确的表示这一张图。

#### 6. 骨架特征

骨架又称为中轴，它的定义来自于烧草模型和最大圆盘模型。骨架不仅包含了物体的几何特征，也描述了物体的拓扑结构。经典的骨架化算法有中轴变

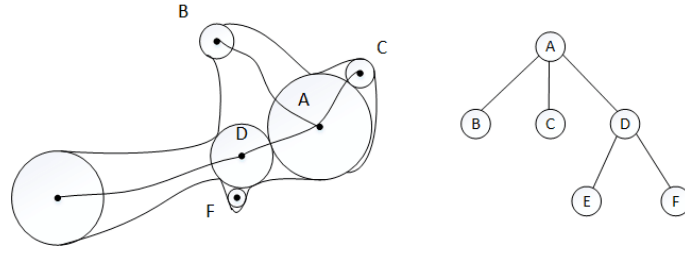


图 3-7 骨架树

换、细化算法、Voronoi 图算法等。在骨架的基础上可以通过奇点图、骨架树等对物体的结构特征进行描述。

骨架树：将骨架映射到一个树状结构（如图 3-7）。骨架上的点可以分为三类：分叉点、端点和连接点。其中当骨架点的邻接点的数目为 1 时被称为端点，邻接点数目为 2 个或 2 个以上时被称为分叉点，除此之外的点被称为连接点。根据骨架树深度优先搜索产生的节点序列中的所有节点用其孩子数替换，替换后得到的新序列即为树描述符。例如，图 3-7 的树描述符为 (3, 2, 0, 0, 0, 0)。

#### 7. 形状上下文特征

形状上下文 (Shape Context, SC) 由 Belongie 等人<sup>[35]</sup>提出，是基于物体轮廓样本点进行描述的。下面给出形状上下文求解的具体步骤。首先，对一幅图像（图 3-8(a)）中的目标进行边缘提取，得到结果如图 3-8(b) 所示。然后将这一轮廓看作是一定数目的采样点的集合（图 3-8(c)），利用这些采样点来描述目标形状。通过对数极坐标空间（图 3-9(a)）计算每个采样点与形状其他点之间的位置关系，采用直方图表示。假设目标轮廓上采样点的集合定义为  $O = \{o_1, \dots, o_n\}$ ，其中  $n$  表示采样点的个数。假设目标轮廓是分段平滑的，当  $n$  充分大时，可以基本与物体实际轮廓贴合。若第  $i$  个采样点  $o_i$  作为参考坐标原点（图 3-9(b)），建立对数极坐标映射表示剩余的  $n - 1$  个采样点的分布情况，可以看到周围与它相邻的点（在极坐标覆盖的范围之内）落于不同的小格子内，表示不同的相对向量，这些相对向量刻画了整个形状与参考点的相对位置，因而就成为这个点的形状上下文，用直方图  $h_i$  表示。

$$h_i(k) = \#\{j \neq i : (o_j - o_i) \in \text{bin}(k)\} \quad (3-21)$$

以  $o_i$  为极坐标中心进行坐标变换，以该中心画一个圆，并把圆等分成  $k$  份，计算圆范围内的点落在每一份中的个数，这样对每一个点就得到一个  $k$  维的直方图。 $\theta$  为极角， $r$  为极坐标半径，对于每一个由  $\theta$  和  $\log r$  确定的极坐标区域，如果该区域包含的点越多，那么在由  $\theta$  和  $\log r$  确定的对应形状直方图

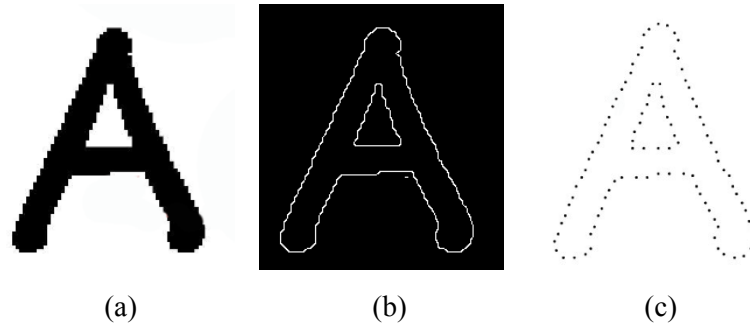


图 3-8 形状描述

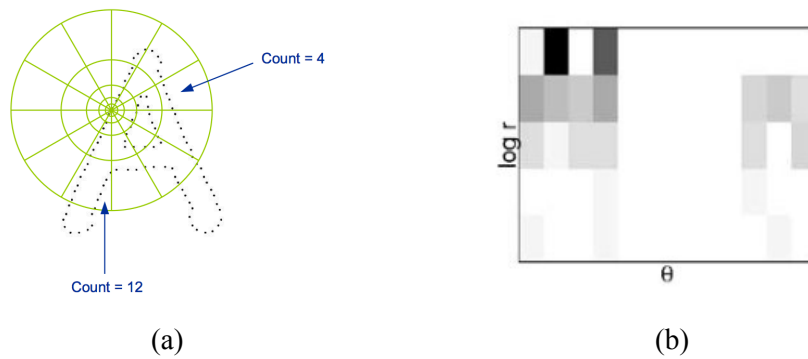


图 3-9 形状上下文的直方图表示

中的数值就越大。

只要每个点的形状上下文出来了，那么对所有点的形状上下文组合起来就可以形成整个物体的形状上下文了。

在物体分类识别的应用中，是这样利用形状上下文特征的：首先对图像数据集中的每一类别都找一幅模板图像，该模板能较好地反应其所属类别图像的共性。假设对一幅图像的物体轮廓采样了 100 个点，就得到一个  $100 \times k$  的矩阵。同样，对模板也能得到一个  $100 \times k$  的矩阵。再计算图像中的每个点与模板中的每个点的距离，得到图像中的点和模板中的点的一一对应关系，结果是一个  $100 \times 100$  的矩阵。最后根据这个一一对应关系计算出一个值，表示该图像和模板之间的相似度。计算出每幅图像和所有类别模板之间的相似度，当做该幅图像的一个特征向量，利用分类算法进行训练，可以生成一个分类器。

### 3.1.3 其它特征

#### 1. 局部对称性

局部对称性的原理是基于周围像素的灰度梯度，计算在给定点  $p$  处的对称性程度。如图 3-10所示，点  $p$  处的对称性可以通过比较位于  $p_i$  和  $p_j$  处的像素对  $i$  和  $j$  的灰度梯度来度量，其中  $p = (p_i + p_j)/2$ 。每对像素对对点  $p$  处的

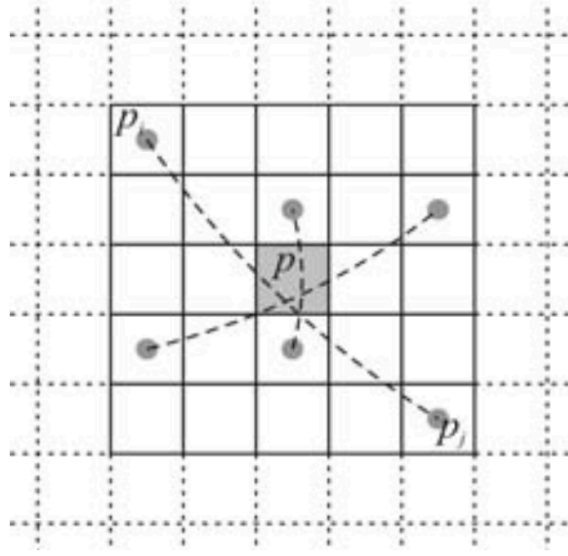


图 3-10 用对称性算子来对像素对的梯度进行比较的三个例子

局部对称性的贡献为

$$c(i, j) = d(i, j, \sigma) \cdot p(i, j) \cdot m_i \cdot m_j \quad (3-22)$$

$m_i$  是像素点  $i$  的梯度的模,  $d(i, j, \sigma)$  是当标准差为  $\sigma$  时, 两个像素点间的距离的高斯加权函数, 对称性

$$p(i, j) = (1 - \cos(\gamma_i + \gamma_j)) \cdot (1 - \cos(\gamma_i - \gamma_j)) \quad (3-23)$$

其中,  $\gamma_i = \theta_i - \alpha$ , 见图 3-11。当点  $p_i$ 、 $p_j$  关于点  $p$  对称时,  $\gamma_i + \gamma_j = \pi$ , 式 (3-23) 中的第一项取得最大值。如果只用第一项, 会使得在直线边缘上的点取得较大的值, 而这些点通常并不具有对称性。为了避免这个问题, 又加入了第二项, 表示具有相似梯度方向的像素对。最后, 对所有像素对的贡献进行求和得到点  $p$  处的 isotropic symmetry 值

$$M^{iso}(x, y) = \sum_{(i, j) \in \Gamma(p)} c(i, j) \quad (3-24)$$

为了使 symmetry operator 对有多个对称轴的对称图案更敏感, Reisfeld 等人<sup>[36]</sup> 又提出了 radial symmetry operator, 作为 isotropic symmetry operator 的扩展。首先, 像素对的贡献的方向用  $\varphi(i, j) = (\theta_i + \theta_j)/2$  来计算。然后, 对最大贡献为  $c(i, j)$  的像素对  $(i, j)$ , 对称方向被确定为  $\phi(p) = \varphi(i, j)$ 。这个值也用来提升有不相似方向的像素对的贡献。

$$M^{rad} = \sum_{(i, j) \in \Gamma(p)} c(i, j) \cdot \sin^2(\varphi(i, j) - \phi(p)) \quad (3-25)$$

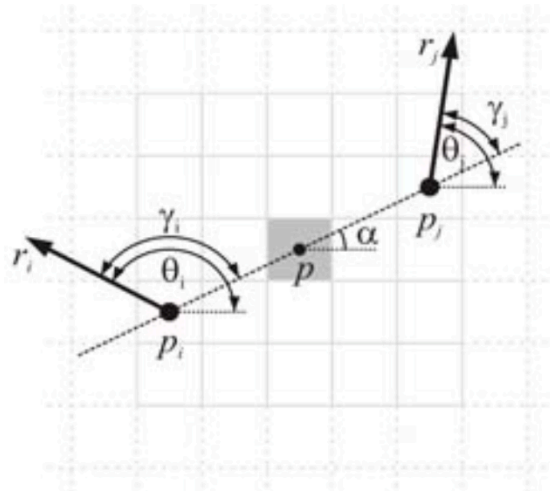


图 3-11 每对像素对的贡献的几何表示

## 2. 灰度直方图

灰度直方图是针对灰度图像的，如果是彩色图像，可以求取其颜色直方图。若灰度级的范围为  $0 \sim 255$ ，则灰度直方图表示这 256 个灰度值中的每一个在图像中出现的个数。

## 3.2 特征选择

2.4.3 和 3.1 节的介绍中基本涵盖了近年来图像处理和机器视觉领域常用的一些特征，其中有些特征适用于浮游动物图像的分类问题，有些并不适用，要对这些特征进行筛选。本文中的特征选择过程是通过对单个特征及特征组合的理论分析与实验结合，判断一个特征或一组特征对浮游动物图像分类效果的影响。其中实验部分是对 2.2 节中介绍的包含 13 类浮游动物的数据集分别用两种基本的分类算法（支持向量机和随机森林）进行分类，对分类结果利用  $K$  折交叉验证方法进行评价，看在数据集上的分类准确率和错误率有没有分别提高和降低。

对前面所总结的所有特征进行实验，我们最终选择了三类特征：PkID 中用到的特征、LBP 特征以及内距离形状上下文特征。相对于其它特征而言，这三类特征在对浮游动物图像的分类上效果较好，其中 PkID 中用到的特征数目较多，可对其进行进一步筛选。

### 3.2.1 PkID 中的 22 个特征

由 2.4.3 节可以看到，PkID 中共采用了 67 个特征，分为 6 大类：位置特征、尺寸特征、灰度特征、形状特征、生物统计特征和自定义特征。这些特征都是带有一定意义的，通过实验和分析，我们得出结论：位置特征不能表征浮游动物的属

性，会降低分类准确率。而由于我们采用的数据集是经过扫描得到的，不同浮游动物类别有透明和不透明之分，表现在扫描图像上就是灰度值高和灰度值低，因而灰度特征可作为浮游动物图像分类的一个重要特征。由于数据集中对每幅图像都做了标尺，不同类别的浮游动物在尺寸上应该是互不相同的，但由于处于不同成长期的浮游动物以及拍摄角度（正面和侧面）的影响，会使得尺寸特征失去了本应有的效果。形状特征是表征一类物体的较为重要的特征，对于不同种类有一定的区分度。生物统计特征近年来在分类识别中使用的也比较多，效果较好，但其物理意义不是十分明确，难以从理论上说明其有用的原因，只能从实验结果上反应。

表 3-1 对 PkID 中的 6 类特征用支持向量机和随机森林方法进行分类的分类准确率

	支持向量机	随机森林
位置特征	15.3%	30.23%
尺寸特征	26.04%	56.95%
灰度特征	37.6%	61.3%
形状特征	37%	63.1%
生物统计特征	52.4%	69.9%
自定义特征	68.05%	69.03%

我们对这 6 类特征分别用支持向量机和随机森林分类算法进行实验，得到的分类准确率如表 3-1 所示。效果比较好的是灰度特征、形状特征、生物统计特征和自定义特征，基本符合我们对于特征的分析，从而可以将位置特征和尺寸特征筛除掉。对剩下的 4 类特征进行进一步实验，不断地挑选一定数量的特征，用支持向量机和随机森林分类算法进行验证，最终我们从 67 个 PkID 特征中挑出了 22 个，如表 3-2 所示，可以看到只选取了生物统计特征、形状特征、自定义特征三种，其中一部分自定义特征是根据其他几类特征来自定义产生的，所以这里没有选取灰度特征，分类效果依然很好。

表 3-2 PkID 中的 22 个特征

生物统计特征	Mean、StdDev、Skew、Kurt、CV、SR
形状特征	Fractal、Skelarea、Circ、Symetrieih、Symetriev、Symetriehc、Symetrievc、Elongation2、Circexc、Convperim、Convarea
自定义特征	MeanPos、PerimAreaexc、FeretAreaexc、PerimFeret、CDexc

### 3.2.2 局部二值模式

对 LBP 特征向量进行提取的步骤:

1. 首先将检测窗口划分为  $16 \times 16$  的单元 (cell)。
2. 对于每个单元中的一个像素, 将相邻的 8 个像素的灰度值与其进行比较, 若周围像素值大于中心像素值, 则该像素点的位置被标记为 1, 否则为 0。这样,  $3 \times 3$  邻域内的 8 个点经比较可产生 8 位二进制数, 即得到该窗口中心像素点的 LBP 值。
3. 然后计算每个单元的 LBP 值直方图, 即每个数字 (假定是十进制数 LBP 值) 出现的频率, 然后对该直方图进行归一化处理。
4. 最后将得到的每个单元的统计直方图进行连接成为一个特征向量, 也就是整幅图的 LBP 纹理特征向量。

### 3.2.3 内距离形状上下文

本文 3.1.2 节中介绍了关于形状上下文的内容, 由于形状上下文不能很好地解决物体类内部之间的变形, 后期产生了改进的基于内部距离的形状上下文<sup>[37]</sup>, 内距离形状上下文中采用内距离代替欧式距离。

采用内距离形状上下文提取图像特征的基本步骤如下:

1. 先从图像中挑选  $m \times n$  张图像作为模板 (每类浮游动物中选取  $m$  张, 共  $n$  类)。
2. 采用内距离形状上下文分别计算训练集和测试集中所有图像和上一步骤中  $m \times n$  张图像间的距离。
3. 将上面计算得到的距离矩阵作为训练集和测试集的特征, 输入到 ELM 中进行学习和分类。



## 4 分类器设计

### 4.1 常用的分类技术

#### 4.1.1 支持向量机方法

支持向量机（Support Vector Machine, SVM）<sup>[38]</sup> 是一种有监督的统计学习方法，能够最小化经验误差和最大化几何边缘，被称为最大间隔分类器，可用于分类与回归分析。SVM 的基本原理是在高维空间中寻找一个分类超平面，将不同类别的数据样本点分开，使不同类别的点之间的间隔最大，该分类超平面即为最大间隔超平面，对应的分类器称为最大间隔分类器。针对二分类问题，图 4-1 可以直观地描述 SVM 的空间特征。

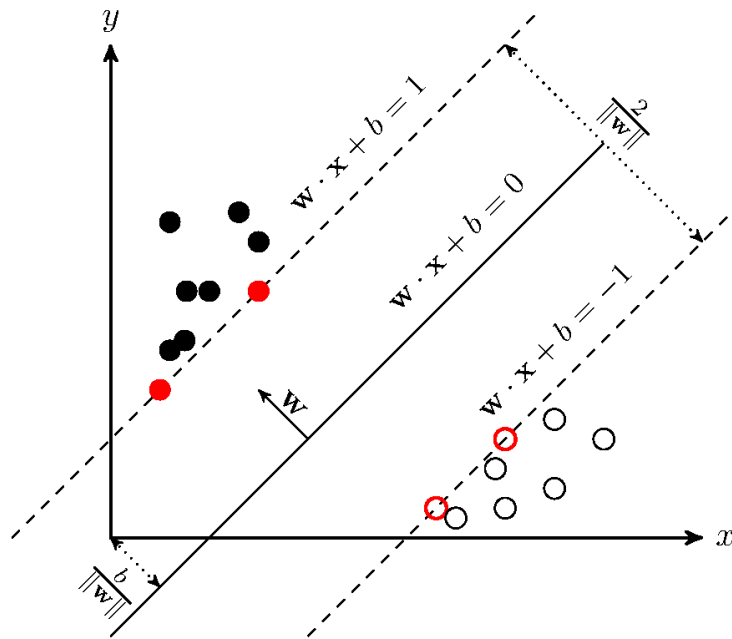


图 4-1 SVM 的训练数据样本及空间特征

假设数据样本为  $x_1, x_2, \dots, x_n$ ，分类超平面可以表达为：

$$w_T x - b = 0 \quad (4-1)$$

其中， $x$  为分类超平面上的点； $w$  为垂直于分类超平面的向量； $b$  为位移量，用于改善分类超平面的灵活性，超平面不用必须通过原点。

两个分类超平面之间具有最大间隔，需要知道训练样本中的支撑向量、距离

支撑向量最近的平行超平面，这些平行超平面可以表示为：

$$\begin{cases} w^T x - b = 1 \\ w^T x - b = -1 \end{cases} \quad (4-2)$$

其中， $w$  是分类超平面的法向量，长度未定；1 和 -1 只是为了计算方便而取的常量，其他常量只要互为相反数亦可。

如果给定的训练样本是线性可分的，那么就可以找到两个间距最大的平行超平面，并且这两个超平面之间没有任何训练样本，它们之间的距离为  $2/\|w\|^2$ 。所以最小化  $\|w\|^2$ ，就可以使这两个超平面之间的间隔最大化。

$$\min \frac{1}{2} \|w\|^2 \quad (4-3)$$

$$s.t. y_i(w^T x_i + b) \geq 1 \quad (4-4)$$

通过将目标函数转化，求解 SVM 可以变成一个标准的凸优化问题。因为目标函数是二次的，而约束条件是线性的，所以这是一个凸二次规划问题。虽然这个问题是一个标准的二次规划问题，但同时具有它的特殊结构，通过将此二次规划问题转化为其对偶问题，可以找到更加有效的方法进行求解。整个对偶问题可表示为：

$$\max_a \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \langle x_i \cdot x_j \rangle \quad (4-5)$$

$$s.t. 0 \leq \alpha_i \leq C, i = 1, \dots, n \quad (4-6)$$

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (4-7)$$

其中， $\alpha$  为引入的拉格朗日乘子。经计算，最优权值向量  $w^*$  和最优偏置  $b^*$  分别为

$$w^* = \sum_{i=1}^n \alpha_i^* y_i x_i \quad (4-8)$$

$$b^* = y_i - \sum_{i=1}^n y_i \alpha_i^* (x_i \cdot x_j) \quad (4-9)$$

其中， $i \in \{i | \alpha_i^* > 0\}$ 。因此得到最优分类面  $(w^* \cdot x) + b^* = 0$ ，而最优分类函数为

$$f(x) = \text{sgn}\{(w^* \cdot x) + b^*\} = \text{sgn}\left\{\left(\sum_{i=1}^n \alpha_i^* y_i (x_i \cdot x_j)\right) + b^*\right\} \quad (4-10)$$

#### 4.1.2 随机森林方法

在介绍随机森林方法之前，首先介绍决策树的概念，决策树是一种树，树中每个节点用于表示某个对象，每个分叉路径用于代表某个可能的属性值，从根节

点到某叶节点所经历的路径所表示的对象值则对应该叶节点。

为了提高识别精度，必须要通过某种衡量准则选择特征，使得分割后的数据集的标签信息增益最大。信息增益是用于衡量给定属性区分训练样本能力的度量值，它的求解是用原始数据集标签熵减去分割后的数据集标签熵，如果熵变小了，则信息增益变大，表示给定属性对数据集的分类带来了一定的信息量。熵是信息论中广泛使用的一个度量标准，用于表示一件事情的不确定性，当要搞明白一件事所要求的信息量越高，这件事的不确定性越大，因而熵越大。假设  $S$  是包含关于某个目标概念的样本集，目标属性具有  $m$  个不同的种类，那么  $S$  相对于这  $m$  个状态类别的熵定义为：

$$Entropy(S) = \sum_{i=1}^m -p_i \log_2 p_i \quad (4-11)$$

其中， $p_i$  为  $S$  中不同类别样本的比例。

一个属性的信息增益可以看作由于使用这个属性分割样本从而导致的期望熵降低，一个属性  $A$  相对样本集合  $S$  的信息增益  $Gain(S, A)$  计算公式为：

$$Gain(S, A) = Entropy(S) - \sum_{v \in V(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (4-12)$$

其中， $V(A)$  是属性  $A$  的值域， $S$  是样本集合， $S_v$  是  $S$  在属性  $A$  上值等于  $v$  的样本集合。

整个处理流程如下：首先计算每个属性的信息增益，然后依照信息增益最大原则选取属性，作为分割样本的属性。接着构建决策树，采用递归思想依次分割下去，直到执行完成就构建好了决策树。最后依据决策树进行分类，给出序列化决策树。

随机森林是一种采用决策树作为基预测器的集成学习方法，2001年由 Breiman 提出<sup>[39]</sup>，通过自助重采样技术，从总数量为  $N$  的原始训练样本集中有放回地重复随机选取  $N$  个样本，由此得到一个自助训练数据集，按此方法重复得到一组自助训练数据集。对每个自助训练集，用决策树学习方法（如 C4.5、ID3、CART 算法）创建一颗完全生长的无剪枝处理的决策树，进而得到一组决策树的集合。

对每个节点按固定概率分布来随机选择数目固定为  $m$  的特征（假设训练样本的总特征数目为  $M$ ， $m \ll M$ ），作为在该节点寻找最佳分裂的依据。在预测时，根据所有决策树对新数据预测结果的平均值或根据投票多少来决定。

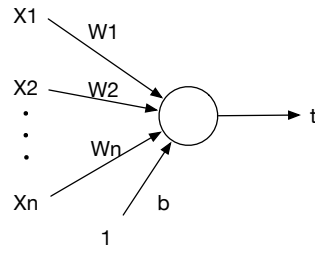


图 4-2 单隐藏层反馈神经网络

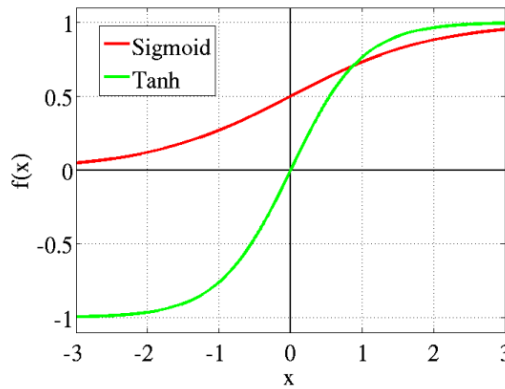


图 4-3 sigmoid 和 tanh 函数的图像

#### 4.1.3 极限学习机方法

极限学习机 (Extreme Learning Machine, ELM) 是一种新型神经网络算法, 最早是由新加坡南洋理工大学的黄广斌教授于 2004 年提出来的<sup>[40]</sup>。极限学习机的前身是单隐藏层反馈神经网络, 在此基础上加上了一种新型的快速学习算法, 使得学习速度得到了很大程度的提高。

神经网络的最小单位为“神经元”, 一个最简单的神经网络可以仅仅由一个“神经元”组成, 如图 4-2所示。

其中, 第一列为输入向量, “1”表示截距,  $b$  表示偏置, 箭头上的  $w$  为买个输入向量分量对应的权重值, 最右端为输出。其输入-输出关系用数学公式表示为  $y = g(w \cdot x + b)$ , 其中函数  $g$  叫做“激励函数”, 通常采用 sigmoid 函数或 tanh 函数, 这两个函数的表达式分别如下:

$$g(z) = \frac{1}{1 + \exp(-z)} \quad (4-13)$$

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (4-14)$$

它们对应的图像分布见图 4-3。

可以看出, 这两个函数都是逻辑函数, 因而这里的输入-输出关系是一种逻辑

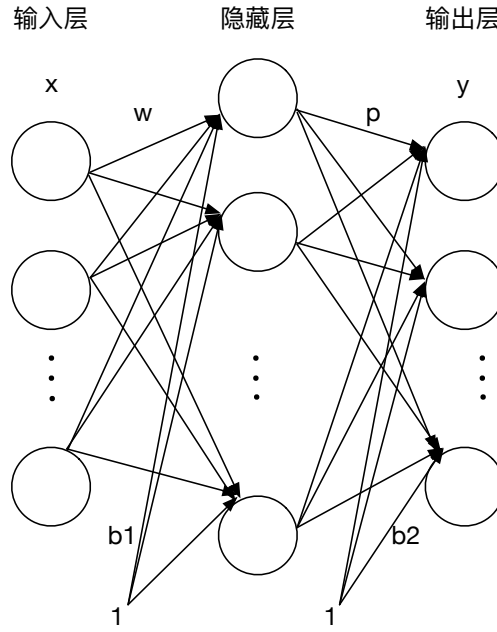


图 4-4 单隐藏层反馈神经网络

回归。

稍微复杂一点的神经网络是联结不只一个的神经元，常见的三层神经网络（单隐藏层反馈神经网络）如图 4-5 所示。

其中最左边一层叫做输入层，中间的一层叫做隐藏层，最右边一层叫做输出层。只有输入层和输出层可以跟外界进行联系，而隐藏层中的内容就对应一个黑盒子，外界无法观测到其值。从左往右来看，左边一层的输出总是连接着右边一层的输入，因而输出层第  $k$  个神经元的响应可以表示为：

$$y[k] = [g(w_1 \cdot x + b_1) \cdot g(w_2 \cdot x + b_2) \cdot \dots \cdot g(w_L \cdot x + b_L)] \cdot p_k + b_2[k], k = 1, \dots, m \quad (4-15)$$

其中， $w_i$  是一个向量，表示隐藏层第  $i$  个神经元的输入权重， $p_i$  也是一个向量，表示输出层第  $k$  个神经元的输入权重。

单隐藏层反馈神经网络的主要优点为：1) 可以充分拟合出任何复杂的输入-输出关系；2) 每一层中单个的神经元以及一条连接对整个神经网络的作用是很小的，当少量神经元和连接发生错误时，对整个网络的影响不大，这样就增强了网络的容错性。其缺点是学习速度较慢，在时间消耗上远高于我们所能接受的最多时间。为了解决这一问题，很多研究者对单隐藏层反馈神经网络进行了进一步的探索，从而在运行时间上有了改进。

黄广斌教授通过对单隐藏层反馈神经网络的进一步研究发现，当“激励函数”  $g$  能够在任何一个区间内都满足无限可微的条件时，我们就可以随机地生成  $w_i$  和

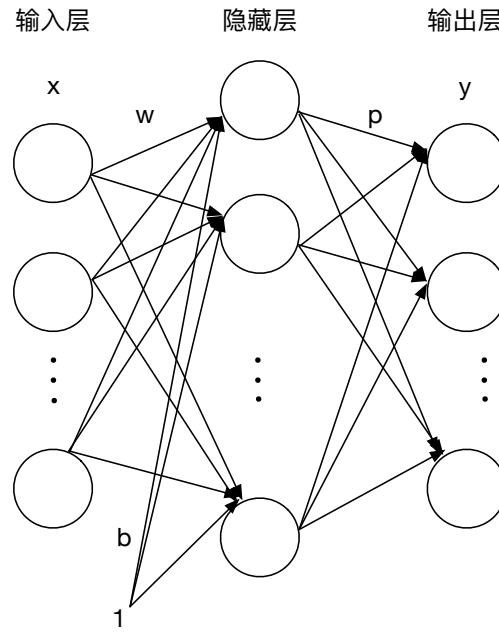


图 4-5 极限学习机网络

$b_i$  的值，不用对它们进一步调整了。并且在某种情况下  $\|T - Hp\|_F = 0$  一式是成立的，因而输出层的偏置也就可以忽略了。这样就可以得到一个新的单隐藏层反馈神经网络，叫做极限学习机网络，如图 4-5 所示。

与单隐藏层反馈神经网络相比，极限学习机具有以下几个优点：

1. 忽略了输出层偏置，并且通过随机的方式产生隐藏层的偏置和输入层的权重，需要计算的就只剩下输出层的权重了。比单隐藏反馈网络要求解的内容少了很多，因而学习时间大大降低。在实际实验中，采用极限学习机算法进行训练通常几秒就完成了，而用单隐藏层神经网络进行训练往往需要几十甚至几百倍的时间。
2. 泛化能力更强。
3. 避免了所求解的局部最优问题以及过拟合的问题。

## 4.2 分类算法选择

目前分类算法有很多种，而对于挑选出来的三类特征，我们将通过具体实验来选出适用每一类特征的分类算法。从论文<sup>[30]</sup>中，我们知道对 PkID 中的特征，采用随机森林分类算法的分类准确率最高，因而对于我们挑选出的 PkID 中的 22 个特征，随机森林算法也应该最适用。在图像处理和机器视觉领域，LBP 特征一般使用支持向量机分类算法进行分类，内距离形状上下文一般使用极限学习机进行分类。下面通过具体实验来进行验证，对这三类特征分别用 4.1 节介绍的三种分

表 4-1 PkID 的 22 个特征采用随机森林进行分类的结果

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2106	0	166	41	10	16	0	0	57	0	152	177	0	2725	0.773	0.242
Bubb	0	624	0	0	5	0	0	14	0	0	0	82	0	725	0.861	0.023
Chae	173	0	1491	0	2	11	0	0	7	1	41	19	0	1745	0.854	0.126
Clad	4	0	0	3856	22	1	13	0	0	42	37	650	0	4625	0.834	0.176
Cope	2	0	1	15	6076	208	0	0	3	0	193	851	21	7370	0.824	0.234
Deca	6	0	0	0	453	2027	0	3	0	4	86	203	3	2785	0.728	0.237
Doli	0	0	0	1	0	0	1061	0	0	215	0	148	0	1425	0.745	0.226
Egg	0	10	0	13	2	7	0	1322	0	23	12	328	3	1720	0.769	0.127
Fiber	75	0	11	4	2	2	0	0	1081	1	55	286	3	1520	0.711	0.23
Gela	5	0	0	148	10	6	194	34	0	1810	89	714	0	3010	0.601	0.297
Multi	253	1	27	125	463	209	14	27	62	129	914	949	2	3175	0.288	0.534
Nonb	153	4	10	475	770	144	89	115	193	348	367	12646	76	15390	0.822	0.267
Pter	0	0	0	0	119	25	0	0	0	3	16	200	722	1085	0.665	0.13
Total	2777	639	1706	4678	7934	2656	1371	1515	1403	2576	1962	17253	830	47300	0.73	0.219

类进行分类，从而选出其最适用的分类算法。

首先对于 PkID 中的 22 个特征用支持向量机、随机森林、极限学习机分类算法分别进行分类，对分类结果用  $K$  折交叉验证进行评价，结果是用随机森林算法分类效果最好，结果如表 4-1 所示，分类准确率达到 73%，错误率为 21.9%。

然后对训练集中图像提取 LBP 特征，并用支持向量机、随机森林、极限学习机分类算法分别进行训练，生成的分类器对测试集的分类结果中效果最好的是支持向量机，如表 4-2 所示，分类正确率为 64%，错误率为 32.8%。

最后，对内距离形状上下文特征进行实验，我们用到的数据集中一共有 13 个不同种类的浮游动物，其中同一类中的图像在形状方面也有可能相差较大，这是由于数据集中的类别有的是在“门”、“属”、“目”等不同级别，有的级别下面还包含很多子级，不同的子级之间有可能差别很多，所以在提取内距离形状上下文特征选取模板时，通常不能每类只选一幅图像作为模板，这样会导致该类下的某些子级被忽略了，不能被正确分类。经过实验，我们得出结论：当从每类浮游动物图像中选取 3 幅及以上的图像作为一类的模板时，分类效果会比较好。对于提

表 4-2 LBP 特征采用支持向量机进行分类的结果

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	1892	0	259	3	3	11	0	0	194	0	162	201	0	2725	0.694	0.399
Bubb	0	666	0	22	0	0	0	9	0	2	3	23	0	725	0.919	0.028
Chae	197	0	1462	0	0	3	0	0	37	0	41	5	0	1745	0.838	0.279
Clad	0	0	0	2962	200	9	21	2	0	53	36	1342	0	4625	0.64	0.348
Cope	27	0	5	208	5268	202	5	1	4	15	227	1354	54	7370	0.715	0.35
Deca	50	0	15	32	400	1568	0	3	2	14	262	413	26	2785	0.563	0.332
Doli	0	0	0	70	5	0	1008	0	0	100	10	232	0	1425	0.707	0.252
Egg	1	4	0	9	0	10	7	1467	2	18	1	201	0	1720	0.853	0.166
Fiber	257	0	156	1	4	1	0	5	819	0	55	222	0	1520	0.539	0.406
Gela	1	0	0	117	46	13	120	37	0	1826	84	766	0	3010	0.607	0.339
Multi	377	0	87	142	533	195	29	24	75	152	680	875	6	3175	0.214	0.662
Nonb	306	15	35	976	1446	204	156	210	238	578	426	10693	107	15390	0.695	0.357
Pter	41	0	9	3	199	130	0	1	8	4	23	297	370	1085	0.341	0.343
Total	3149	685	2028	4545	8104	2346	1346	1759	1379	2762	2010	16624	563	47300	0.64	0.328

取的内距离形状上下文特征，我们分别利用支持向量机、随机森林、极限学习机分类器进行训练并在测试集上验证，最终发现极限学习机分类算法最适用于这一特征。当对每一类选取的模板数目增多时，分类效果会有所提升，但同时学习速度也会变慢，两者成反比的关系。

对此，我们进行了一些实验作说明。当每类浮游动物图像中挑选 3 张作为模板，一共选取 39 幅时，采用 ELM 进行训练，在测试集上的分类结果评价如表 4-3，其分类准确率为 59.4%。选取 65 幅图像作为模板时，评价结果如表 4-4，其分类准确率为 65%。模板数目增加到 104 张时，评价结果如表 4-5，其分类准确率为 65.6%。可以看到，当模板数目增多时，分类准确率是有所提高的，但同时出现的问题是运行速度会下降很多，要在既保证分类准确率又保证运行效率的情况下设定最佳模板数目。而对内距离形状上下文特征使用支持向量机和随机森林分类算法时效果都没有极限学习机好。

上述三类实验验证了每类特征所对应有效的分类算法，至此，分类器设计完



表 4-3 提取内距离形状上下文特征并采用极限学习机进行分类的结果（39 张图像作为模板）

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2092	0	90	2	17	43	3	1	105	25	102	234	11	2725	0.768	0.315
Bubb	0	258	1	5	5	1	0	310	0	4	1	140	0	725	0.356	0.5
Chae	164	0	1413	2	5	5	2	0	104	19	14	17	0	1745	0.81	0.172
Clad	0	0	0	3525	97	21	9	4	1	19	15	932	2	4625	0.762	0.264
Cope	12	0	3	98	5742	124	6	0	8	49	145	1179	4	7370	0.779	0.266
Deca	23	0	1	120	303	1639	5	16	2	46	53	572	5	2785	0.589	0.27
Doli	1	0	0	36	19	3	645	0	1	83	30	607	0	1425	0.453	0.388
Egg	0	216	0	65	10	6	4	948	1	33	11	425	1	1720	0.551	0.378
Fiber	168	0	147	0	11	5	1	4	887	17	87	191	2	1520	0.584	0.378
Gela	13	15	1	81	136	10	57	110	7	1546	112	894	28	3010	0.514	0.338
Multi	289	1	10	157	377	114	82	14	118	146	506	1361	0	3175	0.159	0.65
Nonb	247	26	25	679	1065	253	239	114	189	337	356	11807	53	15390	0.767	0.365
Pter	44	0	16	21	38	22	1	2	2	10	13	227	689	1085	0.635	0.133
Total	3053	516	1707	4791	7825	2246	1054	1523	1425	2334	1445	18586	795	47300	0.594	0.34

成。

表 4-4 提取内距离形状上下文特征并采用极限学习机进行分类的结果（65 张图像作为模板）

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2268	0	40	0	4	34	0	0	73	6	79	218	3	2725	0.832	0.292
Bubb	0	202	0	7	6	0	0	364	5	1	1	139	0	725	0.279	0.423
Chae	145	0	1514	2	1	6	0	1	54	1	7	14	0	1745	0.868	0.15
Clad	0	0	0	3829	44	31	4	3	0	12	9	692	1	4625	0.828	0.213
Cope	7	0	0	56	5972	78	5	0	7	27	161	1053	4	7370	0.81	0.218
Deca	16	0	0	106	115	1865	1	3	2	22	47	606	2	2785	0.67	0.232
Doli	0	0	0	10	6	0	1011	4	0	48	23	323	0	1425	0.709	0.256
Egg	1	109	0	61	5	13	9	1136	11	24	11	340	0	1720	0.66	0.332
Fiber	140	0	158	0	9	3	1	5	902	3	101	198	0	1520	0.593	0.33
Gela	13	7	0	89	132	11	60	82	7	1652	101	847	9	3010	0.549	0.234
Multi	281	0	3	123	367	118	62	10	99	100	537	1473	2	3175	0.169	0.633
Nonb	280	32	56	581	953	261	206	93	186	246	379	12085	32	15390	0.785	0.336
Pter	51	0	10	2	20	8	0	0	2	14	6	212	760	1085	0.7	0.336
Total	3202	350	1781	4866	7634	2428	1359	1701	1348	2156	1462	18200	813	47300	0.65	0.286

表 4-5 提取内距离形状上下文特征并采用极限学习机进行分类的结果（104 张图像作为模板）

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2279	0	42	0	5	35	0	1	50	7	78	228	0	2725	0.836	0.275
Bubb	0	164	0	5	1	0	0	382	5	1	2	165	0	725	0.226	0.423
Chae	138	0	1530	0	3	1	0	0	52	0	8	13	0	1745	0.877	0.114
Clad	2	0	0	3844	49	33	0	3	0	12	8	674	0	4625	0.831	0.202
Cope	2	0	0	59	5987	66	6	1	5	31	156	1053	4	7370	0.812	0.209
Deca	11	0	0	90	109	1972	11	1	2	20	43	526	0	2785	0.708	0.215
Doli	0	0	0	13	7	2	1018	5	2	43	14	320	1	1425	0.714	0.25
Egg	1	82	1	66	6	11	11	1185	10	22	12	313	0	1720	0.689	0.33
Fiber	141	0	99	0	10	1	3	5	939	10	95	217	0	1520	0.618	0.32
Gela	7	5	0	78	138	18	59	81	3	1643	97	873	8	3010	0.546	0.234
Multi	263	0	8	125	344	110	55	9	97	106	516	1538	4	3175	0.163	0.629
Nonb	260	33	37	536	886	253	194	96	206	241	352	12259	37	15390	0.797	0.333
Pter	41	0	10	1	28	10	0	0	9	10	10	198	768	1085	0.708	0.066
Total	3145	284	1727	4817	7573	2512	1357	1769	1380	2146	1391	18377	822	47300	0.656	0.277

## 5 多特征多分类器组合

### 5.1 基本原理

本文选择了三类特征，对这三类特征分别用不同的分类算法进行分类，效果都不是很好，针对单个分类器不能完全表达浮游动物分类信息的情况，我们提出了一种基于多特征多分类器组合的浮游动物图像分类算法，框架如图 5-1 所示。首先，对浮游动物图像进行预处理以去除噪声。然后，对预处理后的图像提取三种不同类型的特征作为低层次特征。再将这些低层次特征输入到相应的机器学习算法，得到训练集中的每一幅图像在每一特征下被分到各个浮游生物图像类别中的概率，这个概率就作为中层次特征。最后，将三种类型特征对应的三种中层次特征串联起来并用支持向量机方法进行训练，得到最终的分类器。

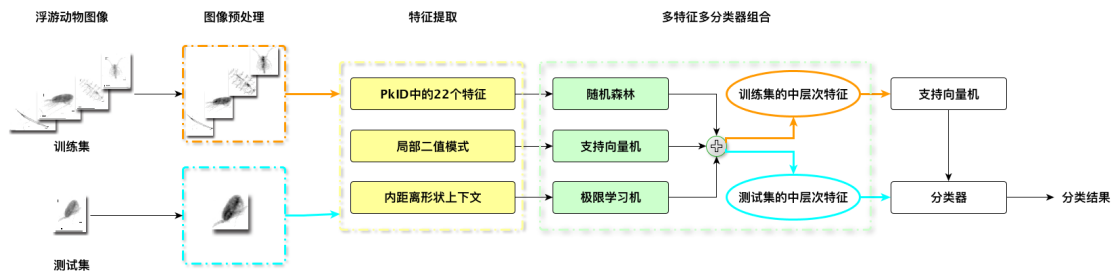


图 5-1 基于多特征多分类器组合的浮游动物图像分类算法框架

在实验中，我们往往希望提取尽可能多的特征，以增强对图像或物体类别的描述，使分类效果更理想。但实际上特征数目与分类有效性之间并不是线性关系，很多特征本身可能只包含了少量有用信息，亦或者是带有一些冗余信息，这些都会对分类效果产生不好的影响。在第 3.1 和 2.4.3 节中，我们整理和总结了各种不同的特征，通过实验与分析，本文最终选取出的三类特征分别是 PkID 中的特征、局部二值模式和内距离形状上下文。

选取出特征后，我们采用多个分类器组合的技术对提取的不同类型的特征作不同的处理，每一类型的特征都用相应的适合该类型特征的分类算法来进行训练。如果训练集中有  $M$  幅图像， $n$  个类别，我们提取了  $m$  种不同类型的特征，每一种类型的特征对应一种相应的分类算法，那么经过每一个分类算法除了能得到一个分类器外，还能得到训练集中的每一幅图像被分为每个类别的概率，对于整个数据集就会得到一个  $M \times n$  的矩阵，因为有  $m$  种不同类型的特征，所以最终会得到一个  $M \times (n * m)$  的矩阵。将这个矩阵作为对训练集图像提取的中层次特征，选取



图 5-2 预处理前和预处理后的图像对比

相应的分类算法进行训练，得到最终的分类器。

## 5.2 图像预处理

由于成像过程中受到的各种各样的干扰，我们所收集到的浮游图像不可避免地会带有一些噪声信号，这些噪声的引入，会直接导致图像的不清晰，造成识别准确率的下降。图 5-2(a) 显示的是一幅原始的浮游动物图像，可以看到除了浮游动物目标外，周围还混杂了一些悬浮物、污点等噪声信息，对图像进行预处理的目的是去除这些噪声信息，而保留和增强图像中对分类有用的信息。本文所采取的预处理方法是先对图像进行二值化，然后去除面积小于一定值的连通区域，图 5-2(b) 显示的是经过预处理后的图像，除浮游动物外的一些污点噪声已经被去除了。

## 5.3 实验对比与分析

在本文中，我们将所提出的基于多特征多分类器组合的浮游动物图像分类方法与 ZooScan 系统进行了对比，选用的数据集是 2.2 节中介绍的 13 类浮游动物图像集，评价方法采用的是  $K$  折交叉验证。

PkID 软件中，有 7 种机器学习算法供选用，将所提取出的 67 个特征都用上，并且机器学习算法定为随机森林时，分类准确率最高，约为 75%，见表 5-1。

将我们的算法中得到的三个分类器特征（中层次特征）串联起来，输入到支持向量机分类算法中，得到最终的分类器。这里我们仍然对其中的内距离形状上下文的参数选取进行了一些实验。当内距离形状上下文特征中模板数目设为 39 时，得到的分类器对测试集的分类结果评价如表 5-2 所示，其分类准确率为 77.1%。当内距离形状上下文特征中模板数目设为 65 时，得到的分类器对测试集的分类结果评价如表 5-3 所示，其分类准确率为 77.7%。而当内距离形状上下文特征中模板

表 5-1 PkID 软件中随机森林分类器分类结果

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2194	0	111	8	1	19	0	0	57	1	157	177	0	2725	0.805	0.204
Bubb	0	606	0	0	0	0	0	29	0	0	3	87	0	725	0.836	0.105
Chae	137	0	1559	0	5	2	0	0	7	1	18	14	2	1745	0.893	0.098
Clad	0	0	0	4084	22	0	7	0	0	7	19	486	0	4625	0.883	0.104
Cope	3	0	0	9	6205	145	0	0	0	1	234	745	28	7370	0.842	2.211
Deca	4	0	0	0	339	2225	0	0	0	0	87	126	4	2785	0.799	0.215
Doli	0	0	0	2	1	0	1065	0	0	156	7	194	0	1425	0.747	0.142
Egg	0	66	0	13	2	1	1	1277	0	28	3	320	9	1720	0.742	0.14
Fiber	59	0	18	0	0	0	0	0	1127	0	62	254	0	1520	0.741	0.207
Gela	6	0	0	72	10	5	93	39	0	2081	79	624	1	3010	0.691	0.204
Multi	267	0	23	51	491	202	9	23	45	91	1028	941	4	3175	0.324	0.509
Nonb	86	5	17	321	716	189	66	117	185	246	390	12937	115	15390	0.841	0.245
Pter	0	0	1	0	70	49	0	0	0	3	8	235	719	1085	0.663	0.185
Total	2756	677	1729	4560	7862	2837	1241	1485	1421	2615	2095	17140	882	47300	0.75	0.216

数目设为 104 时，得到的分类器对测试集的分类结果评价如表 5-4 所示，其分类准确率并没有变化，仍然为 77.7%。所以最终我们一般对内距离形状上下文特征中的模板数目设为 65，这样既能在一定程度上提高分类准确率，又不会使运行速度变得过慢。

对比表 5-1 和表 5-3 可以看到，本文中所提出的基于多特征多分类器组合的浮游动物图像分类算法与 ZooScan 系统中的识别软件相比，在识别准确率上有了明显的提高。并且我们的方法在使用的特征数量上有所降低，只采用了 24 个特征。这 24 个特征又可以分为三类，对这三类特征分别利用不同的分类算法进行训练得到的单个分类器的分类效果比组合这三个分类器的效果要差，这是因为组合分类器正确利用了单个分类器提供的信息，从不同的角度反映了图像的特性，最终的结果表明本文提出的多特征多分类器组合的方法是有效的。

表 5-2 多特征多分类器组合（内距离形状上下文特征采用 39 张图像作为模板）

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2250	0	116	14	7	5	0	0	34	0	154	145	0	2725	0.826	0.197
Bubb	0	651	0	1	5	0	0	2	0	0	1	65	0	725	0.898	0
Chae	151	0	1544	0	1	0	0	0	8	0	31	10	0	1745	0.885	0.105
Clad	0	0	0	4020	30	2	5	0	0	14	29	525	0	4625	0.869	0.136
Cope	1	0	0	16	6190	158	0	0	2	0	230	763	10	7370	0.84	0.217
Deca	5	0	0	1	369	2120	0	0	0	1	96	193	0	2785	0.761	0.191
Doli	0	0	0	7	0	0	1156	0	0	114	0	148	0	1425	0.811	0.123
Egg	0	0	0	6	5	6	0	1462	0	15	4	222	0	1720	0.85	0.075
Fiber	66	0	11	1	4	0	0	0	1134	0	65	239	0	1520	0.746	0.223
Gela	3	0	0	99	10	3	86	17	0	2040	76	676	0	3010	0.678	0.24
Multi	223	0	30	93	481	175	9	13	67	122	997	965	0	3175	0.314	0.529
Nonb	100	0	18	395	730	128	62	87	214	376	419	12803	58	15390	0.832	0.244
Pter	2	0	7	0	72	25	0	0	0	2	13	185	779	1085	0.718	0.182
Total	2801	651	1726	4653	7904	2622	1318	1581	1459	2684	2115	16939	847	47300	0.771	0.182

表 5-3 多特征多分类器组合（内距离形状上下文特征采用 65 张图像作为模板）

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2253	0	106	14	4	5	0	0	35	0	165	143	0	2725	0.827	0.192
Bubb	0	658	0	0	5	0	0	0	0	0	1	61	0	725	0.908	0.005
Chae	119	0	1582	0	0	3	0	0	4	0	21	16	0	1745	0.907	0.095
Clad	0	0	0	4063	29	3	3	0	0	16	17	494	0	4625	0.878	0.135
Cope	1	0	0	14	6216	145	0	0	0	1	201	784	8	7370	0.843	0.217
Deca	8	0	1	1	367	2122	0	0	0	2	85	198	1	2785	0.762	0.186
Doli	0	0	0	6	0	0	1181	0	0	108	1	129	0	1425	0.829	0.116
Egg	0	1	0	15	6	10	1	1438	0	19	7	223	0	1720	0.836	0.076
Fiber	51	0	12	1	3	1	0	0	1137	0	66	249	0	1520	0.748	0.215
Gela	4	0	2	94	12	2	75	12	0	2066	86	657	0	3010	0.686	0.238
Multi	230	0	23	99	473	166	14	18	74	115	994	969	0	3175	0.313	0.51
Nonb	122	2	18	391	762	125	62	88	197	380	370	12813	60	15390	0.833	0.243
Pter	2	0	4	0	63	25	0	0	1	4	13	181	792	1085	0.73	0.08
Total	2790	661	1748	4698	7940	2607	1336	1556	1448	2711	2027	16917	861	47300	0.777	0.177



表 5-4 多特征多分类器组合（内距离形状上下文特征采用 104 张图像作为模板）

	Appe	Bubb	Chae	Clad	Cope	Deca	Doli	Egg	Fiber	Gela	Multi	Nonb	Pter	Total	Recall	1-Pre
Appe	2250	0	116	13	7	6	0	0	46	0	148	139	0	2725	0.826	0.19
Bubb	0	660	0	0	5	0	0	0	0	0	1	59	0	725	0.91	0.002
Chae	120	0	1574	0	1	1	0	0	6	0	30	13	0	1745	0.902	0.105
Clad	0	0	0	4076	25	1	5	0	0	15	18	485	0	4625	0.881	0.142
Cope	0	0	0	15	6206	139	0	0	0	1	222	775	12	7370	0.842	0.215
Deca	7	0	0	0	381	2127	0	0	0	1	79	187	3	2785	0.764	0.175
Doli	0	0	0	8	0	0	1154	0	0	108	0	155	0	1425	0.81	0.11
Egg	0	0	0	11	5	11	0	1444	0	24	7	217	1	1720	0.84	0.079
Fiber	47	0	15	1	4	0	0	0	1150	0	67	236	0	1520	0.757	0.219
Gela	3	0	0	106	17	2	68	18	0	2075	80	641	0	3010	0.689	0.237
Multi	229	0	25	106	470	147	12	21	74	122	999	969	1	3175	0.315	0.522
Nonb	120	1	22	415	725	120	58	85	196	372	424	12785	67	15390	0.831	0.241
Pter	1	0	7	0	57	24	0	0	0	2	14	187	793	1085	0.731	0.096
Total	2777	661	1759	4751	7903	2578	1297	1568	1472	2720	2089	16848	877	47300	0.777	0.179

## 6 总结与展望

### 6.1 总结

浮游动物图像自动分类识别是浮游动物检测系统中的关键技术，随着浮游动物图像获取技术的不断提高，越来越多的图像数据被采集获取到，仅仅依赖人为的种类鉴别手段已经难以满足对大量数据进行实时快速处理的需要。基于机器学习方法的浮游动物图像分类方法，与其它方法相比，识别准确率更高，因此越来越受到重视。

本文围绕机器学习理论和多特征多分类器组合技术，进行了相关的研究和探索，主要完成了以下工作：

1. 以浮游生物成像系统和自动识别技术为基础，研究了国内外的发展现状，并且介绍了为什么要进行浮游动物图像分类识别的研究。
2. 对所要分类的浮游动物图像数据集进行了分析，对每类浮游动物的特点做了一个大致的整理，从而能够人为地对这些图像进行分类。
3. 对图像处理领域常用的特征以及在各种分类竞赛中所采用的一些经典特征做了总结，通过理论和实验的结合，从大量的特征中选择适用于浮游动物图像的特征。
4. 针对不同类型的特征，选择适用于该特征的分类算法进行分类器的设计。
5. 提出了一种基于多特征多分类器组合的浮游动物图像分类方法，从而能够综合不同分类器的信息，提高分类准确率。

### 6.2 展望

尽管本文的算法在浮游动物的识别准确率上有所提高，并且对特征和分类器都进行了总结，做了大量的实验来验证哪些特征和分类器适用于解决浮游动物图像分类这类问题。但是还存在着一些不足，需要做进一步的研究：

1. 我们的算法是基于 Matlab 语言的，但 Matlab 在计算速度上仍然不够快，将来希望用 C/C++ 语言进行编程，进一步提高算法的运行速度。
2. 本文采用的是对分类结果进行融合的方法，还有很多特征融合的方法没有进行尝试，下一步会针对特征融合进行研究，看看对识别效果有没有改善。
3. 每个类别图像数目相差较多时，识别准确率较低，当每个类别数目大致相同时，识别准确率会有所提高，后续会对这一现象继续研究，从理论和实验两

方面来解释说明。

4. 由于我们的实验都是基于同一数据集的，可能会带有数据集偏见，今后的研究希望能不依赖于数据集，要找到在各种浮游动物图像数据集上都能取得较好效果的特征和分类器。

## 参考文献

- [1] Davis C S, Thwaites F T, Gallager S M, et al. A three-axis fast-tow digital video plankton recorder for rapid surveys of plankton taxa and hydrography. *Limnology and Oceanography: Methods*, 2005, 3:59–74.
- [2] Remsen A, Hopkins T L, Samson S. What you see is not what you catch: a comparison of concurrently collected net, optical plankton counter, and shadowed image particle profiling evaluation recorder data from the northeast gulf of mexico. *Deep Sea Research Part I: Oceanographic Research Papers*, 2004, 51(1):129–151.
- [3] Davis C, Gallager S, Berman M, et al. The video plankton recorder (vpr): design and initial results. *Arch. Hydrobiol. Beih*, 1992, 36:67–81.
- [4] Benfield M C, Grosjean P, Culverhouse P F, et al. Rapid: Research on automated plankton identification. *Oceanography*, 2007, 20(2):172–187.
- [5] Watson J. Holomar: A holographic camera for subsea imaging of plankton. *Sea Technology*, 2004, 45(12):53–55.
- [6] Katz J, Donaghay P, Zhang J, et al. Submersible holocamera for detection of particle characteristics and motions in the ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 1999, 46(8):1455–1481.
- [7] Pfitsch D, Malkiel E, Ronzhes Y, et al. Development of a free-drifting submersible digital holographic imaging system. *Proceedings of OCEANS 2005 MTS/IEEE*, volume 1, 2005. 690–696.
- [8] 戴君伟, 王博亮, 谢杰镇, et al. 海洋赤潮生物图像实时采集系统. *高技术通讯*, 2006, 16(12):1316–1320.
- [9] 于翔, 宋家驹, 于连生. 水下全自动显微成像仪. *海洋技术*, 2009, 28(4):14–16.
- [10] Tan-Feng H, Yao-Wu C, et al. 实时浮游生物图像目标智能识别系统设计. *计算机工程*, 2012, 38(15):183–186.
- [11] Zhou Z, Yu X, Sui M, et al. Design and evaluation of optical image sensing device for marine plankton monitoring. *Proceedings of International Conference of Optical Instrument and Technology*, 2008. 71570–71570.
- [12] Jeffries H, Berman M, Poularikas A, et al. Automated sizing, counting and identification of zooplankton by pattern recognition. *Marine Biology*, 1984, 78(3):329–334.
- [13] Hicks Y A, Marshall D, Rosin P L, et al. A model of diatom shape and texture for analysis, synthesis and identification. *Machine Vision and Applications*, 2006, 17(5):297–307.
- [14] Dimitrovski I, Kocev D, Loskovska S, et al. Hierarchical classification of diatom images using ensembles of predictive clustering trees. *Ecological Informatics*, 2012, 7(1):19–29.
- [15] Wilkinson M, Urbach E, Jalba A, et al. Diatom identification with mathematical morphology. *Mathematical Morphology: From Theory to Applications*, 2010. 355–365.

- [16] Embleton K, Gibson C, Heaney S. Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method. *Journal of Plankton Research*, 2003, 25(6):669–681.
- [17] Culverhouse P F, Williams R, Reguera B, et al. Do experts make mistakes? a comparison of human and machine identification of dinoflagellates. *Marine Ecology Progress Series*, 2003, 247(5):17–25.
- [18] Blaschko M B, Holness G, Mattar M A, et al. Automatic in situ identification of plankton. *Proceedings of IEEE Workshops on Application of Computer Vision*, volume 1, 2005. 79–86.
- [19] Hu Q, Davis C S. Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction. *Marine Ecology Progress Series*, 2006, 306:51–61.
- [20] Sosik H M, Olson R J. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnology and Oceanography: Methods*, 2007, 5:204–216.
- [21] Hense B A, Gais P, Jütting U, et al. Use of fluorescence information for automated phytoplankton investigation by image analysis. *Journal of Plankton Research*, 2008, 30(5):587–606.
- [22] Verikas A, Gelzinis A, Bacauskiene M, et al. Phase congruency-based detection of circular objects applied to analysis of phytoplankton images. *Pattern Recognition*, 2012, 45(4):1659–1670.
- [23] Culverhouse P, Williams R, Reguera B, et al. Expert and machine discrimination of marine flora: A comparison of recognition accuracy of field-collected phytoplankton. *Proceedings of International Conference on Visual Information Engineering*, 2003. 177–181.
- [24] Jeffrey E, Hongyu L, Mark D O. Quantifying california current plankton samples with efficient machine learning techniques. *Proceedings of OCEANS 2015 MTS/IEEE Washington*, 2015. 1–9.
- [25] 王明时, 王学民, 张磊. 显微图像分析技术在赤潮生物识别中的应用. *天津大学学报: 自然科学与工程技术版*, 2004, 37(5):392–395.
- [26] 汪振兴, 余焱, 姜建国. 赤潮藻类图像自动识别的研究. *海洋环境科学*, 2007, 26(1):42–44.
- [27] 骆巧琦, 李雪松, 梁君荣, et al. 基于形状特征的硅藻显微图像自动识别. *厦门大学学报: 自然科学版*, 2011, 50(4):690–696.
- [28] 魏雅娟. 暗视场浮游动物图像自动识别方法研究 [硕士学位论文]. 青岛: 中国海洋大学, 2013.
- [29] Bograd S J, Checkley D A, Wooster W S. Calcofi: A half century of physical, chemical, and biological research in the california current system. *Deep Sea Research Part II: Topical Studies in Oceanography*, 2003, 50(14):2349–2353.
- [30] Gorsky G, Ohman M D, Picheral M, et al. Digital zooplankton image analysis using the zooscan integrated system. *Journal of Plankton Research*, 2010, 32(3):285–303.
- [31] Lowe D G. Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, 1999. 1150–1157.
- [32] Lowe D G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 2004, 60(2):91–110.

- [33] Dalal N, Triggs B. Histograms of oriented gradients for human detection. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, 2005. 886–893.
- [34] Ojala T, Pietikainen M, Harwood D. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. Proceedings of the 12th IAPR International Conference on Pattern Recognition, 1994. 582–585.
- [35] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4):509–522.
- [36] Reisfeld D, Wolfson H, Yeshurun Y. Context-free attentional operators: the generalized symmetry transform. International Journal of Computer Vision, 1995, 14(2):119–130.
- [37] Ling H, Jacobs D W. Shape classification using the inner-distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(2):286–299.
- [38] Cortes C, Vapnik V. Support-vector networks. Machine learning, 1995, 20(3):273–297.
- [39] Breiman L. Random forests. Machine learning, 2001, 45(1):5–32.
- [40] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feed-forward neural networks. Proceedings of IEEE International Joint Conference on Neural Networks, volume 2, 2004. 985–990.

## 致 谢

时光匆匆如流水，转眼便又是毕业时节了，在我三年的硕士研究生生涯里，我所收获的不仅仅是愈加丰厚的知识，更重要的是思维方式和表达能力的转变与提升。我很庆幸在这三年我遇到了很多的良师益友，无论在学习上、生活上还是工作上，都给予了我无私的帮助和悉心的照顾，让我在一个温馨、有爱的环境中度过了我的三年研究生生活。

首先，我要感谢我的导师郑海永，这三年来给我最多教育和帮助的人！从我开始读研开始，郑老师一直对我严格要求，在他的指导下，我不仅树立了远大的学习目标，掌握了基本的研究方法，还明白了很多为人处世的道理。本论文从选题到最终完成，每一步都得到了郑老师的帮助和指点，在我迷茫时给予适时的鼓励，帮助我开拓研究思路，解决问题。感谢师母冯丽颖的关怀和教育，使我认识到自身的不足，从而努力提升自己，做一个坚强、会生活、懂生活的新时代女性。在此，谨向导师和师母表示崇高的敬意和衷心的感谢！

感谢实验室所有兄弟姐妹，没有你们的陪伴，我不会拥有这么美好、难忘的三年回忆。感谢我的师姐赵红苗、师妹王如晨，和我并肩作战，在我的研究课题上给予了很多的支持和帮助，让我有了更大提高。

感谢我的家人，谢谢你们无私的付出，在我身后坚定地支持我、关心我、帮助我，有你们在，我更安心！

感谢国家自然科学基金青年科学基金项目“基于视觉注意结合生物形态特征的海洋浮游植物显微图像分析”（批准号：61301240）、国家自然科学基金项目“基于生物形态特征的中国海常见有害赤潮藻显微图像识别”（批准号：61271406）和中央高校基本科研业务费“海洋浮游动物原位探测与分析系统”（批准号：201562023）的资助。

最后，感谢所有关心和帮助过我的人，祝愿你们幸福安康！

## 个人简历、在学期间发表的学术论文与研究成果

### 个人简历

1990 年 9 月 4 日出生于江苏省泰州市姜堰区。

2009 年 9 月考入中国海洋大学电子工程系电子信息科学与技术专业，2013 年 6 月本科毕业并获得理学学士学位。

2013 年 9 月考入中国海洋大学电子系攻读硕士学位至今。

### 发表的学术论文

- [1] Wang X, Dai J, Zhu Y, Zheng H and Qiao X. Spectral saliency via automatic adaptive amplitude spectrum analysis. Journal of Electronic Imaging, 2016. (SCI 收录)
- [2] Zhu Y, Chang L, Dai J, Zheng H, Zheng B. Automatic object detection and segmentation from underwater images via saliency-based region merging. Proceedings of OCEANS 2016 MTS/IEEE Shanghai, 2016. (EI 收录)

### 在学期间参加的研究项目

- 1. 国家自然科学基金青年科学基金项目“基于视觉注意结合生物形态特征的海洋浮游植物显微图像分析”(批准号: 61301240)
- 2. 国家自然科学基金项目“基于生物形态特征的中国海常见有害赤潮藻显微图像识别”(批准号: 61271406)
- 3. 中央高校基本科研业务费“海洋浮游动物原位探测与分析系统”(批准号: 201562023)