



汕 頭 大 學

SHANTOU UNIVERSITY

# 硕 士 学 位 论 文

题 目 基于深度学习的中华白海豚照片个体识别

英文题目 Photo Identification of Chinese White Dolphin  
Based on Deep Learning

姓 名 冯靖安 学 号 111709028

所在学院 工学院 导师姓名 范衡

专 业 信息与通信工程

入学日期 2017.09 答辩日期 2020.06







## 摘要

长期的照片个体识别可以让我们对动物种群的人口统计数据、生活史、运动方式和社会结构有更深入细微的了解。中华白海豚在其成长过程中有着独特多样的颜色和斑点，是照片个体识别的理想物种。并且与其他鲸豚类动物相比，中华白海豚在地理上的可及性、生物学特性、生态功能以及文化地位，使得它具有更高的研究和社会关注度。为了提高对中华白海豚这一濒危海洋动物的保育研究的效益，一方面，我们需要提高中华白海豚照片个体识别的效率，让海洋生态学家有更多时间进行户外研究，另一方面，我们需要实现数据处理的标准化，促进不同团队间的合作。

针对自动化的中华白海豚照片个体识别这一实际问题，本文首次将基于深度学习的图像处理技术应用于中华白海豚的照片个体识别。首先，我们提出了一个用于中华白海豚的照片个体识别的处理流程，包括四个步骤：背鳍定位、背鳍照片质量识别、背鳍个体识别和背鳍分割。基于该流程，我们整理并公开了首个用于中华白海豚的照片个体识别的大规模图像数据集，该数据集共包括 53513 张照片，来自广西钦州（140 个个体）和广东汕头（20 个个体）。在该数据库上，针对每个步骤，我们选择了多个主流的深度卷积神经网络进行测试，并给出了不同网络的结果。由于过去的照片个体识别都是对单个个体的单张图片进行处理，对于低质量照片的识别效果欠佳，我们提出了一种基于群信息的背鳍个体识别方法，通过对一群图片进行处理，并通过异常检测和异常捞回的方式，显著提高了个体识别对于低质量照片的识别准确率（+7.35%），从而提高了整体识别准确率（+2.63%）。最后，我们设计了一个在 Windows 系统上运行的基于深度学习的中华白海豚识别软件，包括便捷的操作流程和直观的人工校验界面，从而辅助生态学家完成中华白海豚的照片个体识别。

**关键词：**中华白海豚；照片个体识别；深度学习；群信息；异常检测

## Abstract

The long term photo-ID studies could enhance our nuanced understanding of demographics, life history, movement pattern and social structure of animal populations. Among all the marine animals, Chinese white dolphins, one of endangered species, are known well for their special color pattern in the life history and diverse "fingerprint" pigmentation. Compared with other cetaceans, this animal has a higher research and social concern because of its geographic accessibility, biological feature, ecological function and cultural status in coastal waters. In order to improve research and conservation benefit of Chinese white dolphin, on the one hand, we need to improve the efficiency of photo identification of Chinese white dolphin to leave marine biologists more time for field research, on the other hand, we need to implement the standardization of data processing to promote cooperation between the different teams.

In order to solve the problem of automated photo identification of Chinese white dolphin, this paper applies deep learning technology to photo identification of Chinese white dolphin for the first time. First of all, we propose a processing pipeline for photo identification of Chinese white dolphin, including four steps: dorsal fin detection, scoring, matching and segmentation. Based on this pipeline, we collect and publish the first large-scale image dataset for automated photo identification of Chinese white dolphin, which contains 53,513 images captured from Qinzhou (with 140 dolphin individuals) and Shantou (with 20 dolphin individuals). To build a baseline for future algorithm development, state-of-the-art computer vision algorithms are evaluated on this dataset. Since previous photo identification methods work by inputting a single image, which do not work well on low quality photos, we present a photo identification method based on group information. By inputting a group of images, and through the way of anomaly detection and anomaly retrieval, we significantly improve the photo identification accuracy for low quality photos (+ 7.35%), and thus improve the overall accuracy (+ 2.63%). At last, we design a deep learning based photo identification software running on the Windows system, including convenient operation flow and intuitive manual verification interface, so as to assist marine biologists to complete photo identification of Chinese white dolphin.

**Keywords:** *Sousa chinensis*; Photo identification; Deep learning; Group information; Anomaly detection

## 目 录

摘要.....	I
Abstract .....	II
目 录.....	III
第 1 章 绪论 .....	1
1.1 研究背景与意义.....	1
1.2 国内外研究现状.....	2
1.3 主要研究内容.....	3
1.4 章节安排 .....	4
第 2 章 数据集的建立.....	5
2.1 整体流程 .....	5
2.2 数据集描述.....	6
2.2.1 照片的收集与标记 .....	6
2.2.2 背鳍定位.....	7
2.2.3 背鳍照片质量识别 .....	9
2.2.4 背鳍个体识别.....	10
2.2.5 背鳍分割.....	12
2.3 数据集划分.....	12
第 3 章 基于深度学习的照片个体识别.....	13
3.1 背鳍定位 .....	13
3.1.1 目标检测.....	13
3.1.2 实验结果.....	17
3.2 背鳍照片质量识别.....	17
3.2.1 图像分类.....	17
3.2.2 实验结果.....	21
3.3 背鳍个体识别.....	21
3.4 背鳍分割 .....	22
3.4.1 图像语义分割.....	22
3.4.2 实验结果.....	23
3.5 交叉地区验证.....	24
第 4 章 基于群信息的背鳍个体识别.....	26

4.1 介绍.....	26
4.2 相关工作.....	28
4.3 方法.....	30
4.3.1 CNN 预测并提取特征.....	31
4.3.2 异常检测 .....	31
4.3.3 异常捞回 .....	33
4.4 实验与结果 .....	33
4.4.1 数据集.....	33
4.4.2 评价指标 .....	34
4.4.3 结果.....	35
第 5 章 中华白海豚识别软件 .....	37
5.1 软件总体流程 .....	37
5.2 神经网络的构建 .....	38
5.3 执行处理的界面 .....	38
5.4 用户校验的界面 .....	41
第 6 章 总结与展望 .....	43
6.1 本文总结.....	43
6.2 未来工作展望 .....	43
参考文献 .....	44
致 谢 .....	51
攻读硕士学位期间主要的工作成果 .....	52

# 第 1 章 绪论

## 1.1 研究背景与意义

基于动物自然标记的照片个体识别,作为一种非侵入性的生态普查技术,在过去的半个世纪中,已经被人们广泛地应用<sup>[1]</sup>。结合其他数据收集技术,如 GPS、水温等,长期的照片个体识别还可以让我们对动物种群的人口统计数据、生活史、运动方式和社会结构有更深入细微的了解,尤其是对于无法接近的自由生活的动物如鲸豚类动物<sup>[2]</sup>。然而,照片个体识别受到动物自然标记(如独特性、多样性、可变性)、数据收集(如图像质量)、识别算法等因素的影响,可能会导致识别错误和劳力成本的大大增加<sup>[3,4]</sup>。此外,鲸豚类动物的照片个体识别具有资源消耗高的性质,再加上目前缺乏图像数据的标准化限制了研究的合作与协调,这些因素反过来又降低了对那些濒临灭绝的海洋物种的保护效率<sup>[5]</sup>。为了使生态图像处理标准化,并提高其在实地生态研究中的成本效益,我们应该通过多学科协作,优先克服个体识别算法的挑战,从而改善当前的照片个体识别系统。

在所有海洋动物中,中华白海豚以其成长过程中独特多样的颜色和斑点而著称<sup>[6]</sup>,这也让他们成为照片个体识别的理想物种。与其他鲸豚类动物相比,由于中华白海豚在地理上的可及性、生物学特性、生态功能以及文化地位,它具有更高的研究和社会关注度。作为海洋沿海生态系统中最具代表性的标志性物种之一,中华白海豚已经被研究了十多年。尽管其连续分布在不同零散的栖息地,在中国已经有超过 8 个确认的种群被不同的研究团队用不同的研究手段进行研究。然而最近,最大的一个位于珠江口的种群,其种群数量呈下降趋势,可见我们对中华白海豚的保护效率仍然有限,这也暗示着中国其他一些较小的种群的灭绝风险会更高。一方面,人们认为,如此高的灭绝风险可能是由于沿海栖息地的内在脆弱性和外部环境压力的相互作用所致,这与其他海洋哺乳动物类似<sup>[7,8]</sup>。另一方面,由于数据收集的不连续和各研究团队难以合作协调,关于中华白海豚群体连通性甚至群体状况的信息差距仍然存在。随着在中国越来越多的人关注如何转变中华白海豚的灭绝风险并提高对其的保护效率,一项国家行动计划(2017-2026 年)被启动,旨在鼓励学科内和学科间打破业务障碍,合作保护管理。因此,随着不同的学术组织开始在中国建立起一个大型的中华白海豚的照片个体数据库,为了最大程度地提高保护效果,一个亟待解决的问题是数据处理的标准化。我们相信通过共享开源的方式,包括使用大规模合作公开的数据集、各类运算平台、以及不断更新开源的机器学习算法,将提高个体识别的准确率,促进保护的作,并提高对海洋哺乳动物的保护效率。

## 1.2 国内外研究现状

传统的照片个体识别的数据处理过程包括以下步骤：图像收集、特征提取、生物特征编码和数据库匹配<sup>[9]</sup>。

在图像采集中，与陆生动物相比，海洋动物由于其生产环境而更加难以观察和采样，这给大规模数据集的构建带来了困难。此外，在数据库同步上，我们还可以发现一些其他的缺点，可以归类成两类，分别是数据收集方法的不一致，和数据样本在时间空间上的不连续。尽管海洋哺乳动物的研究已经持续了半个多世纪，但只有很少的合作公开的数据库，如 OBIS-SEAMAP<sup>[10]</sup>。OBIS-SEAMAP 是一个是以地理坐标为主的在线数据库，汇总了来自全球的海洋哺乳动物、海鸟、海龟、鳐鱼和鲨鱼的观测数据，其中包括少数几个用于照片个体识别的数据集，但图片质量和图片数量并不利于算法开发。据我们了解，现有的关于海洋哺乳动物的用于照片个体识别算法竞赛的公开数据集只有 3 个，2 个座头鲸的数据集<sup>[11,12]</sup>和 1 个露脊鲸的数据集<sup>[13]</sup>，这两个物种都属于濒危鲸鱼物种。其中，“Lifeclef 2016”数据集<sup>[11]</sup>包括 2005 张座头鲸尾鳍照片，由 CetaMada 非政府组织从 2009 到 2014 年拍摄于马达加斯加地区。每张照片都由人工截取出尾鳍部分，主要用于座头鲸的个体识别。在文献[12]中，该数据集包括 25460 张座头鲸照片，由一些研究机构和公众提供。在文献[13]中，该数据集由 11469 张露脊鲸照片组成，其中分为训练集（4544 张）和测试集（6925 张）。

在图像处理中，无论是目标检测还是生物特征编码，与其他半自动或全自动方法相比，手工方法目前仍然是更为常用的。对于手工方法，不同的研究人员可能使用不同的数据管理软件来进行图像管理，包括一些商业软件，如 OFFICE 或 ENDNOTE，或者是自行开发的软件，如 DISCOVERY<sup>[14,15]</sup>。此外，基于不同的鳍曲线算法，人们针对海洋哺乳动物还开发了几个半自动识别系统，如 DARWIN<sup>[16]</sup>，Finscan<sup>[17]</sup>，和 I3S<sup>[18]</sup>。在文献[16]中，Stanley 等人开发了 DARWIN 通过背鳍照片进行海豚个体识别，该系统基于活动轮廓和二次样条小波。DARWIN 也被应用到其他物种的个体识别上，如大白鲨<sup>[19,20]</sup>和宽吻海豚<sup>[21]</sup>。在文献[17]中，Hillman 等人开发了 Finscan，基于智能剪刀算法（Live Wire）<sup>[22]</sup>和两种匹配方法（Curve Matching & String Matching）。Finscan 则被应用到伪虎鲸<sup>[23]</sup>和宽吻海豚<sup>[24]</sup>的个体识别中。在文献[25]中，Stewman 等人则提出一种用于图像检索的迭代式的开放轮廓对齐的方法，并用于宽吻海豚的个体识别。最近几年，随着数字摄影技术、图像处理算法、运算处理能力的发展，对大规模数据集搭建自动化识别系统成为可能<sup>[26-28]</sup>。在文献[29]中，Hughes 等人提出了一种开放轮廓笔划模型，实现了从背鳍图像中对大白鲨个体的自动化识别。在文献[30]中，Weideman 等人定义了一种积分曲率表示方法，并使用有效的匹配算法通过鳍图像进行鲸鱼或海豚的个体识别。

上述算法都是基于鳍曲线进行个体识别,主要适用于海豚的背鳍<sup>[24]</sup>、大白鲨的背鳍<sup>[29]</sup>和座头鲸的尾鳍<sup>[30]</sup>的识别,这些动物的照片大多数没有斑点或纹理。而对于有丰富斑点或纹理的动物的照片,如露脊鲸<sup>[13]</sup>,可以提取斑点或纹理的特征,再利用这些特征进行个体识别。传统图像处理的特征提取方式是提取出个体中的关键点,再对关键点作比对。在基于关键点比对的照片个体识别中,研究的物种有鲸鲨<sup>[31]</sup>、锥齿鲨<sup>[32]</sup>、蝠鲼<sup>[33]</sup>等。在文献[31]中, Arzoumanian 等人通过 blob extraction、旋转校正和对比度增强提取出鲸鲨身上的斑点,再通过斑点坐标构造多个三角形并匹配不同图片的三角形实现个体识别。在文献[32]中, Van 等人提出了一种基于二维仿射变换来比较两个个体的方法,但该方法需要人工选取用于匹配的关键点。在文献[33]中, Christopher 等人先通过图像增强对个体图片进行处理,再通过 SIFT、SURF 或 ORB 等特征描述子对不同图片进行特征提取与匹配。随着 2012 年 Alexnet<sup>[34]</sup> 的提出,卷积神经网络作为一种特征提取的方式,相比于传统的特征提取有明显的性能优势,其应用也愈发广泛,其中就包括动物的照片个体识别。在基于卷积神经网络的照片个体识别中,有关于海洋动物的,如露脊鲸<sup>[26,35-37]</sup>,海豹<sup>[38]</sup>,也有关于陆地动物的,如大猩猩<sup>[39]</sup>,东北虎<sup>[40]</sup>,大象<sup>[41]</sup>。在文献[26]中, Bogucki 等人通过 CNN 对照片中的露脊鲸头部进行定位、识别关键点完成旋转校正、以及个体识别。在文献[35,36]中, Kabani 等人主要设计了一个新的网络 Wide Net 实现对露脊鲸头部的定位,并提出了一个对露脊鲸头部的定位、旋转校正以及个体识别的流程。在文献[37]中, Polzounov 等人通过 HSV 通道图像进行阈值分割实现头部定位,并提出一个网络 DeepSenseNet 实现个体识别。

### 1.3 主要研究内容

本论文主要是为了解决自动化的中华白海豚照片个体识别这一实际问题,首次将深度学习的图像处理技术应用于中华白海豚的照片个体识别。首先,我们提出了一个用于中华白海豚的照片个体识别的处理流程,并基于该流程整理、公开了首个用于中华白海豚的照片个体识别的大规模图像数据集。然后,在该数据库上测试了主流的深度卷积神经网络,并对比了不同网络的结果。进一步地,提出了一种“群识别”方法,针对现有方法的不足,有效利用了拍摄时的先验信息,提高了个体识别的整体准确率。最后,设计了一个在 Windows 系统上的应用软件,辅助生态学家完成中华白海豚的照片个体识别。主要研究内容如下:

- 1) 根据中华白海豚生态保育过程中的实际需求,提出了一个用于中华白海豚的照片个体识别的处理流程,主要包括:背鳍定位、背鳍照片质量识别、背鳍个体识别和背鳍分割。并针对该流程整理、公开了首个用于中华白海豚的照片个体识别的大规模图像数据集,促进该领域照片个体识别算法的研究。

2) 研究主流的深度卷积神经网络在该数据库下的应用, 包括利用 Faster-RCNN、RFCN、SSD 和 YoloV3 进行背鳍定位, 利用 ResNet-101、InceptionV4、InceptionResnetV2 和 MobilenetV1 进行背鳍照片质量识别和照片个体识别, 利用 FCN 和 DeeplabV3 进行背鳍分割, 并对比不同网络的准确率。

3) 针对现有的照片个体识别方法的不足, 提出了一种基于群信息的光照个体识别后处理方法, 主要包括异常检测和异常捞回两个步骤。并对比了该方法与普通个体识别方法的准确率, 该方法有效利用了拍摄时的先验信息, 提高了个体识别的整体准确率。

4) 设计了一个在 Windows 系统上的应用软件, 将在 GPU 服务器上训练的深度卷积神经网络迁移到 Windows 桌面系统上进行应用。并针对生态学家的实际需求和习惯, 设计友好的用户界面, 辅助生态学家完成中华白海豚的照片个体识别。

## 1.4 章节安排

本文的组织结构安排如下:

第一章: 绪论。简述中华白海豚的照片个体识别的背景与意义, 简述中华白海豚以及类似的海洋鲸豚类动物的照片个体识别方法的研究现状, 并简要概述文章的研究内容。

第二章: 数据集的建立。介绍目前人工处理中华白海豚照片的方式, 以及中华白海豚的照片个体识别的自动化处理流程, 并描述基于该流程整理的首个用于中华白海豚的照片个体识别的大规模图像数据集。

第三章: 基于深度学习的照片个体识别。介绍适用于上述流程中各个步骤的主流深度学习算法, 并在该数据库上测试了不同算法, 给出并对比了不同算法的结果。

第四章: 基于群信息的背鳍个体识别。分析目前个体识别方法的不足, 介绍“群”的概念并将其作为先验信息, 提出了一种“群识别”方法, 并对比了该方法与目前个体识别方法的准确率。

第五章: 中华白海豚识别软件。介绍设计运行于 Windows 系统上的应用软件, 包括深度卷积神经网络的迁移实现, 以及用户界面的功能实现。

第六章: 总结与展望。该部分主要是对本文研究工作的创新与不足的总结, 同时给出未来工作的方向。

## 第 2 章 数据集的建立

### 2.1 整体流程

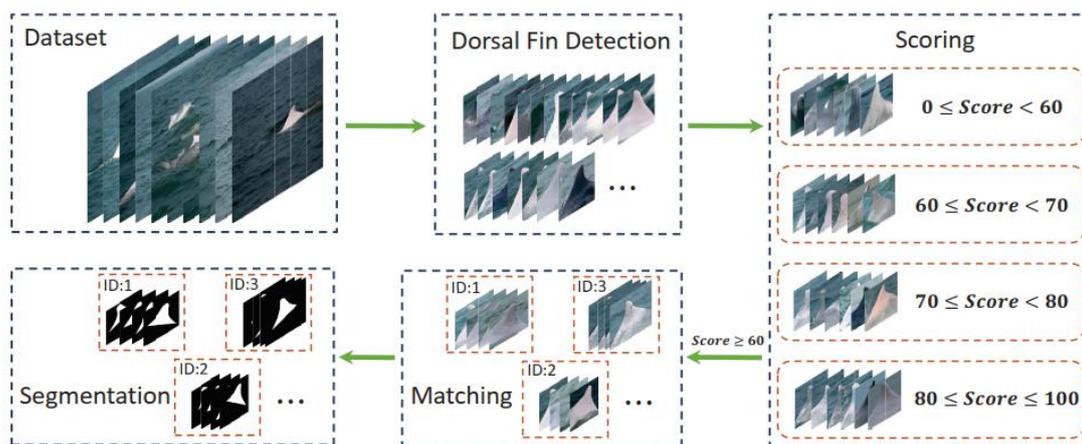


图 2-1 中华白海豚的照片个体识别的整体流程

中华白海豚的照片个体识别的整体流程如图 2-1 所示，主要包括四步：背鳍定位（Dorsal Fin Detection）；背鳍照片质量识别（Scoring）；背鳍个体识别（Matching）；背鳍分割（Segmentation）。其中前三步是手工进行中华白海豚照片个体识别的传统流程。这里先介绍这个流程，对于海洋生物学家出海拍摄的中华白海豚照片，首先要进行背鳍的定位与截取；然后截取出来的背鳍照片根据照片质量被划分成四个质量等级，其中 0-60 分的背鳍照片属于人眼难以辨别个体的照片；接着对高于 60 分的背鳍照片进行背鳍个体识别；最后再对这部分高于 60 分的背鳍照片进行背鳍分割。下面将解释整体流程中每个步骤的意义：

1) 背鳍定位（Dorsal Fin Detection）：大部分情况下，中华白海豚只有背鳍区域露出水面并被拍摄到，而且背鳍区域包含丰富的特征，如背鳍曲线和斑点，研究人员主要根据背鳍上的特征信息辨别不同的中华白海豚个体。

2) 背鳍照片质量识别（Scoring）：定位并截取后的背鳍照片被有经验的海洋生物学家根据照片质量划分为四个质量等级，0-60 分，60-69 分，70-79 分，80-100 分。照片质量的评分标准包括多个维度：背鳍的大小、背鳍的出水程度、是否对焦成功、背鳍是否水平且正对镜头，等等。其中，0-60 分的背鳍照片属于人眼难以辨别个体的照片，将被丢弃，另外的背鳍照片才进行后续的背鳍个体识别和背鳍分割。对于照片个体识别的算法研究而言，背鳍照片质量可以反映出算法在不同质量照片上的效果，从而帮助人们有针对性地提高算法性能。另外，照片个体识别的成功率除了取决于研究人员的经验或算法的性能，还很大程度地取决于照片质量，每张照片的背后都记录

了拍摄者的信息，通过识别背鳍照片质量，能够反映拍摄者的拍摄水平，从而帮助拍摄中提高拍摄水平。

3) 背鳍个体识别 (Matching): 根据背鳍照片进行中华白海豚的照片个体识别, 这种技术作为一种非侵入性的生态普查技术, 在过去的半个世纪中, 已经被人们广泛地应用[1]。结合其他数据收集技术, 如 GPS、水温等, 长期的照片个体识别还可以让我们对动物种群的人口统计数据、生活史、运动方式和社会结构有更深入细微的了解, 尤其是对于无法接近的自由生活的动物如鲸豚类动物<sup>[2]</sup>。

4) 背鳍分割 (Segmentation): 对背鳍照片进行像素级的分割, 主要是为了得到照片中背鳍的像素面积, 该参数结合其他收集的数据, 可以用于进一步研究中华白海豚。手工分割背鳍区域的过程较为繁琐耗时, 难以长期持续进行, 因此传统手工流程并不包含这一步骤。而随着技术进步, 该步骤可以通过计算机快速地完成, 从而快速地得到该参数。

## 2.2 数据集描述

整个数据集分为 2 个地区 4 个步骤的 8 个子数据集, 各个子数据集的图片数如表 2-1 所示。钦州地区的数据集 (Qinzhou Dataset) 和汕头地区的数据集 (Shantou Dataset) 各具特点, 其中 Qinzhou Dataset 的时间跨度较短, 个体数量较多, 照片质量较高。数据集下载地址: <https://github.com/PICWD/PICWD>。

表 2-1 各个子数据集的图片数量

	背鳍定位	背鳍照片质量识别	背鳍个体识别	背鳍分割
Qinzhou Dataset	37297	42949	31093	31093
Shantou Dataset	16216	20048	12656	12656

### 2.2.1 照片的收集与标记

数据集中的照片是海洋生物学家乘坐快艇在沿岸海域, 通过单镜头反光相机搭配长焦镜头拍摄所得, 拍摄场景如图 2-2 所示。拍摄时, 他们会同时记录照片的日期时间、船只的 GPS、水温等。数据集中的照片拍摄于两个地区的沿岸海域, 分别是 Qinzhou (广西省钦州市) 和 Shantou (广东省汕头市), 如图 2-3 所示。

数据集的建立对应中华白海豚的照片个体识别的整体流程, 包含对应四个步骤的四个子数据集, 而且每个子数据集按地区分为 Qinzhou 数据集和 Shantou 数据集。其中前三个子数据集 (背鳍定位、背鳍照片质量识别、背鳍个体识别) 的标记是手工通过 Discovery 软件<sup>[15]</sup> 来完成, 最后一个子数据集 (背鳍分割) 的标记是手工通过 VIA 标记工具<sup>[44]</sup> 来完成。



图 2-2 中华白海豚照片的拍摄场景

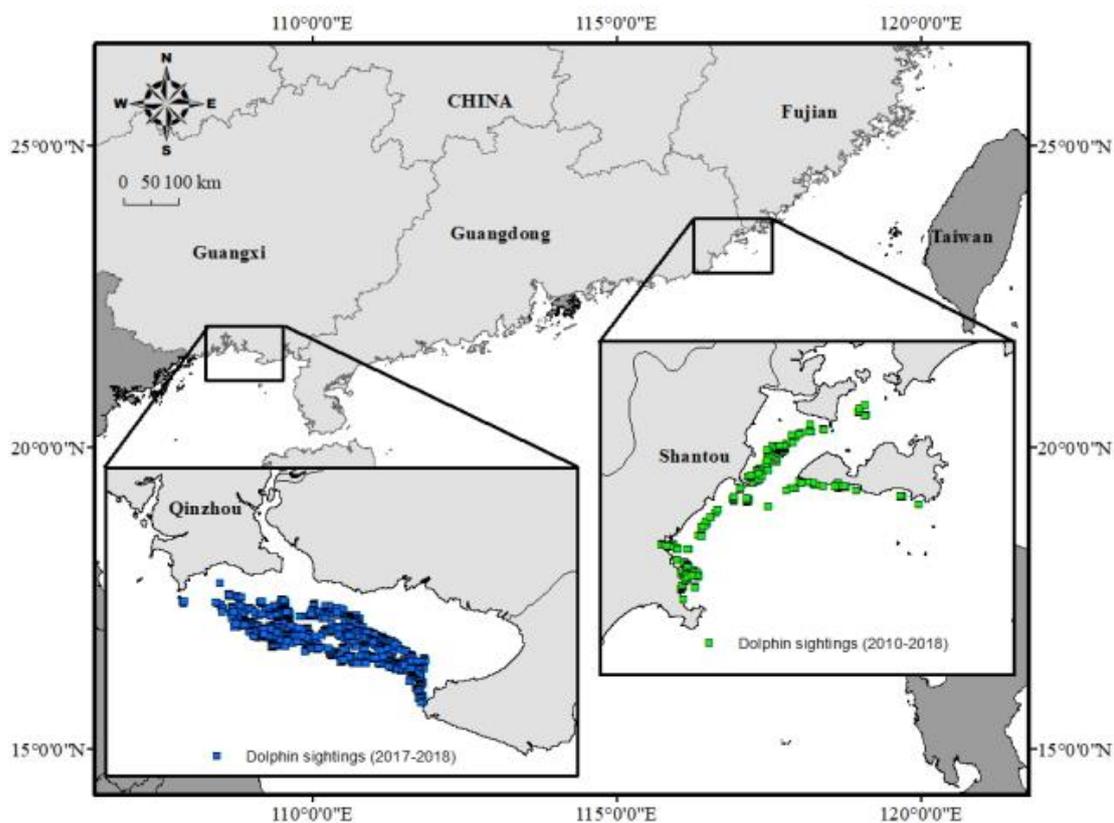


图 2-3 拍摄中华白海豚照片时的地理坐标

## 2.2.2 背鳍定位

背鳍定位子数据集主要包括原图照片和背鳍区域标记框 (bounding box), 如图 2-4 所示。该子数据集按地区分为 Qinzhou Dataset 和 Shantou Dataset, 表 2-1 给出了相应的照片拍摄日期、原图照片数量、原图照片尺寸、背鳍区域标记框数量。参考 Widerface 人脸定位数据集[45], 我们还在表 2-1 中给出了背鳍区域标记框的高度及相应的

数量占比。可见，在 Qinzhou Dataset 中，几乎全部背鳍区域标记框的高度都小于等于 1000，在 Shantou Dataset 中，大多数的背鳍区域标记框的高度都小于等于 500。另外，每张原图照片中都包含 1 到 5 个背鳍区域的标记框，图 2-5 给出了该子数据集中相应于背鳍区域数量的原图照片数量分布图。

表 2-2 背鳍定位子数据集的相关信息

	Qinzhou Dataset		Shantou Dataset	
拍摄日期	2017.10 - 2018.9		2010.8 - 2018.10	
原图照片数量	37297		16216	
原图照片尺寸	5472 × 3648		1936 × 1288 - 5472 × 3648	
背鳍区域 标记框数量	43824		20338	
背鳍区域标记框高 度及数量占比	height ≤ 500	75.6%	height ≤ 200	34.4%
	500 < height ≤ 1000	23.5%	200 < height ≤ 500	60.6%
	height > 1000	0.9%	height > 500	4.7%

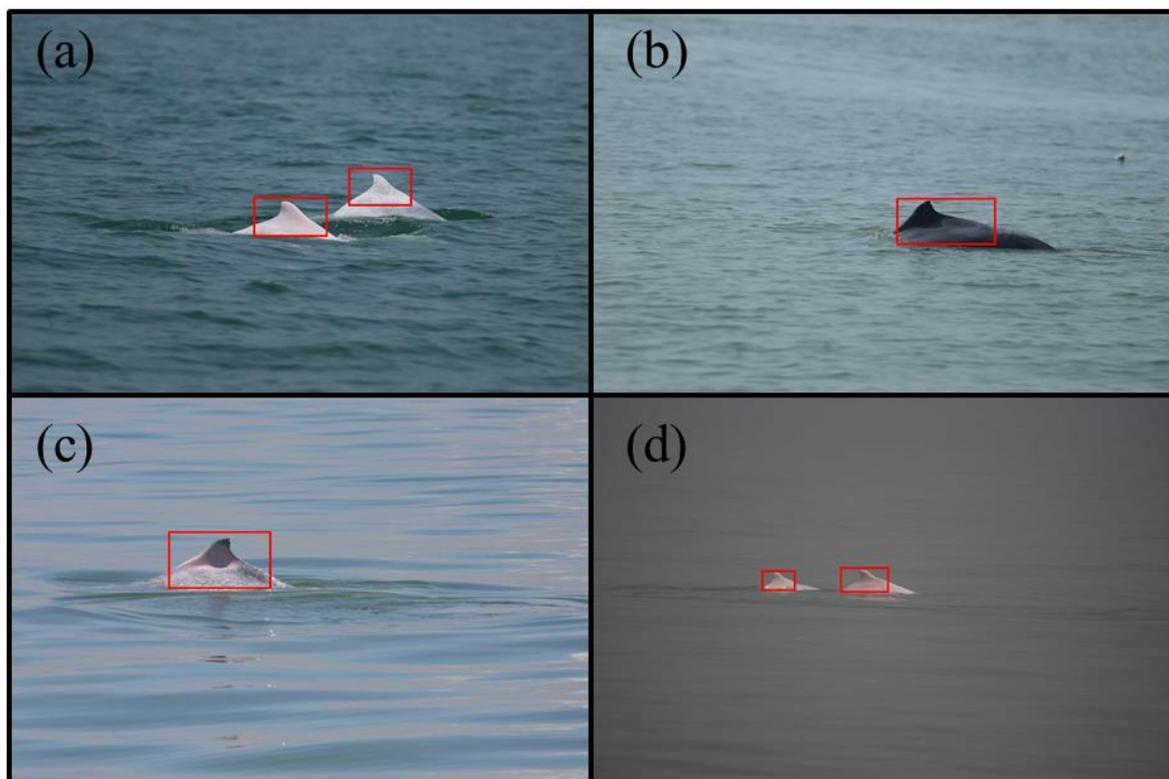


图 2-4 背鳍定位子数据集的示例图（其中图 a、b 来自 Qinzhou Dataset，图 c、d 来自 Shantou Dataset）

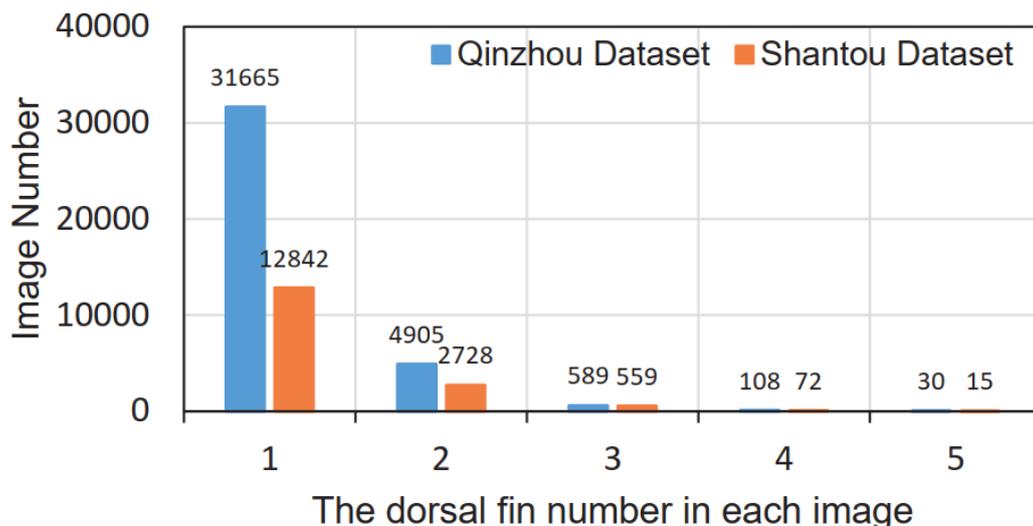


图 2-5 相应于背鳍区域数量的原图照片数量分布图

### 2.2.3 背鳍照片质量识别

背鳍照片质量识别的子数据集主要包括定位截取出的背鳍照片和相应的质量标签，如图 2-6 所示，其中，Qinzhou Dataset 包括 42949 张背鳍照片，Shantou Dataset 包括 20048 张背鳍照片。在全部的背鳍照片中，少数照片由于难以给出合适的照片质量而没有相应的质量标签。各地区各个质量标签下的照片数量如图 2-7 所示，其中，相比于 Shantou Dataset，Qinzhou Dataset 中 80-100 分的照片数量占比更高，60-69 分的照片数量占比更低，可见 Qinzhou Dataset 的整体照片质量更高。

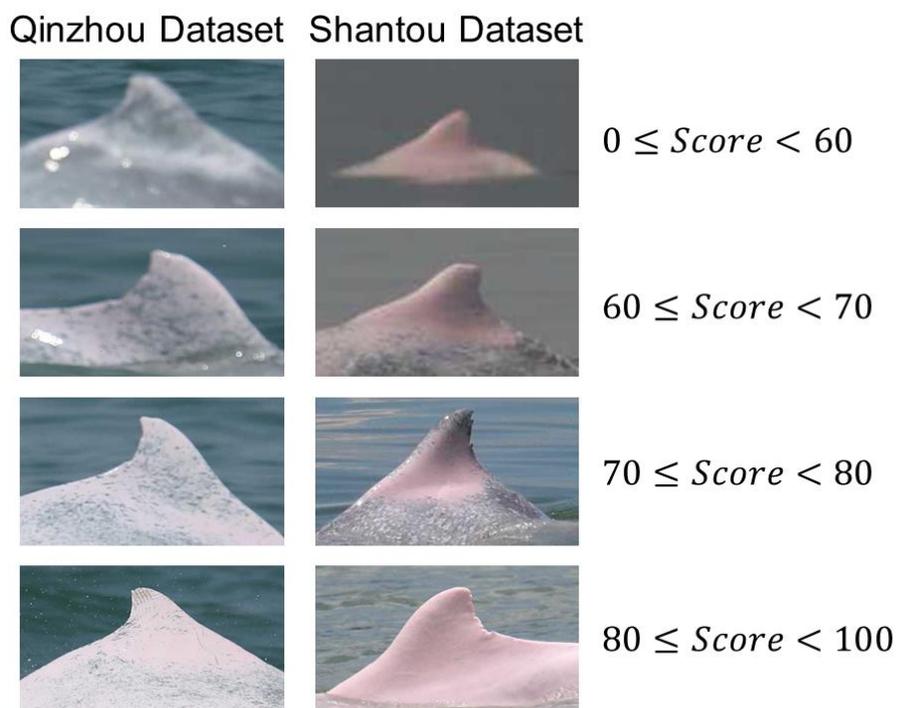


图 2-6 背鳍照片和相应的质量标签

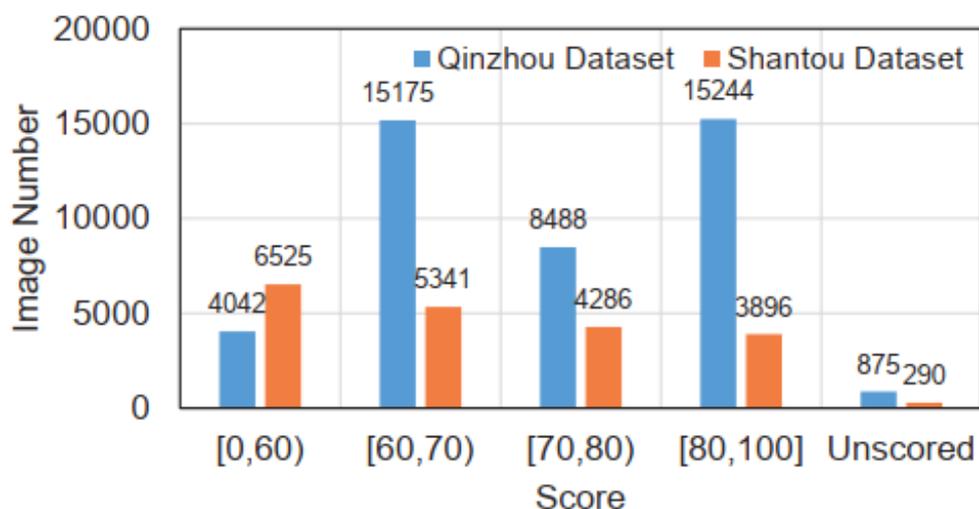


图 2-7 各地区各个质量标签下的照片数量分布图

### 2.2.4 背鳍个体识别

背鳍个体识别子数据集主要包括定位截取出的高于 60 分的背鳍照片和相应的个体标签，其中，Qinzhou Dataset 包括 31093 张背鳍照片和 140 个个体，Shantou Dataset 包括 12656 张背鳍照片和 20 个个体。在全部的高于 60 分的背鳍照片中，少数照片由于难以给出合适的个体标签而没有相应的个体标签。图 2-8 给出了两个地区的数据集中各个个体的背鳍照片数量，存在明显的数据不平衡现象。在两个地区数据集中，我们分别选取了一个个体的多张照片作为示例，如图 2-9 所示。

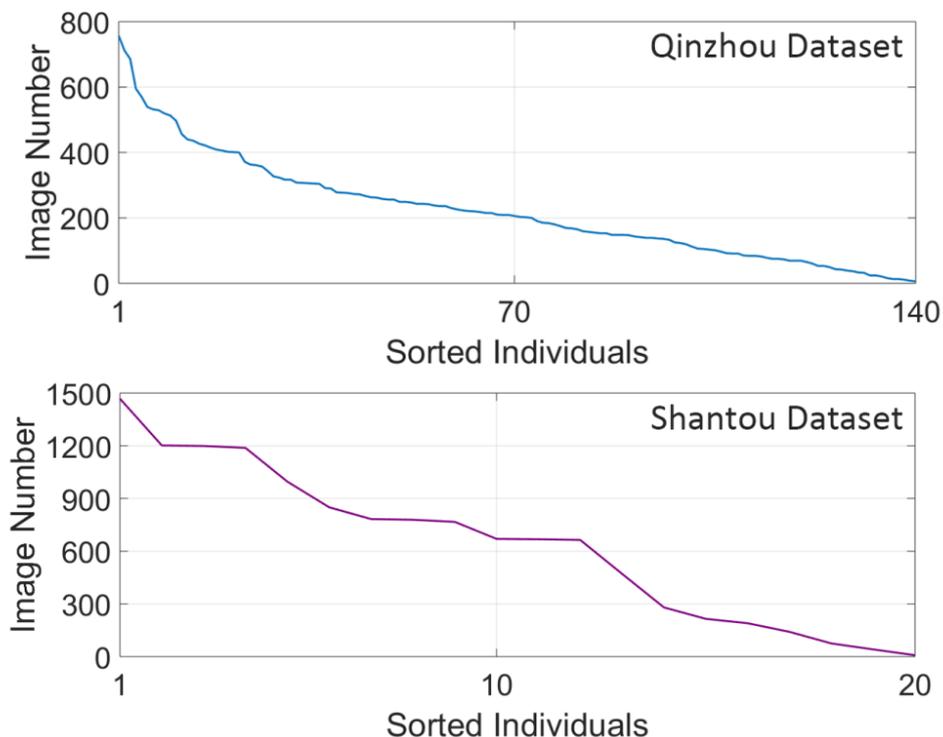


图 2-8 各地区数据集中各个个体的背鳍照片数量分布图

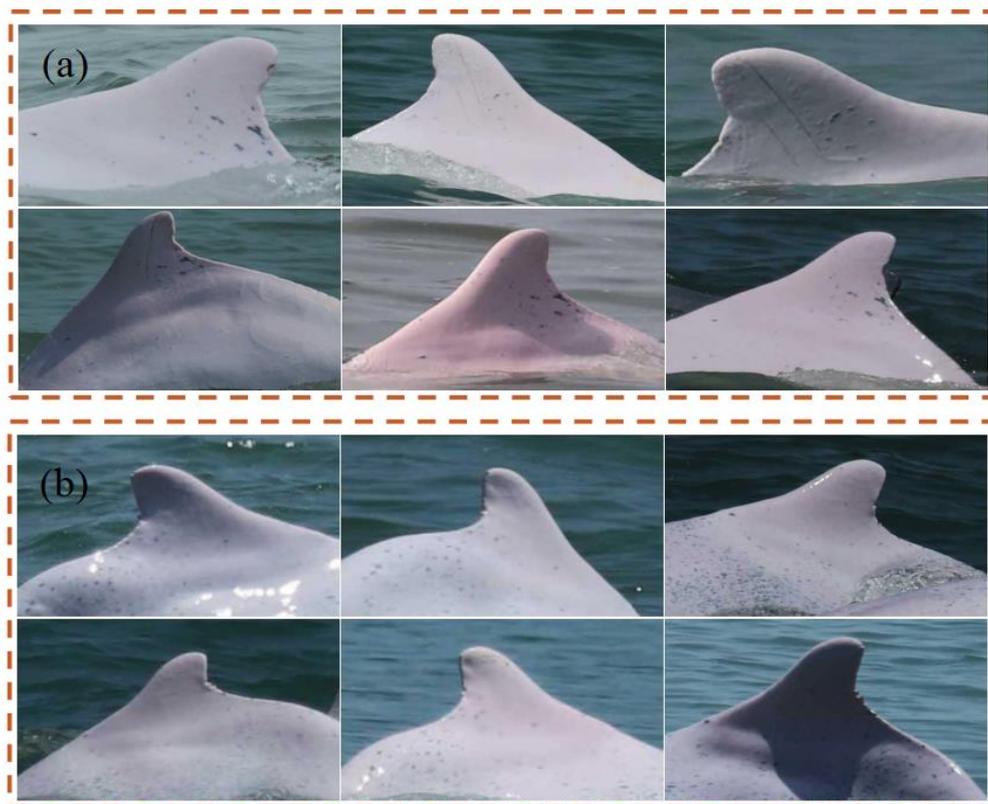


图 2-9 一个个体的多张照片（图 a 所示个体来自 Qinzhou Dataset，图 b 所示个体来自 Shantou Dataset）

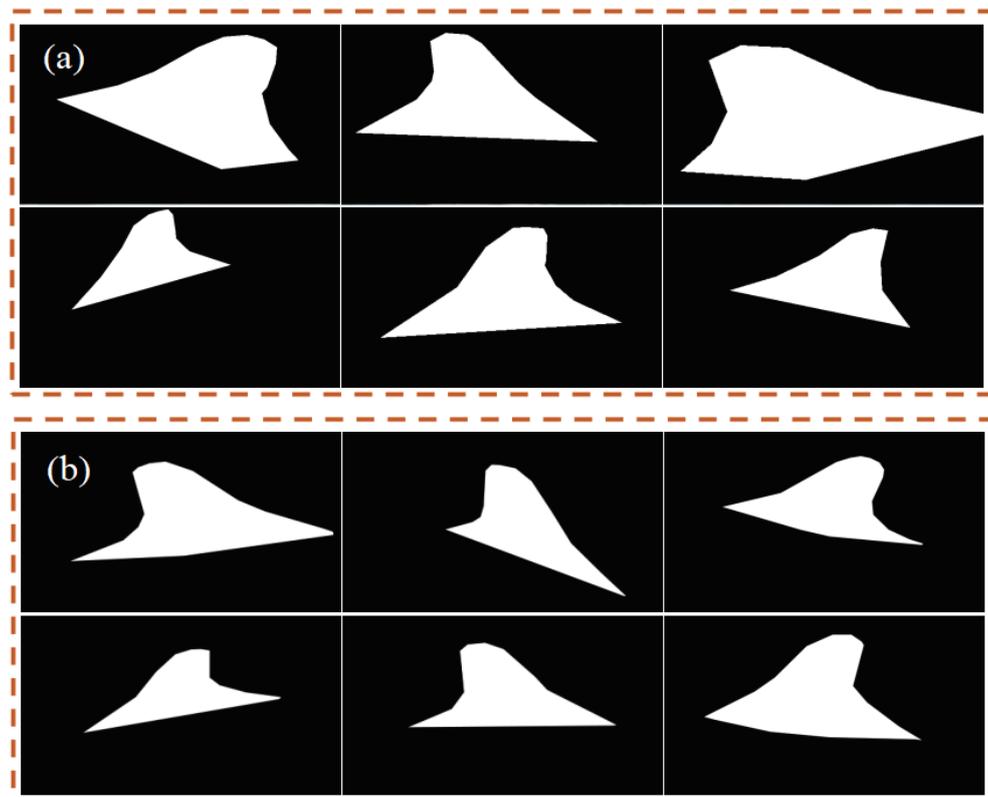


图 2-10 背鳍分割人工分割结果（图 a 对应图 2-9-a 的背鳍照片，图 b 对应图 2-9-b 的背鳍照片）

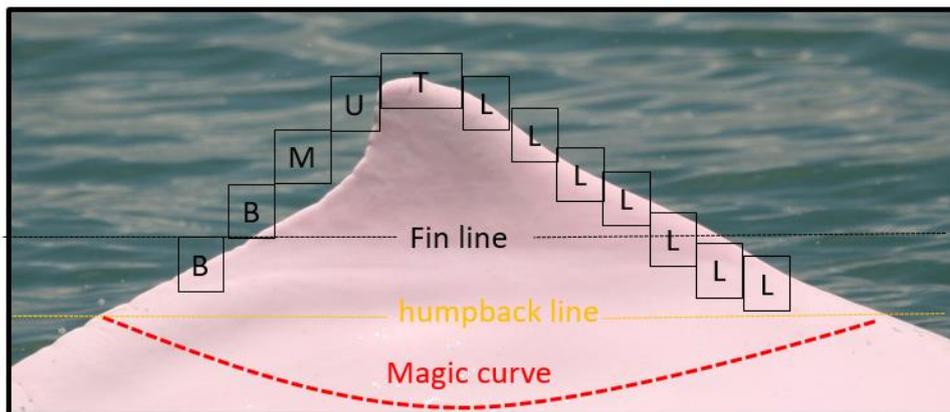


图 2-11 背鳍分割标准

### 2.2.5 背鳍分割

由于部分背鳍照片难以准确分割,背鳍分割子数据集主要包括有个体标签的背鳍照片和相应的背鳍区域二值图,其中,Qinzhou Dataset 包括 31093 张背鳍照片,Shantou Dataset 包括 12656 张背鳍照片。图 2-10 给出了一些人工分割的示例图。人工分割的标准参考图 2-11,背鳍区域是指 Fin line 以上且除去海水背景的部分。

## 2.3 数据集划分

对于各个地区的数据集(Qinzhou Dataset 和 Shantou Dataset),我们选取了 70% 的照片作为训练集,30% 的照片作为测试集。考虑到实际应用过程中,系统从过去的照片进行学习,并应用在将来的照片上。所以,以 Qinzhou Dataset 为例,其中所有的照片先按拍摄时间从前到后进行排序,前面 70% 的照片作为训练集,后面 30% 的照片作为测试集。

## 第 3 章 基于深度学习的照片个体识别

### 3.1 背鳍定位

#### 3.1.1 目标检测

背鳍定位是一个特定场景下的目标检测问题。目标检测 (Object Detection) 作为一个基础性的、有挑战性的计算机视觉问题, 已经被人们研究了几十年。目标检测主要是为了, 在某些给定的图像中确定是否存在给定目录中的物体 (如人、汽车、自行车、狗等), 如果存在, 则给出该物体在图像中的空间位置, 通过包围框 (Bounding Box) 进行输出<sup>[46,47]</sup>。在过去几年, 通过深度学习技术<sup>[48,49]</sup>自动地从图像数据中学习特征, 这已经被证明是有效的方法。同时, 深度学习技术也给目标检测领域带来很大的改进, 如图 3-1 所示<sup>[50]</sup>, 我们可以看到从 2012 年深度学习的出现开始, 目标检测领域的算法性能出现了显著的改进。

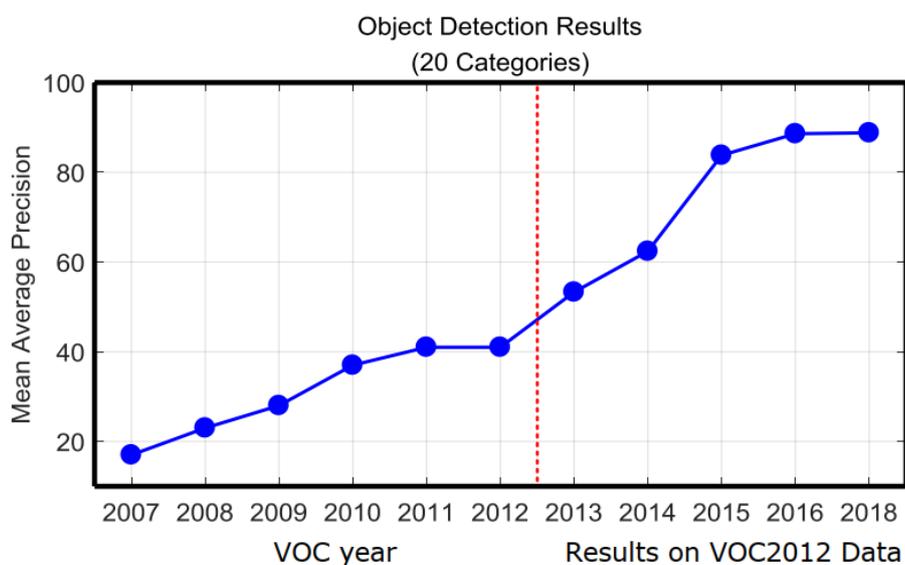


图 3-1 最近几年的目标检测算法性能的提升<sup>[50]</sup> (PASCAL VOC 数据集的结果: VOC2007-2012 参赛者中获胜者的结果)

在最近几年提出来的基于深度学习的目标检测算法中, 有一些算法是具有里程碑意义的, 如图 3-2 所示, 几乎所有的其他提出的算法都是基于上面中的某一算法的框架, 尝试进行改进<sup>[50]</sup>。这些具有里程碑意义的算法, 主要可以分为两大类: 两阶段方法 (Two-stage Methods): 在整个流程中包括一个预处理的步骤用于提取可能的目标区域; 一阶段方法 (Single-stage Methods): 没有把提取可能目标区域这一步与整个流程分离开<sup>[50,51]</sup>。本文在这些具有里程碑意义的方法中选取了 4 种方法进行测试, 包括两种两阶段方法 (Faster R-CNN 和 R-FCN) 和两种一阶段方法 (SSD 和 YOLOV3),

主要原因是：这些方法在几个通用目标检测的数据集（如 PASCAL VOC<sup>[46]</sup> 和 MS COCO<sup>[52]</sup>）上取得很好的效果，这些方法已经在很多其他领域被使用，并证明是有效的。

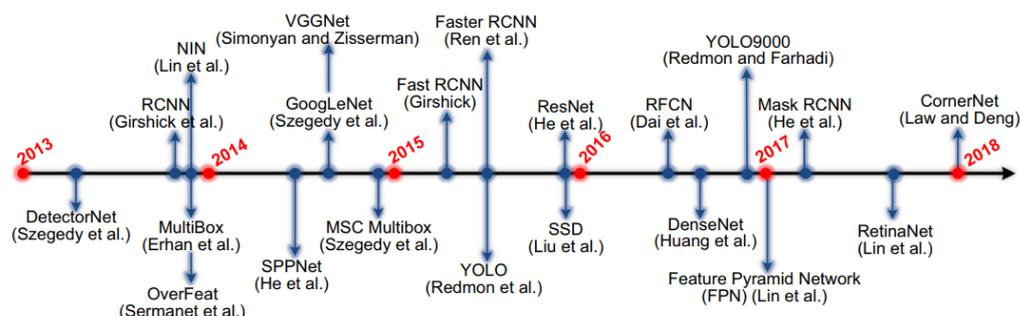


图 3-2 具有里程碑意义的目标检测算法<sup>[50]</sup>

Faster-RCNN<sup>[53,54]</sup>是在 Fast R-CNN<sup>[55]</sup> 的基础上提出来的。虽然 Fast R-CNN 显著加快了检测过程，但它仍然是通过选择性搜索在网络外提取可能的目标区域，这个计算限制了 Fast R-CNN 的速度。后续的研究表明，CNN 具有在 CONV 层中定位对象的能力，而该能力在 FC 层中被削弱。因此，Ren 等人提出了 Faster R-CNN 框架，其中包含一个 RPN 模块用于提取可能的目标区域。这个方法使用了同一个骨干网络和最后一个卷积层的特征，来完成目标区域的提取以及区域目标的分类，如图 3-3 所示，从而实现了高效的目标区域提取。RPN 首先在骨干网络最后的 CONV 层上生成不同尺度和长宽比的  $k$  个参考框(锚点)。锚点位置与图像内容无关，但从锚点提取的特征向量本身与图像内容相关。每个锚点被映射到一个较低维度的向量，该向量被输入到两个同级的 FC 层：一个对象类别分类层和一个 box 回归层。Ren 等人<sup>[53]</sup> 在 2015 年最初提出的 Faster R-CNN 种包含几个交替的训练阶段，后来他们<sup>[54]</sup> 对训练过程进行了简化。使用 VGG16<sup>[56]</sup> 作为骨干网络时，Faster R-CNN 可以在 GPU 上以 5 FPS 进行测试，并且，Faster R-CNN 对每张图片提取 300 个可能的目标区域，可以在 PASCAL VOC 2007 上实现当时最好的目标检测精度。

R-FCN<sup>[57]</sup> 是一种利用全卷积网络进行目标检测的方法，是在 Faster R-CNN 的基础上提出来的。Faster R-CNN 相比于 Fast R-CNN 虽然实现了大幅的速度改进，然而其中每一个 RoI（可能的目标区域）仍然需要通过一个子网络进行计算，而且每张图片一般有几百个 RoI。Dai 等人提出了 R-FCN，在图像上进行完全的卷积计算（没有隐藏的 FC 层），几乎整个图像上的计算都是共享的。如图 3-3 所示，R-FCN 与 Faster R-CNN 仅在 RoI 子网络上有所不同。在 Faster R-CNN 中，RoI 池层之后的计算是无法共享的，所以 Dai 等人提出利用所有的 CONV 层构建一个共享的 RoI 子网络，并在预测前的最后一层 CONV 层提取 RoI。然而，他们发现这种朴素的设计得到的结果精度相当低，他们推测是因为更深的 CONV 层对类别信息更敏感，对位置信息不

敏感。基于这一观察，他们以一组专门的 CONV 层作为 FCN 输出，构建了一组位置敏感的分层图，并在其上添加了位置敏感的 RoI 池化层。他们表明，使用 ResNet101 的 R-FCN 可以达到与 Faster R-CNN 相当的准确性，并且实现更快的运行时间。

SSD<sup>[58]</sup> 是在 Faster R-CNN<sup>[53,54]</sup>和 YOLO<sup>[59]</sup> 的基础上提出来的。为了在不牺牲太多检测精度的情况下保持实时速度，Liu 等人<sup>[59]</sup> 提出了 SSD，其速度快于 YOLO，其精度可与基于区域提取的方法如 Faster R-CNN 相当。SSD 有效地结合了 Faster R-CNN 的 RPN、YOLO 和多尺度的卷积特征<sup>[60]</sup> 中的思想，实现了快速的检测速度，并同时保持了较高的检测质量。与 YOLO 一样，SSD 预测固定数量的边界框和分数，然后使用 NMS 算法产生最终的检测结果。SSD 中的 CNN 网络是全卷积的，其中前面的网络是基于一个标准的架构，如 VGG<sup>[56]</sup>，随后是几个辅助的 CONV 层，大小逐渐减小，如图 3-3 所示。最后一层的信息对于实现准确的定位可能过于粗糙，无法进行精确定位，因此 SSD 通过在多个 CONV 特征图上进行多尺度的检测，每个 CONV 特征图都可以预测类别分数和目标边界框的偏移量。在 VOC2007 测试中，对于输入为 300×300 的图像，Faster R-CNN 以 7 帧/秒的速度实现 73.2% 的 mAP，YOLO 则以 45 帧/秒的速度实现为 63.4%，而 SSD 以 59 帧/秒的速度实现 74.3% 的 mAP。

YOLOV3<sup>[61]</sup> 是在 YOLO<sup>[59]</sup> 和 YOLOV2<sup>[62]</sup> 的基础上提出来的。YOLO<sup>[59]</sup> 在 2016 年被提取，它将目标检测看作是一个回归问题，从图像像素得到边界框坐标和对应类别概率。它抛弃了区域提取的步骤，使用一小组候选区域直接预测结果。与基于区域提取的方法(如 Faster R-CNN)从局部区域的特征进行预测不同，YOLO 使用全局图像的特征。具体来说，YOLO 将图像分成 S×S 网格，每个网格预测 C 类概率、B 个边界框坐标和置信度分数，如图 3-3 所示。通过完全抛弃区域提取的步骤，YOLO 运行速度为 45 帧/秒，而 fast YOLO<sup>[59]</sup> 的速度为 155 帧/秒。YOLO 的缺点在于对小对象的定位效果不好，并且对边框的预测准确度不够高。YOLOV2<sup>[62]</sup> 和 YOLO9000<sup>[62]</sup> 在 2017 年被提出。YOLOV2 主要在 YOLO 的基础上，用 GoogleNet<sup>[63]</sup> 代替简单的 Darknet19，加上批标准化<sup>[64]</sup>，移除全连接层，使用多种大小及长宽比的先验框进行 kmeans 的学习，以及进行多尺度的训练。YOLOV2 在保持 YOLO 运行速度快的优势的同时，明显提升了预测准确度。YOLO9000 则提出了一种联合优化方法，在 ImageNet 分类数据集和 COCO 检测数据集上同时训练，用 WordTree 将多个来源的数据组合起来。这种联合训练允许 YOLO9000 执行弱监督检测，即检测没有包围框标记的类别。YOLOV3<sup>[61]</sup> 在前面的方法的基础上，保持速度优势，提升了预测准确度，尤其是加强了识别小物体的能力。YOLOV3 使用了新的网络结构 Darknet53<sup>[61]</sup>，利用多尺度的图像特征进行检测，并且用 Logistic 的输出代替 Softmax 进行分类预测。YOLOV3 在 MSCOCO 数据集上使用 608 作为输入图像尺寸，mAP 准确度为 57.9%，接近最好的 59.1%，并且预测速度是其 3-4 倍。

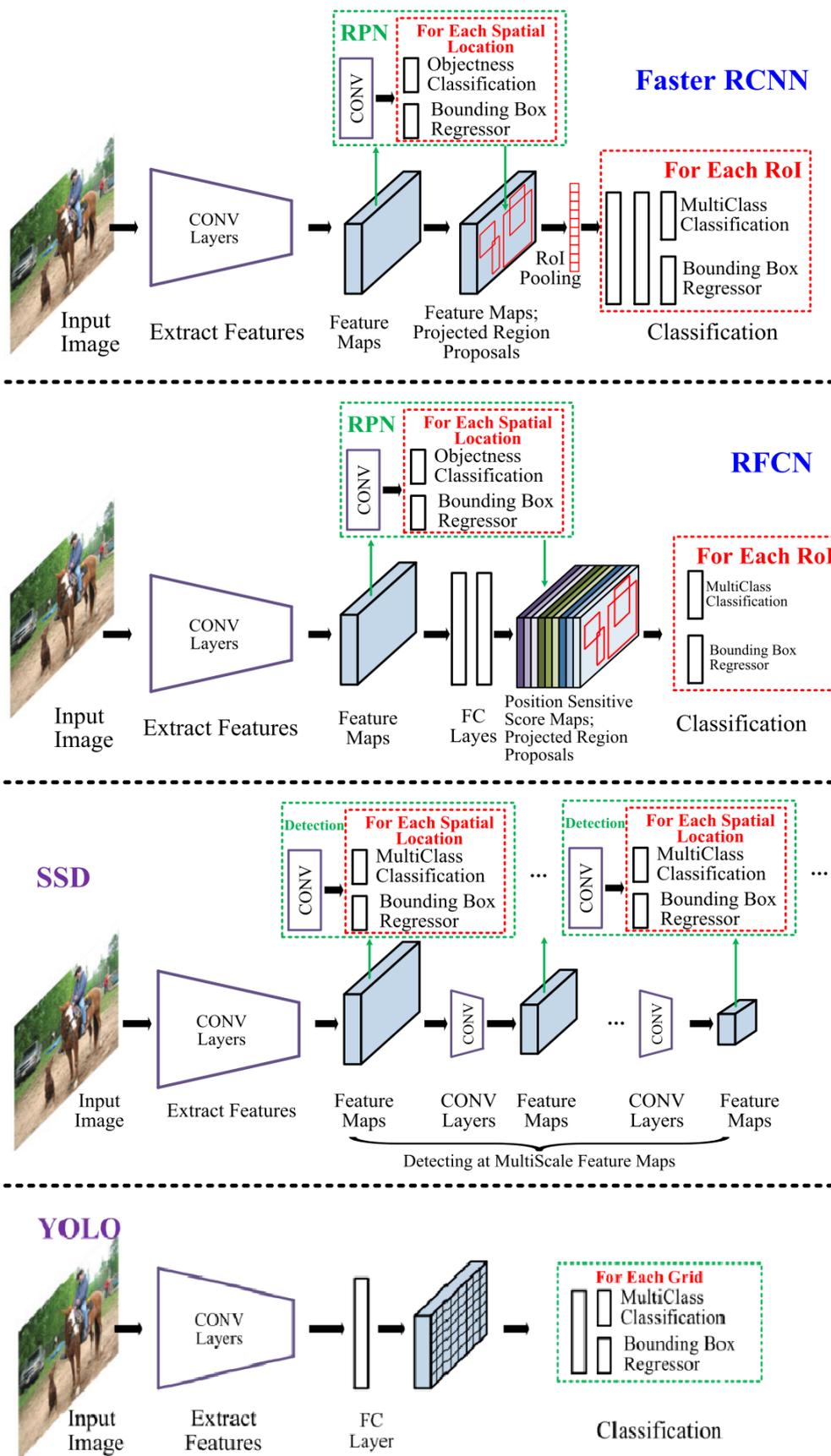


图 3-3 Faster RCNN、RFCN、SSD、YOLO 的流程框图<sup>[50]</sup>

### 3.1.2 实验结果

我们实验所使用的 Faster R-CNN 和 RFCN 都使用了 ResNet-101<sup>[65]</sup> 作为骨干网络，SSD 使用了 MobileNetV1<sup>[66]</sup> 作为骨干网络，而 YOLOV3 则使用 Darknet53 作为骨干网络。

我们选择了 Pascal VOC<sup>[46]</sup> 的评价指标 mAP(mean Average Precision)。Faster R-CNN、R-FCN、SSD、YOLOV3 这四种目标检测算法在背鳍定位的两个地区的数据集上，定位结果如表 3-1 所示。综合两个地区的数据集，可见 Faster R-CNN 效果最好。图 3-4 展示了 Faster R-CNN 的定位结果示例图，可见当前基于深度学习的目标检测算法在背鳍定位的数据集上已经能够实现较好的结果。

表 3-1 背鳍定位的定位结果 (mAP)

	Qinzhou Dataset	Shantou Dataset
Faster R-CNN	0.972	0.970
R-FCN	0.974	0.961
SSD	0.948	0.912
YOLOV3	0.960	0.910

## 3.2 背鳍照片质量识别

### 3.2.1 图像分类

背鳍照片质量识别在本文是作为一个的图像分类问题，输入一张背鳍照片，输出其所属的照片质量类别（0-60 分，60-69 分，70-79 分或 80-100 分）。图像分类在过去主要是通过 SIFT 等手工设计的特征描述子提取图像特征，再通过 SVM 等分类算法进行分类。然而，从 2006 年开始，人们提出了很多方法去克服训练深度卷积神经网络（CNN）时所遇到的问题<sup>[56,67,68]</sup>，并且证明通过深度学习技术<sup>[48,49]</sup>自动地从图像数据中学习特征是更有效的方法。尤其是在 2012 年，Krizhevsky 等人<sup>[68]</sup>提出了一个经典的 CNN 结构 Alexnet，在图像分类的比赛 ILSVRC<sup>[69]</sup>中取得了最好的成绩，并且其大幅领先于其他方法。从 Alexnet 的成功开始，越来越多的 CNN 结构被提出，并且一次又一次地突破 ILSVRC 比赛的结果。随着 2015 年 ResNet<sup>[65]</sup>的提出，深度 CNN 在 ILSVRC 比赛的分类错误率降低到了 3.6%，如图 3-5 所示<sup>[50]</sup>，已经低于人类的分类错误率 5%。本文选取了 4 种目前效果最好，并且在各个领域中得到广泛应用的深度 CNN 网络（ResNet-101<sup>[65]</sup>、InceptionV4<sup>[70]</sup>、InceptionResnetV2<sup>[70]</sup>、MobileNetV1<sup>[66]</sup>），在背鳍照片质量识别的子数据集上进行测试。

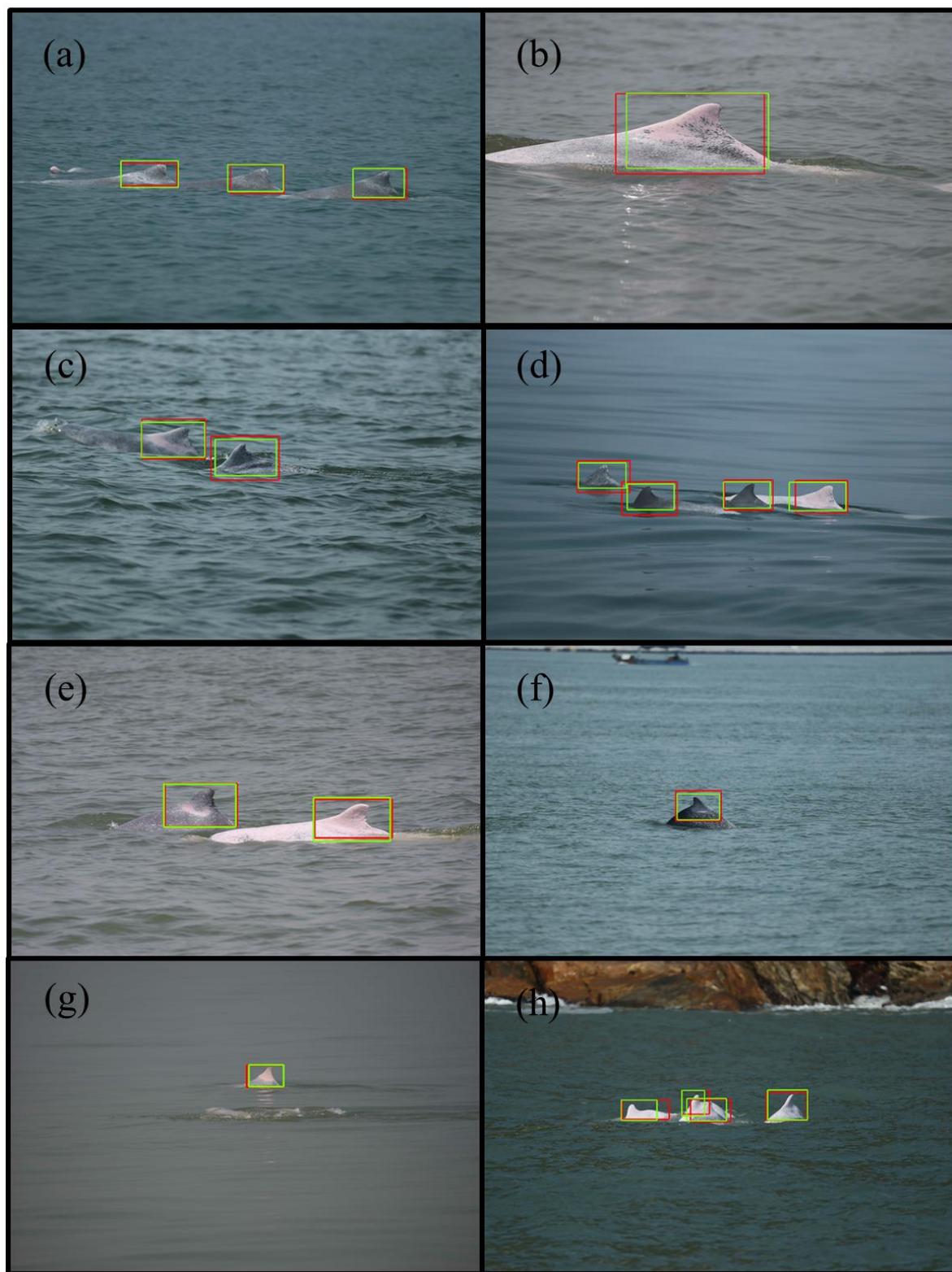


图 3-4 Faster R-CNN 的定位结果 (绿色框表示 Faster R-CNN 的检测框, 红色框表示人工标记框, 图 a-d 来自 Qinzhou Dataset, 图 e-h 来自 Shantou Dataset)

目前已经有很多不同结构的 CNN，但是它们的基本模块是十分类似的。LeCun 等人提出了第一个具有实用意义的 CNN 网络 LeNet-5<sup>[71,72]</sup>，用于手写字体的识别。我们以 LeNet-5 为例介绍 CNN 的基本模块：卷积层、池化层和全连接层，如图 3-6 所示<sup>[73]</sup>。卷积层主要是用于学习得到输入图像的特征。卷积层由多个卷积核组成，用于计算不同的特征图。具体来说，特征图中的每个神经元都连接到前一层相邻神经元的区域，称为感知域。将输入与卷积层进行卷积，然后在卷积结果上应用非线性激活函数（如 sigmoid、tanh 和 ReLU 等），得到新的特征图。激活函数将非线性引入到 CNN 中，这是多层网络检测非线性特征所需要的。池化层的目标是通过降低特征图的分辨率来实现移位不变性。它通常位于两个卷积层之间。池化层的每个特征映射都与前一卷积层对应的特征映射相连接。典型的池化操作是平均池化和最大池化。第 1 个卷积层主要用于检测低层次特征，如边缘和曲线，而更高层次的卷积层则用于对更抽象的特征进行编码。通过叠加几个卷积和池化层，我们可以逐步提取更高层次的特征表示。经过几个卷积和池化层之后，可能会有一个或多个全连接层，目的是执行高级推理。它们将前一层的所有神经元连接到当前层的每个神经元上，生成全局语义信息。CNN 的最后一层是输出层，对于分类任务，通常使用 softmax 和交叉熵损失函数。训练 CNN 是一个全局优化的问题，通过最小化损失函数，我们可以找到最佳的参数。随机梯度下降法是优化 CNN 网络的常用方法。

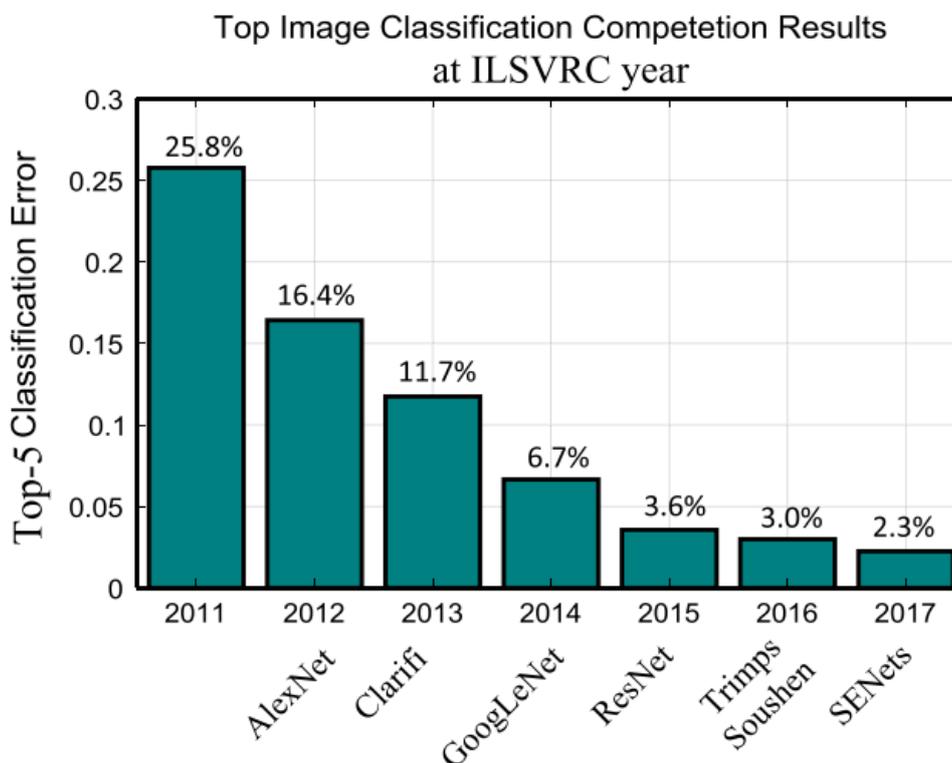


图 3-5 从 2011 年-2017 年 ILSVRC 比赛图像分类任务获胜者的结果<sup>[50]</sup>

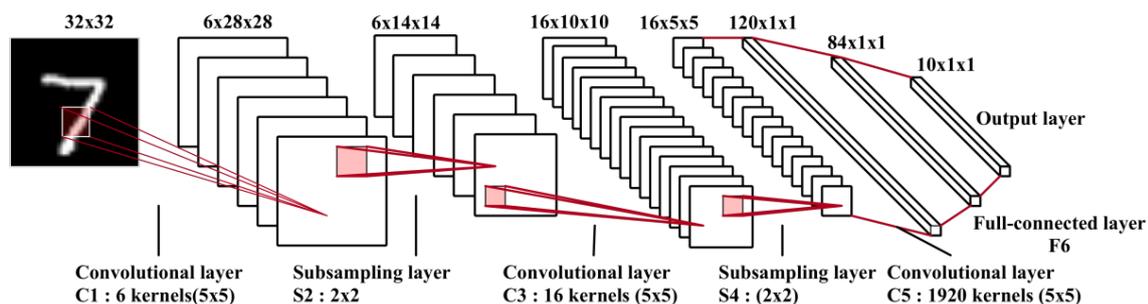


图 3-6 LeNet-5 网络结构<sup>[73]</sup>

ResNet-101<sup>[65]</sup> 是 ILSVRC 2015 年图像分类比赛的冠军。它主要的目的是让 CNN 的识别准确率能够随着网络的加深而提高。从 Alexnet 的成功之后，各种 CNN 结构先后被提出，它们普遍的特点是网络的层级越来越深。然而，He 等人<sup>[65]</sup> 通过实验发现，CNN 在达到一定深度后再增加网络层数，并不能进一步提高分类准确率。他们提出了一种深度残差网络结构，该结构负责学习特征的部分学习的是输出与输入的差值。这种改变不会增加额外的参数，也不会增加计算复杂度，并且可以用带有反向传播的随机梯度下降法进行端到端训练。他们可以通过对该结构进行网络深度的堆叠，便可以实现准确率的提升。

InceptionV4<sup>[70]</sup> 是在 InceptionV1<sup>[74]</sup>、InceptionV2<sup>[75]</sup> 和 InceptionV3<sup>[75]</sup> 的基础上提出来的。其中，InceptionV1 最初被提出时，在 ILSVRC 2014 年获得了最好的分类和检测性能。它主要是利用 1\*1, 3\*3, 5\*5 的小卷积核构造出 Inception module 的基本结构，并在此基础上设计了 22 层的网络结构，这种方式能够大大减少参数量。InceptionV2 则在 InceptionV1 的基础上，在中间的层上，用 1\*n 和 n\*1 的卷积代替 n\*n 的卷积，用两个 3\*3 的卷积代替 5\*5，从而在不提升参数的基础上提升网络的表达能力。InceptionV3 则在 InceptionV2 的基础上，使用 RMSProp 优化方法代替随机梯度下降法，使用 Label Smoothing Regularization (LSR) 正则化，在辅助分类器上使用 BatchNorm，并用 1\*7 和 7\*1 的卷积代替 7\*7 的卷积，从而加快了运算速度，增加了网络非线性，减少网络过拟合。InceptionV4 则在 InceptionV3 的基础上，构造了新的 Inception module，改进了网络结构，最终在 ImageNet 数据集<sup>[69]</sup> 上实现了略优于 ResNet 的分类结果。

InceptionResnetV2<sup>[70]</sup> 是在受到 Resnet 的启发后提出的。它的 Inception module 和 InceptionV4 中的 Inception module 类似，并引入了残差连接。它的计算成本与 InceptionV4 比较接近，并且，它可以在更少的训练迭代次数内达到更高的准确率。

MobileNetV1<sup>[66]</sup> 是主要应用与移动设备的轻量级 CNN 网络。它最大的特点是提出了深度可分离卷积。深度可分离卷积将传统卷积的两步进行分离开来，正常情况下，深度可分离卷积相比于传统卷积，准确率会下降 1%，但是计算量能够减少 8-9 倍。

在 ImageNet 数据集<sup>[69]</sup>上，使用相同的输入分辨率，MobileNetV1 的准确率与 InceptionV1<sup>[74]</sup>和 VGG16<sup>[56]</sup>十分接近，但是参数量和计算量都减少了很多。

### 3.2.2 实验结果

我们选择了 Top-1 准确率<sup>[69]</sup>作为评价指标。ResNet-101、InceptionV4、InceptionResnetV2、MobileNetV1 这四种图像分类算法在背鳍照片质量识别的两个地区的数据集上分类结果如表 3-2 所示。

表 3-2 背鳍照片质量识别的分类结果 (Top-1)

	Qinzhou Dataset	Shantou Dataset
ResNet-101	0.670	0.563
InceptionV4	0.687	0.582
InceptionResnetV2	0.688	0.579
MobileNetV1	0.657	0.560

可见，目前主流的深度 CNN 网络在这个任务上的效果并不太好。主要原因可能在于，照片质量识别的评价标准很复杂，包括背鳍的大小、背鳍的出水程度、是否对焦成功、背鳍是否水平且正对镜头，等等。而且，定位截取的背鳍照片有着相似的模式，如形状和纹理。以上这些因素导致了背鳍照片质量识别难以解决，也为算法的改进提供了空间。

### 3.3 背鳍个体识别

背鳍个体识别在本文是作为一个的图像分类问题，输入一张背鳍照片，输出其所属的个体（Qinzhou Dataset 有 140 个个体，Shantou Dataset 有 20 个个体）。由于与背鳍照片质量识别同为图像分类问题，本文选取了与其相同的 4 个深度 CNN 网络（ResNet-101<sup>[66]</sup>、InceptionV4<sup>[71]</sup>、InceptionResnetV2<sup>[71]</sup>、MobileNetV1<sup>[67]</sup>）进行测试。由于背鳍个体识别的分类类别数较多，我们选择了 Top-1 准确率和 Top-5 准确率<sup>[69]</sup>作为评价指标。这四种图像分类算法在背鳍个体识别的两个地区的数据集上分类结果如表 3-3 所示。另外，我们也给出了四种图像分类算法在背鳍个体识别的两个地区的数据集上，对应各个照片质量的分类结果，如表 3-4 所示。

可见，目前主流的深度 CNN 网络在这个任务上仍然有提升的空间。CNN 识别错误的主要原因可能有：不同个体的背鳍照片差别（类间差别）较小；背鳍个体识别的数据集存在明显的数据不平衡现象，不利于 CNN 网络的训练；CNN 网络对于照片质量较差的背鳍照片识别效果不好。以上这些因素导致目前结果欠佳，同时也是未来算法提升的关键。

表 3-3 背鳍个体识别的分类结果

	Qinzhou Dataset		Shantou Dataset	
	Top-1	Top-5	Top-1	Top-5
ResNet-101	0.806	0.925	0.863	0.966
InceptionV4	0.811	0.924	0.860	0.966
InceptionResnetV2	0.799	0.913	0.862	0.967
MobileNetV1	0.750	0.895	0.839	0.962

表 3-4 各个照片质量的背鳍个体识别的分类结果 (Top-1) (ALL 表示该数据集全部测试照片, 80-100 表示该数据集全部 80-100 分的测试照片, 以此类推)

	Qinzhou Dataset				Shantou Dataset			
	ALL	80-100	70-79	60-69	ALL	80-100	70-79	60-69
ResNet-101	0.806	0.868	0.835	0.682	0.863	0.958	0.898	0.749
InceptionV4	0.811	0.881	0.842	0.671	0.860	0.947	0.879	0.763
InceptionResnetV2	0.799	0.874	0.826	0.654	0.862	0.946	0.886	0.766
MobileNetV1	0.750	0.826	0.773	0.604	0.839	0.939	0.855	0.735

### 3.4 背鳍分割

#### 3.4.1 图像语义分割

背鳍分割在本文是作为一个的图像语义分割问题。图像语义分割也可以看作是像素的分类问题，输入一张图像，输出图像中每个像素点所属的类别。传统的图像分割有基于阈值的、基于分水岭算法的、基于边缘检测的、基于主动轮廓模型的，等等。然而，类似于图像分类和目标检测，在过去几年，通过深度学习技术<sup>[49,50]</sup>自动地从图像数据中学习特征，并基于深度学习得到的特征进行图像分割，这已经被证明是更有效的方法。目前，图像分割领域也有很多网络框架及方法，本文选取了 FCN<sup>[76]</sup> 和 DeeplabV3<sup>[77]</sup> 在背鳍分割子数据集上进行测试。主要原因有：FCN 是深度学习首次应用于图像分割的方法，后来的基于深度学习的图像分割方法都或多或少借鉴其方法；DeeplabV3 在 Pascal VOC 2012<sup>[46]</sup> 图像语义分割数据集上的效果接近最好。

FCN<sup>[76]</sup> 在 2015 年被提出，是深度学习首次应用于图像分割的方法。Jonathan 等人将 CNN 网络最后的全连接层换成了卷积层，并且为了将小分辨率的特征图转化成原图的分辨率大小，引入了反卷积。反卷积又称为解卷积或上卷积，其操作类似于普通卷积，但是它在卷积的输入中进行插值，从而使卷积输出的特征图分辨率增大，如图 3-7 所示。这里的反卷积并非数学意义上的卷积逆运算，只能恢复原图大小，不能恢复原本的数据。FCN 使用了传统的插值方法，如双线性插值等，后来的反卷积更多是进行插值补零，可以提高运行速度，并且使得卷积的权值可以学习。FCN 的优点在

于可以处理任意尺寸的输入图片，并且可以保留原本输入图片中的空间信息。FCN 的缺点在于，由于使用了反卷积操作，得到的结果会比较模糊、平滑，对图像细节不敏感，并且，由于图片中各个像素是分开进行分类的，没有考虑相邻像素的关系，得到的结果缺乏空间一致性和连续性。

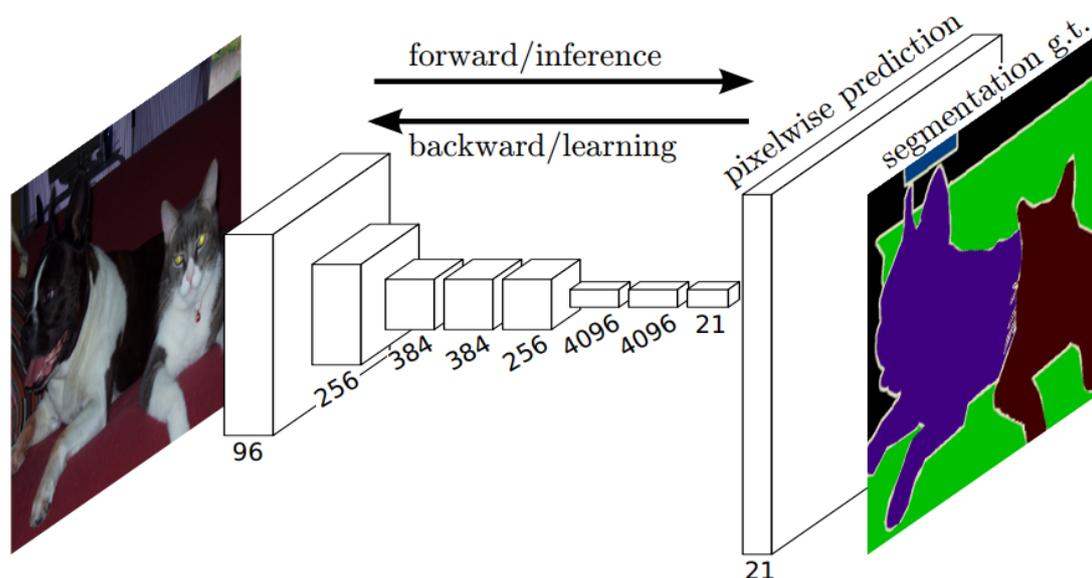


图 3-7 FCN 网络结构

DeeplabV3<sup>[77]</sup> 是在 DeeplabV1<sup>[78]</sup> 和 DeeplabV2<sup>[79]</sup> 的基础上提出来的。在 DeeplabV1 中，为了解决 CNN 多次的池化和下采样导致分辨率下降，位置信息丢失难以恢复，以及 CNN 的空间不变性导致细节信息丢失，引入了空洞卷积（也成为膨胀卷积）和全连接的条件随机场（CRF）。最终，通过简单的结构在当时的 PASCAL VOC 的语义分割数据集上实现了最好的结果。DeeplabV2 在 DeeplabV1 的基础上，为了处理图像中多尺度的物体，利用不同膨胀因子的空洞卷积，通过金字塔结构的池化获取多尺度的物体信息，并进行融合（atrous spatial pyramid pooling/ASPP），在多个数据集上取得了当时最好的效果。DeeplabV3 在 DeeplabV2 的基础上，尝试在更深的结构下使用空洞卷积，在最后一个特征上使用全局平均池化对 ASPP 进行优化，并去掉了 CRF。最终，在 PASCAL VOC 2012 语义分割数据集上实现了当时最好的效果。

### 3.4.2 实验结果

我们实验所使用的 FCN 使用了 VGG<sup>[56]</sup> 作为骨干网络，DeepLabV3 使用了使用了 Xception65<sup>[77]</sup> 作为骨干网络。我们选择 mIoU（mean intersection over union）作为评价指标。这两种算法在背鳍分割的两个地区的数据集上的语义分割结果如表 3-5 所示。图 3-8 展示了 DeepLabV3 的分割效果图。可见，当前基于深度学习的语义分割算法在背鳍分割的数据集上已经能够实现较好的结果。

表 3-5 背鳍分割的图像语义分割结果

	Qinzhou Dataset	Shantou Dataset
FCN	0.816	0.841
DeepLabV3	0.844	0.872

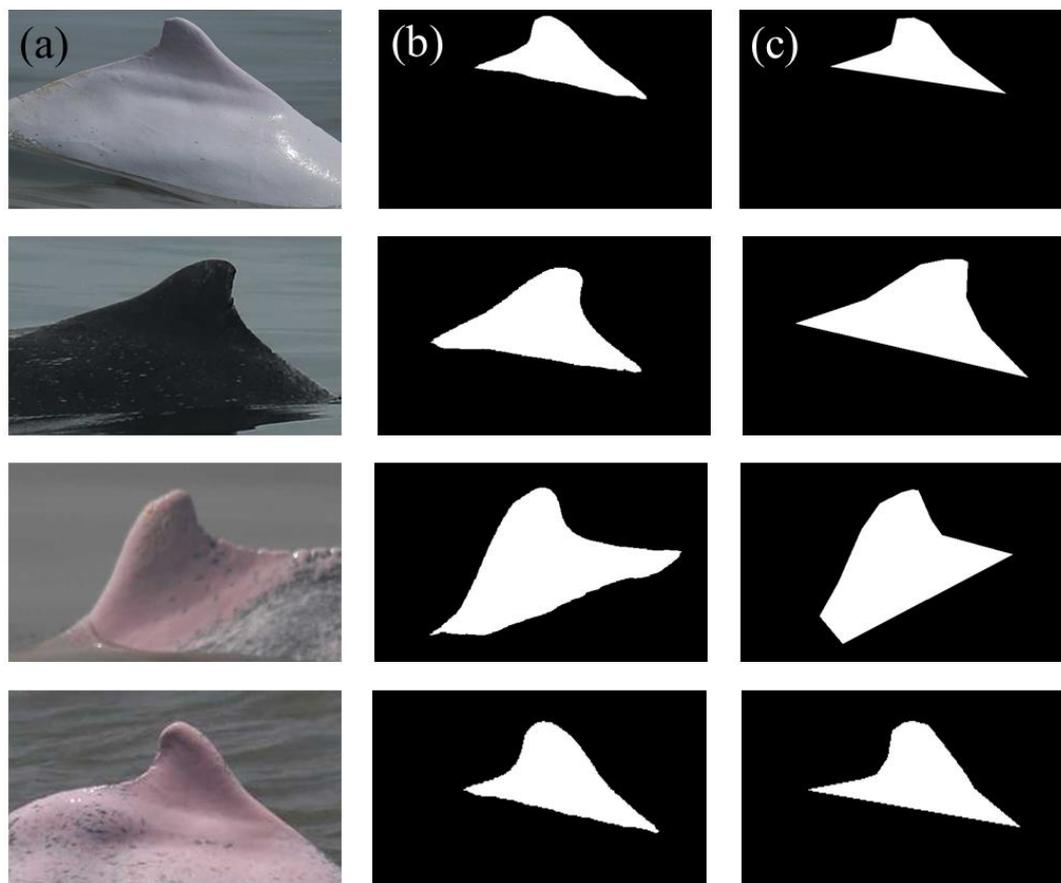


图 3-8 背鳍分割结果示例图 (a)背鳍照片 (b)DeepLabV3 的背鳍分割结果 (c)人工标记 (第 1-2 行照片来自 Qinzhou Dataset, 第 3-4 行照片来自 Shantou Dataset)

### 3.5 交叉地区验证

交叉地区验证, 通过使用一个地区的照片训练网络, 并在另一个地区的照片上测试网络, 以评估不同地区的照片对于背鳍定位、背鳍照片质量识别和背鳍分割的影响。由于两个地区有着完全不同的中华白海豚个体, 因此排除了背鳍个体识别的交叉地区验证。实验的结果如表 3-6 所示。交叉地区训练和测试, 相比于在同一地区的照片中训练并测试, 导致的背鳍定位准确率 (mAP) 下降小于 5%, 导致的背鳍分割准确率 (mIOU) 下降小于 6%, 但是其导致的背鳍照片质量识别准确率 (Top-1) 下降超过 10%。可见, 总体上利用目前的数据集进行训练的 CNN 网络是可以应用到其他地区的中华白海豚上的, 但是背鳍照片质量识别仍然具有挑战性。

表 3-6 交叉地区验证结果

		Shantou Dataset 上训练 Qinzhou Dataset 上测试	Qinzhou Dataset 上训练 Shantou Dataset 上测试
背鳍 定位	Faster R-CNN	0.962	0.921
	R-FCN	0.941	0.941
	SSD	0.843	0.913
	YOLOV3	0.906	0.938
背鳍照 片质量 识别	ResNet-101	0.567	0.463
	InceptionV4	0.565	0.504
	InceptionResnetV2	0.567	0.506
	MobileNetV1	0.555	0.491
背鳍 分割	FCN	0.763	0.809
	DeepLabV3	0.785	0.834

## 第 4 章 基于群信息的背鳍个体识别

### 4.1 介绍

目前关于动物照片个体识别的研究，都是对单个个体的单张图片进行个体识别，即输入一张动物个体的照片，输出该照片中个体对应的个体编号。这方面的研究，具体来说，动物主要包括海洋哺乳动物和陆地动物，而研究内容包括数据集的建立和方法的提出。海洋哺乳动物的数据集，一般只包括动物个体的照片，以及照片中个体的编号。专门用于照片个体识别算法开发及竞赛的公开数据集，据我们了解，只有 3 个，2 个座头鲸的数据集<sup>[11,12]</sup>和 1 个露脊鲸的数据集<sup>[13]</sup>，座头鲸的数据集中是人工截取出的座头鲸尾鳍照片，露脊鲸的数据集中则是航拍的露脊鲸头部照片。海洋哺乳动物的个体识别方法，主要有基于鳍曲线的识别方法、基于关键点比对的方法和基于 CNN 的识别方法，无论如何，这些都是对单个个体的单张图片进行个体识别。基于鳍曲线的识别方法，通过提取动物照片中的背鳍或尾鳍的曲线，根据鳍曲线的特征进行个体识别，研究的对象包括大白鲨的背鳍<sup>[19,20,29]</sup>、海豚的背鳍<sup>[21,24,25,30]</sup>、鲸鱼的尾鳍<sup>[30]</sup>等。基于关键点比对的方法，通过传统图像处理方法提取出个体中的关键点，再与数据库照片的关键点作比对，研究的对象有鲸鲨<sup>[31]</sup>、锥齿鲨<sup>[32]</sup>、蝠鲼<sup>[33]</sup>等。基于 CNN 的识别方法，通过 CNN 进行图像的特征提取，如露脊鲸<sup>[26,35-37]</sup>，海豹<sup>[38]</sup>等。另外，关于陆地动物的数据集和个体识别方法，如大猩猩<sup>[39]</sup>，东北虎<sup>[40]</sup>，大象<sup>[41]</sup>等，也是对单个个体的单张图片进行个体识别。

我们研究的对象是中华白海豚，主要根据其背鳍对其进行照片个体识别。前面我们的背鳍照片个体识别研究，与目前大多数的动物照片个体识别类似，也是对单个个体的单张图片进行个体识别。根据第二章的内容，背鳍个体识别的数据集主要包括定位截取出的高于 60 分的背鳍照片和相应的个体标签。由于中华白海豚的背鳍除了具有鳍曲线的特征，还有着独特多样的颜色和斑点特征，根据第三章的内容，我们把背鳍个体识别作为一个图像分类问题，采用深度 CNN 算法进行处理。对于深度 CNN 算法，输入的是一张定位截取出的高于 60 分的背鳍照片，输出的是相应地区的个体编号（Qinzhou Dataset 有 140 个个体，Shantou Dataset 有 20 个个体）。

这里我们将介绍并引入一个先验信息，群信息，我们有可能利用它提高背鳍个体识别的整体准确率。中华白海豚是一种群居动物，通常会几只或十几只一起活动。为了研究中华白海豚，海洋生态学家通过乘坐快艇在沿岸海域寻找中华白海豚的踪影。当他们发现一群中华白海豚时（通常是几只或十几只），他们会在中华白海豚浮出水面的时候，拍摄尽可能多的照片，其中照片的质量有好有坏。拍摄完这一群中华白海

豚后，他们会记录下这些照片对应的是当天遇到的第几群，因此，在背鳍个体识别的数据集中，每一张背鳍照片都有对应的群信息，如图 4-1 所示。每一群背鳍照片具有这样的特点：每一群背鳍照片中通常只有几个或十几个个体；每个个体通常有高质量的背鳍照片，也有低质量的背鳍照片。

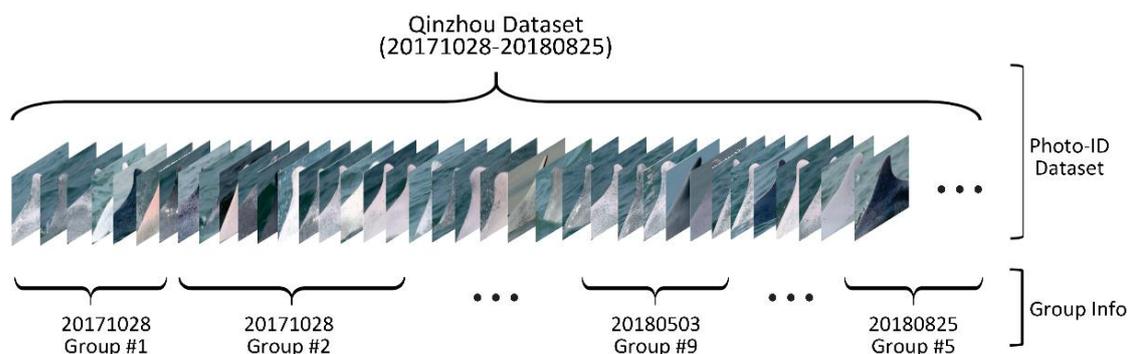


图 4-1 加入群信息后的背鳍个体识别数据集（上半部分代表只有背鳍照片和对应个体编号的个体识别数据集，下半部分代表背鳍照片对应的群信息）

表 4-1 在 Qinzhou Dataset 上 InceptionV4 的背鳍个体识别

背鳍照片质量（分）	80-100	70-79	60-69	ALL
各分数照片占比（%）	49.7%	20.1%	30.2%	100%
InceptionV4 识别准确率（Top-1）	88.11%	84.16%	67.1%	81.1%

根据第三章的内容，在 Qinzhou Dataset (20171028-20180924) 的测试集上，使用效果最好的深度 CNN 网络 InceptionV4 进行背鳍个体识别，结果如表 4-1 所示，可发现照片质量较低（60-69 分）的背鳍图片占了近三分之一，并且这部分照片的准确率较低。我们尝试将个体识别与群信息结合分析，于是选取了 Qinzhou Dataset 的 20180824 的第 3 群的背鳍照片（该群共有 228 张背鳍照片，出现了 11 个不同的个体）进行背鳍个体识别，每张照片逐一输入 InceptionV4，并发现，有一些照片识别出来的个体编号并不在这一群的 11 个个体编号里面。出现这种结果的原因很简单，因为 InceptionV4 是根据 Qinzhou Dataset 的 140 个个体进行训练的，识别出错时，错误的结果有可能是 140 个个体编号中的一个。我们对这一群的出错结果进行了统计，在各个照片质量的照片中，识别出不属于该群的个体数如表 4-2 所示。可发现照片质量较低（60-69 分）的背鳍图片出错情况更严重。

表 4-2 背鳍个体识别出错结果（Qinzhou Dataset 的 20180824 的第 3 群）

背鳍照片质量（分）	80-100	70-79	60-69
InceptionV4 识别出错结果（不属于该群的个体数）	5	5	13

针对低质量照片个体识别的准确率较低、出错结果更多样的现象，我们提出一种基于群信息的背鳍个体识别方法，在普通的背鳍个体识别的基础上，有效地利用群信息，从而提高低质量照片的识别准确率，以及整体的识别准确率。该方法输入的是一群背鳍照片，包括三步：CNN 预测并提取特征，异常检测，和异常捞回，如图 4-2 所示。第一步是将一群照片逐张输入到训练好的 CNN 进行识别，得到预测个体编号和每张照片的 CNN 特征；第二步是通过异常检测算法，根据一群照片的预测个体编号和 CNN 特征，检测出识别错误的照片。第三步是通过 kNN 算法，将检测出来的异常捞回，给异常照片新的个体编号。该方法在 Qinzhou Dataset (20171028-20180825) 的测试集上，相比于普通的背鳍个体识别，将 60-69 分照片的识别准确率提高了 7.35%，将整体识别准确率提高了 2.63%。



图 4-2 基于群信息的背鳍个体识别方法整体流程图

## 4.2 相关工作

我们的方法包括异常检测和异常捞回。异常检测是一个成熟的领域，这里我们主要测试了两种异常检测的方法： $3\sigma$ 原则和 DBSCAN 算法。异常捞回，借鉴了目前主流的人脸识别框架，使用了其中的一个常用算法，kNN 算法。

异常检测(也称离群点检测)是指通过与大多数数据显著不同的方式来识别罕见的项目、事件或观测结果。异常检测主要应用于入侵检测(计算机系统)、欺诈检测(信用卡系统、电信系统、保险索赔)、医疗或公共健康的异常检测、工业损伤检测(系统故障、结构缺陷)和图像处理。异常检测的输出结果有两类：分数或概率，显示某个项可能为异常项的程度；二值结果，给出某个项是否为异常项。异常检测技术有三大类：无监督异常检测技术：假设数据集中大多数的实例是正常的，寻找少数剩余的实例。监督异常检测技术，需要一个训练数据集并带有“正常”和“不正常”的标签，通过训练一个分类器进行检测，其中关键点在于数据具有不平衡的特性。半监督异常检测技术，从给定的正常训练数据集中构造一个表示正常实例的模型，然后利用该模型对测试实例进行处理，输出该测试实例为异常的可能性(概率)<sup>[80]</sup>。

$3\sigma$ 法则<sup>[81]</sup>，是一种基于统计概率模型的异常检测，属于无监督异常检测。在统计学中， $3\sigma$ 法则，也被称为 68-95-99.7 法则或经验法则，用来记住分布在均值附近一个范围内(68.27%、95.45%和 99.73%)的数值的百分比，68.27%、95.45%和 99.73% 的值分别位于均值的一个、两个和三个标准差内。给定一个一维的数据集  $X =$

$\{x_1, \dots, x_n\}$ , 假设 $X$ 满足高斯分布, 均值为 $\mu$ , 标准差为 $\sigma$ , 如果某项 $x \notin (\mu - 3\sigma, \mu + 3\sigma)$ , 该项将被视为异常值, 如图 4-3 所示。另外, 该法则也指出, 即使对于非正态分布的变量, 至少 88.8%的情况应该属于适当计算的  $3\sigma$  区间, 这由契比雪夫不等式推出。

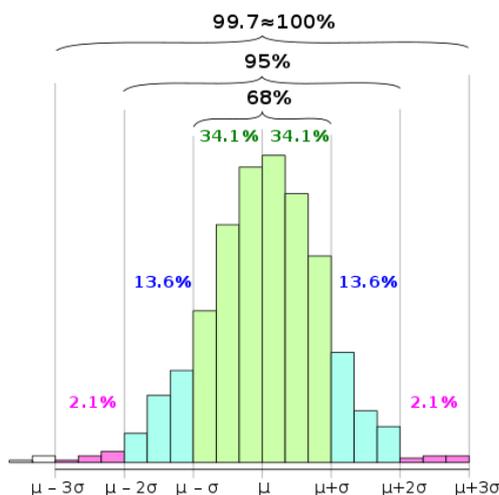


图 4-3  $3\sigma$  法则<sup>[81]</sup> (左侧标记为白色的数据为异常值)

DBSCAN 算法<sup>[82]</sup> 是一个基于密度的聚类算法, 同时也可以作为一种常用的基于密度的异常检测算法。直观地看, DBSCAN 算法可以找到样本点的全部密集区域, 并把这些密集区域当做一个个聚类簇。由于是基于密度进行聚类, 它具有以下主要特点: 不用确定聚类簇的数量, 聚类簇的形状任意, 能检测出噪声。DBSCAN 的算法实现步骤如图 4-4 所示, 在得到聚类簇的同时, 它也得到噪声点, 即异常值。DBSCAN 在计算样本点的密度时, 需要计算样本点之间的距离, 一般使用欧式距离。另外, 它有两个关键参数需要设置, 邻域距离  $\epsilon$  和邻域最小样本个数  $MinPts$ , 这两个参数刻画了什么叫密集。当邻域距离  $\epsilon$  内的点的个数大于最小样本个数  $MinPts$  时, 就是密集。

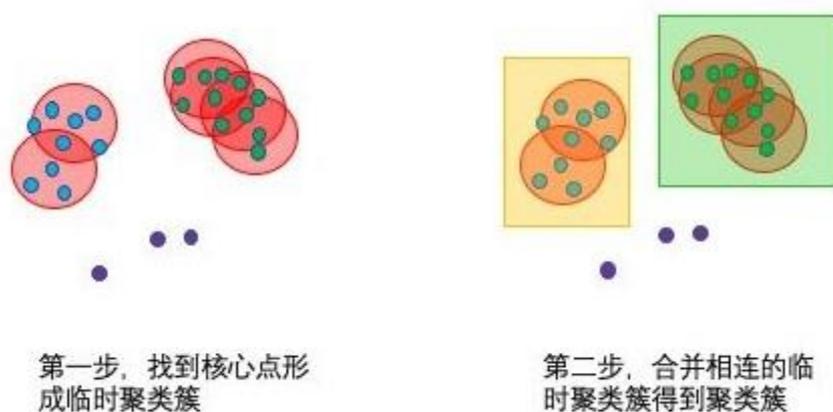


图 4-4 DBSCAN 算法实现步骤

人脸识别主要分为 1:1 任务和 1:N 任务<sup>[83,84]</sup>。1:1 任务是指判断两张人脸图片是否属于同一个人。1:N 任务是指给出一张人脸图片所属的个体编号(ID)。而 1:N 任务又分为封闭集人脸识别和开放集识别,如图 4-5 所示,二者主要的区别在于,前者,待处理图片的所属个体是训练集(Training Set)的所有个体中的一个,后者,待处理图片的所属个体并不在训练集的所有个体中,训练集只是用于训练特征提取器(Feature Extractor),而对比相册(Testing Gallery)中图片所属个体是已知的,待处理图片需要与对比相册中的图片进行特征上的比对,从而得到所属个体。目前,主流的开放集人脸识别算法,在特征提取器后,都是通过 kNN 算法<sup>[85]</sup>得到,通过计算特征间的欧式距离或余弦距离,得到待处理图片的个体编号。kNN 算法的原理如图 4-6 所示,用距离测试样本最近的 k 个样本,取出现次数最多的样本标签,作为测试样本的标签。

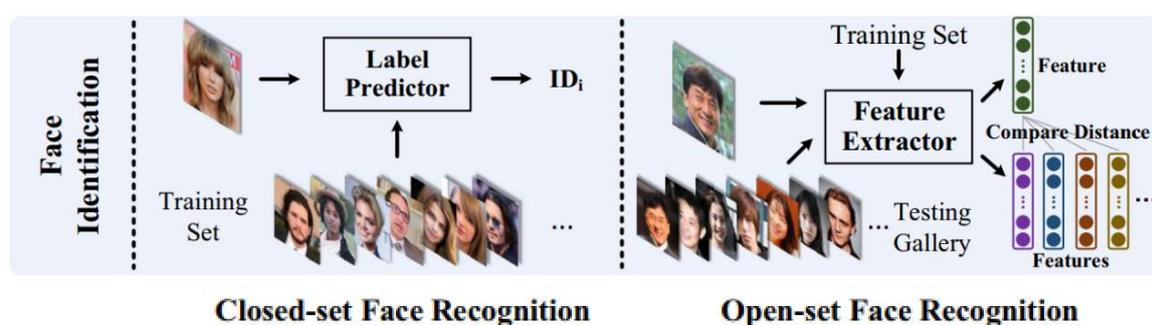


图 4-5 人脸识别(1:N)的流程图<sup>[83]</sup>

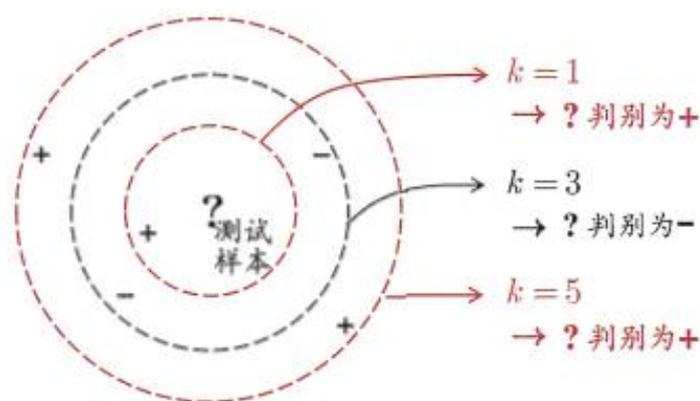


图 4-6 kNN 算法原理<sup>[85]</sup>

### 4.3 方法

这里将介绍我们的基于群信息的背鳍个体识别方法,包括三步: CNN 预测并提取特征,异常检测,和异常捞回。该方法的输入是已经定位截取的待识别的背鳍照片,以及对应的群信息,输出是每张背鳍照片对应的个体编号。

### 4.3.1 CNN 预测并提取特征

根据第三章，普通的背鳍个体识别需要一个训练集，其中包括已经定位截取的高于 60 分的背鳍照片以及对应的个体编号，通过训练一个深度 CNN 网络，实现对新的背鳍照片的识别。这个深度 CNN 网络，每次输入的是一张背鳍照片。

基于群信息的背鳍个体识别方法将使用上面训练好的深度 CNN 网络，区别在于，在测试时将根据群信息将整个测试集进行划分，将属于同一个群的背鳍照片划分在一起，测试时，输入一群照片，利用训练好的深度 CNN 网络逐一处理照片，输出这一群照片的预测个体编号。另外，我们在处理每张照片时，需要提取最后一层 CNN 特征，输出这一群照片的 CNN 特征矩阵。预测个体编号和 CNN 特征矩阵将用于后续的处理。

根据第三章，我们选取了效果最好的深度 CNN 网络 InceptionV4 进行普通的背鳍个体识别，根据 InceptionV4 的网络结构，如图 4-7 所示，选取 Average Pooling 层作为特征，其大小为 1536。假设输入的一群照片有 N 张，输出的预测个体编号为  $N \times 1$  的矩阵，CNN 特征矩阵大小为  $N \times 1536$ 。

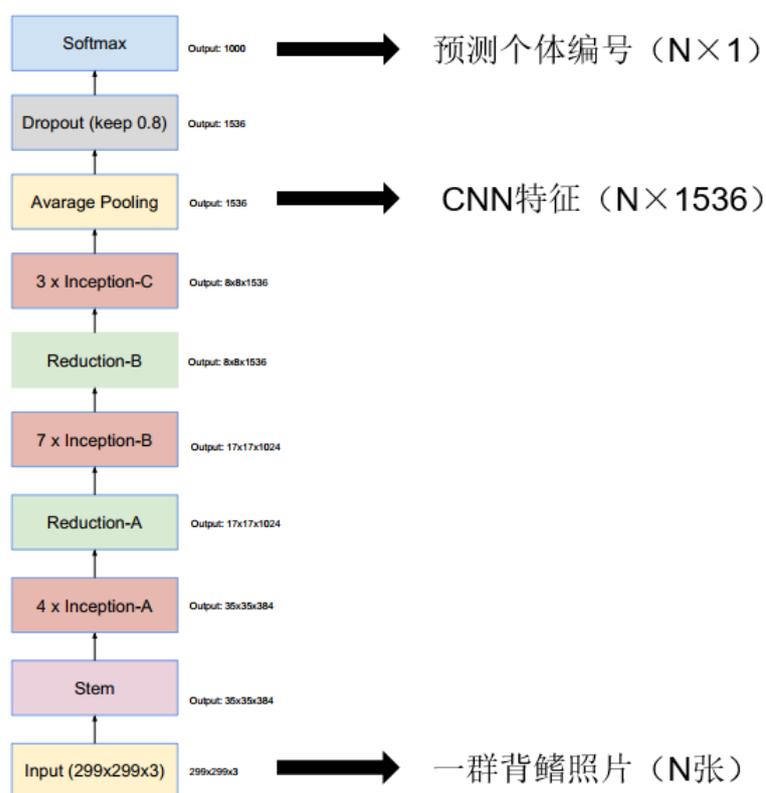


图 4-7 InceptionV4 的输入与输出

### 4.3.2 异常检测

在得到一群背鳍照片的预测个体编号和 CNN 特征矩阵后，将基于这些结果进行异常检测。在这里分别测试了两种方法，统计法和 DBSCAN 算法。在后面的实验部

分，将讨论异常检测的方法的选择。

基于统计法的异常检测，参考了  $3\sigma$  法则，是根据一群背鳍照片的预测个体编号，进行异常检测，每张照片输出一个二值结果。这里，我们借助数据集的其中一群背鳍照片进行说明，Qinzhou Dataset 的 20180824 的第 2 群。该群共有 166 张背鳍照片，来自 4 个个体，个体编号为“A0033”、“A0048”、“A0056”、“C0003”。算法流程如图 4-8 所示，输入的是一群背鳍照片的预测个体编号（也称预测 ID），该列输入中每格代表一张照片的预测 ID，相同颜色代表相同 ID。方法主要有三步：统计每个预测 ID 的样本数（频数），画出频数分布直方图；将频数分布直方图按各个预测 ID 的频数从小到大从左侧开始排序；设置异常样本的占比 THR，根据经验设置为 10%，左侧区间 ( $\leq 10\%$ ) 为异常样本，右侧区间 ( $> 10\%$ ) 为正常样本。在图 4-8 的输出中，直方图的横轴表示所有预测 ID，纵轴表示各个预测 ID 的样本数（频数），虚线左侧区间的预测 ID 对应的每张背鳍照片为异常样本（标记为红色），虚线右侧区间的为正常样本（标记为绿色）。

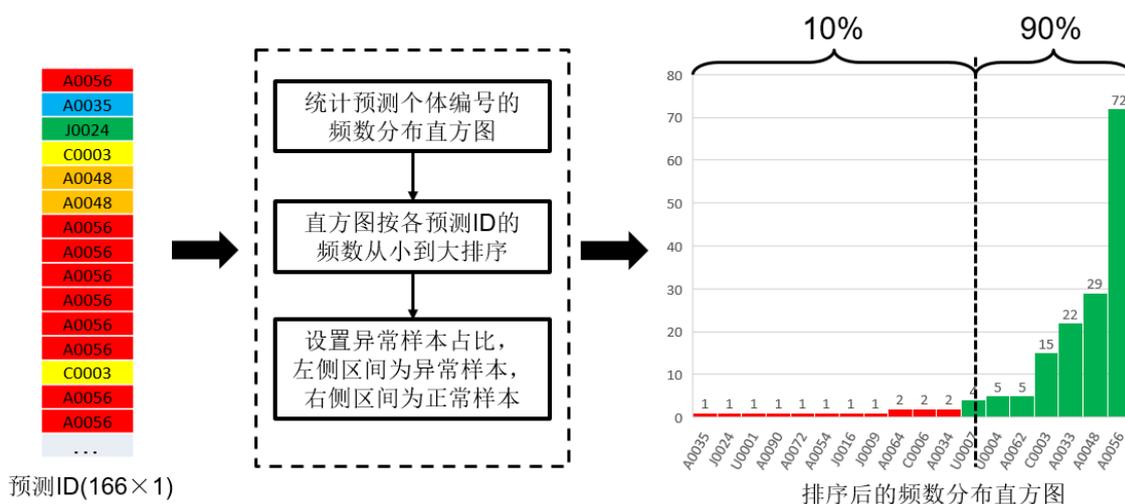


图 4-8 基于统计法的异常检测流程图

基于 DBSCAN 算法的异常检测，是根据一群背鳍照片的 CNN 特征矩阵，进行异常检测，每张照片输出一个二值结果。如图 4-9 所示，输入的是一群背鳍照片的 CNN 特征矩阵。方法主要有三步：对每张背鳍照片的 CNN 特征进行 L2 标准化；设置 DBSCAN 算法的参数，距离计算公式选择了欧式距离，根据观察，邻域距离  $\epsilon$  设为 0.7，邻域最小样本个数 MinPts 设为 5；将标准化后的 CNN 特征矩阵作为 DBSCAN 算法的输入。DBSCAN 的输出为群内每张背鳍照片（每个 CNN 特征）对应的簇（cluster）编号（标为正常），以及不属于任何簇的噪声样本（标为异常）。

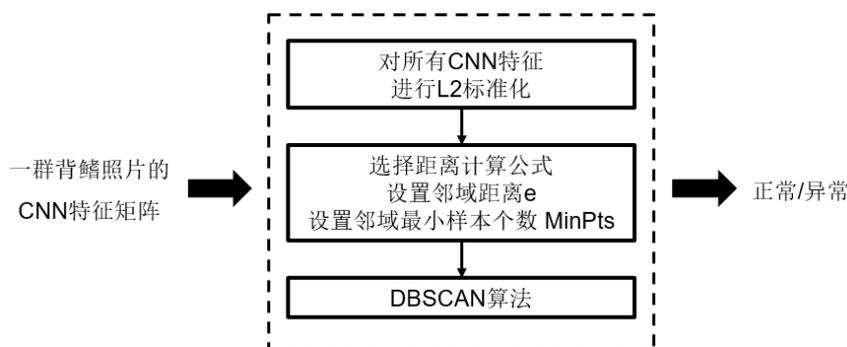


图 4-9 基于 DBSCAN 算法的异常检测流程图

### 4.3.3 异常捞回

异常捞回，参考了开放集人脸识别，根据一群背鳍照片的 CNN 特征矩阵和异常检测的结果，给每张异常照片分配一个新的个体编号。如图 4-10 所示，输入的是一群背鳍照片的 CNN 特征矩阵，以及异常检测中的正常照片的预测个体编号。方法主要有三步：对每张背鳍照片的 CNN 特征进行 L2 标准化；设置 kNN 算法的参数，距离计算公式选择了欧式距离，为了避免噪声对结果的影响，取稍微大的 k 值，根据经验，k 值设为 8；将异常照片作为测试样本，正常照片作为训练样本，使用照片的 CNN 特征进行距离计算，通过 kNN 算法，输出异常照片的新个体编号。另外，在进行 kNN 计算时，如果正常照片的预测个体编号与当前处理的异常照片的预测个体编号相同时，这个预测个体编号是可疑的，忽略这部分正常照片。

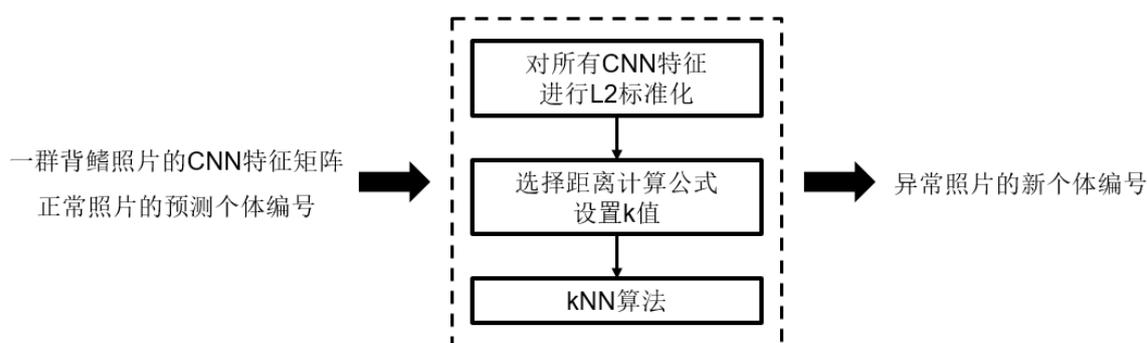


图 4-10 异常捞回流程图

## 4.4 实验与结果

### 4.4.1 数据集

根据第二章，背鳍个体识别有两个地区的数据集 Qinzhou Dataset 和 Shantou Dataset，其中只有 Qinzhou Dataset 的部分照片（20171028 - 20180825）带有群信息。在这一部分照片的测试集中，群信息如图 4-1 所示。该测试集共有 5465 张背鳍照片，共有 99 个群，各个群的群内背鳍照片数量是不平衡的，其数量分布如图 4-11 所示。

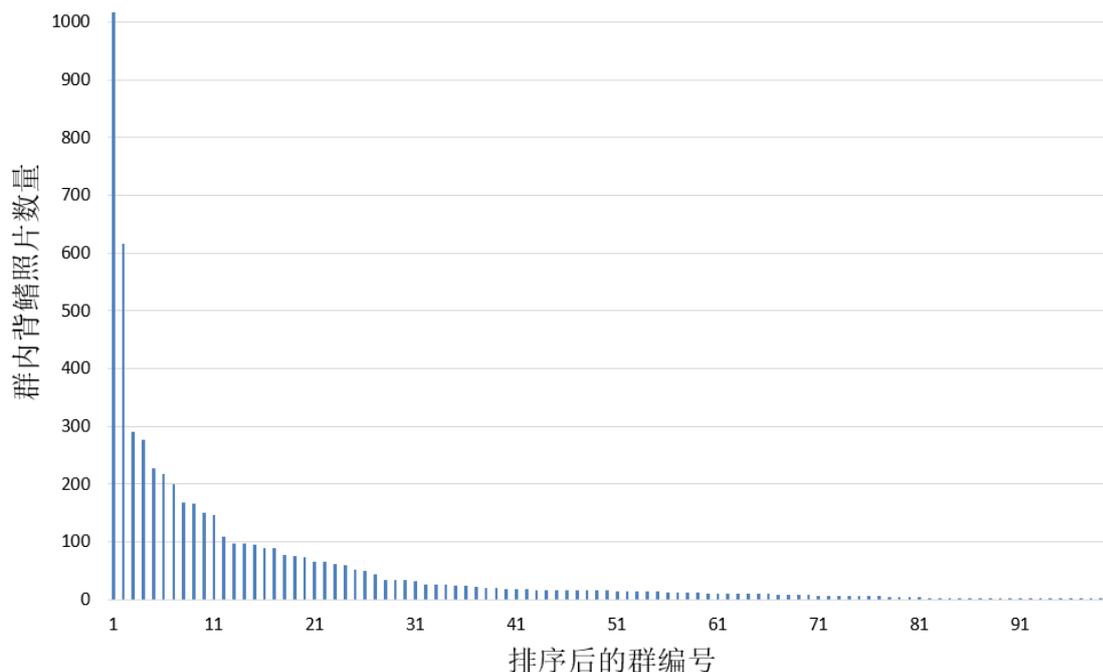


图 4-11 各个群的群内背鳍照片数量

表 4-3 各个指标的含义及其所反映的异常检测算法对于整个方法的影响 (DT\_ID 表示 CNN 直接预测的个体编号, GT\_ID 表示背鳍照片的真实个体编号, AD 表示异常检测的结果, 1 表示正常, 0 表示异常)

评价指标	含义	影响
TP (True Positive)	DT_ID=GT_ID and AD = 1	异常检测正确, 但不会影响最终准确率。
FN (False Negative)	DT_ID=GT_ID and AD = 0	异常检测错误, 直接导致最终结果错误, 使最终准确率下降。
FP (False Positive)	DT_ID≠GT_ID and AD = 1	异常检测错误, 表示异常检测的遗漏, 不会影响最终准确率。
TN (True Negative)	DT_ID≠GT_ID and AD = 0	异常检测正确, 最终准确率还取决于异常捞回的效果。

#### 4.4.2 评价指标

对于异常检测部分的算法, 我们通过 TP、TN、FP、FN 这几个指标进行评价。我们在表 4-3 中给出了各个指标的含义, 以及其所反映的异常检测算法对于整个方法的影响, 从而帮助我们选择更好的异常检测算法。原本基于 CNN 的背鳍分体识别就会得到一个测试集的准确率, 这异常检测和异常捞回是在它的基础上进行结果的修正,

从而改变最终准确率。对于异常检测结果为正常的背鳍照片，我们的方法不会改变原本 CNN 的预测结果，因此，关键的是异常检测结果为异常的背鳍照片，对应的是 FN 和 TN 指标。FN 表示异常检测算法错误地将原本 CNN 预测正确的结果预测为异常，这部分背鳍照片无法通过异常捞回正确识别，必然会使最终准确率下降；TN 表示有效的异常检测结果，最终准确率还取决于在这部分背鳍照片上的异常捞回的效果。在选择异常检测的算法时，应保证  $TN-FN > 0$ ，并且，TN 与 FN 的差值越大越好。

对于整个基于群信息的背鳍个体识别方法，我们结合异常照片的新预测 ID 和正常照片的 CNN 预测 ID，并将其与每张背鳍照片的真实 ID 进行比较，得到整个方法的最终准确率。

### 4.4.3 结果

异常检测的结果如表 4-4 所示，TN-FN 的结果越大，异常检测的效果越好。可见统计法效果更好，而 DBSCAN 即使经过仔细的参数选择也无法得到很好的结果。另外，对于统计法，我们测试了不同的异常区间 THR 的取值，得到的 TN-FN 结果如图 4-12 所示，可见区间范围的选取在 7%-14% 上得到的结果很接近，在这个范围内对参数选择不敏感。

在背鳍个体识别上，我们的方法的识别准确率如表 4-5 所示，相比于普通的 InceptionV4，在质量较低（60-69 分）的背鳍照片上，识别准确率提升较大，提升了 7.35%，整体准确率提升了 2.63%。另外，我们给出了该方法在其中一群（Qinzhou Dataset 的 20180824 的第 2 群）背鳍照片上，对比于普通的 InceptionV4 的识别结果的示例图，如图 4-13 所示。

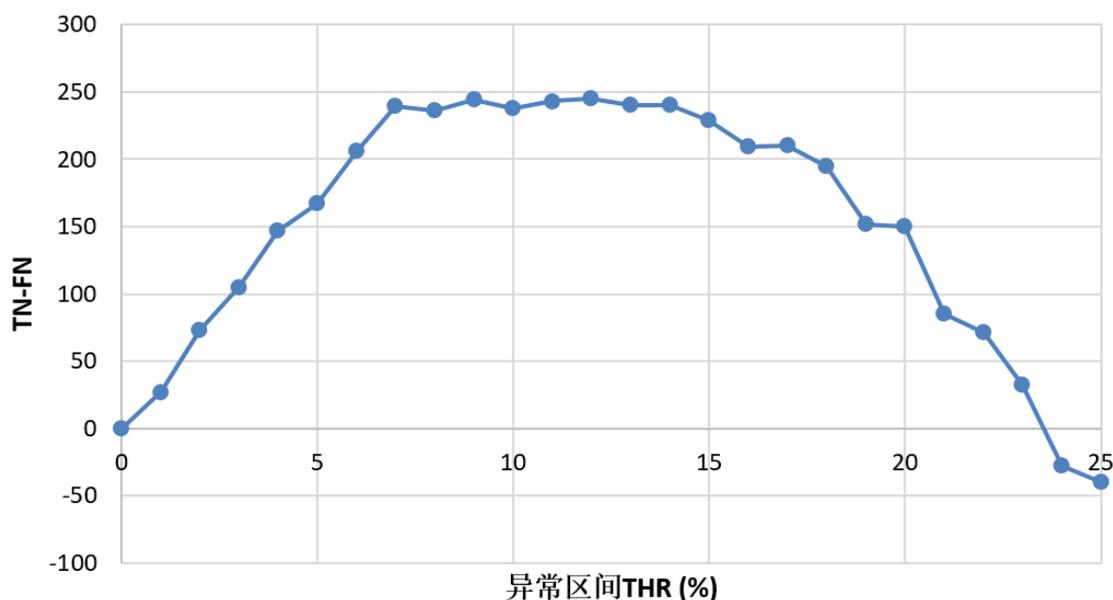


图 4-12 异常区间参数 (THR) 选取对异常检测结果 (TN-FN) 的影响

表 4-4 异常检测算法的结果

	TP	FN	FP	TN	TN-FN
统计法 (THR=10%)	4652	98	379	336	238
DBSCAN (e=0.7, MinPts=5)	4640	110	584	131	21

表 4-5 基于群信息的背鳍个体识别准确率

背鳍照片质量 (分)	80-100	70-79	60-69	ALL
各分数照片占比 (%)	49.7%	20.1%	30.2%	100%
InceptionV4	95.57%	88.97%	73.19%	86.92%
Ours (InceptionV4+统计法)	95.33%	90.71%	80.54%	89.55%

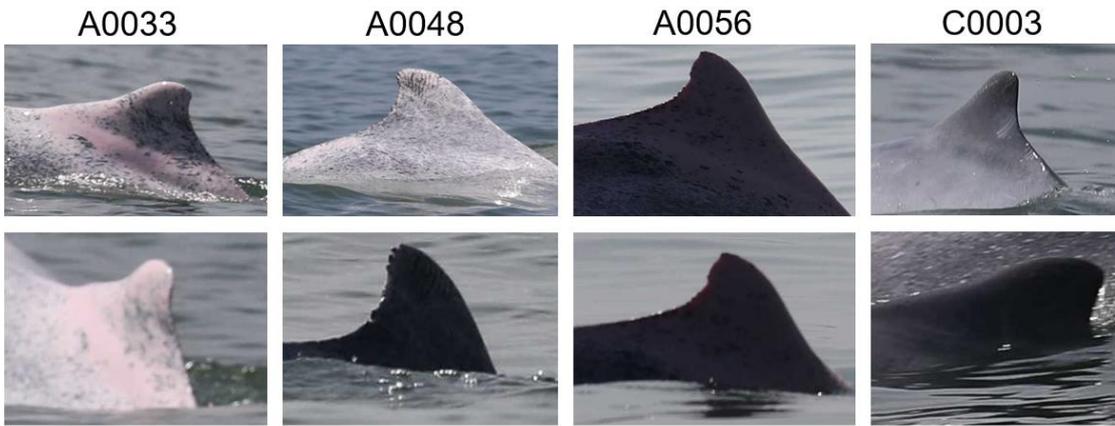


图 4-13 基于群信息的背鳍个体识别结果示例图 (结果取自 Qinzhou Dataset 的 20180824 的第 2 群, 该群共有 4 个个体, 最上方的是个体编号, 第一行是对应各个个体的 80-100 分的照片, 第二行是对应各个个体的 60-69 分的照片。在第二行的照片上, InceptionV4 识别出错, 但我们的方法识别正确。)

## 第 5 章 中华白海豚识别软件

### 5.1 软件总体流程

该软件系统基于开源的深度学习框架 Tensorflow 以及跨平台图形用户界面开发框架 PyQt，前者通常用于各种计算机视觉和自然语言处理任务，通过搭建各类卷积神经网络，能够对目标识别和图像分类处理有出色表现，可以实现对输入的中华白海豚照片进行背鳍定位和背鳍识别，后者是基于 Python 编程环境语言进行界面设计的工具，它可以很好地与 Tensorflow 结合进行软件开发，将基于 Linux 服务器下学习到的深度神经网络模型应用于 Windows 客户端，并展示背鳍图像的识别结果。

另外，在实际使用过程中，需要区分背鳍的左右侧，以方便后续的人工校验（人工进行个体识别结果的校验，需要对比同一侧背鳍的照片），因此软件加入了背鳍左右侧识别。并且，该软件主要用于背鳍个体识别，并不需要用到背鳍分割的结果，该软件没有加入背鳍分割。在实际的背鳍个体识别中，识别前会输入群信息，因此软件加入的是基于群信息的背鳍个体识别。

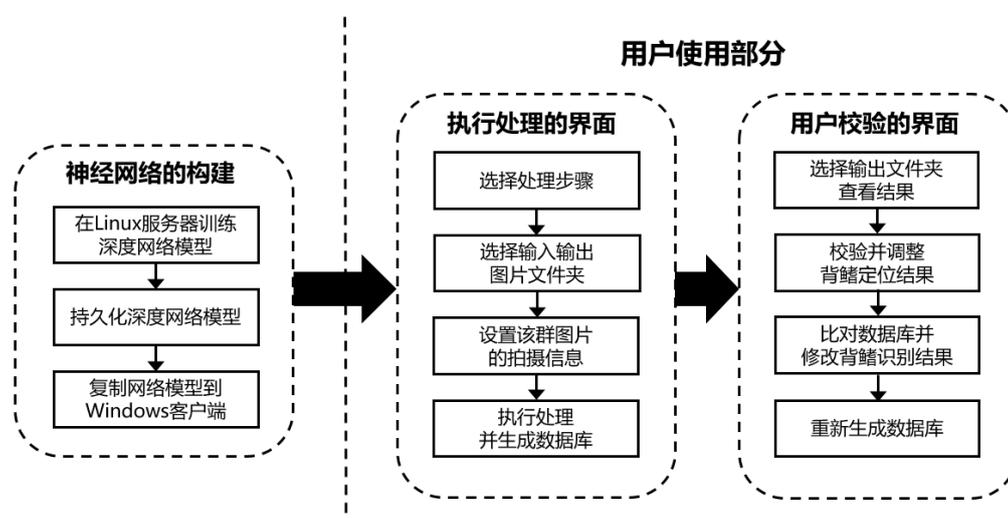


图 5-1 中华白海豚识别软件的总体流程图

该软件系统的总体流程图如图 5-1 所示，其中具体的步骤如下：

1) 用户首先利用 Linux 服务器，分别训练背鳍定位、背鳍左右侧识别、背鳍质量识别、背鳍个体识别这四个深度神经网络模型，再将模型持久化，然后将深度神经网络模型复制到 Windows 客户端上；

2) 用户在 Windows 系统上安装该中华白海豚识别软件，选择硬盘中某一天拍摄的某一群中华白海豚照片，调用复制过来的深度神经网络模型，对这一批照片进行批量处理。

3) 背鳍定位和识别的图片和结果将显示在系统界面上,用户可以重新框选背鳍区域、与数据库再进行人工比对、校验并修改的识别结果,校验后的结果将重新生成数据库。

## 5.2 神经网络的构建

神经网络的构建基于 Tensorflow 平台,其中用于背鳍定位的神经网络源码来自 Tensorflow Object Detection API<sup>[42]</sup>,网络模型为 Faster-RCNN 网络结构,主干网络采用 Resnet101,采用从 COCO 数据集预训练的模型开始训练;用于背鳍左右侧识别、背鳍质量识别、背鳍个体识别的神经网络源码来自 TensorFlow-Slim<sup>[43]</sup>,网络模型为 InceptionV4,采用从 ImageNet 数据集预训练的模型开始训练。

神经网络的训练在 Linux 服务器下完成,该服务器搭载 TitanXP (12GB) 显卡,能够有效提高训练及调试的效率。为了将训练完成的模型的结构和参数保存到一个文件中,并且输出预测结果时无须其他依赖文件,我们需要将训练完成的模型进行持久化。然后将持久化完成的网络模型复制到运行 Windows 系统的电脑里,该电脑只需要搭载显存大于等于 2GB 的显卡,就能利用该神经网络模型对新的图片进行检测。

## 5.3 执行处理的界面

在软件的执行处理界面,如图 5-2 所示,左侧从上到下,最上面是设置原图文件夹、设置背鳍定位截图文件夹、设置最终输出文件夹,中间是设置该群图片的相关信息,如日期、相机型号、拍摄人员、拍摄地区、群编号,最下面是设置将要执行的步骤,包括背鳍定位、背鳍左右侧识别、背鳍质量识别、背鳍个体识别、输出结果的保存等。用户在执行处理前需要先完成以下输入操作:选择处理步骤;选择输入输出图片文件夹;设置该群图片的拍摄信息。

在软件的执行处理界面的中间,如图 5-4 所示,从上到下,上面是执行的按钮,当用户完成输入操作后可点击 Run 开始执行相应步骤,下面则会显示任务处理的进度。

生成的数据库文件以及输出的图片文件夹都会保存在上面设置的输出文件夹中,输出的文件如图 5-5 所示。处理的结果会生成数据库并自动写入 Database.csv 文件中,输出的图片则分类保存在各个文件夹中,背鳍定位的截图保存在 crop 文件夹,背鳍定位整体效果图保存在 overview 文件夹,根据识别的背鳍质量和个体重新构造的文件夹是 output 文件夹,其他文件则是执行处理步骤的中间结果,可供用户查看但不直接展示。

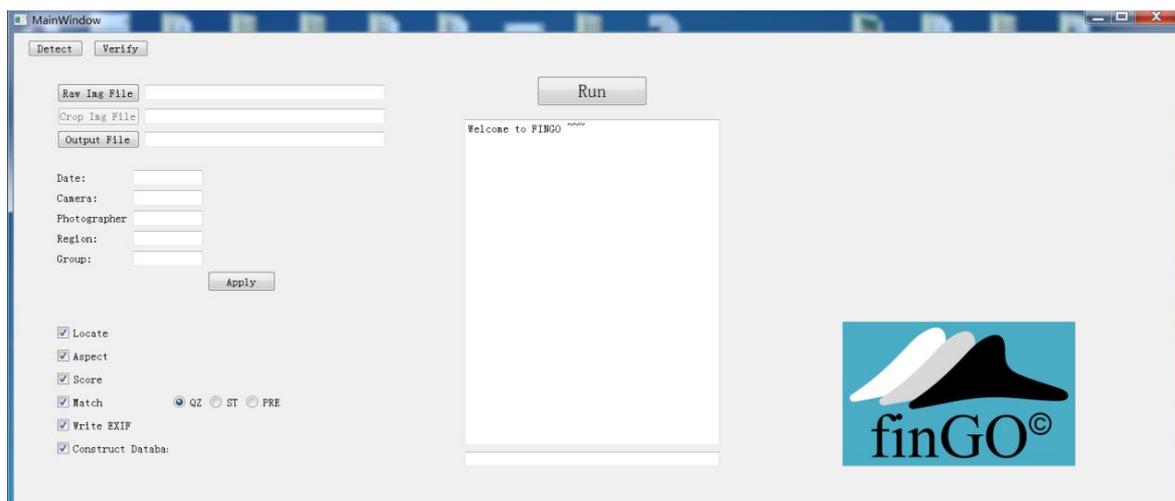


图 5-2 中华白海豚识别软件的执行处理界面

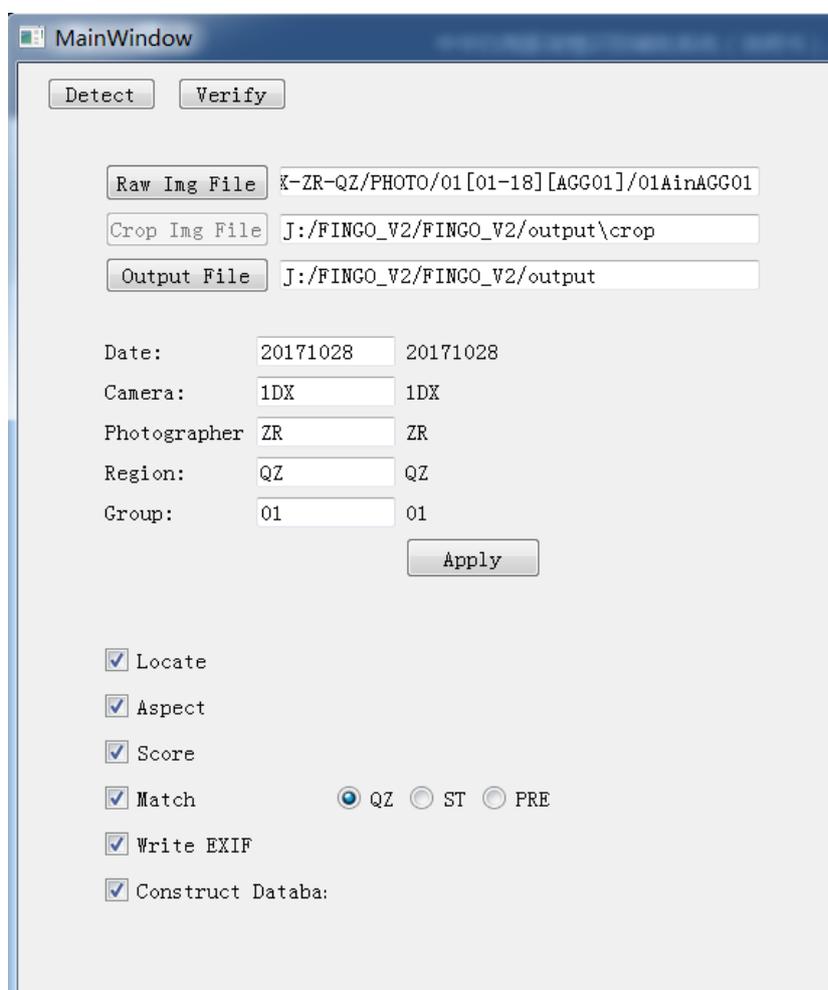


图 5-3 输入参数设置模块

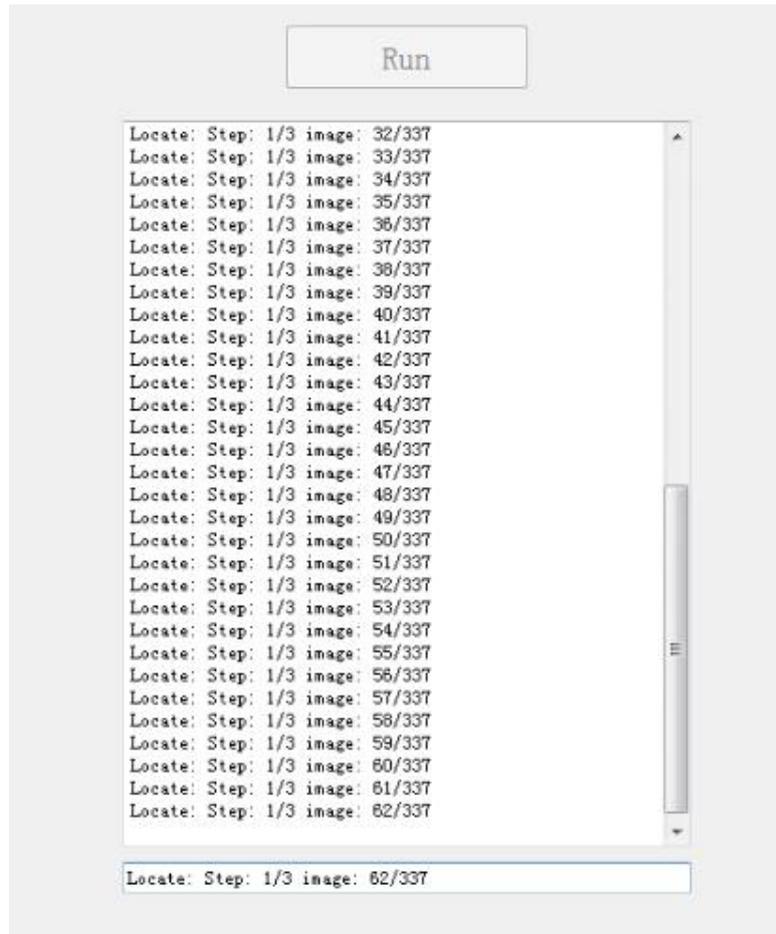


图 5-4 执行进度显示模块

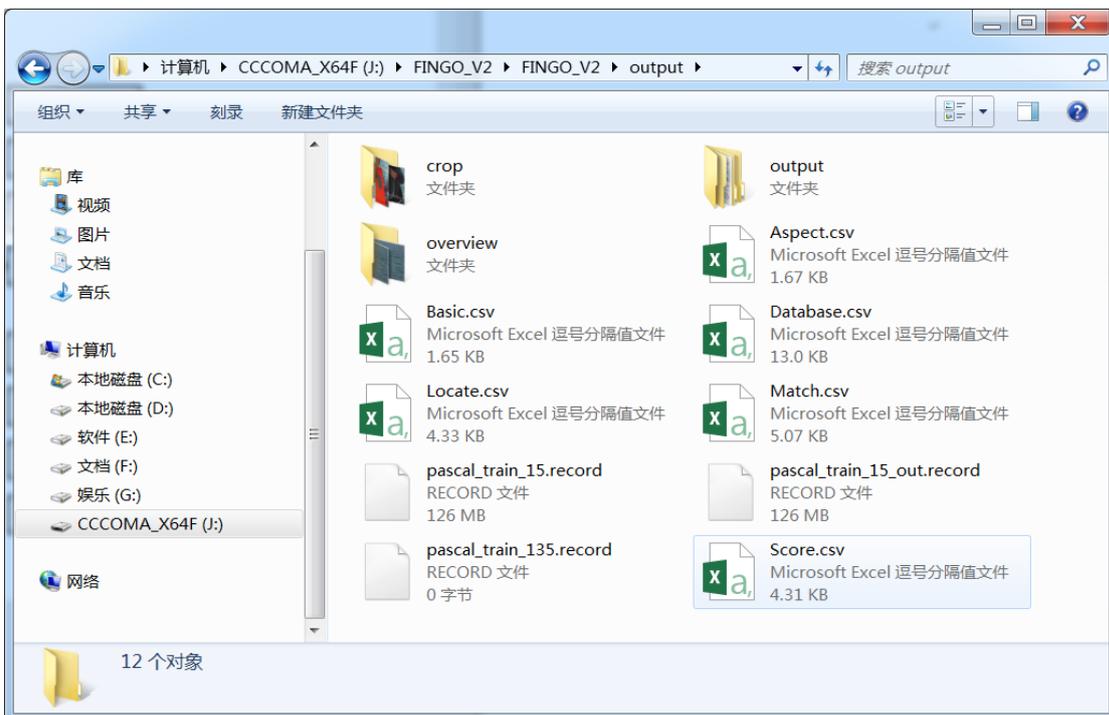


图 5-5 输出文件示例图

## 5.4 用户校验的界面

如图 5-6 所示，展示了处理结果显示的界面，左上方为输出图片的文件夹结构，左下方为背鳍定位的结果，中间上方为当前定位的背鳍，中间下方为当前背鳍的识别结果，包括：背鳍左右侧结果、背鳍质量识别结果和背鳍个体识别结果，右边为数据库中与前背鳍最匹配的三个个体的图片。用户可以快速地翻看软件定位出来的背鳍图片，并查看其左右侧结果、质量识别结果、背鳍个体识别。如果定位的结果出错或遗漏，可以通过界面左下方的按钮进入修改背鳍定位结果的界面，如图 5-7 所示，重新手动框选正确的背鳍区域。如果识别的结果出错，可以在界面的中间下方进行修改。

这个中华白海豚识别软件得到的结果，可以很方便地转换成其他数据格式，导入到其他数据分析平台中，帮助生态学家高效地处理照片，能够让他们专注于户外研究中。并且，该软件的标准处理流程也能够方便地推广到不同的团队，让更多的研究者使用，促进不同研究团队的合作。

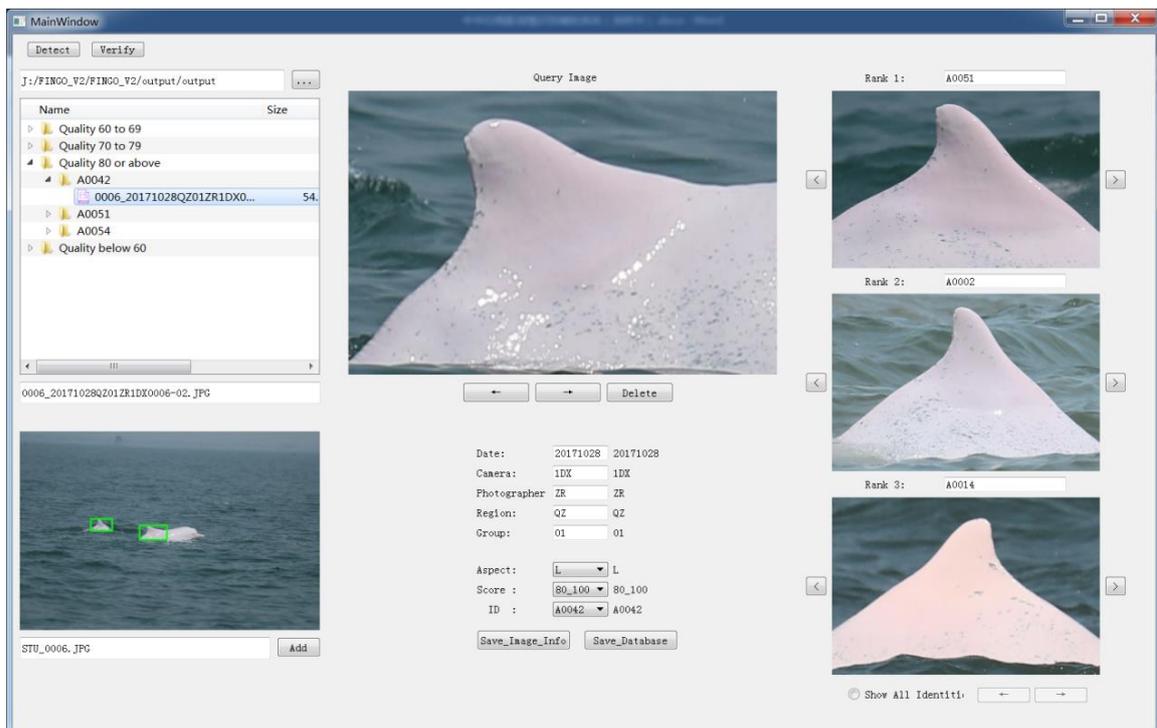


图 5-6 检测结果显示及用户校验的界面

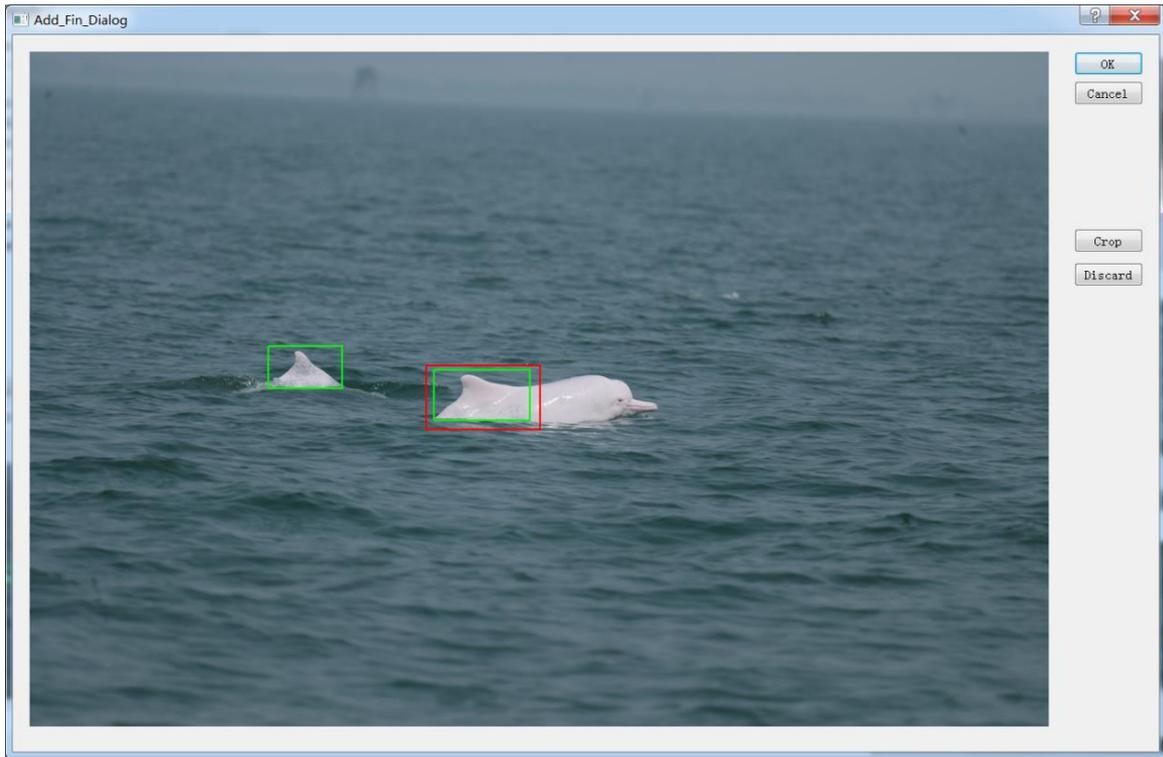


图 5-7 用户修改背鳍定位结果的界面

## 第 6 章 总结与展望

### 6.1 本文总结

长期的照片个体识别还可以让我们对动物种群的人口统计数据、生活史、运动方式和社会结构有更深入细微的了解。中华白海豚在其成长过程中有着独特多样的颜色和斑点，是照片个体识别的理想物种。并且与其他鲸豚类动物相比，中华白海豚在地理上的可及性、生物学特性、生态功能以及文化地位，使得它具有更高的研究和社会关注度。为了提高对中华白海豚这一濒危海洋动物的保育研究的效益，一方面，我们需要提高中华白海豚照片个体识别的效率，让海洋生态学家有更多时间进行户外研究，另一方面，我们需要实现数据处理的标准化，促进不同团队间的合作。

针对自动化的中华白海豚照片个体识别这一实际问题，本文首次将深度学习的图像处理技术应用于中华白海豚的照片个体识别。首先，我们提出了一个用于中华白海豚的照片个体识别的处理流程，并基于该流程整理、公开了首个用于中华白海豚的照片个体识别的大规模图像数据集。然后，在该数据库上测试了主流的深度卷积神经网络，并对比了不同网络的结果。进一步地，提出了一种基于群信息的背鳍个体识别方法，针对现有方法的不足，有效利用了拍摄时的先验信息，提高了个体识别的整体准确率。最后，设计了一个在 Windows 系统上的应用软件，辅助生态学家完成中华白海豚的照片个体识别。

### 6.2 未来工作展望

本文在以下几点仍存在局限性，可作进一步改进：

- 1) 目前基于 CNN 图像分类的照片质量识别准确率不够高；
- 2) 本文提出的基于群信息的背鳍个体识别方法，在异常检测阶段使用的统计法只用到 CNN 的预测 ID 而没有用到 CNN 特征，而 DBSCAN 算法只用到 CNN 特征效果不太好，未来可考虑融合 CNN 预测 ID 和特征进行异常检测；
- 3) 本文的背鳍个体识别仍然是封闭集个体识别，未来可考虑开放集个体识别。

## 参考文献

- [1] Emery L, Wydoski R. Marking and tagging of aquatic animals: an indexed bibliography[J]. US Department of the Interior, Fish and Wildlife Service, 1987, 165: 0-57.
- [2] Wiirsig B, Jefferson T A. Methods of photo-identification for small cetaceans[J]. Reports of the International Whaling Commission, 1990, 12: 43-52.
- [3] Karczmarski L, Cockcroft V G. Matrix photo-identification technique applied in studies of free-ranging bottlenose and humpback dolphins[J]. Aquatic Mammals, 1998, 24(3): 143-147.
- [4] Stevick P T, Palsbøll P J, Smith T D, Bravington M V, Hammond P S. Errors in identification using natural markings: rates, sources, and effects on capture recapture estimates of abundance[J]. Canadian Journal of Fisheries and Aquatic Sciences, 2001, 58(9): 1861-1870.
- [5] Mazor T, Possingham H P, Kark S, et al. Collaboration among countries in marine conservation can achieve substantial efficiencies[J]. Diversity and Distributions, 2013, 19(11): 1380-1393.
- [6] Jefferson T A, Karczmarski L. *Sousa chinensis*[J]. Mammalian species, 2001, 2001(655): 1-9.
- [7] Davidson A D, Boyer A G, Kim H, et al. Drivers and hotspots of extinction risk in marine mammals[J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(9): 3395-3400.
- [8] Huang S, Karczmarski L, Chen J, et al. Demography and population trends of the largest population of Indo-Pacific humpback dolphins[J]. Biological Conservation, 2012, 147(1): 234-242.
- [9] Kumar S, Singh S K, Singh R, Singh A K. Animal Biometrics: Concepts and Recent Application[M]. Animal Biometrics, Springer, Singapore, 2017: 1-20.
- [10] Halpin P N, Read A J, Fujioka E, et al. OBIS-SEAMAP: The World Data Center for Marine Mammal, Sea Bird, and Sea Turtle Distributions[J]. Oceanography, 2009, 22(2): 104-115.
- [11] Joly A, Goeau H, Glotin H, et al. LifeCLEF 2016: Multimedia Life Species Identification Challenges[C]. cross language evaluation forum, Évora, Portugal, Sep, 2016: 286-310.

- [12] Kaggle. Humpback Whale Identification Challenge. Website, 2018. <https://www.kaggle.com/c/whale-categorization-playground>.
- [13] Kaggle. Right Whale Recognition. Website, 2016. <https://www.kaggle.com/c/noaa-right-whale-recognition>.
- [14] Jefferson T A. Population biology of the indo-pacific hump-backed dolphin in Hong Kong waters[J]. Wildlife Monographs, 2000: 1-65.
- [15] Gailey G, Karczmarski L. DISCOVERY: A photo-identification data management system for individually recognizable animals[J]. The Swire Institute of Marine Science, The University of Hong Kong, 2012.
- [16] Stanley R. Darwin: Identifying dolphins from dorsal fin images[D]. Senior Thesis, Eckerd College, 1995.
- [17] Hillman G R, Wursig B, Gailey G A, et al. Computer-assisted photo-identification of individual marine vertebrates: A multi-species system[J]. aquatic mammals, 2003, 29(1):117-123.
- [18] Reijns. I3S: Interactive Individual Identification System. Website, 2020. <https://www.reijns.com/i3s/>
- [19] Chapple T K, Jorgensen S J, Anderson S D, et al. A first estimate of white shark, *Carcharodon carcharias*, abundance off Central California[J]. Biology Letters, 2011, 7(4): 581-583.
- [20] Towner A V, Wcisel M A, Reisinger R R, et al. Gauging the Threat: The First Population Estimate for White Sharks in South Africa Using Photo Identification and Automated Software[J]. PLOS ONE, 2013, 8(6).
- [21] Van Hoey N E. Photo-identification and distribution of bottlenose dolphins[J]. *Tursiops truncatus*, 2013: 2006-2009.
- [22] Mortensen E N, Barrett W A. Interactive segmentation with Intelligent Scissors[J]. Graphical Models and Image Processing, 1998, 60(5): 349-384.
- [23] Baird R W, Gorgone A M, Mcsweeney D J, et al. False Killer Whales ( *Pseudorca crassidens* ) Around the Main Hawaiian Islands: Long-Term Site Fidelity, Inter-Island Movements, and Association Patterns[J]. Marine Mammal Science, 2008, 24(3): 591-612.
- [24] Baird R W, Gorgone A M, McSweeney D J, et al. Population structure of island-associated dolphins: Evidence from photo-identification of common bottlenose dolphins (*Tursiops truncatus*) in the main Hawaiian Islands[J]. Marine Mammal Science, 2009, 25(2): 251-274.

- [25] Stewman J, Debure K R, Hale S A, et al. Iterative 3-d pose correction and content-based image retrieval for dorsal fin recognition[C]. international conference on image analysis and recognition, Póvoa de Varzim, Portugal, July, 2006: 648-660.
- [26] Bogucki R, Cygan M, Khan C B, et al. Applying deep learning to right whale photo identification[J]. Conservation Biology, 2019, 33(3): 676-684.
- [27] Kuhl H S, Burghardt T. Animal biometrics: quantifying and detecting phenotypic appearance[J]. Trends in Ecology and Evolution, 2013, 28(7): 432-441.
- [28] Markowitz T M, Harlin A D, Wursig B, et al. Digital photography improves efficiency of individual dolphin identification. Marine Mammal Science[J]. Marine Mammal Science, 2003, 19(1): 217-223.
- [29] Hughes B, Burghardt T. Automated Visual Fin Identification of Individual Great White Sharks[J]. International Journal of Computer Vision, 2017, 122(3): 542-557.
- [30] Weideman H J, Jablons Z M, Holmberg J, et al. Integral Curvature Representation and Matching Algorithms for Identification of Dolphins and Whales[C]. international conference on computer vision, Venice, Italy, Oct, 2017: 2831-2839.
- [31] Arzoumanian Z, Holmberg J, Norman B, et al. An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*[J]. Journal of Applied Ecology, 2005, 42(6): 999-1011.
- [32] Van Tienhoven A M, Den Hartog J E, Reijns R A, et al. A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*[J]. Journal of Applied Ecology, 2007, 44(2): 273-280.
- [33] Town C, Marshall A, Sethasathien N, et al. Manta Matcher: automated photographic identification of manta rays using keypoint features[J]. Ecology and Evolution, 2013, 3(7): 1902-1914.
- [34] Krizhevsky A, Sutskever I, Hinton G E, et al. ImageNet Classification with Deep Convolutional Neural Networks[C]. neural information processing systems, Lake Tahoe, Nevada, USA, Dec, 2012: 1097-1105.
- [35] Kabani A, Elsakka M R. Improving Right Whale recognition by fine-tuning alignment and using wide localization network[C]. canadian conference on electrical and computer engineering, Windsor, Ontario, Canada, Apr, 2017: 1-6.
- [36] Kabani A, Elsakka M R. North atlantic right whale localization and recognition using very deep and leaky neural network[J]. Math. Appl., 2016, 5(2), 155-170.
- [37] Polzounov A, Terpigova I, Skiparis D, et al. Right whale recognition using convolutional neural networks[J]. arXiv: Computer Vision and Pattern Recognition, 2016.

- [38] Nepovinnikh E. Saimaa ringed seal fur pattern extraction for identification purposes[J]. 2017.
- [39] Brust C, Burghardt T, Groenenberg M, et al. Towards Automated Visual Monitoring of Individual Gorillas in the Wild[C]. international conference on computer vision, Venice, Italy, Oct, 2017: 2820-2830.
- [40] Li S, Li J, Lin W, et al. Amur Tiger Re-identification in the Wild[J]. arXiv: Computer Vision and Pattern Recognition, 2019.
- [41] Korschens M, Barz B, Denzler J, et al. Towards Automatic Identification of Elephants in the Wild[J]. arXiv: Computer Vision and Pattern Recognition, 2018.
- [42] Huang J, Rathod V, Sun C, et al. Tensorflow Object Detection API. Website, 2020. [https://github.com/tensorflow/models/tree/master/research/object\\_detection](https://github.com/tensorflow/models/tree/master/research/object_detection)
- [43] Sergio G. TensorFlow-Slim image classification model library. Website, 2020. <https://github.com/tensorflow/models/tree/master/research/slim>
- [44] Abhishek D, Ankush G, Andrew Z. VGG Image Annotator. Website, 2020. <http://www.robots.ox.ac.uk/~vgg/software/via/>
- [45] Yang S, Luo P, Loy C C, et al. WIDER FACE: A Face Detection Benchmark[C]. computer vision and pattern recognition, Las Vegas, NV, USA, June, 2016: 5525-5533.
- [46] Everingham M, Van Gool L, Williams C K, et al. The Pascal Visual Object Classes (VOC) Challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [47] Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [48] Hinton G E, Salakhutdinov R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [49] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553), 436-444.
- [50] Liu L, Ouyang W, Wang X, et al. Deep Learning for Generic Object Detection: A Survey[J]. International Journal of Computer Vision, 2020, 128(2): 261-318.
- [51] Chen K, Wang J, Pang J, et al. MMDetection: Open MMLab Detection Toolbox and Benchmark.[J]. arXiv: Computer Vision and Pattern Recognition, 2019.
- [52] Lin T, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context[C]. european conference on computer vision, Zurich, Switzerland, Sep, 2014: 740-755.

- [53] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]. neural information processing systems, Montreal, Canada, Dec, 2015: 91-99.
- [54] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [55] Girshick R. Fast R-CNN[C]. international conference on computer vision, Santiago, Chile, Dec, 2015: 1440-1448.
- [56] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]. computer vision and pattern recognition, Columbus, Ohio, USA, June, 2014.
- [57] Dai J, Li Y, He K, et al. R-FCN: Object Detection via Region-based Fully Convolutional Networks[J]. arXiv: Computer Vision and Pattern Recognition, 2016.
- [58] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]. european conference on computer vision, Amsterdam, The Netherlands, Oct, 2016: 21-37.
- [59] Redmon J, Divvala S K, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]. computer vision and pattern recognition, Las Vegas, NV, USA, June, 2016: 779-788.
- [60] Hariharan B, Arbelaez P, Girshick R, et al. Object Instance Segmentation and Fine-Grained Localization Using Hypercolumns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 627-639.
- [61] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[J]. arXiv: Computer Vision and Pattern Recognition, 2018.
- [62] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]. computer vision and pattern recognition, Hawaii, USA, July, 2017: 6517-6525.
- [63] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. computer vision and pattern recognition, Boston, Massachusetts, USA, June, 2015: 1-9.
- [64] He K, Zhang X, Ren S, et al. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification[C]. international conference on computer vision, Santiago, Chile, Dec, 2015: 1026-1034.
- [65] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. computer vision and pattern recognition, Las Vegas, NV, USA, June, 2016: 770-778.

- [66] Howard A, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[J]. arXiv: Computer Vision and Pattern Recognition, 2017.
- [67] Niu X, Suen C Y. A novel hybrid CNN-SVM classifier for recognizing handwritten digits[J]. Pattern Recognition, 2012, 45(4): 1318-1325.
- [68] Krizhevsky A, Sutskever I, Hinton G E, et al. ImageNet Classification with Deep Convolutional Neural Networks[C]. neural information processing systems, Lake Tahoe, Nevada, USA, Dec, 2012: 1097-1105.
- [69] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]. computer vision and pattern recognition, Miami, Florida, USA, June, 2009: 248-255.
- [70] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[C]. national conference on artificial intelligence, Phoenix, Arizona, USA, Feb, 2016: 4278-4284.
- [71] Lecun Y, Boser B E, Denker J S, et al. Handwritten Digit Recognition with a Back-Propagation Network[C]. neural information processing systems, Denver, CO, USA, Nov, 1989: 396-404.
- [72] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [73] Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks[J]. Pattern Recognition, 2018: 354-377.
- [74] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. computer vision and pattern recognition, Boston, Massachusetts, USA, June, 2015: 1-9.
- [75] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]. computer vision and pattern recognition, Las Vegas, NV, USA, June, 2016: 2818-2826.
- [76] Long J, Shelhamer E, Darrell T, et al. Fully convolutional networks for semantic segmentation[C]. computer vision and pattern recognition, Boston, Massachusetts, USA, June, 2015: 3431-3440.
- [77] Chen L, Papandreou G, Schroff F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[J]. arXiv: Computer Vision and Pattern Recognition, 2017.
- [78] Chen L, Papandreou G, Kokkinos I, et al. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs[J]. arXiv: Computer Vision and Pattern Recognition, 2014.

- [79] Chen L, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [80] Chandola V, Banerjee A, Kumar V, et al. Anomaly detection: A survey[J]. ACM Computing Surveys, 2009, 41(3).
- [81] Wikipedia. 68–95–99.7 rule. Website, 2020. [https://en.wikipedia.org/wiki/68%E2%80%9C95%E2%80%9C99.7\\_rule](https://en.wikipedia.org/wiki/68%E2%80%9C95%E2%80%9C99.7_rule)
- [82] Ester M, Kriegel H, Sander J, et al. A density-based algorithm for discovering clusters in large spatial Databases with Noise[C]. knowledge discovery and data mining, Portland, Oregon, USA, Aug, 1996: 226-231.
- [83] Liu W, Wen Y, Yu Z, et al. SphereFace: Deep Hypersphere Embedding for Face Recognition[C]. computer vision and pattern recognition, Hawaii, USA, July, 2017: 6738-6746.
- [84] Wang M, Deng W. Deep Face Recognition: A Survey[J]. arXiv: Computer Vision and Pattern Recognition, 2018.
- [85] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

## 致 谢

时光飞逝，三年的硕士学习生活即将结束，在此，我想真诚地感谢那些曾经帮助我的人。

我要感谢我的导师范衡教授，他学识渊博、待人和蔼，治学严谨科研风格和求真务实的工作作风，是我终身学习的榜样。他给我们提供一个完善的科研平台，并通过循循教导提高了我们的科研能力，让我在思辨和表达的能力上有所提高。另外，即使在疫情期间，也关心着我们的学习和生活情况，他给予我们的支持和信任让我们的学习和生活更顺利。

感谢海洋研究所的郑锐强博士后，虽然研究的领域不一样，但是他对科研充满了热情，在与他合作的过程中能从他身上学到很多知识。另外，他给了我各种学习的机会，比如出海拍摄中华白海豚、参加学术研讨会、撰写基金申报书等，拓宽了我的知识面，让我对学术研究的整个流程有了切身的体会。他为人处事积极主动且乐观，不仅是良师，也是益友。

感谢实验室师兄卢杰威，在与他一同完成实验和论文的过程中，在他身上学习到了很多，比如研究领域知识细节、实验设计的思路、论文撰写的技巧等。另外，在与他一同参加雅思考试的过程中，也得到了他很大的帮助，从他身上学习到了时间管理的能力和认真踏实的作风。

最后感谢我的父母和家人，感谢他们对我的言传身教和多年细心培养，并时刻支持着我的生活和学业，让我在学习上没有顾虑，是我学习的最大动力。

冯靖安

2020年5月

## 攻读硕士学位期间主要的工作成果

### 一、论文

- [1] Jiewei Lu, Zhun Fan, Ce Zheng, Jingan Feng, Longtao Huang, Wenji Li, Erik D. Goodman. Automated Strabismus Detection based on Deep neural networks for Telemedicine Applications. <https://arxiv.org/abs/1809.02940>
- [2] Zhun Fan, Jiewei Lu, Jingan Feng, Ruiqiang Zheng. PICWD: A Large-Scale Dataset for Automated Photo Identification of Chinese White Dolphins.
- [3] Zhun Fan, Jingan Feng, Ruiqiang Zheng, et al. Automated Photo Identification of Chinese White Dolphins Based on Group Information and Anomaly Detection.

### 二、专利

- [1] 冯靖安, 郑锐强, 范衡, 等. 一种基于卷积神经网络的中华白海豚背鳍识别方法. 发明. 申请号: 201911128813.8

### 三、软件著作权

- [1] 中华白海豚背鳍识别系统 (FINGO), 登记号: 2019SR1176922