

文章编号: 1674-5566(2025)02-0307-13

DOI: 10.12024/jsou.20240604587

## 基于深度学习的金枪鱼延绳钓渔获图像识别技术分析

夏超<sup>1</sup>, 陈新军<sup>1,2,3,4</sup>, 刘必林<sup>1,2,3,4</sup>, 孔祥洪<sup>1,3</sup>, 叶旭昌<sup>1,3</sup>

(1. 上海海洋大学 海洋生物资源与管理学院, 上海 201306; 2. 大洋渔业资源可持续开发教育部重点实验室, 上海 201306; 3. 国家远洋渔业工程技术研究中心, 上海 201306; 4. 农业农村部大洋渔业可持续利用重点实验室, 上海 201306)

**摘要:** 为了实现金枪鱼延绳钓渔获物的高效识别和分类, 改善渔业资源监测的准确性, 本研究探索了基于卷积神经网络的鱼类图像识别方法。采用上海海洋大学“淞航”号在中西太平洋公海调查中捕获的3种经济性鱼类及10种兼捕渔获物的图像数据, 运用单发多箱探测器(Single shot multibox detector, SSD)卷积神经网络进行图像分类识别。通过将局部鱼类图像与整体图片数据集进行比较分析, 优化训练数据集, 以提升模型的分类性能。结果表明: 改进后的鱼类图像数据集在SSD模型上的分类准确率达91.6%, 相较于原始数据集提高了6.2%。利用优化后的数据集, SSD模型能够显著提高金枪鱼延绳钓渔获物的识别准确性, 具备更好的稳定性和适应性。本研究为基于卷积神经网络的渔业资源监测提供了有效的技术路径, 尤其在提升金枪鱼延绳钓渔获物自动分类识别精度方面展现了广泛的应用潜力, 对于促进可持续渔业管理和海洋生态保护具有重要意义。

**关键词:** 金枪鱼延绳钓; 渔获物识别; 卷积神经网络; 数据集优化; 中西太平洋

中图分类号: S 977 文献标志码: A

中西太平洋是世界上重要的渔业生产海域, 渔业资源丰富, 特别是金枪鱼种类繁多。金枪鱼延绳钓是该区域内主要的作业方式之一<sup>[1-3]</sup>。金枪鱼渔业受到区域性国际渔业组织的管理, 而渔获物的识别是其中最基础的研究内容之一, 因而受到国际社会的高度关注<sup>[4]</sup>。

在渔业管理中渔获物识别的早期阶段, 船员或观察员主要采用形态学鉴别的方法, 通过观察和比较鱼的外部形态特征, 如鳞片、鳍、体色、体型等来识别不同的鱼种。随着计算机视觉领域卷积神经网络技术的成熟, 国内外学者开始对鱼类图像识别开展相关研究, 以期更为深入地研究鱼类鉴别的方法<sup>[5-8]</sup>。RUM等<sup>[9]</sup>采用VGG16神经网络对马来西亚8种不同种类的淡水鱼图像进行训练分析, 模型的准确率仅达到了60%~80%。ROBILLARD等<sup>[10]</sup>采用改进的Resnet-34架构和

U-Net方法, 对ImageNet数据集中亚马逊河的33属鱼类图像进行预处理, 通过将背景像素归零(变黑)保留鱼类图像来进行训练, 平均准确率达到了97.9%。

前人对卷积神经网络在鱼类识别中的应用研究往往侧重于通过改进模型网络和数据集增强的方式进行。然而, 这些研究通常未结合鱼类自身的形态学特性进行深入分析与探讨。为此, 本研究提出了以下基于计算机技术的鱼种识别优化与改进方法。

基于单次多框检测器(Single shot multibox detector)模型算法, 通过构建中西太平洋金枪鱼延绳钓渔获物图像数据集, 对渔获物图像进行整体与局部的划分, 构建了一套科学合理的中西太平洋金枪鱼延绳钓渔获物鱼种识别技术。研究利用中西太平洋金枪鱼延绳钓捕捞所得的鱼类

收稿日期: 2024-06-28 修回日期: 2024-11-28

基金项目: 上海市高校特聘教授“东方学者”岗位跟踪计划项目(GZ2022011); 农业农村部全球渔业资源调查监测评估(公海渔业资源综合科学调查)专项(D-8025-23-1002)

作者简介: 夏超(1997—), 男, 硕士研究生, 研究方向为人工智能渔业学。E-mail: 694502705@qq.com

通信作者: 陈新军, E-mail: xjchen@shou.edu.cn

版权所有 ©《上海海洋大学学报》编辑部(CC BY-NC-ND 4.0)

Copyright © Editorial Office of Journal of Shanghai Ocean University (CC BY-NC-ND 4.0)

<http://www.shhydxxb.com>

图像,这些图像包括整体(鱼身)和部分(鱼头、中段鱼鳍、鱼尾),实现了对3种经济性金枪鱼鱼种和10种金枪鱼延绳钓兼捕渔获物鱼种的高准确度识别,并提出了相关的技术体系与标准。本研究不仅提高了渔获物种类识别的效率和准确率,还为渔业资源的科学管理和可持续利用提供了坚实的技术支持,对海洋生物监测、生态系统保护也具有重要意义。

## 1 材料与方法

### 1.1 材料来源

实验数据来源于上海海洋大学“淞航”号2023年在中西太平洋公海开展渔业资源调查期间金枪鱼延绳钓所得渔获物图像信息(图1),共收集到4375张13个鱼种的图片作为数据集,并按一定比例进行分层采样,将数据集分为训练集、验证集和测试集,其中训练集占70%(3063张),验证集和测试集各占15%(656张),如表1所示。

### 1.2 研究方法

#### 1.2.1 研究方法与模型构建

本实验使用单发多箱探测器(Single shot

multibox detector, SSD)卷积神经网络,该算法是经典且高效的多目标检测主流算法,广泛应用于多个领域。图像识别过程如下:将待检测图像裁剪为300×300像素,从左至右输入图像数据;在输出特征图中的预测位置生成多个不同大小的默认框,以匹配图像中可能存在的物体;进行物种类型分类处理,并通过卷积预测进行判断,从而实现目标检测的功能。

SSD卷积神经网络是在VGG-16卷积网络的基础上改进的<sup>[11-12]</sup>,该网络直接对整个图像数据进行扫描检测,使用多组金字塔特征图来提取边界框和预测目标类别。网络中每一层的图像特征信息和预测值都是由前一层卷积层经过多次检测和验证提供的。其中,VGG16的fc7层替换为Conv7层,与此同时,增加了多个卷积层:Conv8\_2、Conv9\_2、Conv10\_2、Conv11\_2,以此来加深图像卷积深度,从而在检测过程中获取更多的特征信息。目标检测部分结合了Conv4\_3、Conv7、Conv8\_2卷积层来实现图像目标的定位检测和识别功能。图2为SSD模型应用于中西太平洋金枪鱼延绳钓渔获物图像识别层次结构图。



图1 中西太平洋金枪鱼延绳钓渔获物数据集

Fig. 1 Central and western pacific tuna longline fishing catch dataset

表1 中西太平洋金枪鱼延绳钓渔获物种类  
Tab. 1 Species of western and central Pacific tuna caught by longline fishing

种类 Class	形态学特征 Morphological characteristics	图片数量 Number of images
大眼金枪鱼 <i>Thunnus obesus</i>	体型纺锤形,光滑鳞片覆盖,眼睛大而突出,背部深蓝色,腹部银白色,尾鳍新月形,侧线延伸至尾柄中央	330
黄鳍金枪鱼 <i>Thunnus albacares</i>	体型纺锤状,背部蓝灰色,腹部银白色,侧面有暗色纵纹,背鳍尖端锯齿状,鳃盖有银白色斑块	343
长鳍金枪鱼 <i>Thunnus alalunga</i>	体型纺锤形,背部深蓝色,腹部银白色,侧面有暗色纵纹,背鳍锯齿状,鳃盖有银白色斑块	325
大青鲨 <i>Prionace glauca</i>	体型流线型,背部钢蓝色,腹部白色,头部圆锥形,眼睛大,鳃裂五对,尾鳍分叉,上下叶近乎相等	329
剑鱼 <i>Xiphias gladius</i>	体型延长,背部蓝灰色,腹部银白色,头部剑状,眼睛小,嘴部长而尖,背鳍三角形,尾鳍新月形	336
浅海长尾鲨 <i>Alopias pelagicus</i>	体型纺锤状,背部深蓝色,腹部白色,侧线明显,尾鳍极长超过身体长度,鳃裂多对	342
大眼长尾鲨 <i>Alopias superciliosus</i>	体型纺锤状,背部深蓝色,腹部白色,侧线明显,尾鳍极长超过身体长度,鳃裂多对	335
长鳍鲨 <i>Isurus paucus</i>	体型流线型,背部深蓝色,腹部白色,侧线明显,头部小,眼睛大,嘴部宽,鳃裂多对,尾鳍长	324
帆蜥鱼 <i>Alepisaurus ferox</i>	体型流线型,背部暗褐色,腹部白色,头部小圆锥形,眼睛大,嘴部弯曲半月形,背鳍三角形,尾鳍分叉	341
异鳞蛇鲭 <i>Lepidocybium flavobrunneum</i>	体型流线型,背部暗蓝色,腹部白色,皮肤光滑无鳞片,头部小,眼睛大,嘴部大而尖,鳃裂多对	342
蛇鲭 <i>Gempylus serpens</i>	体型修长侧扁,背部暗蓝色,腹部白色,皮肤光滑无鳞片,头部小,眼睛大,嘴部长而尖,鳃裂多对	325
凹尾长鳍乌鲂 <i>Taractichthys steindachneri</i>	体型侧扁,背部银灰色,腹部白色,皮肤光滑无鳞片,头部小,眼睛中等,嘴部窄,鳃裂多对	343
紫魟 <i>Pteroplatytrygon violacea</i>	体型扁平菱形,背部深紫色,腹部白色或淡黄色,头部小圆形,眼睛侧面,嘴部圆形,尾部有长尾刺	360

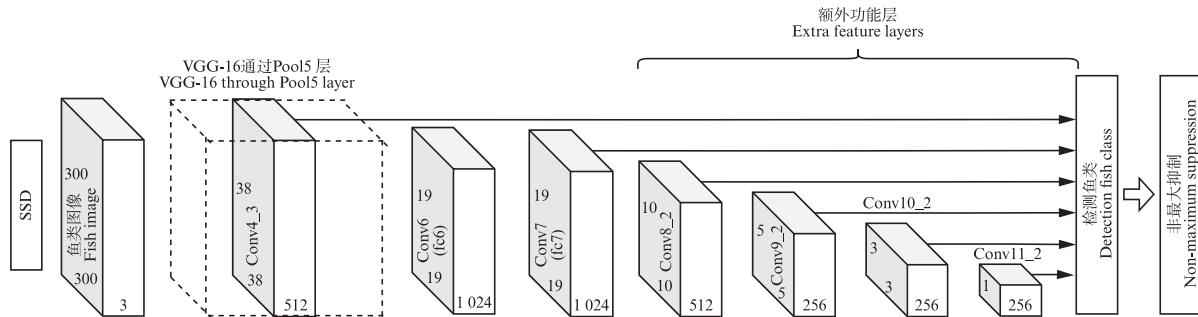


图2 SSD网络结构图  
Fig. 2 SSD network structure diagram

### 1.2.2 数据处理

SSD 卷积神经网络算法是一种监督式的深度学习框架,因此在使用图像数据集进行识别检测时,需要进行标签化注释。在图像数据预处理和标注过程中,通常使用 LabelImg 图像标记工具软件。该工具能够在图像中用矩形帧标记鱼类的位置,并为每个标记的鱼种生成 Pascal VOC 格式的文件,其中的 XML 文件包含了图像的渔获物种类、图像尺寸和矩形框的坐标信息。之后,

将所有 XML 文件转换成统一的 JSON 格式文件。JSON 文件将作为训练模型的输入数据,与原始图像数据一并输入模型进行训练。

使用工具软件创建矩形帧做标记时,LabelImg 会记录下矩形框左上角坐标( $x_{\min}, y_{\min}$ )、矩形框右下角的坐标( $x_{\max}, y_{\max}$ )以及鱼类的宽度(Width)和高度(Height)保存到 Pascal VOC 格式文件中,其中矩形框的中心点作为图像中鱼类的中心坐标( $x_{\text{in}}, y_{\text{in}}$ ),计算公式:

$$\begin{cases} x_{in} = \frac{x_{max} + x_{min}}{2} \\ y_{in} = \frac{y_{max} + y_{min}}{2} \end{cases} \quad (1)$$

然而,由于在金枪鱼延绳钓渔船的复杂捕鱼场景情况下,可能会捕捉到目标鱼类周围遮挡物或者环境信息,其中会包含有大量与目标鱼类无关的背景细节,它们会降低网络模型的训练效率和泛化能力。因此,本研究将渔获物图像划分为

鱼头、中段鱼鳍、鱼尾3部分(图3)。各鱼种局部划分,根据鱼类形态学的解剖结构和鱼体长度测量标准:鱼头部分从鱼嘴延伸至鳃盖后缘;中段鱼鳍从鳃盖后缘至肛门前缘,包括鱼的主要体躯部分;鱼尾则从肛门前缘至尾鳍末端,主要包括尾部和尾鳍。考虑到紫魟(*Pteroplatytrygon violacea*)的独特形态结构无法适用于传统的3部分划分方法,因此未对其进行局部分割,而是采用整体分析的方式处理其图像。



图3 鱼类局部分割示意图

Fig. 3 Fish part segmentation diagram

在本次实验中,共有3组数据集。其中,数据集1中包含13种渔获物整体图像,包含13个渔获物类别,共计4 375张;数据集2为通过数据集1中13种渔获物的相同图像数据依据头部、鱼身、尾部所划分出的局部数据集图像,其中包含37个类别均以渔获物鱼头、渔获物中段鱼鳍、渔获物

鱼尾的形式命名,共计4 375张;数据集3则是选取数据集2中同一鱼种渔获物3个局部部位其中识别准确度最高的特征部位作为划分依据,由此集合生成的图像数据集,共计4 375张,13个渔获物局部类别(表2)。

表2 数据集说明  
Tab. 2 Dataset description

数据集 Dataset	名称 Name	图片数 Number of images
金枪鱼延绳钓渔获物整体图像数据集 Overall image dataset of tuna longline catches	数据集1 Dataset 1	4 375
金枪鱼延绳钓渔获物局部图像数据集 Partial image dataset of tuna longline catches	数据集2 Dataset 2	4 375
金枪鱼延绳钓渔获物高识别度局部图像数据集 High recognition partial image dataset of tuna longline catches	数据集3 Dataset 3	4 375

### 1.2.3 实验配置

本研究使用 Pytorch1.10.0、Cuda 搭建渔获物鱼种识别模型,图形处理器(Graphics processing unit)为 NVIDIA GeForce RTX 3090,操作系统使用 Ubuntu 20.04 实现卷积神经网络模型的训练和测试。

首先,SSD 卷积神经网络模型将输入的图像数据集使用 OpenCV、PIL 等图像处理库,填充或裁剪成  $300 \times 300$  像素的尺寸大小,此调整有助于在保持较高检测精度的同时,提高 SSD 模型推理判断速度。接着,对调整后的图像数据集进行归一化处理,在保证图像数据与训练模型要求一致的情况下输入模型中。在 SSD 模型训练过程中使用 Adam 优化器,优化器可通过训练过程中参数的变化动态调整学习率,使得训练过程稳定性更高及收敛速度更快<sup>[13]</sup>。经过多次尝试以及 SMITH<sup>[14]</sup>提出的学习率和 batch\_size 调整方法,减少超参数调试的难度,在提高训练效率和模型性能的理论下,将初始学习率设置为 0.003,每迭代训练 10 次后,学习率减小到原来的十分之一;同时将 batch\_size 设置为 64,采用联合目标检测损失函数(Joint object detection loss function)作为模型训练的损失函数,根据 LEE 等<sup>[15]</sup>所提出用早停技术来确定合适训练轮次的理论,在各个中西太平洋金枪鱼延绳钓渔获物图像数据集上进行 100 轮次的迭代训练。

### 1.2.4 模型评价与检验

实验结果采取每轮次评估(Epoch-wise evaluation)即每个训练轮次结束后进行一次测试,以直观地观察模型在训练过程中的性能变化与实时反馈。

选择合适的最优数据集,需要使用多个客观的标准来评估模型性能,检测评估指标采用平均精度(Average precision, AP)、平均精度均值

(Mean average precision, mAP)、真实目标数量(Ground truths, gts)、检测到的目标数量(Detections, dets)、召回率(Recall)、精确率(Precision)、损失(Loss, L)和混淆矩阵对实验结果进行评价<sup>[16-17]</sup>,通过以上指标可以全面评估模型的性能,帮助了解模型的优势和不足,指导模型改进和优化工作。

AP 和 mAP 能够评估模型在不同置信度阈值下的准确率。高 AP 和 mAP 表明模型具有较高的检测准确性和稳健性<sup>[18]</sup>。除此以外,gts 和 dets 也常用于评估检测模型的识别准确性。

gts 表示真实目标的数量,即数据集中标记的目标数量,用于评估模型的覆盖范围和对真实目标的识别能力<sup>[17]</sup>。

dets 表示模型检测到的目标数量,即模型成功识别的目标数量,通过比较检测到的目标数量与真实目标数量,可以评估模型的检测能力<sup>[17]</sup>。

mAP 的计算公式:

$$x_m = \frac{\sum_{i=1}^N P_i}{N} \quad (2)$$

式中: $x_m$  为平均精度均值 mAP;  $N$  为对象的类型总数;  $P_i$  为第  $i$  类对象的精度 AP, 表示在给定召回率阈值下的平均准确率。公式描述了对每个对象类别计算的平均准确率,然后将所有类别的平均准确率取平均得到最终的 mAP。

召回率(Recall)衡量了模型成功检测到的真实目标数量与所有真实目标数量的比例。高召回率表示模型对真实目标的覆盖程度较高<sup>[19]</sup>。

召回率的计算公式:

$$x_R = \frac{x_{TP}}{x_{TP} + x_{FN}} \quad (3)$$

式中: $x_R$  为召回率;  $x_{TP}$  为真正例(True positives),即模型将正例正确预测为正例的数量;  $x_{FN}$  为假反例(False negatives),即模型将正例错误预测为反

例的数量。召回率的值在0到1之间,值越高表示模型对于正例的识别能力越强,即模型越能够将真实的正例成功地找出来。

精确率(Precision)衡量的是模型预测为正例的结果中,正确预测的比例。

精确率的计算公式:

$$x_p = \frac{x_{TP}}{x_{TP} + x_{FP}} \quad (4)$$

式中: $x_p$ 为精确率; $x_{FP}$ 为错误预测为正例的反例数量。精确率越高,说明模型预测为正例的结果中,正确的比例越高。

损失(Loss,L)指标衡量了模型在训练过程中预测值与真实标签之间的差异程度以及模型在训练集上的拟合能力的强弱。随着模型训练的进行,损失值会逐渐减小直至达到一个稳定的值或者收敛,此过程表示模型在学习过程中逐渐逼近最优解,Loss值越低,模型的预测越接近真实值,在训练集上的拟合能力越高<sup>[20]</sup>。

在SSD模型中,L通常指的是由联合目标检测损失函数(Joint object detection loss function)计算得到的,将分类损失(Classification loss,L<sub>c</sub>)和定位损失(Localization loss,L<sub>l</sub>)进行加权求和得到总损失,加权系数可以根据任务的重要性和数据集的特点来确定<sup>[11]</sup>,计算公式:

$$L = \alpha \times L_c + \beta \times L_l \quad (5)$$

式中: $\alpha$ 和 $\beta$ 是分类损失和定位损失的权重。

混淆矩阵是一个表格,用于展示模型在不同类别上的预测结果与真实标签之间的关系<sup>[20]</sup>。通过混淆矩阵可以直观地评估模型在不同类别上的表现,包括真正例、假正例、真反例和假反例等。

## 2 结果

### 2.1 金枪鱼延绳钓渔获物高识别度局部图像分布情况

为了验证本研究提出中西太平洋金枪鱼延绳钓渔获物鱼种局部数据集识别模型的有效性,在相同的软硬件环境下,对整体渔获物鱼种数据集和渔获物鱼种局部数据集在同一SSD卷积神

经网络模型中的性能进行了对比分析。保证数据集在训练超参数设置和图像预处理方式上完全一致的实验条件下开始训练,以确保结果的可靠性和公正性。

基于1.2.2节中提到的鱼类图像局部划分方法,除去紫虹,将其余12种渔获物按鱼头、中段鱼鳍、鱼尾3部分,共生成37个类别,用于训练SSD模型,训练所得模型的测试结果(表3)作为数据集3中金枪鱼延绳钓渔获物高识别度局部图像数据集的划分依据。通过分析表3中渔获物各部位的识别情况,对同种类渔获物3个鱼种局部部位的识别平均精度分析比较,综合得出金枪鱼延绳钓渔获物的高识别度局部部位(图4)。

### 2.2 模型性能比较

将图4所示的渔获物高识别度部位数据集3与渔获物整体图像数据集1在相同配置环境下进行SSD模型训练比较。

从图5中可清晰地看到,2个数据集训练的模型在测试中的准确率变化情况。在前20个轮次内,mAP准确率迅速提升,特别是高识别度局部图像数据集的mAP准确率始终高于渔获物整体图像数据集。随着训练轮次的增加,模型的mAP准确率变化幅度逐渐减小,但整体上仍表现出上升趋势。数据集1的mAP准确率在0.83到0.85之间波动,而数据集3的mAP准确率则在0.88到0.91之间波动。模型逐渐收敛后,两条准确率曲线之间存在明显的高度差,表明数据集3在金枪鱼延绳钓渔获物分类识别方面的能力优于数据集1。

图6显示了SSD模型在训练和验证过程中2数据集的识别损失变化曲线。可以看出,在前20轮内,损失快速下降,随后逐渐减缓,并在之后保持轻微波动。整个训练和验证过程中,数据集1和数据集3的损失曲线显示,局部图像数据集的损失率低于整体图像数据集。在最后40到100轮次时,两者的损失相差约0.6。这表明,渔获物高识别度局部图像数据集(数据集3)在SSD模型上具有较好的拟合和泛化能力。

表3 渔获物局部图像识别情况  
Tab. 3 The test results of local images of fishing catches

种类 Class	鱼种局部 Fish species localisation	真实目标数量 Gts	检测目标数量 Dets	召回率 Recall	精确率 Precision	平均精度 AP
大青鲨 <i>Prionace glauca</i>	鱼头	8	23	0.875	0.304	0.693
	中段鱼鳍	8	47	1.000	0.170	0.842
	鱼尾	8	46	1.000	0.174	0.959
剑鱼 <i>Xiphias gladius</i>	鱼头	10	44	0.900	0.205	0.900
	中段鱼鳍	10	54	1.000	0.167	0.900
	鱼尾	10	51	1.000	0.196	0.959
浅海长尾鲨 <i>Alopias pelagicus</i>	鱼头	2	24	1.000	0.083	1.000
	中段鱼鳍	2	25	1.000	0.080	0.950
	鱼尾	2	17	1.000	0.118	0.975
大眼长尾鲨 <i>Alopias superciliosus</i>	鱼头	7	34	1.000	0.206	0.696
	中段鱼鳍	7	42	1.000	0.167	0.875
	鱼尾	7	35	1.000	0.200	0.948
长鳍鲭鲨 <i>Isurus paucus</i>	鱼头	5	11	1.000	0.455	0.911
	中段鱼鳍	5	41	1.000	0.122	0.933
	鱼尾	5	29	1.000	0.172	1.000
帆蜥鱼 <i>Alepisaurus ferox</i>	鱼头	10	22	1.000	0.455	0.967
	中段鱼鳍	10	30	1.000	0.333	0.991
	鱼尾	10	31	1.000	0.323	0.991
大眼金枪鱼 <i>Thunnus obesus</i>	鱼头	8	30	1.000	0.267	0.966
	中段鱼鳍	8	53	1.000	0.151	0.950
	鱼尾	8	50	1.000	0.160	0.917
黄鳍金枪鱼 <i>Thunnus albacares</i>	鱼头	10	40	1.000	0.250	0.824
	中段鱼鳍	10	57	1.000	0.175	0.947
	鱼尾	10	43	1.000	0.233	0.930
长鳍金枪鱼 <i>Thunnus alalunga</i>	鱼头	12	32	1.000	0.375	0.854
	中段鱼鳍	12	48	1.000	0.250	0.906
	鱼尾	12	50	1.000	0.240	0.883
异鳞蛇鲭 <i>Lepidocybium flavobrunneum</i>	鱼头	9	24	1.000	0.375	0.966
	中段鱼鳍	9	43	1.000	0.209	0.916
	鱼尾	9	54	1.000	0.167	1.000
蛇鲭 <i>Gempylus serpens</i>	鱼头	8	20	1.000	0.400	0.884
	中段鱼鳍	8	28	0.875	0.250	0.847
	鱼尾	8	33	0.875	0.212	0.780
凹尾长鳍乌鲂 <i>Taractichthys steindachneri</i>	鱼头	3	17	1.000	0.176	0.960
	中段鱼鳍	3	25	1.000	0.120	1.000
	鱼尾	3	49	1.000	0.061	0.990
紫魟 <i>Pteroplatytrygon violacea</i>	整体	5	12	1.000	0.417	1.000
平均精度均值 mAP						0.919



图4 各渔获物高识别度局部部位  
Fig. 4 Highly recognizable local parts of each catch

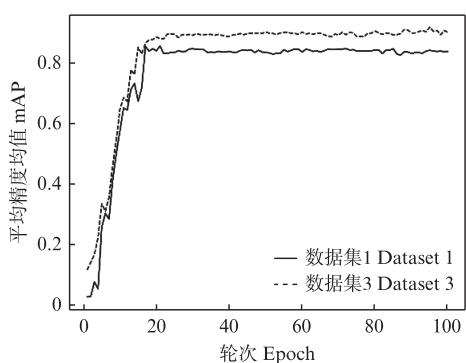


图5 不同数据集训练的SSD模型平均精度均值曲线  
Fig. 5 mAP curves of SSD models trained on different datasets

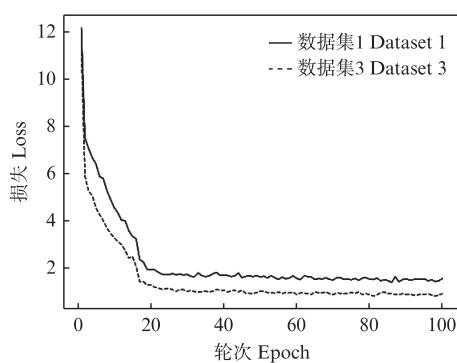


图6 不同数据集训练的SSD模型损失曲线  
Fig. 6 Loss curves of SSD models trained on different datasets

为了更方便地比较分析,图7展示了金枪鱼延绳钓渔获物高识别度局部图像数据集(数据集3)在SSD模型上对测试集各渔获物的识别情况。

图7显示,各鱼种的识别情况相对均衡,没有出现识别率过低或无法识别的情况。

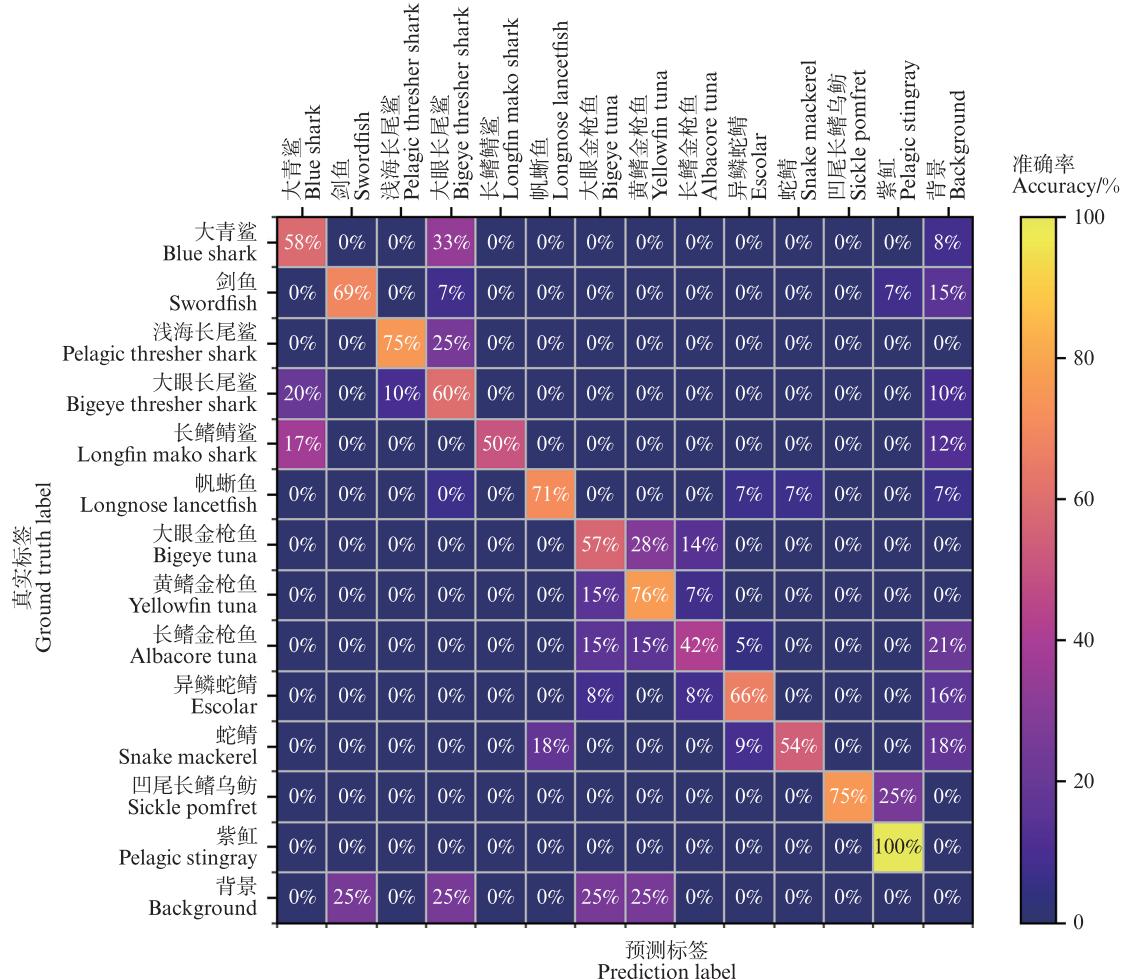


图7 数据集3训练的SSD模型混淆矩阵

Fig. 7 Confusion matrix of the SSD model trained on dataset 3

此外,3种不同数据集训练所得的SSD卷积神经网络所得的模型评估表(表4),用以帮助系统地

分析和评估模型的性能和计算资源需求,选择合适的训练集是提高模型性能和效率的关键。

表4 不同数据集实验结果对比  
Tab. 4 Comparison of experimental results across different datasets

数据集 Dataset	测试准确率 Test accuracy/%	模型训练参数 Model training parameters/( $\times 10^6$ )	浮点运算数 Floating point operations/( $\times 10^9$ )
数据集1 Dataset 1	85.4	25.35	349.36
数据集2 Dataset 2	91.9	28.27	358.94
数据集3 Dataset 3	91.6	27.89	361.42

### 3 讨论

#### 3.1 检验评估不同数据集对实验结果的影响

针对于船上作业这一复杂作业环境,快速、

准确、高效地识别出鱼种是十分重要的,不同数据集训练的实验模型性能差异显著,尤其是对于渔获物高识别度的局部图像数据集(数据集3),其性能明显优于渔获物整体图像数据集(数据集

1),见表5。这种差异主要由于图像信息内容、特征提取、搜索空间等原因形成的:局部图像标注包含了鱼类的关键部位,这些部位呈现出更丰富的细节信息,如鱼类的局部特征形状、颜色和纹理特点,这些关键特征点可以帮助模型更准确地识别鱼类的种类和特征;通过聚焦于局部特征,局部化注意力机制不仅提升了模型对关键特征的敏感度,还减少了处理整个图像的计算负担;SSD模型能够更有效地提取这些集中且显著的特征区域,只需在较小的图像范围内寻找特征,

从而显著减小搜索空间,优化图像识别流程,提高模型的准确度和效率。但是,在图7的混淆矩阵中,数据集3中长鳍金枪鱼中段鱼鳍的识别效果相较于其他渔获物部位较差。由此可以看出,模型在识别3种不同金枪鱼中段鱼鳍部位时存在一定的误判。这些部位特征非常相似,因此后续可以通过提高数据质量、增加样本数量或采用更复杂的模型结构来增强模型的区分能力,从而提升识别精度。

表5 不同数据集SSD模型性能对比  
Tab. 5 Comparison of SSD model performance for different datasets

数据集 Dataset	测试准确率 Test accuracy/%	每秒处理图像帧数 Frames per second/(帧/s)	每张图片检测时间 Detection time per image/s
数据集1 Dataset 1	85.4	29.98	0.032
数据集3 Dataset 3	91.6	31.92	0.027

深度学习方法通过组合目标物体中的多个特征点,构建一个完整的特征集合,这些特征提供了识别所需的信息,从而显著提升了渔获物鱼种分类中的模型性能。在之前的研究中,SPAMPINATO等<sup>[21]</sup>提出了一种结合三维多视图的鱼类检测和分析方法,该方法结合了鱼类的纹理和形状特征,相比传统的形态学鉴别方法具有更高的创新性。XABIER<sup>[22]</sup>在研究中使用Mask R-CNN对热带金枪鱼围网渔业的鱼类图像进行单独分割,并通过ResNet50V2进行物种分类,准确率达到了约70%。

无论是通过人为特征筛选还是运用神经网络进行鱼类图像分割,目的都是通过优化特征来提高识别准确度。然而,提升鱼种识别性能的关键改进方法主要包括以下几个方面:图像数据质量、数据预处理和模型算法改进,高分辨率且光照均匀的清晰鱼类图像数据有助于精细的数据标注,这使得模型在训练过程中能够更快、更准确地掌握特征分布情况;通过有效的数据预处理技术,可以增强特征的可辨识度减少噪声和干扰,从而提升模型的识别能力;改进模型算法中的特征提取区域,使模型更专注于具有区分度的特征,进一步提高识别准确率和效率。总之,综合考虑以上方面,不仅能提升渔获物鱼类图像分类的准确率,还能显著地提高模型的训练效率和性能,从而使卷积神经网络在渔获物鱼类识别应用中取得更优的结果。

### 3.2 模型的应用效果

本研究提出的中西太平洋金枪鱼延绳钓渔获物鱼种识别技术,通过将改进的金枪鱼延绳钓渔获物数据集应用于SSD模型,达到了91.6%的准确率,相比其他数据集取得了更好的识别效果。实验中模型训练了100个轮次,但更多的迭代次数并不一定意味着更好的性能,可能会导致过拟合问题<sup>[23-24]</sup>。更多训练参数通常提高模型灵活性和性能,但也增加资源需求和过拟合风险;更多浮点运算能捕捉细微数据模式,但增加计算复杂度和资源消耗,在资源有限的情况下,减少浮点运算提高效率,但可能降低性能;因此,需在模型性能、计算资源和时间之间取得平衡<sup>[25]</sup>。

渔获物整体图像数据集(数据集1)相对于渔获物高识别度局部图像数据集(数据集3)需要更长的时间和更大的计算资源来处理图像信息。数据集3在准确率、训练参数和浮点运算数之间表现出色,达到了良好均衡。进一步分析数据集3训练所得的神经网络模型,根据鱼类局部识别准确率的不同,可以推断出卷积神经网络在不同金枪鱼延绳钓渔获物局部识别上的差异。

然而,在数据集2的训练过程中,模型识别结果(表4)显示出真实目标数基本小于检测目标数的情况。同时,对于体型较大的渔获物(如大青鲨和大眼长尾鲨),由于远距离拍摄导致分辨率降低和尺寸变化,模型的默认先验框容易在非目标区域检测到不存在的目标,增加了假阳性数

量。因此,针对以上问题,未来的研究方向可以聚焦于优化先验框的大小和比例,使其更好地适应不同尺寸和形状的渔获物,从而提升模型在不同尺度目标检测上的性能。

我们对渔获物高识别度局部图像数据集(数据集3)进行了不同目标检测卷积神经网络模型的分析与比较,结果表明各模型均展现出良好的识别效果(表6)。尽管双阶段(Two-stage)模型中

的Faster R-CNN在精确度上稍高于单阶段(One-stage)模型的YOLO和SSD,但其检测速度略显不足。相比之下,单阶段模型在检测速度上表现优异,尤其适合复杂的金枪鱼延绳钓渔船作业环境。尽管YOLO在检测速度上优于SSD,但在识别准确度上有所牺牲,特别是在识别渔获物局部图像时表现出一定的劣势。

表6 不同卷积神经网络模型性能分析

Tab. 6 Performance analysis of different convolutional neural network models

算法模型 Algorithmic model	测试准确率 Test accuracy/%	每秒处理图像帧数 Frames per second/(帧/s)	每张图片检测时间 Detection time per image/s	权重模型 Weight model/MB
Faster R-CNN	93.2	14.27	0.117	148.67
YOLO	87.2	48.73	0.024	44.73
SSD	91.6	31.92	0.027	92.64

综上所述,本研究提出的金枪鱼延绳钓渔获物局部图像数据集应用于SSD模型,相较于渔获物整体图像数据集相对简单更适用于SSD这样的轻量级模型,同时可以显著提高中西太平洋金枪鱼延绳钓渔获物鱼类图像分类的准确率,表现出更强的鲁棒性。在未来的研究中,处理鱼类局部图像数据时可以引入先验知识,通过手动标注鱼类的特定关键特征部位,或采用局部增强技术,使模型更加关注这些重要区域;对于视觉差异较小的细粒度特征,可以引入注意力机制(如SE模块、CBAM等),以提高模型对细微特征的区分能力,从而进一步提升识别性能。

作者声明本文无利益冲突。

## 参考文献:

- [1] 李勇,戴小杰,朱江峰,等.中西太平洋金枪鱼延绳钓渔获组成及其多样性分析[J].海洋湖沼通报,2011(2): 52-58.  
LI Y, DAI X J, ZHU J F, et al. Species composition and diversity of catches by tuna longline gear from the western and central pacific ocean[J]. Transactions of Oceanology and Limnology, 2011(2): 52-58.
- [2] 栾松鹤,戴小杰,田思泉,等.中西太平洋金枪鱼延绳钓主要渔获物垂直结构的初步研究[J].海洋渔业,2015, 37(6): 501-509.  
LUAN S H, DAI X J, TIAN S Q, et al. Vertical distribution of main species captured by tuna longline fishery in the Western and Central Pacific [J]. Marine Fisheries, 2015, 37(6): 501-509.
- [3] 王啸,王佚兮,刘文俊,等.热带中西太平洋金枪鱼延绳钓渔获物组成及其多样性[J].中国水产科学,2022, 29(5): 732-743.  
WANG X, WANG Y X, LIU W J, et al. Catch composition and species diversity of pelagic longline fishing in the tropical Western and Central Pacific Ocean [J]. Journal of Fishery Sciences of China, 2022, 29(5): 732-743.
- [4] 乐美龙.金枪鱼类渔业管理问题的研究之二:金枪鱼渔业区域性管理组织和其管理新趋势[J].中国水产科学,2008, 15(5): 26-29.  
LE M L. Study on the management of tuna-like fisheries II: regional management organisations of tuna fisheries and new trends in their management [J]. Journal of Fishery Sciences of China, 2008, 15(5): 26-29.
- [5] ALSMADI M K S, OMAR K B, NOAH S A, et al. Fish recognition based on the combination between robust feature selection, image segmentation and geometrical parameter techniques using Artificial Neural Network and Decision Tree [J]. International Journal of Computer Science and Information Security, 2009, 6(2): 215-221.
- [6] 欧利国,王冰妍,刘必林,等.基于计算机视觉的3种金枪鱼属鱼类形态指标自动测量研究[J].海洋学报,2021, 43(11): 105-115.  
OU L G, WANG B Y, LIU B L, et al. Automatic measurement of morphological indexes of three Thunnus species based on computer vision [J]. Haiyang Xuebao, 2021, 43(11): 105-115.
- [7] 刘雨青,李杰,宋利明,等.融合通道剪枝与ByteTrack的轻量化金枪鱼渔获数量实时检测[J].上海海洋大学学报,2023,32(5):1080-1089.  
LIU Y Q, LI J, SONG L M, et al. Tuna catch real-time detection by fusing channel pruning with ByteTrack light

- weight network [J]. Journal of Shanghai Ocean University, 2023, 32(5): 1080-1089.
- [8] 汤永华, 张志鹏, 林森, 等. 基于深度学习的鱼类识别相关技术研究现状及展望 [J]. 海洋渔业, 2024, 46(2): 246-256.
- TANG Y H, ZHANG Z P, LIN S, et al. Review and prospect of fish recognition and related techniques based on deep learning [J]. Marine Fisheries, 2024, 46(2): 246-256.
- [9] RUM S N M, NAWAWI F A Z. FishDeTec: a fish identification application using image recognition approach [J]. International Journal of Advanced Computer Science and Applications, 2021, 12(3): 102-106.
- [10] ROBILLARD A J, TRIZNA M G, RUIZ-TAFUR M, et al. Application of a deep learning image classifier for identification of Amazonian fishes [J]. Ecology and Evolution, 2023, 13(5): e9987.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [12] 张铠臻, 李艳武, 刘博, 等. 基于VGG16网络的超参数调整策略的研究 [J]. 科技与创新, 2021(22): 10-13.
- ZHANG K Z, LI Y W, LIU B, et al. Research on hyperparameter tuning strategy based on VGG16 network [J]. Science and Technology & Innovation, 2021(22): 10-13.
- [13] KINGMA D P, BA J. Adam: a method for stochastic optimization [J]. CoRR, 2014, abs/1412.6980.
- [14] SMITH L N. No more pesky learning rate guessing games [J]. arXiv:1506.01186v2, 2015.
- [15] LEE J, SCHOENHOLZ S S, PENNINGTON J, et al. Finite versus infinite neural networks: an empirical study [C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2020: 15156-15172.
- [16] ERHAN D, SZEGEDY C, TOSHEV A, et al. Scalable object detection using deep neural networks [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 2155-2162.
- [17] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The PASCAL Visual Object Classes (VOC) challenge [J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [18] ZHAO Z Q, ZHENG P, XU S T, et al. Object detection with deep learning: a review [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212-3232.
- [19] SARRAF A, AZHDARI M, SARRAF S. A comprehensive review of deep learning architectures for computer vision applications [J]. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 2021, 77(1): 1-29.
- [20] GU J X, WANG Z H, KUEN J, et al. Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018, 77: 354-377.
- [21] SPAMPINATO C, GIORDANO D, DI SALVO R, et al. Automatic fish classification for underwater species behavior understanding [C]//Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams. New York: Association for Computing Machinery, 2010: 45-50.
- [22] LEKUNBERRI X, RUIZ J, QUINCOCES I, et al. Identification and measurement of tropical tuna species in purse seiner catches using computer vision and deep learning [J]. Ecological Informatics, 2022, 67: 101495.
- [23] CARUANA R, LAWRENCE S, LEE GILES C. Overfitting in neural nets: backpropagation, conjugate gradient, and early stopping [C]//Proceedings of the 13th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2000: 402-408.
- [24] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning [M]. Cambridge, MA: MIT press, 2016.
- [25] ALBEAHDILI H M, HAN T, ISLAM N E. Hybrid algorithm for the optimization of training convolutional neural network [J]. International Journal of Advanced Computer Science and Applications, 2015, 6(10): 79-85.

## Analysis of deep learning-based tuna longline catch image recognition technique

XIA Chao<sup>1</sup>, CHEN Xinjun<sup>1,2,3,4</sup>, LIU Bilin<sup>1,2,3,4</sup>, KONG Xianghong<sup>1,3</sup>, YE Xuchang<sup>1,3</sup>

(1. College of Marine Living Resource Sciences and Management, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Sustainable Exploitation of Oceanic Fisheries Resources, Ministry of Education, Shanghai 201306, China; 3. National Engineering Research Center for Oceanic Fisheries, Shanghai 201306, China; 4. Key Laboratory of Sustainable Utilization of Oceanic Fisheries, Ministry of Agriculture and Rural Affairs, Shanghai 201306, China)

**Abstract:** In order to achieve efficient identification and classification of tuna longline catches and to improve the accuracy of fishery resources monitoring, this study explores a fish image recognition method based on convolutional neural network. The experiments were conducted using image data of three economic fish species and ten bycatch species caught by the Songhang of Shanghai Ocean University during the high seas survey in the western and central Pacific Ocean, and a convolutional neural network (CNN) based on a single shot multiBox detector (SSD) was used to classify and recognise the images. The training dataset is optimised by comparing and analysing the local fish images with the overall image dataset to improve the classification performance of the model. The experimental results show that the classification accuracy of the improved fish image dataset on the SSD model reaches 91.6%, which is a 6.2% improvement compared to the original dataset. The study shows that using the optimised dataset, the SSD model can significantly improve the recognition accuracy of tuna longline catches with better stability and adaptability. This study provides an effective technical path for fishery resource monitoring based on convolutional neural networks, especially in improving the automatic classification and identification accuracy of tuna longline catches, which is of great significance for promoting sustainable fishery management and marine ecological protection.

**Key words:** tuna longlining; catch identification; convolutional neural networks; dataset optimisation; Western and Central Pacific Ocean