arXiv:2301.00596v1 [cs.CV] 2 Jan 2023

一种用于野生鱼类种群个体再识别的对比学习方法

Ørjan Langøy Olsen[1],Tonje Knutsen Sørdalen[2],Morten Goodwin[1],Ketil Malde[3], KristianMuri Knausg˚ard*[4]，以及 Kim Tallaksen Halvorsen[3]

[1]挪威阿格德尔大学人工智能研究中心[2]挪威阿格德尔大学海岸研究中心 [3]挪威卑尔根海洋研究所生态系统声学组[4]挪威阿格德尔大学顶尖机电一体化研究中心

摘要

在陆地和海洋生态学研究中，实体标记都是研究种群动态和行为的常用方法。然而，这类标记技术正日益被基于图像分析的个体再识别所取代。

本文介绍了一种基于对比学习的个体识别模型。该模型采用 Inception v3 网络的前几层，并配有投影头；我们利用对比学习，从一组标准化照片中找出相似或不相似的图像对。我们将这一技术应用于绿隆头鱼，*Sympho-dus melops*，这是一种在生态和商业上都十分重要的鱼类物种。照片是在野生种群中多次捕获同一个体时拍摄的，各次目击记录之间的间隔可能从几天到数年不等。

我们的模型在数据集上的单样本准确率为 0.35，5样本准确率为 0.56，100样本准确率为 0.88。

1 引言

实体标记通过使用外部或内部标记进行个体识别，是监测陆生和水生动物种群的一种广泛使用的方法。对同一个体的重见记录或再捕获信息可用于估算种群规模、存活率和移动模式。然而，大多数标记方法成本高昂、具有侵入性且耗费人力。值得庆幸的是，许多动物具有独特的自然斑纹或形态特征，可用于照片识别，从而无需使用实体标签 [22, 19]。然而，对生态学家而言，研究鱼类可能意味着要追踪种群中的数百甚至数千个体，这使得人工照片识别极具挑战性，甚至难以实现。因此，能够对个体进行全自动或半自动再识别的工具将对生态学家大有裨益。

再识别（re-ID）不同于常规分类，因为它属于少样本学习问题。小样本问题的特点是每个类别的样本很少，但类别数量可能很多，甚至不确定。解决这类问题的一种方法是度量学习：将数据转换为低维嵌入向量，使同一类别的点聚集在一起。随后即可基于这些嵌入向量进行分类。研究已证明，度量学习方法适用于动物物种的再识别 [18]。度量学习方法的一个关键优势是，网络无需重新训练即可添加新类别。

对比学习是一种可用于解决小样本问题的技术。对比学习会比较数据，并判断它们相似还是不同。孪生网络 [1] 是

*通讯作者 kristianmk@ieee.org

https://doi.org/10.7557/18.6824© 作者。被许可方：Septentrio 学术出版，挪威特罗姆瑟。本文为开放获取文章，依据知识共享署名许可协议 (http://creativecommons.org/licenses/by/4.0/) 的条款和条件发布。

最基本的形式，它将两个输入通过具有相同共享权重的同一个网络，并为二者生成嵌入。在训练过程中，它会尝试预测这两个输入是否属于同一类别。它的一大优势是，无需知道某个输入属于哪个类别，也无需知道共有多少个类别。三元组网络 [10]是对孪生网络的改进，包含三个输入。

本研究旨在检验基于图像的再识别分析在一种具有重要商业和生态价值的鱼类物种——绿隆头鱼（*Symphodus melops*）中的适用性。该图像数据集由野生种群中个体被捕获和再捕获时拍摄的标准化照片组成，个体目击记录之间的时间跨度从数天到数年不等。第一步是使用目标检测器在图像中检测鱼，随后采用再识别方法。若精确率足够高，计算机视觉再识别有望取代实体标记，用于个体识别，并可应用于监测存活率、生长、移动和种群规模；这些都是实现可持续管理和保护所需的关键信息 [18] [4]。

2 相关研究

机器学习的进步带来了强大的技术，可从图像和视频数据中提取重要的生态信息。例如，机器学习已成功用于检测鱼类伤口 [5], 、对生物进行计数和分类（基于数字照片和实时视频） [14], [12], 、识别物种， [6], 以及发现并统计数字图像中的生物, [4], 甚至量化它们的行为 [3]。

已有一些关于鱼类再识别的研究，但针对野生硬骨鱼的研究仍然缺乏。Bruslund Haurum 等人 [2] 使用度量学习，在 6 个类别中每类 15 个样本的条件下，对斑马鱼取得了 99% 的 mAP。Mei- dell 和 Sjøblom [16]报告称，在由 715 个个体组成的 225,000 张鲑鱼 ImageNet 图像上，真正例率达到 96%。Li 等人 [13] 使用 10 个个体的 3412 张图像，通过其新颖的 FFRNet 网络取得了 92% 的准确率。这些研究有一个共同点：它们都是在圈养条件下开展的，并且没有使用时间上相互独立的观测结果。换言之，这些个体不会因生长、成熟、衰老或类似的生物过程而改变形态。Moskvyak 等人 [17] 在包含 120 个蝠鲼个体的 1730 张图像的数据集上采用了度量学习方法，取得了 62% 的准确率@1 和 97% 的准确率@10。

3 方法

3.1 数据收集

研究物种*S. melops*是东北大西洋沿海生态系统中具有重要商业和生态价值的物种 [7]。该物种有两种明显不同的雄性形态：体型较大、色彩鲜艳、会筑巢并照料鱼卵的雄鱼，以及体型较小、颜色偏棕、与雌性形态（棕色和灰色）相似的伺机交配雄鱼 [21]。数据集于2018至2021年在挪威西部奥斯特沃尔采集，研究人员将笼式网留在海中过夜，以捕捉绿隆头鱼，并为所有捕获个体标记上带有唯一编码的被动式集成应答器（PIT）标签（11毫米标签, RFID Solutions）。标签植入鱼的腹腔内，完整的采样说明见 [8] 和 [9]。

该方法使我们能够跨时间收集每个个体的独立观测数据，并使数据集涵盖鱼类形态的变化。每次捕获时，都会从鱼的两侧拍摄几张照片，并根据射频识别（RFID）为照片标注编号。拍摄时，鱼的背侧朝上。经过一些筛选后，我们整理出一个可用于该任务的数据集。最终数据集包含来自513个不同个体的2113张照片。另据统计，所有个体首次与最后一次捕获日期之间的平均间隔为230天。图1展示了数据集中的样本。

3.2 个体再识别

再识别系统由一系列不同组件构成，如图2所示。这些组件分为两类：预处理部分和再识别部分。作为流水线预处理步骤的一部分，系统以图像为输入，并将其送入目标检

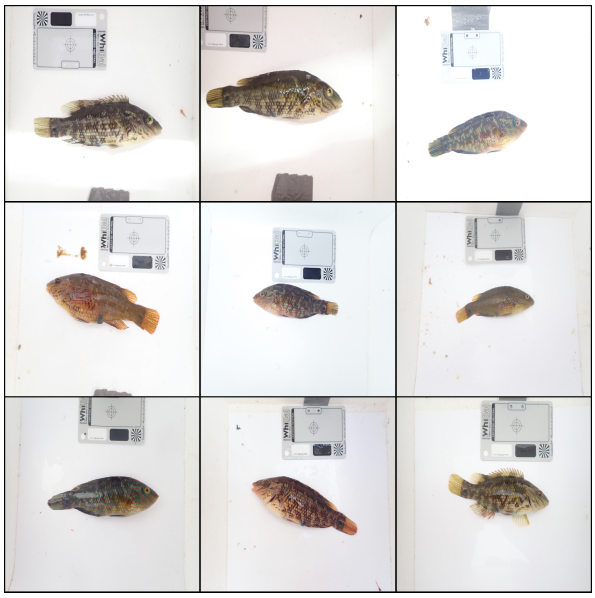


图1：未处理数据集中的样本。

测网络，以获取仅包含画面中鱼的图像裁剪。接着，另一个网络——方向组件——会对鱼是朝向左侧还是右侧进行分类，并将结果作为元数据传递。对于再识别部分，预处理数据会被输入对比学习网络，该网络学习将同一个体的嵌入向量归为一组，并将不同个体的嵌入向量彼此区分。随后即可根据嵌入向量进行分类。通过存储所有先前观察到的个体的嵌入向量，可以利用最近邻方法实现再识别。

目标检测器使用 YOLOv5 [11]，图像尺寸为 416x416，批次大小为 32，训练 50 轮。训练期间，网络使用人工标注的边界框进行训练，这些边界框将鱼框在其中。

方向网络采用 Inception v3 [20]模型，其所有权重均被冻结。网络中添加了一个全局平均池化层、一个包含 32 个神经元的 ReLU激活层，以及一个 sigmoid激活输出。训练所用的数据集是裁剪后仅包含头部的图像。数据集中的方向由人工标注。

嵌入网络由一个 CNN模型和一个投影头组成。其组成部分通过

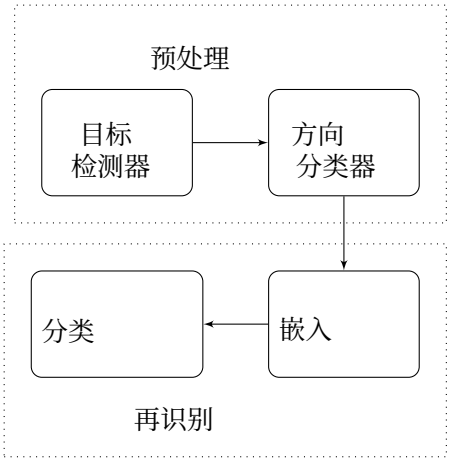


图2：网络流程以一张鱼的图像为输入，并输出该个体的 ID。

实验确定。CNN模型采用在 ImageNet 上预训练的 Inception v3 模型，并移除了第四个拼接层之后的层（第 46 层；若计算激活层，则为第 132 层）。末端附加了一个二维全局平均池化层和一个 128 维线性投影，该投影经过归一化后位于单位超球面上。网络结构图如图3所示。网络的输入尺寸为 224，因此图像在输入网络前会相应调整大小。网络采用加边填充以保持纵横比。在嵌入网络的训练过程中，数据集被划分为训练集和测试集，其中测试集所占比例为 0.3。

嵌入网络采用逐步解冻的方式进行训练。前 100 个训练轮次中，第 29 层之前的层被冻结，学习率为 0.001；接下来的 100 个训练轮次中，第 18 层之前的层被冻结，学习率为 0.0001，总计 200 个训练轮次。选择第 29 层和第 18 层，是因为它们是网络架构中的拼接瓶颈（图3中的绿色节点）。损失函数采用间隔为 1.0 的在线困难三元组挖掘。困难三元组挖掘是一种仅对负样本比正样本更接近锚点的三元组进行反向传播的技术。因此，使用在线挖掘无需采用三个具有共享权重的相同网络。训练样本会随机进行图像增强。色调的调整幅度在 -20 到 20 之间

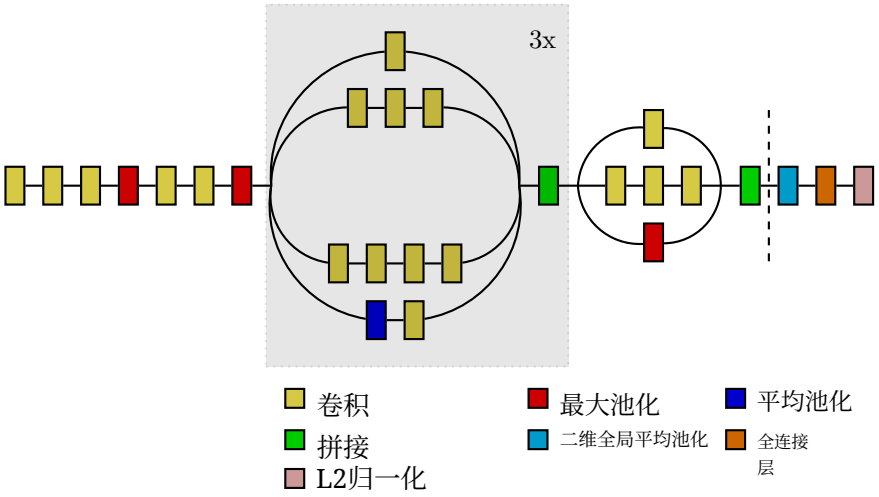


图3：嵌入网络使用了 Inception v3 [20] 的前半部分，并配有自定义投影头。虚线标示了 Inception 部分的结束位置和自定义部分的起始位置。灰色部分重复三次。

并调整饱和度。图像会向任一方向旋转 0 到 0.1 的比例，并进行 0 到 0.1 的缩放变换。批次大小为 32。

分类以及由此延伸而来的再识别采用最近邻方法。在这种情况下，将训练集定义为支持集、测试集定义为查询集很有用。最近邻分类是非参数方法，无需通过优化进行训练。训练步骤只需将支持集图像输入嵌入网络，并存储相应的嵌入向量，供推理步骤使用。要对图像进行分类，只需将查询图像输入嵌入网络，然后选择查询嵌入与支持集嵌入向量中距离最近的点所对应的类别。我们的实现源代码可在 GitHub[1] 上获取。

3.3 实验方法

Symphodusmelops 的头部区域（尤其是鳃盖）具有明显的高对比度图案。因此，探究网络在头部裁剪图像上的表现是否优于在全身裁剪图像上的表现会很有意义。实验通过对裁剪自这两个部位的图像进行训练和评估来开展嵌入网络。

系统还可以将鱼的左右两侧分别视为不同类别，因此，对两侧分别进行推理并在集成分类器中合并结果，可以获得有价值的信息。在此实验中，数据集被划分为左侧图像部分和右侧图像部分，使两部分中的图像成对对应。训练两个模型，一个仅用于左侧图像，另一个仅用于右侧图像。对于每一侧，支持集中的嵌入向量按照与查询图像的距离排序。然后，将两个已排序集合一并考虑，最先出现的类别即为预测类别。

我们还开展了一项实验，以评估系统区分重新观测到的个体与此前从未见过的个体的能力。如果某个查询嵌入与任何支持嵌入之间的距离都大于特定距离阈值，则将其视为新个体。查询集被划分为测试集和验证集。我们通过网格搜索寻找合适的距离阈值，使测试集评估中的 F1分数最大化。本实验的验证数据集包含 317 个样本。

[1] https://github.com/orilan93/SiameseFish

4 结果

我们使用的指标包括准确率@1、准确率@5 和 mAP@5。准确率@1 表示排名最高的类别的正确性，即排名最高的预测类别与真实类别相同的比例。准确率@5 表示排名前五的类别的正确性，即排名前五的类别中包含真实类别的比例。类似地，mAP@5 表示排名前五的类别的精确率，即真实类别中有多少位于排名前五的类别之中。

4.1 再识别

该再识别系统在头部和身体裁剪数据集上均进行了评估。表1展示了准确率@1和准确率@5的结果，表明模型在头部裁剪图像上的表现最佳。图4展示了随着累计尝试次数增加，模型的表现如何变化。该方法在实际应用中至关重要，因为专业生物学家不必在未经排序的图像目录中逐一查找，而是可以查看排序目录，并根据距离度量对最有希望的图像进行排序，在检查了 k 张最有希望的图像后，找到正确个体。数量越大k ，准确率越高；当尝试次数接近支持集中的图像数量时，准确率也接近100%。

表1：头部和身体裁剪图像再识别结果。

Type	准确率@1	准确率@5	mAP@5
Head	0.3534	0.5647	0.4227
Body	0.2043	0.3892	0.2690

表2展示了数据集中随机抽取的四张图像样本，以及训练模型预测为同一个体的图像和真实值。图中还展示了分类排名和嵌入空间中的距离。

为了深入了解模型在进行推理时所关注的内容，我们在图5中展示了一些测试集样本及相应的SHAP图 [15]。彩色区域表明，模型确实捕捉到了鱼的斑纹。

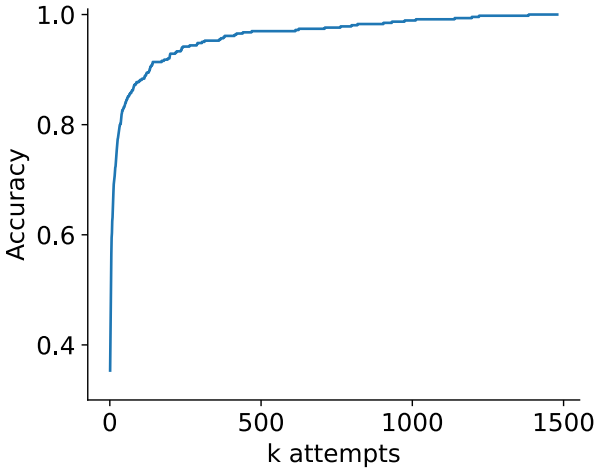


图4：达到特定准确率（accuracy@k）所需的累计尝试次数（k）。

4.2 集成分类器

该实验展示了分别针对鱼的每一侧训练新模型，然后合并各自分类结果的结果。表3显示，这一策略可以显著提升性能。请注意，集成分类器所需的方向组件在使用头部裁剪数据集的验证集上取得了99.38% 的 accuracy@1。

4.3 新观测

正如前面的实验所示，重新识别的效果相对较好。我们希望利用该模型区分新个体与先前观测到的个体。要用该模型识别新个体，需要确定一个嵌入距离阈值。请注意，这与表2中的距离度量有关。通过网格搜索，我们发现阈值为0.820时，模型在验证集上的性能得分最佳。系统将95个个体预测为新的目击记录，在此任务上的准确率@1为62.78%。

5 讨论与结论

表4中总结的实验结果表明，系统在鱼的头部裁剪图像上的表现优于全身。这很可能是因为

表2：查询图像与正确图像的嵌入之间的检索排名和欧几里得距离。

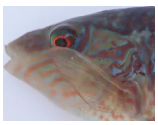
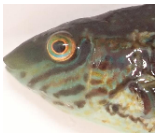
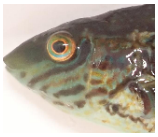
查询	预测结果	真实值	Rank	距离
			1	0.65
			217	1.27
			1	0.61
			1012	1.46

表3：集成分类器的结果。

Type	准确率@1	准确率@5	mAP@5
Left	0.3568	0.5463	0.4243
正确	0.4097	0.5595	0.4623
作为一对	0.5286	0.7533	0.6140

表4：结果摘要。

实验	结果	指标
再识别	0.3534	准确率@1
目标检测器	0.9951	mAP@0.5
方向分类器	0.9937	准确率@1
集成分类器	0.5286	准确率@1
新观测数据	0.6278	准确率@1

头部的纹理最为独特，因此是一项重要特征；调整图像尺寸以适应网络输入尺寸时，算法能够以更高的分辨率获取这一特征。然而，这样做的缺点是，网络能够接触到的数据中的信息会更少。

通过以新的方式利用现有系统，为鱼的每一侧分别训练模型，可以构建一个集成分类器。测试结果显示，该方法将准确率从 35% 大幅提升至 53%。这表明，利用所有可用信息进行准确预测至关重要。

该系统的准确率还不足以支持完全自动化且无需人工介入的系统，而这正是取代生态学研究中实体标签所需要的。不过，我们认为，持续收集数据可以构建一个时间分布更加均衡的数据集，使模型能够考虑个体的生长和衰老。

自动化能够带来诸多益处，正被越来越多的行业采用，生态学领域也不应例外。成功开发出精确率高的再识别算法，可以提供一种改善鱼类福利的新方法，同时成本更低（只需要一台摄像机），而且可能更加准确（不会丢失标签）。展望未来，我们设想再识别技术可以直接应用于

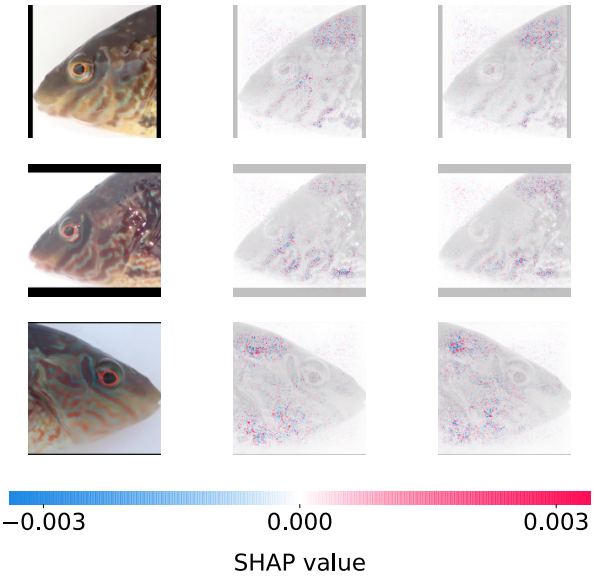


图5：SHAP图显示图像中哪些区域对模型的决策影响最大。

来自水下摄像机的实时流媒体，从而完全免去捕捉和处理鱼类的需要。这将是一种革命性的方法，能够彻底改变我们收集鱼类及其他动物相关关键信息的方式，以促进可持续保护与管理。

致谢

感谢 Torkel Larsen、Anne Berit Skiftesvik、Ovin Holm、Ylva Vik、Nicolai Aasen、Ben Ellis、Vegard Omestad Berntsen 和 Steve Shema 协助在野外采集照片及捕获-再捕获数据。本研究获得了阿格德尔大学人工智能研究中心（CAIR）、海岸研究中心（CCR）和机电一体化顶尖研究中心（TRCM）、海洋研究所（项目15638-01）以及挪威研究委员会的资助（CoastVision，项目编号 325862；CreateView，项目编号 309784）。

参考文献

[1] J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. S¨ackinger 和 R. Shah。使用“孪生”时延神经网络进行签名验证。国际模式识别与人工智能期刊，7(04):669–688，1993 年。

[2]J. Bruslund Haurum、A. Karpova、M. Pedersen、S. Hein Bengtson 和 T. B. Moes- lund。使用度量学习对斑马鱼进行再识别。发表于*IEEE/CVF*冬季计算机视觉应用会议研讨会论文集，第1–11页，2020年。doi：10.1109/WACVW50321.2020.9096922。

[3]E. M. Ditria、E. L. Jinks 和 R. M. Con- nolly。利用图像分类和光流技术自动分析视频中的鱼类啃食行为。《动物行为》，177：31–37，2021年7月。ISSN 00033472。doi：10.1016/j.anbehav.2021.04.018。

[4] M. Goodwin、K. T. Halvorsen、L. Jiao、K. M. Knausg˚ard、A. H. Martin、M. Moyano、R. A. Oomen、J. H. Rasmussen、T. K. Sørdalen 和 S. H. Thorbjørnsen。《释放深度学习在海洋生态学中的潜力：概述、应用与展望》。*ICES*海洋科学期刊，2022年1月。ISSN 1054-3139。doi：10.1093/icesjms/fsab255。

[5] A. Gupta、E. Bringsdal、N. Salbuvik、K. M. Knausg˚ard 和 M. Goodwin。用于养殖鲑鱼伤口检测的精确卷积神经网络方法。载于神经网络工程应用国际会议，第139–149页。Springer，2022年。doi：10.1007/978-3-031-08223-8 12。

[6] A. Gupta、E. S. Kalhagen、Ø. L. Olsen 和 M. Goodwin。应用于鱼类物种的分层目标检测：鱼类物种的分层目标检测。*Nordic Machine Intel-ligence*，2(1)，2022年。doi：10.5617/nmi.9452。

[7]K. T. Halvorsen、T. K. Sørdalen、C. Durif、H. Knutsen、E. M. Olsen、A. B. Skiftesvik、T. E. Rustand、R. M. Bjelland 和 L. A. Vøllestad。筑巢绿隆头鱼（Sympho- dus melops）中雄性偏向的性别体型二态性

：对按体型管制的渔业的影响。*ICES*海洋科学期刊，73（10）：2586–2594，2016年11月。ISSN 10959289。doi：10.1093/icesjms/fsw135。

[8]K. T. Halvorsen、T. K. Sørdalen、L. A. Vøllestad、A. B. Skiftesvik、S. H. Espeland 和 E. M. Olsen。绿隆头鱼（Symphodus melops）的性别和体型选择性捕捞——这种清洁鱼用于鲑鱼养殖。*ICES*海洋科学期刊，74 (3):660–669，2017年3月。ISSN 1054-3139。doi: 10.1093/icesjms/fsw221。

[9] K. T. Halvorsen、T. Larsen、H. I. Browman、C. Durif、N. Aasen、L. A. Vøllestad、A. Cresci、T. K. Sørdalen、R. M. Bjelland 和 A. B. Skiftesvik。温带隆头鱼（隆头鱼科）在一处小型海洋保护区内的移动模式。鱼类生物学期刊，99 (4):1513–1518，2021年7月。ISS N 0022-1112。doi: 10.1111/jfb.14825。

[10] E. Hoffer 和 N. Ailon。使用三元组网络进行深度度量学习。载于国际相似性模式识别研讨会，第 84–92 页。Springer，2015 年。doi：10.1007/ 978-3-319-24̲261-3 7。

[11]G. Jocher、A. Stoken、A. Chaurasia、J. Borovec 等。ultralytics/yolov5：v6.0 - YOLOv5n「Nano」模型、Roboflow 集成、TensorFlow 导出、OpenCV深度神经网络支持，2021 年 10 月。网址：https: //doi.org/10.5281/zenodo.5563715。

[12]K. M. Knausg˚ard、A. Wiklund、T. K. Sørdalen、K. T. Halvorsen、A. R. Kleiven、L. Jiao 和 M. Goodwin。温带鱼类检测与分类：一种基于深度学习的方法。应用智能，52(6):6988–7001，2022年。doi：10.1007/ s10489-020-02154-9。

[13] D. Li、H. Su、K. Jiang、D. Liu 和 X. Duan。基于旋转目标检测的鱼脸识别：数据集与探索。鱼类，7（5）：219，2022年。doi：10.3390/fishes7050219。

[14]V. Lopez-Vazquez、J. M. Lopez-Guede、S. Marini、E. Fanelli、E. Johnsen 和 J. Aguzzi。用于水下动物检测与分类的海底有缆观测站视频图像增强与机器学习流水线。*Sensors*，20(3):726，2020。doi: 10.3390/s20030726。[15] S. M. Lundberg 和 S.-I. Lee。一种解释模型预测的统一方法。神经信息处理系统进展，30，2017。

[16] E. Meidell 和 E. S. Sjøblom。Fishnet：用于鲑鱼识别的统一嵌入。硕士论文，挪威科技大学，2019年。

[17]O. Moskvyak、F. Maire、F. Dayoub、A. O. Armstrong 和 M. Baktashmotlagh。通过学习姿态不变嵌入，根据自然斑纹对蝠鲼进行鲁棒再识别。载于*2021*年数字图像计算：技术与应用（*DICTA*），第1–8页。IEEE，2019年。doi：10.1109/DICTA52665. 2021.9647359。

[18] S. Schneider、G. W. Taylor、S. Linquist 和 S. C. Kremer。运用计算机视觉从相机陷阱数据中对动物进行再识别的过去、现在与未来方法。生态学与进化方法，10(4):461–470，2019。doi: 10.1111/2041-210X.13133。

[19] S. Schneider、G. W. Taylor 和 S. C. Kremer。用于动物个体再识别的相似性学习网络——超越人类观察者的能力。载于*IEEE/CVF* 计算机视觉应用冬季会议研讨会论文集，第 44–52 页，2020。

[20] C. Szegedy、V. Vanhoucke、S. Ioffe、J. Shlens 和 Z. Wojna。重新思考用于计算机视觉的 Inception 架构。载于 *IEEE* 计算机视觉与模式识别会议（*CVPR*）论文集，2016 年 6 月。

[21] I. Uglem、G. Rosenqvist 和 H. S. Wasslavik。绿隆头鱼二态雄鱼之间的表型变异。鱼类生物学期刊，57(1):1–14，2000年7月。ISSN 00221112。doi: 10. 1006/jfbi.2000.1283。

[22] B. G. Weinstein。一种用于动物生态学的计算机视觉方法。动物生态学杂志，87(3): 533–545，2018年。doi: 10.1111/1365-2656.12780。