

机器学习阅读报告

一、文献信息

1. 论文题目: A Deep Learning-based Framework for the Detection of Schools of Herring in Echograms
2. 作者: Alireza Rezvanifar、Tunai Porto Marques、Melissa Cote、Alexandra Branzan Albu、Alex Slonimer、Thomas Tolhurst、Kaan Ersahin、Todd Mudge、Stéphane Gauthier
3. 发表途径: NeurIPS 2019
4. 发表时间: 2019

二、问题意义

1. 研究问题: 如何基于深度学习利用回声图自动检测鲱鱼群

2. 研究背景: 追踪水下物种的丰富度对于了解气候变化对海洋生态系统的影响至关重要。生物学家通常使用回声探测仪监测水下地点, 并将数据可视化为回声图。他们手动或半自动利用经典的机器学习来对这些数据进行分类, 这很费时, 而且容易出现不一致。在计算机视觉的各种任务中, 深度学习已经被证明是非常有效的但是它还没有渗透到超声图像分析中。将深度学习应用于声纳数据中移动海洋生物的检测, 这些数据通常比回声图具有更高的分辨率。

3. 研究意义: 研究水下物种的丰富程度对于研究与气候变化有关的现象引起的水温变化的影响来说是至关重要的。本文提出了一种基于深度学习技术, 从回声图自动检测鱼类, 减少处理时间、所需人力和结果的不一致性。该框架将有助于解释大量原始声学数据, 并且有可能被扩展以处理更多的水下物种数据。

三、思路方法

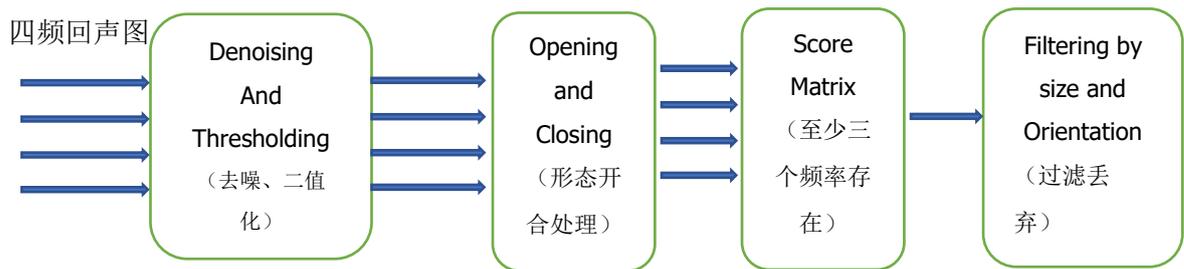
以鲱鱼群作为案例和研究目标。

1. 使用数据集。使用 2015 年在加拿大不列颠哥伦比亚省温哥华岛附近的探索通道内, 在水面上部署的 AZFP 装置收集的数据生成的四频回声图 (67、125、200 和 455 KHz)。四个频率的回声图作为 ROI 提取器的输入, 可以使信息互补。总共有 100 个回声图, 其中 70 个用于训练 (其本身分为 80% 的训练, 20% 的验证) 和 30 个用于测试。丰富的回声图使得数据量不限制对基于深度学习的框架的训练。

2. 设计 ROI 提取器。鲱鱼有两个共同的高水平特征:

- a) 在大多数回声图中可见的强强度核心。
- b) 垂直拉长的形状。

使用 ROI 提取器基于以上特征, 提取 ROI, 即确定可能目标。将回声图进行了以下处理, 得到最终的 ROI 边界框。



3.对提取器的输出图像进行分类。分类阶段决定了每个 ROI 代表的是鲱鱼群还是背景。我们使用基于深度学习的图像分类器对训练集中的 70 个回声图进行训练：将 ROI 边界框与已知正确样本进行比较，计算出重叠区域与并集区域的比值。IoU 得分小于 0.4 的 ROI 边界框作为错误样本，其他的正确正样本。

4.框架性能评估。

ROI 提取器：对于 100 张回声图，在 ROI 提取后使用 DenseNet 进行分类，根据不同的 IoU 阈值将其分为正确样本和错误样本，对 precision 和 recall, f1-core 三个数据进行统计，我们以 recall 为评估性能的依据，尽量防止 ROI 提取器错失样本。（Recall=捕获到的鲱鱼群/总鲱鱼群）

整体框架：在对测试集中的 30 张回声图进行 ROI 提取后，设置 IOU 阈值为 0.4，使用 ResNet、DenseNet 和 InceptionV3 三种体系结构对其进行分类，将测试结果与传统的机器学习方法进行比较，得到其性能好坏。

四、实验结论

作者提出了一个新的框架，以协助利用水声资料研究水下物种的丰富度。这个框架主要有两部分组成：一个新的 ROI 提取器和一个基于深度学习的图像分类器。该框架在 IoU 阈值 0.4 上取得了良好的性能，保证了有意义的检测。

1.ROI 提取器：作者发现，随着 IoU 阈值从 0.4 减小到 0.0，recall 值逐渐从 0.834 上升到 0.931，也就是说 IoU 阈值越小，捕获到的鲱鱼群就越多。但其实这种差异是非常小的，这就很好地证明了大多数 ROI 边界框与基本事实有明显的重叠，设计的 ROI 提取器性能是比较好的。

2.整体框架：DenseNet 体系结构实现了各指标之间的最佳平衡，是最优的。能对 ROI 进行正确的分类，但不排除部分错误。低 IoU 阈值保证了最大召回率，也说明了经过训练的分类器的有效性。但不建议采用低阈值，因为较高的 IoU 阈值保证了检测与分类的可靠性。

五、启发思考

1.本文构造了一个新的框架：一个新的提取器和一个基于深度学习的图像分类器。在未来，可以将 ROI 提取器替换为 ROI 计算系统，简化可检测类的扩展，进一步收集新的视觉样本，得到更好的性能。是改进用于物种丰度跟踪和环境监测的数据处理和解释，并提供可靠支持气候变化影响分析的工具。

2.这个新框架的构成并不复杂，用这个新框架就可以自动地从回声图中得到目标的信息并进行分类，既减少了时间和劳动力，又提高了准确性和一致性。而在之后的研究中，可以将框架中的 ROI 提取器用通用的 ROI 计算系统代替，可以非常简便地扩展可检测类，拓展应用。

3.这项研究体现了机器学习的一大优点，就是省时省力，减少了人力进行判别的主观不确定性。在进行一些图像分析的工作时，人的判断往往是带有主观性的，那么就很容易出现结果的不一致性。基于深度学习的方法在保证高正确率的同时，也恰恰避免了这种结果的不一致性。