

DOI: 10.5846/stxb202002140258

项和雨, 邹斌, 唐亮, 陈维国, 饶凯锋, 刘勇, 马梅, 杨艳. 基于残差注意力网络模型的浮游植物识别. 生态学报, 2021, 41(17): 6883-6892.

Xiang H Y, Zou B, Tang L, Chen W G, Rao K F, Liu Y, Ma M, Yang Y. Phytoplankton recognition based on residual attention network. Acta Ecologica Sinica, 2021, 41(17): 6883-6892.

## 基于残差注意力网络模型的浮游植物识别

项和雨<sup>1</sup>, 邹斌<sup>1</sup>, 唐亮<sup>2</sup>, 陈维国<sup>3</sup>, 饶凯锋<sup>4,5,\*</sup>, 刘勇<sup>2</sup>, 马梅<sup>5,6</sup>, 杨艳<sup>7</sup>

1 湖北大学, 数学与统计学学院, 应用数学湖北省重点实验室, 武汉 430062

2 无锡中科水质环境技术有限公司, 无锡 214024

3 北京工业大学, 北京现代制造业发展基地, 北京 100124

4 中国科学院生态环境研究中心, 环境模拟与污染控制国家重点联合实验室, 北京 100085

5 中国科学院生态环境研究中心, 中国科学院饮用水科学与技术重点实验室, 北京 100085

6 中国科学院大学, 资源与环境学院, 北京 101407

7 武汉晴川学院, 计算机学院, 武汉 430204

**摘要:**浮游植物作为水生生态系统中最重要生物组成部分之一, 对水环境敏感, 在水环境监测中得到了广泛的关注。然而水生环境复杂多样, 准确高效地识别浮游植物是监测工作中的一大挑战。当前浮游植物识别方法可分为经典形态学分类、分子标记和人工智能图像识别三类。前两种方法已被广泛采用, 但费时费力, 不利于监测机构的大规模应用和推广。同样, 利用图像进行自动化分类难以在高准确率与高效率上达到平衡。深度学习技术的发展为此提供了新思路。本文提出一种新的深度卷积神经网络 RAN-11。该网络以残差注意力网络 Attention-56 和 Attention-92 为基础, 凭借通道对齐融合主干上的底层特征与顶层特征, 通过调整注意力模块和残差快个数以精简结构, 并引入了 Leaky ReLU 激活函数代替 ReLU。以太湖 11 个优势属共计 1036 张图像为数据来源进行对比验证。除星杆藻外, RAN-11 对单一优势属的的查准率都在 90% 以上, 并且有 5 个优势属达到 100% 的查准率。RAN-11 的识别准确率为 95.67%, 推理速率为 41.5 帧/s, 不仅比 Attention-92 (95.19% 的准确率, 23.6 帧/s) 更准确, 而且比 Attention-56 (94.71% 的准确率, 41.2 帧/s) 更快, 真正兼顾了准确率与效率。研究结果表明: (1) RAN-11 在查准率、准确率和推理速率上优于原始残差注意力网络, 更优于以词包模型为代表的传统图像识别方法; (2) 融合多尺度特征、精简网络结构和优化激活函数是提高卷积神经网络性能的有力手段。建立在经典分类基础之上, 本文提出新的残差注意力网络来提升浮游植物鉴定技术, 并构建出浮游植物自动化识别系统, 识别准确率高、易于推广, 对于实现水体中浮游植物的自动化监测具有重要意义。

**关键词:**水质监测; 浮游植物识别; 残差注意力网络; 深度学习

## Phytoplankton recognition based on residual attention network

XIANG Heyu<sup>1</sup>, ZOU Bin<sup>1</sup>, TANG Liang<sup>2</sup>, CHEN Weigu<sup>3</sup>, RAO Kaifeng<sup>4,5,\*</sup>, LIU Yong<sup>2</sup>, MA Mei<sup>5,6</sup>, YANG Yan<sup>7</sup>

1 Hubei Key Laboratory of Applied Mathematics, School of Mathematics and Statistics, Hubei University, Wuhan 430062, China

2 CASA (Wuxi) Environmental Technology Company Limited, Wuxi 214024, China

3 School of Economics and Management, Beijing University of Technology, Beijing 100124, China

4 State Key Joint Laboratory of Environment Simulation and Pollution Control, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

5 Key Laboratory of Drinking Water Science and Technology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing

**基金项目:**国家自然科学基金面上项目 (61772011); 中国科学院前沿项目 (QYZDY-SSW-DQC004); 北京市科技计划项目 (Z171100004417025); 山西省智能信息处理重点实验室开放项目 (CICIP2018002)

**收稿日期:**2020-02-14; **修订日期:**2021-03-01

\* 通讯作者 Corresponding author. E-mail: raokf@rcees.ac.cn

100085, China

6 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 101407, China

7 College of Computer, Wuhan Qingchuan University, Wuhan 430204, China

**Abstract:** Phytoplankton, one of the most important biological components in water ecosystem, is sensitive to water environment and has been widely concerned in water environment monitoring. However, the aquatic environment is complex and diverse. Accurate and efficient identification of phytoplankton becomes a challenge in monitoring work. Current phytoplankton recognition methods can be divided into three categories: classical morphological classification, molecular markers, and artificial intelligence image recognition. Although the first two methods have been widely used, they are time-consuming and laborious, which is not conducive to the large-scale application and promotion of monitoring agencies. Similarly, it is difficult to strike a balance between high accuracy and high efficiency for automatic classification based on images. Deep learning provides new insights for phytoplankton identification. We propose a new deep convolutional neural network RAN-11. The network is based on the residual attention network Attention-56 and Attention-92, fusing the bottom and top features of the backbone through channel concatenation, simplifying the structure by adjusting the number of attention and residual module, and adopting the Leaky ReLU activation function rather than ReLU. We used 1036 images of 11 dominant genera of Taihu Lake as data sources to compare and verify our algorithm. The precision of RAN-11 for a single prevailing genus was above 90%, with 5 species achieving 100% precision, except for *Asterionella* sp. The accuracy of RAN-11 was 95.67%, and the inference speed was 41.5 fps (frames per second), which is not only more accurate than Attention-92 (95.19% accuracy, 23.6 fps), but also faster than Attention-56 (94.71% accuracy, 41.2 fps), truly balancing accuracy and efficiency. Results indicate that: (1) RAN-11 is superior to the original residual attention network in terms of precision, accuracy and inference speed, as well as the traditional image recognition method represented by the Bags of Words model. (2) Fusion of multi-scale features, simplification of network structure and optimization of activation function are powerful means to improve the performance of convolutional neural network. Based on the classical classification, this paper proposes a new residual attention network to improve phytoplankton identification technology, and constructs an automatic phytoplankton recognition system, which has high recognition accuracy and easy promotion, and is of great significance to realize the automatic monitoring of phytoplankton in water.

**Key Words:** water quality monitoring; phytoplankton recognition; residual attention network; deep learning

浮游植物的定性定量检测是水环境检测的重要内容之一,对浮游植物进行快速准确识别将有助于对水域环境的监测、评价、开发和利用<sup>[1-3]</sup>。特别地,当水体富营养化引发浮游植物大量繁殖<sup>[4-6]</sup>尤其是蓝藻水华泛滥<sup>[7-8]</sup>,造成公共饮水威胁和巨大经济损失时<sup>[9]</sup>,浮游植物的快速准确识别愈渐凸显其重要性。2017年无锡“供水危机”之后<sup>[10-11]</sup>,国家、省市投入了大量的财力进行太湖综合整治。在蓝藻打捞、控源截污、生态清撤、引江济太等一系列措施下,太湖营养盐的控制取得了显著的成效,水质出现明显的改善<sup>[12]</sup>。尽管太湖水质的不断持续改善,但是太湖蓝藻水华每年仍然如期而至,个别年份暴发的面积和频次甚至出现反复。2019年2月20日生态环境部,就重点湖库水质及蓝藻水华监测发文《关于做好2019年重点湖库蓝藻水华防控工作的通知》,提出2019年1月至5月全国大部分地区气温较常年同期偏高,要求保障饮用水安全,各地应积极采取相关措施防控湖库蓝藻水华灾害。因此,浮游植物监测是一项非常重要而紧迫的工作。长期、有效的浮游植物监测数据积累是揭示富营养化机理和蓝藻水华暴发机制的基础工作,也是防控藻类水华的重要数据支撑。浮游植物生长在复杂的水生环境中,种类繁多且形态不一,准确而高效地识别水体中的浮游植物是监测工作中的重大挑战。

当前浮游植物识别方法主要分为三类<sup>[13-14]</sup>:(1)人工分类,即专家借助显微镜对浮游植物样本进行识别。这种方法最基本最方便,但耗时耗力,识别者往往需要经过专门的训练。(2)分子标记,基于浮游植物特征性

化学组成成分或 DNA 等遗传物质<sup>[15]</sup> 进行分类,如分光光谱法、荧光光谱法、液相色谱法、流式细胞仪法<sup>[16]</sup> 和分子探针技术等。这些方法较为可靠,但需要昂贵的仪器和精细的操作,难以推广。(3) 利用图像处理与模式识别技术分类。尽管人工分类是基于形态学的经典方法,应用最为广泛,但是随着浮游植物监测要求的不断提高,尤其是监测频次的不断增加,这种“高要求”和“人力成本”之间的矛盾愈加凸显。因此,自动化监测作为人工监测的辅助手段被认为是未来浮游植物监测的一个重要研究方向。

利用浮游植物图像进行自动化识别已有很多尝试<sup>[17]</sup>,是实现浮游植物识别自动化的重要途径。经典的图像识别方法包含图像的预处理、图像分割、特征提取和图像分类四个步骤。例如,Stefan 等<sup>[18]</sup> 提取傅里叶描述子和矩不变量等 47 种特征,利用判别分析方法对蓝绿藻进行自动识别;Luo 等<sup>[19]</sup> 通过傅里叶频谱特征得到特征向量,利用 BP 神经网络对圆形硅藻进行自动分类;Mansoor 等<sup>[20]</sup> 利用图像处理技术和多重感知机前馈人工神经网络实现了四种蓝藻的自动识别;Walker 等<sup>[21]</sup> 应用贝叶斯决策函数和荧光辅助图像对湖中两种主要浮游植物进行分类。这些方法的弊端在于需要利用传统的图像分割算法来辅助浮游植物定位,并且需要手工设计特征。面对不同的图像识别任务,需要先设计不同的图像分割算法和特征提取算法,这难以满足现场快速自动测量的要求。深度学习方法克服了这些缺点,得益于精心设计的网络结构,深度卷积神经网络可以自动学习图像的表征,端到端的输出分类结果,兼顾准确率、效率且易于扩展。

目前,利用深度学习方法进行浮游植物识别的研究较少,王永清等<sup>[22]</sup> 基于深度学习框架 Caffe 对三种藻类进行识别,虽达到 100% 的识别率,但识别种类较少,而且未详细描述数据集以及所用神经网络模型的结构和参数。本文以太湖的 11 种优势属为研究对象,基于深度学习中的残差注意力网络模型<sup>[23]</sup>,提出一种新的深度卷积神经网络 RAN-11,并搭建出易推广的浮游植物自动化识别系统,对于提高浮游植物识别准确率和效率具有重要意义。该系统有助于自动化、快速识别水体中的常见浮游植物种类和数量,对于实现水体中浮游植物的自动化监测具有重要价值。

## 1 材料与方 法

### 1.1 浮游植物图像获取

本研究组在太湖多个试点收集水样,然后利用自主研发的数字图像自动采集系统采集了多批浮游植物图像。最后,考虑到浮游植物种类的复杂性和样本数量的不均衡性,仅标注了太湖水体中的优势属,包括微囊藻,游丝藻,沟链藻,长孢藻,裸藻,盘星藻,转板藻,针杆藻,星杆藻,浮丝藻和栅藻共计 11 个属(图 1)。

### 1.2 图像预处理和计算平台搭建

浮游植物的显微镜图像像素高、尺寸大,且在拍摄过程中气泡、杂质、相互遮挡等因素会造成图像清晰度和辨识度难以达到算法识别的标准。为了能够更好地对浮游植物进行标注及算法识别,本研究组开发了一套可以快速对超大图像进行切割的工具软件,该工具可将 2cm×2cm 区域中的 72000×72000 像素的超大图像切割为边长 7200 的矩形图像。专家对切割后的 11 种浮游植物交叉标注,截取边长在 200-260nm 的浮游植物图像共计 1036 张,按照等比例缩放成分辨率为 224×224 的 PNG 格式图片,最后按 3:1:1 的比例将 1036 张图像随机分为训练集、验证集和测试集。

本文采用深度学习框架 Pytorch 搭建网络,在一台 Win10 工作站上进行网络训练和测试。该工作站配备 Intel(R) Xeon(R) Platinum 8268 CPU 和 NVIDIA Quadro GV100 GPU,主频 2.90GHz,显存 32G。

### 1.3 经典图像分类模型概要

近年来,图像分类方法多侧重于学习特征表达<sup>[24]</sup>,代表是词包模型(Bags of Words, BoW)<sup>[25]</sup> 和深度学习模型。其中词包模型在 2005-2012 年期间表现优异,多次斩获 PASCAL VOC 分类竞赛冠军。至 2012 年深度学习模型崛起,逐渐代替词包模型成为图像分类任务的主流方法,屡次刷新 ImageNet 大规模视觉识别竞赛(ImageNet Large Scale Visual Recognition Challenge, ILSVRC)的最高记录。

#### 1.3.1 词包模型

词包模型成熟于自然语言处理领域,主要通过计算文档中单词出现的频率来对文档进行表达和描述。在



图1 浮游植物扫描图

Fig.1 Microphotography of dominant phytoplankton in Lake Taihu (Automatic light scanning system)

图像处理分类问题中,词包模型将每个图像的局部特征看成一个视觉单词,形成词包进行分类,其基本流程如下:

- (1) 底层特征提取。对每一张图像提取特征,如 SIFT 特征或 SURF 特征;
- (2) 构建码本。使用 K-means 算法对提取的特征聚类,用聚类中心构建视觉码本;
- (3) 特征汇聚。利用视觉码本对图像特征编码,获取图像的词包作为新的特征描述;
- (4) 使用分类器对词包进行分类,如逻辑斯蒂回归分类器或支持向量机。

### 1.3.2 残差注意力网络

He 等<sup>[26]</sup>通过实验猜测,深的网络的求解器很难去利用多层网络拟合恒同映射  $g(x) = x$ ,因而在网络层数加深时难以求解目标函数  $H(x)$ 。为此,He 开创性地在卷积层之间加入快捷连接(即在主分支外另起一个分支,两个分支的输入皆为  $x$ ,输出相加),执行恒同映射,形成残差块。这样,原本求  $H(x)$  的问题,凭借快捷连接转化为求解另一个函数  $F(x) = H(x) - x$  的问题。可以发现,在恒同映射恰为目标函数  $H(x)$  的极端情形下,残差块完全避免了拟合恒同映射这一难题。由残差块堆积而成的深度残差网络(ResNet)<sup>[26]</sup>因为足够深并且容易优化,在 ILSVRC 2015 中获得了图像分类任务的冠军。

人类的视觉注意力机制是通过向目标区域投入更多注意力资源,以获取更多所关注目标的细节信息而抑

制其他无用信息。注意力机制可以通过加权来实现。Wang 等<sup>[23]</sup>通过不同的注意力模块来学习针对其特性而专门化的注意力,即令  $H_{i,c}(x) = (1 + M_{i,c}(x)) \times F_{i,c}(x)$ , 其中,  $F_{i,c}(x)$  代表图像的特征图的第  $i$  个位置第  $c$  个通道的特征值,  $M_{i,c}(x) \in [0, 1]$  代表该特征值的权重,  $H_{i,c}(x)$  为纳入了注意力机制的特征输出。由此构建的残差注意力网络 Attention-56 和 Attention-92<sup>[23]</sup> 在三种基准数据集上取得了优异的识别效果。

### 1.4 构建新网络

Wang 等<sup>[23]</sup>把自底向上、自上而下的前馈结构(软掩模分支)作为注意力模块的一部分,以改变特征上的权重,形成越来越精细的注意力。在该模型中先下采样再上采样确保了特征矩阵和权重矩阵的维度一致可以逐点相乘,改变快捷连接的个数和输出尺寸确保各个模块的注意力的精细程度不同,由此奠定了残差注意力网络的基础。

考虑到 Leaky ReLU 激活函数<sup>[27]</sup>在输入为负时仍旧可以产生非零的激活值,从而能促进反向传播过程中网络参数的更新,使网络收敛更快,因此本文使用 Leaky ReLU 激活函数代替 ReLU 激活函数。进一步地,为获取更好的图像表征,本文对浅层特征降采样后与深层特征按通道拼接,并用大步长卷积层代替池化层。但这些改变会增加网络的计算负担。因此,本文通过消融实验改变注意力模块和残差块的个数,并大幅调整相应的通道数以适应特征变换,由此形成卷积神经网络 RAN-11。RAN-11 为 residual attention network-11 的缩写,其中“11”表示该网络主要由表 1 中的 11 种模块组成。

表 1 RAN-11 的网络参数

Table 1 Network parameters of RAN-11

网络层 Network layer	缩写 Abbreviation	输入尺寸 Input size	卷积核参数 Kernel parameter	输出尺寸 Output size
卷积层 Convolution	C	224×224×3	[3×3, 64]	112×112×64
残差块 Residual unit	R	112×112×64	[1×1, 32; 3×3, 32; 1×1, 128]	56×56×128
注意力模块 Attention module	A	56×56×128	A1	56×56×128
残差块 Residual unit	R	56×56×128	[1×1, 64; 3×3, 64; 1×1, 256]	28×28×256
注意力模块 Attention module	A	28×28×384	A2	28×28×384
残差块 Residual unit	R	28×28×384	[1×1, 128; 3×3, 128; 1×1, 512]	14×14×512
注意力模块 Attention module	A	14×14×768	A3×2	14×14×768
残差块 Residual unit	R	14×14×768	[1×1, 256; 3×3, 256; 1×1, 1024]	7×7×1024
残差块 Residual unit	R	7×7×1536	[1×1, 384; 3×3, 384; 1×1, 1536]	7×7×1536
池化层 Average pooling	P	7×7×1536	7×7	1×1×1536
全连接层 Fully connected layer	F	1×1×1536	1536×11	11

#### 1.4.1 构建残差块

残差块由主分支和快捷连接组成,如图 2。其中主分支由 3 层卷积层串联组成,每个卷积层都会配备一个批量归一化层和一个激活函数层,其顺序为:批量归一化层-Leaky ReLU 激活函数层-卷积层。残差注意力网络模型与 ResNet 所采用的恒等映射不同,在残差注意力网络模型的快捷连接中包含额外的卷积块,该卷积块的输入与主分支的输入相同,其输出则与主分支的输出相加后前向传播。

#### 1.4.2 构建注意力模块

根据软掩模分支中快捷连接的个数差异可将注意力模块分为 A1、A2 和 A3 三类,如图 3(网络层以大写英文字母标识,“I”代表上采样层,“S”代表 Sigmoid 激活

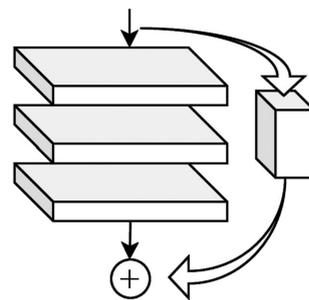


图 2 残差块

Fig.2 Residual block

层,“x”和“+”分别代表逐点相乘和逐点相加,其余缩写见表 1)。以 A1 为例,输入之后的第一个模块和输出之前的最后一个模块均为残差块,二者中间是主分支和软掩模分支。主分支由两个残差块串联组成;软掩模分支主要包含三对“下采样层-上采样层”和两个快捷连接,快捷连接的起点前和终点后皆为残差块。A2、A3 的结构与 A1 类似,两个分支的输出的乘积与主分支的输出之和作为最后一个残差块的输入向前传播。

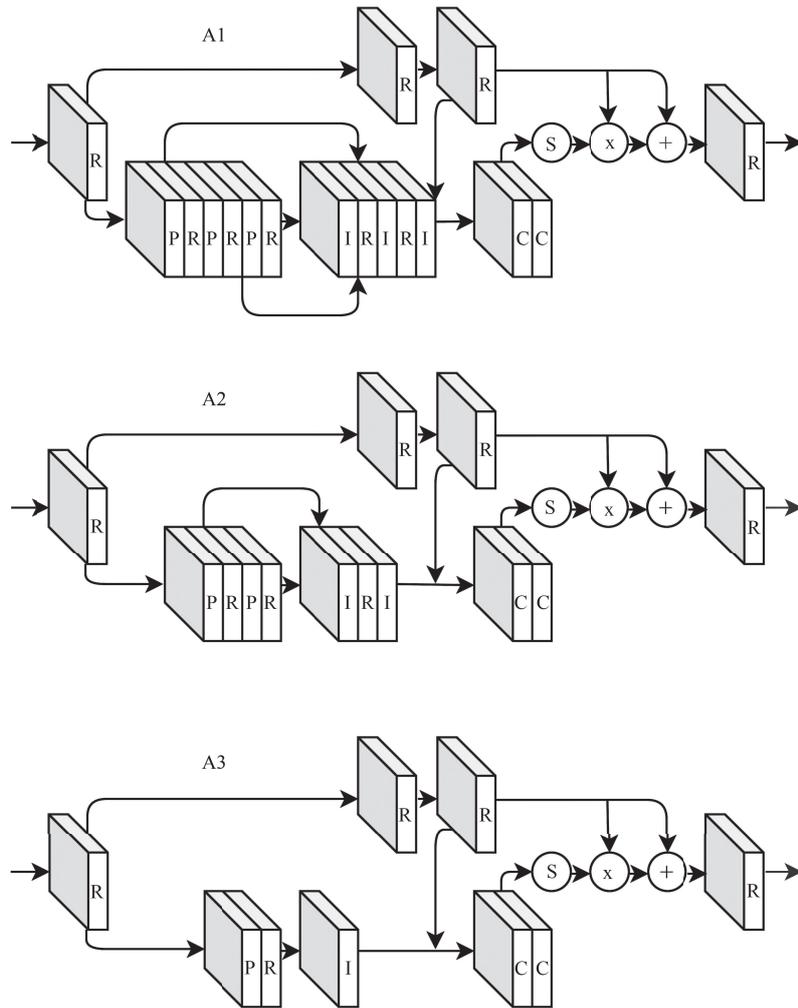


图 3 注意力模块  
Fig.3 Attention modules

### 1.4.3 构建 RAN-11

RAN-11 包含 1 个 A1 模块、1 个 A2 模块和 2 个 A3 模块,注意力模块个数多于 Attention-56 但少于 Attention-92,其网络结构见图 4。RAN-11 主要借助调整残差块参数对特征图进行维度变换。此外,对恰好位于注意力模块之后的残差块,其输出会与恰好位于注意力模块之前的残差块的输出,按通道拼接后再向前传播。

## 2 结果

对 RAN-11、Attention-56<sup>[23]</sup>、Attention-92<sup>[23]</sup>、DenseNet-161<sup>[28]</sup>、BoW<sup>[25]</sup> 五种模型在相同的计算环境下进行对比实验,实验结果见表 2。表 2 中的评价指标分为三部分:(1)模型对单一浮游植物类别的查准率<sup>[29]</sup>; (2)浮游植物的随机测试 Top-k 准确率。模型评价中常用的准确率指标即指 Top-1 准确率,表示预测正确的样本数量在总样本数量中的占比。Top-k 准确率相较于 Top-1 准确率扩大了误差容忍程度,对某一样本,只要

在预测概率为 Top-k 的类别中包含样本的正确类别,就认为该样本预测正确;(3)模型的推理速率,即该模型平均每秒可以处理多少帧图像。

由表 2 知,本文提出的算法 RAN-11 在除星杆藻外其余种类的查准率都在 90% 以上,而在其他模型中,查准率在 90% 以下的类别的数目均不少于两类。同时,RAN-11 是取得 100% 查准率的类别的数目最多的算法,为 5 种。从总体来看,RAN-11 的 Top-1、Top-2 和 Top-3 准确率均是五种算法中最高的,分别为 95.67%、98.56% 和 100%。最后从推理速率来评估,由于深度神经网络普遍具有海量参数,尽管使用 GPU 加速,其速率

还是慢于经典的词包模型。以此为代价,深度神经网络的 Top-k 准确率远高于词包模型,Top-1 准确率普遍比词包模型高 10 个百分点以上。同时,RAN-11 的推理速率为 41.5 fps,是几种深度学习模型中表现最好的。仅从残差注意力网络模型的角度来看,Attention-92 通过增加注意力模块的个数在取得优于 Attention-56 的准确率的同时,却牺牲了推理速率,帧速从 41.2 fps 下降到 23.6 fps。RAN-11 在适当增加注意力模块个数的同时引入多尺度特征和 Leaky ReLU 激活函数,使得准确率和推理速率都获得提升,特别是准确率相较于 Attention-56 提升近一个百分点。

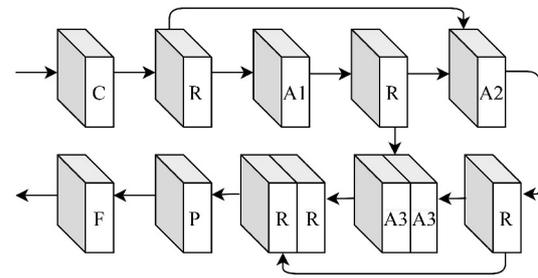


图 4 RAN-11 的网络结构

Fig.4 Network structure of RAN-11

表 2 RAN-11 和其他模型的对比实验

Table 2 Comparison of RAN-11 and other models

模型 Model	残差注意力 网络-11 RAN-11	残差注意力 网络-56 Attention-56	残差注意力 网络-92 Attention-92	稠密连接 网络-161 DenseNet-161	词包模型 BoW
裸藻 <i>Euglena</i> sp.	100.00%	95.65%	95.65%	88.00%	85.71%
转板藻 <i>Mougeotia</i> sp.	100.00%	92.00%	95.65%	92.00%	50.00%
长孢藻 <i>Dolichospermum</i> sp.	91.67%	100.00%	100.00%	100.00%	95.00%
栅藻 <i>Scenedesmus</i> sp.	95.65%	95.65%	95.83%	95.65%	82.35%
盘星藻 <i>Pediastrum</i> sp.	100.00%	100.00%	94.74%	100.00%	73.08%
微囊藻 <i>Microcystis</i> sp.	95.83%	95.83%	95.83%	100.00%	90.48%
浮丝藻 <i>Planktothrix</i> sp.	100.00%	100.00%	100.00%	100.00%	88.89%
游丝藻 <i>Planctonema</i> sp.	90.00%	90.00%	81.82%	90.00%	100.00%
星杆藻 <i>Asterionella</i> sp.	82.61%	86.36%	86.96%	78.26%	78.57%
沟链藻 <i>Aulacoseira</i> sp.	95.65%	95.65%	100.00%	91.30%	84.00%
针杆藻 <i>Synedra</i> sp.	100.00%	95.24%	100.00%	94.74%	62.50%
Top-1 准确率 Accuracy	95.67%	94.71%	95.19%	92.79%	82.69%
Top-2 准确率 Accuracy	98.56%	98.08%	98.08%	96.63%	92.79%
Top-3 准确率 Accuracy	100.00%	99.04%	99.52%	99.04%	95.67%
推理速率 Inference speed/fps	41.5	41.2	23.6	25.4	50.5

在评估算法结果的同时,本文对表 2 中 4 种深度学习算法的训练过程进行可视化分析。深度学习实验中常用回合 (Epoch) 表示训练集中全部样本完成一次前向传播和一次反向传播的过程,用批量尺寸 (Batch size) 表示神经网络每次梯度更新所处理的样本数,称神经网络每处理完一个批量的样本为一次迭代 (Iteration)。回合数和批量通常决定了一个网络要训练多久。本文将所有网络的批量均设置为 16,训练 200 个回合,并利用可视化工具 TensorBoard 记录训练过程中训练集损失随迭代次数、验证集准确率随回合数的变化情况,以确保深度学习算法正常运行。特别地,由于记录的损失数据波动较大不利于观察比较,本文先用三次样条插值

来填充离散数据点,再利用移动平均滤波器对训练损失数据进行平滑后处理。可视化结果见图 5 和图 6。

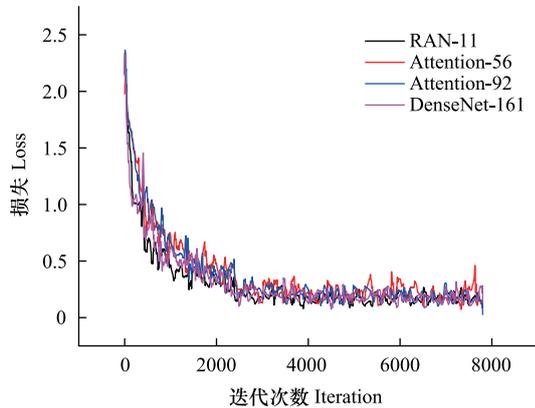


图 5 损失变化曲线

Fig.5 Variation of loss

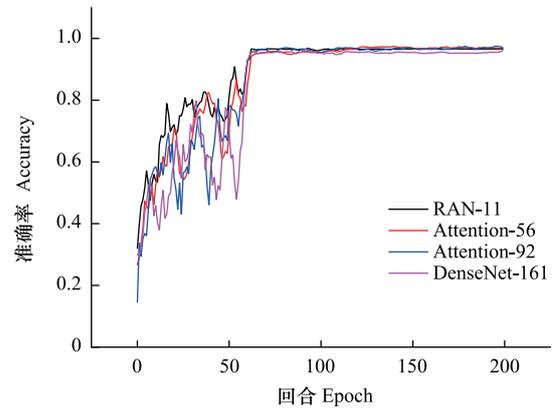


图 6 准确率变化曲线

Fig.6 Variation of accuracy

由图 5 和图 6 知,训练损失在第 3000—4000 次迭代之间便已趋于稳定,验证集准确率在第 75—100 个回合之间便已达到最高,这与现有参数设置下一个回合包含约 40 次迭代相一致。综合损失和准确率变化可知,模型在训练达 100 个回合时便可以停止更新,此时模型的准确率趋于稳定,而训练时间可以压缩掉 1/2,这有利于进行网络调参和模型快速迭代。此外,我们观察到图 5 中蓝色曲线基本位于其他曲线下侧,图 6 中蓝色曲线基本位于其他曲线上侧。这表明 RAN-11 损失下降速率最快、准确率上升速率最快,即算法收敛性好,这得益于我们采用了更好的激活函数 Leaky ReLU 来代替 ReLU,使得即使激活函数的输入值为负时神经元仍然可以更新参数。

### 3 讨论

对于大规模的浮游植物图像识别任务而言,在模型的推理速率相差不是很大的情况下,识别准确率即使提高 1% 也会带来更大实际使用价值,因此 RAN-11 对 Attention56 和 Attention92 的提升是非常有益的,但仍存在改进空间。为进一步提高识别效果,我们可以考虑从以下几方面对现有算法进行改进。

表 2 中算法对星杆藻的最高查准率为 86.96%,其中 RAN-11 算法对星杆藻的查准率仅有 82.61%。查看混淆矩阵<sup>[30]</sup>发现,星杆藻查准率低主要源于相对多的钩链藻被误识别。在部分训练集图像中,破碎的星杆藻不再呈星状,而是和钩链藻一样都呈现黄色和白色交错的链状,后者确实存在较大的被误识别为星杆藻的可能性。在后续工作中,我们将针对性地引入图像的颜色增强以丰富训练集的颜色变化,避免神经网络对这种黄白交错的形态过度学习,从而提升识别查准率和总体准确率。

相对于大型图像识别任务而言,目前浮游植物的有标注数据量还有待增加。浮游植物的显微镜图像中存在大量的相互重叠、遮挡、模糊、物种数量不均匀等现象,导致参与训练的标注图片中某些浮游植物的可鉴别的数量特别稀少,不足以支撑深度学习的训练过程。另外,目前的模型是在有监督的条件下进行的,人工标注的过程非常耗时耗力,因此后续可以考虑引入半监督或者无监督的深度学习方法,将大量的未标注图片中的有用信息加入到训练模型中,丰富样本的种类和数量。

最后,在实际应用中如何将识别过程系统化是一项重要的挑战。实验表明 RAN-11 在浮游植物识别中取得具有竞争力的识别效果,因此可以基于 RAN-11 构建出一种浮游植物自动化识别系统,如图 7。

该系统由三大模块组成:(1)图像获取模块。主要负责抽取试点收集水样,并通过显微镜及图像获取装置,获取浮游植物显微镜图像;(2)图像预处理模块。主要借助计算机软件完成浮游植物图像的切割、标注等,形成可供网络训练或测试的数据集。其中,如果目的是预测浮游植物类别则不需要标注图像。如果没有

权重则需要先标注图像并输入 RAN-11 训练;(3) 图像识别模块。加载已训练的权重,对预处理的浮游植物图像进行快速自动分类。该系统不依赖于水域环境,不需要手工设计特征,加载已有权重文件和待测试的浮游植物样本,便可以端到端的输出识别结果,具有很好的推广性。鉴于图像预处理模块的工作耗时占比大,尤其是图像标注,需要为每张图片建立标签,在后续工作中,可利用模型推断结果对数据集进行预标注后再进行人工修正,提高浮游植物图像的标注效率,逐步完善该系统。

#### 4 结论

本文以太湖浮游植物的 11 个优势属为数据来源,提出了一种新的残差注意力网络模型 RAN-11,可以以高准确率快速识别浮游植物,并构建了一种具有竞争力的浮游植物自动化识别系统,得到以下结论:

(1) RAN-11 因其精心设计的网络结构可以取得比原残差注意力网络 Attention-56 和 Attention-92 更好的识别结果,在准确率和速度上均优于原算法。

(2) 深度学习方法的识别效果远优于传统的以手工设计特征为主的图像识别方法。虽然传统方法如 BoW 计算量小、速度快,但难以弥补识别效果上的不足。

(3) 基于深度学习的浮游植物识别方法,可以准确而高效地对所关注的水体中的浮游植物进行分类,极大地提高了水质检测的可靠性和易操作性,可有力促进水质监测和保护。

(4) 当前深度学习方法在浮游植物识别领域的应用较少,但具有广阔的应用前景。本文构建的浮游植物图像识别系统,易于在实际应用中推广,对于实现水体中浮游植物的自动化监测具有重要意义。

#### 参考文献 (References):

- [ 1 ] 刘宇, 沈建忠. 藻类生物学评价在水质监测中的应用. 水利渔业, 2008, 28(4): 5-7.
- [ 2 ] Douerelo I, Perona E, Mateo P. Use of cyanobacteria to assess water quality in running waters. Environmental Pollution, 2004, 127(3): 377-384.
- [ 3 ] Salmaso S, Morabito G, Buzzi F, Garibaldi L, Simona M, Mosello R. Phytoplankton as an indicator of the water quality of the deep lakes south of the Alps. Hydrobiologia, 2006, 563(1): 167-187.
- [ 4 ] 吕笑天, 吕永龙, 宋帅, 王铁宇. 气候变化与人类活动双重驱动的冷水湖泊富营养化. 生态学报, 2017, 37(22): 7375-7386.
- [ 5 ] 李夜光, 李中奎, 耿亚红, 胡鸿钧, 殷春涛, 殷阳叶新, 桂建平. 富营养化水体中 N、P 浓度对浮游植物生长繁殖速率和生物量的影响. 生态学报, 2006, 26(2): 317-325.
- [ 6 ] 方涛, 李道季, 余立华, 高磊, 张利华. 光照和营养盐磷对微型及微微型浮游植物生长的影响. 生态学报, 2006, 26(9): 2783-2790.
- [ 7 ] 孔繁翔, 高光. 大型浅水富营养化湖泊中蓝藻水华形成机理的思考. 生态学报, 2005, 25(3): 589-595.
- [ 8 ] Huisman J, Codd G A, Paerl H W, Ibelings B W, Verspagen J M H, Visser P M. Cyanobacterial blooms. Nature Reviews Microbiology, 2018, 16(8): 471-483.
- [ 9 ] Ho J C, Michalak A M, Pahlevan N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature, 2019, 574(7780): 667-670.
- [ 10 ] Qin B Q, Zhu G W, Gao G, Zhang Y L, Li W, Paerl H W, Carmichael W W. A drinking water crisis in lake Taihu, China: linkage to climatic variability and lake management. Environmental Management, 2010, 45(1): 105-112.
- [ 11 ] Yang M, Yu J W, Li Z L, Guo Z H, Burch M, Lin T F. Taihu Lake not to blame for Wuxi's woes. Science, 2008, 319(5860): 158.
- [ 12 ] 张军毅, 庄严, 沈建荣, 江岚, 吴蔚, 罗清吉. 太湖梅梁湾菹草繁殖的原因探讨及其生态管理//2014 中国环境科学学会学术年会论文集(第五章). 成都: 中国环境科学学会, 2014: 4441-4446.
- [ 13 ] Schaap A M. Integrating Optics and Microfluidics to Automatically Identify Algae Species. Eindhoven: Technische Universiteit Eindhoven, 2015: 1-17.

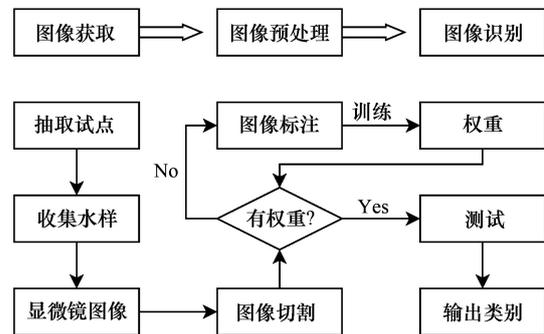


图7 浮游植物自动化识别系统

Fig.7 Automatic phytoplankton identification system

- [14] 高亚辉, 杨军霞, 骆巧琦, 高华, 杨晨辉, 李雪松, 梁君荣, 陈长平. 海洋浮游植物自动分析和识别技术. 厦门大学学报: 自然科学版, 2006, 45(S2): 40-45.
- [15] 郭小路. 三峡库区丰都段回水区营养状态及其浮游植物识别初探. 环境科学导刊, 2018, 37(2): 36-41.
- [16] 张宛宛, 谢玉为, 杨江华, 杨雅楠, 李娣, 张咏, 于红霞, 张效伟. DNA 宏条形码(metabarcoding)技术在浮游植物群落监测研究中的应用. 生态毒理学报, 2017, 12(1): 15-24.
- [17] Santhi N, Pradeepa C, Subashini P, Kalaiselvi S. Automatic identification of algal community from microscopic images. *Bioinformatics and Biology Insights*, 2013, 7: 327-334.
- [18] Thiel S U, Wiltshire R J, Davies L J. Automated object recognition of blue-green algae for measuring water quality-a preliminary study. *Water Research*, 1995, 29(10): 2398-2404.
- [19] Luo Q Q, Gao Y H, Luo J F, Chen C P, Liang J R, Yang C H. Automatic identification of diatoms with circular shape using texture analysis. *Journal of Software*, 2011, 6(3): 428-435.
- [20] Mansoor H, Sorayya M, Aishah S, Mosleh M A A. Automatic recognition system for some Cyanobacteria using image processing techniques and ANN approach//*Proceedings of 2011 International Conference on Environmental and Computer Science*. Singapore: IACSIT Press, 2011: 19.
- [21] Walker R F, Ishikawa K, Kumagai M. Fluorescence-assisted image analysis of freshwater microalgae. *Journal of Microbiological Methods*, 2002, 51(2): 149-162.
- [22] 万永清, 张奇志. 深度学习在藻类分类识别中的应用. 传感器世界, 2019, 25(1): 7-12.
- [23] Wang F, Jiang M Q, Qian C, Yang S, Li C, Zhang H G, Wang X G, Tang X O. Residual attention network for image classification//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 6450-6458.
- [24] 黄凯奇, 任伟强, 谭铁牛. 图像物体分类与检测算法综述. 计算机学报, 2014, 37(6): 1225-1240.
- [25] Zhang Y, Jin R, Zhou Z H. Understanding bag-of-words model: a statistical framework. *International Journal of Machine Learning and Cybernetics*, 2010, 1(1/4): 43-52.
- [26] He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770-778.
- [27] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models//*Proceedings of the 30th International Conference on Machine Learning*. Atlanta: IMLS, 2013, 30(1): 3.
- [28] Huang G, Liu Z, van der Maaten L, Weinberger K Q. Densely connected convolutional networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 2261-2269.
- [29] Akobeng A K. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatrica*, 2007, 96(3): 338-341.
- [30] Stehman S V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 1997, 62(1): 77-89.