

基于 CycleGAN 和注意力增强迁移学习的小样本鱼类识别

刘世晶^{1,2} 刘阳春³ 钱程¹ 郑浩君⁴ 周捷⁵ 张成林¹

(1. 中国水产科学研究院渔业机械仪器研究所, 上海 200092; 2. 中国海洋大学三亚海洋研究院, 三亚 572011;
3. 中国农业机械化科学研究院集团有限公司, 北京 100083; 4. 大连海洋大学航海与船舶工程学院, 大连 116023;
5. 上海市崇明区水产技术推广站, 上海 202150)

摘要: 围绕水产养殖水下目标精准识别的产业发展需求, 针对小样本目标识别精度低、模型算法场景适应能力差等问题, 提出一种基于改进循环对抗网络 (Cycle constraint adversarial network, CycleGAN) 样本扩增和注意力增强迁移学习的小样本养殖鱼类识别方法。利用水下采样装备收集实际养殖场景和可控养殖场景大黄鱼图像, 并以可控场景图像作为辅助样本集。利用 CycleGAN 为基础框架实现辅助样本到实际养殖场景图像的迁移, 并提出一种基于最大平均差异 (Maximum mean discrepancy, MMD) 的迁移模型损失函数优化方法。在迁移学习阶段使用 ResNet50 为基础框架, 并引入 SK-Net (Selective kernel network) 注意力机制优化模型对不同感受野目标的感知能力, 提升模型对无约束鱼类目标的识别精度。试验结果表明, 本文方法有效提升了小样本鱼类目标的识别能力, 鱼类识别召回率达到 94.33%, 平均精度均值达到 96.67%, 为鱼类行为跟踪和表型测量提供了有效的技术支撑。

关键词: 大黄鱼识别; 循环对抗网络; 注意力增强; SK-Net; 迁移学习; 注意力机制

中图分类号: S951.2 **文献标识码:** A **文章编号:** 1000-1298(2023)S1-0296-07

Recognition of Small Sample Cultured Fish Based on CycleGAN and Attention Enhanced Transfer Learning

LIU Shijing^{1,2} LIU Yangchun³ QIAN Cheng¹ ZHENG Haojun⁴ ZHOU Jie⁵ ZHANG Chenglin¹

(1. Fishery Machinery and Instrument Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200092, China
2. Sanya Oceanographic Institution, Ocean University of China, Sanya 572011, China
3. Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China
4. School of Navigation and Naval Architecture, Dalian Ocean University, Dalian 116023, China
5. Shanghai Chongming District Aquatic Technology Promotion Station, Shanghai 202150, China)

Abstract: Focusing on the industrial development needs of accurate underwater target recognition in aquaculture, and aiming at the problems of low target recognition accuracy of small samples and poor adaptability of model algorithm to scenarios, a small sample aquaculture fish recognition method based on improved cycle constraint adversarial network (CycleGAN) sample amplification and attention enhancement transfer learning was proposed. Firstly, the underwater sampling equipment was used to collect the images of the actual and controllable breeding scenes of *Larimichthys crocea*, and the controllable scene images were used as the auxiliary sample set. CycleGAN was used as the basic framework to realize the migration of auxiliary samples to the actual breeding scene images. In particular, an optimization method of the loss function of the migration model based on the maximum mean discrepancy (MMD) was proposed. Then in the transfer learning phase, ResNet50 was used as the basic framework, and SK-Net (selective kernel network) attention mechanism optimization model was introduced to improve the perception ability of different receptive field targets, so as to improve the recognition accuracy of the model for unconstrained fish targets. The experimental results showed that the method proposed effectively improved the recognition ability of fish small sample targets, with a recall rate

收稿日期: 2023-06-30 修回日期: 2023-08-20

基金项目: 青岛海洋科技中心山东省专项经费项目 (2022QNL030001-2) 和中央级公益性科研院所基本科研业务费专项资金项目 (2022XT06)

作者简介: 刘世晶 (1982—), 男, 副研究员, 主要从事渔业信息化、图像处理、模式识别和机器视觉研究, E-mail: liushijing@fmiri.ac.cn

通信作者: 周捷 (1973—), 男, 副研究员, 主要从事数字渔业相关研究, E-mail: 154812721@qq.com

of 94.33% of fish recognition, and an mAP of 96.67%, providing effective technical support for the next step of fish behavior tracking and phenotype measurement.

Key words: large yellow croaker identification; CycleGAN; attention enhanced; SK-Net; transfer learning; attention mechanism

0 引言

水产养殖作为我国重要的农业生产方式之一,为中国粮食安全做出了重大贡献^[1]。水产养殖动物精准识别对及时精准获取水产养殖动物信息、保障水产养殖安全生产具有积极作用,目前已经成为相关领域的研究重点^[2]。目前,水产养殖生物识别研究主要集中在实验室、可控水体等试验场景,相关方法无法满足实际养殖工况水体浑浊、光照条件差异以及游动姿态差异等无约束目标识别要求^[3-5]。而我国水产养殖生物种类多,养殖方式多,养殖区域广,导致不同场景采集的水下生物图像存在域分布差异,降低了图像数据之间的互用效果,限制了识别技术产业化深度应用,产生了基于小样本识别技术发展需求。近年来随着生成对抗网络^[6]以及迁徙学习^[7]的发展,为上述问题提供了可能的解决方案。

在生成对抗网络研究方面,鱼类识别领域应用较少,而在农业其它领域已经开展了相关研究。王志强等^[8]采用 WGAN (Wasserstein generative adversarial networks) 和 MCA-MobileNet (Multi-scale feature fusion and coordinate attention MobileNet) 实现 10 类番茄叶片病害识别,有效改善了番茄病害样本不足导致的识别率较低的问题。STEPHEN 等^[9]提出一种基于 IBS-DGAN (Improved backtracking search algorithm optimized deep generative adversarial network) 的水稻病害预测算法,有效解决了预测过程中出现的不稳定和过拟合问题,试验证明算法识别精度高于 XGBoost (Extreme gradient boosting) 和 SVM (Support vector machine) 方法。生成对抗方法由于其在样本扩增方面的优势已经广泛应用到多个领域,但是由于 GAN (Generative adversarial network) 样本生成过程存在随机噪声、低分辨率和低质量等问题,导致生成样本不稳定,存在模式崩溃的问题^[10]。CycleGAN^[11]由于其双向网络特征,能够解决非成对图像联合训练问题,可以实现将类似分布图像跨域联合应用,目前已经在行人识别、道路场景识别等多个领域得到广泛应用,但尚未开展鱼类识别相关研究。

在迁移学习方面,袁小平等^[12]采用 AlexNet 模型进行迁移学习,验证了模型在 ImageNet 数据集上

迁移学习的有效性。万鹏等^[13]提出一种基于改进 ResNet50 的模型和迁移学习的鱼类识别方法,实现对 6 种典型的大宗淡水鱼精准识别,平均识别率超过 96.2%。上述方法仅仅使用单一的迁移模型识别鱼类目标,虽然取得了较好的识别精度,但仍存在样本量较大,采样环境单一等问题,无法满足无约束环境条件下变场景、小样本鱼类识别需要。为了解决这个问题,ZHANG 等^[14]提出一种基于 AdvFish 的迁移学习方法,实现无约束游动鱼类识别,有效解决了水下图像的噪声背景问题;此外,袁培森等^[15]采用不同的鱼类图像的样本集,证明了基于度量的小样本学习模型结合迁移学习方法的有效性,平均识别精度达到 91.03%,召回率达到 90.78%;相比于传统的深度学习方法,迁移学习有效提升了鱼类识别效果,虽然部分方法优化了算法的环境适应性和种类适应性,但面向实际养殖工况特定鱼类样本无法有效获取的情况,方法的应用仍存在一定局限性。

针对小样本鱼类识别需求,受上述方法启发,本文提出一种基于改进 CycleGAN 和注意力增强迁移学习的小样本鱼类识别方法,并以大黄鱼 (*Larimichthys crocea*) 为试验对象,用以验证本文方法有效性。该方法采用 CycleGAN 作为数据扩增的基础模型,通过引入 MMD^[16]作为损失函数,提升模型自适应拉推域分布的能力;同时将 SK-Net^[17]注意力机制引入到 ResNet50 框架中,并利用扩增样本进行迁移学习,降低迁移学习样本数据分布不均对识别精度的影响,实现小样本鱼类识别;最后利用消融试验和对比试验验证本文方法的有效性。

1 试验材料与方法

1.1 平台与硬件环境

试验采用 Windows 11 64 位操作系统,硬件平台选用的 CPU 处理器为 Intel i7-9700 SRG13, GPU 使用 RTX A5000 显卡提供,显存为 24 GB, CUDA 并行处理核心 8 192 个。程序开发软件平台选择 Pytorch 1.0 和 Python 3.7。

1.2 数据获取

实际养殖工况鱼类图像数据采集自国信 1 号养殖工船 5 号养殖仓,采样时间为 2022 年 10 月 15 日。养殖仓直径 20 m,水深 15 m,养殖品种为大黄鱼。如图 1a 所示,采用滑轨式水下相机获取鱼类图

像,采样深度为水面下 2 m。采样时相机两侧补光灯开启,补光灯额定输出功率设定为 1 W。采用视频采样方式连续获取 120 min 水下图像,由于采样时水体较为浑浊,2 h 共采集可用大黄鱼图像 40 幅。

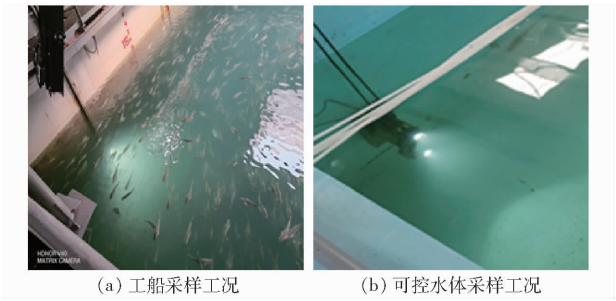


图 1 采样工况
Fig.1 Sampling conditions

可控养殖水体选择工厂化循环水养殖环境,采样时间为 2022 年 7 月。如图 1b 所示,养殖池水深 1 m,暂养不同规格大黄鱼 37 尾。采用拦截网将鱼类固定在 1 m × 4 m 活动范围内,采样相机固定在短边水深 0.5 m 处,采样时相机两侧补光灯开启,补光灯额定输出功率设定为 0.5 W。采用视频采样方式获取自由游动大黄鱼图像 100 幅,构建辅助样本集。

1.3 循环对抗网络

循环对抗网络 (CycleGAN) 是一种用于图像迁移的非对称网络,因其训练过程中不需要使用配对数据,有效降低了网络对图像数据的要求,增强了网络适用范围。循环对抗网络由 2 个生成器和 2 个判别器组成,用于将源域和目标与图像进行非对称转换,其损失函数主要由对抗性损失、循环一致性损失和身份损失 3 部分组成^[11]。其中对抗性损失的作用是迫使生成器生成逼真的图像骗过判别器,对抗性损失由两部分组成,即

$$\mathcal{L}_{GAN}(G,D_T,S,T) = E_{t \sim P_{data}(t)}(\lg D_T(t)) + E_{s \sim P_{data}(s)}(\lg(1 - D_T(G(s)))) \tag{1}$$

$$\mathcal{L}_{GAN}(F,D_S,T,S) = E_{s \sim P_{data}(s)}(\lg D_T(s)) + E_{t \sim P_{data}(t)}(\lg(1 - D_S(F(t)))) \tag{2}$$

式中 $\mathcal{L}_{GAN}$ ——对抗性损失函数

- E ——数学期望
- S ——源域
- s ——源域中图像
- T ——目标域
- t ——目标域中图像
- G ——源域到目标域的生成器
- F ——目标域到源域的生成器
- D_S ——源域判别器
- D_T ——目标域鉴别器

由于单独使用对抗性损失函数不能保证生成器将输入映射到期望的输出,存在“模式崩溃”的问

题^[18],导致生成器生成的图像丢失大量源域图像的信息,因此本文引入循环一致性损失来约束映射函数在目标域和源域周期一致性。循环一致性损失主要通过最小化目标图像与循环图像之间的差异来提高生成器保留原始图像的能力,可以有效提高图像转换的准确性,其整体结构为

$$\mathcal{L}_{cyc}(G,F) = E_{s \sim P_{data}(s)}(\|F(G(s)) - s\|_1) + E_{t \sim P_{data}(t)}(\|G(F(t)) - t\|_1) \tag{3}$$

式中 $\mathcal{L}_{cyc}$ ——循环损失函数

除了对抗性损失和循环一致性损失,身份损失同样被应用到 CycleGAN 中^[19],用以约束图像风格转移过程中输入图像和循环生成图像纹理、颜色等信息的准确性,身份损失函数 $\mathcal{L}_{ID}$ 为

$$\mathcal{L}_{ID} = E_{s \sim P_{data}(s)}(\|(G(s) - s)\|_2) + E_{t \sim P_{data}(t)}(\|(\hat{G}(t) - t)\|_2) \tag{4}$$

1.4 最大平均差异

CycleGAN 主要实现图像域间分布差异迁移,没有考虑域内分布差异对生成特征的影响。而域内分布差异可以为目标识别提供多样化参考数据,这在具有姿态特征变化的目标识别任务中十分重要。最大平均差异 (MMD) 主要通过将变量映射高维空间来度量不同随机变量分布距离,在具有特征分布差异数据相似度评价方面表现出良好的性能^[20]。因此,本研究引入 MMD 来优化循环对抗网络损失函数。

MMD 的思想是找到一个函数,该函数假设对 2 个不同的域分布的不同期望。设 F 是类函数, $S = (s_1, s_2, \dots, s_m)$ 和 $T = (t_1, t_2, \dots, t_n)$ 表示源域 (可控水体鱼类图像) 和目标域 (实际养殖水体鱼类图像), p 和 q 表示源域与目标域的分布。最大平均偏差及其经验估计定义为

$$MMD(F,x,y) = \sup_{f \in F} (E_x(f(s)) - E_y(f(t))) \tag{5}$$

$$MMD(F,S,T) = \sup_{f \in F} \left(\frac{1}{m} \sum_{i=1}^m f(s_i) - \frac{1}{n} \sum_{i=1}^n f(t_i) \right) \tag{6}$$

式中 $\sup$ ——上确界

- m,n ——源域和目标域
- f ——映射函数样本数量

设 H 是函数 f 的再生核希尔伯特空间,并且线性点评估函数映射 f 到 $f(S)$ 是存在且连续的,其中 2 个特征映射之间的内积称为核。选取高斯核为

$$k(s,s') = \exp\left(-\frac{\|s-s'\|^2}{2\sigma^2}\right) \tag{7}$$

当 F 是通用希尔伯特中的单位球时,MMD 将检测 p 和 q 之间的差异。当且仅当 $p = q$ 时,

$MMD(F, S, T) = 0$, 因此可以写为

$$MMD(F, p, q) = \| \mu_p - \mu_q \|^2 \tag{8}$$

式中 μ ——数学期望

则最大平均差异性损失为

$$\mathcal{L}_{MMD} = \left[\frac{1}{m(m-1)} \sum_{i \neq j}^m k(s_i, s_j) + \frac{1}{n(n-1)} \sum_{i \neq j}^m k(t_i, t_j) - \frac{2}{mn} \sum_{i,j=1}^{m,n} k(s_i, t_j) \right]^{\frac{1}{2}} \tag{9}$$

式中 k ——核函数

i, j ——特定域样本坐标

1.5 整体损失函数

整合 CycleGAN 和最大平均差异, 得到本文源域到目标域转换的整体目标函数为

$$\mathcal{L} = \mathcal{L}_{GAN}(G, D_T, S, T) + \mathcal{L}_{GAN}(\hat{G}, D_s, T, S) + \lambda_1 \mathcal{L}_{cyc}(G, \hat{G}) + \lambda_2 \mathcal{L}_{ID} + \lambda_3 \mathcal{L}_{MMD} \tag{10}$$

式中, λ_1 、 λ_2 和 λ_3 为不同损失函数作用权重, λ_1 和 λ_2 分别控制训练过程中周期一致性损失和前景损失的权重, λ_3 控制最大平均差异权重。

1.6 SK-Net

针对自由游动鱼类识别样本扩增问题, 由于鱼类成像距离具有随机性, 导致待识别目标感受野存在较大差异, 传统的框架性模型不具备感知范围自动调整能力。针对此问题本文引入 SK-Net 注意力机制优化模型对不同目标的识别能力。SK-Net 是通过设定不同卷积核, 并利用分割、融合和选择实现对卷积核权重的选择, 能够满足自动训练要求的网络^[21]。SK-Net 网络结构如图 2 所示。

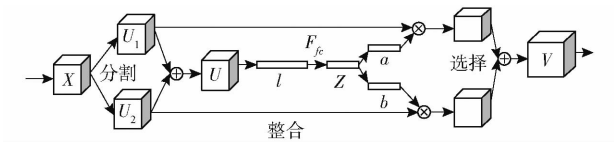


图 2 SK-Net 网络结构

Fig.2 Structure of SK-Net

分割: 针对鱼类特征图像, 选择 3×3 和 5×5 的空洞卷积核进行卷积操作, 并分别得到 2 个输出分支结果 U_1 和 U_2 。

融合: 将 U_1 和 U_2 特征图各元素按照卷积核权重对应求和得到加权图 U , 然后利用全局平均池化生成维度为 $C \times 1$ 通道统计信息 l , 并利用全连接降维生成一个 1 维紧凑特征 Z 来降低维度, Z 为

$$Z = F_{fc}(s) = \delta(\alpha(W_s)) \tag{11}$$

其中 $W_s \in \mathbf{R}^{d \times C}$ $d = \max(C/r, L)$

式中 δ ——ReLU 激活函数

α ——归一化函数

F_{fc} ——全连接层权重函数

W_s ——特征图宽度 r ——缩减比

L —— d 的最小值

为了获得不同通道不同感受野的加权融合信息, 对紧凑特征 Z 应用 softmax 算子计算卷积核权重, 表示为

$$a_c = \frac{e^{A_c Z}}{e^{A_c Z} + e^{B_c Z}} \tag{12}$$

$$b_c = \frac{e^{B_c Z}}{e^{A_c Z} + e^{B_c Z}} \tag{13}$$

其中 $A_c \in \mathbf{R}^{C \times d}$ $B_c \in \mathbf{R}^{C \times d}$

式中 a_c —— U_1 分支卷积核权重

b_c —— U_2 分支卷积核权重

最终通过权重与特征图对应乘积得到输出特征 V 为

$$V = (V_1, V_2, \dots, V_c) \tag{14}$$

其中 $V_c = a_c U_{1c} + b_c U_{2c}$

$$a_c + b_c = 1 \quad V_c \in \mathbf{R}^{H \times W}$$

1.7 SK-Net 优化 ResNet50 模型

根据 ResNet50 框架结构, 本文采用与文献[17] 相同架构, 定义如表 1 所示 SK-Net 优化 ResNet50 (SK-Net50) 框架结构。SK-Net 网络结果中需要确定 M 、 G 、 r 共 3 个超参数, 其中路径数 M 为选择不同卷积核的数量, 分组数 G_p 为每个路径的基数, 缩减比 r 控制操作器中紧凑特征 Z 的维度^[22]。在表 1 显示的结构中, SK-Net50 框架的 4 个阶段分别定义 3、4、6、3 个 SK 单元。

表 1 SK-Net 优化 ResNet50 (SK-Net50) 网络结构

Tab.1 SK-Net optimized ResNet50 (SK-Net50)

network structure	
输出	SK-Net50
256 × 256	7 × 7, 64, stride 为 2
128 × 128	3 × 3 max pool, stride 为 2
128 × 128	$\left[\begin{array}{c} 1 \times 1, 128 \\ SK[M = 2, G_p = 32, r = 16], 128 \end{array} \right] \times 3$
64 × 64	$\left[\begin{array}{c} 1 \times 1, 256 \\ SK[M = 2, G_p = 32, r = 16], 128 \end{array} \right] \times 4$
32 × 32	$\left[\begin{array}{c} 1 \times 1, 128 \\ SK[M = 2, G_p = 32, r = 16], 128 \end{array} \right] \times 6$
16 × 16	$\left[\begin{array}{c} 1 \times 1, 128 \\ SK[M = 2, G_p = 32, r = 16], 128 \end{array} \right] \times 3$
1 × 1	7 × 7 global average pool, 2 分类 fc, softmax

2 结果与讨论

2.1 模型参数与评价指标

本文模型主要包括改进 CycleGAN 和 SK-Net 优

化 ResNet50 两部分,其中 CycleGAN 主要包括超参数 λ_1 、 λ_2 和 λ_3 ,其中 $\lambda_1 = 10$, $\lambda_2 = 5$, $\lambda_3 = 0.8$, MMD 内核超参数随网络一起训练。对于 CycleGAN 生成器网络,设置鱼类图像尺寸为 256 像素 $\times$ 256 像素,网络结构采用与文献[2]中相同的设定,最大池化操作和步长分别设定为 2×2 和 2。针对鱼类识别问题,由于现场工况采样样本和可控工况采样样本鱼类体型成像大小可能存在较大差异,因此在 SK-Net 网络中,采用 3×3 和 5×5 两种卷积尺寸,其中 5×5 卷积核采用 3×3 结合缩放率为 2 的空洞卷积构成。此外,融合后特征向量维度设定为 16。针对 ResNet50 网络为主体的迁移学习网络,使用 SGD 来优化模型,SGD 动量参数统一设为 0.9,权值衰减参数设为 0.000 5,学习率设置为 0.000 2。辅助学习阶段,将全连接层的学习率设置为 0.01,批量大小(batch_size)设为 16,迭代周期(epoch)为 30。

为了评价模型鱼类识别效率,选择召回率(Recall)和平均精度均值(Mean average precision, mAP)^[23]作为模型的评价指标。其中使用召回率来评价模型对样本正确分类的能力,使用 mAP 评价模型分类的综合效果。

2.2 试验结果分析

2.2.1 数据扩增作用

本文提出的基于改进 CycleGAN 的样本扩增方法,通过扩增图像样本改善迁移学习样本分布不均的问题。在实际养殖生产过程中,由于很难构建适用于特定场景的大数据样本,经常会利用其它相似样本扩增数据集,但是由于增加样本采样环境、采样距离以及光照条件不同,扩增样本对识别率提升效果不明显。如表 2 所示,仅仅使用 ResNet50 网络直接扩增后非迁移学习样本召回率仅为 65%,mAP 为 82%,迁移学习作用在小样本上作用不明显,直接扩增样本迁移学习召回率和 mAP 分别提高 5、2 个百分点。使用 CycleGAN 优化损失函数后,由于有效拉近了源域可控环境样本与目标域实际生产环境距离,基于 ResNet50 迁移学习方法召回率达到 82%,mAP 达到 90.33%。

表 2 ResNet50 消融试验结果

方法	ResNet50		迁移学习 + ResNet50	
	召回率	mAP	召回率	mAP
不扩增	58.00	76.77	69.33	84.33
直接扩增	65.00	82.00	70.00	84.00
CycleGAN	80.67	90.33	82.00	90.33
CycleGAN + MMD	86.00	93.00	90.00	93.99

鱼类是一种典型的非刚性运动物体,具备游泳姿态的多样性,CycleGAN 经典损失函数没有考虑域间分布差异,因此,本文通过 IDE 将 MMD 引入到 CycleGAN 中,试图通过减少不同域前景之间的最大平均差异来拉近分布。从表 2 和表 3 中可以看出,使用 MMD 优化后,非迁移和迁移学习目标识别率都得到提升,基于 ResNet50 框架迁移学习召回率达到 90%,mAP 达到 93.99%,而基于 SK-Net50 注意力增强框架的迁徙学习召回率达到 94.33%,mAP 达到 96.67%。

表 3 SK-Net50 消融试验结果

方法	SK-Net50		迁移学习 + SK-Net50	
	召回率	mAP	召回率	mAP
不扩增	67.33	82.33	72.66	86.33
直接扩增	73.33	79.67	91.00	91.30
CycleGAN	86.00	90.00	91.66	90.29
CycleGAN + MMD	89.33	94.67	94.33	96.67

2.2.2 SK-Net 重要性

针对无约束游动鱼类小样本识别问题,由于鱼类游动的随机性,成像存在明显的大小差异。SK-Net 由于其具备感受野权重自适应调整能力,能够有效提升网络对不同大小目标识别效果。因此,本文在 ResNet50 框架中引入 SK-Net 注意力机制,并结合扩增样本提升小样本鱼类目标识别能力。如表 3 所示,当样本数量较少时,SK-Net50 对识别率提升贡献较少。相比于仅使用 ResNet50 网络,经过 SK-Net 注意力增强后直接扩增、使用 CycleGAN 扩增以及 CycleGAN + MMD 扩增样本的识别率均得到一定程度的提升,使用本文方法 SK-Net50 的召回率达到 89.33%,mAP 达到 94.67%。

2.2.3 损失函数权重

参数 λ_1 、 λ_2 和 λ_3 分别控制 3 个目标损失的权重。本文评估了它们对图像迁移的影响。 λ_1 和 λ_2 是 CycleGAN 中的原始参数,文献[11]已经证明参数 $\lambda_1 = 10$ 和 $\lambda_2 = 5$ 为最优选择。本文增加了最大平均差异损失函数,而 λ_3 是控制 MMD 损耗权重的关键参数,因此本节主要比较 λ_3 的敏感性,比较结果如表 4 所示。很明显,与 $\lambda_3 = 0$ 的情况相比,增加 MMD 损失函数对提升识别精度有效。从表 4 可以看出 MMD 损失在图像扩增过程中有积极表现,且当权重较小时,对识别效果影响较明显,如果权重较大,识别效果随权重变化缓慢。因此针对不同数据集,由于数据质量和域分布差异不同, λ_3 的值应该仔细选择。

表 4 不同 λ_3 值对迁移学习 + SK-Net50 识别精度的影响

Tab.4 Effect of different λ_3 values on recognition accuracy

recognition accuracy		%
λ_3	召回率	mAP
0	91.66	90.29
0.4	92.50	94.33
0.8	94.33	96.67
1.2	94.16	96.55
1.6	93.93	96.49

2.2.4 迁移学习的作用

从表 2、3 可以看出,在鱼类识别精度方面,不扩增样本迁移学习后识别精度提升有限,ResNet50 的召回率为 69.33%,mAP 为 84.33%。这主要是由于迁移学习样本存在数据分布差异,当原始样本数量较少时,迁移学习过程不能将已学习特征和目标特征有效关联,降低了鱼类迁移学习效果。联合比较鱼类识别精度和整体识别精度,可以发现,虽然小样本召回率较低,但是 mAP 较高,证明背景识别率较高,说明迁移学习过程有效提取了背景特征。样本扩增后,迁移学习有效提升整体识别精度,特别需要注意的是相比于不扩增样本背景识别效果,由于前景鱼类数据量增大,迁移学习能够有效提升已学习样本的利用率,SK-Net50 框架召回率达到 94.33%,mAP 达到 96.67%,体现了数据扩增、注意力增强以及迁移学习融合的有效性。

2.3 与其它方法比较

为了验证本文方法的有效性,利用本文数据选择目前最先进图像扩增和注意力机制方法进行比较。在图像扩增方面选择 PTGAN^[24]和 CamStyle^[25](Camera Style)方法结合本文的注意力机制优化的 ResNet50 网络进行迁移学习;在注意力机制方面,选择 Squeeze-and-excitation networks (SE-Net)^[26]和 Convolutional block attention module (CBAM)^[27]结合本文数据扩增方法进行迁移学习。试验结果如表 5 所示,其中,PTGAN 通过提取鱼类前景直接对

目标域进行标注,只考虑数据集之间的域差异,没有考虑柔性变化带来的身份损失,识别结果较差;CamStyle 主要使用标签使得正则化变得平滑,以此来降低可能由噪声引起的过拟合风险,但由于同样没有考虑身份损失导致对鱼类目标的识别精度降低;SE-Net 通过学习通道之间的相关性,筛选出了针对通道的注意力,没有考虑目标成像大小变化带来的影响,降低了自由游动目标的识别精度;CBAM 具备通道和空间双重注意机制,来分别在通道和空间维度上优化学习效果,但是由于其中间注意力机制主要关注空间位置对识别效果影响,在存在较多大小变化的目标识别场景时,该算法识别效果不佳。

表 5 不同方法比较结果

Tab.5 Comparison results of different methods

		%
方法	召回率	mAP
PTGAN + SK-Net50	63.00	81.33
CamStyle + SK-Net50	70.67	88.33
CycleGAN + MMD + SE-Net + ResNet50	88.67	92.33
CycleGAN + MMD + CBAM + ResNet50	89.33	94.67
本文方法	94.33	96.67

3 结束语

提出了一种基于改进 CycleGAN 样本扩增结合 SK-Net 注意力机制迁移学习的无约束养殖场景鱼类识别方法。无约束可变场景识别任务仍然面临许多问题,例如,实际生产场景难以获取大量的学习样本,柔性目标数据互用存在域内分布差异等。为了解决这个问题,本文在 CycleGAN 中引入 MMD 损失函数,提升模型对柔性目标特征的识别能力,引入 SK-Net 注意力机制提升模型对不同大小目标的识别能力。同时为了提升技术的实用性,采用迁移学习方法提高样本识别准确率。进行了大量试验,结果验证了本文方法的有效性。与最先进的方法相比,改进的 CycleGAN 方法可以通过简单的框架获得具有竞争力的性能。

参 考 文 献

[1] 刘永新,邵长伟,侯吉伦,等. 中国水产育种研究现状与发展建议[J]. 水产学报, 2023, 47(1): 56–69.
LIU Yongxin, SHAO Changwei, HOU Jilun, et al. Research status and development suggestion of China's aquaculture breeding [J]. Journal of Fisheries, 2023, 47(1): 56–69. (in Chinese)

[2] SANDFORD M, CASTILLO G, HUNG T. A review of fish identification methods applied on small fish[J]. Reviews in Aquaculture, 2020, 12(2): 542–554.

[3] 罗毅,王伟,刘勇,等. 基于 One-Class SVM 的青鳉鱼异常行为识别方法[J]. 河北工业科技, 2022, 39(3): 230–236.
LOU Yi, WANG Wei, LIU Yong, et al. Abnormal behavior recognition method of medaka based on One-Class SVM[J]. Hebei Journal of Industrial Science and Technology, 2022, 39(3): 230–236. (in Chinese)

[4] 陈英义,龚川洋,刘烨琦,等. 基于 FTVGG16 卷积神经网络的鱼类识别方法[J]. 农业机械学报, 2019, 50(5): 223–231.
CHEN Yingyi, LONG Chuanyang, LIU Yeqi, et al. Fish identification method based on FTVGG16 convolutional neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 223–231. (in Chinese)

[5] BARBEDO J. A review on the use of computer vision and artificial intelligence for fish recognition, monitoring, and management

- [J]. *Fishes*, 2022, 7(6): 335.
- [6] 马春光, 郭瑶瑶, 武朋, 等. 生成式对抗网络图像增强研究综述[J]. *信息网络安全*, 2019(5): 10–21.
MA Chunguang, GUO Yaoyao, WU Peng, et al. Review of image enhancement based on generative adversarial networks[J]. *Netinfo. Security*, 2019(5): 10–21. (in Chinese)
- [7] 庄福振, 罗平, 何清, 等. 迁移学习研究进展[J]. *软件学报*, 2015, 26(1): 26–39.
ZHUANG Fuzhen, LUO Ping, HE Qing, et al. Survey on transfer learning research[J]. *Journal of Software*, 2015, 26(1): 26–39. (in Chinese)
- [8] 王志强, 于雪莹, 杨晓婧, 等. 基于 WGAN 和 MCA-MobileNet 的番茄叶片病害识别[J]. *农业机械学报*, 2023, 54(5): 244–252.
WANG Zhiqiang, YU Xueying, YANG Xiaojing, et al. Tomato leaf diseases recognition based on WGAN and MCA-MobileNet [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(5): 244–252. (in Chinese)
- [9] STEPHEN A, PUNITHA A, CHANDRASEKAR A. Optimal deep generative adversarial network and convolutional neural network for rice leaf disease prediction[J]. *The Visual Computer*, 2023, 39:1–18.
- [10] DENG Weijian, ZHENG Liang, KANG Guoliang, et al. Image-image domain adaptation with preserved self-similarity and domain-dissimilarity for person re-identification[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 994–1003.
- [11] ZHU Junyan, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C] // *Proceedings of the IEEE International Conference on Computer Vision*, 2017: 2223–2232.
- [12] 袁小平, 石慧. 基于迁移学习的鱼类识别方法研究[J]. *计算机技术与发展*, 2021, 31(4): 52–56.
YUAN Xiaoping, SHI Hui. Research on fish recognition method based on transfer learning[J]. *Computer Technology and Development*, 2021, 31(4): 52–56. (in Chinese)
- [13] 万鹏, 赵峻威, 朱明, 等. 基于改进 ResNet50 模型的大宗淡水鱼种类识别方法[J]. *农业工程学报*, 2021, 37(12): 159–168.
WAN Peng, ZHAO Junwei, ZHU Ming, et al. Freshwater fish species identification method based on improved ResNet50 model[J]. *Transactions of the CSAE*, 2021, 37(12): 159–168. (in Chinese)
- [14] ZHANG Zhixue, DU Xiujuan, JIN Long, et al. Large-scale underwater fish recognition via deep adversarial learning[J]. *Knowledge and Information Systems*, 2022, 64(2): 353–379.
- [15] 袁培森, 宋进, 徐焕良. 基于残差网络和小样本学习的鱼图像识别[J]. *农业机械学报*, 2022, 53(2): 282–290.
YUAN Peisen, SONG Jin, XU Huanliang. Fish image recognition based on residual network and few-shot learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(2): 282–290. (in Chinese)
- [16] XIA Pengfei, NIU Hongjing, LI Ziqiang, et al. Enhancing backdoor attacks with multi-level MMD regularization[J]. *IEEE Transactions on Dependable and Secure Computing*, 2023, 20(2): 1675–1686.
- [17] LI Xiang, WANG Wenhai, HU Xiaolin, et al. Selective kernel networks[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 510–519.
- [18] GALVAO Y M, PORTELA L, BARROS P, et al. OneFall-GAN: a one-class GAN framework applied to fall detection[J]. *Engineering Science and Technology*, 2022, 35: 101227.
- [19] 刘华超, 张俊然, 刘云飞. 引入特征损失对 CycleGAN 的影响研究[J]. *计算机工程与应用*, 2020, 56(22): 217–223.
LIU Huachao, ZHANG Junran, LIU Yunfei. Influence of identity loss on CycleGAN [J]. *Computer Engineering and Applications*, 2020, 56(22): 217–223. (in Chinese)
- [20] REVATHI R, ANGEL D, ANNAMMAI I. MMD labeling of EASS of cartesian product of two graphs[J]. *Opsearch*, 2023, 60(2): 870–876.
- [21] 王美华, 何海君, 李超. 自适应卷积的残差修正单幅图像去雨[J]. *中国图象图形学报*, 2020, 25(12): 2484–2493.
WANG Meihua, HE Haijun, LI Chao. Single image rain removal based on selective kernel convolution using a residual refine factor[J]. *Journal of Image and Graphics*, 2020, 25(12): 2484–2493. (in Chinese)
- [22] 金彬峰, 许光宇, 于臻, 等. 基于改进 MSR 和 YOLO v5 的雾天行人检测算法[J]. *湖北民族大学学报(自然科学版)*, 2023, 41(1): 58–64.
JIN Binfeng, XU Guangyu, YU Li, et al. Pedestrian detection algorithm in hazy weather based on improved MSR and YOLO v5[J]. *Journal of Hubei Minzu University(Natural Sciences Edition)*, 2023, 41(1): 58–64. (in Chinese)
- [23] 刘世晶, 涂雪滢, 钱程, 等. 基于帧间光流特征和改进 RNN 的草鱼摄食状态分类[J]. *水生生物学报*, 2022, 46(6): 914–921.
LIU Shijing, TU Xueying, QIAN Cheng, et al. Feeding state classification of grass carp based on optical flow and improved RNN[J]. *Acta Hydrobiologica Sinica*, 2022, 46(6): 914–921. (in Chinese)
- [24] WEI Longhui, ZHANG Shiliang, GAO Wen, et al. Person transfer GAN to bridge domain gap for person re-identification[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 79–88.
- [25] ZHONG Zhun, ZHENG Liang, ZHENG Zhedong, et al. Camera style adaptation for person re-identification[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 5157–5166.
- [26] 杜海顺, 张春海, 安文昊, 等. 基于多层信息融合和显著性特征增强的农作物病害识别[J]. *农业机械学报*, 2023, 54(7): 214–222.
DU Haishun, ZHANG Chunhai, AN Wenhao, et al. Crop disease recognition based on multi-layer information fusion and saliency feature enhancement[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(7): 214–222. (in Chinese)
- [27] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3–19.