

基于残差网络和小样本学习的鱼图像识别

袁培森 宋 进 徐焕良
(南京农业大学人工智能学院, 南京 210095)

摘要: 针对鱼种类多、数据采集难度大,且需要细粒度图像识别等问题,提出了一种基于度量学习的小样本学习方法。采用基于度量学习的小样本学习网络以及 ResNet18 的残差块结构,提取鱼图像的深层次特征,并将其映射至嵌入空间进而在嵌入空间判断鱼的种类。为了进一步提升识别准确率,利用小样本学习模型在 mini-ImageNet 数据集进行预训练,然后将训练的结果迁移到 Fish100 细粒度数据集上进行精细化训练,得到最终鱼图像识别的判别模型。使用本文模型与常用的 5 种小样本学习模型,在鱼图像数据集 Fish100 和 ImageNet 上进行对比试验,结果表明本文模型的识别效果最佳,在两个数据集上的识别精度分别达到了 94.77% 和 91.03%,且精度、召回率和 F1 值均明显优于其它模型。

关键词: 鱼; 图像识别; 残差网络; 小样本学习; 迁移学习

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2022)02-0282-09 OSID: 

Fish Image Recognition Based on Residual Network and Few-shot Learning

YUAN Peisen SONG Jin XU Huanliang
(College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: Accurate and effective identification of fish images play important role in the observation of fish populations and the management of the ecological environment. However, there were some issues, such as lots of kinds of fish, difficulty of data collection in the complex environments, and fine-grained fish image recognition. For solving the problem of few image annotation of fish image, a few-shot learning method based on metric learning was proposed. Firstly, the residual block structure of ResNet18 was used to improve the few-shot learning network based on metric learning, for extracting the deep features of fish images, and then they were mapped to the embedding space for obtaining the mean center by clustering skills. Secondly, for further improving the recognition accuracy, the improved few-shot learning model was used for pre-training on the mini-ImageNet dataset, and then the training results were transferred to the Fish100 fine-grained dataset for fine-grained training to get the final discrimination model. Based on this model, comparative experiments were conducted with the existing five few-shot learning models on the fish data set Fish100 and ImageNet. The results showed that the model proposed had the best recognition effect and the recognition accuracy on the two datasets reached 94.77% and 91.03%, respectively, and the accuracy, recall rate, and F1 were significantly better than that of other models. The experiments showed that the method proposed can effectively improve the accuracy of few-shot learning in fish identification with few annotated fish images, which can provide technical support and reference for the application of practical fish image recognition.

Key words: fish; image recognition; residual network; few-shot learning; transfer learning

0 引言

鱼在全球各水域分布广泛,对人类的生产、生活影响极为重要。对鱼类图像数据的识别研究可对鱼种群的观测及其栖息地生态环境的治理起到重要作用,在环境保护、学术研究以及经济生产方面,均有

收稿日期: 2021-02-22 修回日期: 2021-05-14
基金项目: 国家自然科学基金项目(61502236、61806097)和江苏省农业科技自主创新资金项目(SCX(21)3059)
作者简介: 袁培森(1980—),男,副教授,博士,主要从事智能信息处理与海量数据处理研究,E-mail: peiseny@njau.edu.cn
通信作者: 徐焕良(1963—),男,教授,博士生导师,主要从事农业信息化与大数据技术研究,E-mail: huanliangxu@njau.edu.cn

着重大意义^[1]。

近年来,得益于海量数据标注和计算能力的提升,利用深度神经网络进行图像识别,在各领域的研究取得了重大的进展和突破^[2]。然而,基于鱼图像的鱼种类识别进展缓慢。由于水底环境光线较弱,鱼图像采集难度大,获得的鱼图像标注数据集数量较少^[3],无法满足深度神经网络训练所需的大数据集要求。因此,基于小样本学习(Few-shot learning)的方法^[4]仅通过少量样本学习,可以用于鱼图像标注少情况下的识别研究。

最早的小样本学习基于贝叶斯框架^[5]对视觉对象进行学习。由于此方法采用传统学习技术,自动学习能力较弱。研究者提出了语义迁移的方式解决小样本识别问题,MENSINK 等^[6]通过各类别的语义,使用度量学习方法来让模型能够在类别变换时保持稳定。ROHRBACH 等^[7]采用直推式学习,通过构建已知类别空间分布的方式,对未知类别的表示进行预测。这些方法能够让网络模型自动进行学习,但由于需要人为添加语义描述,使用的便捷性仍然有限。SANTORO 等^[8]首次提出了采用外部存储的记忆增强方法,其通过存储部分输入数据的方式,让网络模型将不同时刻的输入建立联系,以学习识别的共性过程。SNELL 等^[9]提出的原型网络方法,通过学习将输入数据映射到嵌入空间,建立类别的均值中心,以便在面对新的数据时通过距离度量判断其所属类别。SUNG 等^[10]所提出的关系网络在原型网络的基础上,将距离度量方式也作为网络学习的内容,以便通过最佳的距离度量方式进行类别判断。CHOI 等^[11]通过让网络学习不同的调节器,使其能够学会不同领域内、差距较大类的小样本识别能力。TSENG 等^[12]在训练阶段利用仿射变换增强图像的特征,模拟不同领域下的特征分布,从而让模型的跨域能力进一步增强。ZHANG 等^[13]提出了一种自适应的网络模型 AdarGCN,以便在样本量不足的情况下,在网络上爬取数据后进行有效地降噪处理,完成样本量更为稀少的小样本图像识别任务。

目前,小样本学习已经被广泛地应用于标注数据稀少的图像识别、目标检测和自然语言处理等领域^[14-15],陈英义等^[16]构建了 FTVGG16 卷积神经网络提高复杂应用场景鱼类目标的识别精度,本文主要针对鱼图像标注稀少情况下的识别质量问题。

为提升小样本情况下的鱼图像准确识别率,本文基于度量学习的小样本学习方法,采用残差块结构作为鱼图像样本深层特征提取器,并将其映射至嵌入空间,形成各类别的均值中心,计算样本与均值

中心的距离,实现鱼图像识别。利用小样本学习和残差网络在 mini-ImageNet 数据集上进行训练,得到识别鱼的初步模型。为准确识别细粒度鱼图像,将前一阶段得到的网络模型利用迁移学习技术^[17],在 Fish100 数据集上进行重新训练,最终得到小样本鱼图像识别模型。为验证本文方法的可行性,运用 5 类常用的小样本学习方法在 Fish100 数据集和 ImageNet 数据集上进行对比分析。

1 试验材料

选用 mini-ImageNet、Fish100 作为试验数据集,数据集 ImageNet 用于测试。

1.1 mini-ImageNet

mini-ImageNet^[18]作为小样本图像识别中常用数据集,选用 ImageNet 数据集中的 100 个类别,包含鱼、鸟等类。文中利用 mini-ImageNet 数据集对小样本学习模型进行预训练,数据集划分如表 1 所示。

表 1 mini-ImageNet 数据集划分
Tab.1 Partition of mini-ImageNet dataset

数据集	类别数	图像数量/幅
训练集	64	38 400
验证集	16	9 600
测试集	20	12 000
总计	100	60 000

mini-ImageNet 数据集中的示例如图 1 所示。



图 1 mini-ImageNet 部分图像示例

Fig.1 Samples of mini-ImageNet's images

1.2 Fish100

Fish100^[19-20]数据集是深度学习图像标注数据集 Image CLEF 中 Marine animal species 的一个子集,包含 100 种鱼,共计 6 358 幅图像。本文所采用的 Fish100 数据集划分详情如表 2 所示。图 2 为 Fish100 数据集中部分图像示例。

1.3 ImageNet

ImageNet^[21]是一个面向机器视觉的大型可视

表 2 Fish100 数据集划分
Tab.2 Division of Fish100 dataset

数据集	类别数	图像数量/幅
训练集	50	3 320
验证集	30	1 917
测试集	20	1 121
总计	100	6 358

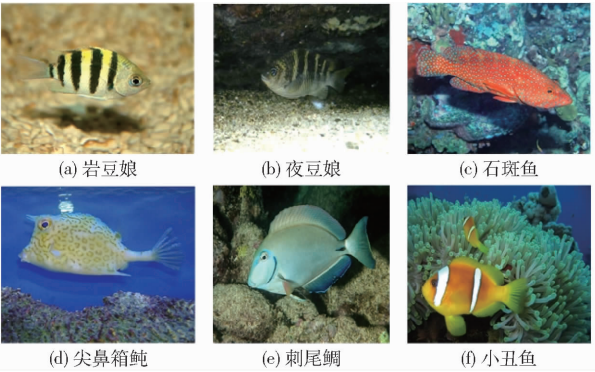


图 2 Fish100 图像示例
Fig.2 Sample of Fish100's images

化数据集,拥有共计超过 1 400 万幅图像,是深度学习领域最为常用的数据集之一,其中包含多种鱼图像数据。

ImageNet 数据集除存在与 Fish100 相同的特点,即类内差异性大、类间相似性高之外,还存在一些背景干扰因素等情况,增加了识别难度。为了检验本文方法的识别能力,本文选用 ImageNet 数据集中的 20 种鱼进行测试,具体如表 3 所示。

表 3 ImageNet 数据集
Tab.3 ImageNet dataset

编号	鱼类名称	图像数量/幅
1	<i>Ictalurus punctatus</i>	647
2	<i>Aulostomus maculatus</i>	693
3	<i>Coreius septentrionalis</i>	122
4	<i>Torpedo nobiliana</i>	460
5	<i>Histiopertus typus</i>	515
6	<i>Zeus faber</i>	560
7	<i>Clupea pallasii</i>	215
8	<i>Diodon nictemerus</i>	115
9	<i>Oncorhynchus mykiss</i>	257
10	<i>Lota lota</i>	612
11	<i>Thunnus thynnus</i>	791
12	<i>Anabas testudineus</i>	156
13	<i>Histiophorus orientalis</i>	258
14	<i>Cololabis saira</i>	267
15	<i>Clupea harengus</i>	356
16	<i>Sphyrænus</i>	101
17	<i>Morone saxatilis</i>	415
18	<i>Fistularia commersonii</i>	552
19	<i>Oncorhynchus kisutch</i>	111
20	<i>Echeneis naucrates</i>	215
总计		7 418

图 3 为 ImageNet 数据集中部分图像示例。



图 3 ImageNet 图像示例
Fig.3 Sample of ImageNet's images

2 试验方法

2.1 数据预处理

预处理操作分为裁剪、格式转换、图像增强等,对于卷积神经网络 (Convolutional neural networks, CNN) 而言,输入图像必须被调整为统一尺寸。本文图像预处理步骤为:①将所有图像转换为 3 通道的 RGB 图像。②将所有图像的尺寸统一调整为 224 像素 × 224 像素。③将调整后的图像以中心为原点,进行随机裁剪。④对所有图像进行正则化,降低网络模型过拟合的可能性。

2.2 网络模型

本文的小样本学习采用基于度量的小样本学习模型,其在面对输入样本时,首先会通过网络模型将其映射至嵌入空间,在此过程中,会通过残差块结构进行特征提取,得到特征向量;随后,通过求平均值的方式,得出类别的均值中心。最后通过度量来进行类别判断。

本文模型共包含 4 个基本单元,即块 (Block),每个块内均包含 1 个残差块 (内含 4 个卷积层)、1 个 Batchnorm 层、激活层和 1 个最大池化层,其整体结构如图 4 所示。

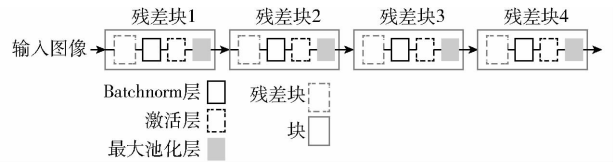


图 4 本文网络模型整体结构

Fig.4 Overall structure of network model

传统的 CNN 网络直接通过训练,学习输入 x 与输出 $H(x)$ 之间的关系,即 $x \rightarrow H(x)$,其中 x 为输入数据, $H(x)$ 为经过网络层的输出。而残差块则致力于使用图 4 所示的有参网络层,来学习输入、输出之间的残差,其中残差块结构^[22-23]如图 5 所示。其计算公式为

$$F(x, W_i) = W_2 \sigma(W_1 x) \tag{1}$$

式中 $W_1、W_2$ ——有参卷积层的权重
 σ ——ReLU 函数
 F ——残差映射函数
 W_i ——第 i 层卷积层的权重矩阵

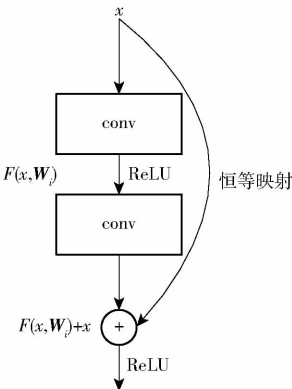


图 5 残差块结构图

Fig.5 Structure of residual block

则输入 x 与输出 $H(x)$ 之间的关系就变为 $x \rightarrow F(x, W_i) + x_0$ 。

由于输入与残差之间的关系较其与输出之间的关系更易表示,因此具有该结构的网络模型,具有更强的图像特征提取能力。

模型中各残差块中卷积层的卷积核数目 (Num output) 如表 4 所示。

表 4 网络模型各卷积层的卷积核数目
Tab.4 Number output of each convolution layers of network model

残差块	卷积层	卷积核数目
残差块 1	卷积层 1	64
	卷积层 2	64
	卷积层 3	64
	卷积层 4	64
残差块 2	卷积层 1	128
	卷积层 2	128
	卷积层 3	128
	卷积层 4	128
残差块 3	卷积层 1	256
	卷积层 2	256
	卷积层 3	256
	卷积层 4	256
残差块 4	卷积层 1	512
	卷积层 2	512
	卷积层 3	512
	卷积层 4	512

模型的其它参数如表 5 所示。

2.3 识别过程

在每次元任务中,采集 $N(K + Q)$ 个数据,其中 N 为类别数量, K 为支撑集样本数量, Q 为验证集样

表 5 网络模型参数

Tab.5 Parameters of network model

参数	数值
卷积核尺寸	3
卷积层步长	2
防止除零参数 ε	1×10^{-5}
动量	0.5
仿射变换参数 γ	1
仿射变换参数 β	0
池化层核尺寸	2
池化层步长	2

本数量。采集完成并进行预处理之后,将支撑集中的样本作为输入图像输入网络模型,在每一个块内通过残差块进行特征提取。当输入图像通过每个卷积层时,其通道数 c 、宽度 w 和高度 h 都会根据网络参数而变化。

$$c = o \tag{2}$$

$$w = (w_0 + 2p - k) / s + 1 \tag{3}$$

$$h = (h_0 + 2p - k) / s + 1 \tag{4}$$

式中 o ——卷积核数量
 p ——边缘填充尺寸
 k ——卷积核尺寸
 s ——步长
 w_0 ——输入图像宽度
 h_0 ——输入图像高度

随后,通过 BatchNorm 层对数据进行归一化处理,公式为

$$y = \frac{x - E(x)}{\sqrt{\text{Var}(x) + \varepsilon}} \gamma + \beta \tag{5}$$

式中 $E(\cdot)$ ——均值函数
 $\text{Var}(\cdot)$ ——方差函数
 y ——归一化值

之后通过激活层,使用 ReLU 函数,通过引入非线性因素的方式,增强神经网络对模型的表达能力。

$$\text{ReLU}(x) = \max(0, x) \tag{6}$$

随后通过最大池化层,对邻域特征点取最大值的方式,对提取到的特征向量进行过滤,降低特征提取的误差。图 6 为最大池化的示意图,特征矩阵中 4 个邻域内分别有 4 个特征点,每个邻域中的最大特征点分别为 7、9、9、8,即通过最大池化层后,每个邻域内只有最大的这 4 个特征点被保留。

池化层的运算方法和卷积层基本相同,只是不再改变图像的通道数。再依次通过网络模型的 4 个块后,输入数据将从原来的 D 维变为 N 维,即被映射至 N 维的嵌入空间。映射完成后,通过将映射完成的各类别所有样本数据取均值,形成各类别的均值中心。

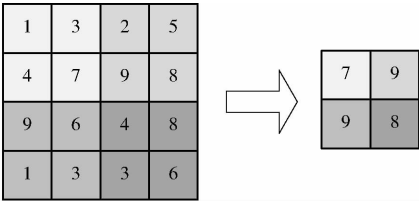


图 6 最大池化示意图
Fig. 6 Illustration of max pooling

$$c_k = \frac{1}{|S_k|} \sum_{(x_i, y_i)} f_{\phi}(x_i) \tag{7}$$

式中 S_k ——类别 k 在支撑集中样本的特征向量
 f_{ϕ} ——嵌入函数
 c_k ——类别 k 的均值中心
 x_i, y_i ——支撑集中第 i 个样本及该样本所属类别

在确定每个类的均值中心后,将验证集数据样本通过嵌入函数映射到嵌入空间,由于训练当中验证集的数据已经标记了类别,将通过其与各类别均值中心的欧氏距离得出其属于其自身类别的概率 P_{ϕ} 为

$$P_{\phi}(y = k | x) = \frac{\exp(-d(f_{\phi}(x), c_k))}{\sum_{k'} \exp(-d(f_{\phi}(x), c_{k'}))} \tag{8}$$

式中 d ——欧氏距离函数
 $c_{k'}$ ——类别 k' 的均值中心

通过在训练过程中重复上述类别判断过程,网络模型将不断地进行优化,这使得同类别样本在映射到嵌入空间后,处于更为相近的位置,由此所得出的均值中心,也将更能代表类别的真实位置。

测试过程如图 7 所示。验证集的样本未标记类别,嵌入函数在通过支撑集中的样本构建类别均值中心后,将验证集中的样本映射到嵌入空间,通过将其与各均值中心进行距离度量,给出其属于各类别的概率,概率最高者即为该样本的预测类别^[24]。

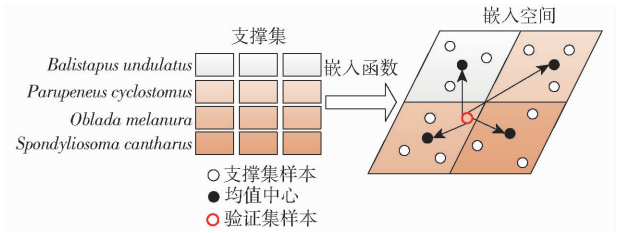


图 7 类别判断过程示意图
Fig. 7 Category recognition process of samples

2.4 迁移学习

迁移学习技术已广泛应用于深层网络模型设计和参数训练,本文的迁移学习过程如图 8 所示。

采用小样本学习中常用的数据集 mini-ImageNet 对模型进行预训练,预训练过程与训练和测试过程保持一致,同样被分解为多个元任务,每个元任务中

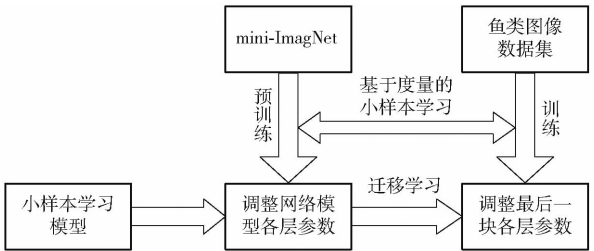


图 8 迁移学习过程
Fig. 8 Transfer learning process

抽取包含 N 个类别的共计 $N(K + Q)$ 个样本,让网络模型学习如何将这这些样本通过距离度量的方式判断类别。在预训练过程中,使用特定的优化器,对网络各层参数,即卷积层的权重 W 以及 Batchnorm 层的 γ 和 β 进行调整。

将预训练后的模型使用鱼图像数据集进行训练,并微调网络最后一个块中的各层。由于本文所设计的模型不通过全连接层输出类别,因此无需进行全连接层的调整。

3 结果与分析

3.1 试验环境

试验平台为 Windows 10,64 位操作系统,CPU 为 i5 - 8300H,2.30 GHz,GPU 为 NVIDIA Geforce GTX 1080ti,5 GB,内存为 16.0 GB。

3.2 试验参数与评价标准

本文中的预训练及训练阶段的试验参数设置如表 6 所示。优化器选用 Adam^[25]。

表 6 试验参数设置	
Tab. 6 Parameters setting of experiment	
参数	数值
每次采样类别数	5
各类别支撑集样本数	15
各类别验证集样本数	15
批量大小	20
最大迭代次数	1 000
图像尺寸/(像素×像素)	224 × 224
图像通道数	3

每次采样抽取 5 个类别,即 way 为 5;每个类别中 15 个样本作为支撑集,即 shot 为 15;15 个样本作为测试集,即 query 为 15,每次采样共抽取 5 × (15 + 15) 幅图像。每次元任务共进行 20 次采样,最大迭代次数为 1 000 次。

本文采用精度、召回率以及 F1 值来衡量模型识别效果,其中精度衡量模型识别的准确性,召回率衡量模型的查全能力,F1 值衡量模型的综合性能。

3.3 模型训练损失值对比

将本文模型与 MAML^[26]、Meta-baseline^[27]、

Meta-learning LSTM^[28]、Prototypical network^[9] 以及 Relation network^[10] 等 5 种常用小样本学习模型在训练阶段的损失值进行对比,结果如图 9 所示。

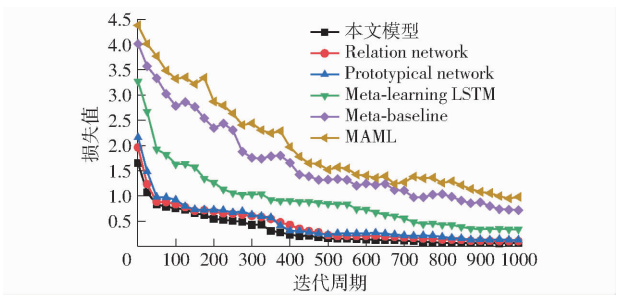


图 9 不同模型训练损失值

Fig. 9 Training loss of different models

从图 9 可以看出,本文模型不仅初始损失值较低,且波动较小,损失值始终低于其它各模型,最终收敛值也最小,为 0.06 左右。Relation network 以及 Prototypical network 两种基于度量的小样本学习方法的表现虽不如本文模型,但总体也较好,整体均呈明显的下降趋势,虽然波动较本文方法更大,但最终收敛到的损失值较为理想,分别为 0.10 和 0.12 左右。Meta-learning LSTM 的表现较前述 3 种模型差,在收敛速度以及损失值方面,都不如前者,最终损失值约为 0.34。Meta-baseline 和 MAML 这两种方法的损失值较高,且处在波动状态,最终的损失值也较其它模型高,分别为 0.71 和 0.97 左右。

3.4 参数对结果的影响

测试各模型在 way、shot 改变情况下的识别能力。分别设置 way 为 5、shot 为 5, way 为 3、shot 为 5, way 为 5、shot 为 3, way 为 3、shot 为 3, 各模型在 3 组参数设置下,在两数据集上的测试精度、召回率和 F1 值分别如图 10 ~ 12 所示。

图 10 为不同参数设置下各模型精度结果。由图 10 可知,各模型在两数据集上的识别精度在不同参数设置下均有明显的区别,其中 way 为 3、shot 为 5 时精度最高; way 为 5、shot 为 5 时次之;随后是 way 为 3、shot 为 3;最后是 way 为 5、shot 为 3。但在不同参数设置下,各模型的识别精度仍保持了相对的差异,且在 Fish100 上的识别精度普遍高于 ImageNet。

图 11 为不同参数设置下各模型召回率结果。由图 11 可知,各模型在两数据集上的召回率在不同参数设置下均有明显的区别,其中 way 为 3、shot 为 5 时的召回率最高; way 为 5、shot 为 5 次之;随后依次为 way 为 3、shot 为 3 和 way 为 5、shot 为 3。在不同参数设置下,各模型的召回率之间也保持了相对差异,且在 Fish100 上的召回率高于 ImageNet。

图 12 为不同参数设置下各模型 F1 值测试结

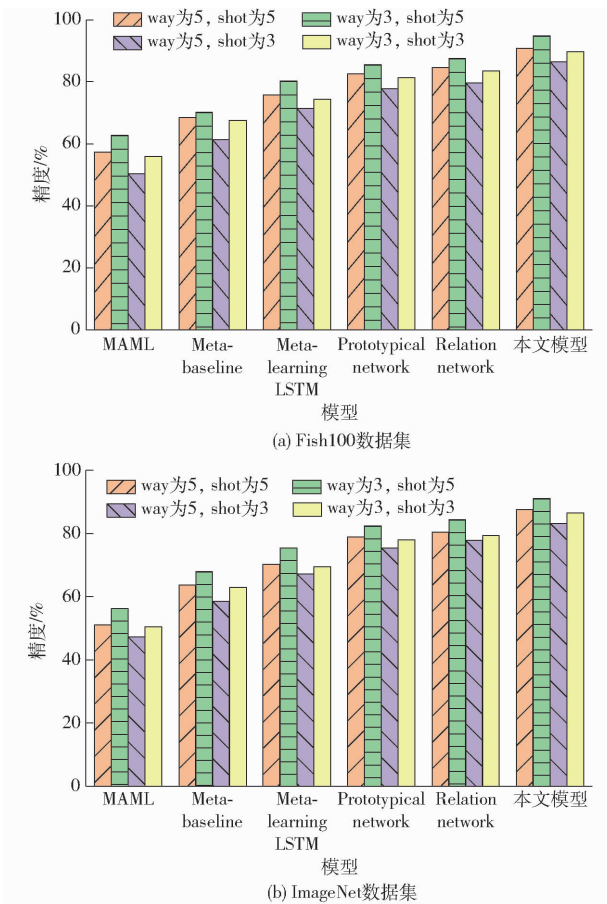


图 10 不同参数设置下各模型精度结果

Fig. 10 Model accuracy results under different parameter settings

果。由图 12 可知,各模型在两数据集上的 F1 值遵循与精度和召回率一样的规律, F1 值由高到低的参数设置依次为 way 为 3、shot 为 5, way 为 5、shot 为 5, way 为 3、shot 为 3 以及 way 为 5、shot 为 3。模型、数据集之间的差异也同精度和召回率。

3.5 way 和 shot 对结果的影响

为进一步确认不同参数设置对模型识别效果的影响,使用本文模型保持 way 为 5, 分别设置 shot 为 1 ~ 6, 在 2 个数据集上分别进行测试, 测试结果如图 13 所示。

由图 13 可知,在 way 相同的情况下,模型识别的精度、召回率和 F1 值随着 shot 的上升而上升,且在 shot 值越小时上升幅度越大。可见在 way 相同的情况下,shot 值与识别效果之间成正比关系。

保持 shot 为 5, 分别设置 way 为 2 ~ 6, 使用本文模型在两个数据集上分别进行识别测试, 结果如图 14 所示。

由图 14 可知,在 shot 相同的情况下,模型的识别精度、召回率和 F1 值随着 way 值的上升而平缓下降。可见在 shot 相同的情况下,way 值与识别效果之间成反比关系。

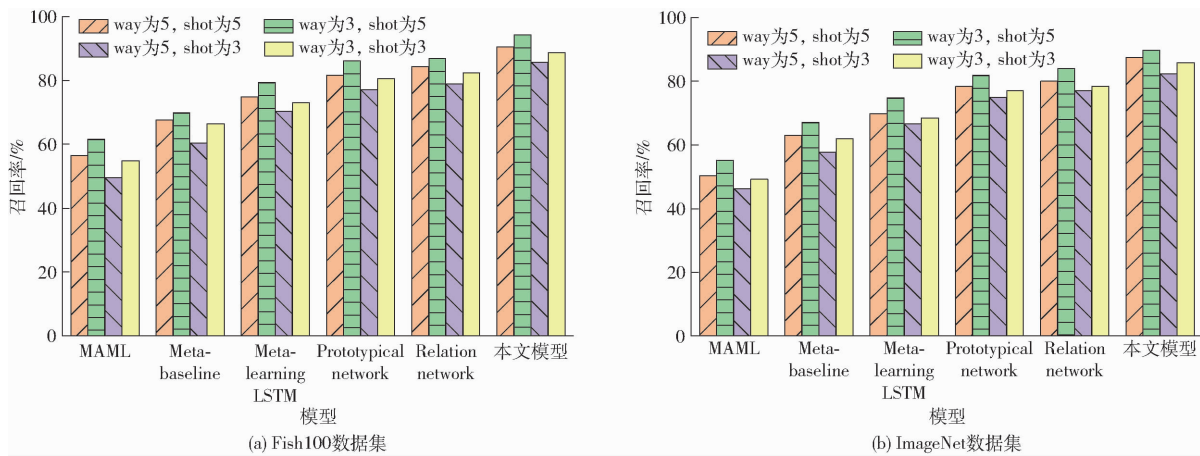


图 11 不同参数设置下各模型召回率结果

Fig. 11 Recall rate results of various models under different parameter settings

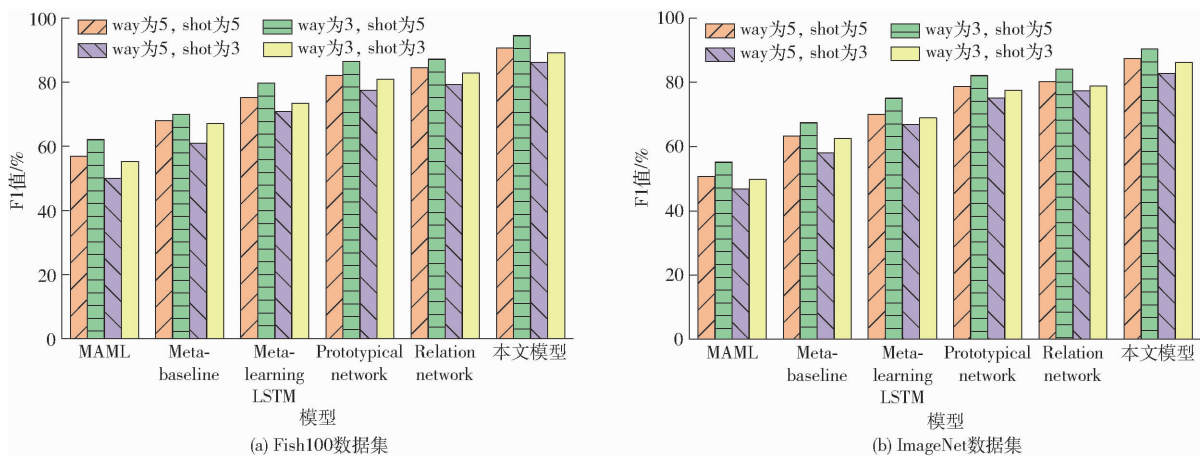


图 12 不同参数设置下各模型 F1 值测试结果

Fig. 12 F1 results of each model under different parameter settings

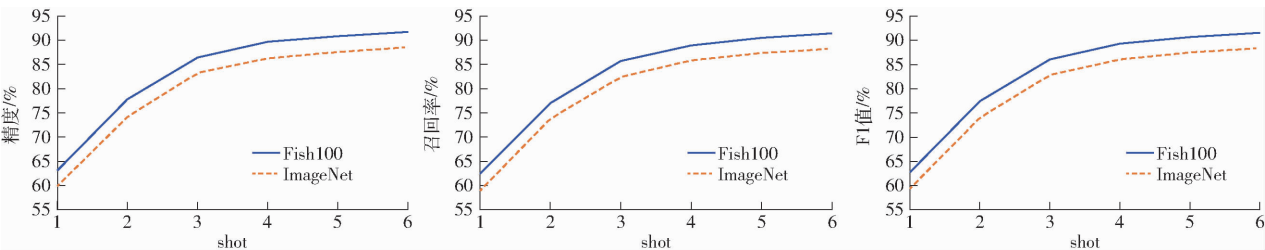


图 13 不同 shot 值下本文模型测试结果

Fig. 13 Test results of proposed model under different shot values

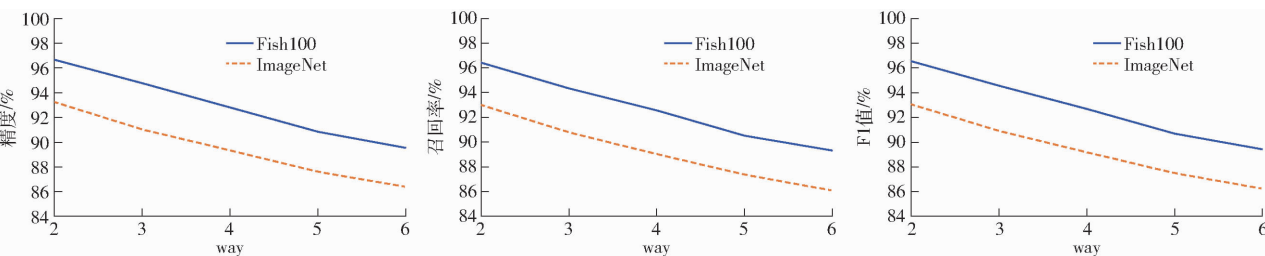


图 14 不同 way 值下本文模型测试结果

Fig. 14 Test results of proposed model under different way values

3.6 模型识别结果对比

将本文模型与 MAML、Meta-baseline、Meta-learning LSTM、Prototypical network 以及 Relation

network 等 5 种小样本学习模型的试验结果进行对比。选择最优参数 way 为 3、shot 为 5 对模型效果进行比较。表 7 为各模型在 Fish100 数据集上的测试

结果。

表 7 Fish100 数据集上 way 为 3、shot 为 5 测试结果

Tab.7 Result on Fish100 dataset while way was 3 and shot was 5 %

模型	精度	召回率	F1 值
MAML	62. 73	61. 54	62. 13
Meta-baseline	70. 19	69. 83	70. 01
Meta-learning LSTM	80. 24	79. 32	79. 78
Prototypical network	85. 43	86. 17	86. 50
Relation network	87. 46	86. 92	87. 19
本文模型	94. 77	94. 35	94. 56

由表 7 可以看出,在 Fish100 数据集上本文模型在精度、召回率以及 F1 值上,均显著优于其它模型,其中精度较其它模型中识别效果最佳的 Relation network 高 7. 31 个百分点,较 Prototypical network 高 9. 34 个百分点,较 Meta-learning LSTM 高 14. 53 个百分点,较 MAML 高 32. 04 个百分点。本文提出利用残差块结构改进小样本学习网络,有效提取鱼图像深层特征,提高了鱼图像识别精度,使得其表现优于其他模型。

各模型在 ImageNet 数据集的识别结果如表 8 所示。由表 8 可以看出,各模型在 ImageNet 鱼图像识别任务中的结果,较 Fish100 数据集均有一定的降低,本文模型的识别精度下降了 3. 74 个百分点,Relation network、Prototypical network、Meta-learning LSTM、Meta-baseline 以及 MAML 则分别下降了 3. 14、3. 00、4. 78、2. 28、6. 43 个百分点。其原因是受 ImageNet 数据集中复杂背景环境因素影响,模型

表 8 ImageNet 数据集上 way 为 3、shot 为 5 测试结果

Tab.8 Result on ImageNet dataset while way was 3 and shot was 5 %

模型	精度	召回率	F1 值
MAML	56. 30	55. 18	55. 24
Meta-baseline	67. 91	67. 03	67. 47
Meta-learning LSTM	75. 46	74. 78	75. 12
Prototypical network	82. 43	81. 80	82. 11
Relation network	84. 32	84. 01	84. 17
本文模型	91. 03	90. 78	90. 90

识别效果均有所降低。即使如此,相比于其它 5 种小样本学习模型,本文模型识别效果仍然最佳,识别精度为 91. 03%,召回率为 90. 78%,F1 值为 90. 90%。综上,本文模型对鱼图像识别具有较好的效果,可为后期小样本鱼图像识别在实际环境中的应用提供技术支撑和参考。

4 结论

(1)本文模型在 Fish100、ImageNet 数据集上均优于其他小样本学习方法,特别在 Fish100 数据集上,试验效果最佳,其中精度、召回率以及 F1 值分别为 94. 77%、94. 35% 和 94. 56%。

(2)在不同参数下,各模型在 Fish100 数据集上的识别效果均优于 ImageNet,表明 ImageNet 数据集识别难度高于 Fish100 数据集。

(3)在小样本学习中,way、shot 的取值会影响模型识别结果,其中 way 的取值与识别结果成反比,而 shot 则与识别结果成正比,且 shot 的影响权重大于 way。

参 考 文 献

[1] LYNCH A J, COOKE S J, DEINES A M, et al. The social, economic, and environmental importance of inland fish and fisheries[J]. Environmental Reviews, 2016, 24(2): 1 – 7.

[2] DARGAN S, KUMAR M, AYYAGARI M R. et al. A survey of deep learning and its applications: a new paradigm to machine learning[J]. Archives of Computational Methods in Engineering, 2020, 27(4): 1071 – 1092.

[3] VAN TREECK R, VAN WICHELEN J, WOLTER C. Fish species sensitivity classification for environmental impact assessment, conservation and restoration planning[J]. Science of the Total Environment, 2020, 708: 135173.

[4] 刘颖, 雷研博, 范九伦, 等. 基于小样本学习的图像分类技术综述[J]. 自动化学报, 2021, 47(2): 297 – 315.

LIU Ying, LEI Yanbo, ZHANG Jiulun, et al. A review of image classification techniques based on few-shot learning[J]. Journal of Automation, 2021, 47(2): 297 – 315. (in Chinese)

[5] FEIFEI L, FERGUS R, PERONA P. One-shot learning of object categories[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4): 594 – 611.

[6] MENSINK T, VERBEEK J, PERRONNIN F, et al. Metric learning for large scale image classification: generalizing to new classes at near-zero cost[C] // European Conference on Computer Vision, 2012: 297 – 310.

[7] ROHRBACH M, EBERT S, SCHIELE B. Transfer learning in a transductive setting[C] // Neural Information Processing Systems, 2013: 46 – 54.

[8] SANTORO A, BARTUNOV S, BOTVINICK M, et al. One-shot learning with memory-augmented neural networks[J]. arXiv, 2016: 1605. 06065.

[9] SNELL J, SWERSKY K, ZEMEL R S. Prototypical networks for few-shot learning: neural information processing systems[C] // Neural Information Processing Systems, 2017: 4077 – 4087.

- [10] SUNG F, YANG Y, ZHANG L, et al. Learning to compare: relation network for few-shot learning: computer vision and pattern recognition[C]//Computer Vision and Pattern Recognition, 2018: 1199 – 1208.
- [11] CHOI Y, PARK J, YI S, et al. Few-shot classification on unseen domains by learning disparate modulators[J]. arXiv, 2019: 2001.08735.
- [12] TSENG H, LEE H, HUANG J, et al. Cross-domain few-shot classification via learned feature-wise transformation[J]. arXiv, 2020:2001.08735.
- [13] ZHANG J, ZHANG M, LU Z, et al. AdarGCN: adaptive aggregation GCN for few-shot learning[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021:3482 – 3491.
- [14] 潘兴甲,张旭龙,董未名,等. 小样本目标检测的研究现状[J]. 南京信息工程大学学报(自然科学版), 2019,11(6): 698 – 705.
PAN Xingjia, ZHANG Xulong, DONG Weiming, et al. Research status of few-shot target detection[J]. Journal of Nanjing University of Information Technology (Natural Science Edition), 2019,11(6): 698 – 705. (in Chinese)
- [15] PENG B, ZHU C, LI C, et al. Few-shot natural language generation for task-oriented dialog[C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings, 2020:172 – 182.
- [16] 陈英义,龚川洋,刘焯琦,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, GONG 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)
- [17] 刘鑫鹏,栾悉道,谢毓湘,等. 迁移学习研究和算法综述[J]. 长沙大学学报, 2018, 32(5): 28 – 31.
LIU Xinpeng, LUAN Xidao, XIE Yuxiang, et al. Research and algorithm review on transfer learning[J]. Journal of Changsha University, 2018,32(5): 28 – 31. (in Chinese)
- [18] SUN Q, LIU Y, CHUA T, et al. Meta-transfer learning for few-shot learning[C]//Computer Vision and Pattern Recognition, 2019: 403 – 412.
- [19] MÜLLER H, CLOUGH P, DESELAERS T, et al. ImageCLEF[M]. Springer Berlin Heidelberg, 2010.
- [20] 冀中,赵可心,张锁平,等. 基于空间变换双线性网络的细粒度鱼类图像分类[J]. 天津大学学报(自然科学与工程技术版), 2019, 52(5): 475 – 482.
JI Zhong, ZHAO Kexin, ZHANG Suoping, et al. Image classification of fine-grained fish based on spatial transformation bilinear network[J]. Journal of Tianjin University (Natural Science and Engineering Technology Edition), 2019,52(5): 475 – 482. (in Chinese)
- [21] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//Computer Vision and Pattern Recognition, 2009: 248 – 255.
- [22] TARG S, ALMEIDA D, LYMAN K. ResNet in resnet: generalizing residual architectures[J]. arXiv, 2016:1603.08029.
- [23] DAI D, YU L, WEI H. Parameters sharing in residual neural networks[J]. Neural Processing Letters, 2020,51(2): 1393 – 1410.
- [24] WANG S, XIA C, WANG T. Feature generation: a novel intrusion detection model based on prototypical network[C]//International Conference on Algorithms and Architectures for Parallel Processing, 2019: 564 – 577.
- [25] KINGMA D P, BA J. Adam: a method for stochastic optimization[C]//International Conference on Learning Representations, 2015: 223 – 236.
- [26] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C]//International Conference on Machine Learning, 2017: 1126 – 1135.
- [27] CHEN Y, WANG X, LIU Z, et al. A new meta-baseline for few-shot learning[J]. arXiv, 2020:2003.04390.
- [28] RAVI S, LAROCHELLE H. Optimization as a model for few-shot learning[C]//International Conference on Learning Representations, 2017: 24 – 26.