

基于改进 Xception 迁移学习的野生菌种类识别研究

陈德刚, 艾孜尔古丽*, 尹鹏博, 陆雅诺, 李顺平

新疆师范大学计算机科学技术学院, 新疆 乌鲁木齐 830054

摘要 由于我国野生菌分布不均, 种类多样, 一般人群难以辨别野生菌是否可食用, 中毒事件时有发生, 亟需一种高效的识别野生菌种类的方法。提出了一种新型野生菌种类图片识别模型 (Dis-Xception-CBAM), 以 Xception 结构为基准, 结合网络权重的迁移学习方法进行特征学习, 加入了注意力机制以进一步提取野生菌显式特征, 同时引入特征图扰动结构, 增加了模型的泛化性能。基于国内常见的野生菌, 构建了 33 种野生菌的图片数据集并进行实验, 当初始学习率为 0.001, 训练迭代次数为 300 时, Top 1 达到 96.32%, Top 5 达到 99.61%。相较于传统图像识别模型, 所提模型取得了较好的结果, 为野生菌识别研究提供了理论依据。

关键词 图像处理; 野生菌; 显式特征; Xception; 迁移学习; CBAM

中图分类号 TP391.41

文献标志码 A

doi: 10.3788/LOP202158.0810023

Research on Identification of Wild Mushroom Species Based on Improved Xception Transfer Learning

Chen Degang, Azragul*, Yin Pengbo, Lu Yanuo, Li Shunping

School of Computer Science and Technology, Xinjiang Normal University, Urumqi, Xinjiang 830054, China

Abstract Due to the uneven distribution of wild mushrooms in China, the diversity of species, and the difficulty for the general population to distinguish whether wild mushrooms are edible, there occur positioning accidents from time to time and thus it is urgent to propose an efficient method to identify wild mushroom species. Here, a new type of wild mushroom specie image recognition model (Dis-Xception-CBAM) is proposed, which takes the Xception structure as a benchmark, combines the transfer learning method with network weights for feature learning, adds the attention mechanism to further extract the explicit features of wild mushrooms, and introduces the feature map distortion structure to increase the model's generalization performance. A dataset of 33 kinds of wild mushroom images is constructed based on the common wild mushrooms in China and the corresponding experiments are conducted. When the initial learning rate is 0.001 and the number of training iterations is 300, the Top 1 reaches 96.32% and the Top 5 reaches 99.61%. Compared with the traditional image recognition model, the proposed model obtains a better result, which provides a theoretical basis for wild mushroom recognition research.

Key words image processing; wild mushrooms; explicit features; Xception; transfer learning; CBAM

OCIS codes 100.2000; 100.2960; 100.3008; 150.1135

1 引言

野生菌主要分为食用菌和毒菌两大类。食用菌即为大多数人常说的“蘑菇”, 而毒菌^[1]亦称毒蘑菇、毒蕈等, 一般是指对人或畜禽有毒的大型真菌物种。

自然界的毒菌估计达 1000 种以上, 而我国至少有 500 种。食用菌作为人们烹饪的重要食材, 在食用安全问题上常常存在很大争议。随着社会的不断发展, 人们追求食用新鲜野生菌, 而菌类分布广泛, 生存环境多样, 大多生长在阴凉潮湿的灌木丛林中, 毒

收稿日期: 2020-10-10; 修回日期: 2020-10-22; 录用日期: 2020-11-05

基金项目: 新疆维吾尔自治区自然科学基金(2017D01A58)

* E-mail: Azragul2010@126.com

菌和食用菌^[2]的宏观特征极其相似,除从事相关专业人员外,普通人群在野外杂草丛生的情况下,极易混淆误食毒菌。因此,高效的识别野生菌种类的方法具有重要意义。

目前,野生菌识别方法主要有常识经验识别、化学检测识别、动物实验检测以及机器学习方法等。利用常识经验识别时,需要观察野生菌的形态、气味、颜色等特征,这对经验背景有极高的要求。而化学检测识别^[3-4]、动物实验检测^[5]等方法通常需要精密、昂贵的实验仪器,实验周期较长,并不适用于一般人群户外识别野生菌种类,具有很大的局限性。文献[6-8]分别提出基于贝叶斯、支持向量机、决策树等机器学习方法进行毒菌种类识别,这些方法虽然易于实现、识别准确率较高,但需要人工进行繁琐的特征标注,不利于野外环境下野生菌种类的实时有效识别,而且这三种方法所使用的实验数据集^[9]均来自于加州大学,不符合我国野生菌种类分布的现状,不具备普适性。

深度学习技术可自动提取数据特征,减少了数据预处理等繁琐工作,且语义特征表达能力强,适用范围广,在自然语言处理、语音处理、图像识别等多个领域有不错表现。在图像识别领域,基于卷积神经网络的 AlexNet^[10]、VGGNet^[11]、残差网络 ResNet^[12]等模型得到了广泛的应用,如田间

杂草识别^[13]、农作物病虫害识别^[14]、马铃薯缺陷检测^[15]等。樊帅昌^[16]等利用残差网络实现了毒菌图像识别,但模型参数量较大,并不适合推广应用。

本文提出一种基于改进 Xception^[17] 迁移学习^[18] 的野生菌种类识别模型 (Dis-Xception-CBAM)。以 Xception 结构为基准,采用迁移学习方法,通过加载网络权重,冻结低层特征,训练高层特征提取层,并加入注意力机制层 (Convolutional Block Attention Module, CBAM)^[19] 和特征图扰动层 (Feature Map Distortion, Disout)^[20], 最终决策分类层的神经元数目与本文野生菌数据集的种类一致。针对国内现存野生菌的分布,采用本文提出的 Dis-Xception-CBAM 模型,能够在不同复杂环境条件下高效识别出菌种类型。

2 材料与方法

2.1 实验材料采集

以中国科学院昆明植物研究所公布的国内常见毒菌种类清单为基础,利用 ImageAssistant 软件,针对我国野生菌种类分布情况,筛选了国内常见的 25 种毒菌和 8 种食用菌。图片数量超过 3500 张,每个类别的数量为 100~200 张,部分野生菌图片样例如图 1 所示。



图 1 野生菌图片样例。(a)致命鹅膏;(b)灰花纹鹅膏;(c)拟卵盖鹅膏;(d)小豹斑鹅膏;(e)红托鹅膏;(f)方形粉褶菌;(g)花边斑褶菇;(h)喜粪裸盖菇;(i)赭鹿花菌;(j)叶状耳盘菌

Fig. 1 Sample data of wild mushroom images. (a) *Amanita exitialis*; (b) *Amanita fuliginea*; (c) *Amanita neovooides*; (d) *Amanita parvipantherina*; (e) *Amanita rubrovolvata*; (f) *Entoloma quadratum*; (g) *Panaeolus sphinctrinus*; (h) *Psilocybe coprophila*; (i) *Gyromitra infula*; (j) *Lonomidotis frondosa*

2.2 数据集扩充

目前,互联网上的野生菌图片数量庞大,但国内野生菌的相关图像数据集较为稀少,本文主要针对

小样本数据集^[21]进行图像识别。为了提高模型的稳健性,增强野生菌种类识别的准确率,本文采用常见的数据增强方式,实现数据集的有效扩充。本文

使用到的数据增强方式有:图片随机角度旋转(0~360°)、水平翻转、随机颜色抖动、添加高斯噪音、图片直方均衡化、随机裁剪。删除部分特征丢失严重

的野生菌图片数据,最终图片数量为 27204 张。以方形粉褶菌为例,不同数据增强方式的效果如图 2 所示。

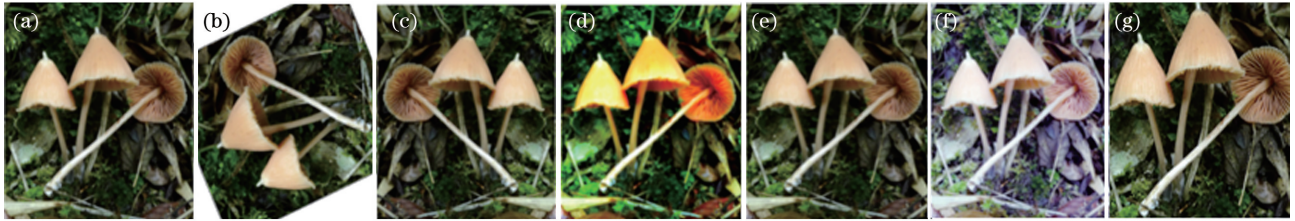


图 2 不同数据增强方式的效果。(a)原始图片;(b)随机旋转;(c)水平翻转;(d)色彩抖动;(e)高斯噪音;(f)直方图均衡化;(g)随机裁剪

Fig. 2 Effects of different data enhancement methods. (a) Origin image; (b) random rotation; (c) horizontal flip; (d) color dither; (e) Gaussian noise; (f) histogram equalization; (g) random cut

2.3 野生菌种类识别模型

本节将介绍基准模型 Xception 的结构,详细阐述蘑菇种类识别模型的流程以及改进 Xception 结构中所采用的注意力机制层和特征图扰动层。

2.3.1 Xception 结构

Xception 结构是基于 Inception 系列^[22-24]和 ResNet 系列^[25-27]发展起来的。Xception 采用 36 层可分离卷积层,各层次结构都有极大的优化,在空间与通道上有效实现了映射的充分解耦合性。同时,它吸取了 ResNet 网络跨层跳跃连接^[28]加快网络训练的经验,将可分离卷积层分为 14 个组件,除前后

两个组件外,其他都采用了残差连接优化网络结构。结合可分离卷积层与 Resnet 网络可获得更深更广的神经网络模型。Xception 主体结构如图 3 所示。A~G 组件结构中神经网络自上而下传递的网络层具体参数设置如表 1、2 所示,其中 Conv 为卷积网络, stride 为卷积滑动窗口步长, Sep Conv 为可分离卷积结构, ReLU 为神经网络激活函数, $\left\{ \begin{matrix} \text{ReLU} \\ \text{Sep Conv } 728, 3 \times 3 \end{matrix} \right\} \times 3$ 表示 Sep Conv 层与 ReLU 层重复三次, MaxPool 与 Global average pool 分别为最大池化层与全局平均池化层。

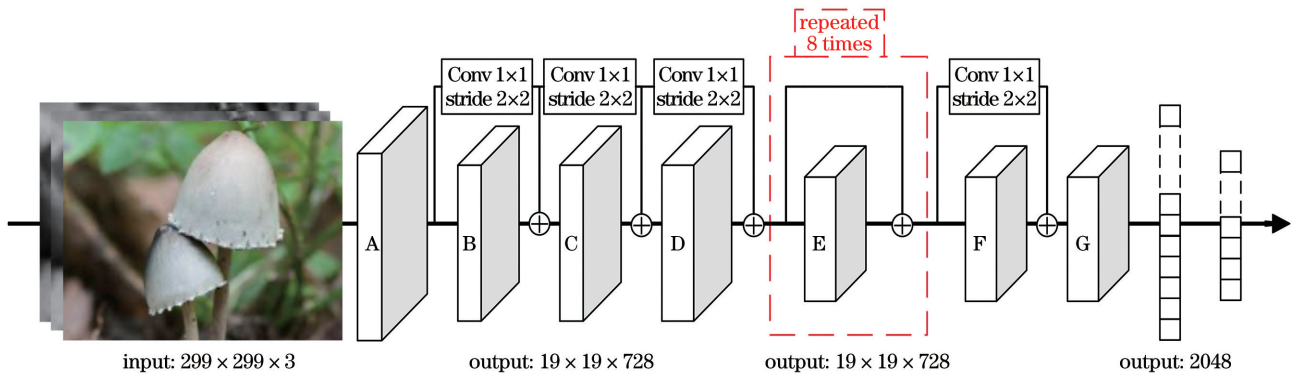


图 3 Xception 结构图

Fig. 3 Structural diagram of Xception

表 1 Xception 中的 A、B、C、D 组件结构

Table 1 Structures of A, B, C, and D components in Xception

A	B	C	D
[Input 299×299×3]	Sep Conv 128, 3×3	$\left\{ \begin{matrix} \text{ReLU} \\ \text{Sep Conv } 256, 3 \times 3 \end{matrix} \right\} \times 3$	$\left\{ \begin{matrix} \text{ReLU} \\ \text{Sep Conv } 728, 3 \times 3 \end{matrix} \right\} \times 3$
Conv 32, 3×3, stride of 2×2	ReLU	MaxPool 3×3, stride of 2×2	MaxPool 3×3, stride of 2×2
ReLU	Sep Conv 128, 3×3	MaxPool 3×3, stride of 2×2	MaxPool 3×3, stride of 2×2
	MaxPool 3×3, stride of 2×2		

表 2 Xception 中的 E、F、G 组件结构

Table 2 Structures of E, F, and G components in Xception

E(repeated 8 times)	F	G
$\left\{ \begin{array}{c} \text{ReLU} \\ \text{Sep Conv } 728, 3 \times 3 \end{array} \right\} \times 3$	ReLU	Sep Conv 1536, 3 × 3
	Sep Conv 728, 3 × 3	ReLU
	ReLU	Sep Conv 2048, 3 × 3
	Sep Conv 1024, 3 × 3	ReLU
	MaxPool 3 × 3, stride of 2 × 2	Global average pool

2.3.2 野生菌种类识别模型的整体架构

本文提出的用于野生菌种类识别的 Dis-Xception-CBAM 模型的实验流程主要如图 4 所示, 首先收集国内常见 33 种野生菌种图片并分类整合。由于国内野生菌不同种类的图片数量有限, 利用深度学习对小样本集^[21] 图像进行识别分析, 并采用适当数据增强方式扩充野生菌图片数据集。本文采用随机裁剪、翻转、添加高斯噪音等数据增强方式增加图片数量, 可以有效提高图像识别的准确率。之后, 对扩充后的数据集进行统一图片大小、格式等相关预处理, 再将其按比例拆分成训练集、验证集及测试集, 训练集和验证集用于网络模型训练, 测试集

用于验证模型实际应用的泛化性能。同时, 将大型图片数据集 ImageNet 输入到预训练模型 Xception 中, 采用迁移学习方法, 将预训练模型的权重参数迁移到本文提出的野生菌种类识别模型 (Dis-Xception-CBAM) 中, 选择冻结 A~E 组件, 仅训练高层特征提取网络 F、G 组件, 后续加入注意力机制层与特征扰动层等, 并将最后决策分类层改为 33 个神经元, 以对应本文所采取的野生菌种类数目, 分类器采用 Softmax, 运用损失交叉熵 (Cross Entropy Loss) 验证损失函数。最后将拆分后的数据集分别输入到神经网络中进行训练与验证, 并获得最终的野生菌分类模型。

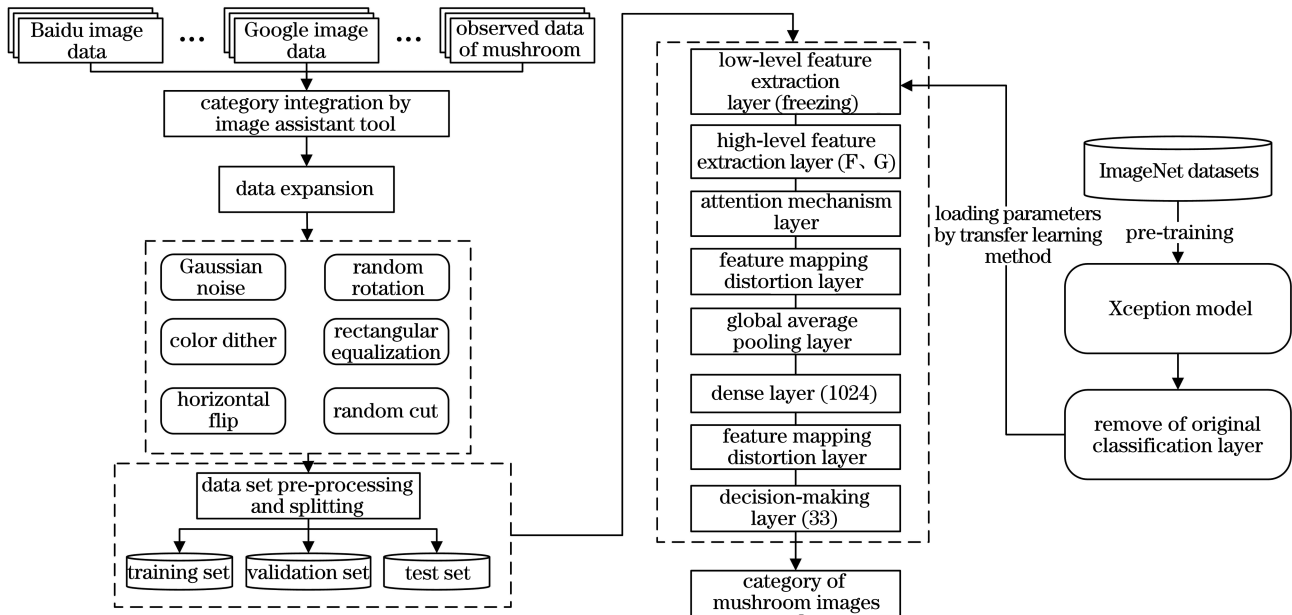


图 4 野生菌种类识别模型的实验流程

Fig. 4 Experimental flow chart of wild mushroom species identification model

2.3.3 注意力机制层

注意力机制在视觉领域里面深受欢迎, 它通过提取图片局部关键特征, 有效减少图像背景等不相关的干扰因素对野生菌种类识别的影响, 在过滤掉无用信息的同时, 有效加快了深度卷积神经网络的

特征提取速度, 提高了模型泛化性能。上一层 Xception 结构中使用了多个卷积核进行图像特征提取, 通道、空间注意力按不同权重提取特征图。本模型使用 CBAM 网络加强通道、空间两个维度上提取特征的能力。该网络优于传统仅在通道维度上提取

权重的通道注意力 (Squeeze-and-Excitation Networks, SENet)^[29] 网络, CBAM 在两个维度上

利用通道、空间之间的关系差异, 学习不同权重网络的特征图。其实现原理过程如图 5 所示。

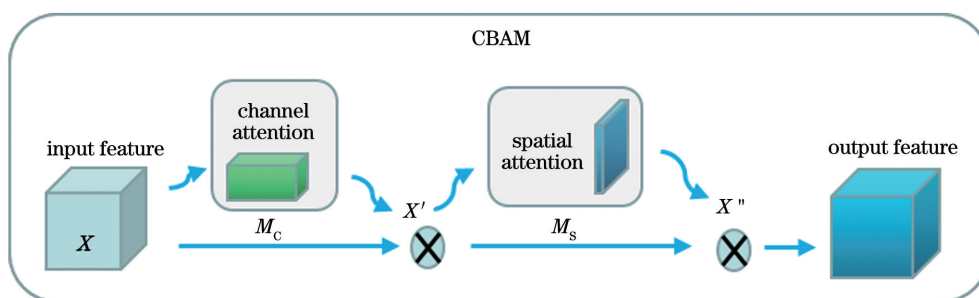


图 5 CBAM 实现原理过程

Fig. 5 Principle diagram of CBAM's realization

CBAM 分通道、空间两个注意力模块来提取特征网络。原始输入特征图 X 经过宽、高尺度的全局最大池化和全局平均化处理, 再利用 M_c 通道进行注意力处理, 获得通道注意力加权的特征图 X' 。下一阶段引入空间注意力模块处理。 M_s 空间注意力处理机制是将通道注意力处理过的 X' 作为该阶段输入, 先在通道维度上分别进行全局最大池化和全局平均化处理, 拼接处理结果后, 再通过卷积降维和 sigmoid 激活, 最终获得通道、空间两个维度上的加权特征图网络 X'' 。本文改进的 Xception 模型是在 G 组

件后面加入 CBAM 注意力模块以提取特征信息。

2.3.4 特征图扰动层

在深度学习中, 神经网络模型训练中常出现数据集太少或训练次数过多等原因导致的模型过拟合情况, 采用特征扰动方法能够显著提高模型的稳健性。将常规网络与使用了特征扰动的方法 (Dropout^[30] 和 Disout 方法) 进行对比, 如图 6 所示, 其中 \circ 代表正常传输的神经元, \otimes 代表已经失活的未被利用的神经元, \bullet 、 \bullet 代表经过指导性特征扰动后仍继续传输的神经元。

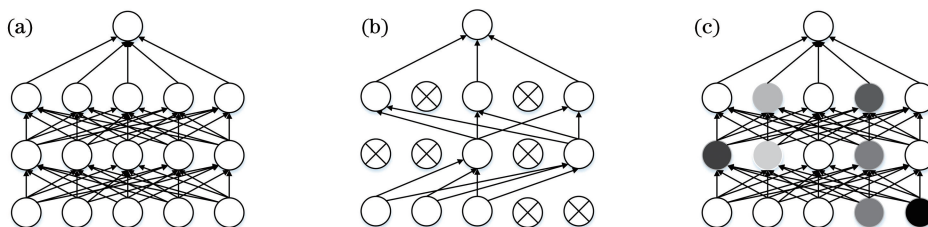


图 6 三种神经网络结构的对比。(a) 传统神经网络; (b) Dropout 神经网络; (c) Disout 神经网络

Fig. 6 Comparison of three kinds of neural network structures. (a) Traditional neural network; (b) Dropout neural network; (c) Disout neural network

传统的解决方法是使用 Dropout 形式, 通过随机丢弃部分神经元来提高神经网络的泛化能力, 相当于在原来的神经网络模型上进行特征扰动, 从而提高模型的容错能力。虽然 Dropout 方法能较好提高模型的泛化能力, 但在扰动值大小固定为原始特征图大小的情况下, 并不一定能够取得全局最优解, 且不具备理论基础支撑。而 Disout 方法提出了一种利用泛化理论即拉德马赫复杂度 (Empirical Rademacher Complexity, ERC) 指导的可学习的特征扰动, 可得到更好的特征图扰动项, 并得到预期风险的上限。但由于神经网络结构复杂, 指导难度大, 且指导形式与每一次网络输出有关, 特征扰动是建立在每一层特征图上的, 因此利用传递性质拉近

ERC 值的计算结果与扰动形式间的距离。

利用中间层复杂度指导邻近层的特征扰动形式, 可得到优化目标。通过优化目标函数, 可以得到合适的扰动值。但需要注意的是, 特征扰动过大会降低特征网络的表征能力, 而过小的扰动则不能起到很好的正则化效果, 因此需要多次训练以获得最优扰动。改进 Xception 模型在 CBAM 层后面和全连接层后面均加入 Disout 特征扰动以增加模型的泛化性能。

3 结果与分析

3.1 实验指标

本文所有图像分类实验均基于控制变量法, 保

证每次实验只有单一变量影响最终实验结果。模型对比实验指标采用图像识别领域中常用的 Top 1, Top 5 衡量模型的性能,平均准确率 Top 1 (T_{Top1}) 表示模型最终预测的概率最大结果与正确标签相同的概率,平均准确率 Top 5 (T_{Top5}) 表示模型输出概率向量前五中有正确标签的概率。计算公式为

$$T_{Top1} = \frac{1}{N} \times \frac{\sum_{i=1}^N t_i}{\sum_{i=1}^N n_i} \times 100\%, \quad (1)$$

$$T_{Top5} = \frac{1}{N} \times \frac{\sum_{i=1}^N T_i}{\sum_{i=1}^N n_i} \times 100\%, \quad (2)$$

式中: N 为野生菌总类别样本数; t_i 为每一类野生菌中预测最大概率向量为正确标签的数量; T_i 为每一类野生菌中预测前五中有正确标签的数量; n_i 为野生菌类别数; i 为第 i 类野生菌,最大值即为 N 。训练时耗和模型参数量计算公式分别为

$$T_{Average} = \sum_{j=1}^n \frac{B_j}{E_j}, \quad (3)$$

$$P = \begin{cases} K^2 \cdot C_{in} \cdot C_{out}, & \text{for convolution layer} \\ W_{in} \cdot W_{out}, & \text{for dense layer} \end{cases}, \quad (4)$$

式中: $T_{Average}$ 为模型达到收敛的平均训练耗时; B_j 为第 j 轮训练耗时; E_j 为第 j 轮训练迭代时间; P 为模型参数量; K 为卷积核的大小; C_{in} 和 C_{out} 分别为卷积层中特征图的输入与输出; W_{in} 和 W_{out} 分别为全连接层中特征图的输入与输出。

3.2 实验平台与参数设置

本文实验平台为图形处理器 Tesla K80,系统为 ubuntu16.04,并行计算架构版本为 10.2,神经网络框架版本为 2.2.0,神经网络库为 2.4.3。野生菌数据被随机划分为训练集、验证集和测试集,其大小比例为 6 : 2 : 2,实验初始学习率为 0.001,优化器为 Adam,训练迭代数(Epoch)为 300,批大小为 32,

表 4 不同训练方式的对比

Table 4 Comparison among different training methods

Experimental number	Training method	Top 1 / %	Top 5 / %	$T_{Average} / s$
1 # #	Random parameters	92.10	98.10	1688.37
2 # #	Freezing all parameters	87.26	98.34	599.38
3 # #	Freezing partial parameters	96.47	99.69	647.91
4 # #	Training all network layers	97.02	99.67	1645.39

实验 1 # # : 不加载 ImageNet 预训练权重,采

用自适应学习率动态调整模型学习步长。在 ResNet 系列中,官方提供的预训练模型权重的输入图像尺寸为 224×224 ,在 Inception 系列中,InceptionResNet-V2 的尺寸为 299×299 。使用较小尺寸的图像输入模型,能减少网络训练时长与内存占用,但其特征图也会有所损失,出现准确率降低、特征丢失等问题。因此,本文按照这两种类型分别调整图片大小。

3.3 不同特征扰动形式的对比

对比不同特征图扰动形式对模型测试集识别准确率的影响,如表 3 所示,实验 1 #、2 #、3 # 分别对应特征扰动,Dropout 和 Disout 模型。

表 3 不同特征图扰动形式对比

Table 3 Comparison among different feature map disturbance forms

Experimental number	Training structure	Top 1 / %	Top 5 / %
1 #	Origin model	94.91	99.30
2 #	Dropout	95.92	99.33
3 #	Disout	96.32	99.61

表 3 表明,引入特征图扰动方法能够有效提高模型的泛化性能。对于未加入特征扰动的原始模型(Origin model),野生菌种类识别准确率 Top 1 仅为 94.91%,Top 5 为 99.33%。对于加入特征扰动的两种方法模型,Top 1 与 Top 5 均有所提高,即模型性能得到改善,容错性得到提高。但由于 Dropout 采用的是随机失活部分神经元,不具备理论支撑,且泛化性能低于 Disout 方法,因此本文采用 Disout 对 Xception-CBAM 模型进行改进。

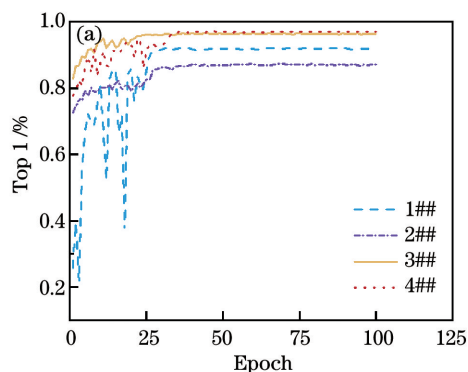
3.4 引入迁移学习的有效性分析

为了验证迁移学习方法对模型性能有促进作用,以 Dis-Xception-CBAM 模型为参照,对比四种模型训练形式下验证集上的识别准确率以及训练时耗,如表 4 所示。四种模型训练形式下的实验分别如下。

用初始权重随机生成方式训练模型。

实验 2 号：采用迁移学习方式，加载 ImageNet 初始权重，但冻结特征提取层的全部参数，只在最后决策分类中加入全连接层以输出野生菌种类。

实验 3 号：与实验 2 号同理，但仅冻结低层



特征提取层的 A~E 组件权重，设置 F、G 组件权重参数为可训练。

实验 4 号：加载所有权重参数，设置所有特征提取层参数可训练。

不同训练方式下模型参数的对比如图 7 所示。

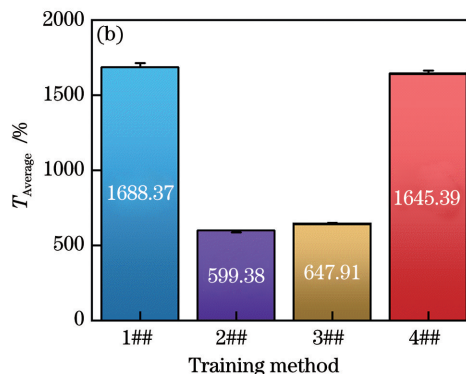


图 7 不同训练方式下模型参数的对比。(a)准确率;(b)平均训练时耗

Fig. 7 Comparison among model parameters for different training methods. (a) Accuracy; (b) average training time

表 4 和图 7 表明，实验 1 号不采用迁移学习时，神经网络模型学习提取野生菌色彩、纹理等显式特征的平均时耗较长，达到 1688.37 s，准确率起伏较大，模型收敛慢，且准确率低于实验 3 号和 4 号。实验 2 号冻结网络全部权重，虽然平均训练时耗最短，也具备了提取野生菌大体显式特征的能力，能够过滤掉环境等干扰因素，但由于野生菌种类繁多，纹理复杂，仅利用最后决策分类层识别野生菌种类，识别效果较实验 1 号差。实验 3 号表明，深度学习图像信息处理方式是分级的，低层到高层特征将会越来越抽象，冻结低层特征能够将模型特征学习的视野聚焦于野生菌本身，仅训练 F、G 组件便足够提取不同野生菌间细节、纹理等显式特征差异，模型 Top 1 率也达到了 96.47%，平均训练时

耗也仅次于实验 2 号。实验 4 号的 Top 1 最高，高于实验 3 号 0.55 个百分点。由于需要从底层开始从头训练，平均训练时耗与不采用迁移学习方式相差不大，但相较于实验 3 号，时长高出约 2.5 倍。因此，本文后续实验均采用实验 3 号中的方法，仅训练高层特征提取层。同时，以上结果也证明了迁移学习方法确实能够有效提高模型的识别性能并减少训练平均时耗。

3.5 不同数据比例的划分

以本文提出的 Dis-Xception-CBAM 模型为载体，将 20% 数据集划分为测试集，如表 5 所示，剩余数据将按照不同比例重新划分训练集与验证集，在测试集上评估模型，探究最佳实验划分比例。

表 5 不同比例划分的对比

Table 5 Comparison among different proportions

Experimental number	Size of training set	Size of validation set	Top 1 / %	Top 5 / %
1-1	5	5	94.23	98.91
2-1	6	4	95.15	98.99
3-1	7	3	95.67	99.36
4-1	8	2	96.32	99.61

传统数据划分均依靠经验，没有实践验证。合理划分数据集能够有效加快神经网络训练，有助于检测模型是否过拟合以及是否需要微调超参数。表 5 表明，训练集与验证集的最佳划分比例为 8 : 2，平均识别率 Top 1、Top 5 分别达到 96.32%

和 99.61%。

3.6 不同模型实验的对比

对图像识别领域常见的预训练模型与本文模型进行对比，如表 6 所示，为了保证实验公平性，以上模型都引入了 ImageNet 数据集预训练过的网络权

重参数。对比模型微调后均加入两层全连接层与 Dropout 层,最后一层加入野生菌种类识别分类层。

Alex 模型是由 Hinton 等设计的,2012 年赢得了 ImageNet 竞赛冠军,从此卷积神经网络在图像识别领域刮起了热潮,具有划时代意义。ResNet 50、101、152 层均出自同一篇论文,该文提出的残差网络结构基于深度学习的浅层网络,为后续模型结构的搭建提供了理论参考,樊帅昌等^[16]也将其应用于毒菌识别。InceptionV1 通过不同卷积结构并行提取特征,增加了卷积核的稀疏性,也减少了模型训练的运算消耗,提高了模型泛化性能。InceptionV3 是 InceptionV1 的改进,利用了因子分解法的思想,将一个较大卷积 $n' \times n'$ 分解成 $1 \times n'$ 和 $n' \times 1$ 的卷积,有效减少了网络参数。InceptionResNetV2^[31] 结合了 ResNet 与 Inception 结构的优点,在提高广度与深度的同时,模型准确率也有所提升。Xception 采用了可分离卷积与残差连接。杨森等^[15]修改了最终分类层,并将其应用于马铃薯外部缺陷分级检测中,该模型即是本文提出的 Dis-Xception-CBAM 模型,是对 Xception 结构的改进,引入了 CBAM 注意力以及 Disout 特征图扰动。

表 6 不同模型实验的对比

Table 6 Comparison among different model experiments

Model	Top 1 / %	Top 5 / %	Number of parameters / 10^6
Alex	86.24	97.94	61.10
ResNet50	89.39	98.64	25.64
ResNet101	91.64	98.69	44.71
ResNet152	92.36	99.32	60.42
InceptionV1	90.64	98.85	13.02
InceptionV3	93.16	99.09	23.93
InceptionResNetV2	95.41	99.28	55.87
Xception	95.58	99.43	22.99
Dis-Xception-CBAM	96.32	99.61	24.04

表 6 表明,本文提出的 Dis-Xception-CBAM 模型在 Top 1, Top 5 指标上均高于现有模型,相较于 Top 1 排名第三的 InceptionResNetV2,模型参数量得到了减少,模型训练收敛速度更快,且平均训练时长缩短。而与基准模型 Xception 相比,Dis-Xception-CBAM 模型参数量仅增加了 1×10^6 ,而 Top 1 已高出 0.74 个百分点。证明本文提出的模型能够为野生菌种类识别研究提供理论依据。

4 结 论

以国内常见的 33 种野生菌种类图像为例,建立了国内常见野生菌数据集。提出了改进的 Xception 模型(Dis-Xception-CBAM),验证了 Disout 特征扰动形式优于传统 Dropout 方法,模型识别准确率更高,相较于未使用特征扰动方法,模型容错性更强。对比了未引入迁移学习与三种不同迁移学习方式下模型的训练准确度和平均训练时耗。通过引入迁移学习方法,并设置全层参数可训练,能够取得最高 Top 1,且平均训练时耗与未引入迁移学习相差不大。但冻结低层网络、仅训练高层特征提取层也能取得不错收益,相较于全层训练,Top 1 只低了 0.55 个百分点,而平均训练时耗降低了将近 1/2.5。比较了不同的数据集划分方式,发现训练集与验证集的最佳划分比例为 8 : 2,此时模型的 Top 1 为 96.32%,Top 5 为 99.61%。在相同的野生菌测试集上,所提出的 Dis-Xception-CBAM 模型的 Top 1 高达 96.32%,相较于参数量巨大的 InceptionResNetV2 与改进的轻量级基准模型 Xception 模型,Top 1 分别提高了 0.91 个百分点和 0.74 个百分点,且模型训练平均时耗较低。综合以上指标,所提出的 Dis-Xception-CBAM 模型在野生菌显式特征识别上优于其他 8 种模型。在后续研究中,将通过扩充数据集、优化模型结构等进一步提高模型稳健性。

参 考 文 献

- [1] Mao X L. Poisonous mushrooms and their toxins in China[J]. Mycosystema, 2006, 25(3): 345-363.
卯晓岚. 中国毒菌物种多样性及其毒素[J]. 菌物学报, 2006, 25(3): 345-363.
- [2] Li L J, Li G Y, Xie Q T. Research progress on poisonous mushroom toxins classification and recognition[J]. Chinese Journal of Food Hygiene, 2013, 25(4): 383-387.
李林静, 李高阳, 谢秋涛. 毒蘑菇毒素的分类与识别研究进展[J]. 中国食品卫生杂志, 2013, 25(4): 383-387.
- [3] Bai W M, Xing R R, Chen L P, et al. DNA barcoding for identification of toxic Amanita species[J]. Food Science, 2021, 42(4): 278-286.
白文明, 邢冉冉, 陈丽萍, 等. 基于 DNA 条形码技术的有毒鹅膏菌属物种鉴别[J]. 食品科学, 2021, 42(4): 278-286.
- [4] Mikola E, Geösel A, Stefanovits-Bányai É, et al.

- Quantitative determination of macro components and classification of some cultivated mushrooms using near-infrared spectroscopy [J]. *Journal of Food Processing and Preservation*, 2020, 44(7): e14540.
- [5] Zheng Y, Fan Q Q, Yu F F, et al. Rapid diagnosis of poisonous mushrooms in animal experiments[C]//Abstracts of 8th Chinese Congress on Biological Toxins, November 12-15, 2007, Nanning, China. [S.l.: s.n.], 2007: 132-133.
郑颖, 范泉水, 余芳芳, 等. 动物实验快速诊断有毒蘑菇的方法[C]//第八届中国生物毒素学术研讨会论文摘要集, 2007 年 11 月 12-15 日, 南宁. [出版地不详: 出版者不详], 2007: 132-133.
- [6] Liu B, Zhang Z D, Zhang T T. Identification of poisonous mushrooms based on Bayesian classification [J]. *Software Guide*, 2015, 14(11):60-62.
刘斌, 张振东, 张婷婷. 基于贝叶斯分类的毒蘑菇识别[J]. *软件导刊*, 2015, 14(11):60-62.
- [7] Fan G, Peng W, Sun S, et al. Discriminant method of mushroom toxicity based on support vector machine [J]. *Chinese Agricultural Science Bulletin*, 2015, 31(19): 232-236.
樊胥, 彭卫, 孙山, 等. 基于支持向量机的蘑菇毒性判别研究[J]. *中国农学通报*, 2015, 31(19): 232-236.
- [8] Zhang C Q. Recognition and research of poisonous mushroom based on machine learning [D]. Taigu, China: Shanxi Agricultural University, 2019.
张超群. 基于机器学习的毒蘑菇识别与研究[D]. 太谷: 山西农业大学, 2019.
- [9] Bache K, Lichman M. UCI machine learning repository [EB/OL]. (2013-12-05) [2020-09-30]. https://www.researchgate.net/publication/272825857_UCI_Machine_Learning_Repository.
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [11] Hu T, Li W H, Qin X X. Semantic segmentation of polarimetric synthetic aperture radar images based on multi-layer deep feature fusion [J]. *Chinese Journal of Lasers*, 2019, 46(2): 0210001.
胡涛, 李卫华, 秦先祥. 基于多层深度特征融合的极化合成孔径雷达图像语义分割[J]. *中国激光*, 2019, 46(2): 0210001.
- [12] Huang S, Li F F, Chen Q. Computed tomography image classification algorithm based on improved deep residual network [J]. *Acta Optica Sinica*, 2020, 40(3): 0310002.
黄盛, 李菲菲, 陈虬. 基于改进深度残差网络的计算断层扫描图像分类算法[J]. *光学学报*, 2020, 40(3): 0310002.
- [13] Zhang L, Jin X, Fu L Y, et al. Recognition method for weeds in rapeseed field based on faster R-CNN deep network [J]. *Laser & Optoelectronics Progress*, 2020, 57(2): 021508.
张乐, 金秀, 傅雷扬, 等. 基于 Faster R-CNN 深度网络的油菜田间杂草识别方法[J]. *激光与光电子学进展*, 2020, 57(2): 021508.
- [14] Song Y Q, Xie X, Liu Z, et al. Crop pests and diseases recognition method based on multi-level EESP model [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(8): 196-202.
宋余庆, 谢熹, 刘哲, 等. 基于多层 EESP 深度学习模型的农作物病虫害识别方法[J]. *农业机械学报*, 2020, 51(8): 196-202.
- [15] Yang S, Feng Q, Zhang J H, et al. Nondestructive classification of defects on potatoes based on lightweight convolutional network [J/OL]. *Food Science*. (2020-07-22) [2020-09-30]. <https://kns.cnki.net/kcms/detail/11.2206.TS.20200722.1413.090.html>.
杨森, 冯全, 张建华, 等. 基于轻量卷积网络的马铃薯外部缺陷无损分级[J/OL]. *食品科学*. (2020-07-22) [2020-09-30]. <https://kns.cnki.net/kcms/detail/11.2206.TS.20200722.1413.090.html>.
- [16] Fan S C, Yi X M, Li J, et al. Toadstool image recognition based on deep residual network and transfer learning [J]. *Chinese Journal of Sensors and Actuators*, 2020, 33(1): 74-83.
樊帅昌, 易晓梅, 李剑, 等. 基于深度残差网络与迁移学习的毒蕈图像识别[J]. *传感技术学报*, 2020, 33(1): 74-83.
- [17] Chollet F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1800-1807.
- [18] Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks? [EB/OL]. (2014-11-06) [2020-09-30]. <https://arxiv.org/abs/1411.1792v1>.
- [19] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [M] // Ferrari V, Hebert M, Sminchisescu C, et al. *Computer Vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 3-19.
- [20] Tang Y H, Wang Y H, Xu Y X, et al. Beyond dropout: feature map distortion to regularize deep neural networks [C] // Proceedings of the AAAI Conference on Artificial Intelligence, February 7-12, 2020, New York. Palo Alto, California, USA: AAAI Press, 2020, 34(4): 5964-5971.

- [21] Liu Y, Lei Y B, Fan J L, et al. Survey on image classification technology based on small sample learning[J]. *Acta Automatica Sinica*, 2021, 47(2): 297-315.
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.
- [22] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 1-9.
- [23] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C] // Proceedings of the 32nd International Conference on Machine Learning, July 6-11, 2015, Lille, France. Cambridge, Massachusetts: MIT Press, 2015.
- [24] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 2818-2826.
- [25] Zhang H M, Quan K, Yang Y N, et al. Super-resolution reconstruction of DEM in mountain area based on deep residual network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(1): 178-184.
张宏鸣, 全凯, 杨亚男, 等. 基于深层残差网络的山区 DEM 超分辨率重构[J]. *农业机械学报*, 2021, 52(1): 178-184.
- [26] Chen H S, Chen J J. Study on construction of IOT network intrusion detection classification model and optimization based on combination of ResNet and bidirectional LSTM network [J]. *Journal of Hunan University (Natural Sciences)*, 2020, 47(8): 1-8.
陈红松, 陈京九. 基于 ResNet 和双向 LSTM 融合的物联网入侵检测分类模型构建与优化研究[J]. *湖南大学学报(自然科学版)*, 2020, 47(8): 1-8.
- [27] Zheng T Y, Wang S Y, Wang G X, et al. High-order natural transition simulation method based on deep residual network [J]. *Acta Physica Sinica*, 2020, 69(20): 204701.
郑天韵, 王圣业, 王光学, 等. 基于深度残差网络的高精度自然转捩模拟方法[J]. *物理学报*, 2020, 69(20): 204701.
- [28] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks [M] // Leibe B, Matas J, Sebe N, et al. *Computer Vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9908: 630-645.
- [29] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 7132-7141.
- [30] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting [J]. *Journal of Machine Learning Research*, 2014, 15: 1929-1958.
- [31] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning [C] // Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA. New York: AAAI Press, 2017: 4278-4284.