

融合多尺度和迁移学习的蝴蝶种类识别

李 飞, 严春雨

贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2022年6月23日; 录用日期: 2022年8月1日; 发布日期: 2022年8月9日

摘 要

蝴蝶在生态系统稳定中发挥着重要作用, 不仅能帮助植物传播花粉, 还能对其生存环境变化做出指示。针对自然环境中蝴蝶种类识别率低的问题, 本文提出一种融合多尺度和迁移学习的识别模型。首先, 使用焦点损失函数解决数据集分布不平衡问题; 其次, 引入迁移学习提升识别准确率、加快模型收敛; 最后, 引入空洞空间卷积池化金字塔, 提取蝴蝶图像不同尺度信息。实验结果表明, 本研究所提方法平均识别准确率达到98.03%, 较原始模型提升了4.88%, 相较其他对比模型也取得较大优势, 可为自然环境中蝴蝶种类识别提供技术支持。

关键词

蝴蝶, 多尺度, 迁移学习, 特征融合

Butterfly Species Recognition Integrated with Multiscale and Transfer Learning

Fei Li, Chunyu Yan

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: Jun. 23rd, 2022; accepted: Aug. 1st, 2022; published: Aug. 9th, 2022

Abstract

Butterflies play an important role in ecosystem stability, not only helping plants spread pollen, but also signaling changes to their environment. Aiming at the low recognition rate of butterfly species in the natural environment, this paper proposes a recognition model that integrates multi-scale and transfer learning. Firstly, the focal loss function is used to solve the problem of imbalanced dataset distribution. Secondly, transfer learning is introduced to improve the recognition accuracy and speed up the model convergence. Finally, ASPP is introduced to enable the model to capture butterfly image information from different scales. The experimental results show that the average

recognition accuracy of the method proposed in this study reaches 98.03%, which is 4.88% higher than the original model. Compared with other comparative models, it also achieves great advantages. It can provide technical support for the identification of butterfly species in the natural environment.

Keywords

Butterfly, Multiscale, Transfer Learning, Fusion of Features

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

蝴蝶隶属于昆虫纲鳞翅目,种类繁多,全世界已知的蝴蝶约有 18000 种[1],其中,中国蝴蝶占 2000 多种[2]。蝴蝶色彩艳丽,具有极高的观赏价值;又因为蝴蝶对其生存环境下的气候变换敏感,是气候变化的指示性生物,可帮助气候学家预测全球气候变化;但是部分蝴蝶对农作物有害,会影响人们的生产与生活。因此,对自然环境中的蝴蝶种类识别具有重要意义。

近年来,使用传统方法对蝴蝶种类进行识别已经取得了一些成果。Kaya 等人[3]使用局部二值模式和灰度共生矩阵对 190 张蝴蝶图像纹理信息进行了提取,然后利用极限学习机进行分类。Li 等人[4]使用多尺度曲率直方图和灰度共生矩阵提取蝴蝶形状及纹理特征,并结合 KNN 算法对蝴蝶种类识别。薛安康等人[5]利用灰度共生矩阵提取蝴蝶纹理信息,结合 KNN 算法进行分类,对 10 种蝴蝶标本照片进行了有效识别。上述方法均是在单一实验背景下进行,但因蝴蝶纹理信息复杂,形状和大小各异,人为设计的特征提取器很难将蝴蝶信息表达出来,使识别准确率受到影响,不能为自然环境中蝴蝶种类识别提供技术支持。

随着计算机视觉技术的发展,研究者们逐渐将深度学习技术应用到蝴蝶种类识别中来,并取得了一定效果。谢娟英等人[6]制作了包含蝴蝶生态照和标本照的中国蝴蝶数据集,使用 FasterR-CNN 进行种类识别,模型取得的最低 mAP 达到 60%。Feng 等人[7]设计了一种基于生成对抗网络和残差网络相结合的方法识别蝴蝶种类,在 ResNet50 上取得了 90.34%的识别准确率。李策等人[8]提出了一种基于迁移学习和可变形卷积网络的蝴蝶识别方法。虽然上述方法已经在蝴蝶种类识别上取得一定效果,但仍然存在识别准确率较低的问题。

综上所述,本研究构建了自然环境中蝴蝶图像数据集,提出了一种融合多尺度和迁移学习的蝴蝶种类识别方法。首先,针对数据集分布不均衡的问题,使用焦点损失函数替换交叉熵损失函数;其次针对蝴蝶图像特征,将在 ImageNet 数据集上训练好的 EfficientNet-B0 [9]模型权重迁移到本文模型,加快网络收敛速度,提升识别准确率;最后,针对蝴蝶尺度多样问题,引入空洞空间卷积池化金字塔,提升网络对各类蝴蝶整体特征提取能力。实验结果表明,本文方法对蝴蝶种类识别的效果显著,平均识别准确率达到了 98.03%,可为自然环境中蝴蝶种类识别提供技术支持。

2. 蝴蝶种类识别模型

2.1. EfficientNet

EfficientNet 网络利用复合模型缩放方法来平衡网络深度、宽度和分辨率,既提高了模型识别准确率

又保证了识别速度。考虑到蝴蝶的拟态性会导致自然环境中的蝴蝶存在复杂的背景干扰, 影响网络对蝴蝶种类的识别, 且本文数据集较小, 因此本研究选用轻量化、高识别率的 EfficientNet-B0 作为基础模型 [10]。EfficientNet-B0 网络结构参数如表 1 所示, 由两个普通卷积、7 个移动倒置瓶颈卷积(Mobile inverted Bottleneck Convolution, MBConv) [11]、平均池化层和全连接层组成。EfficientNet-B0 网络使用残差结构增加网络深度, 可提取到蝴蝶丰富的细粒度特征, 提高识别率。由表 1 可知, EfficientNet-B0 网络的主要模块是 MBConv, 使用深度可分离卷积减小模型参数数量与运算量, 结合 SE 注意力模块, 使网络更加关注蝴蝶所在通道信息, 降低背景干扰权重数值, 进而提升网络特征提取能力, MBConv 模块结构图如图 1 所示。

Table 1. Network structure parameters of EfficientNet-B0

表 1. EfficientNet-B0 网络结构参数

Stage	Block name	Block number	Resolution	Channels
1	Conv 3×3	1	224×224	32
2	MBConv, $k3 \times 3$	1	112×112	16
3	MBConv, $k3 \times 3$	2	112×112	24
4	MBConv, $k5 \times 5$	2	56×56	40
5	MBConv, $k3 \times 3$	3	28×28	80
6	MBConv, $k5 \times 5$	3	14×14	112
7	MBConv, $k5 \times 5$	4	14×14	192
8	MBConv, $k3 \times 3$	1	7×7	320
9	Conv 1×1 & Pooling & FC	1	7×7	1280

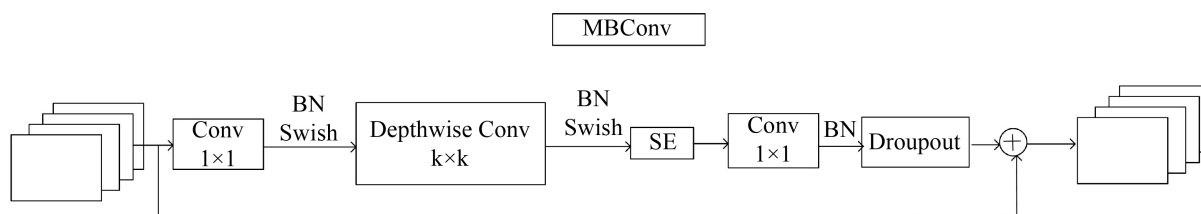


Figure 1. Structure diagram of MBConv

图 1. MBConv 结构图

2.2. 迁移学习

迁移学习利用在其他任务上训练得到的结果辅助新任务学习, 使模型速度收敛加快、减少模型训练时间。本研究将在 ImageNet 数据集上实验得到的模型权重参数迁移到蝴蝶种类识别实验中, 只需要微调参数, 就可加快网络收敛, 提升蝴蝶种类识别准确率。因 EfficientNet 网络为轻量级卷积神经网络, 模型参数量较少, 因此不需要冻结任何层, 直接训练网络中所有参数。

2.3. 空洞空间卷积池化金字塔

因蝴蝶形状大小各异, 原始模型对尺度不一的蝴蝶整体信息提取能力较弱, 因此本文在 EfficientNet-B0 网络的特征提取层后引入空洞空间卷积池化金字塔(ASPP) [12]。ASPP 主要由一个 1×1 的卷积层、采样率分别为 6、12、18 的空洞卷积和全局平均池化并联组成, 其结构图如图 2 所示, 其中

GAP 表示全局平均池化。在 EfficientNet-B0 网络特征提取层后引入 ASPP, 网络可获得多种不同尺度的感受野, 感受野越大, 全局特征更丰富; 感受野越小, 提取到的局部特征和细节特征越丰富, 捕获到的蝴蝶具有更多全局和局部细节特征, 进而网络可实现对任意大小蝴蝶种类识别的目的。

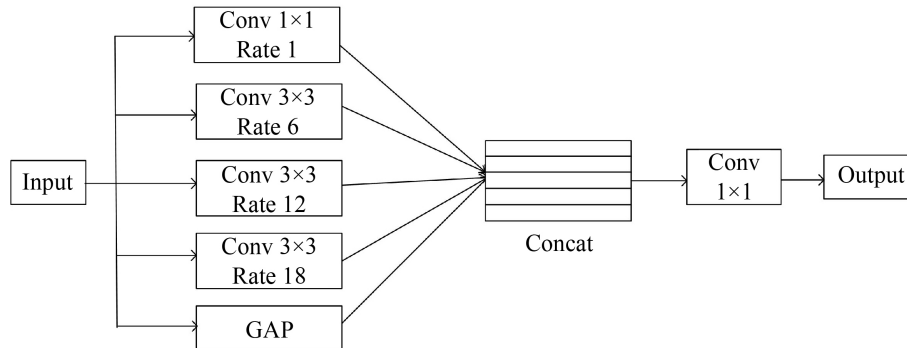


Figure 2. Structure diagram of ASPP
图 2. ASPP 结构图

2.4. FcoalLoss 损失函数

由表 1 知, 本文数据集种类分布不均衡。比如, 最多的钩凤蝶有 826 张, 最少的东方菜粉蝶只有 232 张, 钩凤蝶与东方菜粉蝶数量之比达到 3.6:1。这种情况会导致模型对数量多的蝴蝶识别率高, 对数量少的识别率低, 影响整体识别效果。为此, 使用焦点损失函数(FocalLoss) [13] 替换交叉熵损失函数。与交叉熵损失函数相比, 焦点损失函数会为难区分的蝴蝶种类添加较大权重, 为易区分的蝴蝶添加较小权重:

$$FL_{(p_i)} = -\alpha_i \log(p_i) \quad (1)$$

$$p_i = \begin{cases} p, & \text{if } y = 1 \\ 1 - p, & \text{otherwise} \end{cases} \quad (2)$$

式(1)中 α_i 可平衡正负样本比例不均问题, 对于数量较少的种类, 其值较大。 γ 的值根据样本复杂性进行调节。 p_i 表示经过 Softmax 函数的输出概率。与交叉熵损失函数相比, 焦点损失函数只需调节 α_i 和 γ 的值, 就能在一定程度上解决蝴蝶种类分布不均的问题。

2.5. 模型总体架构

蝴蝶表面的纹理信息十分复杂, 本研究以特征提取能力较强的轻量级 EfficientNet-B0 网络为基础; 将在 ImageNet 数据集上训练好的参数权重迁移到本研究中, 加快模型收敛速度、提升模型识别准确率; 考虑到蝴蝶形状大小多样, 在 EfficientNet-B0 网络的特征提取层后引入 ASPP 以获得不同尺度蝴蝶的特征, 实现多尺度特征融合[14], 以此来增强网络特征提取能力, 提高模型识别率。改进后的模型整体架构如图 3 所示, 改进后的模型命名为 EfficientNet_B0_A。

3. 实验结果与分析

3.1. 制作蝴蝶数据集

现有蝴蝶种类识别研究中大多使用标本图像, 且涉及种类数量较少。针对此, 本文构建了自然环境中蝴蝶数据集, 如图 4 所示。

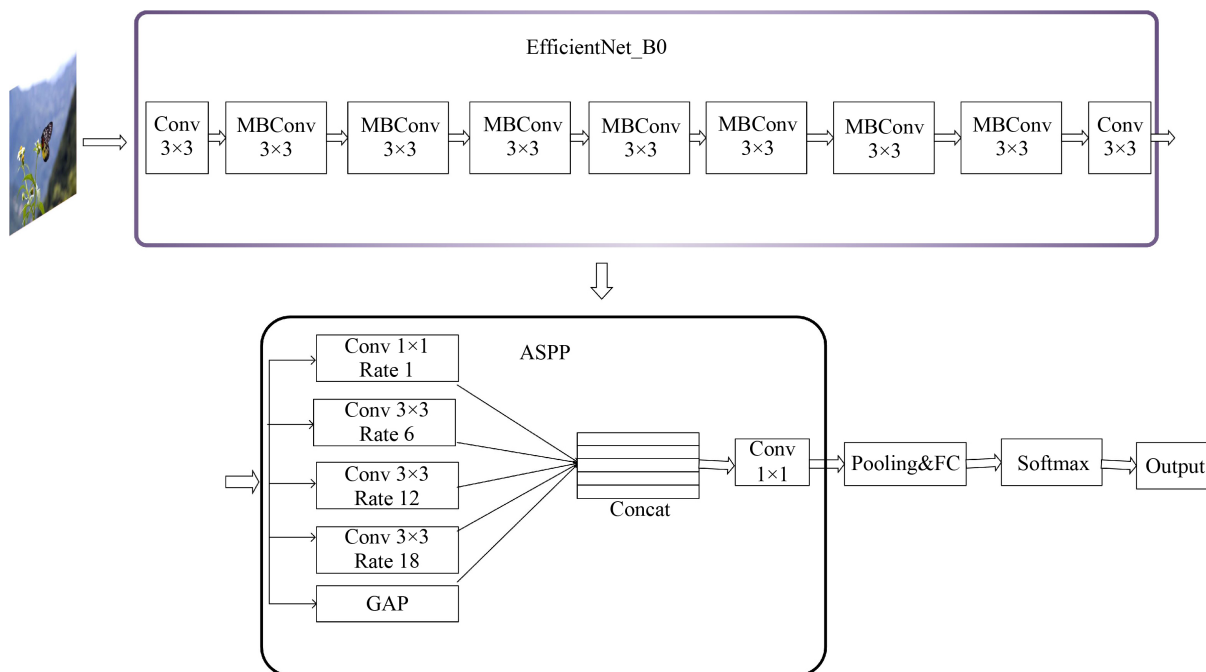
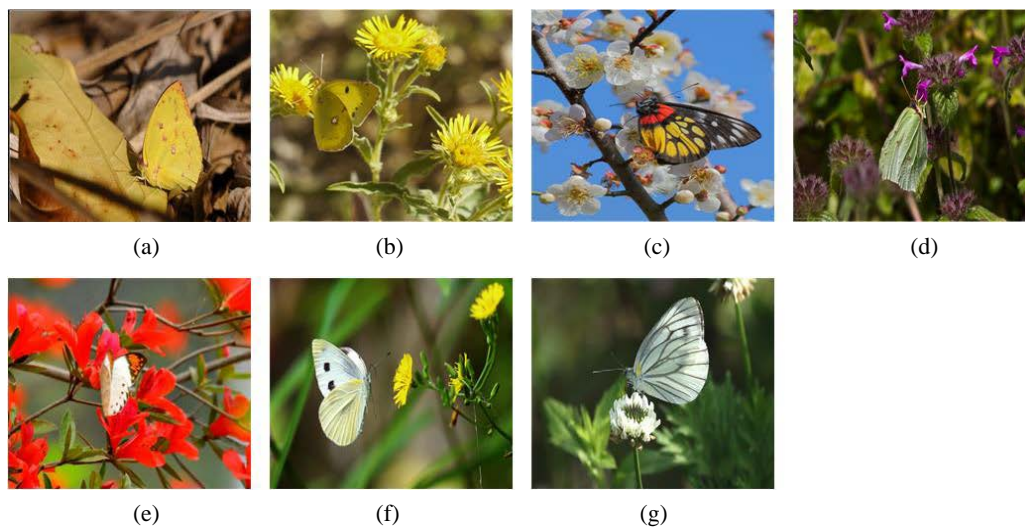


Figure 3. Overall structure of the model

图 3. 模型整体架构



注: (a)~(g)依次为迁粉蝶、斑缘豆粉蝶、报喜班粉蝶、钩粉蝶、鹤顶粉蝶、东方菜粉蝶和黑纹粉蝶。

Figure 4. Butterfly images

图 4. 蝴蝶图像

该数据集由粉蝶科下 7 种蝴蝶生态图像组成, 共 2716 张。数据集来源有: 在贵州省贵阳市使用佳能数码相机拍摄; 从 kaggle 官方数据网站(<https://www.kaggle.com>)获取。本文所用蝴蝶生态图像均经过昆虫学家识别确认, 数据集的详细情况如表 2 所示。

3.2. 图像数据预处理

为满足网络输入图像要求, 像素统一为 224×224 。按 7:3 划分为训练集和验证集。由于数据样本较

少, 为防止训练过程中出现过拟合, 使用椒盐噪声、左右翻转、增强亮度、降低亮度等数据增强方法对训练集进行扩充, 扩充后图像数据集共 9525 张。

Table 2. Butterfly dataset
表 2. 蝴蝶数据集

Butterfly species	Number of images	The number of training sets after data augmentation	Number of validation sets
Catopsilia_pomona	434	1520	130
Colias_erate	253	890	75
Delias_pasithoe	315	1105	94
Gonepteryx_rhamni	826	2895	247
Hebomoia_glaucippe	408	1430	122
Pieris_canidia	232	815	69
Pieris_melete	248	870	74

3.3. 实验环境以及参数设置

实验训练与验证均在同一平台上进行。实验硬件环境: CPU 为 Intel(R) Core(TM) i5-11400F @ 2.6GHZ 2.59 GHZ, 内存为 16 GB; GPU 为 Nvidia RTX3060Ti8GB。软件环境: Pytorch 版本为 1.7.1; Python 版本为 3.8。批量大小(Batch size)设置为 16, 迭代次数(Epoch)设置为 100, 使用随机梯度下降算法(Stochastic Gradient Descent, SGD)作为损失函数优化器, 初始学习率设置为 0.01, 学习率随着训练轮次变化, 每 10 轮次衰减 0.5。

3.4. 评价指标

本研究使用分类问题中常用的平均准确率、精确率、召回率作为评价标准。

1) 平均准确率

平均识别准确率(Average recognition accuracy /%, Arc)指的是验证集分类正确的样本占验证集总样本数的比例, 计算方法如式(3)所示:

$$\text{Arc} = \frac{1}{N} \sum_{i=1}^N \frac{X_i}{Y_i} \times 100\% \quad (3)$$

式中: N 为数据类别总数, 本研究中 $N = 7$; Y_i 为第 i 类的数量; X_i 为第 i 类中分类正确数量。

2) 精确率

精确率(Precision)表示实际正确的蝴蝶数量占全部预测为正的蝴蝶数量的比例, 计算方法如(4)所示:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (4)$$

3) 召回率

召回率(Recall)表示真正正确的蝴蝶数量占实际为正的蝴蝶数量的比例, 计算方式如(5)所示:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (5)$$

4) F1-score

F1-score 综合考虑了精确率和召回率, 是两者的调和平均, 计算公式如式(6)所示,

$$F1\text{-score} = \frac{2TP}{2TP + FP + FN} \times 100\% \quad (6)$$

3.5. 结果分析

3.5.1. 不同识别模型对比实验

为验证本文模型(EfficientNet_B0_A)有效性, 选取 ResNet [15]、DenseNet [16]和 MobileNetV3 [17]模型与本文模型对比, 实验结果如表 3 所示。

Table 3. Comparative experiment of different recognition models

表 3. 不同识别模型对比实验

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score
EfficientNet_B0	93.15	93.02	93.11	93.06
DenseNet-121	93.09	93.01	93.08	93.05
MobileNetV3_large	93.12	93.04	93.10	93.07
ResNet-18	92.65	92.31	92.33	92.32
EfficientNet_B0_A	98.03	97.63	97.62	97.62

由表 3 知, 四种对比网络中 EfficientNet_B0 效果最好, 平均识别率达到 93.15%。本文模型识别率较基础模型提升 4.88%, 达到 98.03%, 且本文模型的精确率、召回率和 F1-score 也均优于四种对比模型。

EfficientNe_B0 与 EfficientNe_B0_A 对比实验的验证集识别准确率和损失值曲线如图 5 所示, 其中绿线代表 EfficientNe_B0, 蓝线代表 EfficientNe_B0_A。由图 5 知, 原模型在 60 轮左右开始收敛, 而改进后模型在 20 轮已经收敛, 说明引入权重迁移后, 模型收敛速度加快。且本文模型在验证集上验证所得平均识别率更高, 损失值更小。

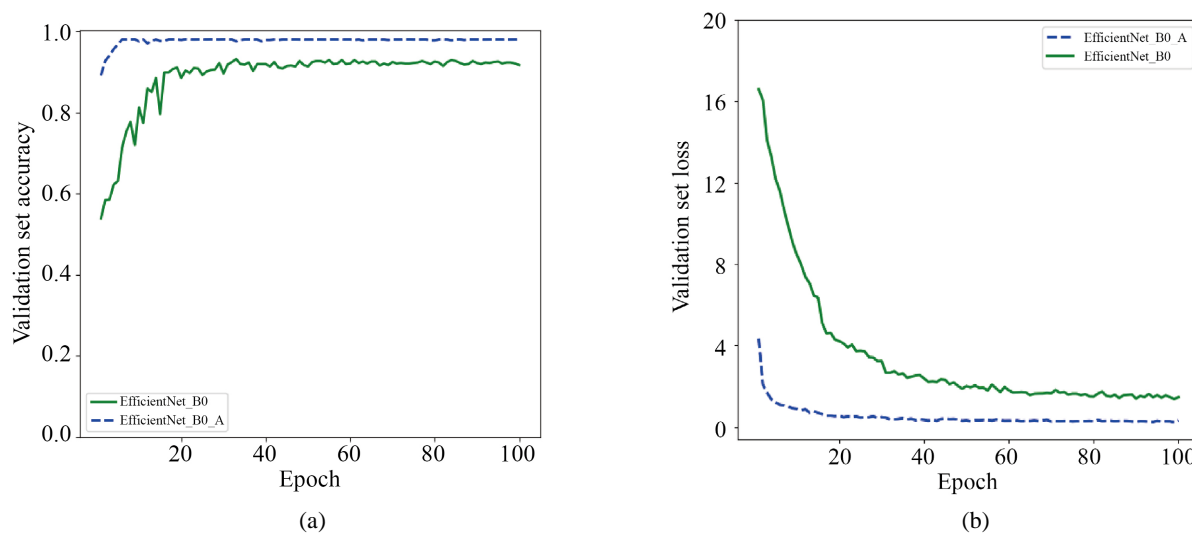


Figure 5. Recognition accuracy and loss value of model validation set. (a) Validation set accuracy; (b) Validation set loss value

图 5. 模型验证集识别准确率和损失值。(a) 验证集准确率; (b) 验证集损失值

3.5.2. 消融实验

为验证 FocalLoss、迁移学习和空洞空间卷积池化金字塔的有效性, 本文在 EfficientNet_B0 网络基础上, 通过逐项增加各模块进行验证, 实验结果如表 4 所示。表中 FL、TL、ASPP 分别表示 FocalLoss 损失函数、迁移学习和空洞空间卷积池化金字塔, 其中组 1 模型使用的是交叉熵损失函数。

Table 4. Ablation experiment

表 4. 消融实验

Group	FL	TL	ASPP	Accuracy (%)
1	—	—	—	93.15
2	√	—	—	93.56
3	√	√	—	97.25
4	√	√	√	98.03

注: √表示使用该方法, —表示没有使用该方法。

由表 4 可知, 组 1 和组 2 对比可知, 引入 FocalLoss 损失函数后, 模型识别准确率较基准模型提升 0.41 个百分点, 说明数据分布不均问题对模型性能的影响得到一定缓解; 由组 2 和组 3 对比可知, 引入迁移学习后, 模型识别准确率大幅度提升, 达到 97.25%, 且由图 5 可知, 模型收敛速度更快; 由组 3 和组 4 对比可得, 引入 ASPP 后网络能从多个尺度提取蝴蝶信息, 进一步提升了识别准确率。

3.5.3. 实验结果分析

本文模型在蝴蝶数据集上得出的混淆矩阵如图 6 所示, 其横坐标为真实标签, 纵坐标为预测标签, 主对角线上数字为预测正确的样本数, 其余数字均为预测错误的样本数。由图 6 知, 迁粉蝶 (Catopsilia_pomona) 较易被误判为其他类别, 因为其分为有纹型和无纹型, 有纹型迁粉蝶易与斑缘豆粉蝶混淆 (Colias_erate), 无纹型迁粉蝶易与钩粉蝶 (Gonepteryx_rhamni) 混淆。

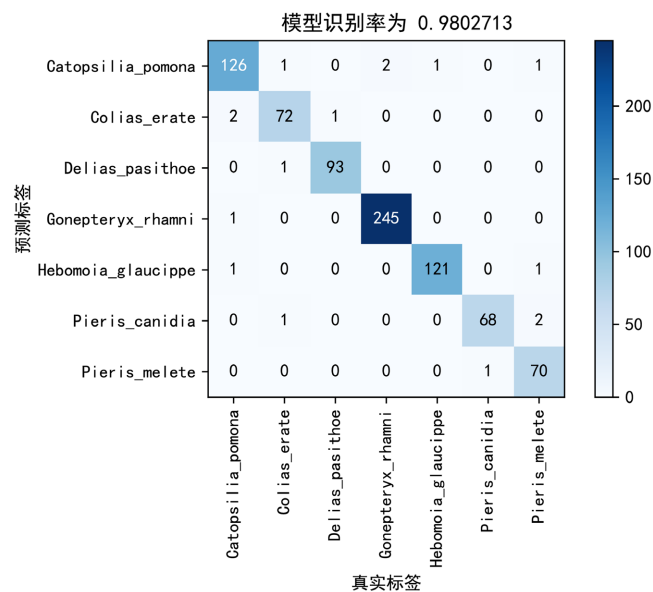


Figure 6. Confusion matrix of the validation set

图 6. 验证集混淆矩阵

4. 结语

针对自然环境中蝴蝶种类识别率低的问题, 本文提出融合多尺度和迁移学习的蝴蝶种类识别模型。实验结果表明, 使用焦点损失函数可以从一定程度上解决样本分布不均衡的问题; 引入迁移学习可加速网络快速收敛并提升模型准确率; 引入 ASPP 后模型能够根据蝴蝶形状和大小从不同尺度捕捉信息, 提取出各类蝴蝶图像整体特征, 提高识别准确率。本研究所提方法的识别准确率、精确率和召回率均优于四种对比网络, 对粉蝶科的 7 种蝴蝶平均识别准确率达到 98.03%。在未来的研究中, 我们会扩充蝴蝶数据集, 并寻找新的改进算法, 提高模型泛化能力。

参考文献

- [1] 谢娟英, 鲁银圆, 孔维轩, 许升全. 基于改进 RetinaNet 的自然环境中蝴蝶种类识别[J]. 计算机研究与发展, 2021, 58(8): 1686-1704.
- [2] 周爱明, 马鹏鹏, 席天宇, 王江宁, 冯晋, 邵泽中, 陶玉磊, 姚青. 基于深度学习的蝴蝶科级标本图像自动识别[J]. 昆虫学报, 2017, 60(11): 1339-1348.
- [3] Kaya, Y., Kayci, L., Tekin, R. and Faruk Ertuğrul, Ö. (2014) Evaluation of Texture Features for Automatic Detecting Butterfly Species Using Extreme Learning Machine. *Journal of Experimental & Theoretical Artificial Intelligence*, **26**, 267-281. <https://doi.org/10.1080/0952813X.2013.861875>
- [4] Li, F. and Xiong, Y. (2018) Automatic Identification of Butterfly Species Based on HoMSC and GLCMoIB. *The Visual Computer*, **34**, 1525-1533. <https://doi.org/10.1007/s00371-017-1426-1>
- [5] Xue, A.K., Li, F. and Xiong, Y. (2019) Automatic Identification of Butterfly Species Based on Gray-Level Co-Occurrence Matrix Features of Image Block. *Journal of Shanghai Jiaotong University (Science)*, **24**, 220-225. <https://doi.org/10.1007/s12204-018-2013-y>
- [6] 谢娟英, 侯琦, 史颖欢, 吕鹏, 景丽萍, 庄福振, 张军平, 谭晓阳, 许升全. 蝴蝶种类自动识别研究[J]. 计算机研究与发展, 2018, 55(8): 1609-1618.
- [7] Feng, M. and Wen, Y.X. (2021) Butterfly Image Generation and Recognition Based on Improved Generative Adversarial Networks. 2021 4th International Conference on Robotics, Control and Automation Engineering (RCAE), Wuhan, 4-6 November 2021, 40-44.
- [8] 李策, 张栋, 杜少毅, 朱子重, 贾盛泽, 曲延云. 一种迁移学习和可变形卷积深度学习的蝴蝶检测算法[J]. 自动化学报, 2019, 45(9): 1772-1782.
- [9] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling Forconvolutional Neural Networks. 36th International Conference on Machine Learning (ICML 2019), Long Beach, 9-15 June 2019, 6105-6114.
- [10] 甘雨, 郭庆文, 王春桃, 梁炜健, 肖德琴, 吴惠彝. 基于改进 EfficientNet 模型的作物害虫识别[J]. 农业工程学报, 2022, 38(1): 203-211.
- [11] Tan M., Chen B., Pang R., Vasudevan, V., Sandler, M., Howard, A., et al. (2019) MnasNet: Platform-Aware Neural Architecture Search for Mobile. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 15-20 June 2019, Long Beach, 2815-2823. <https://doi.org/10.1109/CVPR.2019.00293>
- [12] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L. (2018) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 834-848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- [13] Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollár, P. (2017) Focal Loss for Dense Object Detection. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 22-29 October 2017, 2999-3007. <https://doi.org/10.1109/ICCV.2017.324>
- [14] 谭海, 张荣军, 樊文锋, 张一帆, 徐航. 融合多尺度特征的国产光学影像辐射异常分类检测[J/OL]. 自然资源遥感:1-9. <http://kns.cnki.net/kcms/detail/10.1759.p.20220525.1739.008.html>, 2022-06-14.
- [15] He, K.M., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [16] Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017) Densely Connected Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 2261-2269. <https://doi.org/10.1109/CVPR.2017.243>

- [17] Howard, A., Sandler, M., Chu, G., Wang, W., Chen, L.-C., Tan, M., *et al.* (2019) Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 1314-1324. <https://doi.org/10.1109/ICCV.2019.00140>