

应用迁移学习的提高小数据集蝴蝶种类识别精确度的方法

郭梓凡

沈阳建筑大学, 计算机科学与工程学院, 辽宁 沈阳

收稿日期: 2023年3月16日; 录用日期: 2023年4月17日; 发布日期: 2023年4月27日

摘要

野生蝴蝶品种繁多、分布广泛, 且其对于生存环境的变化较为敏感, 监测范围内的蝴蝶种群生存情况, 可以为监测衡量该地区生态环境变化和衡量生态环境的质量提供帮助。现有的蝴蝶种类数据集大多数数据量较小, 识别的准确率较低。针对该问题, 本文提出了基于VGG16模型的迁移学习及微调方法, 对蝴蝶图像进行种类识别, 以提高识别准确率。首先对蝴蝶种类数据集进行数据增强, 再利用大型图像数据集对VGG16模型进行预训练, 对预训练模型进行参数迁移, 对卷积层和池化层进行“冻结”, 修改全连接层和分类层, 再“解冻”部分卷积层对参数进行微调, 得到识别结果。实验证明: 通过迁移学习加微调的方法, 该网络对于蝴蝶种类识别的准确率得到有效提高, 识别准确率从最初的76.67%, 在迁移及微调后达到83.95%。

关键词

蝴蝶种类, 迁移学习, VGG16网络, 微调

A Method of Improving the Accuracy of Butterfly Species Identification in Small Data Sets Using Transfer Learning

Zifan Guo

School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang Liaoning

Received: Mar. 16th, 2023; accepted: Apr. 17th, 2023; published: Apr. 27th, 2023

Abstract

There are many species and wide distribution of wild butterflies, and they are sensitive to changes

in the living environment. The survival of butterfly populations within the monitoring range can provide help for monitoring and measuring changes in the ecological environment in the region and measuring the quality of the ecological environment. Most of the existing butterfly species data sets have small data volume and low recognition accuracy. In order to solve this problem, this paper proposes a migration learning and fine-tuning method based on the VGG16 model to recognize the butterfly image in order to improve the recognition accuracy. First, the butterfly species data set is enhanced, and then the large image data set is used to pre-train the VGG16 model, migrate the parameters of the pre-training model, “freeze” the convolution layer and pool layer, modify the full connection layer and classification layer, and “thaw” part of the convolution layer to fine-tune the parameters to get the recognition results. The experiment shows that the accuracy of the network for butterfly species recognition has been effectively improved by the method of migration learning and fine tuning, and the recognition accuracy has reached 83.95% after migration and fine tuning from 76.67% at the beginning.

Keywords

Butterfly Species, Transfer Learning, VGG16 Network, Fine Tuning

Copyright © 2023 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

蝴蝶属于节肢动物门，昆虫纲，鳞翅目分类下的锤角亚目。全世界已知的蝴蝶种类大概 18,000 种，中国已知蝶类有 2300 多种[1]。因其轻盈的身体的斑斓多彩的翅膀，被人们形容为“会飞的花朵”。且蝴蝶在生态系统中有着不可或缺作用，能够维护生态系统的稳定，区域内如果蝴蝶的生存状况良好，在一定程度上能够表明该地区的生态环境较为优良。但蝴蝶种类繁多，花纹复杂，野外图像难以获取，如果只靠人工识别，不仅对专业性要求较高，且需要花费大量时间和精力。为了能够较为准确和快速地识别出蝴蝶种类，本文提出了基于 VGG16 网络模型的迁移学习及微调的方法。

最初，往往通过传统的模式识别方法，对蝴蝶图像进行预处理，提取蝴蝶翅膀纹理、形态、颜色等特征，对分类器进行训练，以达到识别蝴蝶种类的目的。张建伟[2]对 43 种蝴蝶标本的蝴蝶前翅及后翅的各 13 种，共 26 种特征进行特征识别，来实现蝴蝶分类。刘芳[3]基于蝴蝶的颜色特征开发了蝴蝶自动识别系统，对 43 种蝴蝶达到较高的识别率。Kang 等[4]使用分支长度相似度(BLS)熵，将不同观察角度的蝴蝶形状特征作为神经网络的输入特征，以实现蝴蝶种类识别。Kaya 等先后通过基于极限学习机(ELM)和 5 个频率 4 个方向的 Gabor 滤波器(GFs) [5]、基于人工神经网络(ANN)和不同角度和距离的灰度共生矩(GLCM) [6]、基于局部二值模式(LBP)和人工神经网络(ANN) [7]的方法对蝴蝶纹理特征进行识别。

自然界中的野生蝴蝶品种繁复，有些品种的蝴蝶雌雄个体间就存在较大差异，有些品种的蝴蝶的翅膀正反面花纹极为不同，而受一些环境或地区的影响，一些不同种的蝴蝶翅膀花纹或者形态会较为相似。因此采用传统的特征提取方式就会存在泛化能力较低、识别精确度不足等问题。

近年来，随着人工智能的飞速发展，深度学习在许多领域都取得了较好的成绩，谢娟英等先基于 Faster R-CNN 深度学习框架实现了对野外蝴蝶种类的识别，模型的 mAP 最低值接近 60% [8]。后引入可变形卷积

增强了目标检测算法 RetinaNet, 在蝴蝶种类识别过程中取得了较为不错的效果[9]。周文进等[10]为提高蝴蝶特征提取的准确度, 提出内嵌通道注意力 MultiSE1D 识别网络, 对 YOLOv3 模型的主干网络进行了改进。

但含有野外蝴蝶照片的数据集较难获取, 且数据集中的数据量较小。使用卷积神经网络对其进行训练容易产生过拟合现象。迁移学习在很多领域都有较广泛的应用, 采用迁移学习可以有效的节省训练时间, 很大程度上减轻过拟合现象的发生, 提高泛化能力, 适合应对数据量较小的问题。

因此, 本文利用 VGG16 模型在 Imagenet 大型数据集上进行预训练, 对预训练模型进行参数迁移, 对数据集进行数据增强, 再通过“解冻”部分卷积层对模型参数进行微调, 以实现小数据集蝴蝶种类识别问题。

2. 数据与方法

2.1. 数据集

本文使用的数据集由 20 种蝴蝶组成, 共 1866 张图像。其中, 每种蝴蝶的图像数量都在 50~200 张之间。且每种蝴蝶图像中都既包括标本图像, 也包括野外蝴蝶图像。如图 1 所示。



Figure 1. Specimen image and field image in data set
图 1. 数据集中的标本图像和野外图像

2.2. VGG16 模型

VGGNet [11]模型由牛津大学几何组(Visual Geometry Group)提出, VGG16 为其中的一种典型网络结构。VGG16 模型除池化层外共有 16 个含权重系数的权重层, 为 13 个卷积层和 3 个全连接层。模型由 5 个模块和最后的 3 个全连接层组成, 前两个模块由 2 个卷积层和 1 个池化层组成, 后三个模块由 3 个卷积层和 1 个池化层组成。卷积层的卷积核均为 3×3 大小, 步长为 1, 五个模块的卷积核数分别为 64,128,256,512,512。池化层的大小是 2×2 , 为最大池化层, 步长为 2。3 个全连接层的输出节点数分别为 4096,4096,1000。最后通过 softmax 函数输出分类结果。所有卷积层和全连接层后都有 ReLU 激活函数, 加快网络收敛速度, 避免梯度消失。VGG16 模型网络结构及参数情况如表 1 所示。

Table 1. Network structure and parameters of VGG16 model
表 1. VGG16 模型网络结构及参数情况

| Layer (type) | Output Shape | Param |
|-------------------------|-----------------|-------------|
| input (InputLayer) | [3, 224, 224] | 0 |
| block1-conv1 (Conv2d) | [64, 224, 224] | 1792 |
| block1-conv2 (Conv2d) | [64, 224, 224] | 36,928 |
| block1-pool (MaxPool2d) | [64, 112, 112] | 0 |
| block2-conv1 (Conv2d) | [128, 112, 112] | 73,856 |
| block2-conv2 (Conv2d) | [128, 112, 112] | 147,584 |
| block2-pool (MaxPool2d) | [128, 56, 56] | 0 |
| block3-conv1 (Conv2d) | [256, 56, 56] | 295,168 |
| block3-conv2 (Conv2d) | [256, 56, 56] | 590,080 |
| block3-conv3 (Conv2d) | [256, 56, 56] | 590,080 |
| block3-pool (MaxPool2d) | [256, 28, 28] | 0 |
| block4-conv1 (Conv2d) | [512, 28, 28] | 1,180,160 |
| block4-conv2 (Conv2d) | [512, 28, 28] | 2,359,808 |
| block4-conv3 (Conv2d) | [512, 28, 28] | 2,359,808 |
| block4-pool (MaxPool2d) | [512, 14, 14] | 0 |
| block5-conv1 (Conv2d) | [512, 14, 14] | 2,359,808 |
| block5-conv2 (Conv2d) | [512, 14, 14] | 2,359,808 |
| block5-conv3 (Conv2d) | [512, 14, 14] | 2,359,808 |
| block5-pool (MaxPool2d) | [512, 7, 7] | 0 |
| FC1 (Linear) | [4096, 1, 1] | 102,764,544 |
| FC2 (Linear) | [4096, 1, 1] | 16,781,312 |
| FC3 (Linear) | [4096, 1, 1] | 4,097,000 |

2.3. 迁移学习与微调

一个神经网络想要达到较好的分类效果，往往需要大量的数据对其进行训练。在数据量较小的情况下，可以采用迁移学习[12]的方法。迁移学习就是将在一项任务中所获得知识用于其他类似任务中，即将预训练好的模型参数迁移到需要训练的模型中，进行再次训练。迁移学习可以极大地节省训练时间，使模型具有较好的初始性能，且有更强的泛化能力。因此，迁移学习广泛应用与诸多领域，孟伟[13]等通过多角度的分析和比较，总结和分析了迁移学习技术在 COVID-19 诊断方面国内外的研究成果。邵良玉[14]通过迁移学习对 5 种花卉图像进行识别，大幅提高了识别准确率。马睿[15]对 4 种卷积神经网络进行迁移学习，对玉米籽粒图像进行分类识别，取得了较高的准确率，证明了应用卷及神经网络和迁移学习的方法的可行性。过程如图 2 所示，使用含有大量数据的数据集 A 对模型 A 进行预训练，得到模型参数较好的预训练模型 A，再将预训练模型 A 的参数迁移到模型 B 中，对数据量较小的数据集 B 进行训练和识别。

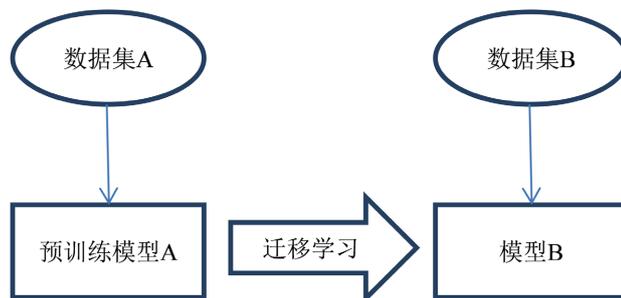


Figure 2. The process of transfer learning
图 2. 迁移学习过程

对于大部分图像来说，他们的图像特征会有部分相似的共性，不同类别的图像又有着各自的特性。在 VGG16 的卷积神经网络中，不同深度的卷积层承担着不同的特征提取任务，浅层的卷积层用于提取图像的共性特征，较高层的卷积层则用于提取图像的深层特征。使用大量数据训练出来的模型参数具有较好的性能，更利于提取图像的浅层特征，微调就是在参数迁移的基础上，“冻结”卷积神经网络中已经训练好的用于提取通用特征的较低层，“解冻”用于提取特殊特征的较高层，对其进行重新训练，以达到较好的识别效果，减少过拟合现象的发生。

在本实验中，首先要用含有大量数据的数据集对 VGG16 模型进行预训练，对训练好的预训练模型进行参数迁移，“冻结”前四组卷积层，“解冻”最后一组卷积层，使用收集到的蝴蝶数据集对最后一组卷积层和改写的全连接层进行训练，以提高识别准确率。

3. 实验

3.1. 实验环境

模型训练环境为 Windows 11 操作系统，Intel(R) Core(TM) i5-12500H 3.10 GHz 处理器，NVIDIA GeForce RTX 3050 的显卡，训练使用 Pytorch 1.2.0 框架，使用 Python3.7.0 作为编程语言。

3.2. 实验方案

3.2.1. 实验一

利用 Imagenet 上的大型数据集对 VGG16 模型进行预训练，将蝴蝶数据集按照训练集、验证集 8:2 的比例进行划分，对数据集进行归一化处理，使图像的大小统一为 $224 \times 224 \times 3$ ，作为 VGG16 模型的输入，保留预训练模型的全部卷积层和池化层，重写最后的 3 个全连接层。前两个全连接层以 ReLU 函数为激活函数，增加 Dropout 层，参数设为 0.5，随机去掉一部分神经元，以增加模型泛化能力，减小模型过拟合，由于蝴蝶数据集的蝴蝶种类数为 20，将最后一个全连接层的节点数改为 20，得到最终识别结果。实验过程如图 3 所示。

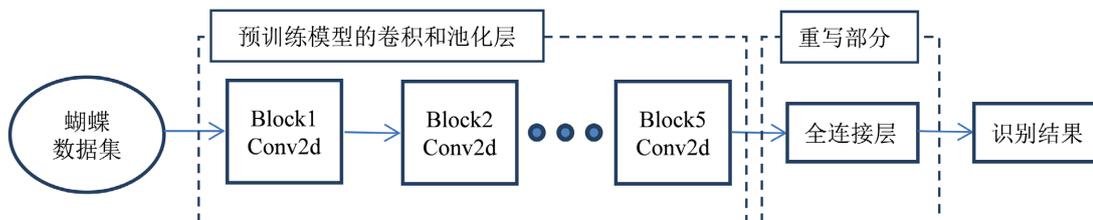


Figure 3. Process of experiment 1
图 3. 实验一流程

3.2.2. 实验二

仍使用 VGG16 的预训练模型，因为数据集的数据量较小，对蝴蝶数据集进行数据增强。将 `batch_size` 设为 64，每次随机选择 64 张图像作为批量输入，再通过使用 Pytorch 框架中的 `transforms.Compose()` 函数，将图像随机旋转 15° ，对一半的图像进行随机水平翻转和垂直翻转，随机调整图像的亮度和对比度等方法，使每批次作为输入的图像都不完全一样，以便于获取更多的图像特征。“冻结”卷积层和池化层，修改全连接层，进行训练，得到最终识别结果。实验过程如图 4。

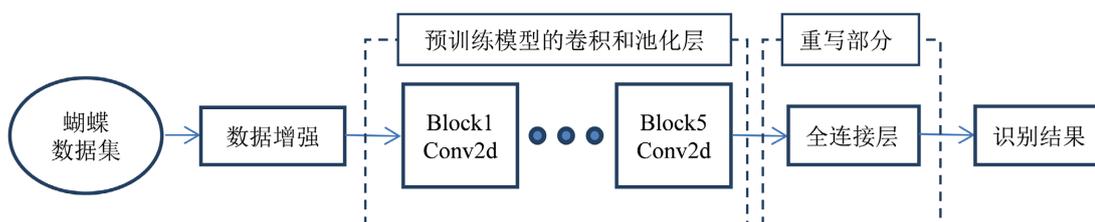


Figure 4. Process of experiment 2

图 4. 实验二流程

3.2.3. 实验三

在实验二的基础上，使用 VGG16 的预训练模型，对数据集进行数据增强，对模型进行微调，“解冻”最后一组卷积层，使“解冻”部分卷积层的参数可以更改，再对这部分卷积层和全连接层进行训练，得到最终识别结果。实验过程如图 5。

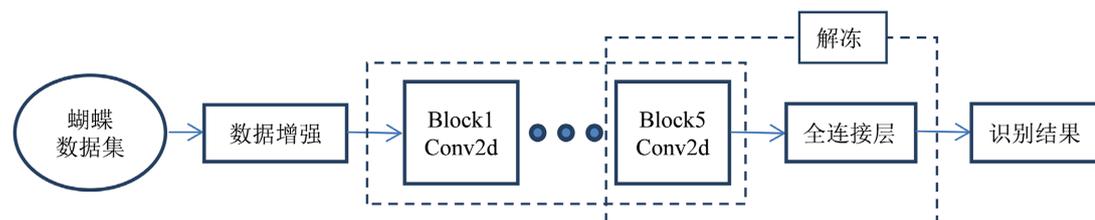


Figure 5. Process of experiment 3

图 5. 实验三流程

4. 实验结果与分析

4.1. 实验一结果与分析

加载在 ImageNet 上预训练好的 VGG16 模型，只加载特征提取部分，即保留卷积层和池化层的参数，不进行训练，不加载全连接层，对应本文数据集中的蝴蝶种类，重写全连接层，只对全连接层进行训练。

设初始学习率为 10^{-4} ，迭代次数为 30，输入批量为 64，结果如图 6 所示，只对图像进行归一化处理，经加载的卷积层和池化层进行特征提取，训练重写的全连接层参数，随迭代次数增加，准确率快速提升，可以看到在迭代次数为 5 时，训集准确率就已经接近 100%，过拟合现象十分严重。验证集准确率只有 76.67%，准确率较低，损失值较大。

4.2. 实验二结果与分析

为提高验证集准确率，减缓过拟合现象，仍使用预训练的 VGG16 模型，通过随机旋转、翻转、调节

亮度获对比度的方式，对蝴蝶数据集进行数据增强，以便于模型能够提取到更多的图像特征，仍“冻结”模型的卷积层，只训练全连接层，结果如图 7 所示。可以看到过拟合现象得到一定的缓解，且验证集准确率有所提升，达到 80.74%，损失值也有所下降。

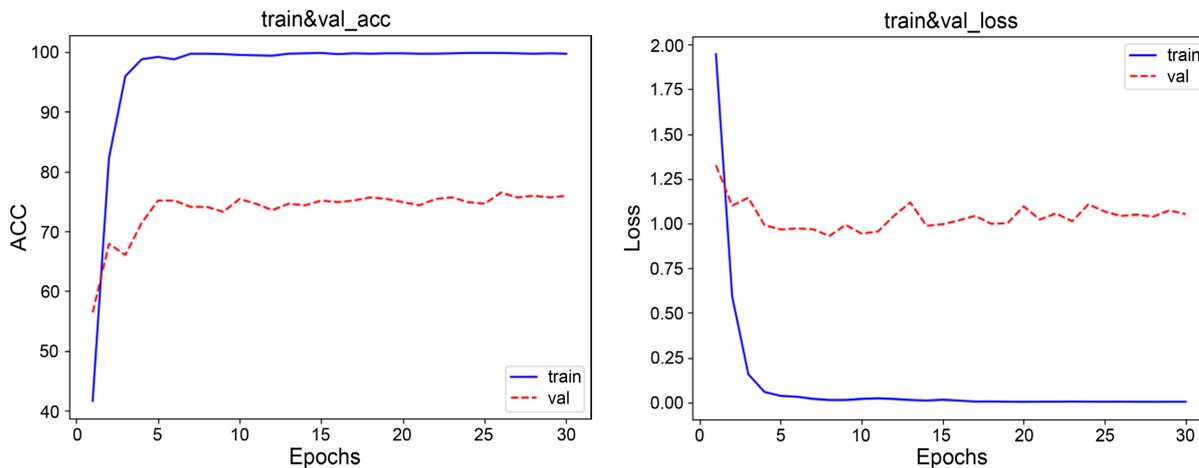


Figure 6. Accuracy and loss of training set and verification set in experiment 1

图 6. 实验一训练集和验证集的准确率和损失

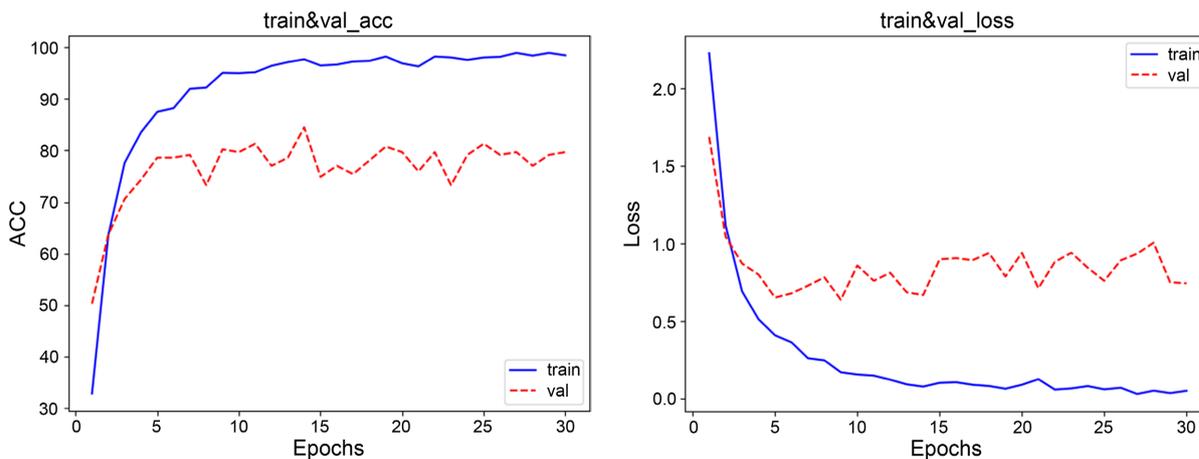


Figure 7. Accuracy and loss of training set and verification set in experiment 2

图 7. 实验二训练集和验证集的准确率和损失

4.3. 实验三结果与分析

为进一步提升模型效果，提高验证集准确率，对模型进行微调，“解冻”模型特征提取部分最后一个模块(block5)的卷积层，允许其参数和全连接层参数同时进行训练，前四模块的参数仍旧保持“冻结”状态，迭代次数设为 40，实验结果如图 8 所示。由图可知，验证集的准确率得到进一步提高，达到 83.95%，损失值进一步下降，过拟合现象得到进一步缓解，模型性能有所提升。

如表 2 所示，实验证明，通过加载在大型数据集上预训练的模型，通过旋转、翻转、亮度、对比度的调节，对数据集进行数据增强，对模型进行参数迁移，“解冻”部分卷积层重新训练，进行模型微调，使模型准确率由最初的 76.67%提升到了 83.95%。降低了训练时间，缓解了过拟合现象，识别准确率得到提升，有效提升了模型性能。

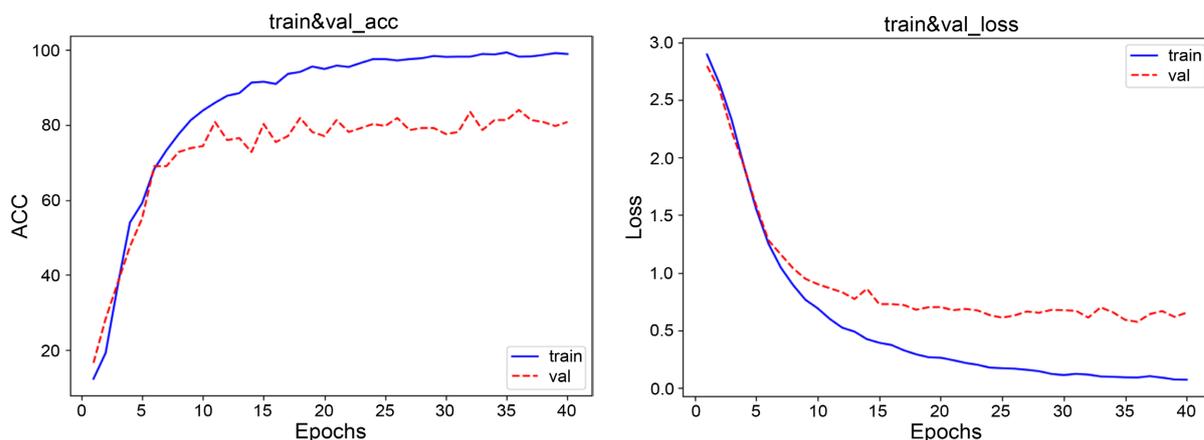


Figure 8. Accuracy and loss of training set and verification set in experiment 3

图 8. 实验三训练集和验证集的准确率和损失

Table 2. Accuracy of training set verification set of three experiments

表 2. 三个实验的训练集验证集准确率

| 模型 | 训练集准确率 | 验证集准确率 |
|-------------------|-----------|--------|
| VGG16 | 99.76176% | 76.67% |
| VGG16 + 数据增强 | 97.85587% | 80.74% |
| VGG16 + 数据增强 + 微调 | 98.80881% | 83.95% |

5. 结论

为解决野外蝴蝶数据集数据量较小，识别准确度低的问题，本文提出了一种通过数据增强和迁移学习及微调的方法，来提升小数据集蝴蝶识别准确度的方法。通过旋转、翻转等数据增强，使得每批次训练的图像都有所不同，这样模型能够学习到更多的图像特征。通过对经大数据集训练过的预训练模型的参数迁移，使模型具有更好的初始性能，经过对最后一组卷积层的“解冻”，对模型进行微调，是模型具有更好的收敛效果，提升了训练速率。实验表明，通过大数据集对 VGG16 模型进行预训练，再迁移到数据量较小的蝴蝶数据集识别中，通过数据增强和迁移微调，可以有效提高识别的准确率。经过三个实验的对比，可以看到未过迁移学习的模型准确率只有 76.67%，而经过数据增强和迁移微调模型的准确率逐步提升，最终达到 83.95%，过拟合现象也得到缓解，证明了通过数据增强和迁移学习及微调的方法，提升小数据集蝴蝶种类识别的方法的可行性。

致谢

本篇论文历时数月终于完成，我要感谢我的导师、朋友和同学为我提供的帮助。

这篇论文的完成离不开单菁老师和王佳英老师的帮助和指导，在我遇到困难时为我指明方向，给予我支持和鼓励，导师对我的每一次指导都让我受益匪浅，为我写论文提供了巨大的帮助，在此对单菁老师和王佳英老师表示深深的感谢。

还要感谢我的朋友们在写论文的过程中对我的陪伴和支持，这是支撑我完成这篇论文的动力，感谢同组的同学为我提出了宝贵的修改意见，使这篇论文更加完善。

参考文献

- [1] Espeland, M., Breinholt, J., Willmott, K.R., *et al.* (2018) A Comprehensive and Dated Phylogenomic Analysis of But-

- terflies. *Current Biology*, **28**, 770-778. <https://doi.org/10.1016/j.cub.2018.01.061>
- [2] 张建伟. 基于计算机视觉技术的蝴蝶自动识别研究[D]: [博士学位论文]. 北京: 中国农业大学, 2006.
- [3] 刘芳, 沈佐锐, 张建伟, 杨红珍. 基于颜色特征的昆虫自动鉴定方法[J]. 昆虫知识, 2008, 45(1): 150-153.
- [4] Kang, S.H., Cho, J.H. and Lee, S.H. (2014) Identification of Butterfly Based on Their Shapes When Viewed from Different Angles Using an Artificial Neural Network. *Journal of Asia-Pacific Entomology*, **17**, 143-149. <https://doi.org/10.1016/j.aspen.2013.12.004>
- [5] Kaya, Y., Kayci, L. and Tekin, R. (2013) A Computer Vision System for the Automatic Identification of Butterfly Species via Gabor- Filter-Based Texture Features and Extreme Learning Machine: GF+ELM. *Tem Journal*, **2**, 13-20.
- [6] Kaya, Y. and Kayci, L. (2014) Application of Artificial Neural Network for Automatic Detection of Butterfly Species Using Color and Texture Features. *Visual Computer*, **30**, 71-79. <https://doi.org/10.1007/s00371-013-0782-8>
- [7] Kaya, Y., Kayci, L. and Uyar, M. (2015) Automatic Identification of Butterfly Species Based on Local Binary Patterns and Artificial Neural Network. *Applied Soft Computing*, **28**, 132-137. <https://doi.org/10.1016/j.asoc.2014.11.046>
- [8] 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.
- [9] 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.
- [10] 周文进, 李凡, 薛峰. 基于YOLOv3和注意力机制的野外蝴蝶种类识别[J]. 郑州大学学报(工学版), 2022, 43(1): 34-40. <https://doi.org/10.13705/j.issn.1671-6833.2022.01.007>
- [11] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. <https://arxiv.org/abs/1409.1556>
- [12] Oquab, M., Bottou, L., Laptev, I. and Sivic, J. (2014) Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks. *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 23-28 June 2014. <https://doi.org/10.1109/CVPR.2014.222>
- [13] 孟伟, 袁艺琳. 迁移学习应用于新型冠状病毒肺炎诊断综述[J]. 计算机科学与探索, 2023, 17(3): 561-576.
- [14] 邵良玉. 基于迁移学习的花类图像分类方法研究[J]. 农业装备与车辆工程, 2022, 60(7): 62-99.
- [15] 马睿, 王佳, 赵威, 郭宏杰, 马德新, 兰进好. 基于卷积神经网络与迁移学习的玉米籽粒图像分类识别[J/OL]. 中国粮油学报: 1-10. <https://kns.cnki.net/kcms/detail/11.2864.TS.20220803.1310.010.html>, 2023-04-21.