

基于改进权重衰减的 EfficientNet 食用菌图像识别

EfficientNet edible fungus image recognition based on improved weight decay

姚芷馨 张太红 赵昀杰

YAO Zhi-xin ZHANG Tai-hong ZHAO Yun-jie

(新疆农业大学计算机与信息工程学院, 新疆 乌鲁木齐 830052)

(College of Computer and Information Engineering, Xinjiang Agricultural
University, Urumqi, Xinjiang 830052, China)

摘要:目的:解决食用菌种类识别传统方法靠人眼判断的不足,降低同一科属性状相似的识别出错率。方法:提出一种基于卷积神经网络的 EfficientNet 食用菌图像分类模型,利用不同设备和拍摄环境采集食用菌图像并建立数据集,通过模型训练技巧和网络技巧对模型性能进行提升,提出一种 YWeight 权重衰减方法来控制有效学习率,通过控制 cross-boundary 来影响模型的泛化性能。结果:该方法在自建数据集 YMushroom 上使 EfficientNet-B0 获得 79.82%(+0.85%) top-1 精度,默认训练过程仅获得 78.97%。在公开数据集 Fungus 上使 EfficientNet-B0 达到 87.62%(+0.78%) 准确率,原始训练准确率为 86.84%。结论:通过调整超参数使模型可接近最优解,通过权重衰减提升了食用菌图像分类模型的性能。

关键词:食用菌;图像分类;YWeight 方法;有效学习率;权重衰减;超参数

Abstract: Objective: In order to solve the problem of edible fungus species identification, an EfficientNet edible fungus image classification model based on convolution neural network is proposed. **Methods:** Firstly, the edible fungus images were collected and the datasets were made according to different equipment and shooting environment, and then the model performance was improved through model training skills and network skills. A YWeight weight attenuation method was proposed to control the effective learning rate, and the generalization performance of the model was affected by controlling the cross-boundary. **Results:** This method makes EfficientNet-B0 obtain 79.82% (+0.85%)

top-1 accuracy on the self built dataset YMushroom, and only 78.97% in the default training process. On the public dataset fungus, the accuracy of EfficientNet-B0 was 87.62% (+0.78%) and the original training accuracy was 86.84%. **Conclusion:** Experiments show that by adjusting the super parameters, the model finds a near optimal solution, and improves the performance of the edible fungus image classification model through weight attenuation, which provides a basis for the automatic management of edible fungus planting base in the future.

Keywords: edible fungus; image classification; YWeight method; effective learning rate; weight decay; hyperparameter

中国是世界上最早进行食用菌栽植的国家之一,拥有丰富的真菌物种资源,目前现已查明的真菌种类达 1 500 种以上,已知的食用菌有 350 多种,其中多属担子菌亚门,已人工驯化栽培成功的有 60 多种^[1]。但目前食用菌种类识别通常靠人工判断,费时、费力、效率低,而且同一科属性状相似的识别出错率高。

近年来,计算机视觉技术在食用菌上的应用研究越来越多,薛雨^[2]利用机器视觉技术,针对食用菌的特征实现精准、稳定的信息采集,通过不同时段的检测数据,对食用菌生产环境进行了精确调节和控制,但只实现了对具有明显形状变化或颜色改变的食用菌进行检测,此方法更适用于实验室环境,对于自然环境中拍摄的食用菌有较大误差;为快速有效地识别野生食用菌,林楠等^[3]基于图像处理方法和机器视觉技术,提出了一种改进的颜色空间融合形态特征提取方法。

研究拟提出一种基于卷积神经网络的 Y-Weight 训练方法,通过控制有效学习率和权重范数来影响模型的泛化性能,以期为常规大型蔬菜市场和大型超市识别菌

基金项目:新疆维吾尔自治区重大科技专项(编号:2017A01002)

作者简介:姚芷馨,女,新疆农业大学在读硕士研究生。

通信作者:张太红(1965—),男,新疆农业大学教授,博士。

E-mail: zth@xjau.edu.cn

收稿日期:2021-11-27 **改回日期:**2022-05-25

种提供帮助,减少人工识别的时间成本,提高工作效率,降低劳动强度。

1 材料与方 法

1.1 数据来源及范围

通过在公共数据集 Fungus^[4]上以及自建数据集 YMushroom 上进行训练。其中,由于 Fungus 数据集类别多样,背景较为复杂,有各种各样的背景造成干扰,因此选用不同科属的 12 类具有明显特征且图片数量较多较均匀的食用菌进行测试,共计 1 093 张图片。其分布情况及部分示例如图 1 所示。

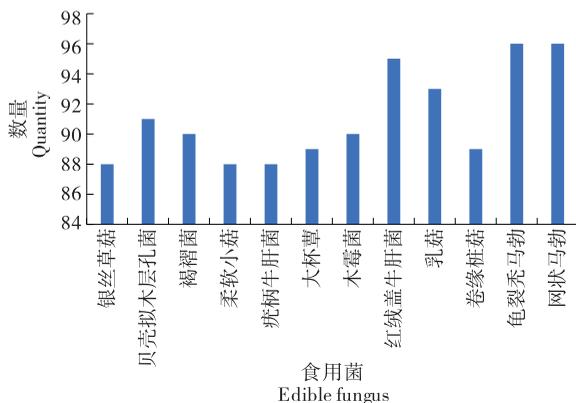


图 1 Fungus(12)数据集分布

Figure 1 Fungus (12) dataset distribution

自建数据集 YMushroom 是通过在常规蔬菜市场及大中型超市中调研,根据 2020—2026 年中国食用菌行业市场行情监测及发展前景展望报告^[5]中提供的市场销售额、整体价格和机器手持采摘的成本等数据,挑选出最具经济价值的 28 种常见食用菌类(香菇、双孢蘑菇、糙皮侧耳、金针菇、黑木耳、银耳、猴头菇、毛头鬼伞、姬松茸、茶薪菇、真姬菇、灰树花、滑菇、刺芹侧耳、盖囊侧耳、毛木耳、竹荪、金顶侧耳、大杯蕈、北冬虫夏草、牛肝菌、松茸、鸡枞、羊肚菌、榛蘑、鸡油菌、红菇、青冈菌)。

由于菌类生长时间与生长环境不同,需在不同时间段进行采集。采集时间从 2019 年 10 月—2021 年 3 月。采集设备分别采用两种不同图像处理方式的智能手机 OPPO Find X2 和 HUAWEI 荣耀 Play4T,使用原相机设置为方形拍摄。参考 Fungus 数据集制定严格的采集标准(尺寸、图像背景、干湿比例、光照强度、不同成熟度等)。其中新鲜食用菌与干货食用菌保持 1:1 的采集比例,每个种类的数据集数量基本保持均衡,并且保证了个体种类具有多样性。因为采集的样本越多,模型的训练精度就越高,所以 28 种常见食用菌平均每种采集了大约 1 784 张图片,共计 49 958 张图片。为提高模型运行速度,将采集后的数据集经过图像处理统一压缩为 600 像

素×600 像素大小,以此建立一个符合食用菌分类的大型数据集。将数据集按照常规训练进行划分,按照 8:1:1 的比例将其中 80% 的样本(约 39 966 幅图像)用作训练集,将 10% 的样本(约 4 996 幅图像)用作验证集,将剩余 10% 的样本(约 4 996 幅图像)用作测试集。之后为投入训练,需要将样本图像文件和对应的标签文件转换成 HDF5 格式文件(train.hdf5、val.hdf5、test.hdf5),保存在固定文件夹中,以便随时调用。自建数据集 YMushroom 分布情况及部分示例如图 2 所示。

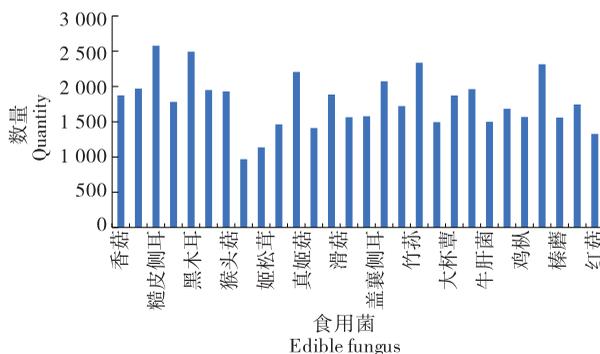


图 2 YMushroom 数据集分布

Figure 2 YMushroom dataset distribution

1.2 研究方法

1.2.1 有效学习率 用 ω 和 $\hat{\omega} = \omega / \|\omega\|_2$ 来表示通道的权重向量, x 表示通道输入大小,则使用欧几里得范数进行批归一化如式(1)所示。

$$BN(\|\omega\|_2 \hat{\omega} x) = BN(\hat{\omega} x). \quad (1)$$

当线性层的输出通道权重向量范数保持不变时,梯度按 $\frac{1}{\|\omega\|_2}$ 缩放,则:

$$\frac{\partial BN(\|\omega\|_2 \hat{\omega} x)}{\partial (\|\omega\|_2 \hat{\omega})} = \frac{1}{\|\omega\|_2} \frac{\partial BN(\hat{\omega} x)}{\partial \hat{\omega}}. \quad (2)$$

当模型缩放尺度保持不变时,权重向量的关键特征就只需要考虑其方向性。当权重随着 step 为 t 、learning rate 为 η 的随机梯度下降进行更新时,下一次 step 的通道权重向量即为:

$$\omega_{t+1} = \omega_t - \eta \nabla L_t(\omega_t). \quad (3)$$

权重向量在方向上的 step 大小按比例变化,如式(4)所示。

$$\hat{\omega}_{t+1} - \hat{\omega}_t \propto \frac{\eta}{\|\omega_t\|_2}. \quad (4)$$

因此,当对所有层使用权重衰减然后进行归一化时,可以防止权重范数的无限增长,从而维持权重方向的步长保持不变,以此提高有效学习率。

1.2.2 权重范数 根据式(1)~式(4)可以看出,卷积层的权重衰减主要是通过约束权重向量范数来完成的。为了研究在训练期间权重向量范数的变化,不考虑卷积层

的权重衰减,在 ImageNet^[6]上使用 EfficientNet-B0 进行训练(100 轮),在卷积层中将权重范数固定为一个常数 $\lambda=0.0001$,而在权重衰减中并不使用定权范数,两者的最佳学习率选择也有所不同,通过使用网格搜索(grid search)^[7]来确定学习率保证最佳性能。如图 3 所示,两者达到了同样的 top-1 精度,这些结果表明批归一化后的卷积层可以忽略权重衰减的影响。

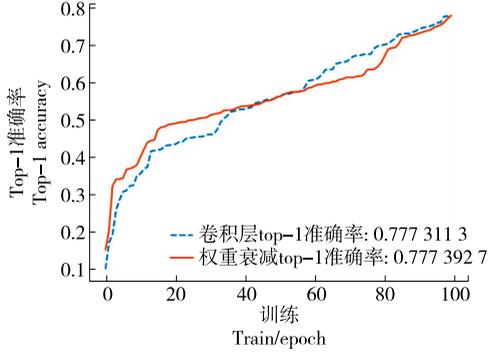


图 3 Conv layers 和 Weight decay 通过网格搜索后的 top-1 精度

Figure 3 Top-1 accuracy of Conv layers and Weight decay after grid search

调整卷积层的权重范数是通过 YWeight 方法定义初始化速率为 0,用 V_0 表示。总的训练步数用 t 表示,初始值为 0。用 x 表示训练样本,用 y 表示对应的标签,动量 momentum 用 μ 表示, τ 表示交叉熵损失函数。则训练速率计算公式如式(5)所示。

$$V_{t+1} = \mu V_t + \nabla \left\{ \sum \tau[f(x; \mathbf{W}_t), y] + \frac{1}{2} \lambda \|\mathbf{W}_t\|_2^2 \right\}. \quad (5)$$

定义一个随机的初始化权重向量 \mathbf{W}_0 , 初始学习率用 c 表示,则有:

$$\mathbf{W}_{t+1} = \mathbf{W}_t - l_r \times \eta_t \times V_{t+1}. \quad (6)$$

下一次的卷积层权重调整公式为:

$$\mathbf{W}_{t+1}^{\text{Conv}} = \frac{\mathbf{W}_{t+1}^{\text{Conv}}}{\|\mathbf{W}_{t+1}^{\text{Conv}}\|_2} \|\mathbf{W}_0^{\text{Conv}}\|_2. \quad (7)$$

1.2.3 全连接层 为研究权重衰减在全连接层中的作用,需要对上述卷积层权重衰减方法进行改进,首先保持其具有尺度不变性,使用普通的应用权重归一化的全连接层来代替原先的全连接层,其次设置 λ 原先的 0.0001 变为 0,最后使用 $\mathbf{W}^{\text{Conv+FC}}$ 来代替原先的 \mathbf{W}^{Conv} ,如式(8)所示。

$$\mathbf{W}_{t+1}^{\text{Conv+FC}} = \frac{\mathbf{W}_{t+1}^{\text{Conv+FC}}}{\|\mathbf{W}_{t+1}^{\text{Conv+FC}}\|_2} \|\mathbf{W}_0^{\text{Conv+FC}}\|_2. \quad (8)$$

原先的全连接层与替换后维持尺度不变性的全连接层进行对比,结构如式(9)、式(10)所示。

$$FC(x; \mathbf{W}^{\text{FC}}) = x^T \mathbf{W}^{\text{FC}}, \quad (9)$$

$$N_{\text{FC}}(x; \mathbf{W}^{\text{FC}}, g) = \frac{x^T \mathbf{W}^{\text{FC}}}{\|\mathbf{W}^{\text{FC}}\|_2} \times g. \quad (10)$$

最后使用 softmax 交叉熵损失函数^[8-9]与式(10)结合,用 τ 表示。用 $s_i = \frac{x^T \mathbf{W}_i}{\|\mathbf{W}\|_2} g$ 表示第 i 类的 log 值,用 p_i 表示第 i 类的概率, k 表示 label 类, j 表示其他类, g 为 N_{FC} 的权值。交叉熵损失函数计算公式如式(11)、式(12)所示。

$$\tau(x, k) = -\log p_k = -\log \frac{e^{s_k}}{\sum_j e^{s_j}}, \quad (11)$$

$$-\frac{\partial \tau(x, k)}{\partial x} = \frac{g}{\|\mathbf{W}\|_2} \sum_{j \neq k} p_j (\mathbf{W}_k - \mathbf{W}_j). \quad (12)$$

在特征空间中,权重向量 \mathbf{W} 的梯度为 x , \mathbf{W}_j 表示其他类别向量, \mathbf{W}_k 表示标签类向量,方向从 \mathbf{W}_j 到 \mathbf{W}_k ,其偏角大小由 p_j 和 g 决定,并且 p_j 也通过 softmax 函数依赖于 g 。当 x 是正确的分类时, g 会持续增长, p_j 会迅速减小并且梯度减弱。 x 更容易偏向于 \mathbf{W}_j 和 \mathbf{W}_k 之间的 class boundary,这种不明确的特征空间容易在训练和测试之间发生偏移,导致泛化性不好。特征空间中训练和测试的偏移如图 4 所示。

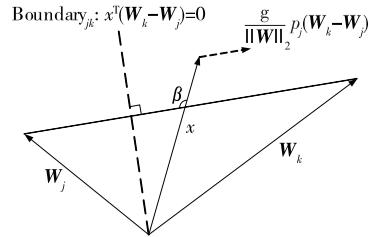


图 4 特征空间中训练和测试的偏移

Figure 4 Offset of training and testing in feature space

基于以上分析,给定一个上限值 α 来限制 g 的大小,用 Y_{FC} 表示,使用 $\sqrt{\#class}$ 来规范化不同数量类别的上限,如式(13)所示。

$$Y_{\text{FC}}(x; \mathbf{W}^{\text{FC}}, g, \alpha) = \frac{x^T \mathbf{W}^{\text{FC}}}{\|\mathbf{W}^{\text{FC}}\|_2} \times \min(g, \alpha * \sqrt{\#class}). \quad (13)$$

使用限制了权值大小的 Y_{FC} 层代替 N_{FC} 层,不改变其他超参数,通过权值选择 α 的大小。总体来说,权重衰减通过约束全连接层的权重范数来影响 cross-boundary risk,最终影响模型的泛化性能。通过使用 Y_{FC} 层可以恢复常规权重衰减的训练的准确率。

1.2.4 训练时 learning rate 和 α 的调整 在 1.2.2 和 1.2.3 中介绍了两种权重衰减的机制:① 讨论有效学习率对除卷积层以外的归一化之后的层的影响;② 对于全连接层 FC,控制 cross-boundary 的影响。 Y_{FC} 将这两种机制结合起来,通过控制超参数 l_r 和 α 直接控制其效果,使用 grid search 网格搜索来确定不同 l_r 和 α 对 top-1 的影响,通过

穷举^[10]的方式针对每种可能的参数组合情况进行训练,并对模型进行评价,根据其评价结果将寻找到的 top-1 最大值时的参数取值情况进行保存,视为最优。

1.3 系统框架

1.3.1 MobileNetV2 MobileNet^[11] 网络结构是一种小巧而高效的 CNN 模型,不仅在准确率和运行速率之间做了折中,还采用了深度可分离卷积模块^[12-13],并设置了两个超参数(宽度乘法和分辨率乘法),通过调整两个超参数来满足应用需求。在 ImageNet 数据集上,MobileNet 与 VGG16 的准确率几乎一样,但参数量只有其 1/32,计算量只有其 1/27。MobileNetV2 引入了 Bottleneck 结构,将其设计成纺锤型,先放大到原来的 6 倍,最后再缩小,形成倒残差模块(Inverted Residual Block)。网络主体结构如表 1 所示。

1.3.2 EfficientNet-B0 EfficientNet^[14-15] 的主干网络继续使用 MobileNetV2 网络中的 MBConv 模块,使用 SENet 中的压缩激励模块(squeeze and excitation)优化网络结构。其中运用到模型尺度化方法,在原始模型基础上将通道数扩大,加深模型深度^[16-17],之后扩大输入图像分辨率,再放大基于通道数、深度、输入图像的组合尺度。EfficientNet 模型尺度化方法如图 5 所示。

表 1 MobileNet 主体结构

Table 1 Main structure of MobileNet

Type/Stride	Filter Shape	Input Size
Conv/s2	3×3×3×32	224×224×3
Conv dw/s1	3×3×32 dw	112×112×32
Conv/s1	1×1×1×64	112×112×32
Conv dw/s2	3×3×64 dw	112×112×64
Conv/s1	1×1×164×128	56×56×64
Conv dw/s1	3×3×128 dw	56×56×128
Conv/s1	1×1×128×128	56×56×128
Conv dw/s2	3×3×128 dw	56×56×128
Conv/s1	1×1×128×256	56×56×128
Conv dw/s1	3×3×256 dw	28×28×256
Conv/s1	1×1×256×512	28×28×256
Conv dw/s2	3×3×256 dw	28×28×256
Conv/s1	1×1×256×512	14×14×256
5×Conv dw/s1	3×3×512 dw	14×14×512
5×Conv/s1	1×1×512×512	14×14×512
Conv dw/s2	3×3×512 dw	14×14×512
Conv/s1	1×1×512×1 024	7×7×512
Conv dw/s2	3×3×1 024 dw	7×7×1 024
Conv/s1	1×1×1 024×1 024	7×7×1 024
Avg Pool/s1	Pool 7×7	7×7×1 024
FC/s1	1 024×1 000	1×1×1 024
Softmax/s1	Classifier	1×1×1 000

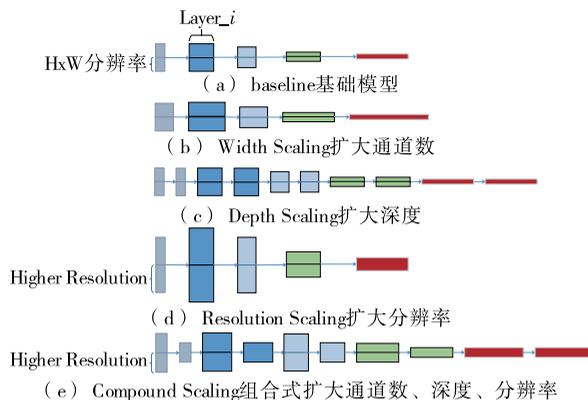


图 5 EfficientNet 模型尺度化方法

Figure 5 Scaling method of EfficientNet model

模型尺度化方法首先执行网格搜索,然后找到在固定资源约束下基线网络(baseline)的不同缩放维度之间的关系^[18]。其次找到每个维度适当的比例系数,最后应用这些系数扩大基线网络,使模型达到预期的大小或资源要求。与传统缩放方法相比,这种模型尺度化方法能够提高模型的精度和运行效率。

2 结果与分析

2.1 试验环境

硬件环境: NVIDIA Tesla P100-PCIE; 显存 16 GB; 显存位宽 4 096 bit; CUDA 版本 11.0。

软件环境: Windows10 操作系统; Python 版本为 3.6.13; Tensorflow 版本为 2.0; Keras 版本为 2.3.1。

2.2 训练方法

通过使用 YWeight 方法,在公共数据集 Fungus 和自建分类数据集 YMushroom 上进行试验。训练时使用 Stochastic Gradient Descent(SGD)随机梯度下降算法^[19],保证不会陷入 original-loss 的奇点和 minibatch-loss 的奇点,decay 设置为 1E-006, momentum 为 0.9。使用一个周期的余弦退火策略(Cosine Annealing),使学习率按照周期变化。由于学习率是神经网络训练中最重要的超参数之一,刚开始训练模型初始化权重是随机的,若选择了一个较大的学习率可能会导致模型震荡,因此前 4 轮选用线性预热(linear warmup)来进行学习率的优化,在预热的小学习率下使模型慢慢趋于稳定。除此之外,还使用了标签平滑(label smoothing)为 0.1 的正则化策略,通过 soft one-hot 来加入噪声,防止过拟合现象,减少真实样本标签的类别在计算损失函数时的权重。

2.3 性能评估指标

评价模型的性能通常使用准确率(Accuracy, A)、精度(Precision, P)、召回率(Recall, R)和 F 值(F-Score, F)等指标来进行衡量,通常被用在二分类测试中。

准确率被用来统计模型正确识别的量度,是正确预

测(真阳性和真阴性)的个数占样本总数的比例,也被称作“兰德精度”,其计算方法如式(14)所示。

$$A = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \times 100\%, \quad (14)$$

式中:

T_P ——将正类预测为正类的样本数,即真阳性;

T_N ——将负类预测为负类的样本数,即真阴性;

F_P ——将负类预测为正类的样本数,即假阳性;

F_N ——将正类预测为负类的样本数,即假阴性。

精度用来计算与真实值的接近程度,即对正类预测的准确性。其计算方法如式(15)所示。

$$P = \frac{T_P}{T_P + F_P} \times 100\%. \quad (15)$$

召回率也被称为灵敏度,为真阳性的数量占总体正类的比例。其计算方法如式(16)所示。

$$R = \frac{T_P}{T_P + F_N} \times 100\%. \quad (16)$$

F 值是综合精度和召回率的一个判断指标,是精度和召回率的调和平均值。其计算方法如公式(17)所示。

$$F = \frac{2PR}{P+R}. \quad (17)$$

由于试验研究的是多分类问题,因此将各个类别的评价指标进行综合计算,求其算术平均值。则宏精度(\bar{P})、宏召回率(\bar{R})、宏 F 值(\bar{F})计算公式如式(18)~式(20)所示。

$$\bar{P} = \frac{1}{n} \sum_{i=1}^n P_i, \quad (18)$$

$$\bar{R} = \frac{1}{n} \sum_{i=1}^n R_i, \quad (19)$$

$$\bar{F} = \frac{2\bar{P}\bar{R}}{\bar{P} + \bar{R}}. \quad (20)$$

2.4 YWeight 方法与贝叶斯优化对比分析

为证明 YWeight 方法的有效性,使用 YWeight 权重衰减和贝叶斯优化(Bayesian Optimization, BO)两种方法,在 CIFAR-10 数据集上使用 ResNet50^[20]网络和 MobileNetV2 网络进行对比试验,结果见表 2。将 ResNet50 训练 120 轮,MobileNet 训练 150 轮,原始模型训练的 top-1 精度为 92.3%和 91.41%,使用贝叶斯优化进行调参后训练的 top-1 精度为 92.37%和 91.53%,使用 YWeight 方法训练后的 top-1 精度为 92.41%和 91.58%。

2.5 数据集试验结果分析

使用 MobileNetV2 和 EfficientNet-B0 两种模型结构进行对比试验,分别在公共数据集 Fungus 和自建数据集 YMushroom 上进行训练。选取这两种网络结构进行研究,是因为在 ImageNet 上,MobileNet 在参数量是 VGG16 的 1/32,计算量是 VGG16 的 1/27 的情况下,与 VGG16 几乎达到了相同的精度。而选取 EfficientNet 网络是因为其延续了 MobileNetV2 中的 MBConv 模块作为模型的主干网络,同时还使用 SENet 中的 squeeze and excitation 方法对网络结构进行了优化,在 8 个广泛使用的数据集上都达到过比较先进的精度。

通过在 Fungus 数据集上对比试验,发现将 MobileNet 与 EfficientNet 两种网络结构通过 YWeight 方法控制学习率和权重范数后,对准确率有一定的提升。使用 YWeight 方法在 MobileNetV2 网络上进行训练的模型命名为 YWeight-MobV2,在 EfficientNet-B0 网络上进行训练的模型命名为 YWeight-EffB0。

通过表 3 可以看出,当 l_r 为 0.2, α 为 4.0E-005 时, MobileNetV2 的准确率为 82.71%,宏 F 值为 81.47%;当使用 YWeight 方法控制 l_r 为 0.4, α 为 8.0 时, YWeight-

表 2 YWeight 与 BO 试验结果

Table 2 YWeight and BO experimental results

网络模型	原始数据			BO			YWeight		
	l_r	λ	top-1/%	l_r	λ	top-1/%	l_r	λ	top-1/%
ResNet50	0.4	1E-004	92.30	0.54	8.5E-005	92.37	1.5	0.6	92.41
MobileNetV2	0.2	4E-005	91.41	0.67	2.4E-005	91.53	0.4	8.0	91.58

表 3 Fungus 数据集上识别结果

Table 3 Identification results on Fungus dataset

模型	轮数	参数	计算力	l_r	α	准确率/%	宏 F 值/%
MobileNetV2	150	3.5M	300M	0.20	4.0E-005	82.71	81.47
YWeight-MobV2	150	3.5M	300M	0.40	8.0	83.59	82.34
EfficientNet-B0	150	5.3M	384M	0.35	4.0	86.84	85.61
YWeight-EffB0	150	5.3M	384M	0.50	8.0	87.62	86.29

MobV2 的准确率为 83.59%，宏 F 值为 82.34%。当 l_r 为 0.35， α 为 4.0 时，EfficientNet-B0 的准确率为 86.84%，宏 F 值为 85.61%；当使用 YWeight 方法控制 l_r 为 0.5， α 为 8.0 时，YWeight-EffB0 的准确率为 87.62%，宏 F 值为 86.29%。由此看出，对于样本数量较少的 Fungus 数据集，在训练次数相同的情况下，通过 YWeight 方法控制不同的学习率以及权重范数，使 MobileNetV2 相比于原有模型识别准确率提高了 0.88%，宏 F 值提高了 0.87%，使 EfficientNet-B0 识别准确率提高了 0.78%，宏 F 值提高了 0.68%。可以看出，MobileNetV2 网络在使用 YWeight 方法后，展现了更大的进步性，top-1 准确率和宏 F 值比 EfficientNet-B0 得到了更大的提升。EfficientNet-B0 网络相比于 MobileNetV2 网络来说其参数较多，模型层数更深，运行速度相对较慢，但 top-1 精度和宏 F 值相对较高。为更清晰直观地了解训练中模型拟合的速度和收敛速度，截取在验证集上损失函数值的变化图像进行对比分析，如图 6 所示。

在 YMushroom 数据集上进行对比试验，使用 YWeight 方法在 MobileNetV2 网络上进行训练的模型命名为 YWeight-MobV2，在 EfficientNet-B0 网络上进行训练的模型命名为 YWeight-EffB0。

通过表 4 可以看出，当 l_r 为 0.3， α 为 4.0 时，MobileNetV2 的准确率为 75.48%，宏 F 值为 74.24%；当使用 YWeight 方法控制 l_r 为 0.4， α 为 8.0 时，YWeight-MobV2

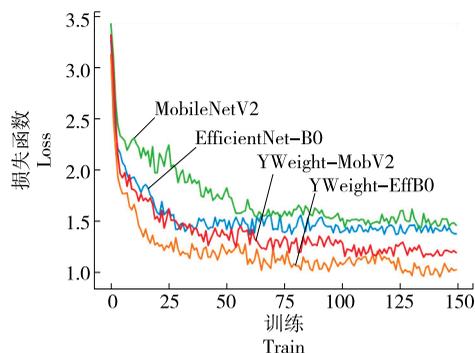


图 6 Fungus 验证集上各模型损失函数的变化曲线
Figure 6 Variation curve of loss function of each model on Fungus verification set

的准确率为 76.35%，宏 F 值为 74.61%。当 l_r 为 0.3， α 为 8.0 时，EfficientNet-B0 的准确率为 78.97%，宏 F 值为 77.74%；当使用 YWeight 方法控制 l_r 为 0.4， α 为 16.0 时，YWeight-EffB0 的准确率为 79.82%，宏 F 值为 78.49%。由此看出，对于比较大型的 YMushroom 数据集，相比于原有模型，在训练次数相同的情况下，通过 YWeight 方法控制不同的学习率以及权重范数，使 MobileNetV2 识别准确率提高了 0.87%，宏 F 值提高了 0.37%，使 EfficientNet-B0 识别准确率提高了 0.85%，宏 F 值提高了 0.75%。可以看出，使用 YWeight 方法后 EfficientNet-B0 在 top-1 准确率和宏 F 值的提升上更具进步性，MobileNetV2 的宏 F 值仅提升了 0.37%，而 EfficientNet-B0 的宏 F 值提

表 4 YMushroom 数据集上识别结果

Table 4 Identification results on YMushroom dataset

模型	轮数	参数	计算力	l_r	α	准确率/%	宏 F 值/%
MobileNetV2	250	3.5M	300M	0.3	4.0	75.48	74.24
YWeight-MobV2	250	3.5M	300M	0.4	8.0	76.35	74.61
EfficientNet-B0	250	5.3M	384M	0.3	8.0	78.97	77.74
YWeight-EffB0	250	5.3M	384M	0.4	16.0	79.82	78.49

升了 0.75%。因此，基于试验要求更高的精度，更适合选取 EfficientNet 网络作为食用菌分类模型的主干网络。4 个模型在 YMushroom 验证集上的损失函数变化图像如图 7 所示。

对比表 3 和表 4、图 6 和图 7 发现，对于小型数据集 Fungus，仅需 150 轮就能够收敛，训练速度较快，但是对于大型数据集 YMushroom 则需要 250 轮才逐渐收敛，达到极值。证明不同的模型拥有不同的最优 l_r ，训练模型在相同的设置下也会有不同的训练结果，需要通过 YWeight 方法找到最优的 l_r 和 α 来提高模型性能。

3 讨论

(1) 目前中国较先进的菌菇研究方法大多是使用传

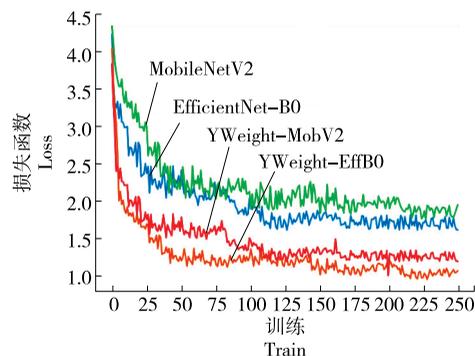


图 7 YMushroom 验证集上各模型损失函数的变化曲线
Figure 7 Variation curve of loss function of each model on YMushroom verification set

统的机器学习方法或者是层数较少的浅层神经网络来进行图像识别,例如 VGG16、AlexNet 等。文中使用的 MobileNet 以及 EfficientNet 均为较深层的卷积神经网络,在模型准确率与运行速率上都有很大的提升。

(2) 目前针对食用菌图像分类的研究分类种数不超过 20 种,且大部分都是基于公共数据集来进行网络训练的,而试验对 28 种最常见的食用菌种类进行了研究,通过 1 年的时间,进行了不同环境、不同时间、不同拍摄设备的食用菌图像采集,构建了 HDF5 格式的大型食用菌数据集,共 49 958 张图片,能够使模型更好的训练。

(3) 大部分食用菌识别用到的网络都是通过迁移学习来完成的,通过修改模型参数来达到更高的准确率,而试验创新使用了一种权重衰减方法,在两种网络结构上试验过后发现均适用,并且比原始的模型提高了准确率和运行效率。

4 结论

采用基于卷积神经网络的 EfficientNet 和 MobileNet 两种网络结构,作为 YWeight 训练方法的食用菌图像分类模型,在两个数据集上进行试验,对模型识别性能进行对比分析,结果表明:

(1) 基于权重衰减策略,提出的训练方法 YWeight 方法,通过一种高效的、鲁棒性的方法选择超参数,通过约束全连接层的权重范数来影响 cross-boundary,最终影响模型的泛化性能。试验证明通过调整超参数使模型在试验中找到了接近最优的解,在公开数据集 Fungus 和自建数据集 YMushroom 中都展现了非常高效的模型准确率,为食用菌自动识别提供了便利。

(2) 研究中发现权重衰减对最后的全连接层具有影响,发现了一种新的机制来补充权重衰减对模型泛化性能的影响,YWeight 方法是简单高效的,并且在大规模数据集上证明了其性能。

(3) 从检测速度上看,基于 YWeight 方法的食用菌图像分类模型能够满足常规蔬菜市场以及大型超市的实时识别要求。

(4) 由于数据集非常庞大,像素比较高,EfficientNet 模型实时识别计算量较大,对设备要求比较高,今后可以尝试其他更轻便高效的网络进行训练,进一步提高食用菌识别性能。

(5) 试验训练时使用的是 SGD 优化器,可以在每次迭代中使用一个样本来更新参数,能够提高在自建的大型数据集 YMushroom 上的训练速度。但是有可能达到局部最优,在目标函数为强凸函数时,SGD 可能无法做到线性收敛,准确率也会下降。因此,后续将继续讨论权重衰减在 Momentum 优化器、Adam 优化器上应用的优点

以及可能存在的问题,并研究学习率、权重衰减和动量之间的联系。

参考文献

- [1] 鲍大鹏. 食用菌科学研究为食用菌产业发展提供越来越重要的科学支撑[J]. 菌物学报, 2021, 40(2): 3 061-3 063.
BAO D P. Scientific research on edible fungi provides more and more important scientific support for the development of edible fungi industry[J]. Mycosystema, 2021, 40(2): 3 061-3 063.
- [2] 薛雨. 基于机器视觉技术的食用菌生长过程监控系统研究[D]. 郑州: 华北水利水电大学, 2018: 55.
XUE Y. Research on monitoring system of edible fungi growth process based on machine vision technology[D]. Zhengzhou: North China University of Water Resources and Electric Power, 2018: 55.
- [3] 林楠, 王娜, 李卓识, 等. 基于机器视觉的野生食用菌特征提取识别研究[J]. 中国农机化学报, 2020, 41(5): 111-119.
LIN N, WANG N, LI Z S, et al. Research on feature extraction and recognition of wild edible fungi based on machine vision[J]. Journal of Chinese Agricultural Mechanization, 2020, 41(5): 111-119.
- [4] FILIPPOVA N, BULYONKOVA T, BANAEV E V, et al. Fungal records database of khanty-mansi autonomous okrug-yugra[J]. BIO Web of Conferences, 2018, 11: 15-20.
- [5] 智研咨询集团. 2020—2026 年中国食用菌行业市场行情监测及发展前景展望报告[EB/OL]. (2020-02-14) [2021-10-17]. <https://www.chyxx.com/industry/202002/833926.html>.
Zhiyan Consulting Group. Report on market monitoring and development prospect of China's edible fungi industry from 2020 to 2026 [EB/OL]. (2020-02-14) [2021-10-17]. <https://www.chyxx.com/industry/202002/833926.html>.
- [6] MARMANIS D, DATCU M, ESCH T, et al. Deep learning earth observation classification using ImageNet pretrained networks [J]. IEEE Geoscience & Remote Sensing Letters, 2016, 13(1): 105-109.
- [7] SON H, FONG Y. Fast grid search and bootstrap-based inference for continuous two-phase polynomial regression models [J]. Environmental, 2021, 32(3): e2664.1-e2664.16.
- [8] 吴慧华, 苏寒松, 刘高华, 等. 基于余弦距离损失函数的人脸表情识别算法[J]. 激光与光电子学进展, 2019, 56(24): 196-202.
WU H H, SU H S, LIU G H, et al. Facial expression recognition algorithm based on cosine distance loss function[J]. Laser & Optoelectronics Progress, 2019, 56(24): 196-202.
- [9] 章东平, 陈思瑶, 李建超, 等. 基于改进型加性余弦间隔损失函数的深度学习人脸识别[J]. 传感技术学报, 2019, 32(12): 1 830-1 835.
ZHANG D P, CHEN S Y, LI J C, et al. Deep learning face recognition based on improved additive cosine interval loss function[J]. Chinese Journal of Sensors and Actuators, 2019, 32(12): 1 830-1 835.
- [10] NGO N, PORTER G A, CHERNG H Q, et al. Development of a color object classification and measurement system using machine

- vision[J]. *Sensors and Materials*, 2019, 31(12): 4 135-4 154.
- [11] LI W, LIU K. Confidence-aware object detection based on Mobile-Netv2 for autonomous driving[J]. *Sensors*, 2021, 21(7): 2 380.
- [12] MALEKABADI A J, KHOJASTEHPOUR M, EMADI B, et al. Development of a machine vision system for determination of mechanical properties of onions[J]. *Computers & Electronics in Agriculture*, 2017, 141: 131-139.
- [13] 郭丽丽, 丁世飞. 深度学习研究进展[J]. *计算机科学*, 2015, 42(5): 28-33.
- GUO L L, DING S F. Research progress of deep learning[J]. *Computer Science*, 2015, 42(5): 28-33.
- [14] ATILA M, UAR M, AKYOL K, et al. Plant leaf disease classification using EfficientNet deep learning model[J]. *Ecological Informatics*, 2021, 61: 101182.
- [15] DUONG L T, NGUYEN P T, DI SIPIO C, et al. Automated fruit recognition using EfficientNet and MixNet [J]. *Computers and Electronics in Agriculture*, 2020, 171: 105326.
- [16] 李文宝, 曹成茂, 张金炎, 等. 基于深度学习的山核桃破壳物料分类识别[J]. *食品与机械*, 2021, 37(9): 133-138, 185.
- LI W B, CAO C M, ZHANG J Y, et al. Classification and recognition of pecan shell breaking materials based on deep learning[J]. *Food & Machinery*, 2021, 37(9): 133-138, 185.
- [17] MA H, CHEN M, ZHANG J W. Study on the fruit grading recognition system based on machine vision[J]. *Advance Journal of Food Science and Technology*, 2015, 8(11): 777-780.
- [18] 赵世达, 王树才, 李振强, 等. 基于 U 型卷积神经网络的羊肋排图像分割[J]. *食品与机械*, 2020, 36(9): 116-121, 154.
- ZHAO S D, WANG S C, LI Z Q, et al. Segmentation of sheep rib image based on U-shaped convolution neural network[J]. *Food & Machinery*, 2020, 36(9): 116-121, 154.
- [19] 张璐, 李卓识, 李玉. 鹅膏属真菌形态特征的主成分与聚类分析[J]. *菌物学报*, 2018, 37(5): 559-564.
- ZHANG L, LI Z S, LI Y. Principal component and cluster analysis of morphological characteristics of *Amanita fungi* [J]. *Mycosystema*, 2018, 37(5): 559-564.
- [20] 王海燕, 张渺, 刘虎林, 等. 基于改进的 ResNet 网络的中餐图像识别方法[J]. *陕西科技大学学报*, 2022, 40(1): 154-160.
- WANG H Y, ZHANG M, LIU H L, et al. Chinese food image recognition method based on improved ResNet[J]. *Journal of Shaanxi University of Science & Technology*, 2022, 40(1): 154-160.

(上接第 81 页)

- [18] 李俊, 王震, 郭晓关, 等. 基于全二维气相飞行时间质谱联用方法分析贵州酱香型白酒挥发性风味成分[J]. *酿酒科技*, 2016(12): 102-106.
- LI J, WANG Z, GUO X G, et al. Analysis of volatile flavoring components in Guizhou Jiangxiang Baijiu by two comprehensive dimensional gas chromatography-time of flight mass spectrometry[J]. *Liquor-making Science & Technology*, 2016(12): 102-106.
- [19] 廖永红, 赵爽, 张毅斌, 等. LLE、SDE、SPME 和 GC-MS 结合保留指数法分析二锅头酒中的风味物质[J]. *中国食品学报*, 2014, 14(6): 220-228.
- LIAO Y H, ZHAO S, ZHANG Y B, et al. Analysis of flavor substances in erguotou wine by LLE, SDE, SPME and GC-MS combined with kovats retention indices[J]. *Journal of Chinese Institute of Food Science and Technology*, 2014, 14(6): 220-228.
- [20] LIU H L, SUN B G. Effect of fermentation processing on the flavor of Baijiu[J]. *Journal of Agricultural and Food Chemistry*, 2018, 66(22): 5 425-5 432.
- [21] 唐平, 山其木格, 王丽, 等. 白酒风味化学研究方法及其酱香型白酒风味化学研究进展[J]. *食品科学*, 2020, 41(17): 315-324.
- TANG P, SHAN Q M G, WANG L, et al. A review of research methods in baijiu flavor chemistry and recent progress in the flavor chemistry of maotai-flavored Baijiu[J]. *Food Science*, 2020, 41(17): 315-324.
- [22] 胡瑾, 马一飞, 罗佳雪, 等. 气相色谱法同时测定白酒中 57 种风味物质[J]. *中国酿造*, 2022, 41(5): 206-211.
- HU J, MA Y F, LUO J X, et al. Simultaneous determination of 57 flavor substances in Baijiu by GC[J]. *China Brewing*, 2022, 41(5): 206-211.

(上接第 85 页)

- [4] 潘丽霞, 徐信贵. 论食品安全监管中的政府信息公开[J]. *中国行政管理*, 2013(4): 29-31, 14.
- PAN L X, XU X G. On government information disclosure in food safety supervision[J]. *China Administration*, 2013(4): 29-31, 14.
- [5] 罗季阳, 张晓娟, 李经津, 等. 形成出口花生黄曲霉毒素风险的外源性因素分析[J]. *食品工业科技*, 2011, 32(10): 406-408.
- LUO J Y, ZHANG X J, LI J J, et al. Analysis of exogenous factors that form the risk of peanut aflatoxin export[J]. *Food Industry Science and Technology*, 2011, 32(10): 406-408.
- [6] 闫海, 潘俊雅. 我国食品安全风险评估的改革重点与法制建议[J]. *政法学刊*, 2019, 36(1): 78-83.
- YAN H, PAN J Y. Reform focus and legal suggestions on food safety risk assessment in China[J]. *Journal of Political Science and Law*, 2019, 36(1): 78-83.
- [7] 罗季阳, 张晓娟, 王欣, 等. 突发食品安全风险的早期识别[J]. *食品工业科技*, 2012, 33(20): 53-55, 59.
- LUO J Y, ZHANG X J, WANG X, et al. Early identification of unexpected food safety risks[J]. *Science and Technology of Food Industry*, 2012, 33(20): 53-55, 59.
- [8] 丰苏, 杜琳, 袁刚, 等. 运用智慧监管理念构建统一的食品安全监管平台[J]. *中国市场监管研究*, 2021(11): 30-34, 51.
- FENG S, DU L, YUAN G, et al. Building a unified food safety supervision platform with the concept of smart supervision[J]. *China Market Supervision Research*, 2021(11): 30-34, 51.