

# 基于轻量级卷积神经网络的植物叶片病害识别方法

贾鹤鸣<sup>1,2</sup>, 郎春博<sup>3\*</sup>, 姜子超<sup>4</sup>

(1. 三明学院 信息工程学院, 福建 三明 365004;

2. 福建省农业物联网应用重点实验室(三明学院), 福建 三明 365004;

3. 西北工业大学 自动化学院, 西安 710129;

4. 东北林业大学 机电工程学院, 哈尔滨 150040)

(\* 通信作者电子邮箱 lang\_chunbo@163.com)

**摘要:**针对目前农业信息领域植物病害识别精度较低、实时性较差的问题,提出了一种基于轻量级卷积神经网络(CNN)的植物叶片病害识别方法。在原有网络中引入深度可分离卷积(DSC)和全局平均池化(GAP)方法,分别用来代替标准卷积运算操作并对网络末端的全连接层部分进行替换。同时,批归一化的技巧也被运用到训练网络的过程中,以改善中间层数据分布并提高收敛速度。为全面而可靠地评估所提方法的性能,在公开的植物叶片病害图像数据集PlantVillage上进行实验,选取损失函数收敛曲线、测试精度、参数内存需求等指标来验证改进策略的有效性。实验结果表明,改进后的网络具有较高的病害识别精度(99.427%)以及较小的内存空间占用(6.47 MB),可见其与其他基于神经网络的叶片识别技术相比具有优势,工程实用性较强。

**关键词:**卷积神经网络;植物叶片病害;图像分类;深度可分离卷积;全局平均池化

**中图分类号:**TP391.4 **文献标志码:**A

## Plant leaf disease recognition method based on lightweight convolutional neural network

JIA Heming<sup>1,2</sup>, LANG Chunbo<sup>3\*</sup>, JIANG Zichao<sup>4</sup>

(1. School of Information Engineering, Sanming University, Sanming Fujian 365004, China;

2. Fujian Key Lab of Agriculture IOT Application (Sanming University), Sanming Fujian 365004, China;

3. School of Automation, Northwestern Polytechnical University, Xi'an Shaanxi 710129, China;

4. College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin Heilongjiang 150040, China)

**Abstract:** Aiming at the problems of low accuracy and poor real-time performance of plant leaf disease recognition in the field of agricultural information, a plant leaf disease recognition method based on lightweight Convolutional Neural Network (CNN) was proposed. The Depthwise Separable Convolution (DSC) and Global Average Pooling (GAP) methods were introduced in the original network to replace the standard convolution operation and the fully connected layer part at the end of the network respectively. At the same time, the technique of batch normalization was also applied to the process of training network to improve the intermediate layer data distribution and increase the convergence speed. In order to comprehensively and reliably evaluate the performance of the proposed method, experiments were conducted on the open plant leaf disease image dataset PlantVillage, and loss function convergence curve, test accuracy, parameter memory demand and other indicators were selected to verify the effectiveness of the improved strategy. Experimental results show that the improved network has higher disease recognition accuracy (99.427%) and smaller memory space occupation (6.47 MB), showing that it is superior to other leaf recognition technologies based on neural network, and has strong engineering practicability.

**Key words:** Convolutional Neural Network (CNN); plant leaf disease; image classification; Depthwise Separable Convolution (DSC); Global Average Pooling (GAP)

## 0 引言

植物病害的产生对农业生产有一定负面影响。如果不及时发现植物病害,就会增加粮食的风险<sup>[1]</sup>,特别是对一些主要的粮食作物,如玉米、水稻、小麦等,它们是满足人民生活需

要、推进生产力发展的关键。因此,探求一种智能化、低成本、高准确的方法来实施植物病害检测有着重要的现实意义。机器学习中的特征提取和模式识别有助于识别植物病害的类型和严重程度。通过植物叶片图像的颜色、形状和大小等特征

收稿日期:2020-09-21;修回日期:2020-11-28;录用日期:2020-11-30。

基金项目:中央高校基本科研业务费专项资金资助项目(2572019BF04)。

作者简介:贾鹤鸣(1983—),男,辽宁辽阳人,教授,博士,CCF会员,主要研究方向:数字图像处理、模式识别;郎春博(1998—),男,辽宁沈阳人,博士研究生,主要研究方向:图形图像处理、计算机视觉;姜子超(1995—),男,黑龙江齐齐哈尔人,硕士研究生,主要研究方向:机器学习、图像分割。

对植物健康状况进行自动质量分析,是提高生产力的一种准确可靠的方法<sup>[2-3]</sup>。

以往的许多研究工作都考虑了图像的认识问题,并采用一种特殊的分类器将图像分为健康图像和病变图像。一般来说,植物叶片是植物病害识别的第一手资料,因为大部分病害的症状最先出现在叶片上。在过去的几十年里,主要病害的识别和分类技术在植物中得到了广泛的应用,包括 $K$ 近邻( $K$ -Nearest Neighbor, KNN)<sup>[4]</sup>、支持向量机(Support Vector Machine, SVM)<sup>[5]</sup>、Fisher线性判别(Fisher Linear Discriminant, FLD)<sup>[6]</sup>、人工神经网络(Artificial Neural Network, ANN)<sup>[7]</sup>、随机森林(Random Forest, RF)<sup>[8]</sup>等。众所周知,经典方法的疾病识别率很大程度上取决于各种算法的病变分割和手工设计的特征,如七不变矩、尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)、Gabor变换、全局-局部奇异值和稀疏表示<sup>[9-11]</sup>等。然而,人工设计的特征需要昂贵的资源条件和专业的知识,同时具有一定的主观性。况且从被提取的目标上很难确定哪些特征是最优的、鲁棒的疾病识别特征。此外,在复杂的背景条件下大多数方法都不能有效地将叶子和相应的病变图像从背景中分割出来,导致实验得到的结果不能有效地用于预测疾病的发生。因此,由于病害叶片图像的复杂性,植物病害图像的自动识别仍然是一项具有挑战性的工作。

近年的深度学习技术,特别是卷积神经网络(Convolutional Neural Network, CNN),正在迅速成为克服上述挑战的首选方法。由于卷积神经网络的尺度不变性,使得它解决的图像问题不受尺度限制,进而在识别和分类方面表现出突出的能力。例如, Mohanty等<sup>[12]</sup>训练了一个深度学习模型来识别14种作物和26种作物病害。Ma等<sup>[13]</sup>利用深度CNN对黄瓜霜霉病、炭疽病、白粉病和目标叶斑病这4种病害的症状进行识别,识别准确率达到了93.4%。Kawasaki等<sup>[14]</sup>提出了一种基于CNN的黄瓜叶病识别方法,达到了94.9%的准确率。同样,本文也利用CNN对植物叶片病害特征进行提取,提出了一种基于VGG-16(Visual Geometry Group-16)的轻量级卷积网络。首先,在原有网络中引入深度可分离卷积(Depthwise Separable Convolution, DSC)<sup>[15]</sup>和全局平均池化(Global Average Pooling, GAP)<sup>[16]</sup>,代替标准卷积运算操作并对网络末端的全连接层部分进行替换。同时,批归一化的技术也被运用到训练网络的过程,以改善中间层数据分布并提高收敛速度<sup>[17]</sup>。改进后网络在植物叶片病害数据集PlantVillage上的实验结果表明,所提出的轻量级卷积网络在识别精度与效率方面有了明显的提高,适用于植物叶片病害识别的任务,具有较强的工程实用性以及较高的研究价值。

## 1 卷积神经网络

### 1.1 卷积神经网络基本理论

卷积神经网络是一种通过多层网络互联并相互传递数据信息的常见深度学习模型。每一层都具有独特的特征来处理输入数据并将其发送到下一层。卷积神经网络从输入待处理图像到输出分类结果的过程,包括:用于获取数据并进行预处理的输入层、提取图像高级特征的卷积层、降采样的池化层、

提高模型非线性的ReLU激活层、整合卷积层提取的高级图像特征的全连接层、在模型最后部分输出各类别预测概率的Softmax层。

#### 1.1.1 输入层

图像输入层是CNN体系结构中的重要组成部分,其将2-D和3-D图像作为主要处理对象。

$$V = \text{Input}(I, D) \quad (1)$$

其中: $\text{Input}$ 代表图像输入层,该层对给定图像 $I$ 进行预处理操作,进而得到满足要求的图像 $V$ 作为后续卷积层输入; $D$ 代表输入图像通道数。

#### 1.1.2 卷积层

该层按照一定的规则在输入图像 $V = m \times n$ 上移动卷积内核窗口 $k$ 。内核窗口 $k$ 每次滑动的距离由步长 $S_d$ 决定,具体计算过程如下:

$$F_m = \text{Conv}(V, k, S_d, R_l) \quad (2)$$

其中: $F_m$ 表示图像特征信息; $R_l$ 表示激活函数,主要用来改善模型的非线性。

#### 1.1.3 批归一化层

该层通过改善中间层的数据分布,进而影响模型优化器的梯度下降过程以实现提高模型收敛性能的目的。它采用了来自前一层卷积的特征图 $F_m$ 并将其激活归一化,该过程可以定义为:

$$F_{m+1} = N(F_m, ax, M) \quad (3)$$

其中: $N(\cdot)$ 表示归一化函数; $ax$ 表示特征图中统一进行均值与方差计算的坐标轴信息; $M$ 表示决定均值与方差之间变化的动量率。

#### 1.1.4 池化层

对从批归一化层获得的每个特征图 $F_{m+1}$ 执行最大池化操作,以减小图像大小。步幅 $S_d$ 的值通常由用户手动选择。池大小 $P_s$ 必须按以下方式赋给该层:

$$F_{m+1} = \text{Pool}(F_{m+1}, P_s) \quad (4)$$

#### 1.1.5 全连接层

全连接层主要作用是归整前面网络层获取的图像信息,将原本稀疏连接的方式转化为完全连接的形式,如式(5)所示:

$$Y_n = F_{\text{connect}} \left( \text{Conv}(V, K, S_d, R_l) + N(F_m, ax, M) + \text{Pool}(F_{m+1}, P_s) \right) \quad (5)$$

#### 1.1.6 Softmax层

该层主要设计用于对全连接层的输出进行分类和预测,并输出每个类别的模型预测结果,其计算式如下:

$$F(X_i) = \frac{e^{X_i}}{\sum_k e^{X_k}} \quad (6)$$

其中, $F(X_i)$ 可以计算全连接层 $Y_n$ 返回的每个类别的概率。

## 1.2 VGG网络

VGG网络(VGGNet)是由牛津大学视觉几何小组和谷歌的DeepMind的研究人员开发的一种深度CNN,致力于研究网络深度和CNN性能之间的联系<sup>[18]</sup>。通过反复堆叠小尺寸卷

积层和池化层,建立了一个网络深度为16~19的深度卷积神经网络模型,并探索了深度对网络性能的影响。与之前最先进的网络架构相比,VGGNet的错误率明显下降,在2014年ILSVRC(ImageNet Large Scale Visual Recognition Challenge)竞赛中获得第二名,在定位项目中获得第一名。在VGGNet的所有研究中都使用了 $3 \times 3$ 卷积内核和 $2 \times 2$ 池内核,通过不断深化网络结构来提高性能。

图1给出了具有5个卷积模块的VGG网络结构,每个卷积模块通常包含2~3次卷积运算。其中,Input、Conv、Pool、FC、Softmax分别代表输入层、卷积层、池化层、全连接层和分类层。

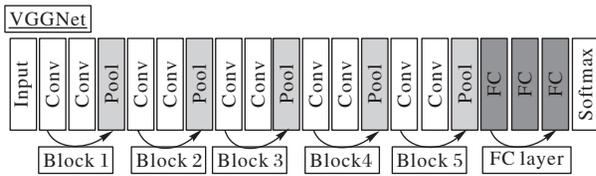


图1 VGG网络框架示意图

Fig. 1 Schematic diagram of VGG network framework

## 2 轻量级卷积神经网络

针对传统CNN参数较多、易过拟合的问题,分别采用深度可分离卷积和全局平均池化策略替换其标准卷积以及全连接层部分。此外,批归一化的方法也被分别用于调整中间层数据分布和防止模型过度拟合。

### 2.1 深度可分离卷积

深度可分离卷积完整过程的第一步是深度卷积部分,也叫通道卷积(图2(b))。给定与特征图中具有相同通道数的过滤器数量,并以通道为单位分别进行卷积运算,如式(7)。

$$DConv(\mathbf{W}, \mathbf{x})_{(i,j)} = \sum_{m,n} \mathbf{W}_{(m,n)} \cdot x_{(i+m,j+n)} \quad (7)$$

深度可分离卷积的第二步是将第一步输出的特征图与通道数和深度卷积相同且大小为 $1 \times 1$ 的卷积核相结合。逐点卷积核的个数在这部分中就代表着提取特征的个数,该操作组合了特征图各维度,因此减小了通道数量,其具体计算过程如式(8)所示:

$$PCConv(\mathbf{W}, \mathbf{x})_{(i,j)} = \sum_k \mathbf{W}_k \cdot x_{(i,j)} \quad (8)$$

其中: $\mathbf{W}$ 是卷积核的权重矩阵,并且是可训练的; $\mathbf{x}$ 表示输入到卷积层的特征图, $(i,j)$ 表示元素的坐标索引; $m$ 、 $n$ 和 $k$ 是卷积核的3个维度。

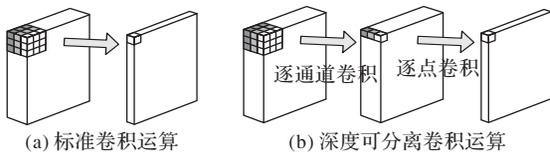


图2 深度可分离卷积与标准卷积对比

Fig. 2 Comparison between depthwise separable convolution and standard convolution

### 2.2 全局平均池化

为了解决传统网络易过拟合的问题,在卷积神经网络的

结构中引入了全局池化层来代替全连接层,将卷积层输出的多个特征图以自身为单位进行映射,映射结果为一个特征点。多个特征点便得到了与上述特征展开相似的一维向量,而后传入Softmax层中进行分类。

设 $x_{ij}^k$ 为卷积层中第 $k$ 个特征图的第 $(i,j)$ 个元素,最后一个卷积层的大小为 $m \times n$ ,GAP过程如下所示:

$$y_{GAP}^k = \frac{1}{mn} \times \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} x_{ij}^k \quad (9)$$

其中 $y_{GAP}^k$ 是GAP层的输出。

### 2.3 批归一化

批归一化层的主要目的是解决网络深入化导致的中间层数据分布不均影响模型学习效率的问题。具体来说,对于具有 $n$ 维输入的层 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 。

首先,对每个单元所表示的特征进行归一化,使其具有0均值和1标准差。

$$\hat{x}_i = \frac{x_i - E[x_i]}{\sqrt{Var[x_i]}} \quad (10)$$

其中: $E[x_i]$ 是每个单元的平均值; $\sqrt{Var[x_i]}$ 表示标准偏差。通过引入参数 $\gamma_i$ 和 $\beta_i$ 对激活值进行移动和缩放来解决原有表征改变的问题,变换式如下:

$$f_i = \gamma_i \hat{x}_i + \beta_i \quad (11)$$

该变换在 $\mathbf{W}x_n$ 的每个维度中独立使用,具有一对独立学习参数 $\gamma_n$ 和 $\beta_n$ , $x_n$ 和 $\mathbf{W}$ 分别表示输入数据集和权重矩阵。

### 2.4 算法架构与分析

图3给出了VGG16网络的架构示意图,可以看到随着特征提取的进行,特征图的尺寸由于池化层的作用而不断减小,维度由于卷积核个数的增多不断增加。算法逐步提取出更多且更高级的图像特征信息,以更好地胜任分类识别任务。此外,改进卷积神经网络的算法各层级信息如表1所示,其中:“3,3”代表卷积核大小;“DSConv”代表深度可分离卷积;“64/S1”代表卷积核个数与卷积步长。以第一层“3,3 DSConv,64/S1”为例,该层参数由逐通道卷积与逐点卷积两部分组成(如2.1节介绍),首先逐通道卷积部分参数为“卷积核大小 $\times$ 输入通道数”即“ $3^2 \times 3$ ”,其次逐点卷积部分为“点卷积核大小 $\times$ 输入通道数 $\times$ 卷积核个数”即“ $(1 \times 1 \times 3) \times 64$ ”。同理可得第二层参数,区别在于输入通道数由3变为64。从表1中可以明显看出,加入全局平均池化策略后节省了传统全连接层大量的参数,如果不采用GAP方法改进该部分,则全连接fc1层的参数为 $8 \times 8 \times 512 \times 4096 = 134\,217\,728$ 。

为解决传统网络进行植物叶片病害识别过程产生的过拟合问题,本文采用深度可分离卷积与全局平均池化策略分别代替标准卷积运算与全连接层进行网络优化。具体来说,深度可分离卷积通过将标准卷积运算拆分为逐通道卷积与逐点卷积以减少卷积层部分参数。如果输出通道的数量为 $o$ ,那么标准卷积层所需的总参数为 $m \times n \times k \times o$ 。深度可分离卷积运算的参数则由逐通道与逐点卷积两部分组成,为 $m \times n \times k + k \times o$ ,它们之间的比例是 $\frac{1}{o} + \frac{1}{m \times n}$ 。而在时间复杂度方

面,深度可分离卷积同样具有一定的优势,设  $M$  和  $K$  分别表示特征映射和卷积核的大小,  $C_{in}$  和  $C_{out}$  分别表示输入和输出通道的数目,则标准卷积的时间复杂度是  $O \sim (M^2 \times K^2 \times C_{in} \times C_{out})$ ,深度可分离卷积的复杂度可表示为  $O \sim (M^2 \times K^2 \times C_{in} + M^2 \times C_{in} \times C_{out})$ ,当输出通道数较大时,方法之间的复杂度差异更加明显。因此,本文采用深度可分离卷积来替代传统的标准卷积运算以减少卷积层部分复杂度,防止过拟合现象产生。

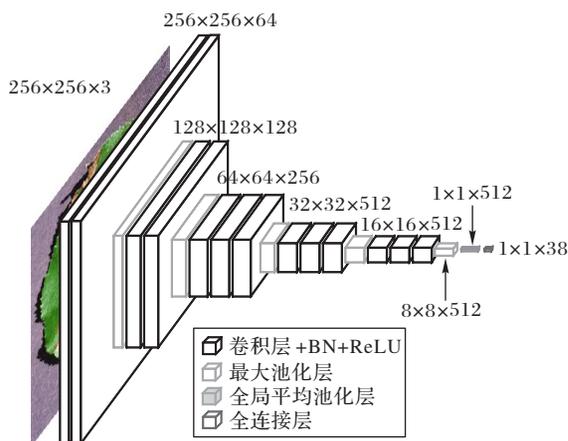


图 3 轻量级卷积神经网络架构

Fig. 3 Architecture of lightweight convolution neural network

表 1 改进网络的各层级信息

Tab. 1 Information of different levels of improved network

层/步长	输出形状	参数
3, 3 DSCConv, 64/S1		$(3^2 \times 3) + (1 \times 1 \times 3) \times 64 = 219$
3, 3 DSCConv, 64/S1	256, 256, 64	$(3^2 \times 64) + (1 \times 1 \times 64) \times 64 = 4\ 672$
2, 2 Maxpool/S2	128, 128, 64	0
3, 3 DSCConv, 128/S1		$(3^2 \times 64) + (1 \times 1 \times 64) \times 128 = 8\ 768$
3, 3 DSCConv, 128/S1	128, 128, 128	$(3^2 \times 128) + (1 \times 1 \times 128) \times 128 = 17\ 536$
2, 2 Maxpool/S2	64, 64, 128	0
3, 3 DSCConv, 256/S1		$(3^2 \times 128) + (1 \times 1 \times 128) \times 256 = 33\ 920$
3, 3 DSCConv, 256/S1	64, 64, 256	$(3^2 \times 256) + (1 \times 1 \times 256) \times 256 = 67\ 840$
3, 3 DSCConv, 256/S1		$(3^2 \times 256) + (1 \times 1 \times 256) \times 256 = 67\ 840$
2, 2 Maxpool/S2	32, 32, 256	0
3, 3 DSCConv, 512/S1		$(3^2 \times 256) + (1 \times 1 \times 256) \times 512 = 133\ 376$
3, 3 DSCConv, 512/S1	32, 32, 512	$(3^2 \times 512) + (1 \times 1 \times 512) \times 512 = 266\ 752$
3, 3 DSCConv, 512/S1		$(3^2 \times 512) + (1 \times 1 \times 512) \times 512 = 266\ 752$
2, 2 Maxpool/S2	16, 16, 512	0
3, 3 DSCConv, 512/S1		$(3^2 \times 512) + (1 \times 1 \times 512) \times 512 = 266\ 752$
3, 3 DSCConv, 512/S1	16, 16, 512	$(3^2 \times 512) + (1 \times 1 \times 512) \times 512 = 266\ 752$
3, 3 DSCConv, 512/S1		$(3^2 \times 512) + (1 \times 1 \times 512) \times 512 = 266\ 752$
2, 2 Maxpool/S2	8, 8, 512	0
GAP	1, 1, 512	0
Dense	1, 1, 38	$512 \times 38 = 19\ 456$
Softmax	1, 1, 38	0

而对于传统算法中参数量庞大的全连接层部分,本文采用全局平均池化策略对其进行优化。全局平均池化层对卷积层输出的特征向量逐通道求取特征均值,通道数对应分类器所需的类别数。其优点之一是增强了提取的特征映射和类别

之间的对应关系,另一个优点在于对特征图求均值的操作不需要额外的参数,与大量参数构成的全连接层相比能够避免训练集过度拟合的现象发生。此外,全局收敛对空间信息进行了总结,因此所构造的特征向量对于输入图像的空间平移具有更强的鲁棒性。在植物叶片病害识别中,使用全局平均池化比全连接操作有更好的识别效果,因为 GAP 可以实现对于数据的降维和参数的压缩,有效地防止过拟合发生并增强模型泛化能力。由于在 GAP 中不需要对参数进行优化,因此可以有效地克服过拟合并减少参数训练的时间。此外,批归一化操作的引入有效地解决了网络中间层数据分布不均的问题,按特征维度求均值与标准差进而进行归一化的操作改善了特征分布情况,对收敛速度与精度均产生了较好的影响。

### 3 植物叶片病害识别实验

#### 3.1 数据集及预处理

本文将深度学习模型架构在植物叶片图像上进行了训练,然后对该模型未见过的图像进行疾病分类和识别。该研究使用了 PlantVillage<sup>[19]</sup> 的公共数据集。PlantVillage 拥有 54 306 张图像,其中包括 14 种农作物上的 26 种常见病害,共 38 类。选取原始数据集中部分图像进行展示,如图 4 所示,分别为:苹果疮痂病、黑腐病、锈病、健康叶;玉米灰斑病、锈病、枯叶病、健康叶;番茄靶斑病、菌斑病、黄曲叶病、健康叶。



图 4 植物叶片病害数据集部分图像

Fig. 4 Some images of plant leaf disease dataset

通过整理数据可以发现,植物叶片数据存在对比度较低、样本分布不平衡等问题,这会在一定程度上影响训练的速度以及收敛精度。因此,分别利用直方图均衡化以及数据扩充技术加以改善(如图 5 所示)。在本文研究中利用 Python 脚本语言编写程序,采用随机旋转角度、随机缩放原图、随机水平/垂直翻转操作实现样本数据的增强。在对少量样本数据类数据扩充的同时也对大量样本数据类进行删减操作,最终使得各类样本数据均为 1 500 张。

#### 3.2 实验环境及设置

实验在图形处理单元(Graphics Processing Unit, GPU)模式下进行,实验所用计算机的详细配置为:内存 16 GB, Nvidia GTX980Ti 显卡,使用的操作系统为 64 位 Ubuntu 16. 04。实验选用的脚本语言为 Python,并安装 OpenCV 图像处理库、Pytorch 深度学习框架。

实验以小批量的形式进行,每批同时用 32 张图片进行训练/测试。损失函数优化器采用随机梯度下降(Stochastic

Gradient Descent, SGD)算法,学习率采用 warmup 更新策略(如图 6 所示)。初始学习率为  $1E - 6$ ,最大学习率为  $1E - 2$ ,预热周期为 3 个 epoch,动量为 0.9,权重衰减系数为  $1E - 3$ 。此外,在优化器的 warmup 阶段结束后,学习速率的更新方式是,如果当前 epoch 的测试精度低于上一个 epoch 的测试精度,学习率衰减到原来的 1/2。遍历测试集中的所有图片一次称为一个 epoch,在训练阶段执行 15 个 epoch。

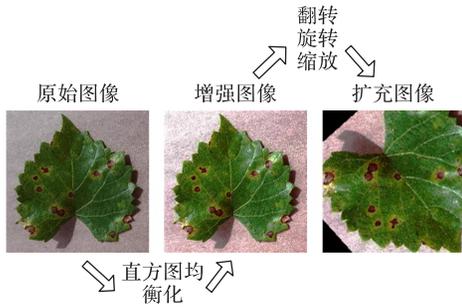


图 5 植物叶片病害图像预处理过程  
Fig. 5 Preprocessing process of plant leaf disease image

### 3.3 实验结果与讨论分析

本节中改进的卷积神经网络被用于解决植物叶片病害识别问题,实验可以被分为五部分,分别探究数据集扩充、改进

表 2 扩充数据集前后算法性能指标对比

Tab. 2 Comparison of algorithm performance indexes before and after extending dataset

算法	原始数据集				扩充后数据集			
	精度/%	损失	查准率/%	F1 评分/%	精度/%	损失	查准率/%	F1 评分/%
原始	97.281	0.0934	97.294	97.367	98.304	0.0524	98.306	98.301
改进	98.476	0.0441	98.271	98.459	99.392	0.0262	99.394	99.392

从表 2 中可以看出,使用数据增强操作处理后的数据集进行实验,传统网络和改进后网络在各指标上均有所提升,提升比例约为 1 个百分点。例如原始模型的测试精度从 97.281% 提升至 98.304%;平均查准率从 97.294% 提升至 98.306%;平均 F1 评分从 97.367% 提升至 98.301%。对于特征信息获取方面,通过随机缩放、旋转、翻转等操作对原始数据进行增强,得到的扩充数据集能够帮助模型学习到更加鲁棒且更具判别性的特征,这类特征的泛化性能更好,进而在训练集上进行测试同样得到了更高的精度。对于数据分布方面,扩充后的数据集中各类训练样本数量相同、分布均匀,这使得模型对于各类别的学习过程不会偏向某一类或某几类,得到的模型泛化能力较强,进而在指标结果中得以体现。

#### 3.3.2 改进策略对测试精度的影响

本节将从测试精度和训练时间两个方面分析改进策略对实验结果的影响。测试精度反映了模型解决当前植物病害分类问题的能力,训练时间反映了模型获得解决当前问题能力所花费的成本。

图 7 显示了传统模型和改进模型在扩展数据集上的精度和损失的比较结果。可以看出,改进模型的收敛精度优于原模型(迭代停止时更高的测试精度以及更低的测试损失),这是因为深度可分离卷积和全局平均池化策略的结合有效地减

少了模型的参数量,避免了模型对当前训练样本的识别能力较好而测试样本的识别较差的过拟合现象发生,提高了其泛化和适应新样本的能力,进而在测试集上获得了更好的收敛精度,而原始模型由于其庞大的参数量导致其在测试集上的泛化性能较差,并不能在学习过程中获得更具判别性的图像特征。从收敛速度方面来看,改进后算法在第 8 个 epoch 开始收敛,而原始算法在第 7 个 epoch 开始收敛,尽管收敛速度较慢于传统模型,但综合精度与速度可以发现,改进后的模型有效地避免了测试集过拟合的现象发生,能够在解空间内获得更为优异的参数。

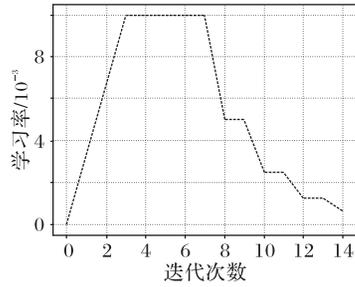


图 6 warmup 学习率调整方式示意图

Fig. 6 Schematic diagram of warmup learning rate adjustment mode

#### 3.3.1 数据集扩充对测试精度的影响

3.1 节介绍了对 PlantVillage 数据集的扩充操作,将样本分布不平衡的原始数据集调整为每类均有 1 500 张图像的数据集。表 2 展示了算法改进前后在原始数据集和扩充后数据集上的评估指标值对比情况。本节实验中选择的评价指标的计算方法是 38 类正确预测样本数与测试集大小的比值。平均查准率和平均 F1 评分为各类别查准率和 F1 评分的平均值。

表 3 给出了不同改进策略组合对模型测试结果的影响,包含骨干网络分别结合深度可分离卷积与全局平均池化后的测试精度与测试损失。从表 3 中可以看出,对标准卷积运算和全连接层的改进都带来了性能的提升,而后的效果更为明显。这是因为导致传统网络过拟合现象产生的关键在于参数量庞大的全连接层,对此部分进行优化能更好地解决该问题。但通过逐通道卷积与逐点卷积相结合的方式代替标准卷积运算的操作也能够一定程度上减少卷积层部分的参数量,并获得较高的测试精度,这也从侧面说明了本文方法将两者结合的可行性与有效性。观察实验结果可以发现,深度可分离卷积与全局平均池化策略的结合使模型在测试阶段获得

了模型的参数量,避免了模型对当前训练样本的识别能力较好而测试样本的识别较差的过拟合现象发生,提高了其泛化和适应新样本的能力,进而在测试集上获得了更好的收敛精度,而原始模型由于其庞大的参数量导致其在测试集上的泛化性能较差,并不能在学习过程中获得更具判别性的图像特征。从收敛速度方面来看,改进后算法在第 8 个 epoch 开始收敛,而原始算法在第 7 个 epoch 开始收敛,尽管收敛速度较慢于传统模型,但综合精度与速度可以发现,改进后的模型有效地避免了测试集过拟合的现象发生,能够在解空间内获得更为优异的参数。

了最高的精度与最低的损失,进一步验证了所提方法的可靠性与高效性。

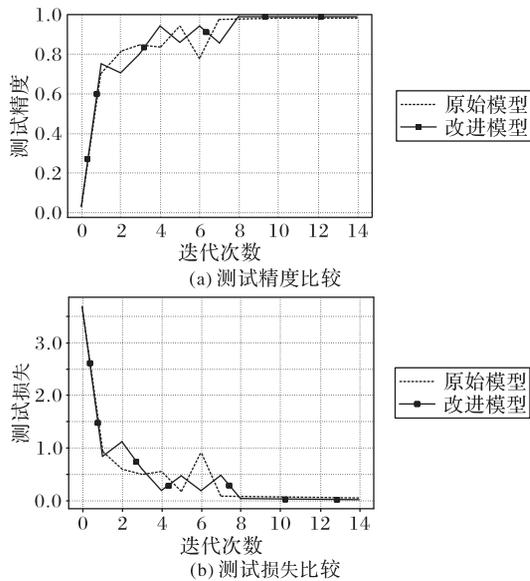


图 7 扩展数据集上传统模型和改进模型的性能比较  
Fig. 7 Performance comparison between traditional model and improved model on extended dataset

表 3 不同改进策略组合的测试结果

Tab. 3 Test results of different combinations of improvement strategies

模型	测试精度/%	测试损失
VGG	98.304	0.0524
VGG+DSC	98.664	0.0407
VGG+GAP	98.817	0.0378
本文模型	99.392	0.0262

表 4 给出了网络改进前后参数量及运行时间的比较,可以看出改进后的网络参数量远小于改进前参数量,这是由于全连接层占据了原始网络的绝大部分参数,而采用的全局平均池化策略替换了原有的全连接层,减少了网络的主要参数;此外,深度可分卷积部分将标准卷积分成逐通道卷积和逐点卷积两部分,进一步降低了原卷积层部分的参数,极大改善了复杂网络的易过拟合特性,在参数量较少的同时,提高了测试精度。从程序运行时间方面看,在每批训练 32 个样本的情况下,原始模型需要 586 s 的时间,而改进模型仅需要 233 s。程序总时间上的差别更加明显,训练 15 个 epoch 花费的时间相差 109.615 min,体现了改进后模型在训练成本方面的优越性。综合来看,改进后的模型参数量显著降低,其成效不但体现在防止过拟合进而提高测试精度方面,同样体现在训练所需的时间成本方面,具有很强的工程实用性。

表 4 算法改进前后参数量及运行时间比较

Tab. 4 Comparison of parameter number and running time before and after improvement

模型	模型大小/MB	参数量	每批次运行时间/s	总时间/min
原始模型	632.79	165 881 958	586	169.866
改进模型	6.47	1 695 873	233	60.251

### 3.3.3 初始化类型与激活函数对测试精度的影响

在本节中,针对不同初始化方法与激活函数类型的模型进行性能测试,进一步优化模型。网络参数的初始化方式对于模型的学习过程以及最终的测试精度均有一定的影响:较好的初始化值能够加快优化算法对其数值调整的速度,进而提高收敛速度;而较差的初始化值不但会减缓优化的进程,甚至会导致算法陷入局部最优。具体来说,本文将初始化类型分为高斯初始化和 Xavier 初始化,激活函数类型分为 ReLU 和 PReLU。

实验结果如表 5 所示,可以发现使用 Xavier 初始化方法与 PReLU 激活函数均能带来性能的提升,而且这两种改进方式的结合使得测试精度进一步提高。

表 5 初始化类型与激活函数对测试结果的影响

Tab. 5 Influence of initialization type and activation function on test results

模型	测试精度/%	测试损失
高斯+ReLU	99.392	0.0262
高斯+PReLU	99.399	0.0236
Xavier+ReLU	99.404	0.0227
Xavier+PReLU	99.427	0.0214

具体来说,Xavier 初始化方法的目标是使得每一层输出的方差应尽量相等,这与 2.3 节中批归一化方法的作用相类似,即改善中间层的分布情况。因此,引入该初始化方法后的模型获得了更为优异的初始分布,优化了模型的学习过程。而 PReLU 函数通过引入参数改善了传统 ReLU 激活函数负半轴激活值为 0 的死区,使网络所学习的特征更加具有鲁棒性与判别性,进一步提高其在测试集上的精度。从结果上看,仅使用 Xavier 初始化可以提升精度 0.012 个百分点,仅使用 PReLU 激活函数可以提升 0.007 个百分点,而结合使用的方式可以提高 0.035 个百分点。实验结果表明,改善初始化方法以及激活函数类型均能够提高模型对于测试样本的识别性能,而且本文所提模型的性能仍具有一定的提升空间,结合近年来神经网络改进策略,有一定的潜力达到更高的识别准确率,更具应用前景。

### 3.3.4 全局池化类型对测试精度的影响

本节探讨改进后网络全局池化部分的不同类型对算法性能所产生的影响。具体来说,改进策略中的全局池化方法分为全局最大池化和全局平均池化,区别在于从给定的特征映射到输出节点的转换过程是计算所有元素的最大值还是平均值,分别对应于显著表达和整体表达。

表 6 给出了全局平均池化以及全局最大池化的测试结果,可以看出,相较于最大池化,全局平均池化在当前网络结构上表现出更好的适应性,得到了更高的测试精度以及更小的测试损失,表明对于当前数据集,提取其整体特征的效果相较于显著特征的效果更好,即获取更具判别性的整体信息意义较大,对样本局部特征的综合与总结有助于模型判别目标,但这不足以说明全局最大池化在解决图像分类问题中的意义和价值,毕竟没有方法可以普遍适用全部模型以及全部问题,因此需探究不同改进策略对网络的影响。

表 6 全局池化类型对测试结果的影响

Tab. 6 Influence of global pooling type on test results

全局池化类型	测试精度/%	测试损失
平均池化	99.392	0.026 2
最大池化	98.971	0.031 6

## 3.3.5 与其他植物病害识别技术的对比

为客观而全面地评估所提网络模型的性能,在本节引入其他学者解决 PlantVillage 数据集图像识别问题的方法进行比较,实验结果如表 7 所示。由表 7 中可以发现,本文方法获得了最高的测试精度以及 F1 评分,提升效果为 1~2 个百分点,且相较于文献[20]方法节省了 88.5% 的参数,兼顾了识别准确率与资源利用率。相较于文献[20]中采用群智能优化算法对网络进行优化的策略相比,本文对于优化器方法部分没有进行改动,仅选择了标准的 SGD 优化算法,但全局平均池化以及深度可分离卷积等策略的引入极大地减少了网络参数,获得了相较于改进优化器方法更好的泛化能力与测试精度。与文献[21]方法相比,本文方法对于过拟合问题的处理采用了 GAP 与 DSC 策略相结合的方法,比单独改进卷积方式效果更为明显,取得了参数量与测试精度间的平衡。

此外,为更好地分析本文网络获得图像特征表达的过程,分别将各模块中卷积层的输出特征图可视化,如图 8 所示。由图 8 可以看出,随着网络层数的提高,所提取的图像特征越来越抽象,特征的纹理性逐渐被更高级的语义性所取代;而本文提出的网络从过程上看获得了较为丰富的边缘信息,从结果上看又取得了较高的识别精度,适用于植物叶片病害的识别。

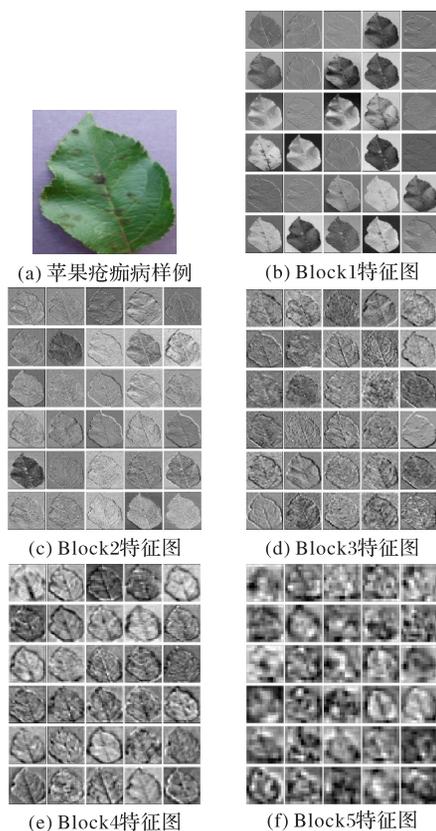


图 8 可视化卷积特征图

Fig. 8 Visual convolution feature maps

表 7 不同方法在 PlantVillage 数据集上的性能对比

Tab. 7 Performance comparison of different methods on PlantVillage dataset

方法	测试精度/%	F1 评分/%	参数量/10 <sup>6</sup>
本文方法	99.427	99.216	1.69
文献[20]方法	98.200	97.000	14.70
文献[21]方法	98.340	—	0.54

## 4 结语

本文提出了一种轻量级网络模型来解决生产实际中植物叶片病害识别困难、智能化较差的问题。该网络将深度可分离卷积与全局平均池化相结合,有效地提高了测试精度,减少了程序运行时间。总结来说,改进后的卷积神经网络不但具有可观的测试精度,并且运行速度、内存占用量方面也优势明显,较好地解决了植物病害图像识别问题,为后续相关研究提供了可靠的技术支持。在未来将会侧重研究调整网络结构对测试结果的影响。此外,其他植物叶片病害数据集的扩展研究也将被考虑在内,测试并改进现有算法,提高模型对于这类问题的泛化能力。

## 参考文献 (References)

- [1] 夏永泉,王兵,支俊,等. 基于随机森林方法的小麦叶片病害识别研究[J]. 图学学报, 2018, 39(1): 57-62. (XIA Y Q, WANG B, ZHI J, et al. Identification of wheat leaf disease based on random forest method [J]. Journal of Graphics, 2018, 39(1): 57-62.)
- [2] 王聘,柴秀娟. 机器学习在植物病害识别研究中的应用[J]. 中国农机化学报, 2019, 40(9): 171-180. (WANG D, CHAI X J. Application of machine learning in plant diseases recognition [J]. Journal of Chinese Agricultural Mechanization, 2019, 40(9): 171-180.)
- [3] 陈国军,程琰,曹岳,等. 基于目标特征的植株深度图像修复[J]. 图学学报, 2019, 40(3): 460-465. (CHEN G J, CHENG Y, CAO Y, et al. Plant depth maps recovery based on target features [J]. Journal of Graphics, 2019, 40(3): 460-465.)
- [4] GUETTARI N, CAPELLE-LAIZÉ A S, CARRÉ P. Blind image steganalysis based on evidential  $k$ -nearest neighbors [C]// Proceedings of the 2016 IEEE International Conference on Image Processing. Piscataway: IEEE, 2016: 2742-2746.
- [5] DEEPA S, UMARANI R. Steganalysis on images using SVM with selected hybrid features of Gini index feature selection algorithm [J]. International Journal of Advanced Research in Computer Science, 2017, 8(5): 1503-1509.
- [6] RAMEZANI M, GHAEMMAGHAMI S. Towards genetic feature selection in image steganalysis [C]// Proceedings of the 2010 7th IEEE Consumer Communications and Networking Conference. Piscataway: IEEE, 2010: 1-4.
- [7] SHEIKHAN M, PEZHMANPOUR M, MOIN M S. Improved contourlet-based steganalysis using binary particle swarm optimization and radial basis neural networks [J]. Neural Computing and Applications, 2012, 21(7): 1717-1728.
- [8] KODOVSKY J, FRIDRICH J, HOLUB V. Ensemble classifiers for steganalysis of digital media [J]. IEEE Transactions on Information Forensics and Security, 2011, 7(2): 432-444.

- [9] GUO Y, HASTIE T, TIBSHIRANI R. Regularized linear discriminant analysis and its application in microarrays [J]. *Biostatistics*, 2007, 8(1): 86-100.
- [10] ZHANG S, WANG Z. Cucumber disease recognition based on global-local singular value decomposition [J]. *Neurocomputing*, 2016, 205: 341-348.
- [11] ZHANG S, WU X, YOU Z, et al. Leaf image based cucumber disease recognition using sparse representation classification [J]. *Computers and Electronics in Agriculture*, 2017, 134: 135-141.
- [12] MOHANTY S P, HUGHES D P, SALATHÉ M. Using deep learning for image-based plant disease detection [J]. *Frontiers in Plant Science*, 2016, 7: Article No. 1419.
- [13] MA J, DU K, ZHENG F, et al. A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network [J]. *Computers and Electronics in Agriculture*, 2018, 154: 18-24.
- [14] KAWASAKI Y, UGA H, KAGIWADA S, et al. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks [C]// *Proceedings of the 2015 International Symposium on Visual Computing, LNCS 9475*. Cham: Springer, 2015: 638-645.
- [15] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. [2020-07-14]. <https://arxiv.org/pdf/1704.04861.pdf>.
- [16] LIN M, CHEN Q, YAN S. Network in network [EB/OL]. [2020-07-21]. <https://arxiv.org/pdf/1312.4400.pdf>.
- [17] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]// *Proceedings of the 2015 32nd International Conference on Machine Learning*. New York: JMLR.org, 2015: 448-456.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2020-07-15]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [19] HUGHES D P, SALATHÉ M. An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing [EB/OL]. [2020-07-17]. <https://arxiv.org/ftp/arxiv/papers/1511/1511.08060.pdf>.
- [20] DARWISH A, EZZAT D, HASSANIEN A E. An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis [J]. *Swarm and Evolutionary Computation*, 2020, 52: Article No. 100616.
- [21] KC K, YIN Z, WU M, et al. Depthwise separable convolution architectures for plant disease classification [J]. *Computers and Electronics in Agriculture*, 2019, 165: Article No. 104948.

This work is partially supported by the Fundamental Research Funds for the Central Universities (2572019BF04).

**JIA Heming**, born in 1983, Ph. D., professor. His research interests include digital image processing, pattern recognition.

**LANG Chunbo**, born in 1998, Ph. D. candidate. His research interests include graphics and image processing, computer vision.

**JIANG Zichao**, born in 1995, M. S. candidate. His research interests include machine learning, image segmentation.