

基于改进VGG16的大豆灰斑病识别

贾倩, 郭俊*, 陈勇明, 梅明林

成都信息工程大学应用数学学院, 四川 成都

收稿日期: 2024年10月21日; 录用日期: 2025年1月2日; 发布日期: 2025年1月10日

摘要

作为重要的粮食作物, 大豆在全球市场中需求巨大。大豆灰斑病严重制约了大豆的优质和高产, 因此, 快速而准确地识别灰斑病对于大豆种植具有重要意义。目前基于深度学习的大豆灰斑病识别方法存在限制, 包括对输入图像尺寸的特定要求、准确率不足, 以及模型参数量大导致计算缓慢。针对这些问题, 本研究提出了一种改进的VGG16图像识别方法。该方法基于VGG16模型进行了优化, 保留了三层卷积层结构, 并集成了Inception模块。每个Inception模块由三个并行分支构成: 1×1 卷积层用于降维, 1×1 后接 3×3 卷积层用于特征提取, 以及 3×3 最大池化层后接 1×1 卷积层以聚合空间信息。此外, 本研究采用全局平均池化层替代了传统的全连接层, 减少了模型复杂度并提升计算效率。实验结果表明, 该结构不仅放宽了对输入图像尺寸的限制, 还显著减少了模型参数, 使得参数个数仅为传统VGG16的1.5%, 而模型准确率达到99.1%。

关键词

大豆灰斑病, 深度学习, VGG16, Inception模块, 全局平均池化

Recognition of Gray Spot of Soybean Based on Improved VGG16

Qian Jia, Jun Guo*, Yongming Chen, Minglin Mei

College of Applied Mathematics, Chengdu University of Information Technology, Chengdu Sichuan

Received: Oct. 21st, 2024; accepted: Jan. 2nd, 2025; published: Jan. 10th, 2025

Abstract

As an important food crop, soybean is in great demand in the global market. Gray spot of soybean seriously restricts the quality and high yield of soybean, so it is of great significance to identify gray

*通讯作者。

文章引用: 贾倩, 郭俊, 陈勇明, 梅明林. 基于改进 VGG16 的大豆灰斑病识别[J]. 人工智能与机器人研究, 2025, 14(1): 43-51. DOI: 10.12677/airr.2025.141005

spot quickly and accurately for soybean planting. Current deep learn-based methods for soybean gray spot recognition have limitations, including specific requirements for input image size, inadequate accuracy, and slow calculation due to the large number of model parameters. To solve these problems, an improved VGG16 image recognition method is proposed in this paper. The method is optimized based on the VGG16 model, retains the three-layer convolutional layer structure, and integrates the Inception module. Each Inception module consists of three parallel branches: a 1×1 convolutional layer for dimensionality reduction, a 1×1 convolutional layer followed by a 3×3 convolutional layer for feature extraction, and a 3×3 maximum pooling layer followed by a 1×1 convolutional layer for aggregation of spatial information. In addition, the global average pooling layer is used to replace the traditional fully connected layer, which reduces the model complexity and improves the computational efficiency. The experimental results show that the structure not only eases the limit on the input image size, but also significantly reduces the model parameters, making the number of parameters only 1.5% of the traditional VGG16, and the model accuracy reaches 99.1%.

Keywords

Gray Spot of Soybean, Deep Learning, VGG16, Inception Module, Global Average Pooling

Copyright © 2025 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

大豆作为中国重要粮食作物之一，市场需求量巨大。大豆灰斑病[1]，见图 1，由大豆尾孢菌引起的真菌性病害，对大豆的种子、叶片、茎部和豆荚构成严重威胁。该病害的发生通常导致大豆减产约 13%，而在严重情况下，减产幅度可达 30%至 50%。大豆灰斑病的防治策略主要包括三个方向：抗病品种的选育、化学防治以及农业防治措施。在这些方法中，化学防治因其效果显著、应用广泛和操作便捷而成为主要的防治手段。因此，准确识别大豆灰斑病对于大豆生产过程中的早期防治至关重要。然而，传统的识别方法不仅耗费大量的人力和物力，而且效率极低，这严重延缓了对大豆灰斑病的及时防控。



Figure 1. Distinctive features of normal versus gray spot-affected soybean leaves

图 1. 正常大豆叶片和患有灰斑病叶片对比

在大豆病害识别领域，金乔[2]提出了一种运用图像背景分割、生成对抗网络[3]、卷积神经网络、迁移学习等图像处理关键技术的大豆叶害识别技术，该模型的识别准确率为 94.84%；蒋丰千等人[4]在针对大豆叶斑病、花叶病、霜霉病和灰斑病利用卷积神经网络设计大豆病害检测系统，准确率达到 78%；Karlekar 等人[5]在对 PDDB 数据集上的大豆病害图片先进行图片背景去除，基于保留下来的图片大小(32 × 32)介绍了一种基于深度学习卷积神经网络的大豆病害识别方法，识别数据集中 16 种大豆病害，模型识

别精度达到 98.14%。现有研究表明,深度学习技术在提高大豆病害识别准确率方面具有巨大潜力。随着研究的不断深入,研究显示对卷积神经网络的结构进行优化是提升模型性能的有效途径。例如,VGG 模型[6]通过适当的改进,不仅可以增强其特征提取的能力,还能够显著减少模型的参数数量,从而提高计算效率。岳有军等人[7]通过引入高阶残差和参数共享反馈子网络改进了 VGG 网络,以提高农作物病害图像识别的准确性和鲁棒性。在复杂自然环境下,该方法能有效降低背景噪声的干扰,提升识别精度。鲍文霞等人[8]开发了基于改进卷积神经网络的苹果叶部病害识别的网络模型。将 VGG16 网络模型从 ImageNet 数据集上学习到的先验知识迁移到苹果病害叶片数据集上,采用选择性核卷积和全局平均池化技术,有效提高了小病斑区域的识别准确率和模型的收敛速度。王美娟等人[9]通过对传统 VGGNet 模型的改进,提高了玉米病害症状的识别准确度至 99.31%。许景辉等人[10]在 VGG16 模型的基础上设计了全新的全连接层模块,并将 VGG16 模型在 ImageNet 图像数据集训练好的卷积层迁移到本模型中来对玉米病害图像进行识别,平均识别准确率为 95.33%。胡骏等人[11]针对水稻稻瘟病图像识别中人工识别率低、识别能力差和识别技术应用不普及的问题,提出了基于改进 VGG16 模型的水稻稻瘟病的识别模型,以减少模型参数,降低运算负荷,提高检测性能,加快模型收敛。尽管这些研究在提高识别准确率方面取得了显著进展,但在大豆灰斑病识别上,现有的模型在识别精度、输入图像尺寸的适应性以及参数数量导致的计算效率方面存在局限性。因此在对大豆灰斑病识别中建立一个高准确率且普遍适用的模型就显得尤为重要。

因此,本研究开发了一种基于 VGG16 架构优化的深度学习模型,旨在同时解决对输入图片大小的要求和参数量大导致运行缓慢的问题。模型的创新之处包括:

1) 保留 VGG16 的三层卷积层: 本研究保留了 VGG16 的三层卷积层,其中卷积核尺寸为 3×3 ,以保持特征提取的效率和效果。

2) 引入优化的 Inception 模块: 在 Inception 模块中,本研究引入了三个并行分支:

a) 1×1 卷积层: 旨在降维,减少计算复杂度。

b) 1×1 卷积层后接 3×3 卷积层: 目的在于特征提取,增强模型对局部特征的捕捉能力。

c) 3×3 最大池化层后接 1×1 卷积层: 用于聚合空间信息,增强模型对全局特征的感知。

3) 使用全局平均池化层: 本研究使用全局平均池化层替代了传统的全连接层,有效地将每个特征图的空间维度降维至 1×1 。

4) 特征映射: 在卷积层输出的特征图上应用全局平均池化,将每个特征图的空间维度降为 1×1 ,然后通过 softmax 函数将特征映射到各个类别上。这种方法实现了特征图中每个位置的特征都被考虑在内,而不仅仅是依赖于全连接层的权重矩阵。

这些优化措施使得模型在保持计算效率的同时,提高了对特征的捕捉能力[12]和分类性能。实验结果表明,改进后的模型准确率高达 99.1%,显著提高了改进前的 98.3%,并且模型参数大大降低,从原来的 1.3×10^8 降低至 2×10^6 ,仅为改进前的 1.5%。归一化混淆矩阵见图 10,灰斑病正确分类实例的比例为 0.95,这一结果突显了模型在大豆灰斑病识别任务中的高预测精度。该指标进一步证实了模型在特定植物病理学诊断中的可靠性与有效性。

2 实验设计

2.1. 实验数据集

1) 数据来源: 本研究的数据集来源于网站“探险家的数据窝”,可通过以下链接访问:
<https://www.dilitanxianjia.com/1974/>。该数据集包含大豆叶片病害检测图像,涉及两种病害类型及健康状态。大豆叶片灰斑病和大豆正常照片见图 2 和图 3。



Figure 2. Manifestations of gray spot of soybean on foliar tissue

图 2. 部分灰斑病图片



Figure 3. Morphological characteristics of healthy soybean leaves

图 3. 部分正常叶片图片

2) 图像裁剪：数据导入后，采用 transform 函数对数据进行裁剪，以满足 VGG16 模型的输入大小。

3) 数据加强：鉴于 VGG16 作为卷积神经网络中的一个重要模型，在大规模图像识别任务中表现优异，因此运用 transforms.RandomHorizontalFlip 模块对图像进行随机水平翻转，增强训练数据的多样性，以提高模型的泛化能力。

4) 标准化处理：为了使得输入数据的分布更加统一[13]，避免模型因为输入数据的尺度的不同而导致对模型训练困难，同时加快模型的收敛速度和提高模型的训练稳定性，在本研究的模型中使用 transforms.Normalize 模块进行数据标准化处理。

2.2. 卷积神经网络

从 1962 年生物学家在对猫脑视觉皮层的研究中初次发现类似神经网络结构之后，经过数十年的研究，神经网络已经形成了一套完整的理论体系。LeCun 等人基于 Fukushima 的研究使用反向传播(BP)算法设计并训练了卷积神经网络(CNN)，在模型识别领域展现出良好的分类性能[14]。

卷积神经网络作为一种深度学习模型，主要用于处理具有网格结构的数据，如图像和视频。CNN 主要由输入层、卷积层、池化层、全连接层和输出层构成，通过卷积层和池化层的交替使用，以及卷积层后接池化层再接卷积层的结构，自动提取和学习输入数据的特征表示。

1) 卷积层：卷积层是卷积神经网络的基本构建单元，它通过对输入数据和一组可学习的卷积核进行卷积操作来提取特征。每个卷积核会在输入数据上滑动，计算局部区域的加权和，从而生成特征图。卷积操作的具体计算公式如下：

$$(f * g)(i, j) = \sum_m \sum_n f(m, n) \cdot g(i - m, j - n)$$

其中 f 是输入图像或特征图， g 是卷积核或滤波器， (i, j) 是输出特征图上的位置， (m, n) 是卷积核上的位置， $*$ 表示卷积操作。

2) 池化层：池化层的目的在于降低特征图的空间维度，从而减少模型的参数数量和计算量，同时保留关键特征。常用的池化操作包括最大池化和平均池化[15]。池化层的工作原理和卷积层类似，但在卷积层中，卷积过程涉及各个相应位置数值的累加，而最大池化则选取相应矩阵中的最大值，平均池化则计算相应矩阵上的平均值。

例如，当使用 2×2 的卷积核和步长为 2 时，池化操作的效果参见图 4 和图 5。

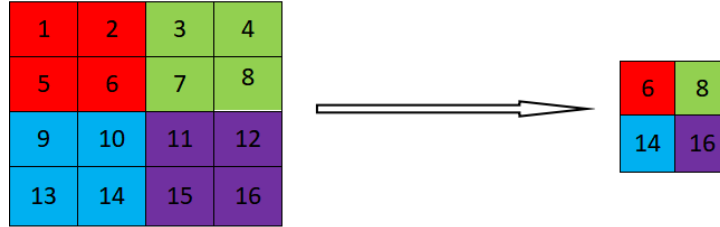


Figure 4. Visualizing the operation of maximum pooling in CNNs

图 4. 卷积神经网络中最大池化操作的可视化

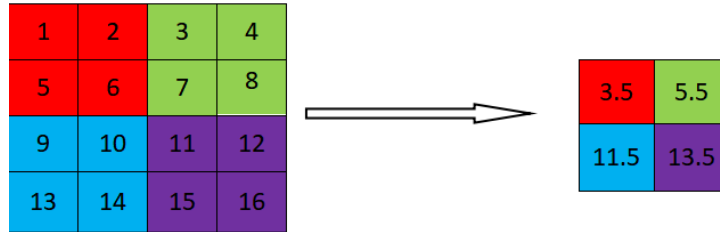


Figure 5. Visualizing the operation of average pooling in CNNs

图 5. 卷积神经网络中平均池化操作的可视化

3) 激活函数：在卷积层之后，通常会包含非线性激活函数，例如 ReLU，用于引入非线性特性，增加模型的表达能力，其数学表达式为：

$$\text{ReLU}(x) = \max(0, x)$$

4) 全连接层：全连接层将前一层的所有特征映射到输出层，用于将卷积层和池化层提取的特征映射到最终的输出类别。

5) 损失函数：用于量化模型预测与真实标签之间的差异。常用的损失函数，如交叉熵损失函数应用于分类任务。在分类问题中，激活函数的常见数学表达式为：

$$L = -\sum_{c=1}^M y_{o,c} \log(\hat{y}_{o,c})$$

其中， M 是类别的总数， $y_{o,c}$ 是第 o 个样本的真实标签， $\hat{y}_{o,c}$ 是模型预测的第 o 个样本属于类别 c 的概率。

CNN 通过反向传播算法进行训练，通过最小化损失函数来调整网络参数，使得模型能够学习到输入数据的有效特征表示，从而实现对各种视觉任务的有效处理。

2.3. VGG16

VGG16，由牛津大学的 Visual Geometry Group 开发，是一个卷积神经网络模型。该模型在 2014 年的 ImageNet 大规模视觉识别挑战赛中取得了优异的成绩，代表了当时深度学习模型架构的先进水平。

VGG16 包含 16 个卷积层和 3 个全连接层，具有深度网络结构。该模型的特征在于所有的卷积层均采用 3×3 的卷积核，步长设置为 1 和填充为 1，同时通过池化层降低特征图的空间维度。鉴于模型的深度和参数量较大，VGG16 通常需要较多的计算资源和更长的训练时间，但在很多视觉任务上表现出了较好的性能。

图 6 展示了 VGG16 模型的结构。VGG16 是一种特定的卷积神经网络模型，输入层接受 224×224 像素大小的 RGB 三色通道图像。模型的卷积层部分由 5 个卷积块组成：第一个卷积块由两个卷积层组成，每个卷积层配备 64 个大小为 3×3 的卷积核，步长为 1，填充为 1，每个卷积层后均接一个 ReLU 激活函数。类似地，第二个卷积块包含两个卷积层，而第三、四和五卷积块各包含三个卷积层。每个卷积块结束后均接一个最大池化层，池化层的卷积核大小为 2×2 ，步长为 2。卷积层之后包含三个全连接层，每个全连接层配备 4096 个神经元。最后一个全连接层作为输出层，配备 1000 个神经元，并且使用 softmax 激活函数。

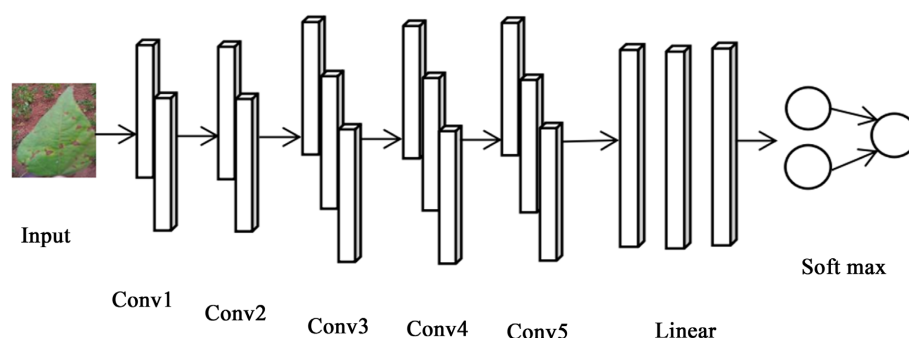


Figure 6. Schematic representation of the VGG16 Neural network model

图 6. VGG16 神经网络模型的结构图示

2.4. 改进的 VGG16

对 VGG16 模型进行了优化，旨在提高大豆灰斑病识别的准确性和效率。改进措施包括：

引入三支 Inception 模块：相较于文献[9]中提出的四分支结构，该设计减少了模型的复杂度，同时保持了对多尺度特征的有效捕捉。

优化特征提取：三支 Inception 模块通过 1×1 卷积层进行降维， 1×1 后接 3×3 卷积层进行特征提取，以及 3×3 最大池化层后接 1×1 卷积层进行空间信息聚合，有效地提取了大豆叶片图像的多尺度特征。

使用全局平均池化层：采用全局平均池化层替代了传统的全连接层，有效地将每个特征图的空间维度降维至 1×1 ，从而减少了模型的参数数量和计算复杂度。

减少参数数量：相比于四分支结构，三支设计减少了模型的参数数量，降低了计算资源的需求，提高了模型的训练和推理速度。改进的 VGG16 模型结构见图 7：

改进后模型的前向传播(forward)过程包括以下步骤：

- 1) 输入的张量 x 经过 VGG16 的特征提取部分(self.features)。
- 2) 随后，这些特征图作为输入传递给 Inception 模块(self.inception)。
- 3) 经过 Inception 模块后，得到的特征图经过自适应平均池化层(self.avgpool)进行全局平均池化操作，将特征图的尺寸调整为 1×1 。
- 4) 调整后的特征图通过 torch.flatten 函数被压平为一维张量。
- 5) 特征张量随后通过两个全连接层。

6) 最终, 对全连接层的输出通过 softmax 函数(`self.softmax`)进行归一化处理, 以得到每个类别的概率分布。

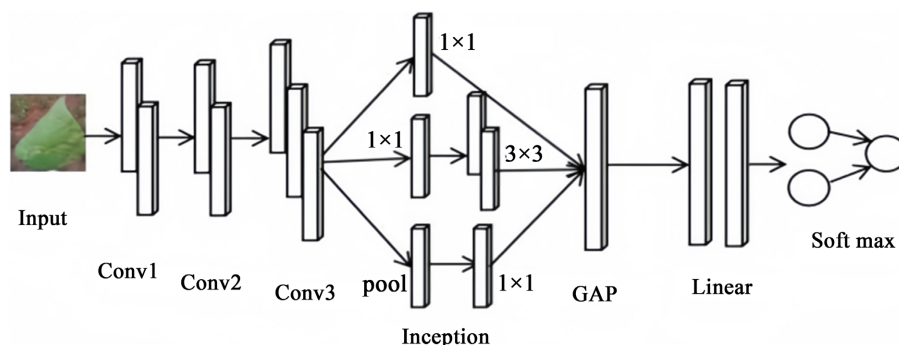


Figure 7. Schematic representation of the improved VGG16 neural network model
图 7. 改进的 VGG16 神经网络模型的结构图

2.5. 模型训练

实验环境配置为 Python3.9, 深度学习框架为 Pytorch1.13, 并配合 CUDA11.6。参照图 8, 在训练网络的过程中, 通过批处理迭代数据, 应用反向传播算法来优化模型参数, 同时计算损失值和训练集准确率, 以实现神经网络模型的训练过程。通过反向传播算法来最小化损失函数, 能够计算损失函数相对于模型参数的梯度, 随后根据这些梯度来更新参数, 促使损失函数值逐渐降低, 进而提高模型对数据的拟合度。具体操作流程见图 8。

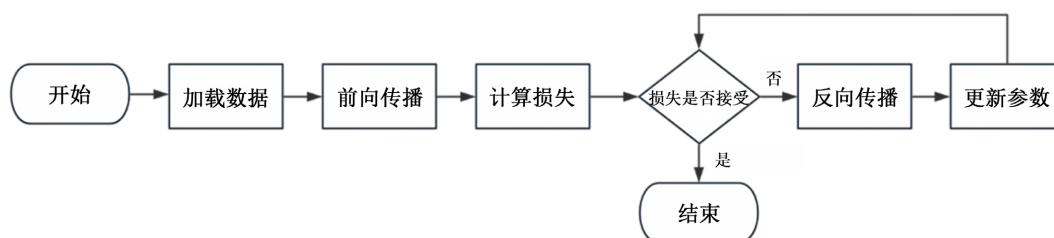


Figure 8. Flowchart of the deep learning model training procedure
图 8. 深度学习模型训练流程图

3. 结果分析

本研究融合了 VGG16 与 Inception 模块的优势, 保留 VGG16 的前三个卷积块以优化提取特征, 然后引入 Inception 模块, 利用不同尺寸的卷积核提取多尺度特征。最终, 采用全局池化代替了一个全连接层, 减少了模型参数数量, 显著降低了计算负担。

如表 1 所示, 改进后的模型准确率高达 99.1%, 显著高于改进前的 98.3%, 同时, 模型参数数量显著减少, 改进后的模型参数数量仅为改进前的 1.5%。此外, 本研究还根据数据特点(图片大小为 500×500)设计了相应的输入尺寸为 $500 \times 500 \times 3$ 的卷积网络结构, 所得模型的准确率达到 98.8%, 但模型的参数量是传统 VGG16 的三倍。

由图 9 所示, 随着训练周期的增加, 模型的训练准确率显著提高, 初始值约为 0.5, 逐渐增加至接近 1.0。这表明模型在训练数据上学得越来越准确。在验证集上, 与训练准确率相比, 验证准确率的起始点较低, 但同样呈现出上升趋势。

Table 1. Comparative analysis of classification accuracy and model parameters
表 1. 分类准确性与模型参数的比较分析

	准确度	模型参数
VGG16	98.3%	134281029
VGG16withInception	99.1%	2072258
MyClassifier	98.8%	492581186

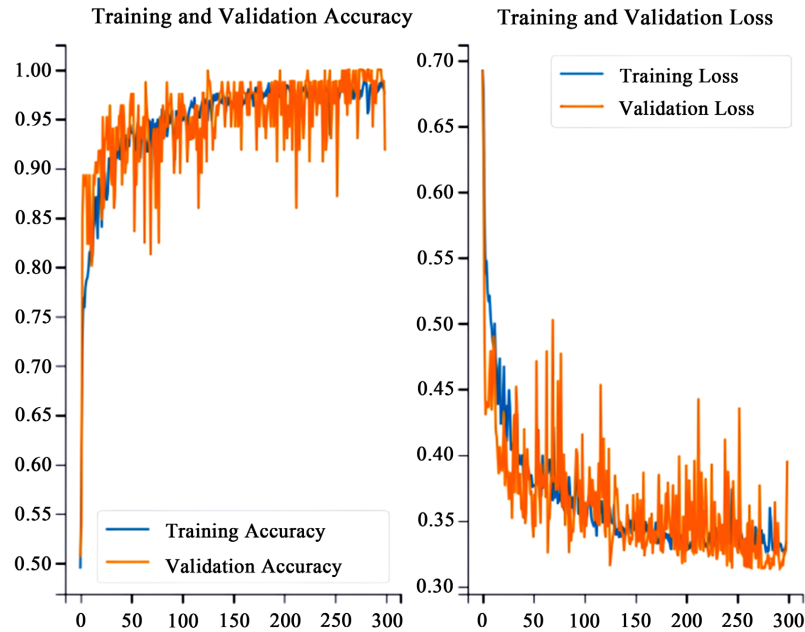


Figure 9. Dynamic changes in loss function throughout the training phase
图 9. 训练阶段损失函数的动态变化

图 10 展示的归一化混淆矩阵显示灰斑病的正确分类率为 0.95，表明模型在识别大豆灰斑病方面具有较高的预测准确性。

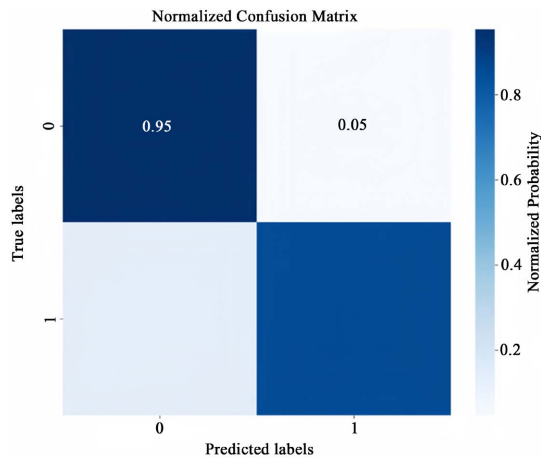


Figure 10. Normalized confusion matrix for model performance evaluation
图 10. 模型性能评估的归一化混淆矩阵

4 总结

本文提出了一种改进的 VGG16 模型,用于大豆灰斑病的识别。该模型不仅提高了识别准确率,同时显著降低了模型的参数数量,进而提高了运算效率。因此,卷积神经网络凭借其强大的特征提取能力,在图像识别领域,特别是在农业病害诊断中,展现出了巨大的应用潜力。本研究只聚焦于大豆叶片单一病害的识别,然而,在实际应用中,大豆叶片可能同时遭受多种病害的侵袭,这些病害的特征可能在形态、颜色和纹理上的特征可能存在显著差异,增加了识别任务的难度。针对这一挑战,未来的研究应致力于开发能够识别和区分多种病害的模型。这些模型应具备高度的泛化能力,以应对不同病害特征的多样性和复杂性。此外,研究还应考虑模型的实时性和准确性,以满足农业病害快速诊断的实际需求。通过融合先进的图像处理技术和深度学习算法,未来的研究有望实现对大豆叶片病害的高效、准确的综合识别,从而为农业生产提供强有力的技术支持。

基金项目

四川省科技厅项目资助(2022JDR0043);四川省数值模拟重点仿真实验室项目(KLNS-2023SZFZ002);成都信息工程大学青年创新项目(KYQN202324)。

参考文献

- [1] 马永波. 大豆灰斑病的发生特点及防治措施[J]. 新农业, 2022(20): 17.
- [2] 金乔. 基于深度学习的大豆病害识别研究[D]: [硕士学位论文]. 长春: 吉林农业大学, 2023.
- [3] 张凯, 陈亚军, 张俊. 生成对抗网络在医学小样本数据中的应用[J]. 内江师范学院学报, 2020, 35(4): 57-60.
- [4] 蒋丰千, 李旻, 余大为, 等. 基于 Caffe 卷积神经网络的大豆病害检测系统[J]. 浙江农业学报, 2019, 31(7): 1177-1183.
- [5] Karlekar, A. and Seal, A. (2020) SoyNet: Soybean Leaf Diseases Classification. *Computers and Electronics in Agriculture*, **172**, Article ID: 105342. <https://doi.org/10.1016/j.compag.2020.105342>
- [6] Liu, Y., Pu, H. and Sun, D. (2021) Efficient Extraction of Deep Image Features Using Convolutional Neural Network (CNN) for Applications in Detecting and Analysing Complex Food Matrices. *Trends in Food Science & Technology*, **113**, 193-204. <https://doi.org/10.1016/j.tifs.2021.04.042>
- [7] 岳有军, 李雪松, 赵辉, 等. 基于改进 VGG 网络的农作物病害图像识别[J]. 农机化研究, 2022, 44(6): 18-24.
- [8] 鲍文霞, 吴刚, 胡根生, 等. 基于改进卷积神经网络的苹果叶部病害识别[J]. 安徽大学学报(自然科学版), 2021, 45(1): 53-59.
- [9] 王美娟, 尹飞. 卷积神经网络的多尺度改进及其在玉米病害症状识别中的应用[J]. 河南农业大学学报, 2021, 55(5): 906-916.
- [10] 许景辉, 邵明烨, 王一琛, 等. 基于迁移学习的卷积神经网络玉米病害图像识别[J]. 农业机械学报, 2020, 51(2): 230-236, 253.
- [11] 胡骏, 陆兴华, 林桢菡, 等. 改进的 VGG16 在水稻稻瘟病图像识别中的应用[J]. 计算机应用, 2023, 43(z2): 196-200.
- [12] Ma, W., Wu, Y., Cen, F. and Wang, G. (2020) MDFN: Multi-Scale Deep Feature Learning Network for Object Detection. *Pattern Recognition*, **100**, Article ID: 107149. <https://doi.org/10.1016/j.patcog.2019.107149>
- [13] Singh, D. and Singh, B. (2020) Investigating the Impact of Data Normalization on Classification Performance. *Applied Soft Computing*, **97**, Article ID: 105524. <https://doi.org/10.1016/j.asoc.2019.105524>
- [14] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- [15] Hyun, J., Seong, H. and Kim, E. (2021) Universal Pooling—A New Pooling Method for Convolutional Neural Networks. *Expert Systems with Applications*, **180**, Article ID: 115084. <https://doi.org/10.1016/j.eswa.2021.115084>