doi:10.6041/j.issn.1000-1298.2024.11.002

基于轻量级密集多尺度注意力网络的小麦叶部锈病识别方法

鲍文霞　赵诗意　黄林生　梁　栋　胡根生

（安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心，合肥 230601）

摘要：人工诊断小麦锈病成本高、效率低，已无法满足现代农业生产的需要。本文提出了一种轻量级密集多尺度注意力网络模型（Mobile－Dense multi-scale attention net，Mobile－DMSANet），用于自动识别田间自然场景中的小麦叶部锈病（条锈病和叶锈病）。该模型在输入层设计了一个快速下采样模块（Fast subsampling block，FSB），它在不增加计算成本的前提下提高模型的特征表达能力。模型的特征提取层使用 3 个轻量级特征提取模块（Dense multi-scale attention，DMSA）来提取小麦叶部锈病的特征。DMSA 模块设计了一个多尺度的 3 路卷积层（Multi-scale three-way convolution，MSTC）用于获得不同尺度感受野，以提高模型的表达能力和对不同尺寸锈病的感知能力。DMSA 模块中 6 个 MSTC 层通过密集连接实现特征重用，不仅大大减少了模型的参数量，而且提高了对这两种相似的小麦叶部锈病的特征提取能力。在 DMSA 模块中还引入了协调注意力机制（Coordinated attention，CA），来提高对病害信息的敏感性，并抑制图像中的背景信息。模型的输出层使用 Softmax 函数实现小麦叶部锈病识别。结果表明，Mobile－DMSANet 模型在测试数据集上的识别准确率为 96.4%，高于经典 CNN 模型（如 ResNet50、AlexNet）和轻量级 CNN 模型（如 ShufflenetV2、DenseNet 系列）。Mobile－DMSANet 参数量为 4.54×10^5，与其他轻量级模型相比大幅下降。本文所设计模型可用于移动端小麦叶部锈病的自动识别。

关键词：小麦条锈病；小麦叶锈病；病害识别；轻量级卷积神经网络；Molile－DMSANet

中图分类号：TP391.41　　**文献标识码**：A　　**文章编号**：1000-1298(2024)11-0021-11　　**OSID**：

Lightweight Dense Multi-scale Attention Network for Identification of Rust on Wheat Leaves

BAO Wenxia　ZHAO Shiyi　HUANG Linsheng　LIANG Dong　HU Gensheng

(*National Engineering Research Center for Agro-Ecological Big Data Analysis and Application*, *Anhui University*, *Hefei* 230601, *China*)

Abstract: Artificial identification of wheat rust is costly and inefficient, and can no longer meet the needs of modern agricultural production. A lightweight dense multi-scale attention network model called Mobile－DMSANet was presented for the automatic identification of rust on wheat leaves (stripe rust and leaf rust) from images of natural scenes taken in the field. In the input layer of the network, a fast subsampling block (FSB) was used to improve the feature expression ability of the network without adding computational cost. In the feature extraction layer, three lightweight blocks called dense multi-scale attention (DMSA) blocks were used to extract the features of rust on wheat leaves. In the DMSA block, a multi-scale three-way convolution (MSTC) layer was designed to get different scales for the receptive fields, in order to improve the expressive ability of the network and its ability to perceive the features of rust disease at different scales. Six MSTC layers were used to achieve feature reuse by dense connections in the DMSA block, an approach that not only greatly reduced the number of parameters of the network but also improved the feature extraction ability for similar diseases. A coordinated attention (CA) block was also introduced to the DMSA block to increase the sensitivity to positional information and suppress background information in the image. The output layer of the network used a Softmax function to classify

收稿日期：2023－12－28　修回日期：2024－03－13

基金项目：安徽省自然科学基金项目（2208085MC60）、安徽省省厅高校科研计划项目（2023AH050084）和国家自然科学基金项目（62273001、32372632）

作者简介：鲍文霞（1980—），女，教授，博士，主要从事农业和生态视觉与模式识别研究，E-mail：bwxia@ahu.edu.cn

通信作者：黄林生（1977—），男，教授，博士，主要从事农业生态大数据分析研究，E-mail：linsheng0808@ahu.edu.cn

rust on wheat leaves. The results showed that the recognition accuracy of Mobile - DMSANet model on the test dataset was 96.4%, which was higher than that of other models. Mobile - DMSANet had only 454 000 parameters, less than for other lightweight models. The proposed model can be used for the automatic identification of rust on wheat leaves using mobile devices.

Key words: wheat stripe rust; wheat leaf rust; disease identification; lightweight convolutional neural network; Molile - DMSANet

0 引言

小麦是一种在世界范围内广泛种植的粮食作物。小麦叶部病害的发生严重影响了小麦的产量和质量[1-4],其中条锈病和叶锈病是我国小麦叶部病害中传播最快、发生面积最广、危害最大的病害[5]。小麦条锈病是由条形锈菌引起的病害,条锈病的小夏孢子堆,颜色为鲜黄色,形状为椭圆形,在叶片上以虚线状沿叶脉纵向排列。小麦叶锈病又称黄疸病,是由小麦隐匿柄锈菌侵染所引起的病害,叶锈病的中等大小夏孢子堆,颜色为黄褐色,形状接近圆形,在叶片上分布散生不规则[6]。这两种病害症状极其相似,在实际生产中难以区分,但这两种病害的发生率对于小麦病害预测预报和综合治理具有重要意义。目前,我国小麦锈病的诊断主要是依靠植保人员目测观察作物的表型,有时会出现误判。因此,研究一种低成本、快速、可以移植到移动端的小麦叶部锈病识别模型具有重要的现实意义[7]。

随着图像识别和机器学习的快速发展,农作物病害识别进入了一个新的时代。MOSHOU 等[8]提出一种多层感知器识别小麦锈病的算法,该算法主要利用了健康植株和患锈病植株的光谱反射率差异。TIAN 等[9]以小麦叶部白粉病、小麦锈病、叶枯病、刺柏病为对象,对颜色特征集、形状特征集和纹理特征集进行聚类分析,提出一种基于支持向量机的小麦叶部病害多分类新策略。鲍文霞等[3]利用椭圆型度量计算特征向量之间的距离,实现不同严重度小麦叶片病害的识别。然而,人为设计的特征具有不完备性,传统机器学习对小麦叶部病害识别性能不佳。

近年来,卷积神经网络模型在图像识别和分类中的应用日益广泛。XIONG 等[10]设计了一种自动图像分割算法,该算法保留作物病斑并自动去除图像中的背景信息,然后结合 MobileNet 进行疾病识别和分类。实验结果表明,该系统对 6 种作物的 27 种病害的识别率在 80% 以上。GAO 等[11]提出一种基于双分支、高效、通道关注(DECA)的作物病害识别 CNN 模型。结果表明,在 PlantVillage、AI Challenger 2018 和黄瓜病害数据集上,DECA 模型准确率分别达到 86.65%、99.74% 和 98.54%。

农作物病害识别大多数情况下是在野外进行,上述模型庞大的参数与计算量使得其无法直接运行在移动设备上。除此之外,田间自然场景中获取的图像由于背景复杂、光照强度不均匀,上述模型的识别精度大大下降。为了克服这些难点,LIN 等[12]提出一种基于坐标关注的多特征融合网络 CAMFFNet 模型,用于田间条件下的烟草病害识别,CAMFFNet 模型参数量只有 2.37×10^6。CHEN 等[13]提出一种用于玉米叶病害分类的轻量级卷积神经网络,精度达到 95.86%。CHEN 等[14]利用改进后的 DenseNet 对作物病害进行了分类,在密集块中使用深度可分卷积代替传统的卷积层,然后嵌入注意力模块来学习通道间关系和空间点对输入特征的重要性,通过这种方式提高了病害识别精度。

上述模型一定程度上解决了背景复杂、模型参数多等难点问题,但小麦叶部锈病识别还存在特征相似、病斑尺寸不一致的难点。为进一步缩减模型参数,提高模型准确率,并且解决以上难点问题,本文设计一种轻量级密集多尺度注意力网络模型 Mobile - DMSANet,用于田间自然场景下对小麦叶部条锈病、叶锈病的自动识别。

1 数据集构建与识别方法

1.1 数据集构建

1.1.1 数据获取

实验研究区域位于安徽省合肥市郭河镇安徽农业大学产学研基地。该试验基地作为国家级科研平台,同时也是国家小麦育种分中心,长期以来一直承担着小麦、水稻等多种农作物育种科研试验任务。小麦叶片图像采集于 2019 年 4 月 26 日至 5 月 10 日 09:00—16:00,采集设备为佳能 EOS 600D 型相机和 Redmi Note5 型手机。相机拍摄的图像分辨率为 3 456 像素 ×2 304 像素和 5 184 像素 ×3 456 像素,手机拍摄的图像分辨率为 4 000 像素 ×3 000 像素。拍摄后经过筛选得到 3 类小麦叶部锈病图像共计 849 幅,其中 386 幅为小麦条锈病,297 幅为小麦叶锈病,166 幅为健康小麦叶片。如图 1 所示,第 1 行为小麦健康叶片图像,标记为“0”;第 2 行为小麦条锈病图像,标记为“1”;第 3 行为小麦叶锈病图像,标记为“2”。图 1 中使用红色框来突出病斑区域,

可看出条锈病和叶锈病极其相似，病斑大小不一致，图像背景复杂，光线不均匀。



图 1　小麦叶片图像

Fig. 1　Images of wheat leaves

1.1.2　数据预处理与数据扩充

1.1.2.1　数据预处理

由于田间自然场景拍摄的图像光照不均匀，本文采用单尺度 Retinex 算法[15-16]对图像进行光照预处理。根据 Retinex 理论，物体颜色由物体对长波（红色）、中波（绿色）、短波（蓝色）光线反射能力决定，而不是由反射光强度绝对值决定[17]。其基本原理为

$$S(x,y)=R(x,y)L(x,y) \tag{1}$$

式中　x——图像中像素点横坐标

y——图像中像素点纵坐标

$R(x,y)$表示物体反射图像，$L(x,y)$表示入射图像，$S(x,y)$表示人眼看到的视觉图像。$R(x,y)$可以变换为

$$R(x,y)=\frac{S(x,y)}{L(x,y)} \tag{2}$$

$$\lg R(x,y)=\lg S(x,y)-\lg L(x,y) \tag{3}$$

入射光通常是一种相对均匀且变化缓慢的噪声，因此 $L(x,y)$可以用低频成分表示为

$$L(x,y)=F(x,y)S(x,y) \tag{4}$$

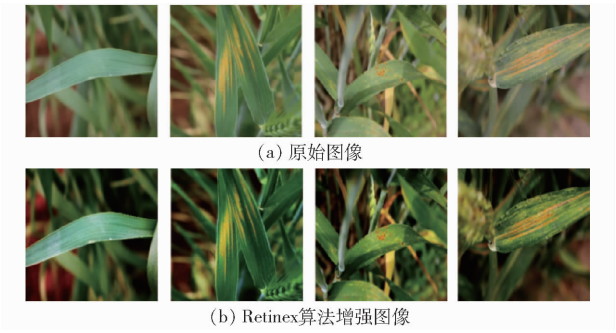
$F(x,y)$表示中心环绕函数，将式(4)代入式(3)中，得到

$$\lg R(x,y)=\lg S(x,y)-\lg(F(x,y)S(x,y)) \tag{5}$$

最后将结果由对数域转换到实数域，从而可以得到物体反射图像 $R(x,y)$，达到提升图像质量的效果。图 2 为使用 Retinex 算法对光照进行预处理后的部分图像。通过对比可看到，去除图像中的低频分量后，图像纹理变得更加清晰。

1.1.2.2　数据扩充

为了增加训练样本多样性，减少模型过拟合，对数据集样本进行了扩充。为了确保训练集、验证集和测试集的分布完全独立，将小麦叶部锈病数据集划分为训练集、验证集和测试集。然后，随机选择旋转、平移、翻转、对比度增强等方法来平衡训练集和验证集数据分布，增加了图像之间的多样性。扩充后样本分布如表 1 所示，样本图像如图 3 所示。样本扩充后训练集共 2 605 幅图像，验证集共 434 幅图像，279 幅测试图像不进行扩充。



(a) 原始图像

(b) Retinex算法增强图像

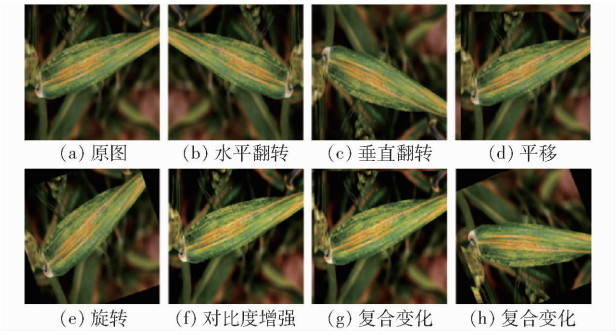
图 2　原始图像和 Retinex 算法图像增强示例

Fig. 2　Original images and enhanced images using Retinex

表 1　扩充后数据分布

Tab. 1　Sample distribution after augmentation

类别	训练集样本数	验证集样本数	测试集样本数
健康	864	144	54
条锈病	888	148	127
叶锈病	853	142	98
总计	2 605	434	279



(a) 原图　(b) 水平翻转　(c) 垂直翻转　(d) 平移

(e) 旋转　(f) 对比度增强　(g) 复合变化　(h) 复合变化

图 3　部分扩充图像

Fig. 3　Examples of augmented images

1.2　识别方法

1.2.1　Mobile-DMSANet 模型

为减少模型参数，提高锈病识别率，并解决田间自然场景中小麦锈病图像背景复杂、两种锈病相似、病斑大小不一等问题，本文设计了一个轻量级密集多尺度注意力网络模型 Mobile-DMSANet，该模型结构如图 4 所示，主要参数如表 2 所示。Mobile-DMSANet 主要由快速下采样模块 FSB、3 个轻量级特征提取模块 DMSA block、2 个过渡层和分类器组成。

在模型输入层，设计一种快速下采样模块 FSB。它的主要功能是实现输入图像空间维度的第一次下采样和通道数的增加。FSB 用卷积和池化融合的双分支结构取代 7×7 的卷积，在减少计算量的

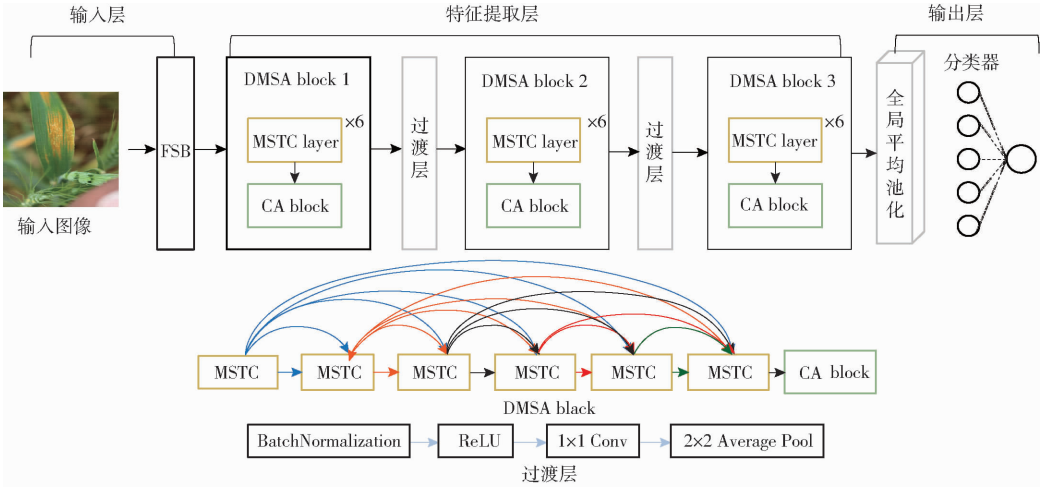


图 4 Mobile－DMSANet 模型结构图

Fig. 4 Structure of proposed Mobile－DMSANet model

表 2 Mobile－DMSANet 主要参数

Tab. 2 Primary parameters of Mobile－DMSANet

层名称	输入尺寸	模块类型	输出尺寸	通道数	重复次数
Input 1	224×224×3	Input layer	224×224		
Input 2	224×224×3	FSB	56×56	32	1
DMSA block 1	56×56×32	MSTC layer	56×56	24	6
DMSA block 1	56×56×176	CA block	56×56	176	1
Transform1	56×56×176	1×1 Conv	56×56	88	1
Transform2	56×56×88	Average Pooling	28×28	88	1
DMSA block 2	28×28×88	MSTC layer	28×28	24	6
DMSA block 2	28×28×232	CA block	28×28	232	1
Transform1	28×28×232	1×1 Conv	28×28	116	1
Transform2	28×28×116	Average Pooling	14×14	116	1
DMSA block 3	14×14×116	MSTC layer	14×14	24	6
DMSA block 3	14×14×260	CA block	14×14	260	1
Classification layer	14×14×260	7×7 Global Average Pool 3D fully-connected, Softmax	Number of categories	1	1

同时，该模块能确保较强的特征表达能力。

模型特征提取层由 3 个轻量级模块 DMSA block 和 2 个过渡层构成。DMSA block 由 6 个 MSTC layer 密集连接，再插入 CA 协调注意力机制[18]组成。DMSA block 通过密集连接实现特征重用，既大幅度减少了模型参数量，缓解了过拟合现象，又提高对两种相似锈病的特征提取能力。MSTC layer 通过 3 个分支来得到不同尺度的感受野，从而提高模型对不同尺寸小麦叶部锈病的感知能力和模型表达能力。CA 协调注意力机制将位置信息嵌入通道注意力机制中，能够有效地提取图像中锈病特征信息，使得模型能够突出病斑区域，抑制背景信息。过渡层[14]用一个 1×1 卷积降低维度和 2×2 平均池化层减小特征图大小，并连接两个相邻的 DMSA block，这进一步有效减少了参数量，达到模型轻量化的效果。在模型输出层使用 Softmax 分类器实现锈病识别。所提出 Mobile－DMSANet 参数量为 4.54×10^5，参数量较少。

1.2.2 DMSA block

如图 4 所示，DMSA block 采用密集连接的方式将所有 MSTC layer 彼此连接。对于每一个 MSTC layer 来说，所有前一层的特征图作为输入，而这一层的特征图用作所有后续层的输入，意味着可以直接从损失函数和原始输入信号中获取梯度[19]。这样，加强了整个模型信息流和梯度传播，从而缓解了条锈病和叶锈病特征差异小而引起的过拟合问题，使模型更易于训练，并具有一定的正则效果。数学上，这种密集连接[14]可以表示为

$$x_l = H_l([x_0, x_1, \cdots, x_{l-1}]) \tag{6}$$

式中 l——模型层数

H_l——非线性转化函数

它是一个组合操作，包括一系列的 BN、ReLU、Pooling 等操作。$[x_0, x_1, \cdots, x_{l-1}]$ 表示 0 层到 $l-1$ 层的特征图拼接。

如果每个函数 H_l 产生 k 个特征图，则第 l 层有 $k_0+k(l-1)$ 个输入特征，其中 k_0 是输入层的通道数，超参数 k 是模型的增长率。每一个 MSTC layer 都可以访问其 DMSA block 中的所有先前的特征图，因此可以访问模型的“集体知识”。每一层向这个“集体知识”添加 k 个自己的特征图。这样，随着模型深度加深，“集体知识”中既包含浅层的形状、纹理信息，又包含更深层的语义信息。小麦锈病由于病原物不同而使得病斑的形状、纹理特征存在差异，条锈病锈斑成行分布，叶锈病锈斑无规则分布。在 DMSA block 中较深的语义信息和浅层细节信息充分融合，且 MSTC layer 用 3 条支路增强感受野、提升模型语义特征强度。所以 DMSA block 能够更好地提取出条锈病与叶锈病的纹理、形状信息特征差异，从而解决了两种小麦锈病相似的难题。

DMSA block 通过特征重用和特征传递获得很高的参数利用率。为进一步实现轻量化，本文通过控制超参数 k，只向模型的“集体知识”中添加一小部分特征图，使得特征图具有非常窄的通道，减少了模型参数数量。此外，在 MSTC layer 中每个卷积层相对较窄，如 1×1 层和 3×3 层，学习到的特征映射数量有限，从而减少冗余，解决模型参数多的难题。

DMSA block 插入 CA 协调注意力机制，它将位置信息嵌入到通道注意力机制中，从而使模型能够有效提取图像中病斑的特征信息。常见的 SEnet 注意力机制[20]只考虑编码通道之间的信息，而忽略了位置信息。CBAM[21]（Convolutional block attention module）注意力机制只能捕捉局部关系，无法解决视觉任务中的远程依赖关系。CA 与它们的不同之处在于它能够捕捉具有精确位置信息的远程空间交互。CA 模块结构如图 5 所示。

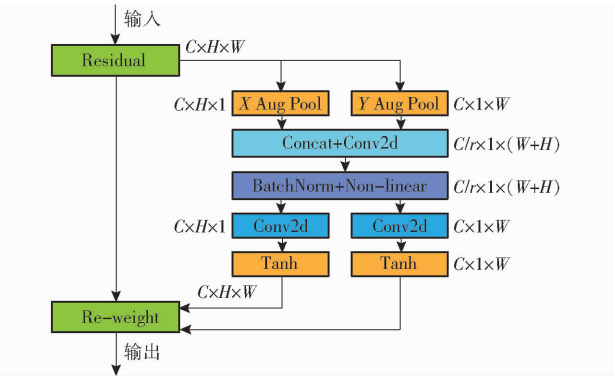


图 5　CA 模块结构图

Fig. 5　Structure diagram of CA model

1.2.3　MSTC layer

小麦叶部锈病图像由于拍摄角度不同、锈病严重程度不同，导致图像病斑大小不一致。受 GoogLeNet[22]和 Inceptionnet[23]的启发，本文设计了一个多尺度 3 路卷积层（MSTC layer）结构来获得不同尺度的感受野，从而提高模型表达能力和模型对不同大小病斑的感知能力。如图 6 所示，MSTC layer 采用 3 个分支来增加模型的深度和宽度，提高模型表达能力。在每一次 3×3 卷积之前引入一个 1×1 卷积，通过控制卷积核维度，以减少 feature-maps 的数量，从而减小参数并提高效率，实现跨通道的交互和信息整合。图 6 中 $2k$，$k/3$ 表示 filter 的个数，超参数 k 是模型增长率，通过控制超参数 k 获得比较窄的特征图，实现轻量化。一个分支使用一个 1×1 卷积和 3×3 卷积进行串联；一个分支中使用了两个 3×3 的卷积，这个分支感受野为 5×5；另一个分支使用了 3×3 的卷积，这个分支的感受野是 7×7；第 4 个分支表示密集连接。最后拼接操作意味着不同尺度特征的融合，图 6 能够同时兼顾小尺寸病害和大尺寸病害。

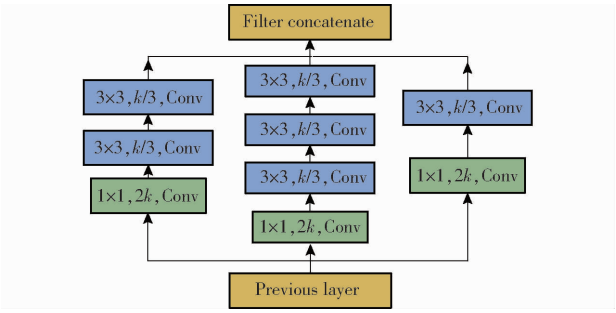


图 6　MSTC layer 结构图

Fig. 6　Structure diagram of MSTC layer

大尺寸卷积核可以带来更大的感受野，但也意味着更多的参数，比如 5×5 卷积核参数是 3×3 卷积核的 25/9 = 2.78 倍。为此，在第 2 个分支中用 2 个连续的 3×3 卷积层（stride 为 1）来代替单个的 5×5 卷积层，在第 3 个分支中 3 个连续的 3×3 卷积层来代替单个的 7×7 卷积层，这样在保持感受野范围的同时又减少参数量。

1.2.4　FSB

受 Inception－V4[23]、DSOD[24]和 Pelee[25]等的启发，本文在第 1 个 DMSA block 之前设计了一个快速降低特征图分辨率的模块 FSB，实现输入图像空间维度的第 1 次下采样和通道数的增加。ResNet[26]和 DenseNet[19]在前两层都是用一个 7×7、stride 为 2 的卷积层和一个 3×3、stride 为 2 的池化层，浅层网络的作用是提取图像的边缘、纹理等信息。FSB 模块的设计就是用比较小的代价来取代 7×7 卷积和池化层。FSB 模块结构如图 7 所示。

在图 7 中，输入图像 X 的尺寸为 $h\times w\times c$，首先使用一个 3×3 卷积进行快速降维，然后使用两个分

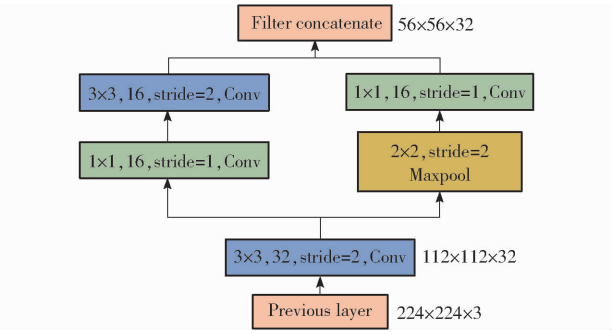


图 7 FSB 结构图

Fig. 7 Structure diagram of FSB

支的结构。一个分支先使用 1×1 卷积来降低特征图的通道数,以此来减少参数量,再用 3×3 卷积在不改变通道数量的前提下提取特征。可以表示为

$$H(X) = T_{nl}\{T_{nl}[T_{nl}(X)]\} \tag{7}$$

式中 $H(X)$——第 2 个 3×3 卷积降低分辨率得到的特征图

T_{nl}——激活函数带来的非线性变换

另一个分支先使用一个 maxpooling,后使用 1×1 卷积降维。目的是通过最大池化先将输入中的部分信息进行传递,保证最终的结果在减少参数量的同时仍具备足够的语义信息,不会造成信息的过度损失。可以表示为

$$P(X) = T_{nl}\{T_l[T_{nl}(X)]\} \tag{8}$$

式中 $P(X)$——经过第 2 个分支得到的特征图,保留了较高分辨率信息

T_l——不使用非线性激活函数的特征变换

最后将两个分支的输出特征图在通道维度进行 concatenate。这种将 3×3 卷积和最大池化进行分支的融合策略可以丰富特征层。此外,结构中 1×1 卷积层的添加也能增加非线性拟合能力。$H(X)$ 和 $P(X)$ 级联后得到融合特征的特征图 Y,可以表示为

$$Y = \mathrm{concat}[H(X); P(X)] \tag{9}$$

特征图 Y 尺寸为 $h/4 \times w/4 \times 2k$。由于使用小的 3×3 卷积来代替大的 7×7 卷积,并且使用与池化层并行连接,所设计的 FSB 会减少部分计算量。

2 实验结果

2.1 实验配置

实验使用 Python 作为编程语言,开源的 pytorch 作为深度学习框架,并使用 Anaconda3 构建的虚拟开发环境进行模型的训练和测试。CNN 参数设置如下:初始学习率为 0.001,如果验证集的损失值在 8 个周期后没有降低,则学习率降低到初始值的 1/10,批量大小为 8,交叉熵(Cross entropy)作为损失函数,采用流行的 Adam 优化器进行优化。

2.2 评价指标

选择准确率(Accuracy)、精度(Precision)、F1 值(F1 score)和召回率(Recall)作为评价指标,评估所提出 Mobile-DMSANet 模型性能。

2.3 消融实验

在小麦叶部锈病数据集上对 Mobile-DMSANet 模型中的 CA 模块、MSTC layer、FSB 分别进行消融实验,结果如表 3 所示。将 MSTC layer 改为仅由一个 1×1 卷积和 3×3 卷积组成一条支路的卷积层,不使用 CA 模块和 FSB,模型识别准确率为 91.3%。加入 CA 模块后,在增加少许参数量的条件下,模型识别准确率达到 92.1%。只加入 MSTC layer 后,模型参数量下降,识别准确率显著提高,达到 94.6%,条锈病和叶锈病的 F1 值提升。同时使用 CA 模块和 MSTC layer,模型能更好地抽取小麦锈病的特征,模型识别准确率达到 95.7%。同时使用 MSTC layer 和 FSB,模型在不增加参数量的同时,提高了准确率。加入 MSTC layer、CA 模块、FSB 后,该模型识别准确率达到 96.4%,比基准模型提高 5.1 个百分点,且参数量下降至 4.54×10^5。

表 3 消融实验结果

Tab. 3 Results of ablation experiments

CA	MSTC	FSB	准确率/%	F1 值/%			参数量
				健康	条锈病	叶锈病	
×	×	×	91.3	93.7	91.1	90.5	6.73×10^5
√	×	×	92.1	95.3	92.4	90.1	7.44×10^5
×	√	×	94.6	98.2	94.0	93.5	3.91×10^5
√	√	×	95.7	98.2	95.2	94.9	4.65×10^5
×	√	√	94.9	98.2	94.4	93.9	3.88×10^5
√	√	√	96.4	98.2	96.1	95.9	4.54×10^5

注:√表示使用该模块,×表示未使用该模块。下同。

为进一步验证 MSTC layer 的有效性,将 MSTC layer 添加到其他骨干网络模型中,实验结果如表 4 所示。骨干网络模型选择 Densenet 201 和 Densenet_BC_100。在 Densenet 201 中加入 MSTC layer 后,模型识别准确率提高 4.7 个百分点,参数量下降 66%。在 Densenet_BC_100 中加入 MSTC layer 后,模型识别准确率提高 1.4 个百分点,参数量也下降接近 50%。参数显著下降是因为 MSTC layer 在 3 条支路中通过控制超参数使得特征图的通道数更少。MSTC layer 融合不同尺度特征,增强感受野,从而提高了模型准确率。

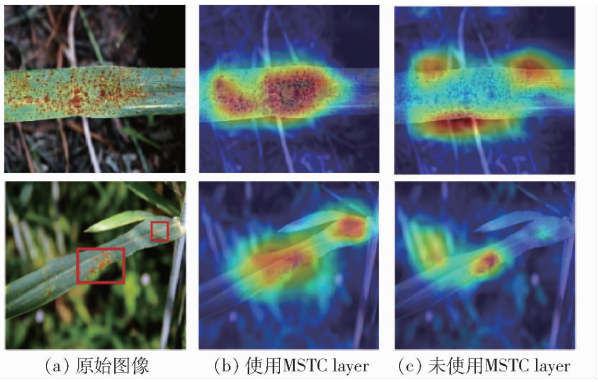
MSTC layer 通过获得不同尺度的感受野,提高对不同大小锈病的感知能力。使用 Grad-CAM 显示图像经过和不经过 MSTC layer 的可视化对比,结果如图8所示。从图8中可以看出,对于大尺寸病

表 4　不同骨干网络添加 MSTC layer 的识别结果

Tab. 4　Recognition results from adding MSTC layer to different baseline networks

基线网络	MSTC	准确率/%	F1 值/%			参数量
			健康	条锈病	叶锈病	
Densenet 201	×	84.2	94.4	84.2	78.5	1.803×10^7
Densenet 201	√	88.9	91.9	89.3	86.7	5.690×10^6
Densenet_BC_100	×	90.7	94.3	90.9	88.4	2.963×10^6
Densenet_BC_100	√	92.1	95.2	92.6	89.8	1.688×10^6
Mobile－DMSANet	√	96.4	98.2	96.1	95.9	4.540×10^5

斑，若不使用 MSTC layer，会因为感受野不足而只关注局部信息，比如边缘、纹理等。对于背景复杂的小尺寸病斑，感受野过大，容易被视为背景而无法被识别。所以，MSTC layer 能够准确地聚焦于小麦叶片的不同尺寸病斑，从而提高锈病识别准确率。

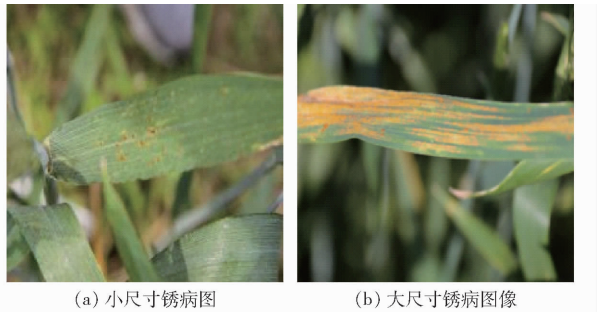


(a) 原始图像　(b) 使用MSTC layer　(c) 未使用MSTC layer

图 8　MSTC layer 的可视化对比结果

Fig. 8　Visualization comparison results of MSTC layer

将不使用 MSTC layer 的卷积层输出的特征图和使用 MSTC layer 输出的特征图进行可视化对比，小尺寸锈病图像和大尺寸锈病图像如图 9 所示。部分结果如图 10(小尺寸锈病特征图)、图 11(大尺寸锈病特征图)所示。图中 MSTC layer 分别是 DMSA block 1 中第 3 个 MSTC layer、DMSA block 2 中第 3 个 MSTC layer、MSTC block 3 中第 1 个 MSTC layer、MSTC block 3 中第 3 个 MSTC layer。如图 10 所示，模型在较浅层提取图像的纹理、边缘、颜色等信息。随着卷积层数加深，特征图中图像视觉信息减少，抽象信息增加。可以看出，在增加 MSTC layer 后，感兴趣区域(图中黄色越明亮，感兴趣程度越深)更加集中在病斑位置。所以，MSTC layer 能够提高对不同大小病斑感知能力，使得感兴趣区域特征更加突出。



(a) 小尺寸锈病图　(b) 大尺寸锈病图像

图 9　不同尺寸锈病图像示例

Fig. 9　Examples of rust disease images of different sizes

2.4　不同注意力机制对比实验

将 Mobile－DMSANet 模型注意力去除的网络作为基线模型(baseline)，然后分别加入 SE、CBAM 以及 CA 注意力进行对比实验，识别结果如表 5 所示。从表 5 可以看出，采用 CA 模块构建的模型识别准确率和 F1 值均高于采用 SE 模块和 CBAM 模块的模型。结果表明，引入 CA 注意力模块可以使模型更好地找到图像中的病变区域，减少背景对识别结果的影响。

使用 Grad－CAM 显示图像得到 SE 模块、CBAM 模块和 CA 模块的可视化结果如图 12 所示。样本图像 1 中，SE 模块和 CBAM 模块关注到了背景区域导致识别结果错误。相比之下，CA 模块专注于感兴趣的区域，模型给出了正确的识别结果。样本图像 2 中，CBAM 模块虽关注到部分病斑，但囊括了很大一部分背景区域。样本图像 3 中，不管是 SE 模块还是 CBAM 模块，虽都关注到了病斑，但并不完整。结果表明，CA 注意力模块可以更好地关注到图像中的病斑区域，减少背景对识别效果的影响。

2.5　对比实验

表 6 为 Mobile－DMSANet 模型与经典 CNN 模型(AlexNet 和 ResNet50)和轻量级 CNN 模型(ShufflenetV2 和 Densenet_BC_100)在小麦叶部锈病数据集上对比实验的结果。从整体的识别率来看，本文模型的准确率达到 96.4%，大幅优于其他 CNN。较 ResNet50 的准确率提升 15.8 个百分点，较轻量级模型 DenseNet121 和 ShufflenetV2[27] 的准确率分别提升 6.5 个百分点和 6.8 个百分点。3 个类别的 F1 值均是 Mobile－DMSANet 模型最高。从参数量来看，Mobile－DMSANet 模型参数量只有 4.54×10^5，大幅度小于其他网络模型。

表 7 为 Mobile－DMSANet 模型与其他轻量型作物病害识别文献中提出的模型在本文小麦叶部锈病数据集上进行对比实验的结果。其中 MS－Dnet 是 CHEN 等[14]提出的一种采用深度可分离卷积并用于作物病害识别的轻量级网络模型。MS－Dnet 模型参数量较少，但该模型采用深度可分离卷积层提取特征时，通道之间的信息没有交互，存在信息的屏蔽和阻塞，特别是在涉及全局信息以及复杂背景的情况下，存在一定的局限性。CA_Densenet_BC_40 模型是 BAO 等[28]提出的一种用于田间自然条件下

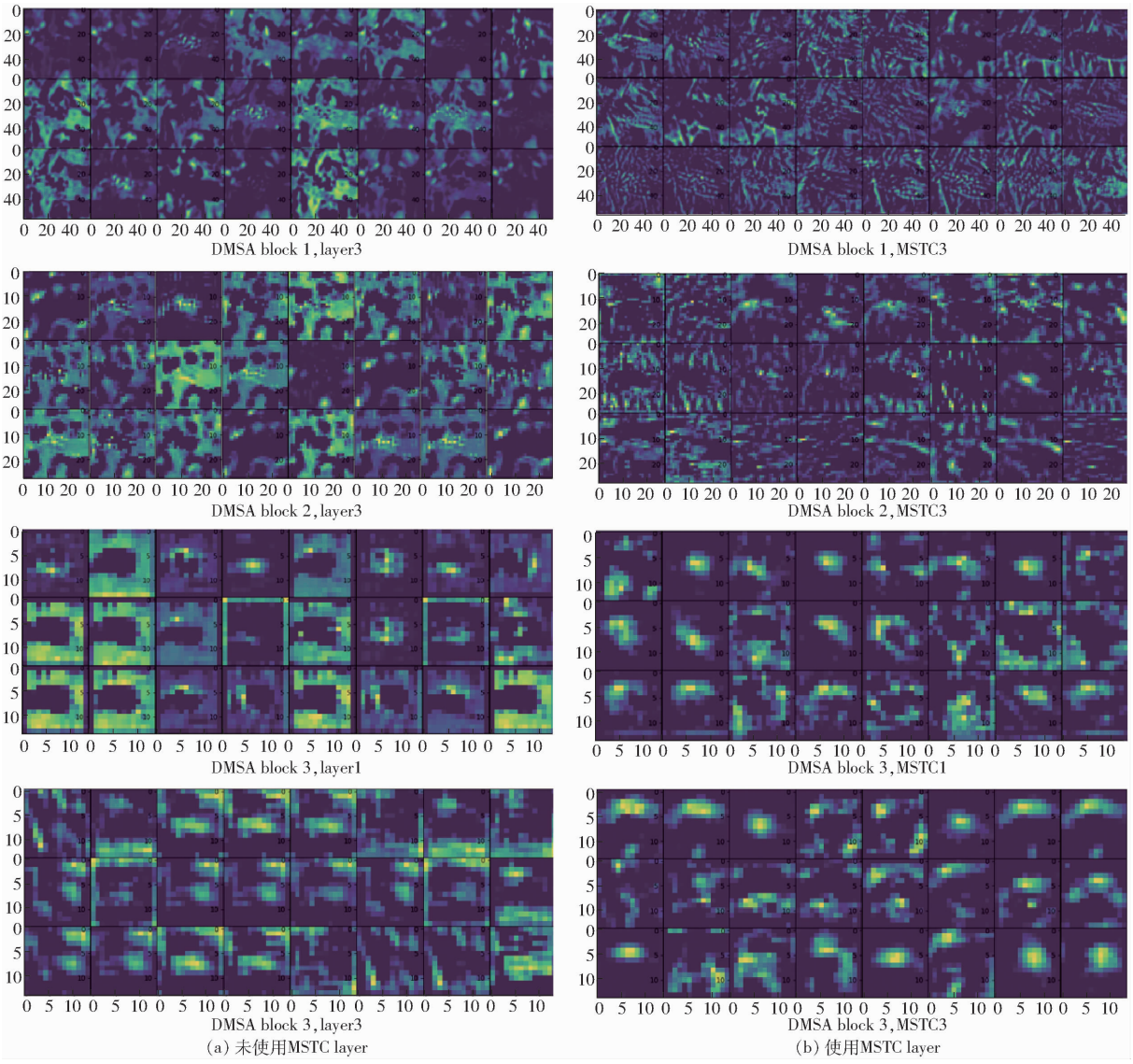


图 10 小尺寸锈病特征图的可视化结果

Fig. 10 Characterization maps for small area of disease

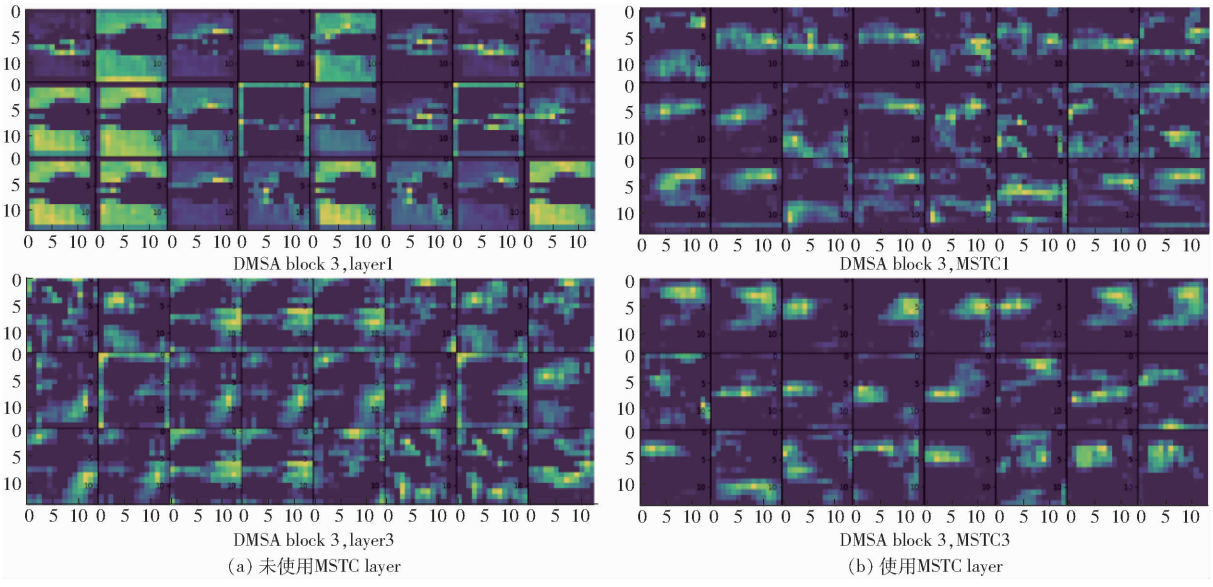


图 11 大尺寸锈病特征图的可视化结果

Fig. 11 Characterization maps for large area of disease

表 5　不同注意力机制的实验结果

Tab. 5　Experimental results from different attention mechanisms　%

模型	准确率	精确率	召回率	F1 值
baseline	94.9	95.1	95.9	95.5
+ SE	95.3	95.4	96.2	95.8
+ CBAM	95.3	95.4	96.0	95.6
+ CA	96.4	96.5	97.0	96.7

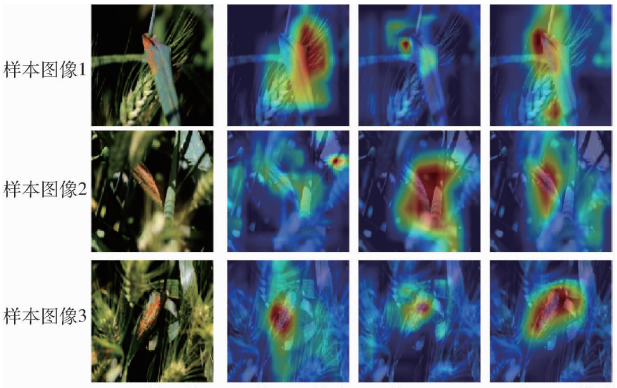


图 12　叠加不同注意力机制的可视化结果

Fig. 12　Visualization results for different attention mechanisms

表 6　本文模型与其他卷积神经网络模型对比结果

Tab. 6　Comparison of proposed model with other networks

模型	类别	精确率/%	召回率/%	F1 值/%	准确率/%	参数量
AlexNet	健康	94.3	92.6	93.5	76.3	$5.701\,3\times10^7$
	条锈病	72.4	82.7	77.2		
	叶锈病	71.6	59.2	64.8		
ResNet50	健康	86.0	90.7	88.3	80.6	$2.350\,9\times10^7$
	条锈病	77.9	85.8	81.6		
	叶锈病	81.7	68.4	74.4		
ShufflenetV2	健康	96.2	94.4	95.3	89.6	$5.354\,0\times10^6$
	条锈病	91.7	87.4	89.5		
	叶锈病	83.8	89.8	86.7		
DenseNet121	健康	96.4	98.1	97.2	89.9	$6.961\,0\times10^6$
	条锈病	91.1	88.2	89.6		
	叶锈病	85.1	97.8	86.4		
Densenet_BC_100	健康	96.2	92.6	94.3	90.7	$2.969\,0\times10^6$
	条锈病	91.3	90.6	90.9		
	叶锈病	87.1	89.8	88.4		
Mobile－DMSANet	健康	96.4	100	98.2	96.4	$4.540\,0\times10^5$
	条锈病	96.1	96.1	96.1		
	叶锈病	96.9	94.9	95.9		

棉蚜危害程度的自动检测和分类模型。该模型中注意力机制能有效抑制背景信息，但将其用于小麦叶部病害识别时不能很好地适应不同大小的病斑。轻量级模型 MoblienetV2[29] 已广泛应用于作物病害识别，该模型的倒残差结构提高了梯度的跨层传播能力，防止梯度的消失。但是值得注意的是在下采样过程造成对病害信息的部分丢失，特别是对于特征差异小的数据集。田间小麦叶部锈病识别存在背景复杂、特征相似和病斑大小不一致等难点问题，本文提出的模型通过密集连接实现特征重用，通过 MSTC layer 获得不同大小的感受野，从而在参数量较少的情况下提高了准确率。由表 7 可知，Mobile－DMSANet 准确率较 MS－Dnet 提升 4.6 个百分点，较 CA_Densenet_BC_40 提升 4.3 个百分点，较 MoblienetV2 提升 7.9 个百分点。3 个类别 F1 值均是 Mobile－DMSANet 最高。

表 7　本文模型与其他文献结果对比

Tab. 7　Comparison of proposed model with other literatures

模型	类别	精确率/%	召回率/%	F1 值/%	准确率/%	参数量
MS－Dnet[14]	健康	94.5	96.3	95.4	91.8	6.030×10^5
	条锈病	91.3	90.6	90.9		
	叶锈病	90.8	90.8	90.8		
CA_Densenet_BC_40[28]	健康	96.2	94.4	95.3	92.1	7.440×10^5
	条锈病	94.3	90.6	92.4		
	叶锈病	87.5	92.9	90.1		
MobilenetV2[29]	健康	96.1	90.7	93.3	88.5	2.229×10^6
	条锈病	88.2	88.2	88.2		
	叶锈病	85.1	87.8	86.4		
Mobile－DMSANet	健康	96.4	100	98.2	96.4	4.540×10^5
	条锈病	96.1	96.1	96.1		
	叶锈病	96.9	94.9	95.9		

不同模型在测试集上的混淆矩阵如图 13 所示，ROC 曲线如图 14 所示。从图 13 可以看出，Mobile－DMSANet 模型在条锈病和叶锈病上的识别错误率大大下降，这也验证了 Mobile－DMSANet 模型对小麦叶部锈病有很好的识别效果。ROC 曲线中，曲线越靠近左上角，表明分类效果越好，从图 14 可以看出 Mobile－DMSANet 模型性能优于其他模型。

3　结束语

以田间自然环境下的小麦叶部锈病为对象构建数据集。为减少模型参数，提高锈病识别率，并解决背景复杂、两种锈病相似、病斑大小不一等问题，设计了一个轻量级密集多尺度注意力网络模型 Mobile－DMSANet，用于小麦叶部锈病的自动识别。首先，采用单尺度 Retinex 算法对图像进行光照预处理，以减少光照不均匀对识别结果的影响。Mobile－DMSANet 主要包含输入层、特征提取层和输出层。在输入层设计了一个 FSB 模块，该模块使用卷积和池化融合的双分支结构取代了传统下采样方式，在减少计算量的同时，提高了模型特征表达能力。在特征提取层设计一个 DMSA block，它由 6 个 MSTC layer 密集连接，再插入 CA 协调注意力机制组成。

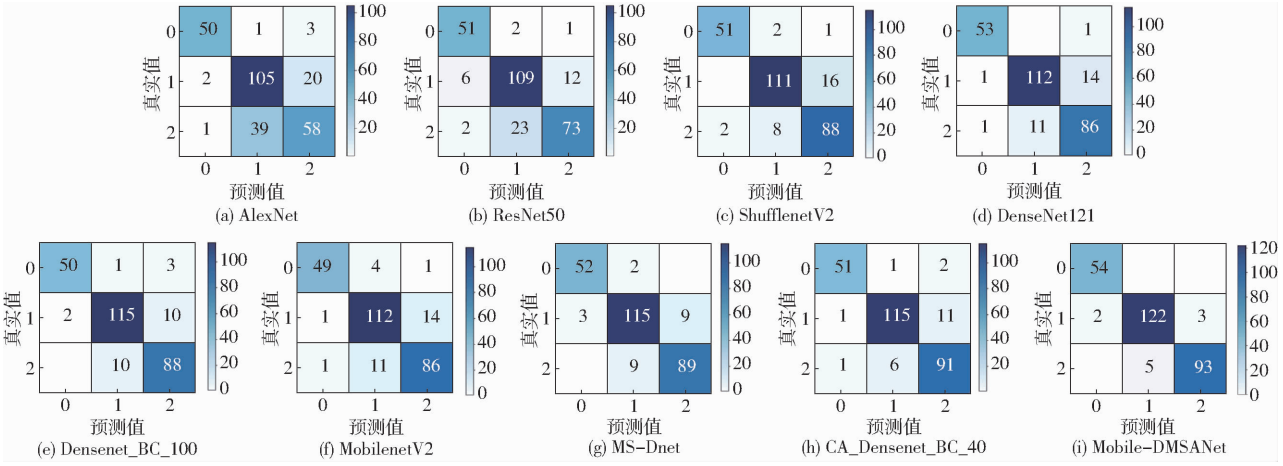


图 13 不同模型在测试集上的混淆矩阵

Fig. 13 Confusion matrices for different models on test set

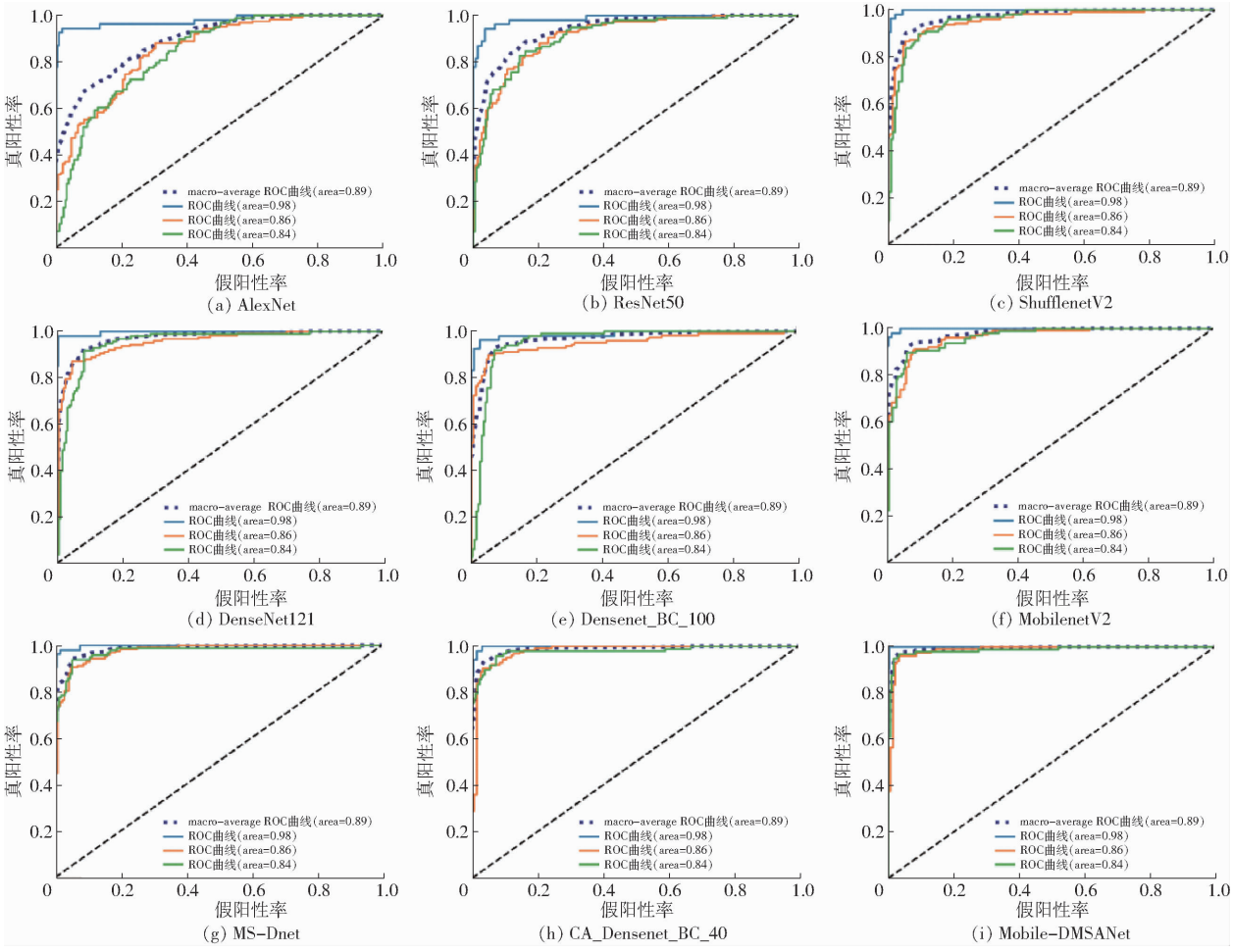


图 14 不同模型在测试集上的 ROC 曲线

Fig. 14 ROC curves for different models on test set

MSTC layer 3 条支路的设计提高了模型对不同尺寸锈病的感知能力。各个 MSTC layer 彼此连接,加强了整个模型的信息流和梯度传播,缓解了过拟合,并能够很好地提取两种相似锈病的特征。CA 模块的引入使得模型更好地关注病害信息,抑制复杂图像中的背景信息。在输出层使用 Softmax 分类器实现小麦叶部锈病识别。实验结果表明,所提出的 Mobile-MSANet 模型在田间自然环境下的小麦叶部锈病识别任务中有着较好性能,准确率达到 96.4%,优于其他模型。Mobile-MSANet 模型参数量只有 4.54×10^5,显著少于其他模型。Mobile-DMSANet 在准确率和参数数量之间取得了很好的平衡,该模型可以移植到移动设备上,以准确识别小麦叶部锈病。

参 考 文 献

[1] ZHANG P, YAN X, GEBREWAHID T W, et al. Genome-wide association mapping of leaf rust and stripe rust resistance in wheat accessions using the 90K SNP array[J]. Theoretical and Applied Genetics,2021,134(4): 1233 – 1251.

[2] 鲍文霞, 林泽, 胡根生, 等. 基于 RSTCNN 的小麦叶片病害严重度估计[J]. 农业机械学报,2021,52(12):242 – 252.
BAO Wenxia, LIN Ze, HU Gensheng, et al. Severity estimation of wheat leaf diseases based on RSTCNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(12):242 – 252. (in Chinese)

[3] 鲍文霞, 赵健, 张东彦, 等. 基于椭圆型度量学习的小麦叶部病害识别[J]. 农业机械学报,2018,49(12):20 – 26.
BAO Wenxia, ZHAO Jian, ZHANG Dongyan, et al. Recognition of wheat leaf diseases based on elliptic metric learning[J]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(12):20 – 26. (in Chinese)

[4] HAYIT T, ERBAY H, VARN F, et al. The classification of wheat yellow rust disease based on a combination of textural and deep features[J]. Multimedia Tools and Applications,2023,23:1 – 19.

[5] HEIDARIAN DEHKORDI R, EL JARROUDI M, KOUADIO L, et al. Monitoring wheat leaf rust and stripe rust in winter wheat using high-resolution UAV-based red-green-blue imagery[J]. Remote Sensing,2020,12(22):3696.

[6] WANG H L, JIANG Q, SUN Z Y, et al. Identification of stripe rust and leaf rust on different wheat varieties based on image processing technology[J]. Agronomy,2023,13(1):260.

[7] BAO W X, YANG X H, LIANG D, et al. Lightweight convolutional neural network model for field wheat ear disease identification[J]. Computers and Electronics in Agriculture,2021,189:106367.

[8] MOSHOU D, BRAVO C, WEST J, et al. Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks[J]. Computers and Electronics in Agriculture,2004,44(3):173 – 188.

[9] TIAN Y, ZHAO C J, LU S L, et al. Multiple classifier combination for recognition of wheat leaf diseases[J]. Intelligent Automation and Soft Computing,2011,17(5):519 – 529.

[10] XIONG J W, MIN W. Identification of cash crop diseases using automatic image segmentation algorithm and deep learning with expanded dataset[J]. Computers and Electronics in Agriculture,2020,177(1):105712.

[11] GAO G, LI Z X, HUAN Z, et al. Human behavior recognition model based on feature and classifier selection[J]. Sensors, 2021,21(23):7791.

[12] LIN J W, CHEN Y, PAN R Y, et al. CAMFFNet: a novel convolutional neural network model for tobacco disease image recognition[J]. Computers and Electronics in Agriculture,2022,202:107390.

[13] CHEN J, WANG W, ZHANG D, et al. Attention embedded lightweight network for maize disease recognition[J]. Plant Pathology,2020,70(3):630 – 642.

[14] CHEN W, CHEN J, DUAN R, et al. MS – DNet: a mobile neural network for plant disease identification[J]. Computers and Electronics in Agriculture,2022,199:107175.

[15] 冀汶莉, 刘洲, 邢海花. 基于 YOLO v5 的农田杂草识别轻量化方法研究[J]. 农业机械学报,2024,55(1):212 – 222,293.
JI Wenli, LIU Zhou, XING Haihua. Lightweight method for identifying farmland weeds based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(1):212 – 222,293. (in Chinese)

[16] XU Fenglin, MIAO Yubin, ZHANG Ming. Navigation image enhancement based on color weighted guided image filtering-retinex algorithm[J]. Journal of Shanghai Jiaotong University,2019,53(8):921.

[17] SUN J, YANG Y, HE X, et al. Northern maize leaf blight detection under complex field environment based on deep learning [J]. IEEE Access,2020,8:33679 – 33688.

[18] 左昊轩, 黄祺成, 杨佳昊. 基于改进 YOLO v5s 的作物黄化曲叶病检测方法[J]. 农业机械学报,2023,54(增刊 1): 230 – 238.
ZUO Haoxuan, HUANG Qicheng, YANG Jiahao. Improved YOLO v5s-based detection method for crop yellow leaf curl virus disease[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(Supp. 1):230 – 238. (in Chinese)

[19] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),2017:2261 – 2269.

[20] HU J, SHEN L, ALBANIE S, et al. squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(8):2011 – 2023.

[21] WOO S H, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// 15th European Conference on Computer Vision (ECCV),2018:3 – 19.

[22] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2015:1 – 9.

[23] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception – v4, Inception – ResNet and the impact of residual connections on learning[C]// 31st AAAI Conference on Artificial Intelligence,2016:4278 – 4284.

[24] SHEN Z Q, LIU Z, LI J G, et al. DSOD: learning deeply supervised object detectors from scratch[C]// 16th IEEE International Conference on Computer Vision (ICCV),2017:1937 – 1945.

[25] WANG R J, LI X, LING C X. Pelee: a real-time object detection system on mobile devices[C]// 32nd Conference on Neural Information Processing Systems (NIPS),2018:31.

[26] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016:770 – 778.

[27] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]// 15th European Conference on Computer Vision (ECCV),2018:122 – 138.

[28] BAO W X, CHENG T, ZHOU X G, et al. An improved DenseNet model to classify the damage caused by cotton aphid[J]. Computers and Electronics in Agriculture,2022,203:107485.

[29] WESSNER R N, FROZZA R, DA SILVA BAGATINI D, et al. Recognition of weeds in corn crops: system with convolutional neural networks[J]. Journal of Agriculture and Food Research,2023,14:100669.