doi:10.6041/j.issn.1000-1298.2022.08.036

基于颜色掩膜网络和自注意力机制的叶片病害识别方法

于　明[1]　李若曦[2]　阎　刚[1]　王　岩[3]　王建春[4]　李　扬[4]

(1. 河北工业大学人工智能学院, 天津 300401; 2. 河北工业大学电子信息工程学院, 天津 300401;
3. 天津商业大学信息工程学院, 天津 300401; 4. 天津市农业科学院, 天津 300190)

摘要: 为了提取到更加准确、丰富的叶片病斑的颜色特征和空间特征,解决病害严重程度细粒度分类粗糙、识别准确率低等问题,提出一种融合颜色掩膜网络和自注意力机制(Fusion color mask and self-attention network, FCMSAN)的病害识别方法。FCMSAN 由颜色掩膜网络(Color mask network, CMN)和融合通道自适应的自注意力网络(Channel adaptive self-attention network, CASAN)构成。CMN 通过学习叶片病斑颜色区域信息提高模型提取颜色特征的能力;CASAN 能够提取全局范围内的病斑特征,同时加入病斑的位置特征和通道自适应特征,可以精确、全面定位叶片病斑区域。最后通过特征转换融合模块(Transfer fusion layer, TFL)将 CMN 和 CASAN 进行融合。经实验证明,FCMSAN 在 61 类农作物病虫害细粒度识别中,Top-1 的分类准确率达到 87.97%,平均 F1 值达到 84.48%。最后通过可视化分析,验证了本文方法在病害识别中的有效性。

关键词: 农作物病害; 自注意力机制; 细粒度识别; 特征提取; 颜色掩膜网络

中图分类号: TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2022)08-0337-08　　**OSID**:

Crop Diseases Recognition Method via Fusion Color Mask and Self-attention Mechanism

YU Ming[1]　LI Ruoxi[2]　YAN Gang[1]　WANG Yan[3]　WANG Jianchun[4]　LI Yang[4]

(1. *School of Artificial Intelligence*, *Hebei University of Technology*, *Tianjin* 300401, *China*
2. *School of Information Engineering*, *Hebei University of Technology*, *Tianjin* 300401, *China*
3. *College of Information Engineering*, *Tianjin University of Commerce*, *Tianjin* 300401, *China*
4. *Tianjin Academy of Agricultural Sciences*, *Tianjin* 300190, *China*)

Abstract: To reduce the loss of crop diseases, a large number of chemicals are used for disease control. However, due to the untimely and inaccurate judgment of the disease, chemical agents are abused, which also has a great impact on the ecological environment and food safety. Therefore, it is urgent to develop accurate crop disease recognition based on images. In the image recognition of crop diseases, the shape, region, and color of leaf spots are the main indexes to distinguish different types of diseases. In order to extract more accurate and rich color features and spatial features of leaf spots, and solve the problems of coarse fine-grained classification of disease severity and low recognition accuracy, a fusion color mask and self-attention network (FCMSAN) was proposed. FCMSAN was composed of a color mask network (CMN) and a channel adaptive self-attention network (CASAN). CMN can improve the ability of color feature extraction by learning the color features of leaf spots. CASAN can extract the disease spot features in the global scope and add the location features and channel adaptive features of the disease spot, which can locate the leaf disease spot area accurately and comprehensively. Finally, the outputs of CMN and CASAN were fused by transfer fusion layer (TFL). The experimental results showed that the classification Top-1 accuracy reached 87.97% and the average F1 value can reach 84.48% in the fine-grained identification of 61 types of crop diseases. Finally, the effectiveness of the proposed method for crop disease recognition was verified by visualization experiments.

Key words: crop diseases; self-attention mechanism; fine-grained identification; feature extraction; color mask network

收稿日期: 2021-08-28　修回日期: 2021-10-25

基金项目: 国家自然科学基金重大研究计划项目(91746207)和国家重点研发计划项目(2018YFC08)

作者简介: 于明(1964—),男,教授,博士生导师,主要从事图像处理与模式识别研究,E-mail: yuming@hebut.edu.cn

0 引言

为了减轻农作物病害的损失,大量化学药物被用于病害防治[1]。但由于对病害的判断不及时且不准确,导致了化学药剂的滥用,也对生态环境和食品安全造成了巨大影响。因此,开展精准的基于图形图像的农作物病害识别刻不容缓。

传统方法通过机器学习相关模型手动提取病斑区域的颜色、形状、纹理、边缘等特征后进行分类[2-5]。但传统机器学习模型需要针对不同作物的病害人工设计特征提取方法,不仅设计的过程复杂,还容易受到光线、拍摄角度、背景等因素的干扰,而且模型的泛化能力弱,难以迁移到其他类型的识别任务中去。

为解决以上问题,研究人员借助了深度学习方法[6-9]。目前病害识别主要基于 AlexNet[10]、VGG16[11]、GoogLeNet[12]、ResNet50[13]等模型。宋余庆等[14]为了解决各种病害区域大小不一的问题,提出多层次增强高效空间金字塔卷积深度学习模型,提取病害图像高层语义特征。由于农作物病害表征多样,蒲秀夫等[15]采用多尺度卷积核提取特征,解决了特征多样性不足的问题。

注意力机制能够捕捉到关键信息,近年来,已被成功应用于病害识别任务中。金晓芙[16]提出 Inception - v3 + Att + Salie 方法,通过在网络中加入注意力机制可以更好关注细节,在病害的细粒度分类上取得更好的结果。王美华等[17]提出 I_CBAM(Improved convolutional block attention module)将 CBAM[18]中两个注意力模块的串行连接方式改为并行连接方式,来解决 CBAM 在病害识别中表现不稳定等问题。自注意力机制是对注意力机制的改进,减少了模型对外部信息的依赖,更加关注捕获特征的内在联系,针对特征间的相关性建模,在全局范围内根据特征的重要程度分配权重,从而有效提高模型效率。目前,自注意力机制主要有多头自注意力(MHSA)[19],以及 Non-Local[20]、SA-axial[21]、CA[22]、ViT[23]等模型,这些模型提出用不同形式的自注意力机制(局部、全局、矢量、轴向等)来替换主干网络中的部分卷积,在各类视觉任务中取得了较好的结果。但是这些自注意力机制模型主要针对空间信息建模,无法自适应学习来自不同通道间的特征信息。

叶片病斑的颜色特征和形态特征是病害识别的重要依据,因此设计颜色掩膜网络(Color mask network,CMN)针对叶片病斑颜色特征进行学习;然后将所设计的通道自适应自注意力机制嵌入 ResNet50,得到融合通道自适应的自注意力网络(Channel adaptive self-attention network, CASAN)自适应得到病斑的全局、位置和通道上下文信息;最后针对农作物病害识别中病害类别易混淆的问题,将 CMN 和 CASAN 结合为融合颜色掩膜网络和自注意力机制的双流网络 FCMSAN,同时提取病斑的颜色特征、全局空间特征和通道特征,旨在提升农作物病虫害细粒度识别效果。

1 本文方法

本文提出的 FCMSAN 整体框架如图 1 所示。其中 CMN 的网络结构为堆叠的卷积神经网络,可以抑制非病斑颜色区域的特征,提取病斑颜色区域的特征。CASAN 在 ResNet50 中引入所设计的通道自适应自注意力机制,能够建立叶片病害的长距离依赖,同时提取叶片病害的位置特征,从而

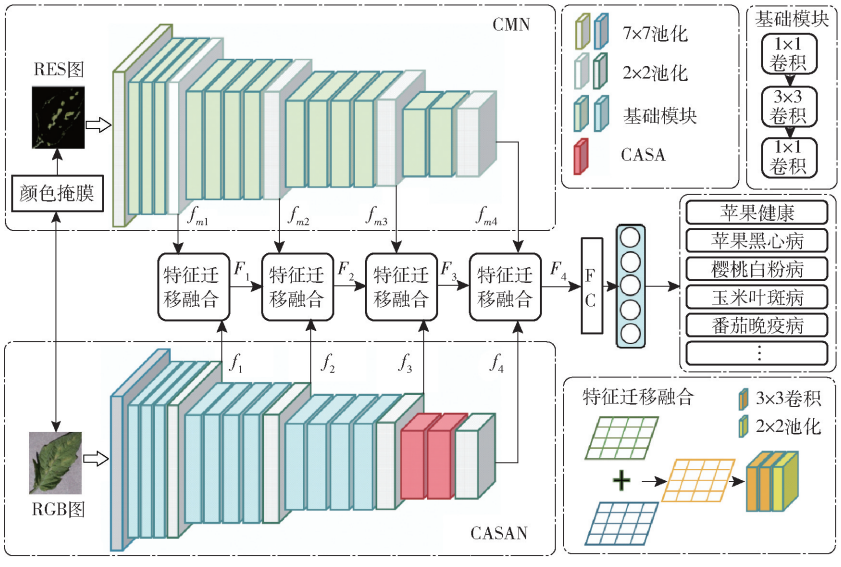


图 1 FCMSAN 整体框架图

Fig. 1 Framework of FCMSAN

获得全局范围内病斑特征的空间信息。另外，在自注意力机制中加入的通道自适应，能够获取重要的通道维度信息。最后，将双流特征输入到特征转换融合模块（TFL）中进行融合，得到预测的病害类别。

1.1　颜色掩模网络

颜色掩模网络 CMN 的输入为构建掩膜提取病斑颜色特征后的残差 RES 图，该过程如图 2 所示。先将原始 RGB 图像转换到 HSV 颜色空间，查得 HSV 颜色空间中表示绿色的数值范围，利用二值化操作将绿色置 0，构建掩膜 Mask 图像

$$M=\begin{cases}0 & (x\in[35,43,46]\cup[77,255,255])\\ 255 & (\text{其他})\end{cases} \tag{1}$$

最后将掩膜 Mask 图像和原始 RGB 图像作像素与操作，得到 RES 图。

经 7×7 卷积核对 RES 图进行卷积，提取局部特征。然后对由基础模块堆叠而成的卷积块进行卷积操作。基础模块每堆叠若干次后，通过 2×2 最大池化层进行下采样，进一步缩小图像的分辨率，得到高层低分辨率的特征。

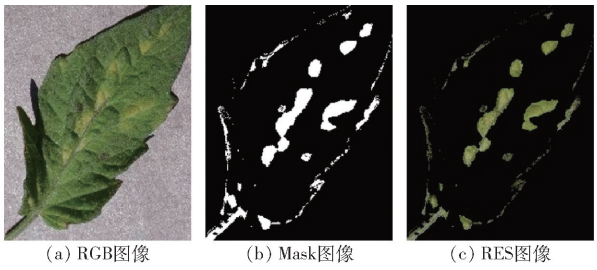


图 2　RES 图像构建过程

Fig. 2　Process of constructing RES

1.2　融合通道自适应的自注意力网络

融合通道自适应的自注意力网络（CASAN）将 CMN 的最后两个堆叠的基础模块中的 3×3 卷积替换为 CASA 模块。CASA 模块的结构如图 3 所示。

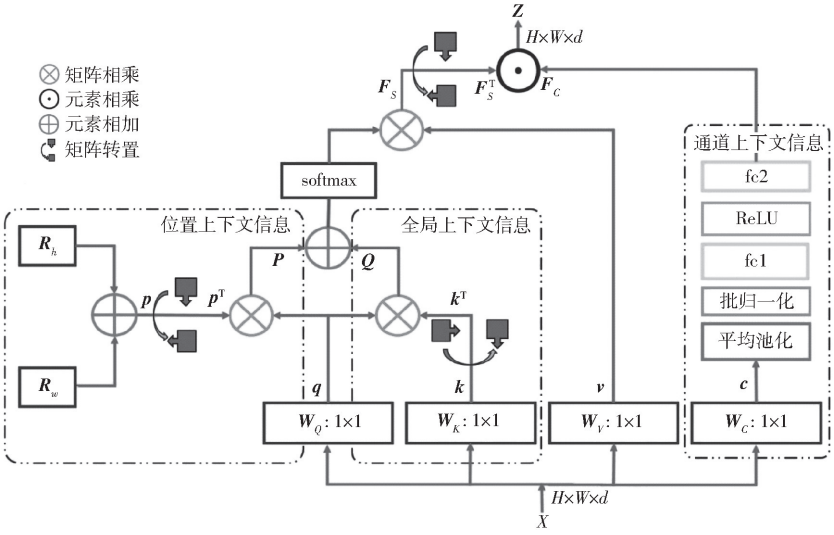


图 3　CASA 结构图

Fig. 3　Structure diagram of CASA

CASAN 提取 3 类特征：全局上下文信息、位置上下文信息和通道上下文信息。给定输入特征 $X\in\mathbf{R}^{H\times W\times d}$，将 X 输入 4 个权重矩阵 $\boldsymbol{W}_Q\in\mathbf{R}^{1\times1\times d}$、$\boldsymbol{W}_K\in\mathbf{R}^{1\times1\times d}$、$\boldsymbol{W}_V\in\mathbf{R}^{1\times1\times d}$、$\boldsymbol{W}_C\in\mathbf{R}^{1\times1\times d}$，分别获得查询矩阵 $\boldsymbol{q}\in\mathbf{R}^{H\times W\times d}$、键矩阵 $\boldsymbol{k}\in\mathbf{R}^{H\times W\times d}$、值矩阵 $\boldsymbol{v}\in\mathbf{R}^{H\times W\times d}$ 和通道矩阵 $\boldsymbol{c}\in\mathbf{R}^{H\times W\times d}$，计算式为

$$\boldsymbol{q}=\boldsymbol{W}_Q(X) \tag{2}$$

$$\boldsymbol{k}=\boldsymbol{W}_K(X) \tag{3}$$

$$\boldsymbol{v}=\boldsymbol{W}_V(X) \tag{4}$$

$$\boldsymbol{c}=\boldsymbol{W}_C(X) \tag{5}$$

1.2.1　全局上下文信息

将查询矩阵 $\boldsymbol{q}\in\mathbf{R}^{H\times W\times d}$ 与矩阵 $\boldsymbol{k}^{\mathrm{T}}\in\mathbf{R}^{d\times H\times W}$ 相乘，得到全局语义信息 $\boldsymbol{Q}$ 为

$$\boldsymbol{Q}=\boldsymbol{q}\otimes\boldsymbol{k}^{\mathrm{T}} \tag{6}$$

可以得到矩阵特征与键矩阵的关系，捕获特征图任意两点间的长程距离依赖，从而获取到全局信息。

1.2.2　位置上下文信息

分别对特征图的高度 H 和宽度 W 归一化，得到 $\boldsymbol{R}_h\in\mathbf{R}^{H\times1\times d}$、$\boldsymbol{R}_w\in\mathbf{R}^{1\times W\times d}$，将 $\boldsymbol{R}_h$ 和 $\boldsymbol{R}_w$ 在元素级相加，得到位置信息 $\boldsymbol{p}$ 为

$$\boldsymbol{p}=\boldsymbol{R}_h\oplus\boldsymbol{R}_w \tag{7}$$

式中　⊕——元素相加运算符

然后将查询矩阵 $\boldsymbol{q}\in\mathbf{R}^{H\times W\times d}$ 和矩阵 $\boldsymbol{p}^{\mathrm{T}}\in\mathbf{R}^{H\times W\times d}$ 相乘，得到位置信息 $\boldsymbol{P}$ 为

$$\boldsymbol{P}=\boldsymbol{q}\otimes\boldsymbol{p}^{\mathrm{T}} \tag{8}$$

式中　⊗——矩阵相乘运算符

该过程通过在空间维度设置可训练参数，获取

在高度 H 和宽度 W 上与特征相关的二维坐标。通过这种方式将提取到的特征赋予相应的位置信息，极大丰富了特征在空间上的语义。将提取到的全局上下文信息 $\boldsymbol{Q}$ 与位置上下文信息 $\boldsymbol{P}$ 相加，通过 softmax 函数后与值矩阵 $\boldsymbol{v}$ 相乘，得到结合全局内容信息与位置信息的空间注意力特征 $\boldsymbol{F}_S \in \mathbf{R}^{H\times W\times d}$，$\boldsymbol{F}_S$ 为

$$\boldsymbol{F}_S = \mathrm{softmax}(\boldsymbol{Q}\oplus\boldsymbol{P})\otimes\boldsymbol{v} \tag{9}$$

1.2.3 通道上下文信息

将通道矩阵 $\boldsymbol{c}$ 输入平均池化层，经批次正则化加快模型的训练速度。通过全连接层 1(fc1)、激活函数 ReLU 和全连接层 2(fc2)进行非线性映射，得到通道注意力特征 $\boldsymbol{F}_C \in \mathbf{R}^{d\times 1\times 1}$。该过程使得模型自适应不同通道特征。最后，得到通道上下文信息 $\boldsymbol{Z}$ 为

$$\boldsymbol{Z} = \boldsymbol{F}_S^{\mathrm{T}} \odot \boldsymbol{F}_C \tag{10}$$

式中 ⊙——元素相乘运算符

1.3 双流特征融合

为增强特征的传播和表达能力，如图 1 所示，将 CMN 的输出 f_{m1}、f_{m2}、f_{m3}、f_{m4} 与 CASAN 的输出 f_1、f_2、f_3、f_4 进行特征融合。将 f_{m1} 与 f_1 输入到 TFL 中，先进行元素相加，再输入到两个卷积核为 3×3、步长为 1、填充为 1 的卷积层，通过步长为 2 的 2×2 最大池化层下采样，得到 F_1。然后将 f_{m2} 与 f_2 输入到 TFL 中进行相同的操作，再与 F_1 相加得到 F_2。与获取 F_2 的过程类似，将 f_{m3} 与 f_3 输入 TFL 中，再加上 F_2 得到 F_3，将 f_{m4} 与 f_4 输入 TFL 中，再加上 F_3 得到 F_4。

最后，将 F_4 输入到全连接层中，得到 61 类叶片病害的分类结果。

2 实验结果与分析

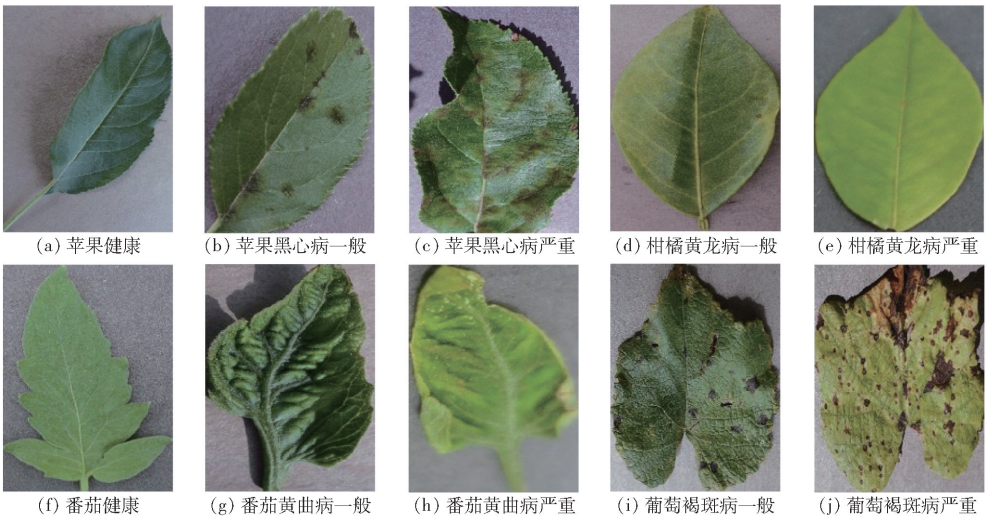
2.1 实验数据集

实验所用的数据集为 AI Challenger 2018[24]，共包含 31 721 幅训练集图像和 4 540 幅验证集图像。有苹果、草莓、樱桃等 10 个物种，36 类不同病害，每种病害分为健康、一般和严重 3 个等级，共 61 类病害，标签为 0～60。表 1 为该数据集部分分类标签。图 4 为数据集中部分病害图像。

表 1 AI Challenger 数据集部分分类标签

Tab. 1 Labels for AI Challenger

标签	分类	标签	分类
0	苹果健康	25	柑橘黄龙病一般
1	苹果黑心病一般	26	柑橘黄龙病严重
2	苹果黑心病严重	41	番茄健康
22	葡萄褐斑病一般	58	番茄黄曲病一般
23	葡萄褐斑病严重	59	番茄黄曲病严重



(a) 苹果健康 (b) 苹果黑心病一般 (c) 苹果黑心病严重 (d) 柑橘黄龙病一般 (e) 柑橘黄龙病严重

(f) 番茄健康 (g) 番茄黄曲病一般 (h) 番茄黄曲病严重 (i) 葡萄褐斑病一般 (j) 葡萄褐斑病严重

图 4 AI Challenger 数据集部分图像示例

Fig. 4 Some samples of AI Challenger

2.2 实验环境与参数设置

服务器操作系统为 Windows 10，处理器为 Intel(R) Xeon(R) Silver 4210 CPU@2.20 GHz 2.19 GHz(2 核处理器)，运行内存 128 GB，GPU 为 NVIDIA GTX2080Ti。

实验参数设置：输入图像尺寸为 256 像素×256 像素，每次训练样本数为 16，共训练 1 982 次，进行 40 次迭代，使用交叉熵损失函数与 Adam 优化器，初始学习率为 0.000 3，动量因子为 0.9，每迭代 15 次，学习率衰减一次，衰减因子取 0.1。

2.3 评价指标

在实验中引入 Top－k 准确率，用来计算预测结果中概率最大的前 k 个结果，包含正确标签的占比。准确率(Accuracy)是预测正确的样本数占总样本个数的比例。

由于 AI Challenger 2018 数据集类别分布不均

衡，如柑橘黄龙病图像为 3 627 幅，番茄健康类别图像为 1 028 幅，苹果黑心病图像 361 幅，而西红柿疮痂病图像只有 2 幅。因此引入精确率、召回率、F1 值对本文方法的识别性能进行全面评估[25]。

2.4　实验结果分析

2.4.1　单双流网络性能比较

为验证本文方法对于病害识别的有效性，将 2 个单流网络 CMN、CASAN 与双流网络 FCMSAN 进行实验对比，如表 2 所示，表中的准确率、召回率、F1 值均为所有类别的算术平均值。从实验结果可以看出，FCMSAN 在各项评价指标上的结果都优于 CMN 和 CASAN。在 Top－1 准确率上，FCMSAN 相较于 CMN 和 CASAN，分别提升了 2.31、0.35 个百分点；在 F1 值上，FCMSAN 比 CMN 和 CASAN 分别提升了 3.47、0.61 个百分点。单流网络与双流网络在每个子类上的识别准确率对比如图 5 所示。可以看出，CMN 在某些子类的识别准确率相较 CASAN 更高，FCMSAN 在这些分类上的识别准确率均较高；同样 CASAN 也可以在某些子类上提升识别准确率。CMN 和 CASAN 在识别病害类别时能够相互补充，因此得到的融合网络 FCMSAN 的识别性能最优。

表 2　单双流网络性能对比

Tab. 2　Comparison of single network with double stream networks　%

模型	Top－1 准确率	Top－5 准确率	精确率	召回率	F1 值
CMN	85.66	99.01	81.22	80.80	81.01
CASAN	87.62	99.85	84.40	83.26	83.87
FCMSAN	87.97	99.88	85.21	83.77	84.48

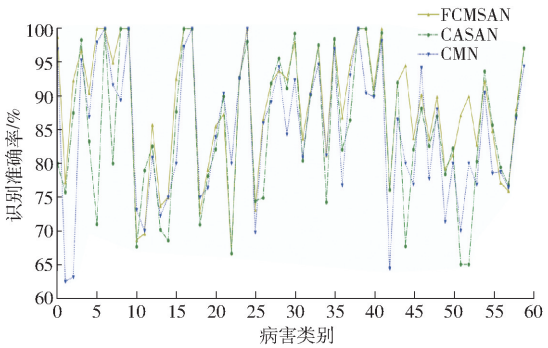


图 5　单双流网络在每个子类上的识别准确率

Fig. 5　Precision of single and dual networks on each subclass

2.4.2　基础模块堆叠数的消融实验

FCMSAN 的单流网络结构基于 ResNet50，由堆叠若干基础模块后下采样，重复 4 次构成。表 3 堆叠数量中，每个数字代表堆叠的基础模块数量，[3,4,6,3]即为 ResNet50 原有的堆叠结构。保留前两次的堆叠参数，对后两次的数量进行调整。实验结果显示，当 FCMSAN 的基础模块堆叠数为[3,4,5,2]时，模型的各项指标均达到最高，能够获得最好的识别性能。

表 3　基础模块堆叠数量不同时的识别效果

Tab. 3　Comparison of basic blocks with different stacked numbers　%

堆叠数量	Top－1 准确率	Top－5 准确率	精确率	召回率	F1 值
[3,4,4,2]	87.47	99.82	84.61	83.23	84.18
[3,4,5,2]	87.97	99.88	85.21	83.77	84.48
[3,4,6,2]	87.22	99.78	84.27	83.22	83.66
[3,4,4,3]	87.37	99.80	84.42	83.43	83.78
[3,4,5,3]	87.24	99.78	84.32	83.26	83.73
[3,4,6,3]	87.12	99.78	84.18	83.17	83.46

2.4.3　模块输入数的消融实验

对特征转换融合模块输入数进行消融实验，结果如表 4 所示。将 CMN 输出的 f_{m1}、f_{m2}、f_{m3}、f_{m4} 与 CASAN 输出的 f_1、f_2、f_3、f_4 对应，输入数表示从后向前选取的数量，可以看出，当融合模块的输入数为 3，即选择 f_{m2}、f_{m3}、f_{m4} 与 f_2、f_3、f_4 作为 TFL 的输入时，模型的 Top－1 准确率与 F1 值最高，因此选择此模块的输入数。

表 4　融合模块输入数不同时的识别效果

Tab. 4　Comparison of fusion module with different inputs　%

输入数	Top－1 准确率	Top－5 准确率	精确率	召回率	F1 值
4	87.48	99.83	84.22	83.24	83.80
3	87.97	99.88	85.21	83.77	84.48
2	87.52	99.88	84.27	83.22	83.86
1	87.46	99.82	84.12	83.34	83.77

2.4.4　通道自适应自注意力机制的性能分析

为了验证所提出的自注意力机制的性能，将经典自注意力机制 CBAM[12]、SA[26]、CA[16]、Non-Local[14]和 CASA 分别嵌入到 ResNet50 的 conv4_x 与 conv5_x 之间，实验结果如表 5 所示。可以看出，CASA 在病害识别任务上的各项指标优于其他注意力机制，与全局建模感受野的自注意力机制 Non-Local 相比，CASA 的 Top－1 准确率提升了 0.57 个百分点；与轴向建模的自注意力机制 CA 相比，CASA 的 Top－1 准确率提升了 0.27 个百分点；比通道建模的注意力机制 CBAM 与 SA 的 Top－1 准确率高 0.23、0.35 个百分点。实验证明，CASA 可以注意到特征间的关系，将更多权重赋给重要特征，从而提高模型的识别性能。为进一步探究 CASA 的有

效性,本文利用激活热图(Grad - CAM)[27]对不同模型进行可视化分析比较,如图 6 所示。对比分析输入图像,可知第 1 行、第 3 行的病害呈圆形斑点,大量分布在叶片中;第 2 行病害以斑状分布在各处,但是以最上方边缘处、正中偏右处最为严重;第 4 行病害以块状体现在叶片左上边缘。通过热力激活图可以看出,与 CBAM、SA、CA、Non - Local 等模型相比,CASA 模型对叶片的病害区域定位更加准确、广泛。说明 CASA 模型能够有效提取病害特征。

表 5 各种注意力机制对比

Tab. 5 Comparison results of different attention mechanisms %

模型	Top - 1 准确率	Top - 5 准确率	精确率	召回率	F1 值
CBAM	87.27	99.81	83.67	83.24	83.46
SA	87.15	99.85	83.69	83.26	83.38
CA	87.23	99.85	83.56	83.22	83.44
Non - Local	86.93	99.78	83.59	82.66	83.13
CASA	87.50	99.88	84.40	83.46	83.97

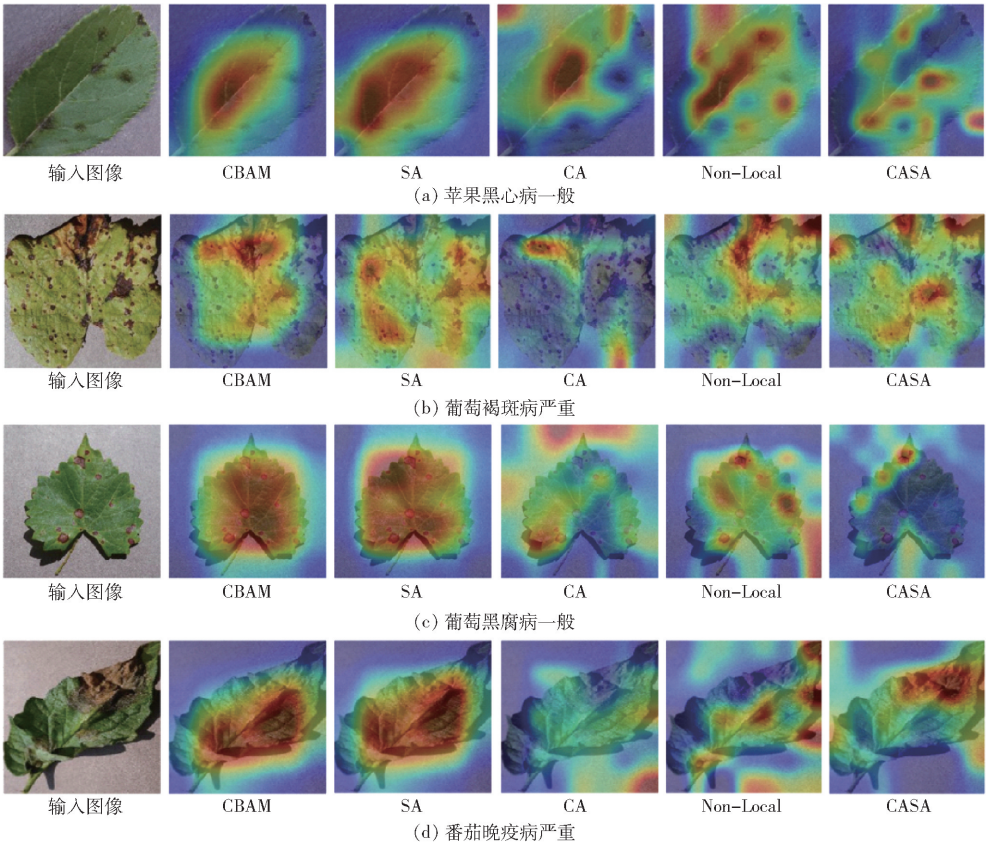


图 6 各注意力机制热力激活图对比

Fig. 6 Comparison of thermal activation diagrams with different attention mechanisms

2.4.5 与其他方法对比

将本文方法与其他方法进行比较,结果如表 6 所示。可以看出,FCMSAN 的各项指标均表现良好。

表 6 与其他方法识别性能对比

Tab. 6 Comparison recognition performance with other methods %

模型	Top - 1 准确率	Top - 5 准确率	精确率	召回率	F1 值
AlexNet[10]	85.36	99.77	81.63	80.79	81.38
VGG16[11]	85.45	99.58	81.75	81.02	81.52
GoogLeNet[12]	86.65	99.78	83.22	82.87	83.05
ResNet50[13]	87.06	99.81	83.19	83.38	83.29
Inception - v3 + Att + Salie[16]	87.48				
I_CBAM[17]	86.98				
FCMSAN	87.97	99.88	85.21	83.77	84.48

其中 Top - 1 准确率达到 87.97%,分别比 AlexNet[10]、VGG16[11]、GoogLeNet[12]、ResNet50[13]、Inception - v3 + Att + Salie[16]和 I_CBAM[17]提升 2.61、2.52、1.32、0.91、0.49、0.99 个百分点;F1 值分别比 AlexNet、VGG16、GoogLeNet、ResNet50 提高 3.1、2.96、1.43、1.19 个百分点,证明在样本分布不均衡情况下,FCMSAN 对各类病害的识别具有较好的效果。

本文提出的方法具有较高的识别准确性,进一步对图像病害识别检测效率进行对比实验,如表 7 所示,可以看出本文方法的检测时间略长于基线模型的检测时间。这是因为 FCMSAN 集成了双流网络,模型参数更大;且 CASA 模块在提取到更丰富的病害特征的同时,也引入了更大的计算量,使得

表 7　与其他方法检测速度对比

Tab. 7　Comparison detection speed with other methods

ms

模型	单幅图像训练时间	单幅图像测试时间
AlexNet[10]	7.5	10.32
VGG16[11]	6.8	12.8
GoogLeNet[12]	9.3	16.8
ResNet50[13]	19.18	13.89
FCMSAN	23.5	25.3

FCMSAN 的检测速度较基线网络在可接受的范围内有所下降。

3　结束语

针对当前植物叶片病害细粒度分类粗糙、识别准确率低等问题，提出基于颜色掩膜与通道自适应自注意力机制的病害识别网络。该网络能够根据叶片颜色变化、病斑范围与病斑间的位置关系，有效学习叶片病害颜色特征和位置特征。在公开数据集上进行大量实验证明了 FCMSAN 网络的有效性，对病害识别达到了较好的效果。

参　考　文　献

[1] 王翔宇, 温皓杰, 李鑫星, 等. 农业主要病害检测与预警技术研究进展分析[J]. 农业机械学报, 2016, 47(9): 266 - 277.
WANG Xiangyu, WEN Haojie, LI Xinxing, et al. Research progress analysis of mainly agricultural diseases detection and early warning technologies[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 266 - 277. (in Chinese)

[2] 孙红, 李松, 李民赞, 等. 农业信息成像感知与深度学习应用研究进展[J]. 农业机械学报,2020, 51(5) : 1 - 17.
SUN Hong, LI Song, LI Minzan, et al. Research progress of image sensing and deep learning in agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5):1 - 17. (in Chinese)

[3] PARRAGA-ALAVA J, CUSME K, LOOR A, et al. RoCoLe: a robust coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition[J]. Data in Brief, 2019, 25: 104414.

[4] SHRIVASTAVA V K, PRADHAN M K. Rice plant disease classification using color features: a machine learning paradigm[J]. Journal of Plant Pathology, 2021, 103(1): 17 - 26.

[5] ANAGNOSTIS A, ASIMINARI G, PAPAGEORGIOU E, et al. A convolutional neural networks based method for anthracnoe infected walnut tree leaves identification[J]. Applied Sciences, 2020, 10(2): 469 - 496.

[6] 翟肇裕, 曹益飞, 徐焕良, 等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报, 2021, 52(7): 1 - 18.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1 - 18. (in Chinese)

[7] 余小东, 杨孟辑, 张海清, 等. 基于迁移学习的农作物病虫害检测方法研究与应用[J]. 农业机械学报, 2020,51(10): 252 - 258.
YU Xiaodong, YANG Mengji, ZHANG Haiqing, et al. Research and application of crop diseases detection method based on transfer learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(10): 252 - 258. (in Chinese)

[8] 贾少鹏, 高红菊, 杭潇. 基于深度学习的农作物病虫害图像识别技术研究进展[J]. 农业机械学报,2019, 50(增刊): 313 - 317.
JIA Shaopeng, GAO Hongju, HANG Xiao. Research progress on image recognition technology of crop pests and diseases based on deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 313 - 317. (in Chinese)

[9] 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.
XU Jinghui, SHAO Mingye, WANG Yichen, et al. Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 230 - 236, 253. (in Chinese)

[10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25: 1097 - 1105.

[11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.

[12] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// Proceedings of the IEEE Conferenceon Computer Vision and Pattern Recognition, 2015: 1 - 9.

[13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 - 778.

[14] 宋余庆, 谢熹, 刘哲, 等. 基于多层 EESP 深度学习模型的农作物病虫害识别方法[J]. 农业机械学报, 2020,51(8): 196 - 202.
SONG Yuqing, XIE Xi, LIU Zhe, et al. Crop pests and diseases recognition method based on multi-level EESP model[J].

Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 196 - 202. (in Chinese)
[15] 蒲秀夫, 宁芊, 雷印杰, 等. 基于二值化卷积神经网络的农业病虫害识别[J]. 中国农机化学报, 2020, 41(2): 177 - 182.
PU Xiufu, NING Qian, LEI Yinjie, et al. Identification of agricultural pests and diseases based on binarization convolutional neural network[J]. Journal of Chinese Agricultural Mechanization, 2020,41(2): 177 - 182. (in Chinese)
[16] 金晓芙. 图像分类技术在农作物病虫害分类问题中的研究与应用[D]. 北京:北京邮电大学, 2020.
JIN Xiaofu. Research and application of image classification technology in classification of crop diseases and insect pests[D]. Beijing: Beijing University of Posts and Telecommunications, 2020. (in Chinese)
[17] 王美华, 吴振鑫, 周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. 农业机械学报, 2021, 52(4): 239 - 247.
WANG Meihua, WU Zhenxin, ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 239 - 247. (in Chinese)
[18] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 - 19.
[19] VASWAMI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv preprint arXiv:1706.03762, 2017.
[20] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7794 - 7803.
[21] WANG H, ZHU Y, GREEN B, et al. Axial-deeplab: stand-alone axial-attention for panoptic segmentation[C]// European Conference on Computer Vision. Springer, Cham, 2020: 108 - 126.
[22] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713 - 13722.
[23] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16 × 16 words: transformers for image recognition at scale[J]. arXiv preprint arXiv: 2010.11929, 2020.
[24] YANG M, HE Y, ZHANG H, et al. The research on detection of crop diseases ranking based on transfer learning[C]//2019 6th International Conference on Information Science and Control Engineering (ICISCE). IEEE, 2019: 620 - 624.
[25] NGOA C, PHAN R C, SEE J. Spontaneous subtle expression recognition: imbalanced databases and solutions[C]// Proceedings of the Asian Conference on Computer Vision. Heidelberg: Springer, 2014: 33 - 48.
[26] ZHANG Q L, YANG Y B. Sa-net: shuffle attention for deep convolutional neural networks[C]//2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). IEEE, 2021: 2235 - 2239.
[27] SEVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 618 - 626.