doi:10.6041/j.issn.1000-1298.2021.12.026

基于 RSTCNN 的小麦叶片病害严重度估计

鲍文霞[1]　林　泽[1]　胡根生[1]　梁　栋[1]　黄林生[1]　杨先军[2]

(1. 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;
2. 中国科学院合肥物质科学研究院, 合肥 230031)

摘要: 以小麦叶片条锈病和白粉病为研究对象,针对同类型病害的不同严重度之间的图像颜色及纹理特征差异较小,传统方法病害严重度估计准确率不高的问题,提出一种基于循环空间变换的卷积神经网络(Recurrent spatial transformer convolutional neural network,RSTCNN)对小麦叶片病害进行严重度估计。RSTCNN 包含 3 个尺度网络,并由区域检测子网络进行连接。每个尺度网络以 VGG19 作为基础网络以提取病害的特征,同时为了统一区域检测过程中前后特征图的维度,在全连接层前引入空间金字塔池化(Spatial pyramid pooling,SPP);区域检测子网络则采用空间变换(Spatial transformer,ST)有效提取尺度网络特征图中病害的注意力区域。小麦叶片病害图像通过每个尺度网络中卷积池化层得到的特征图,一方面可作为预测病害严重度类别概率的依据,另一方面通过 ST 进行注意力区域检测并将检测到的区域作为下一个尺度网络的输入,通过交替促进的方式对注意力区域检测和局部细粒度特征表达进行联合优化和递归学习,最后对不同尺度网络的输出特征进行融合再并入到全连接层和 Softmax 层进行分类,从而实现小麦叶片病害严重度的估计。本文对采集的患有条锈病和白粉病的小麦叶片图像结合数据增强方法构建病害数据集,实验验证了改进后的 RSTCNN 在 3 层尺度融合的网络对病害严重度估计准确率较佳,达到了 95.8%。相较于基础分类网络模型,RSTCNN 准确率提升了 7~9 个百分点,相较于传统的基于颜色和纹理特征的机器学习算法,RSTCNN 准确率提升了 9~20 个百分点。结果表明,本文方法显著提高了小麦叶片病害严重度估计的准确率。

关键词: 小麦; 叶片病害; 严重度估计; 循环空间变换卷积神经网络

中图分类号: S435.121; TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2021)12-0242-11　　**OSID**:

Severity Estimation of Wheat Leaf Diseases Based on RSTCNN

BAO Wenxia[1]　LIN Ze[1]　HU Gensheng[1]　LIANG Dong[1]　HUANG Linsheng[1]　YANG Xianjun[2]

(1. *National Engineering Research Center for Agro-ecological Big Data Analysis and Application*, *Anhui University*, *Hefei* 230601, *China*
2. *Hefei Institutes of Physical Science*, *Chinese Academy of Sciences*, *Hefei* 230031, *China*)

Abstract: Accurately estimating the severity of wheat leaf diseases can reduce planting costs and agricultural ecological environment pollution by the targeted application of pesticides, and contribute to the precise prevention and control of diseases in wheat field, while reducing the cost of the pesticide and the pollution of agricultural ecological environment. The stripe rust and powdery of wheat leaves were taken as the research object. For the problems of small differences in color and texture characteristics of images with different severities of the same diseases, and low classification accuracy of traditional methods, an improved convolutional neural network was proposed based on recurrent spatial transform (Recurrent spatial transformer convolutional neural network, RSTCNN), which was conductive for severity assessment of wheat leaf diseases. RSTCNN consisted of three scale networks which were connected by regional detection sub-networks. In each scale network, VGG19 was used as the basic classification sub-network to extract the disease features. Furthermore, in order to fix the dimensions of the front and behind feature maps in the region detection process, the spatial pyramid pooling (SPP) was introduced before the fully connected layer. Spatial transform (ST) was used by region detection sub-

收稿日期: 2020-12-28　修回日期: 2021-04-02

基金项目: 国家自然科学基金项目(61672032、41771463)和安徽省科技重大专项(16030701091)

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

通信作者: 杨先军(1980—),男,研究员,博士,主要从事智能信号处理与检测技术研究,E-mail: xjyang@iim.ac.cn

network to effectively extract the attention region of the upper scale network feature map. The feature map of the disease image obtained by the convolutional pooling layer can be used as a key for predicting the category probability of the severities at multiple scales. Besides, ST were performed for the detection of region attention to serve as the input of the next scale network. The joint optimization and recursive learning of attention region detection and local fine-grained feature representation were carried out by means of alternating promotion. Finally, the output features of the different scale networks were merged, and then incorporated into the fully connected layer and the Softmax layer for classification, so as to realize the estimation of wheat leaf disease severity. The disease dataset of wheat leaf images with stripe rust and powdery mildew was built by data enhancement methods. Through experiments, results were found that the improved RSTCNN had a better accuracy in estimating the severity of network fused at three layer scales, reaching an accuracy of 95.8%. Compared with the basic classification network model, the accuracy rates were increased by 7 ~ 9 percentage points, which effectively enhanced the classification ability of the disease areas in the images. Compared with traditional machine learning algorithms based on color and texture features, the accuracy rates of RSTCNN were improved by 9 ~ 20 percentage points. The results showed that the proposed method significantly improved the estimation accuracy of wheat leaf disease severities.

Key words: wheat; leaf disease; severity estimation; RSTCNN

0　引言

准确地估计小麦病害的严重度,有助于麦田病情的精准防控,同时针对性施用农药,减少种植成本和农业生态环境污染,对于病害的科学防治尤为关键[1]。

目前小麦病害严重度估计主要依靠有经验的工作人员到现场进行目视观察,人工方法耗时且费力,而且依赖于经验。利用可见光、多光谱或者红外传感器可以获得作物病害信息,结合计算机视觉和机器学习技术实现高效监测作物病害的严重程度,从而实时科学地精细化管控农业生产,提升作业效率,对于实施“智慧农业”有着积极意义。染病组织的光谱反射率有效地反映作物病害发生过程中的内部结构和形态等差异,因而利用光谱成像技术可以进行小麦病害的严重度估计[2-3],但该方法的计算成本较大,而且野外环境对高光谱图像的采集造成一定影响。与传统的人工检测和高光谱成像相比,红外和可见光图像技术在小麦病害严重度估计工作中应用更广泛[4-5]。通过采集的红外热成像探明病害过程中的温度变化规律,以及对 RGB 图像的分析和处理,可实现对小麦叶锈病的快速检测和分级。目前基于经典机器视觉的病害严重度估计方式特征提取难,由于小麦病害叶片图像易受复杂的背景环境干扰,尤其是同类别病害的不同严重度之间的颜色纹理特征相似,细粒度差异较小,传统方法难以准确地对小麦叶片病害进行严重度估计。

随着深度学习技术的发展,卷积神经网络(Convolutional neural networks,CNN)在图像处理和分类中的应用日益广泛[6-9]。与传统机器学习算法相比,CNN 能够自动检索图像的全局信息和局部特征,不需要人工检测和提取,具有更好的性能[10-12]。

相比于作物病害的识别[13-15],病害的严重度估计对方法的有效性和精确度的要求更高。由于作物病害不同严重度之间的特征差异小,准确进行病害严重度估计往往需要大量的数据对网络进行训练[16-19]。上述严重度估计方法均是直接对原始数据集进行学习以实现对作物病害的严重度估计,学习的对象是病害图像的全局信息,缺乏作物病害图像局部区域之间的相互关联性[20]。

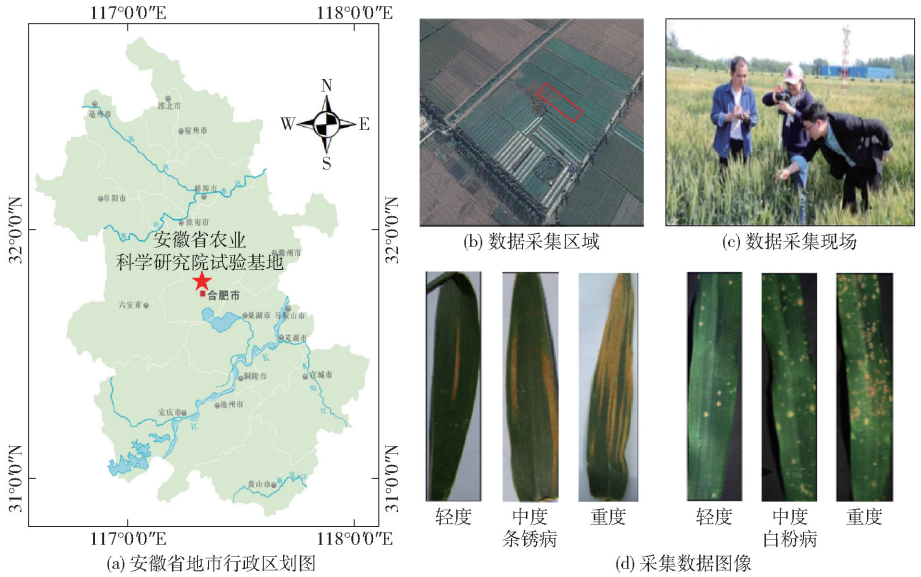
小麦叶片病害的严重度等级之间的颜色纹理特征差异较小,可以归属于细粒度分类问题。目前,细粒度分类问题的解决方式可以分为两类:基于图像注意力区域检测的方法[21-22]和基于图像细粒度特征表达的方法[23-24],分别通过弱监督的信息检测图像中的注意力区域和使用高维图像特征对图像信息自适应提取,实现精细化分类。文献[25]研究表明,针对细粒度图像的处理和分类问题,形态、轮廓特征的重要性表现较小,而局部细节特征则至关重要。针对小麦条锈病和白粉病症状主要位于叶片的局部区域的特征,运用传统分类方式对全局信息进行卷积操作,势必会影响最终的分类准确率。本文通过注意力区域检测和局部细粒度特征表达的联合优化和递归学习,构造基于循环空间变换的卷积神经网络(Recurrent spatial transformer convolutional neural network,RSTCNN)方法用于小麦叶片病害的严重度估计。将图像由卷积池化后生成的特征图,通过空间变换(Spatial transformer,ST)层实现叶片的病害区域检测,提取出病害特征的关键区域,接着以交替促进的方式对注意力区域检测和局部细粒度

特征表达进行联合优化,由分类子网络实现图像分类并依次递归学习。最后对各尺度全连接层的输出进行特征融合,再并入到 Softmax 层预测最终的概率,以实现病害的严重度估计。

1 材料与方法

1.1 数据来源

本文所用小麦叶片病害图像来源于安徽省农业科学研究院试验基地(31°89′N,117°25′E),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,长期以来一直承担小麦、水稻等多种农作物育种科研试验任务。叶片病害数据采集于 2019 年 4 月 26 日至 2019 年 5 月 10 日 09:00—16:00,天气晴朗且少云,相对湿度和温度分别在 30% ~50% 和 18 ~25℃之间。此时该基地的小麦叶片正值青翠繁茂期,主要遭受条锈病、白粉病的侵扰。拍摄设备为 EOS80D 型佳能数码相机(图像尺寸 6 016 像素 × 4 000 像素,焦距 4 mm,光圈 f/22,曝光时间 1/2 000 s)。采集并经过筛选后得到两类小麦叶片病害图像各计 180 幅。研究区域和数据采集工作图如图 1 所示。

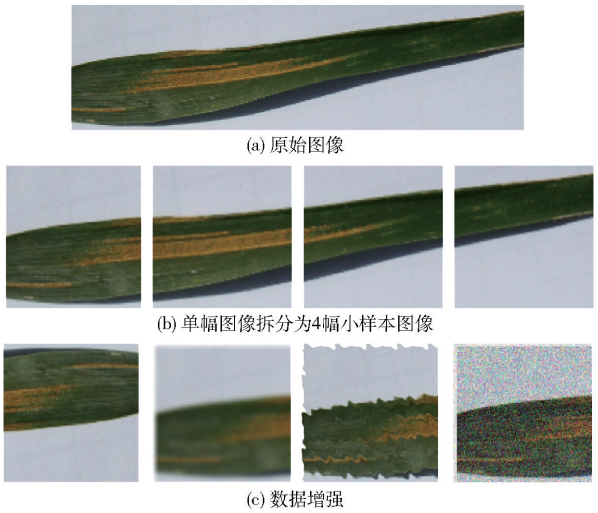


(a) 安徽省地市行政区划图 (b) 数据采集区域 (c) 数据采集现场 (d) 采集数据图像

图 1 研究区域和数据采集工作图

Fig. 1 Schematics of research area and data collection

为了更好地检验模型性能,针对获取的数据集样本有限和不平衡的情况,先将原始样本通过裁剪的方式,切分图像的长边为 4 等份,每幅图像可扩充得到 4 幅病害图像,如图 2 所示。接着对扩充后的病害图像按照病斑面积占叶片总面积的百分比 R 来界定病害严重程度[26]:$0 < R \leqslant 30\%$ 为轻度,$30\% < R \leqslant 50\%$ 为中度,$50\% < R$ 为重度,从而得到 2 种病害 6 类不同严重度的数据样本。最后从各类严重度的原始样本中分别随机选取 40 幅样本作为测试集,其余的通过旋转、模糊、扭曲和加噪等数据增强处理分别扩充至 600 幅,共计 3 600 幅,再随机划分训练集和验证集的数目比值为 9∶1,以提高样本多样性和模型鲁棒性,防止过拟合现象的发生。数据集样本分布见表 1。



(a) 原始图像

(b) 单幅图像拆分为4幅小样本图像

(c) 数据增强

图 2 数据集制作流程

Fig. 2 Dataset manufacturing process

表 1 数据集样本分布

Tab. 1 Sample distribution of dataset 幅

数据集	白粉病			条锈病			总计
	轻度	中度	重度	轻度	中度	重度	
原始样本	335	179	86	276	203	121	1 200
扩充样本	600	600	600	600	600	600	3 600
训练集	540	540	540	540	540	540	540
验证集	60	60	60	60	60	60	360
测试集	40	40	40	40	40	40	240

1.2 小麦病害严重度估计

针对同种病害不同严重度间的颜色、纹理特征差异较小以及细粒度图像难以分类问题,本文构建 RSTCNN 网络模型,该网络由图像分类和区域检测

两个子任务的多层结构构成，运用区域检测和图像分类两个子网络交替训练的思想，通过交替促进的方式对注意力区域检测和局部细粒度特征表达进行联合优化和递归学习，从而达到精确估计小麦叶片病害严重度的目的。

1.2.1　RSTCNN

RSTCNN 主要包含 3 个尺度网络（可根据需要叠加多个尺度），每个尺度网络的结构都一样，均以基础分类网络模型为框架，并插入空间金字塔池化（Spatial pyramid pooling，SPP）层固定输出特征图维度，各尺度网络间通过 ST 层连接并进行区域检测。本文实验的每个尺度网络以 VGG19 作为子网络，图像通过卷积池化层生成的特征图，一方面输入该尺度网络的 SPP 层和全连接层，另一方面由空间变换层进行区域检测后，作为下一个尺度网络的输入。RSTCNN 网络架构如图 3 所示，其中 $a_s(s=1,2,3)$ 表示各尺度网络（Scale 1、Scale 2、Scale 3），以 VGG19 作为框架，其网络结构如图 4 所示；b_s 表示各尺度网络的卷积层和池化层，从输入病害叶片图像中提取特征；c_s 表示各尺度网络的 SPP 层，用来在全连接层前统一输出不同特征图维度；d_s 表示各尺度网络的全连接层和Softmax 层，可得到单个分类

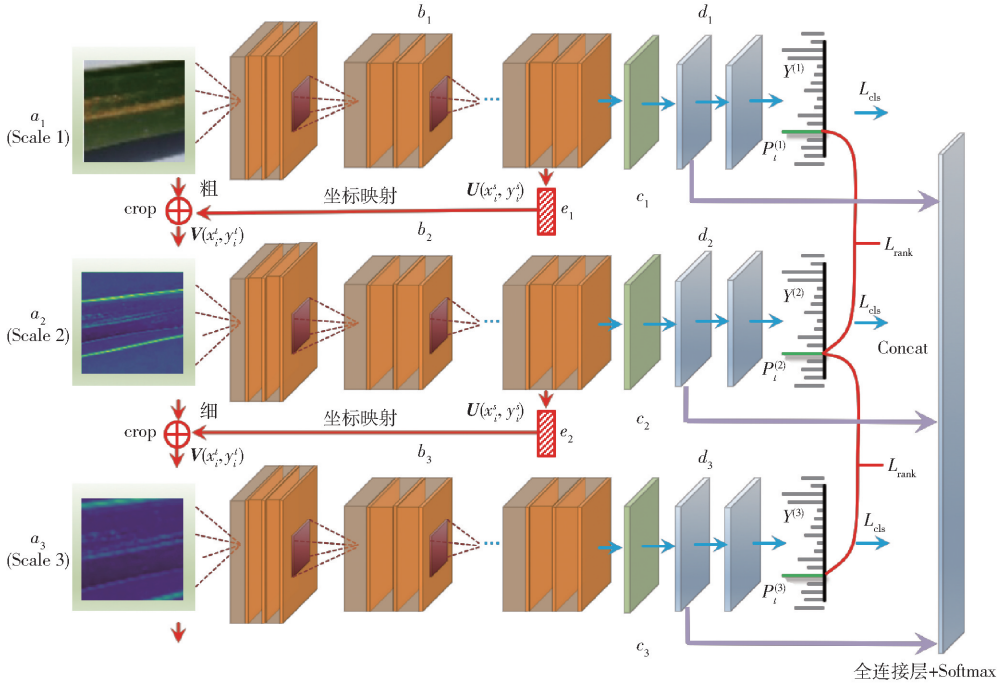


图 3　RSTCNN 网络架构

Fig. 3　Network architecture of RSTCNN

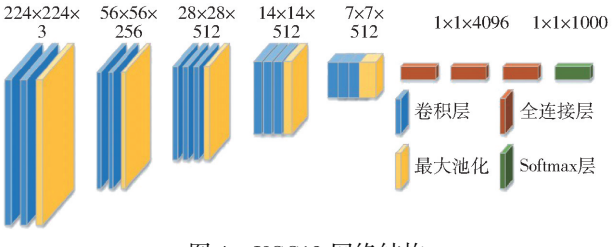


图 4　VGG19 网络结构

Fig. 4　Network architecture of VGG19

子网络的输出结果。

在更高尺度的网络中，以上层尺度网络卷积池化后的特征图作为研究对象，一方面由 ST 层（e_1、e_2）进行区域检测，由粗到细迭代生成病害细粒度表征的注意力区域并提取，另一方面通过分类子网络，送入 SPP 层和全连接层。接着将注意力区域特征图以循环的方式在下一个尺度网络再次进行迭代训练，然后对全连接层的输出规范一致化后以 Concat 方式对特征图通道合并以达到特征融合[27]的目的，最后并入到 Softmax 层预测该图像病害严重度类别的概率，获得更佳的分类效果。算法流程如下：

（1）将数据集通过分类子网络提取特征并分类，然后在全连接层前将提取特征图送入 ST 层区域，检测得到叶片病害的注意力区域。

（2）将注意力区域生成的特征图在分类子网络的全连接层前通过 SPP 层，统一规范化特征图的尺寸，并送入下级尺度网络，再次进行局部细粒度特征的学习和区域检测。

（3）通过尺度内分类损失和尺度间排序交叉熵损失联合优化，以相互强化的方式对图像注意力区域检测和局部细粒度特征表达进行递归学习，并可根据需要决定尺度和迭代次数。

（4）综合各尺度网络全连接层的输出，规范一致化后通过 Concat 进行特征融合，再并入到 Softmax

层得到最终预测值,达到更佳的分类结果,实现叶片病害严重度的精准估计。

1.2.2 区域检测和提取

区域检测子网络位于各个尺度网络之间,通过空间变换有效检测出上级尺度网络特征图中的注意力区域,提取并作为下级尺度网络的输入。ST 层基本架构如图 5 所示,通过区位网络和采样映射机制[28-29],在网络中对数据进行空间变换处理。其核心为空间转换模块,主要由 3 部分组成,其中区位网络用来计算空间变换的参数,坐标映射器得到输入特征图 $\boldsymbol{U}$ 到输出特征图 $\boldsymbol{V}$ 的映射位置对应关系 $\boldsymbol{T}_\theta$,图像采样器则根据输入特征图 $\boldsymbol{U}$ 和对应映射关系 $\boldsymbol{T}_\theta$,生成最终的输出映射。

(1)区位网络

区位网络以特征图作为输入,仿射变换后经一系列的隐藏网络层学习并输出变换参数 $\boldsymbol{\theta}$。对于一个局部网络,设输入前一层的特征图为 $\boldsymbol{U}$,则区位网络的输出表示为

$$\boldsymbol{\theta}=f_{loc}(\boldsymbol{U}) \tag{1}$$

式中 f_{loc}——卷积神经网络(局部网络)

$\boldsymbol{\theta}$——网络输出的变换矩阵,用作转换器的训练参数(本文实验为二维仿射变换,$\boldsymbol{\theta}$ 为六维矩阵,元素分别记作 $\theta_{\alpha\beta}(\alpha=1,2;\beta=1,2,3)$)

(2)坐标映射器

采用逆变换的方法获得输出特征图,并生成网格上每个像素点的对应值。根据区位网络的参数,坐标映射器构建逆变换后的图像与输入图像上采样网格 G_i 的坐标映射,得到输入特征图 $\boldsymbol{U}\in\mathbf{R}^{H\times W\times C}$ 到输出特征图 $\boldsymbol{V}\in\mathbf{R}^{H'\times W'\times C}$ 的各映射位置对应关系 $\boldsymbol{T}_\theta$。其中 W、H 是对应特征图的宽、高,C 为通道数。通过仿射变换,将输入和输出的横纵坐标范围规范一致化,得到变换后输出的特征图。

假设特征图像 $\boldsymbol{U}$ 的像素坐标为 (x_i^u,y_i^u),$\boldsymbol{V}$ 的像素坐标为 (x_i^v,y_i^v),空间变换函数 $\boldsymbol{T}_\theta$ 为二维仿射变换函数,则特征图 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 对应关系为

$$\begin{bmatrix}x_i^u\\y_i^u\end{bmatrix}=\boldsymbol{T}_\theta(G_i)=\boldsymbol{\theta}\begin{bmatrix}x_i^v\\y_i^v\\1\end{bmatrix}=\begin{bmatrix}\theta_{11}&\theta_{12}&\theta_{13}\\\theta_{21}&\theta_{22}&\theta_{23}\end{bmatrix}\begin{bmatrix}x_i^v\\y_i^v\\1\end{bmatrix} \tag{2}$$

根据式(2)计算出每个输入特征图像素点的对应输出像素点位置。因 (x_i^u,y_i^u) 有可能落在输入特征图像素点的中间区域,故需通过双线性插值推算出相应点,由逆变换得到输出特征图 $\boldsymbol{V}(x_i^v,y_i^v)$ 的像素点值。

(3)图像采样器

由前两部分传输变换获得相应的网格和坐标映射关系 $\boldsymbol{T}_\theta$ 后,将特征图和坐标映射关系同时作为生成器的输入获得输出。

在原始图中采样像素复制到对应图中,生成输出映射。将原图像像素坐标仿射变换为目标图像的像素点坐标,表达为

$$\boldsymbol{V}_i^c=\sum_n^H\sum_m^W\boldsymbol{U}_{n,m}^c k(x_i^s-n;\boldsymbol{\Phi}_x)k(y_i^s-m;\boldsymbol{\Phi}_y)$$
$$\forall i\in(1,2,\cdots,HW)\quad\forall c\in(1,2,\cdots,C) \tag{3}$$

式中 $\boldsymbol{V}_i^c$——目标图像的位置坐标

$\boldsymbol{U}_{n,m}^c$——位于颜色通道 c 中坐标为 (n,m) 的像素值

k——核函数,表示实现重采样功能的线性插值

x_i^s、y_i^s——原图像像素坐标

$\boldsymbol{\Phi}_x$、$\boldsymbol{\Phi}_y$——采样内核 k 的输入插值参数

综上,ST 层是一个可微模块,既可对输入特征图进行空间变换,也可作为网络模块插入到任务需求网络的任意层中,使得 CNN 能够自主在空间上转换特征映射,进行区域检测和获取。

对检测区域进行提取,设提取出的病害特征区域均为正方形,表示为 (t_x,t_y,t_l),可得

$$(t_x,t_y,t_l)=\boldsymbol{g}(W_cX) \tag{4}$$

式中 $\boldsymbol{g}(\cdot)$——ST 层网络函数

X——输入图像

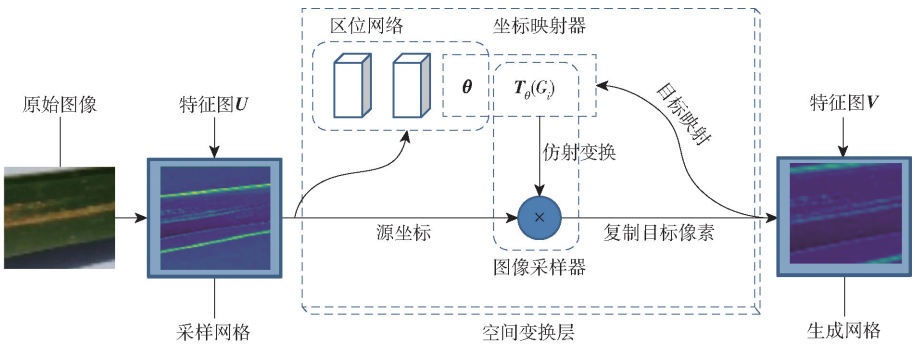


图 5 ST 层架构

Fig. 5 ST layer architecture

W_c——分类子网络，从 X 中提取特征

t_x、t_y——检测区域中心点坐标

t_l——检测区域边长

特征图的注意力区域检测示意图如图 6 所示。

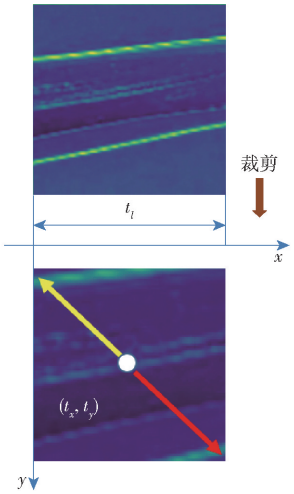


图 6　区域检测示意图

Fig. 6　Schematic of region detection

假设特征图的左上角是坐标系的起点，用符号 C_l 表示左上角点，符号 C_r 表示右下角点。则根据 ST 层网络的 t_x、t_y 和 t_l，可计算出裁剪后区域的 x 轴坐标 $t_{x(C_l)}$、$t_{x(C_r)}$ 和 y 轴坐标 $t_{y(C_l)}$、$t_{y(C_r)}$ 分别为

$$\begin{cases} t_{x(C_l)} = t_x - t_l/2 \\ t_{y(C_l)} = t_y - t_l/2 \\ t_{x(C_r)} = t_x + t_l/2 \\ t_{y(C_r)} = t_y + t_l/2 \end{cases} \tag{5}$$

为实现整体网络端到端优化训练，需要定义掩膜函数为连续函数以易于在逆向传播中优化，这里使用二维 boxcar 函数的变体作为注意掩膜进行近似裁剪操作，避免了直接根据坐标和边长信息生成 (0,1) 的掩膜，表达式为

$$M = (h(x - t_{x(C_l)}) - h(x - t_{x(C_r)})) \cdot (h(y - t_{y(C_l)}) - h(y - t_{y(C_r)})) \tag{6}$$

其中

$$h(x) = \frac{1}{1 + \exp(-kx)} \tag{7}$$

式中　M——掩膜函数

h——Sigmoid 激活函数

通过坐标关系，将掩膜函数 M 和输入图像 X 的数组元素依次相乘提取区域 X^{att}，表达式为

$$X^{att} = X \otimes M(t_x, t_y, t_l) \tag{8}$$

1.2.3　空间金字塔池化

传统 CNN 的全连接层是神经元对于所有输入的连接权重，若输入尺寸不一致，则会使得全连接层参数无法固定。CNN 要求输入维度一致是导致随机尺寸的图像识别率低的关键。RSTCNN 的各尺度网络需要特征图经过 ST 层进行区域检测和特征提取，会使特征图的尺寸发生改变。若采用传统的数据集预处理的方式，对特征图进行裁剪或者放大，使其维度相同，则会带来较大的弊端，一方面裁剪掉的区域可能会包含病害目标区域，另一方面特征图的放大也会导致图像几何方面的失真和扭曲[30]。

为了解决特征图经过 ST 层提取区域后特征图尺寸会改变的问题，RSTCNN 在各尺度网络的全连接层前引入 SPP 层[31-32]，如图 7 所示。不同尺寸的输入特征图规范化为一致性向量，以保存特征信息的同时，改进模型的识别精度。SPP 通过空间像素级的层次关系，将空间的第 l 层级进行划分（$l \in \mathbf{N}^+$），其中每层有 2^l 个子集，横纵坐标则依次切分成 2^l 个网格，获得 4^l 块特征图。对于任意大小的输入特征图，每个子网络上的 3 个池化层分别以 4^l 个网格将特征图分为 4^l 块，并采用最大池化以维持图像特征的细节信息，得到相应的 4^l 维特征向量。最后将各池化层提取的特征融合，使任意尺寸特征图均统一规范为同维的特征向量，不再固定限制输入图像的尺寸。从而避免了特征图因适应传统 CNN 的方式需要进行裁剪或扭曲，导致图像部分信息丢失，特征提取难等问题。

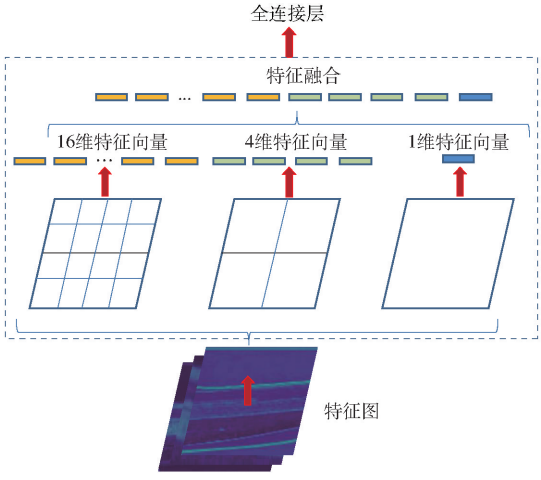


图 7　SPP 示意图

Fig. 7　Schematic of SPP

1.2.4　训练策略

整个网络模型基于注意力区域检测和局部细粒度特征表达联合优化和递归学习的思想，以相互促进的方式迭代训练，算法流程如图 8 所示。

具体的训练步骤如下：

(1) 初始化分类子网络中的卷积层和全连接层参数。

(2) 初始化 ST 层网络参数，用分类子网络最后一层卷积层中注意力特征权值最高的区域初始化区位像素坐标。通过将输入特征图像 $\boldsymbol{U}$ 像素坐标

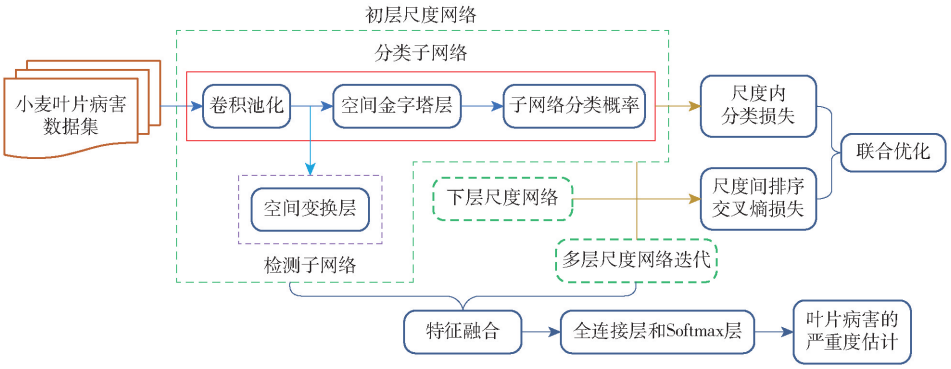


图 8 算法流程图

Fig. 8 Algorithm flow chart

(x_i^s, y_i^s)映射到输出特征图像 $\boldsymbol{V}$ 像素坐标(x_i^t, y_i^t)，调整 d_1、d_2 中的参数对原始图像进行区域范围选择。同理使用类似的方法输出更高尺度上的特征图，得到更精细的注意力区域。

(3)设置学习率，先通过取值为 0 来固定分类和检测子网络参数。接着保持 ST 子网络参数不变，优化分类子网络直到 L_{cls} 收敛。然后固定分类子网络参数，并切换到尺度间的排序交叉熵损失来优化两个 ST 子网络，直到 L_{rank} 收敛。

(4)返回步骤(3)，直到两类子网络损失均收敛。

RSTCNN 的损失函数包含两部分，由 $Y^{(s)}$ 和 Y^* 构成分类损失 L_{cls}，其中 $Y^{(s)}$ 为预测的类别概率，Y^* 为真实类别。以及表示尺度间排序的交叉熵损失 L_{rank}，Scale 1 网络和 Scale 2 网络，以及 Scale 2 网络和 Scale 3 网络分别构成两个不同 L_{rank}，表达式为

$$L(X) = \sum_{s=1}^{3} L_{\text{cls}}(Y^{(s)}, Y^*) + \sum_{s=1}^{2} L_{\text{rank}}(P_t^{(s)}, P_t^{(s+1)}) \tag{9}$$

其中

$$L_{\text{rank}}(P_t^{(s)}, P_t^{(s+1)}) = \max\{0, P_t^{(s)} - P_t^{(s+1)} + m_{\text{argin}}\} \tag{10}$$

式中 $L(X)$——RSTCNN 的损失函数

$P_t^{(s)}$——Scale $s(s=1,2,3)$ 网络的实际标签概率

t——真实标签类别

m_{argin}——定值参数

从式(10)可以得到，当更高尺度网络的 P_t 大于相邻的前一尺度网络的 P_t 时，损失较小，实验中设置 m_{argin} 为 0.05。网络训练时使用损失函数交替训练的策略，并需保证 $P_t^{(s+1)} > P_t^{(s)} + m_{\text{argin}}$。

综上，通过 RSTCNN 获取到的小麦图像输出的不同尺度特征表示为$\{F_1, F_2, \cdots, F_i, \cdots, F_N\}$。其中 N 表示尺度数目，F_i 表示第 i 个尺度网络的全连接层输出。通过独立地对每个 F_i 规范化统一后，以 Concat 方式实现特征融合，再并入到 Softmax 层得到最终的分类结果。

2 结果与分析

2.1 实验环境和评价指标

本文实验软件环境为 Linux Ubuntu 16.04 操作系统、NVIDIA GTX2070 GPU 上的 CUDA8.0 平台，使用 Pytorch 的开源软件框架和 Python 编程语言实现。实验中采用批量训练方式，将训练集和测试集划分为多个批次，以便更好对比观察实验效果。所有模型训练的迭代周期(Epoch)均设置为 300，通过 L2 正则化与 Dropout 策略减少模型过拟合程度，整个训练过程是迭代进行的，并采用损失函数交替训练的策略来优化网络参数。

为了验证基于 RSTCNN 模型对小麦病害叶片进行严重度估计算法的有效性，分别对网络模型集成方式和参数选取进行对比实验，同时也将传统的机器学习分类器和流行的 CNN 进行对比。模型评价指标有精确率(Precision)、召回率(Recall)、F1 值(F1 scores)以及准确率(Accuracy)。其中，精确率针对预测结果，实际正样本占预测为正样本的比例。本文的正样本指属于某一类别的样本，而负样本指不属于某一类别的样本。召回率针对原样本，预测为正样本的数量占实际正样本的比例。F1 值为精确率与召回率的调和均值。准确率为预测正确的结果占总样本的比例。

2.2 不同分类网络及其作为 RSTCNN 子网络时严重度估计

运用本文的 RSTCNN 模型，分别选取常用的 VGG19、GoogleNet、ResNet50、AlexNet 基础网络作为分类子网络，并与区域检测子网络分别构成 RSTCNN_VGG19、RSTCNN_GoogleNet、RSTCNN_ResNet50、RSTCNN_AlexNet 网络模型。通过和基础分类网络进行比较实验，验证 RSTCNN 模型通过分类实现对小麦叶片病害严重度估计的提升效果。

表 2 展示了不同分类网络模型对小麦叶片病害严重度估计的结果，表明相较于传统的分类网络，所提出基于 RSTCNN 模型改进网络增强了基础分类网络对图像中病害区域的检测和分类能力。通过交替促进的方式对注意力区域检测和局部细粒度特征表达进行联合优化和递归学习，根本性解决了细粒度图像识别中，局部细节特征起主导作用的难点，各评价指标均有一定的提升。

表 2　不同分类网络模型的严重度估计结果

Tab. 2　Severity estimation results of different network models　%

网络模型	精确率	召回率	F1 值	准确率
VGG19	86.7	89.0	87.8	88.1
GoogleNet	84.2	84.7	85.5	85.8
ResNet50	85.8	84.5	85.1	83.3
AlexNet	80.7	82.3	81.5	80.7
RSTCNN_VGG19	94.3	94.9	94.6	95.8
RSTCNN_GoogleNet	93.9	93.2	93.5	92.5
RSTCNN_ResNet50	92.3	91.1	91.7	90.6
RSTCNN_AlexNet	91.3	91.9	91.6	89.4

图 9 为不同分类网络模型的严重度估计准确率变化曲线，直观地对比出 RSTCNN 模型在基础分类网络上改进后，各个训练周期时精度的提升状况，并减少了模型达到稳定的训练周期，为小麦叶病的严重度估计提供了高效精准的指导。其中 RSTCNN 网络模型与 VGG19 联合构造的准确率最高，因此本文所有的对比验证实验中，RSTCNN 网络模型均采用 VGG19 作为分类子网络。

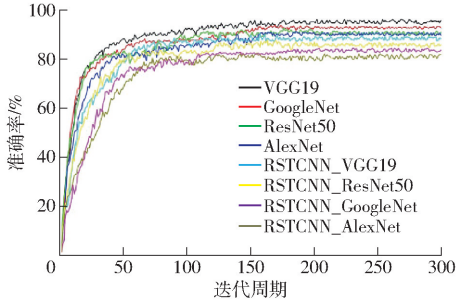


图 9　不同分类网络和 RSTCNN 的严重度估计准确率

Fig. 9　Severity estimation accuracies of different basic classification networks and RSTCNN

2.3　特征可视化和不同区域检测子网络对比

本文运用多尺度网络的优势，表现为模型训练时可逐渐聚焦到小麦叶片病害的显著区域，并增强其特征信息表达能力，有利于提高严重度估计效果。另外针对分类网络和 ST 网络设计 2 个损失，通过限制其中一个网络的参数来交替训练另一个网络，实现图像的注意力区域检测和局部细粒度特征表达的联合优化和递归学习。由于该模型通过多尺度网络对图像及其特征图进行了多次区域检测和再分类，从某种意义上也起到了扩充数据集的效果。

对 RSTCNN 模型学习过程中生成的特征进行可视化，以 VGG19 作为分类子网络，通过 3 个尺度网络的构造形式，得到一系列特征图如图 10 所示，可以发现随着 RSTCNN 尺度的增加，通过 ST 层的区域检测和提取能力，逐渐定位并精细化了图像中具有叶片病害的注意力区域，达到了提高病害严重度估计准确率的目的。

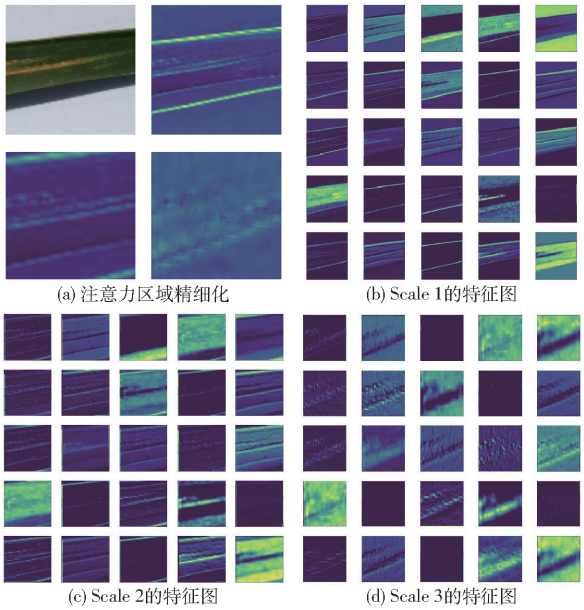


图 10　特征图可视化

Fig. 10　Feature map visualizations

为验证 RSTCNN 模型的 ST 层子网络进行注意力区域检测的优越性，本文选取了常用的区域检测网络进行对比实验，分别得到小麦叶片病害严重度估计的准确率。其中，全卷积自适应网络(FCAN)[33]由图像域适应网络和特征自适应网络组成，以解决域适应背景下的语义分割难题。多注意力卷积神经网络(MACNN)[34]利用特征图不同通道关注的视觉信息不同会导致峰值响应区域差异性的特点，聚类响应区域接近的通道获得注意力区域。多样化视觉注意网络(DVAN)[35]追求最大限度收集差异性信息，并结合图像的不同区域和尺度的信息来建立动态的类内分类表示。

此外，本文还对各尺度(Scale 2、Scale 3、Scale 4)下的单独严重度估计结果以及融合不同尺度特征的严重度估计结果(Scale 1 +2、Scale 1 +2 +3、Scale 1 +2 +3 +4)分别进行了实验。图 11 为不同注意力区域获取方式和 RSTCNN 各个尺度下的病害严重度估计准确率，通过实验结果可发现，随着融合尺度的增加，网络模型更加复杂，但准确率由较高提升逐渐

变缓至甚微,综合衡量 3 层尺度融合的网络准确率为最佳。Scale 4 网络效果不及 Scale 3 网络效果,主要原因在于 Scale 4 丢失了较多的全局信息,导致部分病害区域的信息欠缺。

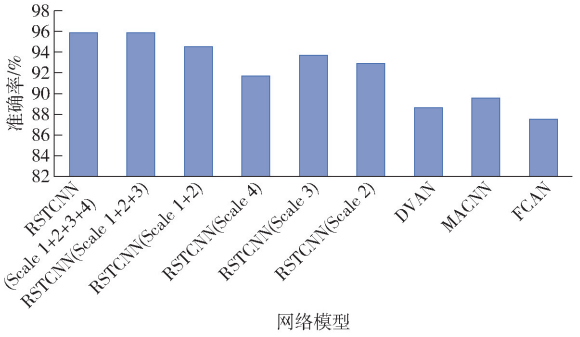


图 11 不同区域检测子网络的严重度估计准确率

Fig. 11 Severity estimation accuracies of different regional detection methods

2.4 病害严重度估计结果混淆矩阵

在使用深度学习识别图像时,混淆矩阵作为有效的可视化工具,将数据划分为输入类和预测输出类的样本数量,直观地把分类结果的精度展示于一幅矩阵图中,比较出分类结果和实际测得值间的差异,因此常被用在多分类问题中评价模型的精度。

实验对小麦叶片的条锈病和白粉病进行病害等级分类为轻度、中度、重度,其中条锈病的轻度、中度、重度分别记为 A、B、C,白粉病的轻度、中度、重度分别记为 D、E、F,共计 6 类。VGG19 基础网络模型和将 VGG19 作为分类子网络的 RSTCNN 网络模型的生成严重度估计混淆矩阵如图 12 所示。图中每行为小麦病害的实际种类,其总数Sum_lin表示归属该种类的实际数据。每列为小麦病害的预测种类,其总数 Sum_col 表示该种类的预测数据。行列交叉处的数值表示小麦病害被预测为相应行种类的数目,而对角线上的值为预测正确的结果总数。根据结果,直观展示出叶片病害严重度的真实数据和预测数据混淆的样本数目,以及各类病害严重度样本的分类错误数据类型,表明 RSTCNN 网络模型可以实现更精准的严重度估计效果。

2.5 与传统算法对比

为验证本文方法相比于传统机器学习算法的优越性,实验将本文方法与 3 种在农作物病害严重度估计上的传统机器学习算法进行了比较分析。这 3 种传统机器学习算法分别运用 K 近邻(KNN)分类[36]提取图像颜色特征,利用随机森林(RF)分类[37]提取图像纹理特征和通过支持向量机(SVM)分类[38]提取图像的组合特征。

实验中,通过开源库 Scikit - learn2.2 调用 3

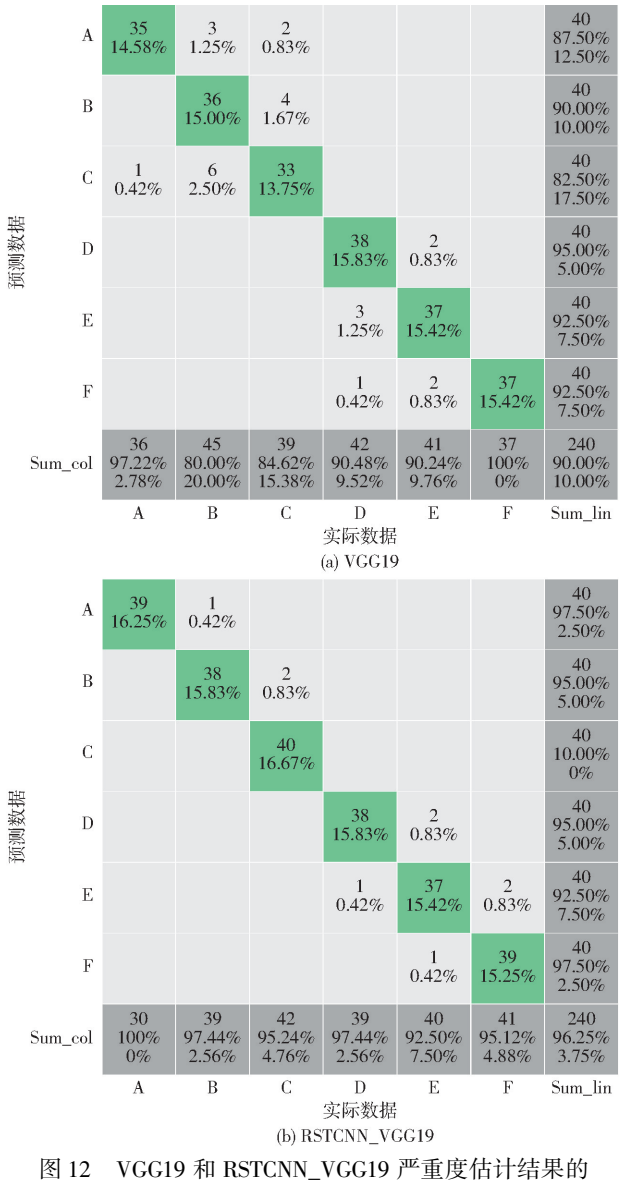


图 12 VGG19 和 RSTCNN_VGG19 严重度估计结果的混淆矩阵

Fig. 12 Confusion matrixes of severity estimation results in VGG19 and RSTCNN_VGG19

种经典机器学习分类器。其中,KNN 的样本点相似性度量采用欧氏距离,待预测样本附近标签的个数在通过交叉验证法后,选取 5 为最优值用于比较实验。RF 选取对样本有放回抽样进行构建,所有参数均先采用默认值再调参,应用十折交叉验证发现预测变量值为 81 且最大特征数为 6 时取得最佳结果。SVM 选取分类库 NuSVC 用于病害严重度的多分类,经调参优化,得到训练误差上界和支持向量下限的值为 0.5,核函数为 RBF 核,函数最高次幂为 3 时,效果最佳。在经上述参数调优设置后,不同经典算法的平均严重度估计准确率如图 13 所示,可以发现相较于传统的机器学习算法,本文提出的 RSTCNN 网络模型显著提高了小麦叶片病害严重度估计的平均准确率。

不同方法的精确率、召回率和F1值如表3所

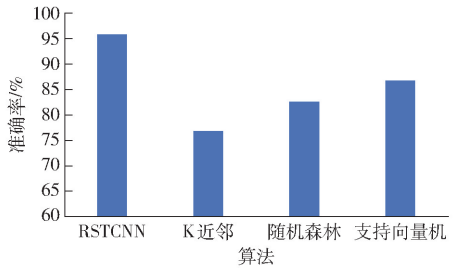


图 13　本文方法与传统机器学习算法的严重度估计准确率对比

Fig. 13　Comparison of severity estimation accuracy with traditional machine learning algorithms

表 3　与传统机器学习算法的各类指标对比结果

Tab. 3　Comparison of various indicators with traditional machine learning algorithms　%

算法	精确率	召回率	F1 值
RSTCNN	94.3	94.9	94.6
K 近邻	75.9	78.4	77.1
随机森林	77.6	81.5	79.5
支持向量机	87.8	84.5	85.9

示，可以得到，相比于其他 3 种传统机器学习方法，本文所提的 RSTCNN 方法具有更佳的评价指标。

3　结论

（1）针对病害等级间特征差异较小、细粒度图像难以分类的问题，RSTCNN 运用区域检测和图像分类 2 个子网络交替训练的思想，基于注意力区域检测和局部细粒度特征表达的联合优化和递归学习的方式，逐渐聚焦到图像中具有叶片病害的注意力区域，减少无关信息的干扰，然后对检测出的病害注意力区域自适应提取，并送入下一级的分类网络，依次迭代交替训练，最后通过多尺度特征融合，达到更好的分类效果，实现小麦叶片病害的严重度估计。与传统分类网络相比，RSTCNN 网络模型的精确率为 94.3%，召回率为 94.9%，F1 值为 94.6%，病害严重度估计准确率达到 95.8%，各评价指标均有所提高，表明该模型显著增强了基础分类网络对图像中病害区域的分类能力。

（2）与常见的区域检测网络进行对比实验，RSTCNN 模型的 ST 层子网络比其他方法严重度估计准确率提高了 4 ~ 5 个百分点。另外还对各尺度下的单独严重度估计结果以及融合不同尺度特征的严重度估计结果分别进行了实验，发现融合尺度下的严重度估计准确率更高。但随着融合尺度的增加，网络变得更复杂，准确率却只上升 0.01 个百分点，因此本文设计的 3 层尺度融合的网络严重度估计准确率为最佳。

（3）与农作物病害严重度估计的传统机器学习算法进行比较分析，相比于传统基于颜色和纹理特征的方式，RSTCNN 的分类准确率提高了 9 ~ 20 个百分点，显著提高了小麦叶片病害严重度等级的估计准确率。

参　考　文　献

[1] OKSANA K, ROMAN D, OLGA T, et al. Designing a methodological concept for the diagnosis of early development of the main wheat diseases pathogens[C]//BIO Web Conferences, 2020, 21: 00002.

[2] 雷雨, 韩德俊, 曾庆东, 等. 基于高光谱成像技术的小麦条锈病病害程度分级方法[J/OL]. 农业机械学报, 2018, 49(5): 226 - 232.
LEI Yu, HAN Dejun, ZENG Qingdong, et al. Grading method of disease severity of wheat stripe rust based on hyperspectral imaging technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 226 - 232. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20180526&journal_id = jcsam. DOI: 10.6041/j.issn.1000-1298.2018.05.026. (in Chinese)

[3] ZHAO J, FANG Y, CHU G, et al. Identification of leaf-scale wheat powdery mildew combining hyperspectral imaging and an SVM classifier[J]. Plants, 2020, 9: 936.

[4] 朱文静,陈华,李林,等. 基于红外热成像边缘检测算法的小麦叶锈病分级研究[J/OL]. 农业机械学报, 2019, 50(4): 36 - 41, 48.
ZHU Wenjing, CHEN Hua, LI Lin, et al. Grading of wheat leaf rust based on edge detection of infrared thermal imaging[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 36 - 41, 48. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20190404&flag = 1. DOI: 10.6041/j.issn.1000-1298.2019.04.004. (in Chinese)

[5] SWETHA V, JAYARAM R. A novel method for plant leaf malady recognition using machine learning classifiers[C]//2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2019: 1360 - 1365.

[6] 杨洋, 张亚兰, 苗伟, 等. 基于卷积神经网络的玉米根茎精确识别与定位研究[J/OL]. 农业机械学报, 2018, 49(10): 46 - 53.
YANG Yang, ZHANG Yalan, MIAO Wei, et al. Accurate identification and location of corn rhizome based on Faster R - CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(10): 46 - 53. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20190420&flag = 1. DOI: 10.6041/j.issn.1000-1298.2019.04.020. (in Chinese)

[7] PEI Y, HUANG Y, ZOU Q, et al. Effects of image degradation and degradation removal to CNN-based image classification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(4): 1239 - 1253.

[8] UMIT A, MURAT U, KERNAL A, et al. Plant leaf disease classification using EfficientNet deep learning model[J]. Ecological Informatics, 2021, 61: 101182.

[9] 杨森, 冯全, 张建华, 等. 基于深度学习与复合字典的马铃薯病害识别方法[J/OL]. 农业机械学报, 2020, 51(7): 22 - 29.
YANG Sen, FENG Quan, ZHANG Jianhua, et al. Identification method for potato disease based on deep learning and composite dictionary[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 22 - 29. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=20200703&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.07.003. (in Chinese)

[10] NIE X, WANG L, DING H, et al. Strawberry verticillium wilt detection network based on multi-task learning and attention [J]. IEEE Access, 2019, 7: 170003 - 170011.

[11] ENNADIFF E, LARABA S, VINCKE D, et al. Wheat diseases classification and localization using convolutional neural networks and GradCAM visualization[C]//2020 International Conference on Intelligent Systems and Computer Vision (ISCV), 2020: 1 - 5.

[12] BEUKERT U, THORWARTH P, ZHAO Y. Comparing the potential of marker-assisted selection and genomic prediction for improving rust resistance in hybrid wheat[J]. Frontiers in Plant Science, 2020, 11: 594113.

[13] SLADOJEVIE S, ARESENOVIE M, ANDERTA A, et al. Deep neural networks based recognition of plant disease by leaf image classification[J]. Computational Intelligence and Neuroscience, 2016,2016: 3289801.

[14] 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/OL]. 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/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 230 - 236, 253. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? file_no=2002025&flag=1. DOI:10.6041/j.issn.1000-1298.2020.02.025. (in Chinese)

[15] LIN Z, MU S, HUANG F, et al. A unified matrix-based convolutional neural network for fine-grained image classification of wheat leaf diseases[J]. IEEE Access, 2019, 7: 11570 - 11590.

[16] WANG G, SUN Y, WANG J. Automatic image-based plant disease severity estimation using deep learning[J]. Computational Intelligence and Neuroscience, 2017, 2017: 2917536.

[17] LIANG Q, XIANG S, HU Y, et al. PD2SE - Net: computer-assisted plant disease diagnosis and severity estimation network [J]. Computers and Electronics in Agriculture, 2019, 157: 518 - 529.

[18] JOSE G E, RENATO A K, JOSE A V. Deep learning for classification and severity estimation of coffee leaf biotic stress[J]. Computers and Electronics in Agriculture, 2020, 169: 105162.

[19] XIANG S, LIANG Q, SUN W, et al. L - CSMS: novel lightweight network for plant disease severity recognition[J]. Journal of Plant Diseases and Protection, 2021, 128: 557 - 569.

[20] FU J, ZHENG H L, MEI T. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 4438 - 4446.

[21] GE W, LIN X, YU Y. Weakly supervised complementary parts models for fine-grained image classification from the bottom up [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 3034 - 3043.

[22] ZHENG H, FU J, ZHA Z, et al. Looking for the devil in the details: learning trilinear attention sampling network for fine-grained image recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 5012 - 5021.

[23] WEI X, WANG P, LIU L, et al. Piecewise classifier mappings: learning fine-grained learners for novel categories with few examples[J]. IEEE Transactions on Image Processing, 2019, 28(12): 6116 - 6125.

[24] MAJID N, BABAR M, SABIH Q, et al. Structural rietveld refinement and magnetic features of prosademium (Pr) doped Cu nanocrystalline spinel ferrites[J]. Ceramics International, 2019, 45(8): 10187 - 10195.

[25] ZHENG H, FU J, MEI T, et al. Learning multi-attention convolutional neural network for fine-grained image recognition[C]// Proceedings of the IEEE International Conference on Computer Vision (ICCV),2017:5209 - 5217.

[26] ZHANG D Y, XIE X H, LIANG D, et al. Comparative analysis of identification methods for wheat leaf diseases[J]. International Agricultural Engineering Journal, 2016,25(3):197 - 203.

[27] ZENG Q, LI X, LIN H. Concat convolutional neural network for pulsar candidate selection[J]. Monthly Notices of the Royal Astronomical Society, 2020, 494(3): 3110 - 3119.

[28] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[J]. Advances in Neural Information Processing Systems, 2015, 1(2): 2017 - 2025.

[29] HUANG W, HUA G, YU Z, et al. Recurrent spatial transformer network for high-accuracy image registration in moving PCB defect detection[J]. The Journal of Engineering, 2020, 13(7): 438 - 443.

(下转第 263 页)