

基于双池化与多尺度核特征加权 CNN 的典型牧草识别

肖志云^{1,2} 赵晓陈^{1,2}

(1. 内蒙古工业大学电力学院, 呼和浩特 010080; 2. 内蒙古工业大学内蒙古机电控制重点实验室, 呼和浩特 010051)

摘要: 针对自然背景下牧草难识别的问题, 提出一种基于双池化与多尺度核特征加权的卷积神经网络牧草识别方法。双池化特征加权结构通过将卷积层输出的特征图分别进行最大值池化和均值池化得到两组特征图, 引入特征重标定策略, 依照各通道特征图对当前任务的重要程度进行加权, 以增强有用特征、抑制无用特征; 多尺度核特征加权结构通过在卷积层中同时使用 3×3 和 5×5 两种卷积核, 并将网络的前几层特征复用后进行加权, 以提高重要特征的利用率。对 10 类牧草图像进行识别实验, 结果表明, 该方法识别率为 94.1%, 比 VGG-13 网络提高了 5.7 个百分点, 双池化与多尺度特征加权有效提高了牧草识别精度。

关键词: 牧草识别; 卷积神经网络; 特征重标定

中图分类号: TP391.41 文献标识码: A 文章编号: 1000-1298(2020)05-0182-10

OSID:



Typical Forage Recognition Based on Double Pooling and Multi-scale Kernel Feature Weighted Convolution Neural Network

XIAO Zhiyun^{1,2} ZHAO Xiaochen^{1,2}

(1. College of Electric Power, Inner Mongolia University of Technology, Huhhot 010080, China

2. Inner Mongolia Key Laboratory of Mechatronic Control, Inner Mongolia University of Technology, Huhhot 010051, China)

Abstract: In order to solve the problem of forage recognition under natural conditions, a convolutional neural network method based on double-pooling feature weighting and multi-scale convolution kernel feature weighting structure was proposed. The spatial information and significance information of the image were fully utilized by using the dual-pooling feature weighted structure. Two groups of feature graphs were obtained by max-pooling and mean-pooling of feature graphs output from the convolution layer, and then these two groups of features were spliced. Finally, a feature re-calibration strategy was introduced to weight the importance of current tasks according to the feature graphs of each channel, so as to enhance useful features and suppress useless features. Image information was more fully mined by using multi-scale feature weighting structure. The 3×3 and 5×5 convolution kernels were used at the same time, and the features of the first several layers of the network were spliced with the features of the current layer to improve feature utilization rate. Feature re-calibration strategy was also introduced to weight features. The recognition experiments of ten pasture images showed that the recognition rate of the method was 94.1%, which was 5.7 percentage points higher than that of VGG-13 network, the double pooling and multi-scale feature weighting structure effectively improved the recognition accuracy.

Key words: forage recognition; convolutional neural network; feature recalibration

0 引言

内蒙古自治区是我国重要的草原畜牧业生产基地, 全区草场面积达 8 666.7 万 hm^2 , 占内蒙古土地面积的 76.5%, 约占全国天然草场面积的 $1/4$ ^[1]。深入研究牧草是合理利用草地资源、实现国家可持

续发展的重要途径。实现牧草生长趋势的精准检测、研究牧草物种多样性等都需要对牧草进行准确分类。目前, 牧草识别主要由人工完成, 费时、费力, 且精度较低。

随着机器视觉技术的发展, 众多研究者以叶片的颜色、纹理、形状等视觉特征研究植物叶片识别方

收稿日期: 2019-08-28 修回日期: 2019-12-10

基金项目: 国家自然科学基金项目(61661042)

作者简介: 肖志云(1974—), 男, 教授, 博士, 主要从事机器视觉在农业中的应用研究, E-mail: xiaozhiyun@imut.edu.cn

法^[2-7]。首先使用特定的图像分割算法将叶片与背景分离,然后将提取到的叶片的颜色、纹理、边缘等多个特征相融合的方式研究植物叶片识别方法,利用多种聚类和分类算法建立识别模型,获得了较好的识别结果。上述方法是从图像分割与叶片特定特征的提取出发,对植物叶片进行分类识别,而当存在光照变化、背景复杂、叶片重叠等自然因素影响时,图像分割效果会受到影响,提取到的特征也会发生较大变化。因此,基于图像分割与提取特征的识别方法在自然环境下效果不理想。

自 ALEX 等^[8]提出 AlexNet 卷积神经网络模型以来,卷积神经网络大量应用于图像分类研究中^[9-13]。文献[14]在对稻飞虱与非稻飞虱进行图像分类研究时,首先使用迁移学习在 ResNet-50 框架上训练,然后通过 Mask R-CNN 对其进行分类识别,平均识别精度达 92.3%;文献[15]在对花卉品种进行分类时,提出一种将卷积神经网络与随机森林相结合的方法,相比手工提取特征方法,有效提高了识别率。与手动提取特征的方法相比,基于深度卷积网络的图像识别方法无需图像分割,可自动提

取特征,且提取的特征不受光照变化、背景复杂等环境的影响,适于自然背景下的图像识别。

上述基于卷积神经网络的识别方法虽然提高了识别精度,但采用的都是一种池化方式,且对特征没有按重要性进行区分,不能充分利用图像所包含的信息。本文针对自然背景下牧草图像难识别的问题,充分利用牧草图像信息,设计一种基于双池化与多尺度核特征加权的卷积神经网络结构,以有效提高网络的识别精度。

1 样本集获取

通过两种方式获取牧草图像:首先在自然场景下使用相机对牧草图像进行采集;其次通过网络爬虫爬取互联网上的图像对数据集进行补充。通过对上述方法获得的牧草图像进行筛选后,共保留紫花苜蓿、红三叶、鸡眼草、紫云英等 10 类 2 110 幅牧草图像,且每幅牧草图像的背景各不相同,牧草区域有大有小,光照强弱也有不同。图 1 为本文需识别的 10 类牧草图像,表 1 为数据集中各类牧草的数量。

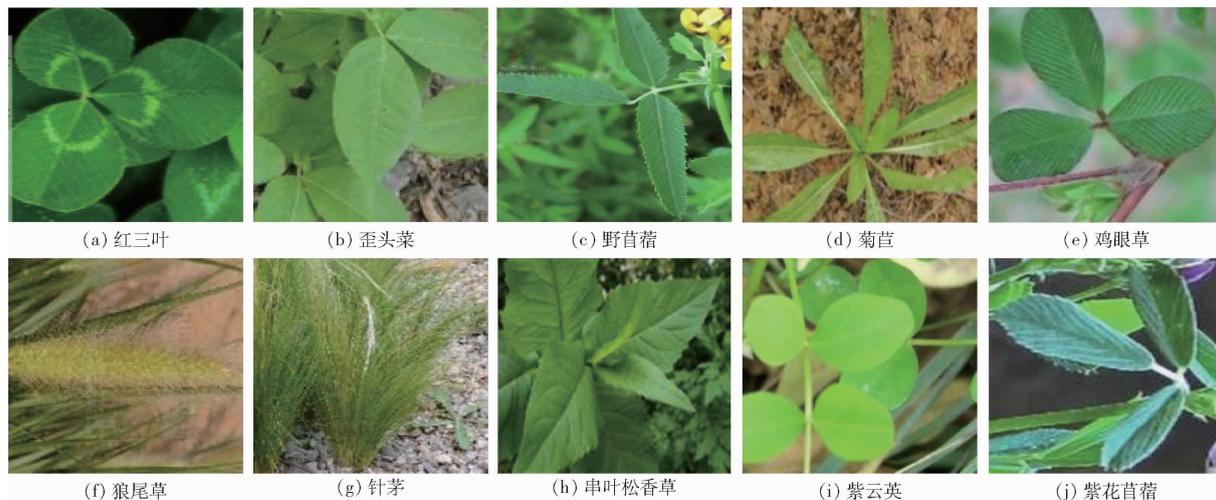


图 1 本文需识别的牧草种类

Fig. 1 Forage species images to be identified

表 1 样本集中各类样本数量

Tab. 1 Number of pastures in sample set

种类	红三叶	歪头菜	野苜蓿	菊苣	针茅
数量	194	265	227	181	141
种类	鸡眼草	狼尾草	紫花苜蓿	串叶松香草	紫云英
数量	218	281	205	194	204

2 卷积神经网络基本框架及优化

2.1 卷积神经网络基本框架

本文网络是在 VGG-Net^[16]结构基础上改进而来,VGGNet-13 结构如图 2 所示,图中,Conv 代表卷积核,Max Pooling 代表最大值池化,Avg Pooling

代表均值池化,FC 代表全连接层。网络采用了卷积层堆叠的结构且在每 2 个卷积层后接 1 个池化层,网络共包括 10 个卷积层和 5 个池化层,其中卷积核尺寸为 3×3,池化层采用 Max Pooling。

2.2 函数选取及网络优化

为提高网络对特征分布的拟合能力,卷积神经网络通常在卷积层和全连接层后引入非线性激活函数。常用的激活函数为 ReLU 激活函数,其表达式为

$$f(x) = \max(0, x) \tag{1}$$

式中 x ——输入特征值

$f(x)$ ——输出特征值

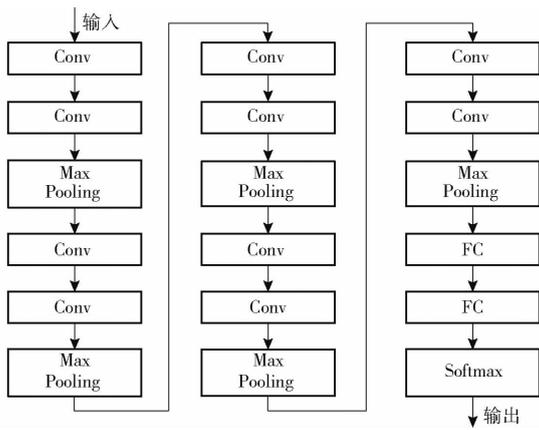


图2 VGGNet-13 结构

Fig. 2 Network layer parameter

由式(1)可知,当 $x < 0$ 时 ReLU 函数的输出为 0,训练时会造成神经元失活,因此本文网络的非线性激活函数采用 ELU^[17] 函数,其表达式为

$$f(x) = \begin{cases} x & (x > 0) \\ \alpha(e^x - 1) & (x \leq 0) \end{cases} \quad (2)$$

式中 α ——伸缩因子

α 控制 ELU 负值部分何时饱和,在本文中取值为 1。

由式(2)可知,ELU 函数在 $x \leq 0$ 部分是一个指数函数,因此不会出现神经元失活问题。

在最后 1 个池化层后引入了 Global Average Pooling^[18] 层,可以大量减少网络参数,有效地防止过拟合,甚至在很多网络中都使用它来替代全连接层。为提高网络非线性同时保证识别精度,本文并没有取消全连接层而是将其参数设为 512,之后是 Softmax 层用于对牧草图像分类。

采用 Adam 优化器最小化损失函数。由于数据集中的样本之间存在差异,必然存在网络容易学习的样本和难以学得的样本,因此采用 Focal loss^[19] 代替交叉熵作为网络的损失函数。以二分类为例,交叉熵表达式为

$$L = \begin{cases} -y \lg \hat{y} & (y = 1) \\ -(1 - y) \lg(1 - \hat{y}) & (y = 0) \end{cases} \quad (3)$$

式中 y ——样本标签 \hat{y} ——网络输出值

L ——交叉熵损失函数值

Focal loss 的表达式为

$$L_f = \begin{cases} -y(1 - \hat{y})^\gamma \lg \hat{y} & (y = 1) \\ -\hat{y}^\gamma(1 - y) \lg(1 - \hat{y}) & (y = 0) \end{cases} \quad (4)$$

式中 L_f ——Focal loss 值

γ ——指数调控因子

由式(4)可知,Focal loss 在交叉熵的基础上引入了 γ 因子,当 $\gamma > 0$ 时 \hat{y} 值越大的样本有更小的损失函数值,因此 Focal loss 可以使网络在训练时更加关注于难以学得的样本,经实验测定,当 $\gamma = 5$ 时效果最佳。

3 双池化与多尺度核特征加权结构

3.1 双池化特征加权结构

池化层是卷积神经网络的基本结构,池化可以在保留主要特征的同时降低特征维度,减少卷积神经网络的运算量,最常用的有 Max Pooling 和 Average Pooling。

Max Pooling 选取池化区域内最大值作为特征值,其表达式为

$$h_{\max} = \max_{i,j \in C} k(i,j) \quad (5)$$

式中 C ——池化窗口内 (i,j) 的取值集合

$k(i,j)$ ——池化窗口内位置为 (i,j) 的像素值

h_{\max} ——池化后的值

Average Pooling 选取池化区域内的平均值作为特征值,其表达式为

$$h_{\text{avg}} = \frac{1}{m^2} \sum_{i,j \in C} k(i,j) \quad (6)$$

式中 h_{avg} ——池化后的值

m ——池化窗口尺寸

研究表明 Max Pooling 能够更多地保留图像纹理等显著性信息^[20],而 Average Pooling 则更多地保留了图像的局部空间信息^[21]。目前大多数卷积神经网络只选取一种池化方式,会造成更大的信息损失,为充分利用特征信息可以同时使用这两种池化方法。

结合 SeNet(Squeeze-and-excitation networks)^[22] 提出了一种双池化卷积神经网络结构,其结构如图 3

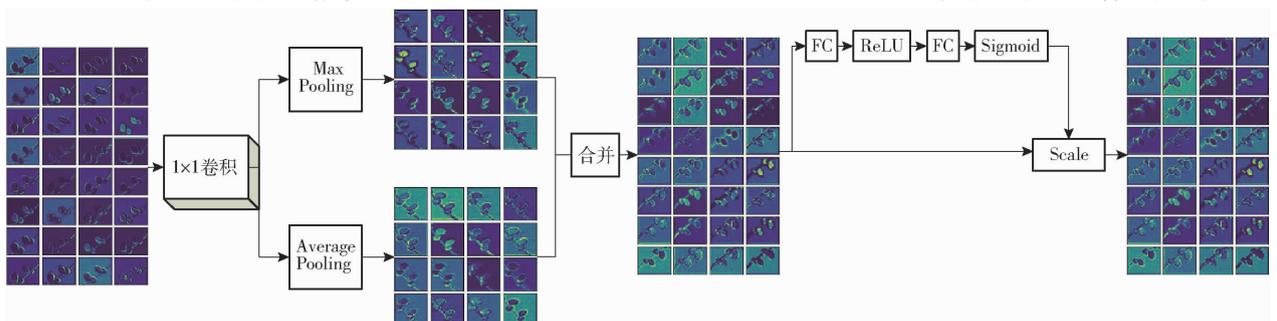


图3 双池化结构

Fig. 3 Double pool structure

所示。首先对上一个卷积层输出的特征图进行 1×1 卷积操作,将特征图的个数降为原来的 $1/2$,然后对降维后的特征图分别进行最大池化和平均池化操作得到两组特征图,之后将这两组特征图进行拼接。

图 4 为两种池化方式单独使用时测试集的认识率随训练次数的变化曲线,从图中可知,两种池化对识别结果的影响不同,因此将拼接后的特征图直接输入到下一个卷积层不合理。为解决这个问题,引入 SeNet 中的特征重标定策略,即通过学习的方式自动获取每个特征的重要程度,然后依照这个重要程度对特征进行加权,从而提升有用的特征并抑制对当前任务作用较小的特征。

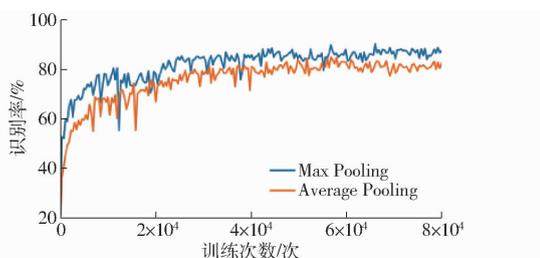


图 4 两种池化方式识别率对比

Fig. 4 Accuracy rate graph of two different pooling methods

如图 3 右半部分所示,首先对拼接后特征图进行 Global Average Pooling 操作,将 $H \times H \times C$ 的特征图变换为 $1 \times 1 \times C$ 的向量,其中 H 为特征图的尺寸, C 为特征图的个数。然后将此向量先后输入到 2 个全连接层,第 2 个全连接层输出向量中的元素即为池化层拼接后对应特征图的权重,之后将该权重与原特征图相乘得到新的特征图,最后将新特征图输入到下一个卷积层。在此过程中引入全连接层的目的在于使网络能够在训练过程中通过学习不断地调整特征图的权重,直到其最大程度地满足于当前任务。

如图 5 为训练过程中权重分布图,由图可知,在训练开始阶段权重分布变化幅度比较大,随着训练的进行权重分布逐渐趋于稳定,说明在本文双池化结构中引入特征重标定策略后其权重在网络训练过

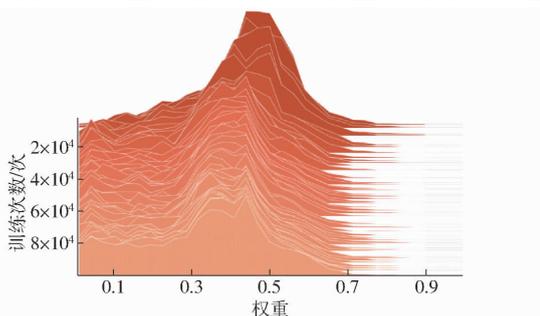


图 5 网络训练过程中权重分布变化

Fig. 5 Weight distribution change diagram during network training

程中不断调整,直到其收敛于最优权重。

网络训练完成后,若最大池化层产生的特征图的最小权重大于平均池化计算出的特征图的最大权重,则说明相对最大池化,平均池化层产生的特征图对当前任务的贡献很小,这种情况下引入平均池化将失去意义。

表 2 为网络训练完成输入 1 幅图像后,提取的第 1 个双池化结构中 Max Pooling 和 Average Pooling 各特征图的权重,即各特征图对当前任务的重要程度,结合图 4 可知,虽然在 2 种池化单独使用时 Max Pooling 的整体效果优于 Average Pooling,但并不是 Max Pooling 中所有特征图对当前任务的重要程度都高于 Average Pooling,说明本文方法是可行的。

表 2 两种池化方式所得各特征图权重

Tab.2 Weight values of each feature map obtained by two pooling methods

特征图序号	权重	
	Max Pooling	Average Pooling
1	0.462 832	0.030 5303
2	0.070 543	0.436 524
3	0.602 405	0.299 334
4	0.380 751	0.117 953
5	0.328 779	0.397 692
6	0.263 189	0.400 007
7	0.331 436	0.241 204
8	0.453 233	0.293 562
9	0.117 131	0.247 180
10	0.381 899	0.420 472
11	0.326 777	0.502 506
12	0.143 810	0.326 579
13	0.373 160	0.334 339
14	0.538 489	0.409 004
15	0.092 623	0.392 140
16	0.136 549	0.534 684

3.2 多尺度核特征加权结构

在卷积神经网络中,不同尺寸的卷积核对图像的感受野是不同的,小尺寸的卷积核与图像进行卷积操作得到的特征能更好地反映图像的局部特征,大尺寸的卷积核则能更好地反映图像的全局特征^[23],为增强网络适应性,本文使用两种尺寸卷积核的组合来对牧草进行识别,但网络的运算量是随着卷积核尺寸的增大而增大的,卷积核太大则会造成海量的运算量。因此本文只对比了尺寸为 3×3 、 5×5 、 7×7 的卷积核与 3×3 卷积核组合后对网络识别率的影响,实验结果显示 3 种组合的识别率分别为 92.76%、94.10%、92.81%。

由以上结论可知,在 3×3 卷积核中引入 5×5 的卷积核后识别率提升了 1.34 个百分点,而在

3×3卷积核中引入7×7的卷积核后网络识别率提升并不明显,因此本文使用3×3卷积核与5×5卷积核组成Conv block结构对牧草图像进行识别。

为充分利用特征信息,引入了Densenet^[24]中的特征复用思想,即网络任意两层之间都有直接连接,每一层的输入都是前面所有层输出的并集,而该层所学习的特征图也会被直接传递给其后面所有层作为输入。其表达式为

$$x_l = H([x_0, x_1, x_2, \dots, x_{l-1}]) \quad (7)$$

式中 x_l ——第 l 层特征图

[]——拼接运算

$H(\cdot)$ ——ELU 激活函数

这样在提高特征利用率的同时减少了参数量^[25];网络更容易训练,并且稠密连接的正则化作用在少量的训练集中能够有效降低过拟合^[26];加速卷积神经网络反向传播时梯度的传递,在一定程度上

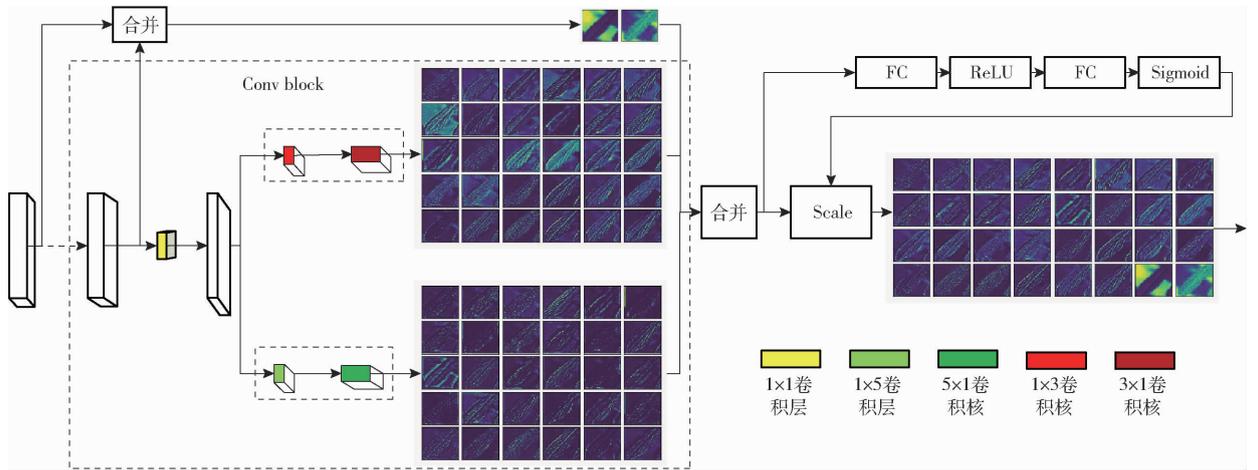


图6 多尺度特征加权结构图

Fig.6 Multi-scale feature weighted structure

图7为网络训练完成后提取的各层输出的特征图,由于相邻2层提取的特征图区别较小,因此每2个Conv block模块提取1次特征图。由图可以看出,网络Conv block2模块输出的特征图包含了牧草图像空间、颜色等混合信息,但图像的背景部分仍然存在;Conv block4和Conv block6的特征图更多地包含了牧草图像的纹理和轮廓等不变性特征,而图像背景部分已经基本消除;Conv block8层则更多保留了各类牧草之间有区分性的特征,而Conv block10输出的特征图则代表了与类别相关的信息。以上分析表明,随着网络层数的加深,特征的抽象程度和稀疏程度也越来越高,且提取到的图像特征主要是牧草特征而非背景,因此本文网络能够有效提取牧草特征。

上缓解了随着网络的加深而产生的梯度消失问题和模型退化问题。上述Conv block和复用特征组成的结构如图6所示,具体操作如下。

Conv block操作:首先将上一层得到的特征图输入到1×1卷积层,目的是减少特征图的数量,以减少运算量;然后将降维后的特征图分别输入到3×3卷积层和5×5卷积层,同样为减少计算量将3×3卷积核分解为1×3卷积核和3×1卷积核,5×5卷积核分解为1×5卷积核和5×1卷积核。

特征复用操作:将经上一个Pooling block层或输入层输出的特征图和与当前Conv block之间所有Conv block输出的特征图进行拼接,之后经过一组1×1的卷积核进行降维。

经过上述操作后,共产生3组特征图,之后的操作是将这3组特征图进行拼接。同样,在此也引入特征重标定策略,对拼接后的特征图进行加权以增强有用特征,抑制无用特征。

4 实验与结果分析

对紫花苜蓿、针茅、狼尾草、紫云英等10种典型牧草进行识别研究,所选软件平台为Anaconda3中的spyder,使用Python语言和Tensorflow深度学习框架进行图像预处理和卷积神经网络的搭建;硬件平台为计算机,处理器为IntelCore i5-8400,GPU为GTX1060,6GB显存。

4.1 数据集扩充与图像预处理

首先将获得的样本尺寸调整到128像素×128像素,然后按4:1的比例将初始数据集随机分为训练集和测试集。由于卷积神经网络参数较多,在训练过程中极易发生过拟合,因此在对网络训练之前需要对初始样本的训练集进行扩充。常用的数据扩充方法有平移、旋转、亮度变换、添加噪声等。数据扩充应结合具体研究目标的特点,不能随意扩充,

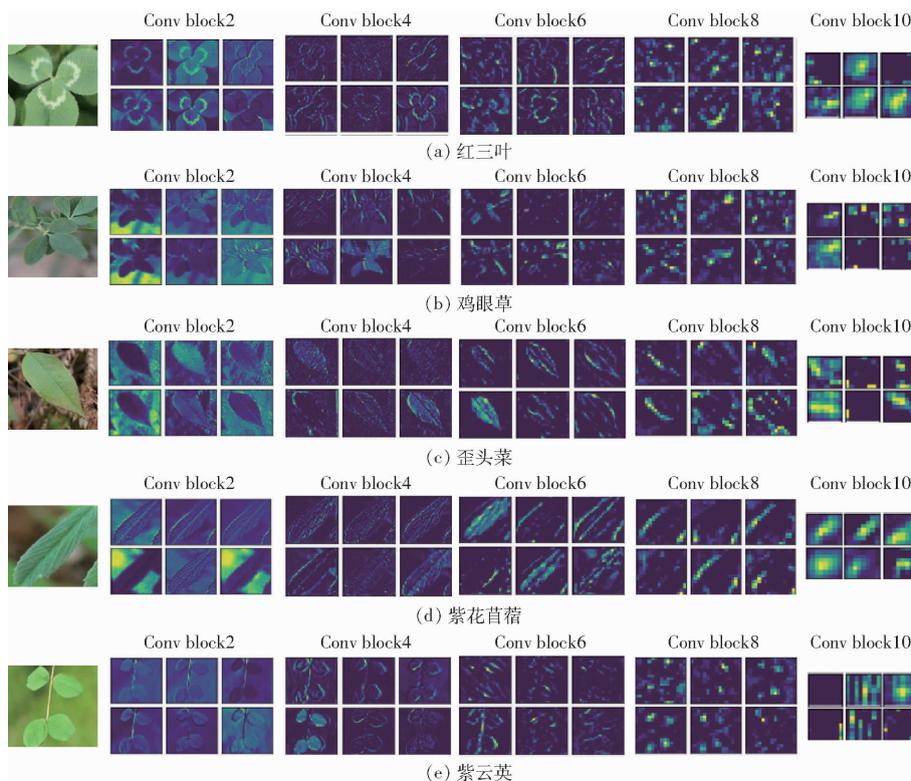


图 7 网络各层提取的特征图

Fig. 7 Feature map extracted from each layer of network

例如形状是牧草主要表征之一,在数据扩充时不应改变牧草形状。另外扩充的图像与原图像之间应有一定差距,但同时要尽可能地保留原始图像中牧草的相关特征。例如在对图像进行平移时,若平移的像素点太少,则平移后的图像与原图无异,在这种情况下反而会加重网络过拟合,若平移过多则会丢失较多的牧草信息。因此在对图像进行平移变换时,将平移的范围限制在 10~30 像素,在旋转变换时只进行了 90°、180°、270° 的旋转,其他数据扩充方式也遵

循以上思想。数据扩充分 3 次进行:第 1 次,通过旋转、平移、裁剪和填充将数据扩充为原训练集的 7 倍。第 2 次,考虑到在自然条件下光照强度对识别结果影响较大,因此对上述扩充后的图像进行亮度变换扩充至原训练集的 14 倍。第 3 次,对原始训练集进行添加噪声和水平镜像变换。最终扩充后的训练集是原训练集的 16 倍,且每次扩充后网络识别率均有提升。对样本集进行编号分类并制作标签,扩充后各类牧草数量及各类牧草的标签如表 3 所示。

表 3 扩充后的训练集中各类样本数量

Tab. 3 Number of samples in expanded training set

样本	紫花苜蓿	串叶松香草	野苜蓿	狼尾草	菊苣	歪头菜	红三叶	针茅	鸡眼草	紫云英	总计
标签	0	1	2	3	4	5	6	7	8	9	
数量	2 624	2 464	2 912	3 616	2 320	3 376	2 476	1 810	2 798	2 624	27 020

在使用卷积神经网络对图像进行处理之前首先对图像进行标准化操作,表达式为

$$X^* = \frac{X - \mu}{\sigma} \quad (8)$$

式中 X ——图像样本 μ ——样本均值
 σ ——样本标准差
 X^* ——标准化的样本

采用 Adam 优化器对网络进行训练,批处理容量为 60,采用指数衰减的学习率,基础学习率 0.002 2,学习率衰减系数为 0.99, L_2 正则化系数为 0.000 7,模型迭代次数为 100 000。

4.2 实验结果分析

图 8 为使用 VGG-13 对牧草数据集训练时的相关曲线,由图 8a 可知,训练集识别率在训练 10 000 次时已经达到 100%,而测试集识别率仅为 70%。由图 8b 可知,训练集上的损失值在训练 50 000 次后已经接近 0,而测试集上的损失值在 0.3 附近,由此得知在直接将 VGG-13 在牧草图像数据集上训练时会发生过拟合,为验证本文网络的有效性,本文对 VGG-13 网络各层参数进行了调整,使其适应于牧草数据集。

图 9 为调整后的 VGG-13 与本文网络在测试

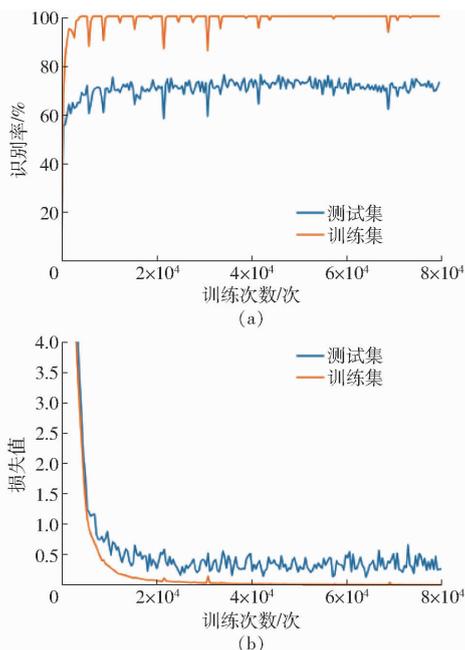


图8 VGG-13在牧草图像数据集上相关曲线

Fig. 8 Correlation curves of VGG-13 on pasture image dataset

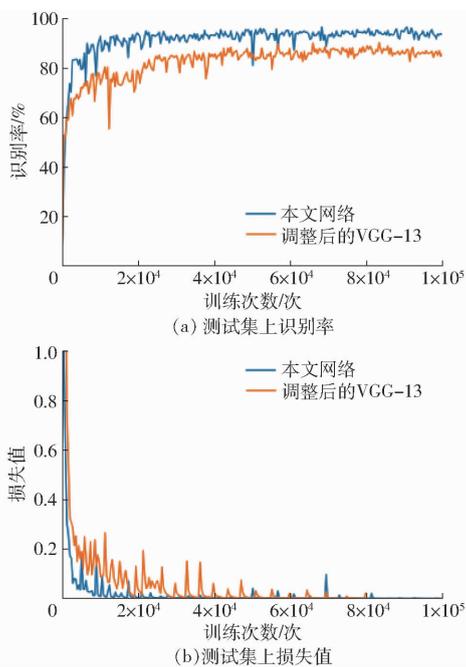


图9 本文网络与调整后的VGG-13相关曲线

Fig. 9 Comparison of network and adjusted VGG-13 correlation curves

集上的识别率和在训练集上的损失值曲线对比图,结合图8a知,调整后的VGG-13相比调整前识别率有了明显提高,而相比调整后的VGG-13,本文网络的识别率更高。由图9b可知,本文网络的损失值在迭代20000次时已经接近0,而调整后的VGG-13在迭代近40000次才接近0,因此本文网络明显收敛更快。

为验证本文方法的有效性,将调整后的VGG-

13、调整后的VGG-13+双池化结构和本文网络在单类样本测试集上进行了对比,其识别率如表4所示。由表4可得知,调整后的VGG-13结构只有5类牧草识别率超过90%;VGG-13+双池化结构有8类牧草图像识别率超过90%,其中有3类牧草地识别率超过了95%,紫云英的识别率达到100%;本文网络除鸡眼草外,识别率均超过了90%,其中有5类牧草超过了95%,针茅的识别率达到100%。由以上分析可知,3种网络中本文网络在对单类样本识别中表现最好,其次是调整后的VGG-13+双池化结构,调整后的VGG-13网络表现最差。结果表明本文提出的牧草识别方法在单类样本上是有效的。

表4 各网络在训练集中各类样本上的识别率

Tab. 4 Recognition rates of various networks on various samples in training set

样本	数量	网络识别率/%		
		调整后的VGG-13	调整后的VGG-13+双池化特征加权结构	本文网络
紫花苜蓿	41	80.05	92.34	90.24
串叶松香草	40	85.10	91.68	90.26
野苜蓿	45	94.34	95.72	98.81
狼尾草	55	87.92	98.34	98.81
菊苣	36	90.49	87.17	93.35
歪头菜	54	87.60	85.51	90.70
红三叶	39	95.86	93.16	98.34
针茅	28	95.10	93.82	100
鸡眼草	43	83.60	90.02	85.27
紫云英	40	94.68	100	96.44

为进一步分析识别率差异的原因,本文提取了部分分类正确和分类错误的图像类激活热力图,分别如图10、11所示。

由图10可知,对于能够正确分类的样本,其对分类有贡献的区域主要集中在牧草所在区域,尤其是与各牧草有显著差异的部分,如红三叶图像中“V”形白色纹理部分,紫云英叶基部分,紫花苜蓿的叶尖部分。

由图11可知,对于错分样本,可以分为3种情况:第1种情况是图像对分类起作用的区域是背景而不是牧草本身,如图11b、11e、11f、11g、11i,原因是在原图像中背景叶片的特征更为明显,网络在分类时误将背景叶片作为目标叶片,进而导致错分。第2种情况是图像中牧草叶片个数过多,单个叶片区域过小,使网络无法提取有效特征,如图11a、11h。第3种情况是图像对分类起作用的区域是牧草叶片部分,但却将类别错分,如图11c、11d,原因是原图像中牧草区域专有特征不明显,对分类起作用的区域与其他牧草相似,从而导致分类错误。由以

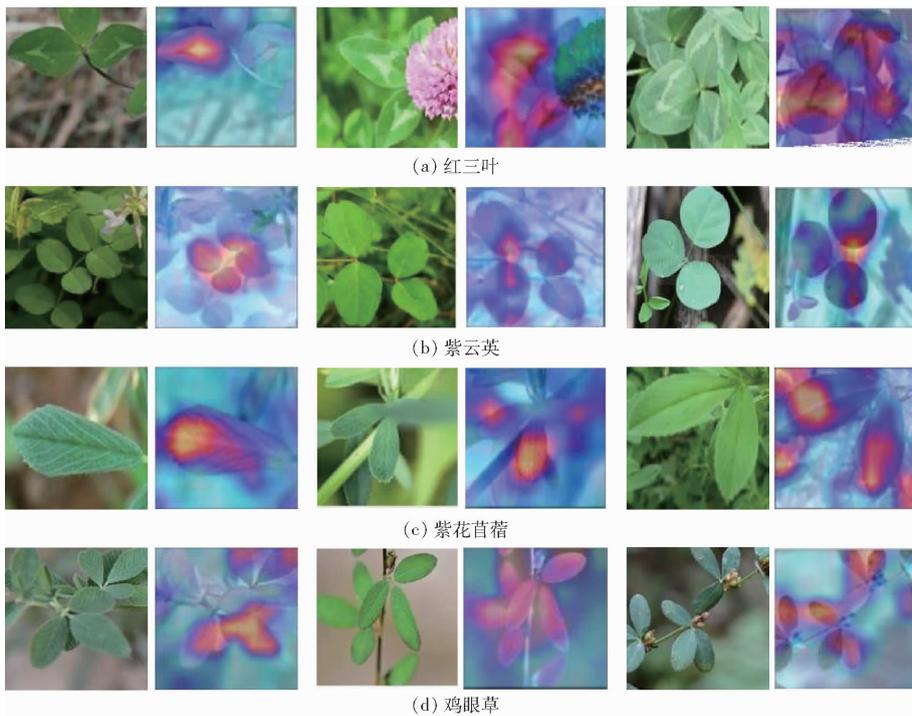


图 10 分类正确图像的类激活热力图

Fig. 10 Class activation thermal maps of correct image classification

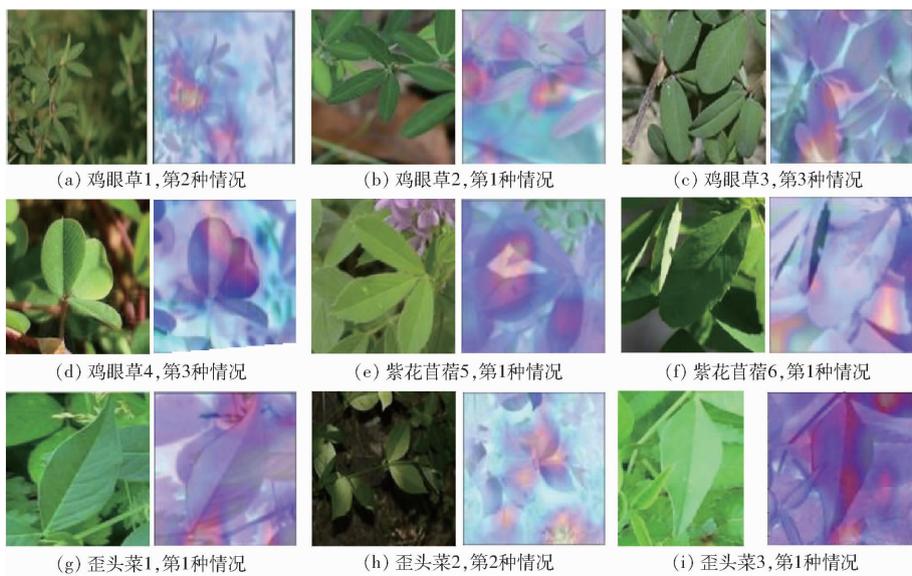


图 11 分类错误图像的类激活热力图

Fig. 11 Class activation thermal map of classification error image

上分析可知,歪头菜识别率较低主要由第 1 和第 2 种情况造成,紫花苜蓿则由第 1 种情况造成,而在鸡眼草被错分的样本中均存在这 3 种情况。

表 5 为各网络在单幅图像运行时间和在测试集上平均识别率对比。由表 5 可知,VGG-13 网络的识别率最低且所用时间最长。本文网络比调整后的 VGG-13 网络在训练集上的平均识别率高 5.7 个百分点,但运行时间却只多了 0.035 s,证明本文提出的方法是有效可行的。网络各层参数设置如

表 6 所示。

表 5 各网络测试集平均识别率与运行时间对比
Tab. 5 Comparison of average recognition rate and running time of each network test set

网络	识别单幅图像 所用时间/s	识别 率/%
VGG-13	0.358	74.00
调整后的 VGG-13	0.092	88.40
调整后的 VGG-13 + 双池化结构	0.106	92.76
本文网络	0.127	94.10

表6 网络各层参数设置
Tab.6 Network layer parameters

网络层	参数		网络层	参数	
Conv block1	Conv3 × 1,16	Conv5 × 1,16	Pooling block	Conv1 × 1,32	
	Conv1 × 3,16	Conv1 × 5,16		Max Pooling	Avg Pooling
Conv block2	Conv1 × 1,16		Conv block7	Conv1 × 1,32	
	Conv3 × 1,16	Conv5 × 1,16		Conv3 × 1,30	Conv5 × 1,30
	Conv1 × 3,16	Conv1 × 5,16		Conv1 × 3,30	Conv1 × 5,30
pooling block	Conv1 × 1,16		Conv block8	Conv1 × 1,32	
	Max Pooling	Avg Pooling		Conv3 × 1,30	Conv5 × 1,30
Conv block3	Conv1 × 1,16		Pooling block	Conv1 × 1,32	
	Conv3 × 1,16	Conv5 × 1,16		Conv3 × 1,30	Conv5 × 1,30
Conv block4	Conv1 × 1,16		Conv block9	Conv1 × 1,32	
	Conv1 × 3,16	Conv1 × 5,16		Conv3 × 1,60	Conv5 × 1,60
Conv block4	Conv3 × 1,16	Conv5 × 1,16	Conv block10	Conv1 × 3,60	Conv1 × 5,60
	Conv1 × 3,16	Conv1 × 5,16		Conv1 × 1,60	
Pooling block	Conv1 × 1,16		Conv block10	Conv3 × 1,60	Conv5 × 1,60
	Max Pooling	Avg Pooling		Conv1 × 3,60	Conv1 × 5,60
Conv block5	Conv1 × 1,32		Global Average Pooling		8 × 8
	Conv3 × 1,30	Conv5 × 1,30	FC		256
Conv block6	Conv1 × 3,30	Conv1 × 5,30	FC		256
	Conv1 × 1,32		Softmax		10
Conv block6	Conv3 × 1,30	Conv5 × 1,30			
	Conv1 × 3,30	Conv1 × 5,30			

5 结论

(1)提出的方法能够更充分地利用图像包含的信息,且在测试集上的平均识别率达到了94.1%,明显优于VGG-13网络。

(2)提出的网络结构在训练时,网络的收敛速

度是VGG-13网络的2倍以上,说明本文网络可以大幅提高网络收敛速度。

(3)对10类牧草进行单独识别时,本文提出的方法识别率优于VGG-13网络,且有9类牧草识别率超过90%,其中有5类牧草识别率在95%以上,证明本文所提出的方法能够对牧草图像进行准确识别。

参 考 文 献

- [1] 徐玲玲,钱拴. 内蒙古天然草场牧草产量的气象估算模型[J]. 中国草地学报,2011,33(1):54-58.
XU Lingling, QIAN Shuan. Meteorological estimation model for forage yield in natural grassland in Inner Mongolia[J]. Chinese Journal of Grassland, 2011, 33(1):54-58. (in Chinese)
- [2] WANG H, LI G, MA Z, et al. Application of neural networks to image recognition of plant diseases [C] // International Conference on Systems and Informatics, 2012:2159-2164.
- [3] ISNANTO R R, ZAHRA A A, JULIETTA P. Pattern recognition on herbs leaves using region-based invariants feature extraction [C] // 2016 3rd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). Los Alamitos: IEEE Computer Society Press, 2016:455-459.
- [4] 阎庆,梁栋,张晶晶. 基于Fisher变换的植物叶片图像识别监督LLE算法[J/OL]. 农业机械学报,2012,43(9):179-183.
YAN Qing, LIANG Dong, ZHANG Jingjing. Recognition method of plant leaves based on Fisher projection-supervised LLE algorithm[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(9):179-183. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20120933&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2012.09.033. (in Chinese)
- [5] LUKIC M, TUBA E, TUBA M. Leaf recognition algorithm using support vector machine with Hu moments and local binary patterns [C] // IEEE International Symposium on Applied Machine Intelligence & Informatics. Los Alamitos: IEEE Computer Society Press, 2017:000485-000490.
- [6] 韩丁,武佩,张强,等. 基于颜色矩的典型草原牧草特征提取与图像识别[J]. 农业工程学报,2016,32(23):168-175.
HAN Ding, WU Pei, ZHANG Qiang, et al. Feature extraction and image recognition of typical grassland forage based on color moment[J]. Transactions of the CSAE, 2016, 32(23):168-175. (in Chinese)
- [7] 肖志云,刘洪. 马铃薯典型病害图像自适应特征融合与快速识别[J/OL]. 农业机械学报,2017,48(12):26-32.
XIAO Zhiyun, LIU Hong. Adaptive features fusion and fast recognition of potato typical disease images[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(12):26-32. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20171203&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.12.003. (in Chinese)

- [8] ALEX K, ILYA S, GEOFFREY E. ImageNet classification with deep convolutional neural networks[C] // Advances in Neural Information Processing Systems 25, vol. 2: 26th Annual Conference on Neural Information Processing Systems 2012, Lake Tahoe, Nevada, USA. 2012; 1097 - 1105.
- [9] 高震宇, 王安, 刘勇, 等. 基于卷积神经网络的鲜茶叶智能分选系统研究[J/OL]. 农业机械学报, 2017, 48(7): 53 - 58. GAO Zhenyu, WANG An, LIU Yong, et al. Intelligent fresh-tea-leaves sorting system research based on convolution neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(7): 53 - 58. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170707&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.07.007. (in Chinese)
- [10] 肖雄, 王健翔, 张勇军, 等. 一种用于轴承故障诊断的二维卷积神经网络优化方法[J]. 中国电机工程学报, 2019, 139(15): 4558 - 4568. XIAO Xiong, WANG Jianxiang, ZHANG Yongjun, et al. A two-dimensional convolutional neural network optimization method for bearing fault diagnosis[J]. Proceedings of the CSEE, 2019, 39(15): 4558 - 4568. (in Chinese)
- [11] 尼加提·卡斯木, 师庆东, 刘素红, 等. 基于卷积网络的沙漠腹地绿洲植物群落自动分类方法[J/OL]. 农业机械学报, 2019, 50(1): 217 - 225. NIJAT Kasim, SHI Qingdong, LIU Suhong, et al. Automatic classification method of oasis plant community in desert hinterland based on VGGNet and ResNet models[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(1): 217 - 225. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190124&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.01.024. (in Chinese)
- [12] HU J, CHEN Z, YANG M, et al. A multi-scale fusion convolutional neural network for plant leaf recognition[J]. IEEE Signal Processing Letters, 2018, 25(6): 853 - 857.
- [13] 毕松, 高峰, 陈俊文, 等. 基于深度卷积神经网络的柑橘目标识别方法[J/OL]. 农业机械学报, 2019, 50(5): 181 - 186. BI Song, GAO Feng, CHEN Junwen, et al. Detection method of citrus based on deep convolution neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 181 - 186. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190521&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.05.021. (in Chinese)
- [14] 林相泽, 朱赛华, 张俊媛, 等. 基于迁移学习和 Mask R - CNN 的稻飞虱图像分类方法[J/OL]. 农业机械学报, 2019, 50(7): 201 - 207. LIN Xiangze, ZHU Saihua, ZHANG Junyuan, et al. Rice planthopper image classification method based on transfer learning and Mask R - CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 201 - 207. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190721&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.07.021. (in Chinese)
- [15] GOGUL I, KUMAR V S. Flower species recognition system using convolution neural networks and transfer learning[C] // 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN). Los Alamitos: IEEE Computer Society Press, 2017: 1 - 6.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C] // Proceedings of International Conference on Learning Representations 2015. San Diego: ICLR, 2015: 463 - 476.
- [17] CLEVERT D A, DJROK-ARNE, UNTERTHINER T, et al. Fast and accurate deep network learning by exponential linear units (ELUs)[C] // International Conference on Learning Representations, 2016: 1 - 14.
- [18] LIN M, CHEN Q, YAN S. Network in network[C] // International Conference on Learning Representations, 2014: 1 - 10.
- [19] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C] // 2017 IEEE International Conference on Computer Vision (ICCV), 2017: 2999 - 3007.
- [20] 时增林, 叶阳东, 吴云鹏, 等. 基于序的空间金字塔池化网络的人群计数方法[J]. 自动化学报, 2016, 42(6): 866 - 874. SHI Zenglin, YE Yangdong, WU Yunpeng, et al. Crowd counting using rank-based spatial pyramid pooling network[J]. Acta Automatica Sinica, 2016, 42(6): 866 - 874. (in Chinese)
- [21] 孙俊, 谭文军, 毛罕平, 等. 基于改进卷积神经网络的多种植物叶片病害识别[J]. 农业工程学报, 2017, 33(19): 209 - 215. SUN Jun, TAN Wenjun, MAO Hanping, et al. Recognition of multiple plant leaf diseases based on improved convolutional neural network[J]. Transactions of the CSAE, 2017, 33(19): 209 - 215. (in Chinese)
- [22] JIE H, LI S, ALBANIE S, et al. Squeeze-and-excitation networks[C] // Proc. of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2018: 7132 - 7141.
- [23] 余烨, 金强, 傅云翔, 等. 基于 Fg - CarNet 的车辆型号精细分类研究[J]. 自动化学报, 2018, 44(10): 1864 - 1875. YU Ye, JIN Qiang, FU Yunxiang, et al. Fine-grained classification of car models using FG - CarNet convolutional neural network[J]. Acta Automatica Sinica, 2018, 44(10): 1864 - 1875. (in Chinese)
- [24] HUANG G, LIU Z, MAATEN L V D, et al. Densely connected convolutional networks[C] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Los Alamitos: IEEE Computer Society Press, 2017: 2261 - 2269.
- [25] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述[J]. 计算机科学, 2019, 46(3): 63 - 73. CHEN Chao, QI Feng. Review on development of convolutional neural network and its application in computer vision[J]. Computer Science, 2019, 46(3): 63 - 73. (in Chinese)
- [26] 熊鹏, 汤宝平, 邓蕾, 等. 基于动态加权密集连接卷积网络的变转速行星齿轮箱故障诊断[J]. 机械工程学报, 2019, 55(7): 52 - 57. XIONG Peng, TANG Baoping, DENG Lei, et al. Fault diagnosis for planetary gearbox by dynamically weighted densely connected convolutional networks[J]. Journal of Mechanical Engineering, 2019, 55(7): 52 - 57. (in Chinese)