

本文引用:刘伟,邹伟红,卢彦杰,胡为,刘塔斯. 基于迁移学习和模型融合的中药材图像识别研究[J]. 湖南中医药大学学报,2022,42(5): 809-814.

# 基于迁移学习和模型融合的中药材图像识别研究

刘伟<sup>1</sup>,邹伟红<sup>1</sup>,卢彦杰<sup>1</sup>,胡为<sup>1</sup>,刘塔斯<sup>2\*</sup>

(1.湖南中医药大学信息科学与工程学院,湖南长沙 410208;2.湖南中医药大学药学院,湖南长沙 410208)

**[摘要]** **目的** 为了提高中药材图像识别的准确率,针对中药材形状不规则、纹理特征细微、种类繁多等特点,结合深度学习方法开展中药材图像识别研究。**方法** 通过爬虫获取中药材图像并进行数据预处理,构建中药材图像数据集,使用 Xception、DenseNet 作为主干网络提取中药材特征,通过迁移学习、数据增强、模型融合等方法优化网络结构,并提出一种基于 Xception 和 DenseNet 融合的中药材图像识别模型 DxFusion。**结果** 通过迁移学习、数据增强和模型融合,DxFusion 在 60 种中药材测试集上分类精度可达 99.65%,效果优于已有分类模型。**结论** 基于迁移学习和多模型融合的深度卷积神经网络可以提取图像中不规则的中药材特征,有效提升中药材图像识别的准确率。

**[关键词]** 深度学习;中药材识别;卷积神经网络;迁移学习;数据增强;模型融合

**[中图分类号]** R282;TP751

**[文献标志码]** A

**[文章编号]** doi:10.3969/j.issn.1674-070X.2022.05.019

## Research on image recognition of Chinese medicinal materials based on transfer learning and model fusion

LIU Wei<sup>1</sup>, ZOU Weihong<sup>1</sup>, LU Yanjie<sup>1</sup>, HU Wei<sup>1</sup>, LIU Tasi<sup>2\*</sup>

(1. School of Information Science and Engineering, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China;

2. School of Pharmacy, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China)

**[Abstract]** **Objective** In order to improve the accuracy of Chinese medicinal materials image recognition, the research of Chinese medicinal materials image recognition is carried out based on deep learning according to the characteristics of irregular shape, subtle texture characteristics, and various types. **Methods** Images of Chinese medicinal materials were obtained through crawlers and data preprocessing was performed, image data sets of Chinese medicinal materials were established, Xception and DenseNet was used as the backbone network to extract the characteristics of Chinese medicinal materials, and the network structure was optimized by transfer learning, data augmentation and model fusion. A method called DxFusion for image recognition of Chinese medicinal materials based on the fusion of Xception and DenseNet was proposed. **Results** Through transfer learning, data augmentation and model fusion, the classification accuracy of DxFusion was reached 99.65% on 60 kinds of Chinese medicinal materials test sets, which was better than the existing classification models. **Conclusion** The deep convolutional neural network based on transfer learning and multi-model fusion can extract the irregular characteristics of Chinese medicinal materials in the image, and effectively improve the accuracy for the image recognition of Chinese medicinal materials.

**[Keywords]** deep learning; recognition of Chinese medicinal materials; convolutional neural network; transfer learning; data augmentation; model fusion

**[收稿日期]** 2021-10-26

**[基金项目]** 湖南省重点研发计划项目(2018SK2112);湖南省教育厅科学研究项目(20C1435)。

**[第一作者]** 刘伟,男,博士,副教授,研究方向:中医药信息学和人工智能。

**[通信作者]** \*刘塔斯,女,教授,E-mail:liutasi@126.com。

中药材是中医药事业传承和发展的物质基础。随着科学技术的飞速发展,计算机技术已应用于中药材的溯源、鉴定、质量控制。近年来,人工智能在自然语言处理、计算机视觉等领域均取得了重要突破。

中药材图像识别是通过图像处理技术对中药材进行识别,可应用于中药材自动分拣和快速鉴定等环节。在中药材领域,深度学习相关研究不断涌现。庄奕珊<sup>[1]</sup>使用 SqueezeNet、GoogLeNet 集成学习提取中药材特征进行分类;黄方亮等<sup>[2]</sup>使用 AlexNet 模型应用于 5 种中草药分类;王寒迎<sup>[3]</sup>使用 ZCA 白化减少中药材微性状图像间的冗余性,改进卷积神经网络提取特征;吴冲等<sup>[4]</sup>使用 YOLO 提取图像中的中药饮片,使用 ResNet50 和 VGG16 提取中药饮片特征然后进行分类;史婷婷等<sup>[5]</sup>在 GoogLeNet 网络增加 2 个辅助 Softmax 进行仿野生种植金银花遥感识别;陶益等<sup>[6]</sup>使用卷积神经网络和深度自动编码器预测干燥黑枸杞中总花色苷、总黄酮和总酚类物质含量。

由于中药材具有形状不规则、纹理特征细微、种类繁多等特点,使用传统机器学习分类算法精度不佳。因此,本文使用 MobileNet、Inception、Xception、ResNet、DenseNet 卷积神经网络提取中药材特征,使用 Softmax 归一化分类,通过模型融合、Dropout、BN 等方法来优化卷积神经网络,在 60 种中药材类别上取得了 99.65% 的准确率,提升了中药材图像识别的质量。

## 1 资料与方法

### 1.1 中药材图像数据集构建

1.1.1 中药材类别挑选 本文选取 60 种常用的中药材用于识别研究,具体类别名见表 1 所示。

1.1.2 中药材图片采集 在 Python 语言中,提供了 requests、lxml 等大量爬虫工具库,借助这些模块,可以快速实现根据关键词爬取百度图片和谷歌图片的爬虫代码。本文使用多线程和正则表达式等技术实现了根据关键词爬取中药材图片。

1.1.3 中药材图片清洗 在中药材图片采集阶段共获得了 15 287 张图片,由于存在一些非对应类别、含有大量水印、较为模糊、图像特征不明显的图像数据,因此需要进行人工剔除无用的数据操作。经过数据清洗后,最终选取的中药材数据集图片数量为 9987 张,对应 60 个类别,且各类别间图像数量较为均衡。

表 1 本文所识别的中药材类别名一览表

序号	中药材名	序号	中药材名	序号	中药材名
1	侧柏叶	21	楮实子	41	草果
2	八角茴香	22	槟榔	42	草乌
3	冰片	23	沉香	43	草豆蔻
4	北沙参	24	白前	44	菝葜
5	北豆根	25	白及	45	苍术
6	半夏	26	白扁豆	46	苍耳子
7	半边莲	27	白术	47	蓖麻子
8	川木通	28	白果	48	萆薢
9	川木香	29	白矾	49	薄荷
10	川楝子	30	白芍	50	蝉蜕
11	川乌	31	白花蛇舌草	51	蟾酥
12	川牛膝	32	白芷	52	补骨脂
13	川贝母	33	白茅根	53	赤小豆
14	巴戟天	34	白薇	54	赤芍
15	巴豆	35	白藜	55	车前子
16	常山	36	白附子	56	车前草
17	斑蝥	37	白头翁	57	阿胶
18	板蓝根	38	白鲜皮	58	阿魏
19	柏子仁	39	百合	59	陈皮
20	柴胡	40	百部	60	鳖甲

注:中药材名来源于香港浸会大学中药材图像数据库。

1.1.4 图像归一化 由于卷积神经网络模型对输入尺寸要求为 224 或 299,而采集的图片输入尺寸不一,因此需要对图像进行归一化操作,即将其处理为统一输入尺寸,调整为卷积神经网络标准输入形式 224 和 299。

1.1.5 数据集划分 为了保证实验变量的唯一性,本文按照 6:2:2 的比例对原始数据集和增强数据集进行划分。其中,训练集、验证集和测试集均无交集。

### 1.2 基础理论知识

1.2.1 深度学习与卷积神经网络 深度学习是机器学习的一个重要研究方向。它通过学习样本表现特征和内在规律,从而对新的数据进行解释推理,解决了很多过去无法实现的模式识别复杂难题<sup>[7]</sup>。

卷积神经网络(convolutional neural network, CNN)是目前深度学习计算机视觉(computer vision, CV)领域最常用的算法<sup>[8-9]</sup>,它可以大幅提升图像分类精度,其基本结构如图 1 所示。

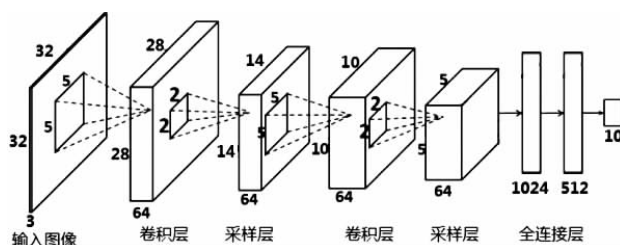


图 1 卷积神经网络结构

1.2.2 激活函数与优化算法 神经网络中每一层的神经元都会接受上一层神经元的输出作为本层输入,再将本层神经元的输出传递给下一层。如果输入与输出之间不存在函数关系,则神经网络只是简单的线性组合关系,表达能力很弱。因此,研究人员在输入与输出之间添加了函数关系,使神经网络的表达能力大大增强,这样的函数称为激活函数<sup>[10]</sup>。

机器学习问题即为最优化问题,使用优化算法来最优化损失函数。按照梯度类型可分为有梯度算法和无梯度算法,其中,有梯度算法包括梯度下降算法<sup>[11-12]</sup>、动量法、RMSProp<sup>[13]</sup>、Adam<sup>[14]</sup>等。

1.2.3 迁移学习与数据增强 迁移学习是一种机器学习方法,把某个领域下学习的知识迁移到另外一个领域,使其取得更好的学习效果<sup>[15]</sup>。通常,当源域与目标域具有相似特征,且源域样本数据充足,目标域样本数据较小时,非常适合通过迁移学习来提高样本不足任务的学习效果。有实验表明,迁移学习在小样本下效果显著<sup>[16]</sup>。

数据增强是为了减少由于数据集较小而导致模型过拟合现象而提出来的方法<sup>[17]</sup>。通过对图像进行各种变化后作为训练集,进而提高模型的泛化能力。常用的数据增强方法包括图像旋转、缩放、平移、噪声和尺寸变化等<sup>[18]</sup>。

1.2.4 模型融合 模型融合即按照某种方法训练多个模型。可以通过数学证明,随着个体分类器数目的增多,集成的错误率将呈指数级下降,最终可以趋于零。

### 1.3 经典网络结构

1.3.1 MobileNet MobileNet<sup>[19]</sup>具有轻量级的特点,为了使网络参数下降,它使用了一种深度可分离卷积,相比于传统卷积操作大大减少了参数数量,真正达到轻量级的目的。

1.3.2 Inception Inception<sup>[20]</sup>通过寻找一个密集成分来代替最优局部稀疏结构,引入了大量的卷积操作进行降维,使得 Inception 可以高效拟合网络的稀疏部分,如图 2 所示。

1.3.3 Xception Xception<sup>[21]</sup>网络是对 InceptionV3 网络的一种改进,Xception 想要解耦通道相关性和空间相关性,使用了深度可分离卷积来代替 Inception 原有的卷积操作。即先使用深度方向卷积,对每个通道进行卷积后连接,再使用逐点卷积,对深

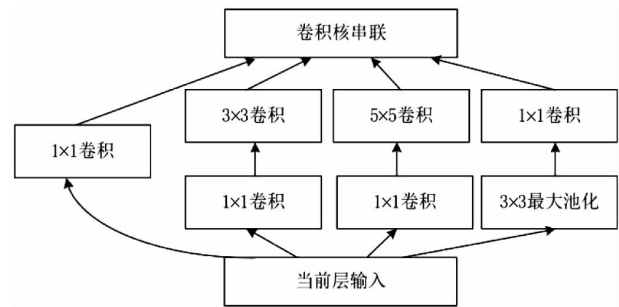


图 2 Inception 模块结构图

度方向卷积连接的结果进行一次卷积。

1.3.4 ResNet 在残差网络没有出现之前,在计算机视觉领域运用 CNN 时,当 GoogleNet 在原有 22 层或 VGG 在原有 19 层再次增加网络层数后,网络会出现退化现象,损失由最初的逐渐下降趋于平稳或不断上升,不再收敛,准确率也不断下降。在 ResNet<sup>[22]</sup>中,通过残差网络来解决这一问题。残差网络的基本模块是残差块,其结构如图 3 所示。

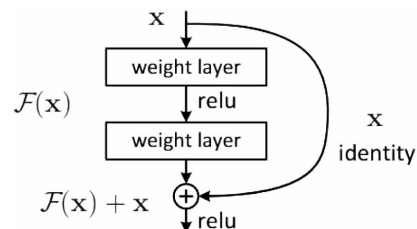


图 3 残差块结构图

1.3.5 DenseNet DenseNet<sup>[23]</sup>最大的特点是在常规池化层之后使用了一个被称之为 Dense Block 的结构连接。DenseNet 被称为密集连接网络的原因在于:Dense Block 中每一层 Bottle Neck 的输入均来源于上一层 Bottle Neck 的输出和原始输入的堆叠。

### 1.4 DxFusion 融合模型

本研究对 DenseNet 和 Xception 经过全局池化得到的特征向量使用 Concatenate 层来进行特征融合,再使用全连接层保证足够多的特征,然后加入 Dropout 层防止过拟合,最后再使用 Softmax 进行分类。在本文中,将这个新的网络结构称为 DxFusion (DenseNet Xception Fusion),其网络结构如图 4 所示。

## 2 结果与分析

### 2.1 图像分类评价指标

在深度学习计算机视觉领域,对于多分类任务而言,常用评价指标包括准确率(accuracy)、召回率(recall)、精确率(precision)和 F1-Score 值等。

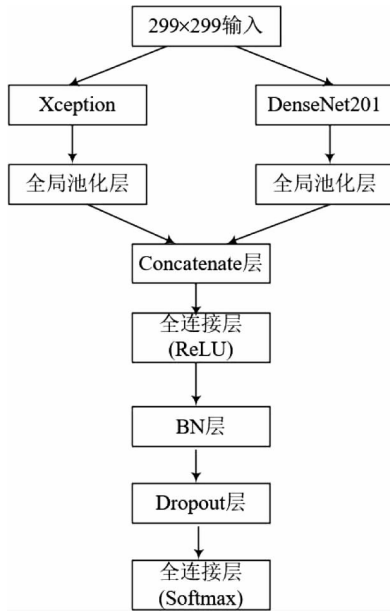


图 4 DxFusion 网络结构

对于多分类问题而言,平均精确率、召回率和 F1-Score 值等价于准确率,因此本文选择测试集准确率作为模型评价指标。具体计算公式如下:

$$Total = \sum_{i=1}^k (TP_i + FN_i) = \sum_{i=1}^k (TP_i + FN_i) \quad \text{式(1)}$$

$$Total_{TP} = \sum_{i=1}^k TP_i \quad \text{式(2)}$$

$$Accuracy = \frac{Total_{TP}}{Total} \quad \text{式(3)}$$

### 2.2 迁移学习实验结果

分别对 MobileNet、Inception、Xception、ResNet、DenseNet 等网络结构进行从头开始和迁移学习两种训练方式,目的在于比较迁移学习和从头训练在中药材数据集上的准确率和收敛速率,通过比较分析,得出实验结果。实验分为两个阶段,即从头训练和微调训练,基本训练参数设置如表 2 和表 3 所示,训练结果如图 5 和图 6 所示。由图 6 可知,从头训练收敛速率较慢,且由于样本不够广泛,验证集上模型的准确率不够高,在中药材数据集上难以训练得到一个较好的模型;而迁移学习下的模型收敛速率很快,验证集上也拥有较高的准确率。

表 2 迁移学习训练参数设置

训练参数	从头训练	微调训练
学习率	0.001	0.000 1
优化函数	Adam	Adam
迭代次数	50	50
批处理大小	64	64
数据集	原始数据集	原始数据集
训练层数	全部层	原有部分顶层、自行设计顶层
损失函数	多分类交叉熵	多分类交叉熵

表 3 各个网络的微调细节参数设置

模型	训练层
MobileNetV3	解冻所有层
InceptionV3	从第 106 层开始解除冻结
Xception	解冻所有层
ResNet152	从第 205 层开始解除冻结
DenseNet201	从第 224 层开始解除冻结

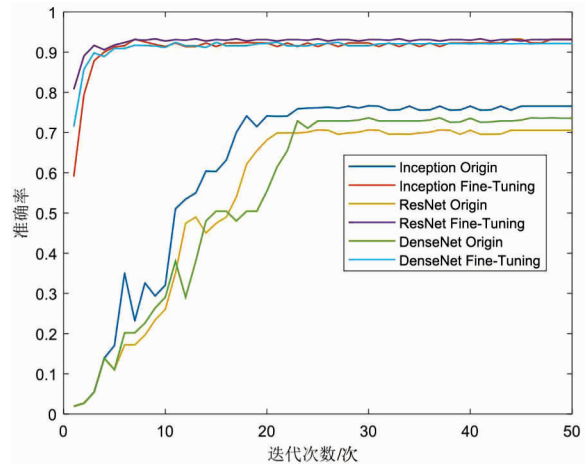


图 5 迁移学习验证集上准确率对比

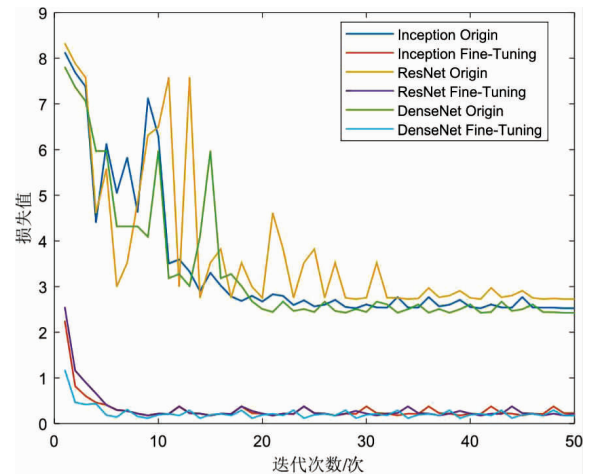


图 6 迁移学习验证集上损失值对比

### 2.3 数据增强实验结果

对原始数据集通过垂直翻转、水平横移、垂直横移、随机旋转和随机缩放等方式进行增强,得到增强数据集。原始数据集为经过数据清洗后的 9987 张图片,增强数据集图片数量为 34 040 张。训练参数与表 2 微调训练一致,训练结果如图 7 和表 4 所示。由结果数据可知,各个网络模型在数据增强之后,在验证集、测试集上的准确率均有一定提升。

### 2.4 模型融合实验结果

本文使用在数据增强实验阶段表现最好的两个网络 DenseNet 和 Xception 进行模型融合,共同提取特征,构建了 DxFusion 融合模型。训练参数与表 2 微

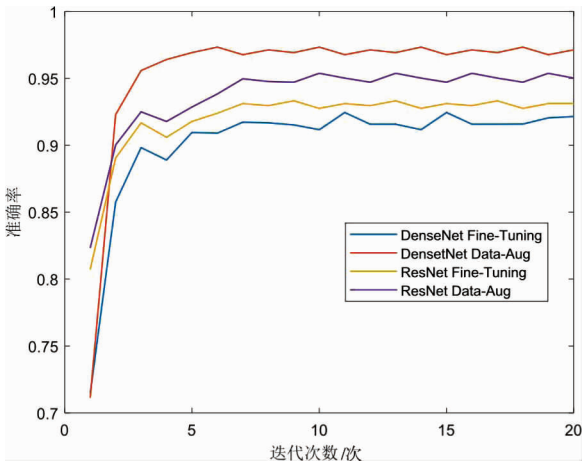


图7 数据增强验证集准确率对比图

表4 数据增强测试集准确率对比结果

模型	原始数据集/%	增强数据集/%	相比提升
MobileNetV3	94.18	95.84	1.76
InceptionV3	92.29	94.79	2.71
Xception	93.58	96.65	3.28
ResNet152	92.71	94.78	2.23
DenseNet201	93.46	96.50	2.57

调训练一致,使用数据增强得到增强数据集,训练曲线如图8所示,评价结果如表5所示。由表5中的数据可知,融合后的模型 DxFusion 在测试集上的准确率(99.65%)均优于单个 DenseNet 网络(96.50%)和 Xception 网络(96.65%),说明模型融合可以进一步提高模型的识别准确率。

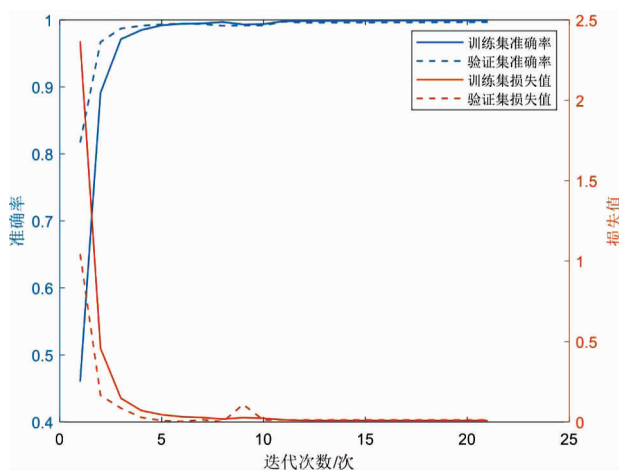


图8 DxFusion 训练曲线

表5 DxFusion 模型评价

模型	训练集		验证集		测试集	
	准确率/%	损失值	准确率/%	损失值	准确率/%	损失值
DxFusion	99.93	0.008	99.67	0.013	99.65	0.25

### 2.5 方法比较

不同方法的中药材图像识别准确率对比结果如

表6所示。消融实验的结果表明,加入迁移学习、数据增强、模型融合等优化策略之后,本文的中药材识别的准确率取得了显著提升。为了比较不同网络结构在不同训练方式下中药材图像识别的差异,本文以测试集准确率作为评价指标。由表6可知,从头训练阶段,中药材识别准确率为76.98%;通过迁移学习,中药材识别准确率为94.18%;通过数据增强,准确率提升为96.56%;通过模型融合,准确率提升至99.65%。

表6 中药材识别准确率对比结果

模型	从头训练/%	迁移学习/%	数据增强/%
MobileNetV3	72.97	94.18	95.84
InceptionV3	76.98	92.29	94.79
Xception	74.97	93.58	96.65
ResNet152	70.75	92.71	94.91
DenseNet201	75.98	93.46	96.50
DxFusion(Ours)	—	—	99.65

与已有的中药材图像识别研究相比<sup>[1-5]</sup>,本文基于迁移学习和模型融合开展中药材图像识别研究,取得了良好的效果,如表7所示。

表7 中药材识别方法比较结果

模型	中药材类别数/种	优化方法	准确率/%
AlexNet <sup>[1]</sup>	98	集成学习	85.00
SqueezeNet <sup>[1]</sup>	98	集成学习	90.50
GoogleNet <sup>[1]</sup>	98	集成学习	93.45
AlexNet <sup>[2]</sup>	5	数据增强	87.50
ResNet18 <sup>[3]</sup>	8	—	72.56
ResNet50 <sup>[4]</sup>	11	目标检测	97.88
GoogleNet <sup>[5]</sup>	2	—	97.50
DxFusion(Ours)	60	模型融合	99.65

### 3 讨论

目前,在中药材图像识别领域存在缺少大规模样本、相关研究文献较少等问题,找不到公开的中药材数据集,本文通过网络爬虫自行构建了数据集。实验结果说明,迁移学习、数据增强、模型融合等方式可以提高中药材识别的准确率,利用 Xception 和 DenseNet 进行模型融合,在60种类别的测试集上达到了99.65%的准确率,表明了卷积神经网络可以很好地提取出中药材的特征,较好地完成了60种中药材自动识别,具有准确性、高效性等特点,从而减少了人为中药材辨别的工作量。本文所提出的智能化中药材图像识别方法可降低人为识别中药材的难度和成本。

本文将深度学习迁移学习、数据增强、模型融合等技术并应用于中药材图像识别,通过 Dropout、BN 算法改进卷积神经网络,为中药材识别研究提供了一种新思路,主要贡献和创新点如下:(1)使用了目前较为先进的 CNN 网络结构(DenseNet 和 Xception),并对它们提取的全局特征进行融合,提出了一种新的网络结构 Dx Fusion,构建了新型融合模型;(2)运用数据增强技术扩充中药材图像数据集,构建了超过 34 000 张图片的中药材图像增强数据集,可广泛应用于中药材图像研究;(3)已有研究存在识别中药材类别数目较少或识别准确率较低等问题,本文在 60 种中药材类别上取得了 99.65% 的准确率,在识别类别数目和准确率两个指标上均优于已有大多数中药材图像识别方法。

在后续的研究工作中,将进一步增加中药材类别并扩充数据集。同时充分考虑中药材纹理底层特性,通过集成学习来进一步提升中药材识别准确率。此外,还将利用生成对抗网络来产生对抗样本,对模型进行对抗训练,提升图像识别模型的稳健性和抗干扰能力。

## 参考文献

- [1] 庄奕珊.基于深度神经网络的中药材识别[D].广州:华南理工大学,2018.
- [2] 黄方亮,俞磊,沈同平,等.基于 AlexNet 深度学习模型的中草药植物图像分类研究与实现[J].齐鲁工业大学学报,2020,34(2):44-49.
- [3] 王寒迎.基于深度学习的中药材鉴别方法研究[D].桂林:桂林电子科技大学,2019.
- [4] 吴冲,谭超群,黄永亮,等.基于深度学习算法的川贝母、山楂及半夏饮片的智能鉴别[J].中国实验方剂学杂志,2020,26(21):195-201.
- [5] 史婷婷,张小波,郭兰萍,等.基于深度卷积神经网络的仿野生种植金银花遥感识别方法研究[J].中国中药杂志,2020,45(23):5658-5662.
- [6] 陶益,陈林,江恩赐,等.人工智能和工业 4.0 视域下高光谱成像技术融合深度学习方法在中药领域中的应用与展望[J].中国中药杂志,2020,45(22):5438-5442.
- [7] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [8] SCHMIDHUBER J. Deep learning in neural networks: An overview[J]. Neural Networks, 2015, 61(1): 85-117.
- [9] ULLAH S, HALIM Z. Imagined character recognition through EEG signals using deep convolutional neural network[J]. Medical & Biological Engineering & Computing, 2021, 59(5): 1167-1183.
- [10] QIAN S, LIU H, LIU C, et al. Adaptive activation functions in convolutional neural networks[J]. Neurocomputing, 2018, 272(1): 204-212.
- [11] BOTTOU L. Large-scale machine learning with stochastic gradient descent[C]//Proceedings of COMPSTAT'2010, Physica-Verlag, 2010: 177-186.
- [12] KETKAR N. Deep Learning with Python[M]. Berkeley: Apress, 2017: 113-132.
- [13] XU D P, ZHANG S D, ZHANG H S, et al. Convergence of the RMSProp deep learning method with penalty for nonconvex optimization[J]. Neural Networks, 2021, 139(7): 17-23.
- [14] BARAKAT, BIANCHI P. Convergence and dynamical behavior of the ADAM algorithm for nonconvex stochastic optimization[J]. SIAM Journal on Optimization, 2021, 31(1): 244-274.
- [15] 杨强,张宇,戴文渊,等.迁移学习[M].北京:机械工业出版社,2020.
- [16] 胡胜利,吴季.一种基于迁移学习的小样本图像分类方法[J].湖北理工学院学报,2021,37(2):27-32.
- [17] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 1-48.
- [18] MIKOLAJCZYK A, GROCHOWSKI M. Data augmentation for improving deep learning in image classification problem[C]//2018 International Interdisciplinary PhD Workshop, IEEE, 2018: 117-122.
- [19] LI G, ZHANG H X, E L N, et al. Recognition of honeycomb lung in CT images based on improved MobileNet model[J]. Medical Physics, 2021, 48(8): 4304-4315.
- [20] LIU Z Y, YANG C, HUANG J, et al. Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer[J]. Future Generation Computer Systems, 2021, 114(1): 358-367.
- [21] CHEN B J, JU X W, XIAO B, et al. Locally GAN-generated face detection based on an improved Xception[J]. Information Sciences, 2021, 572(9): 16-28.
- [22] PHAWINEE S, CAI J F, GUO Z Y, et al. Face recognition in an intelligent door lock with ResNet model based on deep learning[J]. Journal of Intelligent and Fuzzy Systems, 2021, 40(4): 1-11.
- [23] HEMALATHA J, ROSELINE S A, et al. An efficient DenseNet-based deep learning model for malware detection[J]. Entropy, 2021, 23(3): 344.

(本文编辑 匡静之)