

DOI: 10.13288/j.11-2166/r.2019.10.007

基于深度学习与迁移学习的中医舌象提取识别研究

刘梦, 王曦廷, 周璐, 谭丽博, 李杰, 关静, 李峰*

(北京中医药大学中医学院, 北京市朝阳区北三环东路 11 号, 100029)

摘要 目的 探讨基于 Faster R-CNN 算法的深度学习与迁移学习技术在中医齿痕舌与裂纹舌局部特征提取识别中的应用效果。方法 将 500 例受试者舌象图片使用 LabelImg V 1.6 软件进行齿痕舌与裂纹舌区域标注, 划分训练数据集与测试数据集, 通过 Faster R-CNN 的深度学习技术与迁移学习技术中的微调 (fine-tune) 方法构建模型, 最后使用准确率、精确率、召回率对模型效果进行评价。结果 对于裂纹舌, 模型识别真阳性 99 例, 假阴性 3 例, 真阴性 45 例, 假阳性 3 例, 准确率 0.960、精确率 0.971、召回率 0.971。对于齿痕舌, 模型识别真阳性 60 例, 假阴性 20 例, 真阴性 69 例, 假阳性 1 例, 准确率 0.860、精确率为 0.983、召回率 0.750。模型宏观准确率 0.910、宏观精确率 0.977、宏观召回率 0.860。图像识别结果显示, 模型不受舌象中病理变化所在位置的影响, 对舌象局部特征提取方面也具有较强的适应性特点。结论 基于深度学习与迁移学习的方法, 可以较好地完成中医舌象局部特征辨识任务, 具有较好的迁移能力。

关键词 舌象; 裂纹舌; 齿痕舌; 舌象识别

中医通过“望闻问切”的四诊方式, 获取患者与疾病相关的体征信息, 然后基于辨证论治的基本理论对疾病进行诊断。舌诊作为中医望诊的主要组成部分之一, 在临床诊断中起着重要的作用。中医学认为, 舌与脏腑、经络、气血、津液有着密切的关系, 舌象的变化能较客观准确地反映病情, 可作为诊断疾病、了解病情的发展变化和辨证施治的重要依据之一。

传统的中医舌象辨识依赖于医生的肉眼观察及经验判断, 其诊断结果既受医生主观的知识水平、思维能力和诊断技能的限制, 又受光线、温度等外部环境的影响, 缺乏客观、定量、标准的评价及应用。随着技术进步, 利用信息技术来探索舌诊指标的信息化^[1]、中医舌象辨识^[2]、舌象诊断及病证关联的研究越来越广泛而深入。李晓宇等^[3]利用支持向量机和粗糙集等算法, 就舌象分析进行研究, 对舌象采集、舌体外观自动分析、舌体区域分割识别、舌象特征自动分析等方面进行了研究。许家佗等^[4]对舌象进行了较全面的研究, 包括基于颜色、纹理、裂纹和点刺等特征的探索研究^[5]。

Obafemi-Ajayi 等^[6]提出了一种基于几何图形与多项式方程的特征提取方法, 用于自动识别分类舌象图片形状。

尽管中医舌象识别的研究取得了一定进展, 但是目前舌象客观化研究主要基于整体、舌质与舌苔全局特征提取与分析研究, 对于舌象不同部位特征提取与应用研究的报道较少^[7]。关于中医舌象辨识仍存在着样本不平衡、图像处理算法繁重、难以自动化高通量处理、关联研究迁移能力弱等问题。得益于机器学习方法中以深度神经网络为模型的图像识别技术的成功, 深度学习被广泛应用于学术与工业领域。例如, Oren 等^[8]学者强调深度学习是一种重要的预测分析工具, 使用深度卷积神经网络 (DeepLoc) 分析乳腺癌细胞图像, 在亚细胞定位的分类上显示了很好的效果。Kandaswamy 等^[9]学者使用深度神经网络 (DNNs), 对化合物的作用机制进行了分类。由此可见, 深度学习技术受到了众多科研者的关注, 并且表现出极大的应用潜力。

本研究基于 Faster R-CNN 的深度学习与微调 (fine-tune) 的方法进行舌象局部特征提取的研究^[10], 针对中医齿痕舌与裂纹舌舌象, 我们构建了用于训练、测试与验证的标准数据库, 并使用 LabelImg 工具进行舌象标注。使用 ZF 网络^[11]作为

基金项目: 北京市自然科学基金(7162124); 2018 北京中医药大学青年教师项目(2018-JYBZZ-JS007)

* 通讯作者: lifeng95@vip.sina.com

特征提取网络构建深度学习模型,实现对中医齿痕舌与裂纹舌舌象的局部综合特征进行一体化识别。采用迁移学习的方法提高数据样本的训练效率,防止过拟合,以解决中医舌象样本数据少、信息量大、特征众多等问题。现报告如下。

1 方法

本研究首先部署 Caffe 框架^[12],随后建立齿痕舌及裂纹舌中医舌象特征检测模型。实验的具体流程为:收集并筛选舌象图片,进行特征标注,构建舌象数据集,部署开源深度学习框架,采用 VOC 2007 数据集进行预训练,采用 fine-tune 方法进行舌象识别模型的训练与测试。

1.1 舌象筛选与预处理

本研究数据基于课题组既往使用松下 GH4 相机所采集的临床舌象照片,照片格式为 JPG,像素为 2848×4288 。共收集 600 例舌象图片。参照全国中医药行业高等教育“十二五”规划教材《中医诊断学》^[13]舌诊部分对于齿痕舌和裂纹舌进行判断,并交由两位中医诊断学教授进行一致性评价确认,共筛选出齿痕舌和裂纹舌舌象图片 500 例。其中符合齿痕舌诊断图片 260 例,裂纹舌诊断图片 300 例,齿痕舌与裂纹舌并存的图片 60 例。对筛选后的舌象图片数据进行统一编码,编码形式为顺序 6 位数字,起点为 100101。对筛选后的舌象图片进一步裁剪,留存包含舌象的区域并进行预处理,形式如图 1。



图 1 预处理后的舌象图片

1.2 舌象标注

使用基于 python 的 LabelImg v 1.6 图像标注软件,对裂纹舌与齿痕舌区域进行标注,见图 2。设定齿痕舌标签为“chi_hen_she”,裂纹舌标签为“lie_wen_she”。生成对应的记录有标注区域信息的 xml 文件,并由上述教授进行确认。将所有标注完成的数据按照 7 : 3 的比例随机分为训练集 350 例和测试集 150 例。数据标注所用的随机方法为:使用 python 语言,安装并载入 Numpy 库,调用 numpy.random.randint() 函数。该函数可以返回一个在区间 (low, high) 中离散均匀抽样的数

组。本实验中,该函数设置参数为 numpy.random.randint(100101, 100501, (1, 350), dtype = “l”)。

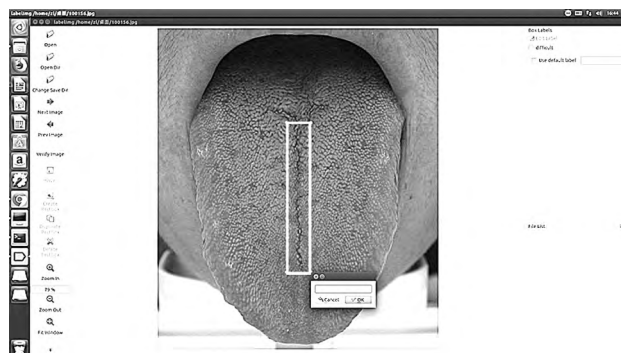


图 2 使用 LabelImg 进行舌象标注操作界面

1.3 模型构建

采用 Faster R-CNN 算法与 fine-tune 方法相结合,可使小样本训练数据所建立的中医舌象模型有良好表现。冻结预训练模型的部分卷积层(通常是靠近输入的多数卷积层),训练剩下的卷积层(通常是靠近输出的部分卷积层)和全连接层。在结构上, Faster R-CNN 已经将特征抽取、区域建议网络、边框回归、分类都整合在了一个网络中,使得综合性能有较大提高,在检测速度方面尤为明显。Faster R-CNN 模型中预制了两种预训练网络结果,分别为 VGG16 与 ZF 网络。使用 VOC2007 数据集完成模型预训练。

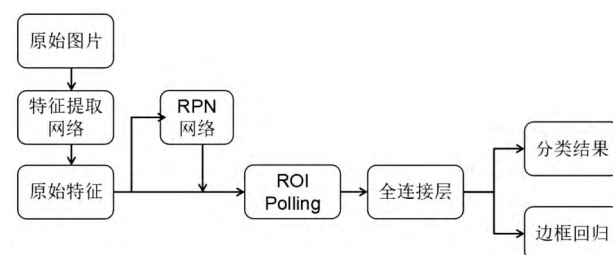


图 3 Faster R-CNN 算法框架图

1.3.1 特征提取网络模型构建 如前文所述, Faster R-CNN 是一种基于卷积神经网络之上的多目标标注算法,集特征提取与目标识别于一身,可直接利用卷积层对中医舌象进行特征提取、ROI 层锁定反映病理变化的区域、CLS 层鉴定对目标区域舌象类别 3 个过程,其算法框架见图 3。本研究,选择 ZF 网络作为特征提取网络, ZF 网络是一个 8 层结构的卷积网络模型,其参数见表 1,其结构见图 4。

ZF 网络属小型网络。由于所选用的中医舌象图片总数为 500 份,在这种样本较少的情况下,较倾向使用层数较浅的 ZF 网络。同时,在 Faster

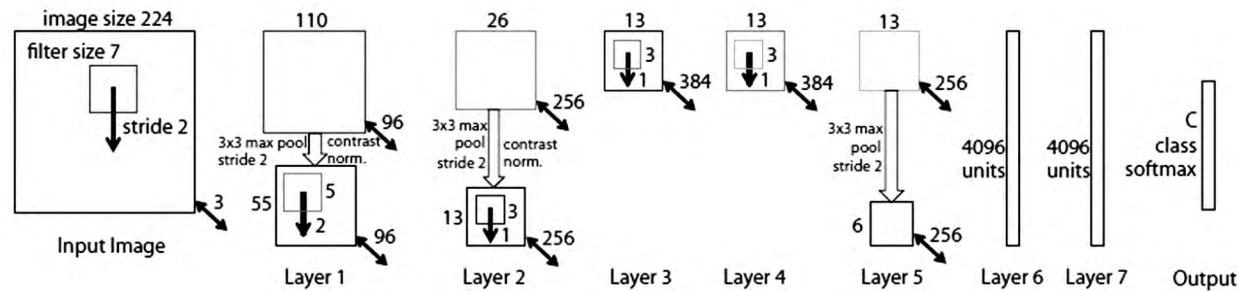


图 4 ZF 网络结构图

表 1 ZF 网络参数表

类型/层数	卷积核数量	卷积核大小	步长	输出
CONV1	3@ 227 × 227	7 × 7	2	96@ 111 × 111
POOL1	96@ 111 × 111	3 × 3	2	96@ 55 × 55
CONV2	96@ 55 × 55	5 × 5	2	256@ 26 × 26
POOL2	256@ 26 × 26	3 × 3	2	256@ 13 × 13
CONV3	256@ 26 × 26	3 × 3	1	512@ 13 × 13
CONV4	512@ 13 × 13	3 × 3	1	1024@ 13 × 13
CONV5	1024@ 13 × 13	3 × 3	1	512@ 13 × 13
FC6	86528@ 1 × 1	1 × 1	1	4096@ 1 × 1
FC7	4096@ 1 × 1	1 × 1	1	4096@ 1 × 1
FC8	4096@ 1 × 1	1 × 1	1	1000@ 1 × 1

R-CNN 原作者提供的算法支持中，也将 ZF 网络作为默认的预训练模型，所提供的参数具有较好的参考性。而且，流程中使用 ZF 网络采取默认配置，大大加快了本实验的研究进度。此外，在目前使用 Faster R-CNN 的研究中，均有使用 ZF 网络^[4]或 VGG16^[5]作为特征提取网络。在一些 Faster R-CNN 的应用研究中，其结果表明使用不同的特征提取网络差异较小^[6-17]。故本研究使用 ZF 网络作为特征提取网络。

1.3.2 fine-tune 方法的应用 迁移学习可以理解为通过某一任务学习到的知识可以被迁移到另外一个任务^[8]。迁移学习不仅节约了数据训练的时间，也能有效防止深度学习网络的过拟合。一般的迁移学习是将训练好的一个网络的前 n 层复制到目标网络的前 n 层，对目标网络的其余层随机初始化，然后开始训练任务。其中，在做反向传播时，有两种方式可以选择：一种是把迁移过来的前 n 层冻结起来，即在训练时不改变这 n 层的权重；另一种是不冻结这前 n 层，而是不断地调整他们的权重值，称为 fine-tune^[9]。

中医舌象样本数据少，而信息量大，其识别涉及舌质、舌形、舌体等众多特征。因此，使用 fine-tune 技术可有效减少训练数据，这也是近年来深度学习中的一种高效的训练方式^[20]。借助 fine-

tune 技术既可减少所需的训练样本，又可防止过拟合的发生，具有一举两得的优点。本研究采用将最后一层全连接层前的所有层进行迁移的方法，修改最后一层全连接层的分类信息，用已经训练好的 ZF 模型中 fine-tune 的权重来初始化权重。由于本研究借助 Caffe 框架实现 Faster R-CNN，实现方案参照文献 [10] 所提供的开源代码 (<https://github.com/rbgirshick/py-faster-rcnn/>)，因此在训练 fine-tune 中所采用的参数参照开源代码中的默认设置，前 60 000 次迭代，学习率参数为 0.001；后 20 000 次迭代，学习率为 0.0001，momentum 设置为 0.9，weight decay 设置为 0.0005。

1.3.3 基于 Caffe 实现模型构建 为了验证算法的有效性及其并行优化后的高效性，建立实验仿真环境：CPU 处理器为 i7 四代处理器，固态硬盘 250G，内存为 16GB，GPU 处理器为 NVIDIA 1080ti。实验平台为 Ubuntu 16.04。

首先参照 Caffe 实现 Faster R-CNN 模型的方法构建本研究算法框架。下载 Faster R-CNN 的预训练模型并顺利运行 Demo。然后使用经预处理的 350 份舌象图片进行模型训练。具体的训练流程为：加载已经训练 ZF 网络，初始化权重；使用 ZF 网络初始化权重，并开始抽取特征；初始化识别框参数；判断边框位置与类别是否与训练集合相符；反向传递更新参数；完成训练，生成中医舌象识别模型。使用测试集进行测试，评估模型准确性。

1.4 模型效果评价

本次模型采用准确率、精确率、召回率参数进行评价，以评价模型的性能表现。这些参数的计算涉及模型与中医诊断学专家共同的判别结果。例如，针对某张舌象图，若专家的结论是图像属于某一类别，分类算法给的结果与专家相同则称为真阳性 (TP)，若结果不同则称为假阴性 (FN)；若专家判断某样本不属于某一类别，而分类算法与专家结论相同则称为真阴性 (TN)，若与专家结论相反

则称为假阳性 (FP)。计算公式为: 准确率 =

$$\frac{TP + TN}{TP + TN + FP + FN}, \text{精确率} = \frac{TP}{TP + FP}, \text{召回率} =$$

$$\frac{TP}{TP + FN}。$$

2 结果

使用测试集的 150 例舌象图片对模型进行评估。对于裂纹舌, TP = 99, FP = 3, TN = 45, FN = 3, 模型准确率为 0.960、精确率为 0.971、召回率为 0.971。对于齿痕舌, TP = 60, FP = 20, TN = 69, FN = 1, 模型准确率为 0.860、精确率为 0.983、召回率为 0.750。对于多目标任务的模型, 通常采用宏观准确率、宏观精确率、宏观召回率进行模型评估, 其结果通常为各项子任务性能指标的均值。因此, 本研究中的模型宏观准确率为 0.910、宏观精确率为 0.977、宏观召回率为 0.860。

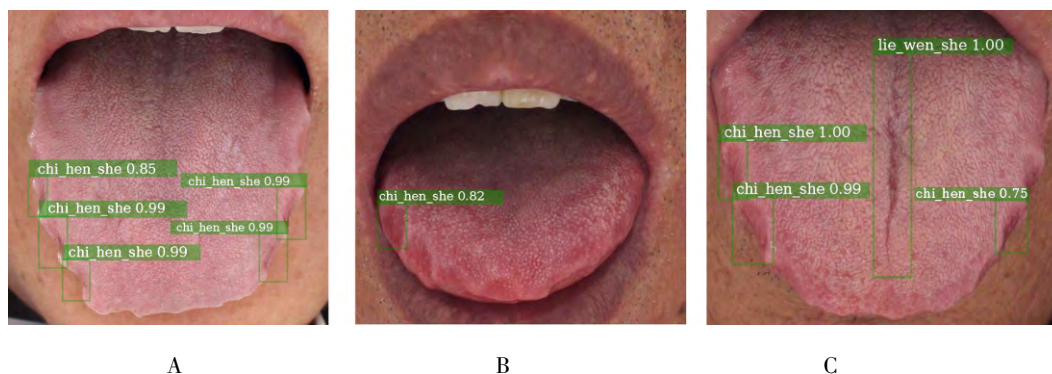
从实验结果来看, 该模型不受舌象中病理变化所在位置的影响, 在位置不一时, 仍可以进行良好的标注。但在全面标注上, 虽然仍有部分反映病理特点的特征未被识别, 但已经标注的部分足以揭示病理变化, 如图 5 中方框所示部分。

从具体舌象图片特征提取来看, 本算法在舌象

局部特征提取方面也具有较强的适应性特点。局部舌象的识别, 受到周围色彩、纹理的影响较大, 这往往是将某一局部舌象作为一个小部分进行特征提取识别的重点和难点。本系列算法依据区域建议网络 RPN (region proposal networks), 快速生成候选区域, 然后经多轮训练, 让 RPN 与 Faster R-CNN 共享权值。从实验结果看, 本算法有助于解决这一问题。例如, 对于齿痕舌象识别而言, 图 5-A 所识别出的齿痕, 其舌质色偏淡, 舌象图片属于居中拍摄; 而图 5-B 所识别出的齿痕, 其舌质色偏红, 而舌象图片拍摄角度偏俯视; 这表明, 本模型可以较好地识别不同舌质色、不同拍摄角度的齿痕舌。从识别裂纹舌来看, 本模型可以识别不同位置、不同纹理的裂纹舌。综合实验结果表明, 本模型可以较好地解决局部舌象识别的科学问题。同时, 由图 5A 和图 5-C 可以看出, 本模型也可较好地完成多目标识别任务。

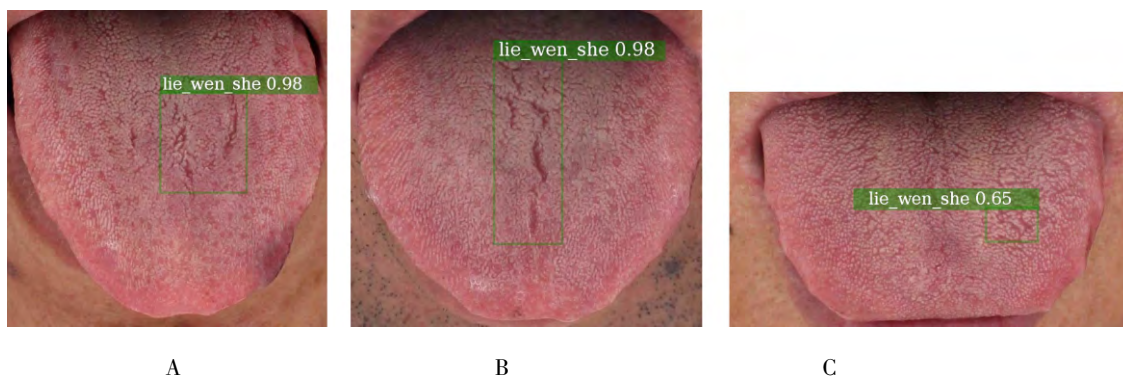
3 讨论

舌诊是中医临床诊断中的重要内容之一, 有关舌象特征的鉴定与疾病关联分析是完成中医舌诊的重要环节。信息技术在中医舌象客观化研究中发挥



注: 图中数字代表置信度 (检出率)

图 5 中医舌象识别模型的齿痕舌、裂纹舌识别结果举例



注: 图中数字代表置信度 (检出率)

图 6 中医舌象识别模型的齿痕舌、裂纹舌识别结果待改进部分

着越来越重要的作用。本研究所提出的模型具有很好的泛化能力,相对于传统机器学习构建的舌象分类模型,本模型无需其他算法介入图像预处理,仅依靠卷积层便可完成特征提取,集舌象的识别、标注于一体,可有效降低模型设计的难度。本研究为信息技术在舌象客观化与舌象自动化分析的研究提供了借鉴,即借助深度学习与迁移学习在中医舌象辨识方面所具有的独特优势,其算法构架精炼,工序精简,可实现自动化高通量处理,可以较好地解决局部舌象识别的相关问题。

在中医临床方面,基于信息技术下舌象识别的实时传播优势,可为中医远程医疗提供舌象相关分析参数,增加参考分析报告;基于数据分析算力优势,可为大数据下疾病系统的舌象关联性分析提供自动化、快速的舌象识别方法学支持。在人工智能与大数据背景下,中医诊断信息化是亟待深入研究的领域。本研究为解决中医局部舌象智能识别做出了尝试,在中医证候诊断与风险评估、构建智能化的中医健康管理服务云平台,实现对亚健康、慢性病人群的个性化、动态化健康管理中具有潜在的参考及运用价值^[2]。

虽然本方法具有一定优点,但实验也存在着一一定的局限性。例如,从模型在验证数据集上的表现可以看出模型仍具有一定的改进空间。由图 6-A 与图 6-B 可以看出,在裂纹的纹理提取中,在非狭长形的断续裂纹舌象中,本模型也可以进行提取识别,但出现了将左右两个裂纹在距离较大时提取成一个裂纹舌舌象的情况。这可能与边框回归的收敛能力有关,提示后期可以进一步调整模型的参数。再如,一般舌裂纹模型具有较好的识别能力,但是也会存在一些错误判断,如图 6-C 所示,对于一些舌苔有条纹而非舌裂纹的情况,在舌象图片放大时,模型会误判为裂纹舌。对于一些舌面有明显凹陷,色彩偏暗的部位,模型也会误判为裂纹舌;值得说明的是,虽然模型给出了判断,但检出率不高,即模型认为其为此种舌象的概率不高。此外,对于特征提取网络所提取的特征,目前仍难以对其做出理论性解释。并且对于中医学中舌苔的厚薄,仍需要借助传图的机器视觉技术,通过其颜色进行量化。因此,解决中医舌象识别等相关问题的系统性,仍需要采用多种方法进一步尝试。

注:刘梦和王曦廷对本文有同等贡献。

参考文献

[1] 马旭翔,李勇枝,许家伦. 舌诊信息化技术在证候诊断

及疗效评价中的应用 [J]. 中国中医基础医学杂志, 2018, 24(12): 1716-1719.

[2] 汤一平,王丽冉,何霞,等. 基于多任务卷积神经网络的舌象分类研究 [J]. 计算机科学, 2018, 45(12): 255-261.

[3] 李晓宇,张新峰,沈兰荪. 基于支撑向量机的中医舌色苔色识别算法研究 [J]. 北京生物医学工程, 2006, 25(1): 43-46.

[4] 许家伦,周昌乐,方肇勤,等. 舌像颜色特征的计算机分析与识别研究 [J]. 上海中医药大学学报, 2004, 18(3): 43-47.

[5] 屠立平,许家伦,张志枫,等. 一种自然光图像颜色校正方法在舌色图像校正中的应用 [J]. 南京中医药大学学报, 2011, 27(1): 15-18.

[6] OBAFEMI-AJAYI T, KANAWONG R, XU D, et al. Features for automated tongue image shape classification [C] // IEEE International Conference on Bioinformatics & Biomedicine Workshops. IEEE Computer Society, 2012: 273-279. doi: 10.1109/BIBMW.2012.6470316.

[7] 黄景斌,许家伦,张志枫,等. 基于舌象图像分区指标的糖尿病疗效评价 [J]. 中华中医药杂志, 2017, 32(3): 1148-1151.

[8] OREN ZK, BEN TG, JIMMY B, et al. Automated analysis of high-content microscopy data with deep learning [J]. Mol Syst Biol, 2017, 13(4): 924.

[9] KANDASWAMY C, SILVA LM, ALEXANDRE LA, et al. High-content analysis of breast cancer using single-cell deep transfer learning [J]. J Biomol Screen, 2016, 21(3): 252-259.

[10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE T PATTERN ANAL, 2017, 39(6): 1137-1149.

[11] ZEILER MD, FERGUS R. Visualizing and understanding convolutional networks [C] // European Conference on Computer Vision. Springer International Publishing, 2013: 818-833. doi: 10.1007/978-3-319-40590-4_53.

[12] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding [C] // Proceeding of the 22nd Acm International Conference on Multimedia. ACM, 2014: 675-678. doi: 10.1145/2647868.2654889.

[13] 李峰,董昌武. 中医诊断学 [M]. 北京: 科学出版社, 2018: 42-68.

[14] 朱虹,翟超,吕志,等. 基于 Faster-RCNN 的智能家居行人检测系统设计与实现 [J]. 工业控制计算机, 2018, 31(4): 68-70.

[15] 王黎,陆慧娟,叶敏超,等. Faster-RCNN 的癌症影像检测方法 [J]. 中国计量大学学报, 2018, 29(2): 136-141.

[16] 桑军,郭沛,项志立,等. Faster-RCNN 的车型识别分析

- 刊. 重庆大学学报, 2017, 40(7): 32-36.
- [7] 史凯静, 鲍泓, 徐冰心, 等. 基于 Faster RCNN 的智能车道路前方车辆检测方法 [J]. 计算机工程, 2018, 44(7): 36-41.
- [8] PAN SJ, YANG Q. A survey on transfer learning [J]. IEEE T KNOWL DATA EN, 2010, 22(10): 1345-1359.
- [9] 周曼, 刘志勇, 陈梦迟, 等. 基于 AlexNet 的迁移学习在流程工业图像识别中的应用 [J]. 工业控制计算机, 2018, 31(11): 80-82.
- [20] WANG G, LI W, ZULUAGA MA, et al. Interactive medical image segmentation using deep learning with image-specific fine tuning [J]. IEEE Trans Med Imaging, 2018, 37(7): 1562-1573.
- [21] 王忆勤. 中医诊断技术发展及四诊信息融合研究 [J]. 上海中医药大学学报, 2019, 33(1): 1-7.

Study on Extraction and Recognition of Traditional Chinese Medicine Tongue Manifestation: Based on Deep Learning and Migration Learning

LIU Meng, WANG Xiting, ZHOU Lu, TAN Libo, LI Jie, GUAN Jing, LI Feng

(School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100029)

ABSTRACT Objective To explore the application effect of deep learning and migration learning technology based on Faster R-CNN algorithm in local feature extraction and recognition of teeth-printed tongue and fissured tongue in traditional Chinese medicine (TCM). **Methods** The tongue images of 500 subjects were labeled in the areas of teeth-printed tongue and fissured tongue with the Labellmg V 1.6 software, and the training data set and test data set were divided. The deep learning technique of Faster R-CNN and the fine-tune in the migration learning technology were used to construct the model, and finally the precision rate, the accuracy rate, and the recall rate were used to evaluate the model effect. **Results** For the fissured tongue, the model identified 99 cases of true positive, 3 cases of false negative, 45 cases of true negative, 3 cases of false positive. The precision rate was 0.960, the accuracy rate was 0.971, and the recall rate was 0.971. For the teeth-printed tongue, the model identified true positive 60 cases, and false negative 20 cases, true negative 69 cases, and false positive 1 case. The precision rate was 0.860, the accuracy rate was 0.983, and the recall rate was 0.750. The model has a macro precision rate of 0.910, a macro accuracy rate of 0.977, and a macro recall rate of 0.860. The image recognition results showed that the model was not affected by the location of the pathological changes in the tongue image, and it also had strong adaptability to the local feature extraction of the tongue image. **Conclusion** Based on the method of deep learning and migration learning, the local feature recognition task of TCM tongue image can be completed well and has good migration ability.

Keywords tongue image; teeth-printed tongue; fissured tongue; tongue image recognition

(收稿日期: 2018-08-27; 修回日期: 2019-02-27)

编辑: 贾维娜

欢迎邮购 《中医杂志》 过刊

1988、1989、2000、2001 年合订本, 每本 100 元, 2002 年合订本 105 元, 2003 年合订本 105 元, 2004 年合订本 115 元, 2005 年合订本 125 元, 2007—2010 年合订本每年 140 元。2011(缺 5、6 期)—2013 年每册 9.80 元, 2014(缺 21 期)—2017 年(缺 3、5、15、18 期) 每册 15 元, 2018 年每册 20 元(缺 5、7、9 期)。2011—2013 年合订本(每年 4 本), 每年 280 元, 2014、2015、2016、2017 年合订本(每年 4 本), 每年 400 元。2018 年合订本, 600 元(共 4 本)。《中医杂志》2006—2008 年增刊每册 40 元。以上免费邮寄, 如需挂号每件另加 3 元。

欲购者请汇款至北京市东直门内南小街 16 号, 收款人: 中医杂志社

或关注《中医杂志》官方微信 (zyzzgzh) 购买。邮编: 100700, 电话: (010) 64089195。