

DOI: 10.16781/j.CN31-2187/R.20240166

· 论 著 ·

## 双模型策略在指甲病图像智能诊断中的应用

陈军晓<sup>1</sup>, 殷杰<sup>1</sup>, 扈东营<sup>2</sup>, 吴兆<sup>1</sup>, 朱秀艳<sup>1</sup>, 王士勇<sup>1\*</sup>

1. 海军军医大学(第二军医大学)第三附属医院信息科, 上海 200438

2. 海军军医大学(第二军医大学)第二附属医院皮肤科, 上海 200003

**[摘要]** **目的** 探索一种在小数据量条件下提高医学诊断神经网络模型准确率和泛化能力的方法, 解决在指甲病图像计算机辅助诊断中由于训练数据量小而导致神经网络模型性能不佳的问题。**方法** 提出融合实例分割与细粒度特征分类的双模型策略, 采用第一届全国数字健康创新应用大赛健康医疗大数据主题赛基于图像的指甲病智能诊断模型任务数据集训练基于双模型策略的神经网络模型, 该任务数据集涵盖甲母痣、甲沟炎、银屑病甲、甲真菌病、甲下出血、甲黑线、甲周疣、甲黑素瘤 8 类指甲病, 各类别不平衡。评估双模型策略的诊断性能, 并与相同软、硬件训练条件下单模型策略[图像分类模型(ResNet50、Swin Transformer)和基于快速区域卷积神经网络(Faster R-CNN)的目标检测模型]进行比较。**结果** 纳入任务数据集包括甲母痣 210 例、甲沟炎 186 例、银屑病甲 69 例、甲真菌病 203 例、甲下出血 149 例、甲黑线 71 例、甲周疣 93 例、甲黑素瘤 67 例共 1 048 例样本, 其中 90% 的样本用于训练不同策略的模型, 10% 用于评估模型。基于 ResNet50 的图像分类模型的 micro F1 值为 0.324, 基于 Swin Transformer 的图像分类模型为 0.381, 基于 Faster R-CNN 的目标检测模型为 0.572, 基于双模型策略的 Mask R-CNN+Swin Transformer 模型为 0.714。双模型策略预测各指甲病的准确度为甲母痣 80.95% (17/21)、甲沟炎 89.47% (17/19)、银屑病甲 100.00% (7/7)、甲真菌病 70.00% (14/20)、甲下出血 73.33% (11/15)、甲黑线 14.29% (1/7)、甲周疣 55.56% (5/9)、甲黑素瘤 42.86% (3/7)。双模型策略在该任务 1 000 例测试集中的 micro F1 值为 0.844。**结论** 双模型策略可以有效结合功能不同的模型, 更好地完成小数据量训练条件下的指甲病图像智能诊断任务。

**[关键词]** 指甲病; 智能诊断; 神经网络; 实例分割; 细粒度特征分类

**[引用本文]** 陈军晓, 殷杰, 扈东营, 等. 双模型策略在指甲病图像智能诊断中的应用[J]. 海军军医大学学报, 2024, 45 (8): 981-989. DOI: 10.16781/j.CN31-2187/R.20240166.

### Application of dual-model strategy in image intelligent diagnosis of nail diseases

CHEN Junxiao<sup>1</sup>, YIN Jie<sup>1</sup>, HU Dongying<sup>2</sup>, WU Zhao<sup>1</sup>, ZHU Xiuyan<sup>1</sup>, WANG Shiyong<sup>1\*</sup>

1. Department of Information, The Third Affiliated Hospital of Naval Medical University (Second Military Medical University), Shanghai 200438, China

2. Department of Dermatology, The Second Affiliated Hospital of Naval Medical University (Second Military Medical University), Shanghai 200003, China

**[Abstract]** **Objective** To explore a method to improve the accuracy and generalization ability of medical diagnostic neural network models under conditions of small data volumes, and to address the issue of poor neural network model performance in computer-aided diagnosis of nail diseases due to limited training data. **Methods** A dual-model strategy integrating instance segmentation with fine-grained feature classification was proposed. The neural network model based on dual-model strategy was trained using the dataset of Image-Based Intelligent Diagnosis of Nail Disease Model task of the first National Digital Health Innovation Application Competition & Health and Medical Big Data Theme Competition. This dataset covered 8 types of nail diseases, including nail matrix nevi, paronychia, nail psoriasis, onychomycosis, subungual hemorrhage, melanonychia, periungual warts, and nail melanoma, with class imbalance present. The diagnostic performance of the dual-model strategy was evaluated and compared with single-model strategies (image classification models [ResNet50 and Swin Transformer] and target detection model based on faster region-based convolutional neural network [Faster R-CNN]) under the same hardware and software training conditions. **Results** The dataset included 1 048 samples, including 210 cases of nail matrix nevi, 186 cases of paronychia, 69 cases of nail psoriasis, 203 cases of onychomycosis, 149 cases of subungual hemorrhage, 71 cases of melanonychia, 93 cases of periungual warts, and 67 cases of nail melanoma, with 90% used for training various models and 10% for evaluation. The micro F1 score was 0.324 in the image classification model based on

[收稿日期] 2024-03-15

[接受日期] 2024-06-27

[作者简介] 陈军晓, 硕士, 助理工程师. E-mail: chenjunxiao@smmu.edu.cn

\*通信作者( Corresponding author). Tel: 021-81875561, E-mail: luohong\_wsy@126.com

ResNet50, 0.381 in the image classification model based on Swin Transformer, 0.572 in the target detection model based on Faster R-CNN, and 0.714 in the dual-model strategy model Mask R-CNN+Swin Transformer. The accuracy rates for diagnosing different nail diseases in the dual-model strategy were: nail matrix nevi 80.95% (17/21), paronychia 89.47% (17/19), nail psoriasis 100.00% (7/7), onychomycosis 70.00% (14/20), subungual hemorrhage 73.33% (11/15), melanonychia 14.29% (1/7), periungual warts 55.56% (5/9), and nail melanoma 42.86% (3/7). The micro F1 score for evaluating the dual-model strategy on a test set of 1 000 cases was 0.844. **Conclusion** The dual-model strategy can effectively combine models with different functions to well accomplish the task of intelligent diagnosis of nail diseases under small data volume training conditions.

[ **Key words** ] nail disease; intelligent diagnosis; neural network; instance segmentation; fine-grained feature classification

[ **Citation** ] CHEN J, YIN J, HU D, et al. Application of dual-model strategy in image intelligent diagnosis of nail diseases[J]. Acad J Naval Med Univ, 2024, 45(8): 981-989. DOI: 10.16781/j.CN31-2187/R.20240166.

人工智能技术是计算机辅助诊断 (computer-aided diagnosis) 的重要技术方法之一<sup>[1]</sup>。深度学习作为人工智能的关键分支,已经在医学领域展现其强大的能力和潜力<sup>[2]</sup>,其中卷积神经网络 (convolutional neural network, CNN) 和 Transformer 等神经网络模型的表现尤为突出<sup>[3]</sup>。神经网络模型经过医学数据训练后能全自动提取数据特征、挖掘数据规律、输出预测结果。CNN 通过模拟人类视觉系统的机制,利用卷积层对图像进行处理,有效识别出图像中的局部特征,并通过池化层 (pooling layer) 降低特征维度,从而实现对图像特征的高效提取和学习。CNN 在目标检测、实例分割任务上经历了从区域卷积神经网络 (region-based convolutional neural network, R-CNN) 到快速区域卷积神经网络 (faster region-based convolutional neural network, Faster R-CNN) 再到掩码区域卷积神经网络 (mask region-based convolutional neural network, Mask R-CNN) 的演进<sup>[4]</sup>。由于 Mask R-CNN 的出色性能,其作为基础模型被改进后用于检测或分割医学图像中的器官、异常组织、细胞等,取得了良好的应用效果<sup>[5-8]</sup>。Transformer<sup>[9]</sup>是一种基于自注意力机制的神经网络模型,其核心特征的自注意力机制 (self-attention mechanism) 提高了对整个序列的全局理解能力。Swin Transformer<sup>[10]</sup>在 Transformer 架构上引入了层级结构和移位窗口机制,使得模型能够以更高效的方式处理图像中的局部和全局信息。层级结构使网络能够逐渐扩大感受野,移位窗口则提供了一种灵活的方式来聚焦于图像的不同部分。这些特性使 Swin Transformer 特别适合处理细粒度特征分类任务。已经有研究将 Swin Transformer 用于基于临床数据的全自动疾病分类,如甲状腺结节分类、肺癌病理亚型图像识

别、癌症分期等,均取得了理想效果<sup>[11-13]</sup>。

神经网络模型降低了诊断的主观性,提高了诊断效率,但是一般需要大量有效的数据使神经网络模型学习相关的特征规律。而在医学领域,尤其是对于特定类型的疾病,由于人群中发病率较低和医学伦理等因素,不容易收集相应病种的大规模且有专家标注的高质量数据集,可用于训练神经网络模型的医学数据规模有限。在基于图像智能诊断指甲病任务中,还面临着许多挑战。一方面,由于任务提供的指甲照片背景复杂多变,模型算法在识别病变区域时容易受到干扰,影响诊断的准确性。另一方面,指甲病变类型之间的细微差别要求模型具有更高的细粒度特征区分能力。

为克服以上的任务难点,提出一种融合实例分割与细粒度特征分类的双模型策略。该策略将复杂的诊断任务分解为实例分割与细粒度特征分类2个简单任务,降低训练优质神经网络模型的数据量需求,大幅提高预测性能。首先通过实例分割模型,从输入图像的复杂背景中提取出指甲区域,最大限度减小背景对诊断的噪音影响。随后,细粒度特征分类模型对提取的指甲区域进行深度分析,准确识别和区分健康指甲和各类病甲。最后,汇总各指甲区域结果,输出图像的诊断意见。双模型策略在第一届全国数字健康创新应用大赛健康医疗大数据主题赛<sup>[14]</sup>基于图像的智能诊断指甲病模型任务中得到验证,能够在指甲病图像训练集规模不大的情况下训练出具备良好准确性和泛化能力的神经网络模型,完成基于图像的诊断任务。

## 1 材料和方法

### 1.1 训练数据

1.1.1 数据来源 训练数据来自第一届全国数字健

康创新应用大赛健康医疗大数据主题赛基于图像的指甲病智能诊断模型任务数据集。任务要求为通过真实环境拍摄的手/脚照片诊断所患的指甲病。任务数据集为真实情况下用手机相机拍摄的8类指甲病照片共2000张,包括甲母痣210例、甲沟炎186例、银屑病甲1021例、甲真菌病203例、甲下出血149例、甲黑线71例、甲周疣93例、甲黑素瘤67例,照片的拍照视角、背景及大小均不一致。由于患者病甲数量与其所患指甲病有关联性,照片中常见甲母痣、甲黑素瘤、甲下出血影响1个

指甲,甲沟炎、甲周疣影响1~2个指甲,甲真菌病、甲黑线影响1~3个指甲,银屑病甲影响多个指甲。任务数据集包含经过皮肤病专家标注的每例照片的指甲病类型,每张照片有且仅有1个指甲病标注。任务数据集样本示例见图1。数据集中指甲作为判断指甲病的主体区域,最小仅占5%,最大达95%。该任务另外有测试集1000张照片,用于一次性测试模型性能。任务测试集的样本分布与任务数据集一致。

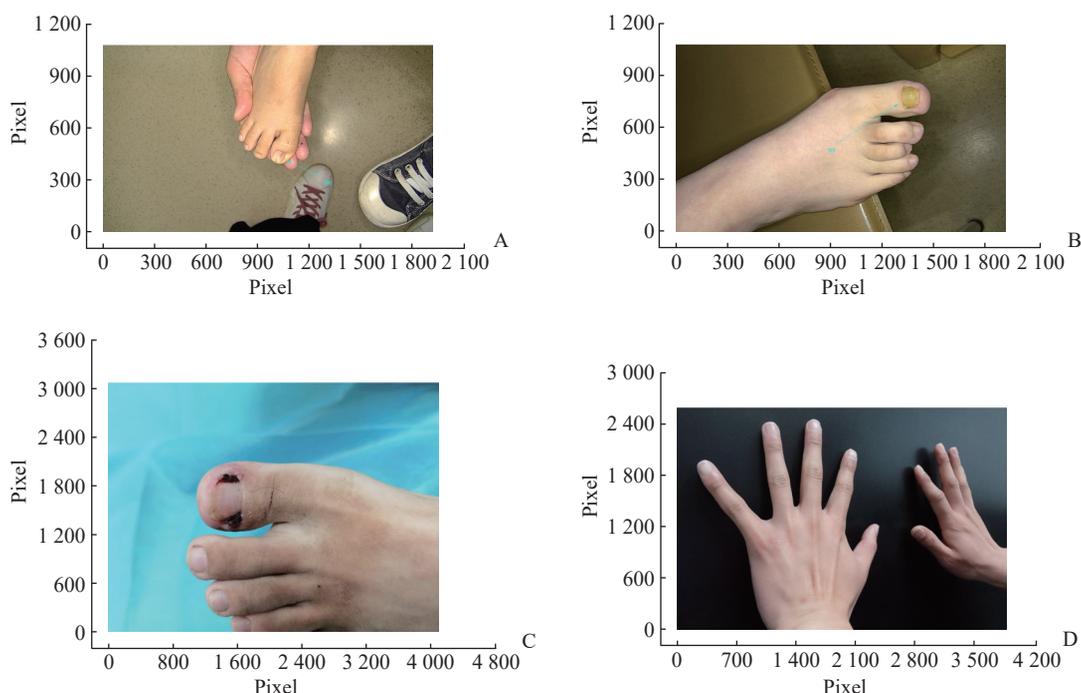


图1 诊断指甲病的训练数据示例

Fig 1 Training data examples for diagnosing nail diseases

A: Nail matrix nevi (arrow); B: Onychomycosis (arrow); C: Paronychia; D: Nail psoriasis.

纳入任务数据集中甲母痣210例、甲沟炎186例、银屑病甲69例、甲真菌病203例、甲下出血149例、甲黑线71例、甲周疣93例、甲黑素瘤67例共1048例样本为本研究数据集。考虑任务时间限制及早期实验中双模型对银屑病甲的预测仅需40例样本训练就能达到100%的准确率,因此未纳入任务数据集中的全部银屑病甲样本。

1.1.2 数据集准备 双模型策略将指甲病智能诊断任务分解为实例分割与细粒度特征分类2个串行的简单任务,需要分别准备数据集用于实例分割模型和细粒度特征分类模型的训练、验证和测试。数据集加工的操作流程见图2,主要步骤为:(1)由于

每类指甲病样本数量不平衡,依照分层抽样法划分90%的数据943例样本为训练集,10%的数据105例样本为测试集。(2)在训练集原图的指头区域用多边形掩码(mask)标注,原图和多边形掩码标注构成实例分割训练集。(3)用多边形掩码标注裁剪训练集原图,每张原图裁出若干张指头图像,并用黑色像素填充指头图像长或宽,使长宽比为1:1,每张指头图像标注9种样本类型中的1类(8类指甲病和健康指甲)。健康指甲图像为原图中未见异常的指头区域经二次加工获得。(4)指头图像和指类标注构成细粒度特征分类训练集。

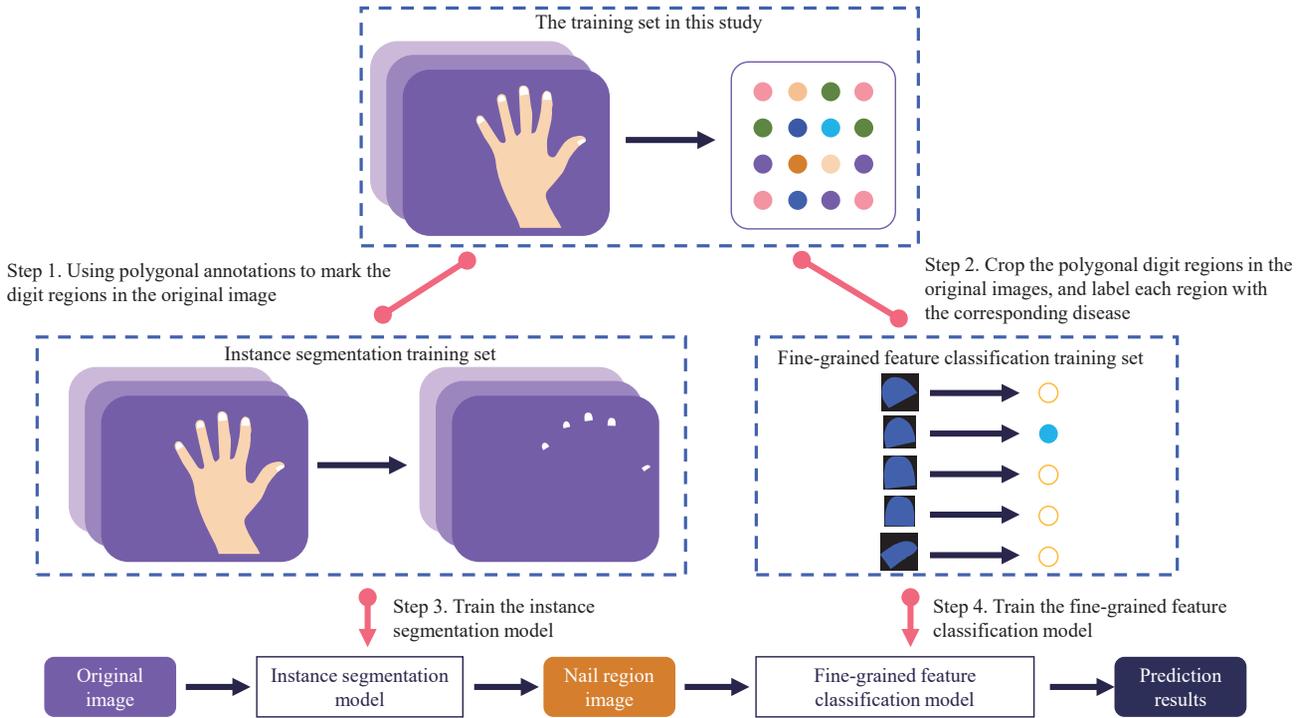


图2 数据集的加工流程

Fig 2 Processing workflow of dataset

1.2 双模型策略 双模型策略基于2个阶段的分析策略：实例分割和细粒度特征分类。第一阶段，实例分割模型负责分离输入原图中的诊断区域（指头）与非诊断区域，以此减少背景噪声并突显病变特征。第二阶段，分割得到的指头图像被送入细粒

度特征分类模型，该模型识别和区分细微的病变特征，输出疾病分类。这2个阶段紧密结合共同构成了完整的诊断流程。最后，合并同一张图像中所有指头图像的分类结果，得到诊断预测结果。见图3。

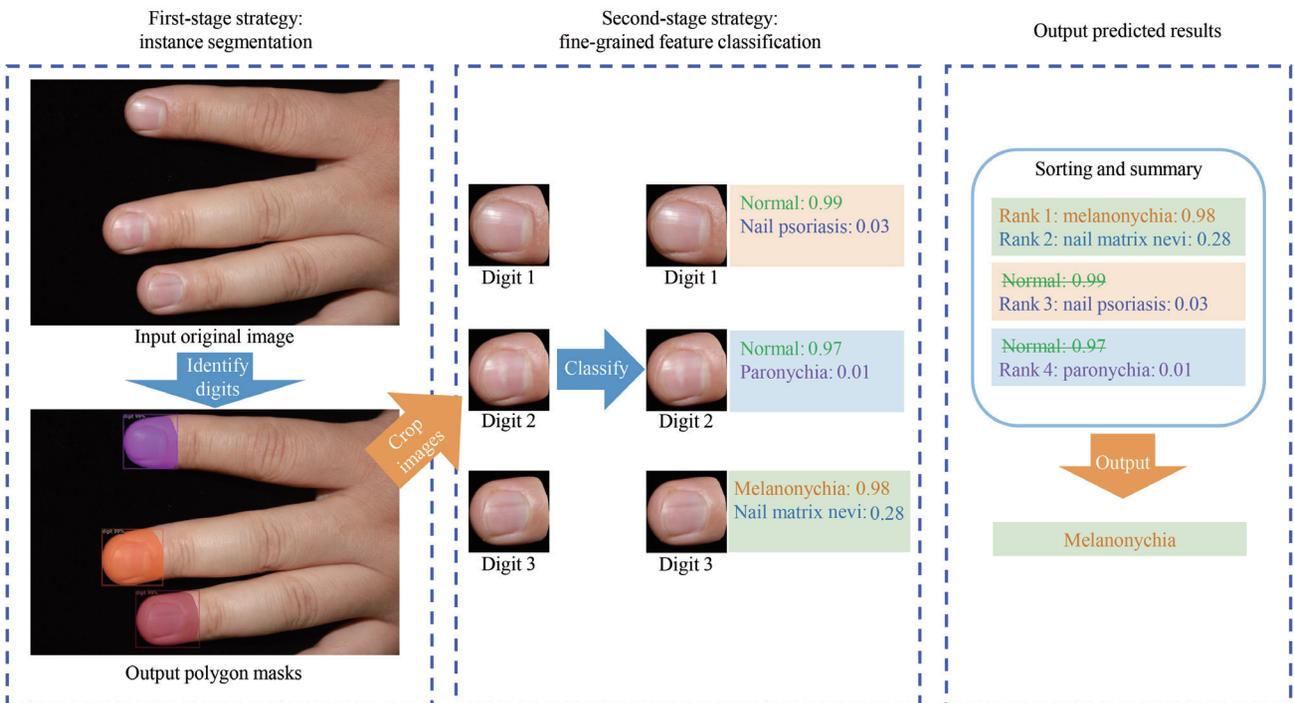


图3 双模型策略的2个分析预测阶段和预测结果输出

Fig 3 Two analytical prediction stages and prediction result output of dual-model strategy

1.3 策略设计 为了充分利用有限的的数据资源并提高模型的泛化能力,在双模型策略中采用迁移学习法构建实例分割和细粒度特征分类模型。迁移学习是一种有效的深度学习技术,其通过将已在大型数据集(如 Microsoft COCO、ImageNet)上预训练的模型应用于相关但数据规模较小的任务,显著提高了学习效率和性能。

1.3.1 Mask R-CNN 实例分割模型 在双模型策略的第一阶段,实例分割模型的选择至关重要,因为它决定了能否从原图中捕获正确的指头区域,直接影响到后续分类模型的输入质量。本研究采用开源平台 Detectron2 上的基于残差网络(residual network, ResNet)和特征金字塔网络(feature pyramid network, FPN)的 Mask R-CNN 模型(Mask R-CNN R-50 FPN 3x)作为实例分割模型。该模型结合了 ResNet 的特征提取能力和 FPN 的多尺度特征整合机制,通过扩展了常规训练周期 3 倍的训练策略能够更深入地学习和适应数据集的特性。该模型从图像中精确分割出每个指头区域并生成高质量的区域掩码,为后续的细粒度特征分类奠定了基础。

1.3.2 Swin Transformer 细粒度特征分类模型 在双模型策略的第二阶段,采用 PyTorch 图像模型库(PyTorch image model, TIMM)上的微型版 Swin Transformer 模型(Swin\_S3\_Tiny\_224)执行细粒度特征分类任务。Swin Transformer 是一种基于 Transformer 架构的新型网络,其引入了层级结构和移位窗口机制,在处理图像任务时更加高效。Swin\_S3\_Tiny\_224 作为 Swin Transformer 的一个轻量级版本,尺寸较小,耗费计算资源较低,但仍然保持了高效的特征提取能力,能够捕捉图像中的细微差异,适用于细粒度特征分类任务。在本研究中,Swin\_S3\_Tiny\_224 模型用于对实例分割后得到的指头图像进行详细的特征分析和疾病分类,以准确区分不同类型的指甲病。

1.3.3 双模型融合 在双模型策略中,实例分割模型和细粒度特征分类模型以低耦合形式进行融合。实例分割模型将原图的指头区域分割出来,随后细粒度特征分类模型逐个分析指头区域,预测指头区域的疾病类型。在综合每个指头区域的指甲病预测结果后,排除正常指甲,置信度最高的预测类型为原图的指甲病诊断结果。这种低耦合的融合策略既

充分利用了 2 个模型在各自领域的优势,即实例分割模型的空间精度和细粒度特征分类模型分类精度,又使 2 个模型可以独立训练,从而更佳地完成任务。

## 2 实验

2.1 实验条件 实验服务器硬件配置为 Intel Xeon Silver 16 核 CPU、64 GB 内存、NVIDIA Tesla V100 16 GB GPU、Ubuntu 20.04 LTS 操作系统,并安装有 CUDA 11.7、Python 3.8、JupyterLab 3.6 等软件。

### 2.2 实验过程

2.2.1 数据标注 在 Labelme 5.3.1 标注工具中使用多边形对训练集原图中的指头区域进行精细标注,该标注用于训练实例分割模型,使模型学习如何精确地将诊断区域与非诊断区域分离。由于人工标注多边形指头区域比较耗时,本研究采取自学习(self-training)策略。首先人工标注 200 张图像中的指头区域,训练初步的实例分割模型后,用其辅助标注剩余的训练集原图,最后人工修正多边形标注,总体耗时相较于单纯人工标注缩短了 90%。训练集原图和多边形标注转换为 COCO 数据集格式后构成第一阶段实例分割训练集。

将标注文件中照片的多边形区域复制一份,在 Labelme 5.3.1 标注工具中标注该指头区域的指甲病类型。用 OpenCV 工具沿多边形标注裁剪原图得到的指头图与对应的指甲病标注构成第二阶段细粒度特征分类数据集。

2.2.2 模型训练 实例分割模型训练采用 Detectron2 平台的 Mask R-CNN R-50 FPN 3x 框架和预训练权重。划分训练集 20% 的数据用于验证,其他用于训练模型。优化器为随机梯度下降(stochastic gradient descent)算法,初始学习率为 0.002 5,训练批量大小为每批次 2 个样本,总迭代次数为 1 400 次。此外,为了更好地调节和防止过拟合,采用了早停机制,并在每 100 次迭代保存 1 次模型,以便从中选取最优模型。训练过程中,每 10 次迭代进行 1 次可视化,及时监控模型学习进度。在 500 次迭代时进行学习率衰减,以细化模型的调整。使用图像分割(segmentation) 75% 交并比(intersection over union) 阈值下的平均精度 AP75 作为主要的验证精度指标,每训练迭代 200 次后通过 COCOEvaluator 进行模型评估,确保模型的有效

性和准确性。模型训练结果见图4,在第800次迭代训练时,验证精度开始稳定在0.98,损失值下降趋势大幅减缓,考虑避免过拟合,选择第800次迭代的模型为最佳实例分割模型。

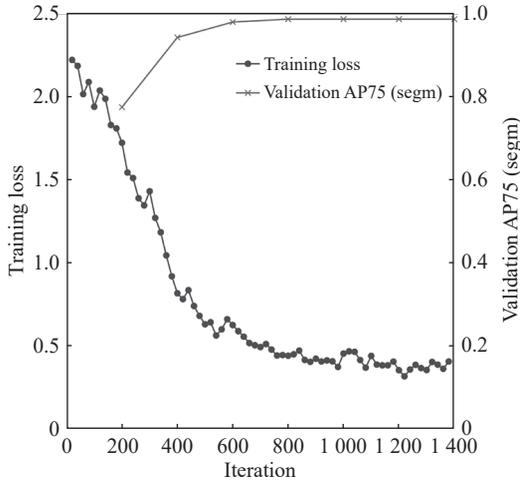


图4 实例分割模型训练结果

Fig 4 Training results of instance segmentation model

AP75: Average precision at intersection over union threshold of 0.75; segm: Segmentation.

细粒度特征分类模型训练采用TIMM上的Swin\_S3\_Tiny\_224模型作为基础模型。同样划分训练集20%的数据用于验证,其他用于训练模型。优化器使用自适应矩估计(adaptive moment estimation)算法,初始学习率为0.0001,为避免过拟合设置权重衰减为0.001,损失函数为交叉熵损失(cross-entropy loss function)。由于指甲区域的各指甲病数量不平衡,每个训练周期(epoch)采用分层随机加载的策略,使模型充分训练每类数据。训练过程中由ReduceLROnPlateau调度器动态调整学习率,当4个周期内模型性能没有改善,学习率将降低80%。对输入的指甲区域图像进行预处理和数据增强,包括白平衡增强、统一尺寸(224像素×224像素)、50%概率水平翻转和垂直翻转、10°以内随机旋转,然后根据ImageNet数据集均值和标准差进行标准化处理,提高模型对图像的理解能力、增加数据的多样性及改善模型在面对真实世界数据时的性能和鲁棒性。每批次训练32个样本,共20个训练周期,每个训练周期保存检查点,方便训练后选择最优模型。模型训练结果见图5,在第14个训练周期后准确率稳定在0.89附近,且训练损失不再下降,选择其为最佳细粒度特征分类模型。

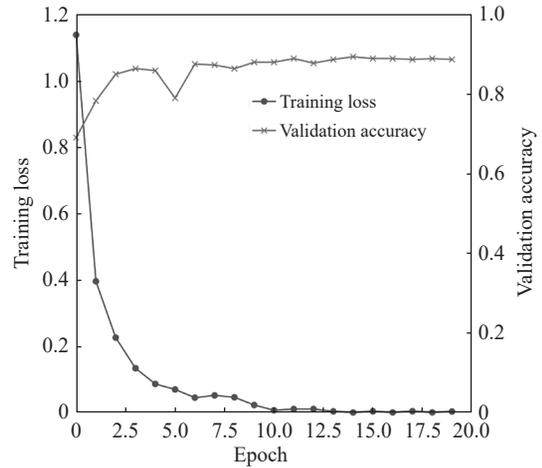


图5 细粒度特征分类模型训练结果

Fig 5 Training results of fine-grained feature classification model

2.2.3 模型预测 训练得到最佳的实例分割模型和细粒度特征分类模型按照图2的组合形式完成指甲病图像诊断任务。模型预测步骤如下:(1)将预测原图输入Mask R-CNN实例分割模型,得到原图中用于诊断的指甲区域的多边形掩码;(2)用OpenCV工具按照多边形掩码裁剪原图并扩展,得到若干包含指甲的正方形图;(3)将指甲图像输入Swin Transformer细粒度特征分类模型,得到每个指甲图像的指甲病预测结果;(4)汇总原图中每个指甲图像的指甲病预测结果,去除正常类型,取置信度最高的结果作为此图的疾病诊断。

### 2.3 实验结果

2.3.1 双模型策略与单模型策略的性能比较 采用单模型策略和双模型策略在本研究训练集中训练基于图像的指甲病智能诊断模型,并比较它们在本研究测试集上的性能。单模型策略包括基于ResNet50的图像分类模型、基于Swin Transformer的图像分类模型、基于Faster R-CNN的目标检测模型。其中目标检测模型需要额外人工标注,在原图上用矩形标注指甲区域,每个区域同时指明该指甲属于9种类型(8种指甲病和健康指甲)中的哪一种。训练各模型的图像数据类型分布见表1。使用单模型策略ResNet50或Swin Transformer图像分类模型执行该指甲病诊断任务的模型评估指标micro F1值分别仅为0.324和0.381,基于Faster R-CNN的目标检测模型的micro F1值为0.572;双模型策略独立训练了2个功能模型,2个模型协同完成指甲病诊断任务,

micro F1 值达到 0.714, 性能超出了实验中最好的单模型策略——基于 Faster R-CNN 的目标检测模型 24.8%。

表 1 各模型的训练集中指甲病的类型分布

Tab 1 Distribution of nail disease types in training set for each model

Type of nail diseases	Single-model strategy			Dual-model strategy	
	ResNet50 (original images)	Swin Transformer (original images)	Faster R-CNN (rectangular region of digits)	Mask R-CNN (polygonal region of digits)	Swin Transformer (polygonal region of digits)
Nail matrix nevi	189	189	201	201	201
Paronychia	183	183	191	191	191
Nail psoriasis	167	167	231	231	231
Onychomycosis	134	134	277	277	277
Subungual hemorrhage	84	84	144	144	144
Melanonychia	64	64	71	71	71
Periungual warts	62	62	113	113	113
Nail melanoma	60	60	60	60	60
Normal			878	878	878

R-CNN: Region-based convolutional neural network.

2.3.2 双模型策略在不同规模大小数据集上的性能比较 为了观察双模型策略在不同规模数据集上的表现, 分别采用本研究测试集 105 例样本和任务测试集 1 000 例样本测试双模型策略的性能。本研究测试集的测试对象为在本研究训练集 20%、40%、60%、80%、100% 规模上训练的基于双模型策略的模型。不同规模的数据集产生方式为有放回地分层抽取。由于任务限制, 任务测试集仅用于一次性测试, 测试对象为 100% 规模训练的模型。模型评估指标为 micro F1 值。由图 6 可见, 双模型策略在本研究训练集 20% 规模训练时 micro F1 值为 0.525, 已经优于训练集 100% 规模训练的图像分类模型。训练规模增至 100% 时 micro F1 值提升至 0.714。双模型策略在任务测试集上的 micro F1 值为 0.844。

2.3.3 双模型策略对不同指甲病的预测结果 使用本研究测试集 105 例数据测试双模型预测各类指甲病的准确度, 结果为甲母痣 80.95% (17/21)、甲沟炎 89.47% (17/19)、银屑病甲 100.00% (7/7)、甲真菌病 70.00% (14/20)、甲下出血 73.33% (11/15)、甲黑线 14.29% (1/7)、甲周疣 55.56% (5/9)、甲黑素瘤 42.86% (3/7)。为了探索双模型策略对不同类型、不同训练样本数量的指甲病的预测准确程度, 对本研究训练集上双模型训练数据的指甲病类型分布和本研究测试集上双模型的预测准确度进行分析, 结果见图 7。双模型策略对各指甲病的预测准确度受到训练样本量的影响较

显著, 甲母痣、甲沟炎、银屑病甲、甲周疣、甲黑素瘤的预测准确度与训练样本量呈正相关。

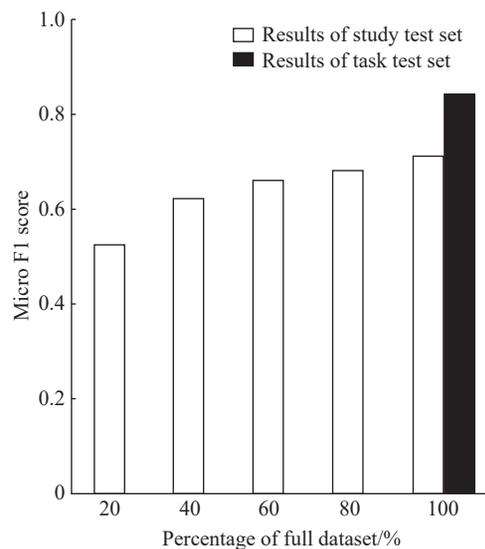


图 6 不同规模测试集上双模型策略的性能

Fig 6 Performance of dual-model strategy in test sets with different scales

### 3 讨论

皮肤病学是一门高度依赖皮损直观表现和形态的学科, 相关疾病的临床诊断很大程度上受到皮肤科医生知识储备和临床经验的影响。然而, 由于皮肤病种类繁多, 患者数量庞大, 皮肤科专科医生相对短缺, 单纯依靠医生肉眼观察进行诊断存在一定的误诊率。组织病理学检查是皮肤病诊断的金标准, 但由于是有创操作, 部分患者依从性较差。

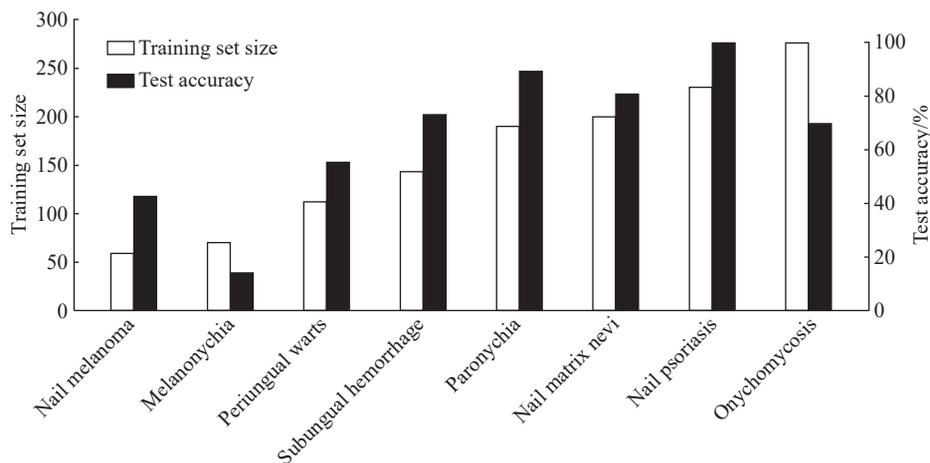


图7 双模型策略的各类型指甲病的训练样本量及测试集预测准确度

Fig 7 Training sample sizes and test set prediction accuracy of dual-model strategy for each nail disease

近年来,人工智能技术蓬勃发展,其已在医学图像识别、疾病诊断、药物研发等多个领域表现出重要价值,而皮肤病学由于具有更大的临床、皮肤镜和组织病理学图像数据库,人工智能应用前景和潜力巨大。指甲是重要的皮肤附属器,指甲病临床种类繁多,一种指甲病可以有多种不同的表现,多种指甲病又可能呈现相同的体征,导致指甲病诊疗难度较大;与此同时,指甲病几乎没有种族差异的特点是采用人工智能技术进行诊断的先天优势<sup>[15]</sup>。Yilmaz等<sup>[16]</sup>通过使用VGG16和InceptionV3模型对氢氧化钾试验检测的显微图像进行分析,证明了深度CNN在快速自动检测指甲真菌感染方面的有效性,并与16名皮肤科医生的诊断结果比较,发现其具有较高的准确率和特异性。Folle等<sup>[17]</sup>开发了一种基于Transformer的深度神经网络,用于自动化量化修改版银屑病甲严重程度指数评分,该研究通过自动关键点检测系统从1154张指甲照片中提取数据,并证明其展示出与人工评分高度相关的精确预测结果。

本研究选取甲母痣、甲沟炎、银屑病甲、甲真菌病、甲下出血、甲黑线、甲周疣及甲黑素瘤等8种指甲病数据训练基于双模型策略的模型,用于进一步验证神经网络模型用于指甲病智能诊断的可行性。

双模型策略在指甲病图像智能诊断任务中展现出显著的优势。传统的单模型策略在数据规模有限的情况下处理复杂背景和细粒度特征时表现不佳。双模型策略通过将复杂任务分解成2个简单任务,有效解决了这一问题,相较于单模型策略性能

大幅提高。实例分割模型在第一阶段发挥关键作用,它能准确地从复杂的背景中分离出指头区域,为后续的疾病诊断提供了清晰、专注的输入。这一步骤减少了非诊断区域对最终诊断结果的干扰,使模型聚焦分析病变区域。细粒度特征分类模型在第二阶段充分发挥作用,有效区分了各种细微的病变特征。这一步骤体现了细粒度特征分类模型在处理细节层面上的精准性,特别是有效识别和分类多种肉眼不易区分的指甲病变类型。此外,迁移学习在有限数据环境下提高了模型训练的效率和效果。预训练模型的引入显著加快了学习速度,同时保持了模型对新数据的高度灵敏性。

在本研究测试集上,基于双模型策略的模型micro F1值为0.714,优于3种单模型策略。分析照片原图可知指头区域以外的图像会对模型分析指甲病造成干扰,在数据集规模不大的情况下,图像分类模型难以从整张图像中挖掘出疾病特征。基于Faster R-CNN的目标检测模型由于进行了额外的目标标注,模型学习从原照片寻找各类指甲的空间位置,性能比图像分类稍好,但区分指甲病类型的能力较弱,其micro F1值为0.572。双模型策略独立训练了2个功能模型,其中Mask R-CNN实例分割模型学习寻找所有指头区域的空间位置和精确形状,Swi Transformer细粒度特征分类模型学习对指头精细区域进行分类,2个模型协同较好地完成了指甲病的诊断任务,micro F1值超出实验中最好的单模型策略(基于Faster R-CNN的目标检测模型)24.8%。

训练样本量增加对双模型策略有积极影响。

指甲病的特点也是预测准确率的影响因素。任务数据集中甲黑线、甲真菌病的外观特征不如其他指甲病典型,特征模型较难学习,容易误判。甲黑素瘤对指甲的形状、颜色改变影响显著,虽然样本量较小,但是相对容易识别。在任务测试集上,基于双模型策略模型的micro F1值为0.844,优于在本研究测试集上的结果,主要原因是2个测试集中银屑病甲例数占比不同。由于银屑病甲训练样本量足够用于双模型策略模型学习,因此双模型策略对银屑病甲的预测准确率高,在任务测试集上有更好的成绩。

双模型策略的输入数据为单模态,与实际线下诊疗进行多维度全方位诊查不同,适合作为医疗条件较差的地区群众快速自我预检指甲病的辅助手段。根据模型对各指甲病的预测准确率,其用以诊断甲母痣、甲沟炎、银屑病甲更容易使人信服,暂不适合用于诊断甲黑线、甲黑素瘤、甲周疣。应该增加数据量较小的指甲病样本,扩大指甲病类型,增加输入模态,使模型挖掘更全面的指甲病规律。此外,模型的性能与标注指头区域指甲病类型的质量相关,需要加入专家知识指导标注,提升模型性能。同时,应结合实际临床测试,确保模型在真实世界环境中的可靠性和实用性。通过不断的优化,双模型策略有望在医学图像诊断领域发挥更大的作用,为患者提供更准确、更高效的服务。

## [参考文献]

- [1] CHAN H P, HADJIISKI L M, SAMALA R K. Computer-aided diagnosis in the era of deep learning[J]. *Med Phys*, 2020, 47(5): e218-e227. DOI: 10.1002/mp.13764.
- [2] 郭芳琪,刘晟,徐磊,等.基于AI-SONICTM Thyroid 5.3.3.0的超声图像分析对甲状腺结节恶性风险的预测价值[J]. *海军军医大学学报*, 2024, 45(1): 29-36. DOI: 10.16781/j.CN31-2187/R.20230027.
- [3] GUO F, LIU S, XU L, et al. Ultrasound image analysis based on AI-SONICTM Thyroid 5.3.3.0 to predict malignant risk of thyroid nodules[J]. *Acad J Naval Med Univ*, 2024, 45(1): 29-36. DOI: 10.16781/j.CN31-2187/R.20230027.
- [4] CHEN X, WANG X, ZHANG K, et al. Recent advances and clinical applications of deep learning in medical image analysis[J]. *Med Image Anal*, 2022, 79: 102444. DOI: 10.1016/j.media.2022.102444.
- [5] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. *IEEE Trans Pattern Anal Mach Intell*, 2020, 42(2): 386-397. DOI: 10.1109/TPAMI.2018.2844175.
- [6] CHOI S J, KIM D K, KIM B S, et al. Mask R-CNN based multiclass segmentation model for endotracheal intubation using video laryngoscope[J]. *Digit Health*, 2023, 9: 20552076231211547. DOI: 10.1177/20552076231211547.
- [7] ZHANG R, CHENG C, ZHAO X, et al. Multiscale mask R-CNN-based lung tumor detection using PET imaging[J]. *Mol Imaging*, 2019, 18: 1536012119863531. DOI: 10.1177/1536012119863531.
- [8] PRANGEMEIER T, WILDNER C, FRANÇANI A O, et al. Yeast cell segmentation in microstructured environments with deep learning[J]. *Biosystems*, 2022, 211: 104557. DOI: 10.1016/j.biosystems.2021.104557.
- [9] 王琮智,许梓璧,马祥园,等.基于数据扩增和迁移学习的Mask R-CNN脑CT图像自动分割研究[J]. *中国生物医学工程学报*, 2021, 40(4): 410-418. DOI: 10.3969/j.issn.0258-8021.2021.04.04.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB/OL]. 2017: arXiv: 1706.03762. (2023-08-02)[2024-03-15]. <http://arxiv.org/abs/1706.03762>.
- [11] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 9992-10002. DOI: 10.1109/ICCV48922.2021.00986.
- [12] BAIMA N, WANG T, ZHAO C K, et al. Dense Swin Transformer for classification of thyroid nodules[J]. *Annu Int Conf IEEE Eng Med Biol Soc*, 2023, 2023: 1-4. DOI: 10.1109/EMBC40787.2023.10340827.
- [13] WANG Y, LUO F, YANG X, et al. The Swin-Transformer network based on focal loss is used to identify images of pathological subtypes of lung adenocarcinoma with high similarity and class imbalance[J]. *J Cancer Res Clin Oncol*, 2023, 149(11): 8581-8592. DOI: 10.1007/s00432-023-04795-y.
- [14] 王润媛,陈星材,吴蔚,等.基于增强CT图像和Swin Transformer网络的食管癌T分期智能诊断模型的构建与评估[J]. *陆军军医大学学报*, 2023, 45(16): 1770-1778. DOI: 10.16016/j.2097-0927.202302021.
- [15] 陈军晓,殷杰.第一届全国数字健康创新应用大赛健康医疗大数据主题赛.赛题7:基于图像的指甲病智能诊断模型[EB/OL]. (2023-11-10)[2024-03-15]. <http://www.nhc.gov.cn/mohwsbwstjxxzx/s2908/202311/7a60fd225f7342179efe4c848808b4e5/files/bf213c1e9f4946cc8f46a4430aa3a577.pdf>.
- [16] LIM S S, OHN J, MUN J H. Diagnosis of onychomycosis: from conventional techniques and dermoscopy to artificial intelligence[J]. *Front Med*, 2021, 8: 637216. DOI: 10.3389/fmed.2021.637216.
- [17] YILMAZ A, GÖKTAY F, VAROL R, et al. Deep convolutional neural networks for onychomycosis detection using microscopic images with KOH examination[J]. *Mycoses*, 2022, 65(12): 1119-1126. DOI: 10.1111/myc.13498.
- [18] FOLLE L, FENZL P, FAGNI F, et al. DeepNAPSI multi-reader nail psoriasis prediction using deep learning[J]. *Sci Rep*, 2023, 13(1): 5329. DOI: 10.1038/s41598-023-32440-8.