硕士学位论文

基于临床图像的红斑鳞屑类皮肤疾病辅助诊断研究

Study on the Assistant Diagnosis of Erythematous Scaly Skin Diseases Based on Clinical Images

学　科　专　业　　控制科学与工程

学　科　方　向　　模式识别与智能系统

作　者　姓　名　　何小宇

指　导　教　师　　谢斌 副教授

中　南　大　学

2019 年 5 月

中图分类号 TP391.4　　学校代码 10533
UDC 621.3　　学位类别 学术学位

硕士学位论文

基于临床图像的红斑鳞屑类皮肤疾病辅助诊断研究

Study on the Assistant Diagnosis of Erythematous Scaly Skin Diseases Based on Clinical Images

作者姓名：　何小宇
学科专业：　控制科学与工程
学科方向：　模式识别与智能系统
研究方向：　图像处理与模式识别
二级培养单位：　自动化学院
指导教师：　谢斌 副教授

论文答辩日期_____　　答辩委员会主席_____

中　南　大　学
2019 年 5 月

学位论文原创性声明

本人郑重声明，所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了论文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得中南大学或其他教育机构的学位或证书而使用过的材料。与我共同工作的同志对本研究所作的贡献均已在论文中作了明确的说明。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

作者签名：_____ 日期：_____年___月___日

学位论文版权使用授权书

本学位论文作者和指导教师完全了解中南大学有关保留、使用学位论文的规定：即学校有权保留并向国家有关部门或机构送交学位论文的复印件和电子版；本人允许本学位论文被查阅和借阅；学校可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用复印、缩印或其它手段保存和汇编本学位论文。

保密论文待解密后适应本声明。

作者签名：_____ 导师签名_____

日期：_____年___月___日 日期：_____年___月___日

基于临床图像的红斑鳞屑类皮肤疾病辅助诊断研究

摘要：红斑鳞屑类皮肤疾病是一类临床表现极其相似的皮肤疾病，其鉴别诊断历来是皮肤科的难题。本文提出一项基于临床图像的红斑鳞屑类皮肤疾病辅助诊断研究计划，研究病种包括扁平苔藓、副银屑病、红斑狼疮、湿疹和银屑病，旨在通过计算机视觉的相关算法，开发一套对各类红斑鳞屑类皮肤疾病进行准确分类和定位的辅助诊断系统。具体过程如下：

首先，本文与合作单位湘雅医院一起构建了一个大型的红斑鳞屑类皮肤疾病数据集，包含有 4377 张临床图像，每张图像都含有病变区域的标记以及病人全部病史资料作为支持，是目前公开领域最大的红斑鳞屑类皮肤疾病数据集。

基于此数据集，本文首先尝试了基于视觉特征的红斑鳞屑类皮肤疾病分类。本文采用支持向量机对 SIFT、HOG、CH、CN 以及 SIFT+CN 等图像特征的分类效果进行了评估。实验结果表明，基于传统视觉特征的红斑鳞屑类皮肤疾病的鉴别准确率普遍都不理想，准确率最高仅达 52.19%，这说明了依靠传统视觉特征的方法无法很好地鉴别各类红斑鳞屑类皮肤疾病。

受到卷积神经网络在医学诊断领域取得的一系列成果的鼓舞，本文还尝试了基于卷积神经网络的红斑鳞屑类皮肤疾病分类。本文引入了 4 种主流的 CNN 结构：InceptionV3、InceptionResNetV2、DenseNet121 和 Xception。本文采用上述 4 个 CNN 模型在是否采用迁移学习的实验条件下进行了 8 组对比实验。实验结果表明，基于 ImageNet 预训练权重的 Xception+ft 模型分类准确率可以达到最高的 81.60%，远优于同样实验条件下的传统视觉特征的分类效果。此外，我们还组织了 CNN 模型与 30 名专业皮肤科医生的“人机对战”实验。实验结果表明，CNN 的总体准确率略微优于医生，但是医生对于每类疾病的识别水平更为平均。值得一提的是，本实验结果只代表了二者对于临床图片的鉴别水平，并不代表实际的诊断能力。

本文还开展了基于多任务学习的红斑鳞屑类皮肤疾病病变区域检测研究，它包括疾病区域定位和疾病分类两个任务。针对皮肤疾病图像具有不同尺度和细粒度分类的特点，本文提出了基于 Faster R-CNN 的多尺度特征融合的改进策略。针对不同的特征融合方式，我们设计了 9 组对比实验。实验结果表明，最优的特征融合组合为

Conv1+Conv3+Conv5，其均值平均精度可以达到 73.60%，这说明了利用卷积神经网络的更广范围的卷积特征的融合能更有效地提升目标检测任务的检测性能。

最后，本文设计并实现了一个红斑鳞屑类皮肤疾病的辅助诊断系统，提供基于临床图像的疾病分类、病变区域检测以及诊断错误信息上报的功能。本系统不仅能为基层医生和患者提供疾病诊断意见，同时还能积累更加全面的患者数据来实现诊断算法的更新和迭代。

图 32 幅，表 8 个，参考文献 58 篇

关键词：红斑鳞屑；辅助诊断；视觉特征；卷积神经网络；迁移学习；多任务学习
分类号：TP391.4

Study on the Assistant Diagnosis of Erythematous Scaly Skin Diseases Based on Clinical Images

Abstract: The erythematous scaly skin disease is a kind of skin disease with extremely similar clinical manifestations, and its differential diagnosis has always been a difficult problem in dermatology. In this paper, we propose an image-based assistant diagnosis research program for erythematous scales skin diseases, which includes lichen planus, parapsoriasis, lupus erythematosus, eczema and psoriasis. The aim is to develop an assistant diagnosis system for the accurate classification and location of various erythematous scales skin diseases through computer vision related algorithms. The specific process is as follows:

Firstly, a large dataset of erythematous scaly skin diseases was constructed with Xiangya Hospital, which contains 4377 clinical images. Each image contains markers of lesion areas and all patient history data as support. It is the largest dataset of erythematous scaly skin diseases in the open field.

Based on this dataset, this paper first attempts to classify erythematous scaly skin diseases based on visual features. In this paper, support vector machine is used to evaluate the classification effect of SIFT, HOG, CH, CN and SIFT+CN image features. The experimental results show that the identification accuracy of erythematous scaly skin diseases based on traditional visual features is generally unsatisfactory, and the highest accuracy is only 52.19%. This shows that the traditional visual features can not be used to identify all kinds of erythematous scaly skin diseases.

Encouraged by a series of achievements of convolution neural network in the field of medical diagnosis, this paper also attempts to classify erythematous scaly skin diseases based on convolution neural network. This paper introduces 4 main CNN structures: Inception V3, Inception ResNetV2, DenseNet121 and Xception. In this paper, 8 groups of comparative experiments were conducted on the 4 CNN models under the experimental conditions of whether or not transfer learning was used. The experimental results show that the Xception+ft model based on ImageNet pre-training weight can achieve the highest classification accuracy of 81.60%, which is much better than the traditional visual feature classification under the same experimental conditions. In addition, we also organized a "man-machine battle" experiment between CNN model and 30 professional dermatologists. The experimental results show that the overall

accuracy of CNN is slightly better than that of doctors, but the recognition level of doctors for each kind of disease is more average. It is worth mentioning that the results of this experiment only represent the level of differentiation between the two clinical pictures, and do not represent the actual diagnostic ability.

This paper also carried out a multi-task learning-based skin lesion detection study of erythematous scaly skin diseases, which includes two tasks: disease location and disease classification. Aiming at the characteristics of skin disease images with different scales and fine-grained classification, this paper proposes an improved multi-scale feature fusion strategy based on Faster R-CNN. For different feature fusion methods, we designed 9 groups of comparative experiments. The experimental results show that the optimal combination of feature fusion is Conv1+Conv3+Conv5, and its average accuracy can reach 73.60%. This shows that the fusion of convolution features with a wider range of convolution neural networks can effectively improve the detection performance of target detection tasks.

Finally, this paper designs and implements an assistant diagnosis system for erythematous scaly skin diseases, which provides the functions of disease classification based on clinical images, lesion detection and reporting of diagnostic error information. This system can not only provide disease diagnosis advice for primary doctors and patients, but also accumulate more comprehensive patient data to update and iterate the diagnosis algorithm.

Keywords: erythematous scales; assistant diagnosis; visual features; convolutional neural networks; transfer learning; multitask learning
Classification: TP391.4

目 录

摘要.....I
Abstract.....III
1 绪论.....1
1.1 研究背景.....1
1.2 国内外研究现状.....2
1.3 研究目的和内容.....4
1.3.1 研究目的.....4
1.3.2 研究内容.....4
1.4 本章小结.....5
2 红斑鳞屑类皮肤疾病数据集.....7
2.1 现有皮肤疾病临床图像数据集.....7
2.2 湘雅皮肤疾病临床图像数据集构建.....9
2.2.1 数据集构建流程.....9
2.2.2 数据采集和清理.....10
2.2.3 数据标注和裁剪.....11
2.3 本数据集的特点及优势.....12
2.4 本章小结.....12
3 基于视觉特征提取的红斑鳞屑类皮肤疾病分类.....13
3.1 视觉特征.....13
3.1.1 SIFT 特征.....13
3.1.2 HOG 特征.....16
3.1.3 CH 特征.....18
3.1.4 CN 特征.....19
3.2 实验方案设计.....21
3.3 实验结果分析.....22
3.4 本章小结.....23
4 基于卷积神经网络的红斑鳞屑类皮肤疾病分类.....24
4.1 基于卷积神经网络的图像分类.....24
4.1.1 卷积神经网络理论.....24
4.1.2 迁移学习理论.....26
4.1.3 本文采用的几种主流的卷积神经网络介绍.....26
4.2 实验方案设计.....30

4.3 实验结果分析.....31
4.3.1 CNN 分类模型对比实验.....31
4.3.2 CNN 模型与皮肤科专家医生对比实验.....33
4.4 本章小结.....34
5 基于多任务学习的红斑鳞屑类皮肤疾病病变区域检测.....35
5.1 基于多任务学习的 Faster R-CNN.....35
5.2 基于多尺度特征融合的 Faster R-CNN.....37
5.2.1 改进动机.....37
5.2.2 网络结构.....39
5.3 实验方案设计.....40
5.4 实验结果分析.....42
5.5 本章小结.....44
6 红斑鳞屑类皮肤疾病辅助诊断系统的设计与实现.....45
6.1 系统功能.....45
6.2 系统设计与实现.....47
6.3 本章小结.....50
7 总结与展望.....51
7.1 论文总结.....51
7.2 工作展望.....52
参考文献.....53
攻读学位期间主要的研究成果.....57
致谢.....59

1 绪论

1.1 研究背景

皮肤疾病是指发生在皮肤及其附属器官疾病的统称。皮肤是人的身体器官中最大的器官，皮肤疾病的类型众多且发病率极高，全球约 11%人口患有各类皮肤疾病[1]。皮肤疾病不但影响身体健康，而且引起恐慌与社会歧视，对患者的心理也造成了严重的负担，已成为困扰全社会的重大公共卫生问题。

红斑鳞屑类皮肤疾病主要是指一组病因尚未完全明了的以红斑或者鳞屑为主要特征的皮肤病，主要包含扁平苔藓、副银屑病、红斑狼疮、湿疹和银屑病等皮肤疾病。这类疾病的临床特征极为接近，但治疗手段差异很大，其鉴别诊断历来是皮肤病科的一个难题，一旦误诊或错过最佳治疗时机很可能会使病情进一步加重。

目前，临床上的红斑鳞屑类皮肤疾病的诊断需要结合临床表现、皮肤镜诊断以及组织病理学检查（活检）等。由于红斑鳞屑类皮肤疾病是一组临床表现相似的皮肤疾病，因此很难单依靠临床表现对患者做出诊断，且对医生诊断水平要求较高。皮肤镜是一种无创的放大检查方法，可以穿过表皮角质层，成像出表皮和浅表真皮的色素沉着和血管等的细微结构。皮肤镜检查最初用于色素瘤的诊断，近年来，皮肤镜已扩展到红斑鳞屑类皮肤疾病的诊断上。例如，利用皮肤镜鉴别银屑病与湿疹[2]：前者的表现为亮红色背景下一致性分布的点状血管或球状血管，可见弥漫分布的白色鳞屑；后者表现为深红色的背景下的局灶性分布，可见黄斑。组织病理学检查，指的是通过环钻或小手术取下小部分病灶皮肤，经过固定、染色等技术在显微镜下或电脑上观察其变化的诊断技术。在皮肤病学领域，组织病理的检查结果一般可以作为患者确诊的“金标准”，但是同时病理学检查会对病人产生轻微的创伤和不适[3]。

由上所述，我们可以知道专业的皮肤科医生资源和医学诊断设备对于红斑鳞屑类皮肤疾病的诊断极为重要。然而，我国却普遍存在着医患比例失衡和医疗资源分配不均的问题。据统计，我国注册的皮肤科专业医生仅 2 万余人，而我国皮肤疾病病人却有近 1.5 亿人。同时，优质医疗资源分布也很不均匀，基层医疗机构几乎没有专业的皮肤科医生。一些皮肤疾病涉及患者个人隐私，患者在没有充分医疗知识的背景下，很容易受到各类营销广告的影响，选择了一些专业医疗技术不足的盈利性私营医院。这些因素都导致了我国大量红斑鳞屑类疾病患者无法得到及时有效的治疗。

随着人工智能技术的进步，各类计算机辅助诊断系统是解决上述医疗供需不

平衡问题的一条有效途径。“人工智能+医疗”已经成为当下人工智能领域最具发展潜力和应用前景的交叉学科之一。本文将先进人工智能技术应用于发病率高、人工鉴别困难的红斑鳞屑类疾病的辅助诊断领域，选择获取方便的临床图像作为研究对象，旨在开发一套仅依靠患者的临床图像就能对各类型红斑鳞屑类皮肤疾病进行准确、稳定分类的辅助诊断系统。

1.2 国内外研究现状

早期的皮肤疾病辅助诊断多采用手动提取特征结合分类器的方式。常见的手工设计特征有：应用于诊断有色的皮损的 ABCD 准则[4]、鉴别黑色素瘤的七点检查表准则(7-point checklist)[5]、6 类具有皮肤疾病临床特征描述能力的图像特征[6]等。这些特征从图像中提取之后，需要经过机器学习中的常用方法进行分类。如支持向量机[7]、随机森林[8]、K-近邻算法[9]、决策树[10]等。这类手工设计特征的方式的优点是其设计的特征都是有临床诊断意义的，但是这种方法往往针对要鉴别的疾病设计特定的特征，不具有推广性，当面对复杂的分类任务时甚至难以设计出有效的鉴别特征。

近年来，GPU 计算能力的提升推动了卷积神经网络（Convolutional Neural Networks, CNN）的快速发展，使得训练深层神经网络成为可能。皮肤疾病辅助诊断领域出现了很多关于利用卷积神经网络进行研究的例子。最具有代表性的是 2017 年的斯坦福团队 Esteva[11]等人的研究。他们利用 InceptionV3 的 CNN 结构在一个约 130000 幅、超过 2000 种皮肤疾病的临床图片数据集上进行训练，在角质形成细胞癌 VS 良性脂溢性角化病、恶性黑色素瘤 VS 普通的痣的鉴别任务上证实 CNN 分类模型可以达到专业皮肤科医生媲美的水平。受此鼓舞，许多基于卷积神经网络的研究涌现出来，且在各种皮肤疾病医学图像分析领域都有应用。

在临床图像方面，Han[12]等人针对皮肤癌的良性与恶性临床图像，利用 ResNet152 的 CNN 结构训练出了 12 种良恶性皮肤癌的分类模型。与 16 名专业皮肤科医生进行对比，CNN 模型在某些疾病的分类水平上超过了人类医生。Sun[13]等人收集并整理了美国的公开皮肤科网站的临床图像，实现了 CaffeNet 和 VGGNet 在此数据集上的 128 种以及 198 种皮肤病的基准分类。其缺点是分类准确率不高，其总数只有 6584 张，对于 198 类皮肤疾病而言，每类皮肤疾病的图像数量不到 50 张。Liao[14]等人提出了 CNN 模型在不同的皮肤疾病数据集上存在泛化能力差的问题。即在一个数据集上测试效果好，但是在另外的数据集上却不尽如人意，尤其是在不同人种间的数据集上。

在皮肤镜图像方面，Yu[15]等人提出了一个全卷积的残差网络和深度残差网络相结合的方法用于鉴别皮肤镜图像是否为黑色素瘤，在 ISBI 2016 的数据集上，

该方法分类准确率为 0.855，在该比赛中排名第一。Yang[16]等人提出了一种结合区域平均值池化的线性分类器的方法，任务同样是对皮肤镜图像判定是否为黑色素瘤，在 ISBI 2017 的公共数据集上，该方法的准确率可达到 0.830，通过与其他 12 种算法的对比证明了其网络的有效性。Xie[17]等人针对黑素瘤的皮肤镜图像的良恶性二分类问题提出了一种基于自生成神经网络和集成神经网络的方法，该方法在由中国人民解放军空军总医院收集的黄种人数据集和国外的白种人数据集上的测试准确率分别为 94.17%和 91.11%，结果验证了该方法优于其他集成模型。Wang[18]等人的研究内容是针对色素痣和脂溢性角化的皮肤镜图像分类，文章采用了 ResNet50 的 CNN 结构在 3015 幅色素痣和 2079 幅脂溢性角化的皮肤镜图像数据集上进行二分类模型的训练，在 30 例色素痣和 30 例脂溢性角化的测试集上与 95 名专业皮肤科医生的判别准确率进行了对比。最终的结果是 CNN 的识别率可以达到 88.33%，略高于医生平均水平的 84.47%。文章采用的数据集都是来自于北京协和医院皮肤科门诊收集的数据，标签都经过了专业医生的严格筛选。

在皮肤病理图像方面,Wang[19]等人利用构建的 8 层卷积神经网络实现了对组织病理图像中性粒的识别，并通过建立实验表明此种特征优于手工特征。而 Pal[20]等人在对银屑病组织病理图像中芒罗小脓肿的识别问题上，先使用 U-Net 网络模型对角质层进行分割，之后采用 CapsDeMM 网络从提取的切片中识别中性粒细胞。

通过总结上面的一些文献，我们可以总结出以下几点研究现状：

（1）从研究病种上来看，目前本领域的研究对象多基于皮肤肿瘤，尤其以黑素瘤居多，缺少关于红斑鳞屑类皮肤疾病的诊断研究。我们认为产生这种现象的主要原因可能是红斑鳞屑类皮肤疾病的视觉特征比黑色素瘤更为复杂。从图 1-1 中我们可以看到，黑色素瘤与周围皮肤疾病的对比度很高，大多数颜色为黑色。但是红斑鳞屑类皮肤疾病的颜色更为复杂。在皮损的形状方面，黑色素瘤的病变形状多呈圆形，但红斑鳞屑类皮肤疾病的皮损形状一般并不规则。

（2）从研究的图像的种类上来看，目前的研究多以皮肤镜图像的研究为主。我们认为这主要是因为皮肤镜图像由专业设备进行获取，背景简单，都是基于皮损部位拍摄，且放大倍数大，对病变特征的观察十分清楚。而临床图像背景比较复杂，其拍摄角度和距离的不可控也为其后续的识别造成困难。

（3）从研究方法上来看，借助大型皮肤疾病影像数据库进行迭代训练的深度学习方法占据主流。相比较之前的依靠人工设计特征的机器学习方法，卷积神经网络能自动提取皮肤疾病的特征，并且不受疾病种类和分类任务的限制。

图 1-1 红斑鳞屑类皮肤疾病与黑素瘤的临床视觉表征比较

因此，本文基于以上的研究现状，提出了基于临床图像的红斑鳞屑类皮肤疾病的辅助诊断研究计划，旨在开发一套仅依靠患者的临床图像就能对各类型红斑鳞屑类皮肤疾病进行准确、稳定分类的辅助诊断系统。

1.3 研究目的和内容

1.3.1 研究目的

红斑鳞屑类皮肤疾病是一类临床表现极其相似的一类皮肤疾病，这类疾病的临床特征极为接近，但治疗手段差异很大，其鉴别诊断历来是皮肤科的一个难题。目前，临床上的红斑鳞屑类皮肤疾病的诊断需要结合临床表现、皮肤镜诊断以及组织病理学检查等，对医生诊断水平和医疗设备要求都较高。本文以获取方便的皮肤疾病临床图像为研究对象，旨在通过计算机视觉的相关算法开发一套仅依靠患者的临床图像就能对各类型红斑鳞屑类皮肤疾病（包括五类：扁平苔藓、副银屑病、红斑狼疮、湿疹和银屑病）进行准确、稳定分类的辅助诊断系统。本系统可以有效辅助医生进行诊断，提升基层医生对于红斑鳞屑类皮肤疾病的诊断水平，同时还能为患者提供智能自诊服务，具有很强的科研价值以及社会意义。

1.3.2 研究内容

本文主要研究了如何利用计算机视觉中的相关算法对红斑鳞屑类皮肤疾病的临床图像进行有效的辅助诊断。基于对于目前医学图像鉴别诊断系统的调研，本文的研究内容具体如下：

（1）本文与合作单位湘雅医院一起构建了一个大型的基于中国人种的皮肤疾病数据集，数据集中每张图像都带有病变区域的标记信息，且具有患者的全部医学资料作为支持；

（2）研究图像视觉特征对于红斑鳞屑类皮肤疾病临床特征的描述能力，通过提取与疾病表征相关的颜色、形状和纹理特征，并利用机器学习方法实现基于视觉特征的红斑鳞屑类皮肤疾病分类；

（3）研究基于卷积神经网络的红斑鳞屑类皮肤疾病的分类方法。构建不同的卷积神经网络结构，尝试通过神经网络的训练自动学习疾病鉴别有关的特征，并探究迁移学习的方法对于皮肤疾病分类任务是否具有帮助作用；

（4）研究基于多任务学习的红斑鳞屑类皮肤疾病病变区域检测的方法。尝试利用数据集中的病变区域标记的信息，基于深度学习的方法训练出一个多任务学习的模型，即病变区域的定位任务和疾病类别的分类任务。这里我们将根据本任务的特点对现有的目标检测算法进行改进，使其适应于皮肤疾病的检测；

（5）设计并实现了一个红斑鳞屑类皮肤疾病的辅助诊断系统，根据前面设定的研究对象和研究内容，本系统的输入为红斑鳞屑类皮肤疾病的临床图像，可以提供红斑鳞屑类皮肤疾病的图像分类、病变区域检测以及诊断错误信息上报的功能。本系统计划能为基层医生提供相应的诊断决策服务，为患者提供智能自诊服务，同时还能积累更加全面的患者数据实现诊断算法的更新和迭代。

1.4 本章小结

本文将先进人工智能技术应用于发病率高、人工鉴别困难的红斑鳞屑类疾病的辅助诊断领域，选择获取方便的临床图像作为研究对象，旨在通过计算机视觉中的先进算法构建一套仅依靠患者的临床图像就能对各类型红斑鳞屑类皮肤疾病进行准确、稳定分类的辅助诊断系统。

本文共分七个部分，论文的组织结构如下：

第一章为绪论部分，介绍了本文进行红斑鳞屑类皮肤疾病辅助诊断研究的背景和意义，之后对皮肤疾病辅助诊断领域的研究现状进行综述。最后，对本文的研究目的和研究内容进行简要阐述。

第二章对我们首先对目前在皮肤疾病辅助诊断研究领域现有的临床图像数据集进行了介绍。之后，介绍了我们与合作单位湘雅医院一起构建的湘雅皮肤疾病数据集，重点描述了数据集的采集、清理和裁剪的过程。最终，我们得到了可用于红斑鳞屑类皮肤疾病辅助诊断研究的原始图像数据集——湘雅 ES-Raw 数

据集和更加集中于皮损部位的病变区域数据集——湘雅 ES-Patch 数据集。

第三章对皮肤疾病的临床图像的视觉特征进行分析，主要介绍了 5 类图像视觉特征及其计算方式。基于湘雅 ES-Patch 数据集，本章采用机器学习的特征提取和分类器的方法对不同视觉特征的分类效果进行了对比实验。根据对比实验结果，我们针对皮肤疾病分类任务的特点做出了讨论和分析。

第四章的主要内容是基于卷积神经网络的红斑鳞屑类皮肤疾病分类。本章首先对卷积神经网络和迁移学习的理论基础进行了概述，之后对后面实验用到的几种主流的卷积神经网络结构进行了介绍。基于湘雅 ES-Patch 数据集，本章采用对比实验的方式进行了不同 CNN 模型的疾病分类实验，并与第三章中基于视觉特征的疾病分类效果进行对比。除此之外，我们还在 100 张图像的测试集上将 CNN 分类模型与 30 名专业皮肤科医生进行了读图识别水平的对比，通过组织这一“人机对战”的实验得出 CNN 模型对于红斑鳞屑类皮肤疾病临床图像的鉴别能力的相对水平。

第五章的主要内容是基于多任务学习的红斑鳞屑类皮肤疾病病变区域检测。不同于第二章和第三章的疾病分类实验，本章的研究内容包括疾病区域定位和疾病类型分类两个任务。本章首先介绍了多任务学习的相关理论，接着介绍多任务学习在深度学习领域的典型的应用——Faster R-CNN。针对皮肤疾病病变区域检测的特点，本文对原有的 Faster R-CNN 的特征提取网络提出了多尺度融合的改进策略。实验部分，本章设置了不同特征融合策略下的对比实验，利用目标检测领域的评价指标对实验结果做出分析和讨论。

第六章的主要内容是设计并开发了一款红斑鳞屑类皮肤疾病的辅助诊断系统。本章介绍了系统的主要功能以及具体的设计和实现过程。基于前文的诊断算法研究，本系统能够承担红斑鳞屑类皮肤疾病分类、病变区域检测和诊断错误信息上报等功能。本系统的客户端能很方便地安装在智能手机或平板等移动端设备上，为广大的基层医生和患者等提供智能诊断服务，同时还能实现模型诊断算法的自我更新。

第七章对本文进行总结，并指出本文的优势及创新点。

2 红斑鳞屑类皮肤疾病数据集

本文的研究内容是针对皮肤疾病临床图像的辅助诊断工作，因此我们的前期准备工作是要得到一个图像质量高且标签真实可靠的数据集。本章中，我们首先对目前在皮肤疾病辅助诊断研究领域现有的几个临床图像数据集进行介绍。经过调研之后，我们发现这些数据集的来源主要以欧美人种为主，图像上带有数据集网站的水印，且包含红斑鳞屑类皮肤疾病的图片较少。因此，我们与合作单位湘雅医院一起构建了一个大规模的基于中国人种的皮肤疾病临床图像数据集，下面将具体介绍目前现有数据集的不足以及本数据集的构建过程。

2.1 现有皮肤疾病临床图像数据集

通过对本领域的相关文献的调研，我们发现目前被研究者们广泛采用的公开临床图像数据集主要有以下 7 个：

Dermofit（http://licensing.eri.ed.ac.uk/i/software/dermofit-image-library.html）数据集由英国爱丁堡大学的研究者所提供，数据质量高，被研究者们使用得较为广泛，但是不免费公开。Dermofit 包括 10 类皮肤疾病：日光性角化病（45 张），基底细胞癌（239 张），色素痣（331 张），脂溢性角化（257 张），鳞状细胞癌（88 张），上皮内癌（78 张），化脓性肉芽肿（24 张），血管瘤（97 张），皮肤纤维瘤（65 张）和恶性黑色素瘤（76 张），总计 1300 张。从这些疾病上来看，这 10 类都属于肿瘤类皮肤疾病，该数据集与本文的研究目标不符。

ISIC（www.isic-archive.com）数据集来自于国际皮肤影像协作组织（The International Skin Imaging Collaboration, ISIC），该组织旨在促进皮肤影像数据的诊断能力。ISIC 数据集共包含 23,906 张影像数据，分为 16 个类别，数据即包含皮肤镜图像又包含临床图像，成像质量高且没有水印。数据集内每张图像都包含了患者的一些基本属性，包括患者大约年龄、性别、病史特征和病变尺寸等。然而，数据集中的临床图像总数仅有 100 张，包括 37 个黑素瘤，40 个基底细胞癌和 23 个鳞状细胞癌的图像，这 3 种疾病也都属于皮肤肿瘤这一大类，与本文的研究对象红斑鳞屑类皮肤疾病无关。

Derm101（www.derm101.com）是一个为临床医生提供各种在线资源的综合性网站,具有针对各种皮肤状况的最新治疗建议，以及高质量的诊断图像。Derm101 数据集共包含 525 个类别的 22,979 个临床图像，每张图像包含基本的临床诊断标签和皮损发生的部位信息，这些数据对学术研究是免费公开的。Derm101 数据集最大的优点在于其图像没有加水印信息，方便我们进行后续图像分析和处理。

Dermnet（www.dermnet.com）也是一个提供各种皮肤医学资源的综合性网站，通过文章，照片和视频致力于在线医学教育。Dermnet 图像库有近 18,974 张临床图像，626 种皮肤病，这些数据对学术研究是免费公开的。但是，此数据集中的每个图像只有诊断名称的标签，没有其他标签。

DermIS（www.dermis.net）数据集是由德国海堡大学整理公开的免费公开数据集，该数据库中有 7,172 张图片,共分为 735 类。数据集内每张图像除了有常规的诊断名称标签之外，还有人种信息、病变发生的部位信息以及年龄信息等文字说明。该数据集的缺点是其分类太多，每一类的图像数量并不多，且图像上带有数据集网站的水印。

AtlasDerm（www.atlasdermatologico.com.br）是来自于巴西的一个皮肤疾病临床图像数据集网站，该数据库共有 534 类，9613 张图片。由于地域的限制，该数据集收集的患者大多以黑色人种为主，每张图像只有诊断名称的标签，且图像上带有数据集网站的水印。

Danderm（www.danderm-pdv.is.kkh.dk）是来自于丹麦的一个皮肤疾病临床图像数据集网站。该数据集共有 91 类，1110 张的临床图像。Danderm 数据集存在和 AtlasDerm 一样的缺点，该数据集收集的患者大多是白色人种，同样也只有诊断名称的标签，且含有网站的水印信息。

综上所述，我们将上面提到的数据库的一些特性总结如表 2-1。

表 2-1 皮肤疾病临床图像数据集调研整理

数据集	疾病类别	图像数量	数据来源	有无水印	数据集中红斑鳞屑类皮肤疾病的数量				
					扁平苔藓	副银屑病	红斑狼疮	湿疹	银屑病
Dermofit	10	1300	英国	无	0	0	0	0	0
ISIC	3	100	欧洲	无	0	0	0	0	0
Derm101	525	22979	美国	无	60	0	59	249	270
Dermnet	626	18974	欧洲	有	271	0	207	312	132
DermIS	735	7172	德国	有	212	9	160	245	239
AtlasDerm	534	9613	巴西	有	175	15	141	47	128
Danderm	91	1110	丹麦	有	16	0	17	23	43
湘雅 Raw	543	54098	中国	无	364	348	513	534	2618
湘雅 Patch	543	150223	中国	无	900	900	900	900	900

由表 2-1 可知，对于本文的研究任务，目前现有的皮肤疾病数据集普遍存在着一些明显的缺陷：

1）目前现有数据集的来源主要以欧美人种为主，如白色人种和黑色人种，缺乏大规模的以黄色人种为主的皮肤疾病数据集。而不同人种之间的皮肤疾病发病率、皮肤疾病特点、肤色背景等都存在差异；

2）目前现有数据集的图像上大多都带有各自数据集网站的水印，这些水印信息对于皮肤病变区域的识别和分析会造成干扰；

3）目前现有数据集中包含红斑鳞屑类皮肤疾病的图像较少，尤其是副银屑病，7 个公开的数据集一共都只有 24 张图像，这样的数据集的数量太小了，无效进行后续的疾病辅助诊断研究。

因此，本文与湘雅医院合作自行构建了一个大规模的基于中国人种的皮肤疾病临床图像数据集，并在此基础上开展了红斑鳞屑类皮肤疾病辅助诊断的研究。本文后面研究工作都是基于湘雅 Raw 数据集（原图数据集）和 Patch 数据集（病变区域数据集），这两个数据集的数量也总结在表 2-1 中。其中，出于数据集平衡的考虑，我们在由湘雅 Raw 数据集到湘雅 Patch 数据集的裁剪阶段，将每类皮肤疾病的数量都保持一致即都为 900 张。

可以看到的是，我们的湘雅 Raw 和 Patch 数据集所包含的 5 类红斑鳞屑类皮肤疾病的数量比目前现有的公开数据集都要更大，且都是由高清采集设备采集得到，数据集内每张图像都包含病理诊断和病人全部病史的支持。下文，我们将详细介绍这 2 个数据集的获取过程。

2.2 湘雅皮肤疾病临床图像数据集构建

2.2.1 数据集构建流程

从 2016 年 1 月至 2018 年 12 月，我们通过合作单位湘雅医院在皮肤科门诊对患者数据进行收集，经过 3 年多的时间，我们完成了对现有的湘雅皮肤疾病数据集的采集、整理和标注等工作。最终，我们得到的湘雅 Raw 数据集一共含有 54098 幅，543 类皮肤疾病的临床图像，由 Raw 数据集裁剪得到的病变区域构成的 Patch 数据集共有 150223 幅。本数据集内每幅图片都有真实病例作为支持，含有病人的病理和病史的全部诊断记录。最终，我们从这 2 个数据集中取出红斑鳞屑类的 5 类皮肤疾病的图像，组成了最终用于我们后文实验的湘雅 ES-Raw 和 ES-Patch 数据集，数据集构建的流程如图 2-1 所示。

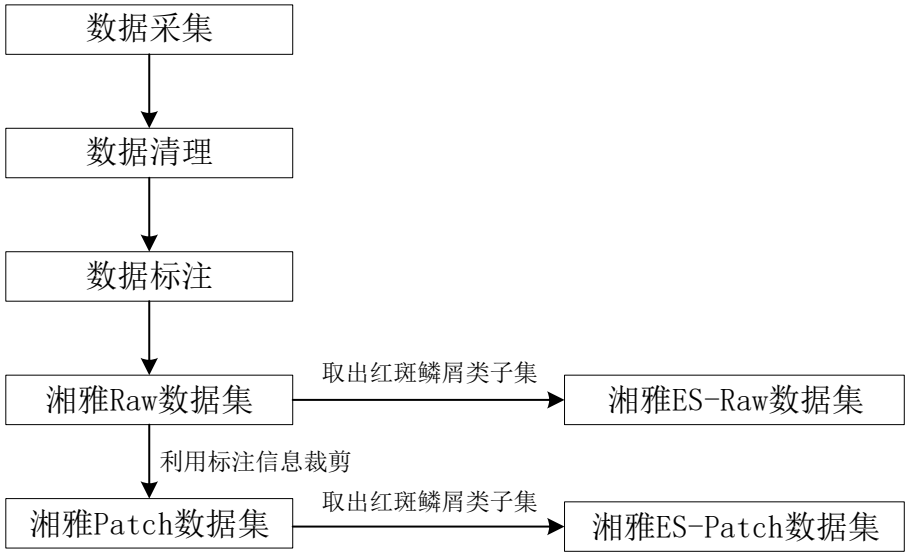


图 2-1 数据集构建的流程

2.2.2 数据采集和清理

数据采集工作由湘雅医院的 9 名专业皮肤科医生完成，以患者的病理检查的最终确诊结果作为图像的真实诊断标签。我们采用的图像采集设备为专业的高清数码相机，拍摄区域以皮损发生的主要部分为图像中心。对于单个病人的一个皮损部位，我们会根据皮损的具体大小调整拍摄的距离以确保采集得到 3 幅不同缩放程度的图像，具体如图 2-2 所示，图中从左到右表示拍摄距离越来越近。采取不同缩放程度的图像作为模型训练数据，能使模型对于不同尺度下的图像识别效果更加稳定。另外，为了排除复杂背景的干扰对于卷积神经网络识别算法的影响，我们的拍摄的背景为纯黑色或者其他纯色背景。

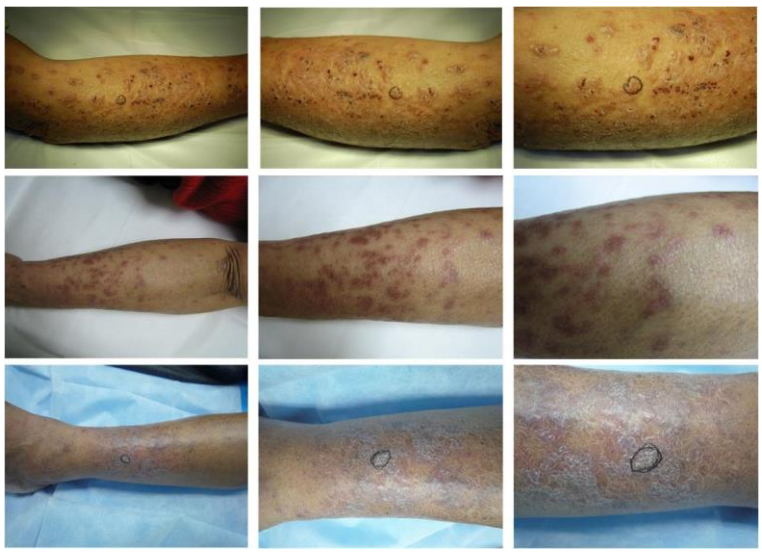


图 2-2 不同缩放程度的图像

数据清理工作由数位皮肤病专家和深度学习算法专家合作完成。首先，为了确保数据标签的有效性和准确性，我们对于收集得到的图像数据进行了严格地清理。我们剔除了其中成像质量低（图像模糊不清）、标签错误（标签与病人的病史不符）和包含大部分明显的人体部位信息（如鼻子、眼睛）的图像。

2.2.3 数据标注和裁剪

数据的标注工作是指对于上一步清理过后的图像利用开源的图像标注软件labelImg（https://github.com/tzutalin/labe-lImg）进行病变区域的再次框选，以获得更加精确的皮肤皮损区域，数据标注后的图像的保存操作如图 2-3 所示。我们通过对图像数据进行筛查，图像中还是可能出现的毛发，衣物等干扰因素，这些与这两类疾病鉴别无关的干扰因素都将会在这一阶段被尽量去除。

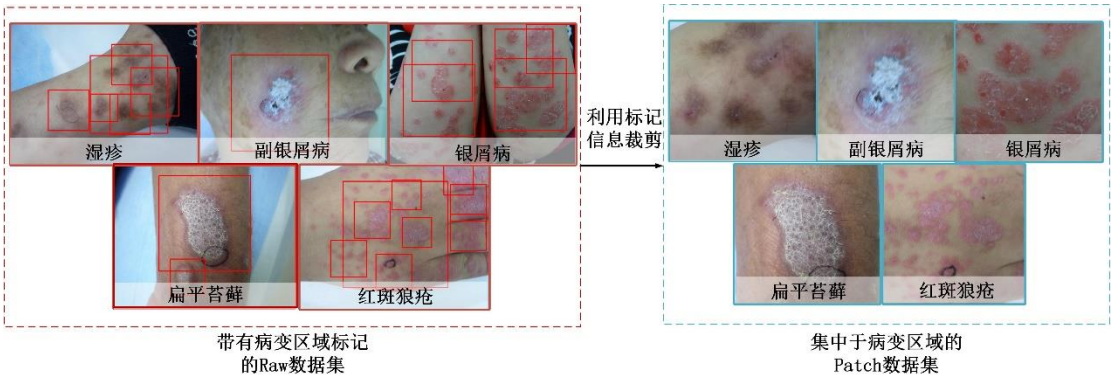


图 2-3 数据裁剪过程

我们将裁剪前后的数据各类疾病的数量总结于表 2-2 中。我们可以看到，裁剪前的 ES-Raw 数据集中的银屑病的数量较多，这是因为我们合作单位湘雅医院皮肤科是银屑病的专科门诊，因此接诊银屑病的病人较多。经过裁剪之后，我们得到了在图像上更加集中于病变区域的湘雅 ES-Patch 数据集，为了保持类别均衡，我们把 5 类皮肤疾病的图片数量都保持在 900 张。可以看到，湘雅 ES-Patch 数据集类别上更为均衡，且图像上更为集中在病变区域，可以用来直接做细粒度的图像分类，后文第三章和第四章的工作都是基于湘雅 ES-Patch 数据集。

表 2-2 湘雅 ES-Raw 数据集和湘雅 ES-Patch 数据集

数据集	扁平苔藓	副银屑病	红斑狼疮	湿疹	银屑病
湘雅 ES-Raw	364	348	513	534	2618
湘雅 ES-Patch	900	900	900	900	900

2.3 本数据集的特点及优势

我们的数据集相对已知的数据集有着明显的优势以及特点，现总结几点如下：

1）图像数据总量大，病种齐全：由表 2-1 中的统计结果我们可以知道，本数据集的原图数量有 54098 张，Patch 数据集有 150223 张，覆盖了 543 类皮肤疾病。据我们所知，本数据集是目前公开的皮肤疾病临床图像数据集中规模最大的。在 5 类红斑鳞屑类皮肤疾病数量上，本数据集的也是最大的；

2）图像数据来源以中国人种为主：本数据集的采集单位是在湘雅医院的皮肤科门诊部，就诊的病人多为中国患者，因此收集的数据多为中国患者的图像。前面介绍的目前公开的数据集大多为欧美地区的人种，本数据集的提出有助于弥补中国人种的皮肤疾病数据库的缺失；

3）图像数据质量好，类型丰富：本数据集的采集过程由专业皮肤科医生严格把控，采集设备为高清数码相机，采集得到的图像数据对于皮肤病变区域的呈现准确且清楚。另外，本数据集还包含有不同拍摄距离的皮损部位图像以及同一个病人各个患病时期的皮损图像，对于疾病辅助诊断的研究价值极高；

4）图像数据标签信息齐全：经过我们的前期调研，我们发现目前现有的公开数据集都只有图像的诊断标签或者很少的病史信息（病变部位等），图像的背景都没有经过处理。本数据集完全按照病人为单位进行组织，数据集内每张图像都含有对应病人的全套病史病理资料，以及由专业医生标注的病变区域的标签，这些病变区域的标签能辅助我们进行皮肤疾病的目标检测任务的研究。

2.4 本章小结

本章首先介绍了目前在皮肤疾病辅助诊断研究领域常被采用的几个现有的临床图像数据集，介绍了它们的规模，主要人种来源以及标签信息等。最后得出的结论是这些现有的数据集来源主要以欧美人种为主，且图像上带有水印，数据集包含的红斑鳞屑类皮肤疾病的图片也较少。因此，接下来就引出本文自行构建的湘雅皮肤疾病临床图像数据集，然后具体介绍了数据集的采集、清理、标注和裁剪的过程。最终，我们得到了可用于后文辅助诊断研究的两个基准数据集——湘雅 ES-Raw 数据集和湘雅 ES-Patch 数据集。

3 基于视觉特征提取的红斑鳞屑类皮肤疾病分类

视觉特征描述了皮肤疾病的一些临床表现，包括皮损部位的颜色、纹理和形状等，这些特点也都是临床诊断中医生考虑的重要依据。基于此，本章选择了4类能够描述上述特性的视觉特征，包括：尺度不变特征变换（SIFT）特征、方向梯度直方图（HOG）特征、颜色直方图（CH）特征和颜色命名（CN）特征。接下来，本文将对不同的特征提取方法进行介绍。

3.1 视觉特征

3.1.1 SIFT 特征

尺度不变特征变换（Scale Invariant Feature Transform, SIFT）[21]是一种图像纹理特征描述子提取方法，被广泛应用于图像拼接，识别，检索等邻域。SIFT 特征是一种对图像旋转、缩放以及视角变化等都能保持较好的不变性的特征描述子，其提取步骤简要描述如下：

（1）尺度空间的极值检测：利用二维高斯函数$G(x_i, y_i, \sigma)$与原始灰度图像$I(x, y)$卷积运算得到不同尺度的图像，其中二维高斯函数表达式如式（3-1）所示，图像尺度空间的表达式如式（3-2）所示：

$$G(x_i, y_i, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma^2} \right) \tag{3-1}$$

$$L(x, y, \sigma) = G(x_i, y_i, \sigma) * I(x, y) \tag{3-2}$$

（2）特征点定位：对于每个感兴趣的点，比例和位置由精确拟合模型确定，关键点的选择规则为这些点的稳定程度。拟合模型通常是三维二次的拟合函数，不仅消除了不稳定的边缘响应点和低对比度的点，而且还提高了匹配稳定性和抗噪性。

（3）特征方向赋值：基于图像局部的梯度方向，将一个或多个方向分配给每个特征点位置，并且所有后续操作变换特征点的方向，以便提供这些特征的不变性。通常，我们可以通过梯度直方图统计来确定特征点的方向。 梯度的模数和方向的计算由等式（3-3）和（3-4）给出：

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{3-3}$$

$$\theta(x,y)=\tan^{-1}\left(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right) \quad (3\text{-}4)$$

其中，L 为特征点所在的尺度空间值。完成特征点的梯度计算之后，使用直方图计算特征点附近邻域内像素的方向和梯度。梯度直方图的峰值代表该特征点邻域梯度的方向，最大值代表该特征点的主方向。算法为了提高匹配的鲁棒性，将大于主方向峰值 80%的方向设定为特征点的辅方向。特征点有多个方向的情况下，实际上是在此位置上有多个关键点，不同的方向可以增强匹配的鲁棒性。

（4）特征点描述：特征点的描述即用一组向量对提取的特征点进行描述。这组向量不单包括特征点，还包括特征点周围对其有贡献的像素点。基本实现方法是对特征点周围的像素区域分块，计算块内的梯度直方图，生成具有独特性的向量。如图 3-1 所示，我们对每个特征点使用 4×4 共 16 个种子点来描述，这样一个特征点就可以产生 128 维的 SIFT 特征向量。

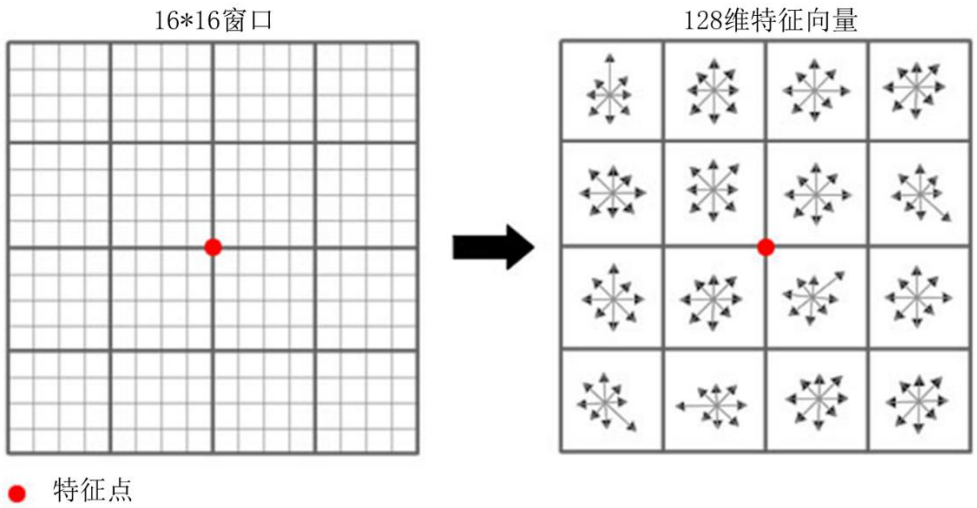


图 3-1 SIFT 特征特征描述生成

由于每张图像中的特征点个数不一样，假设我们对于一张图像提取得到 n 个特征点，则我们提取得到该图像的特征维度为 n*128，由于每张图片的特征数目不同，这样无法输入到后文的 SVM 分类器进行训练。这里我们采用了 SIFT 特征提取+K-Means 聚类+词袋模型的方法[22]，可以将每张图片的特征维度都表示成 k*1 维。具体流程图如图 3-2 所示：

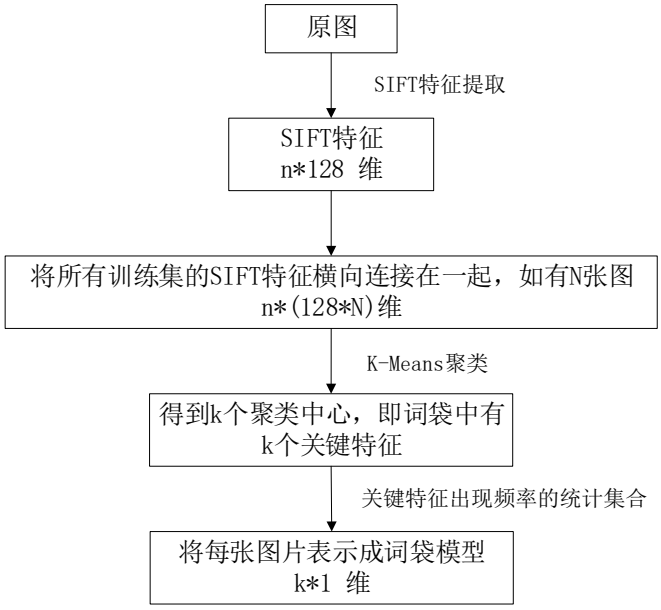


图 3-2 SIFT 特征特征描述生成

如上图所示，这里需要指出的是这里的特征连接指的是将训练集内的每一个类别的 SIFT 特征进行横向连接，之后通过 K-Means 聚类得到 k 个关键特征，即我们的词袋模型有 k 个关键词。然后，我们再回到训练集中对每张图像进行关键特征出现频率的统计，将这一 k 维的统计量作为每张图像最终的特征。SVM 训练完毕后，测试集图像的特征提取过程同样也是采取上述的方式。

图 3-3 SIFT 特征可视化展示

本文针对皮肤疾病图像进行 SIFT 特征提取，我们将 SIFT 特征可视化如图 3-3 所示。整体来说，SIFT 特征能较好的捕捉病变部位的纹理特征，如图中所示的红斑和白色鳞屑等，这些都是红斑鳞屑类皮肤疾病特有的图像特征。

3.1.2 HOG 特征

方向梯度直方图（Histogram of Oriented Gradient, HOG）[23]是一种图像处理邻域常用的纹理特征描述子，在物体检测领域得到了很好的应用。HOG 特征通过计算图像局部区域的梯度方向直方图来构成特征，其实现过程是将图像分成小的局部区域，然后分别进行计算局部区域中每个像素点的梯度方向直方图，最后将这些直方图组合起来构成图像的 HOG 特征。主要包括以下几个步骤：

（1）梯度的计算：对图像在不采用高斯平滑处理的情况下，使用 x 方向上的分梯度模板$[-1,0,1]$和 y 方向上的分梯度模板$[-1,0,1]^T$对整幅图像逐像素地遍历，遍历过后，每个像素点都具有了水平和垂直方向上的分梯度值。具体梯度的模值和方向的计算方法如式（3-5）和式（3-6）所示。

$$gradient(x,y)=\sqrt{sub-gradient_x(x,y)^2+sub-gradient_y(x,y)^2} \tag{3-5}$$

$$angle(x,y)=\tan^{-1}\left(\frac{sub-gradient_y(x,y)}{sub-gradient_x(x,y)}\right) \tag{3-6}$$

（2）分局部区域投票：我们可以将图像分为多个图像块即 block，然后在图像块内再划分更细的细胞单元即 cell。这样划分之后，我们可以通过计算 cell 中每个像素的梯度直方图，得到 cell 维度的特征。将 block 中的特征串联起来，构成 block 的特征。为了使特征向量空间对光照，阴影和边缘变化具有鲁棒性，我们需要对 block 块内的 HOG 特征向量进行归一化。一般采用的归一化函数有以下四种：

$$L2-norm, v\leftarrow v/\sqrt{\|v\|_2^2+\varepsilon^2} \tag{3-7}$$

$$L2-Hys, v\xleftarrow{v_{\max}=0.2} v/\sqrt{\|v\|_2^2+\varepsilon^2} \tag{3-8}$$

$$L1-norm, v\leftarrow v/(\|v\|_1+\varepsilon) \tag{3-9}$$

$$L1-sqrt, v\leftarrow \sqrt{v/(\|v\|_1+\varepsilon)} \tag{3-10}$$

（3）最终我们通过以上的步骤可以得到输入图像对应的 HOG 特征向量，长度为$\beta*\varsigma*\eta$，其中β表示每个 cell 中的特征向量长度，ς,η分别表示图像中划分的 block 的个数和一个 block 中 cell 的个数。

这里值得注意的是，SIFT 特征对于不同图像的特征维度是不一致，与检测到的特征点的个数有关，而 HOG 特征维度对于同样大小的输入图像是一致的。如，对于 64*128 的图像，我们取每个 cell 包含 8*8 个像素，每 2*2 个 cell 组成一个 block。因为每个 cell 有 9 个特征，所以每个 block 内有 4*9=36 个特征。以 8 个像素作为窗口的扫描步长，水平和垂直方向将分别有 7 个和 15 个的扫描窗口。即对于 64*128 的图片，最终提取的 HOG 特征维度为 36*7*15=3780。

本文针对皮肤疾病图像进行 HOG 特征提取，我们将提取到的 HOG 特征可视化如图 3-4 所示。整体来说，HOG 特征的效果不如 SIFT 特征，而且 HOG 特征生成过程冗长，提取速度慢。因此本文都是将输入图像缩放到 64*128，然后再进行特征提取，提取到的 HOG 特征维度都为 3780。另外由于梯度的性质，该描述子对噪声点相对比较敏感。从图 3-4 我们可以看出 HOG 特征给出了很多与疾病特征无关的描述子，这些噪声点会在分类阶段造成很大干扰，影响疾病分类效果。

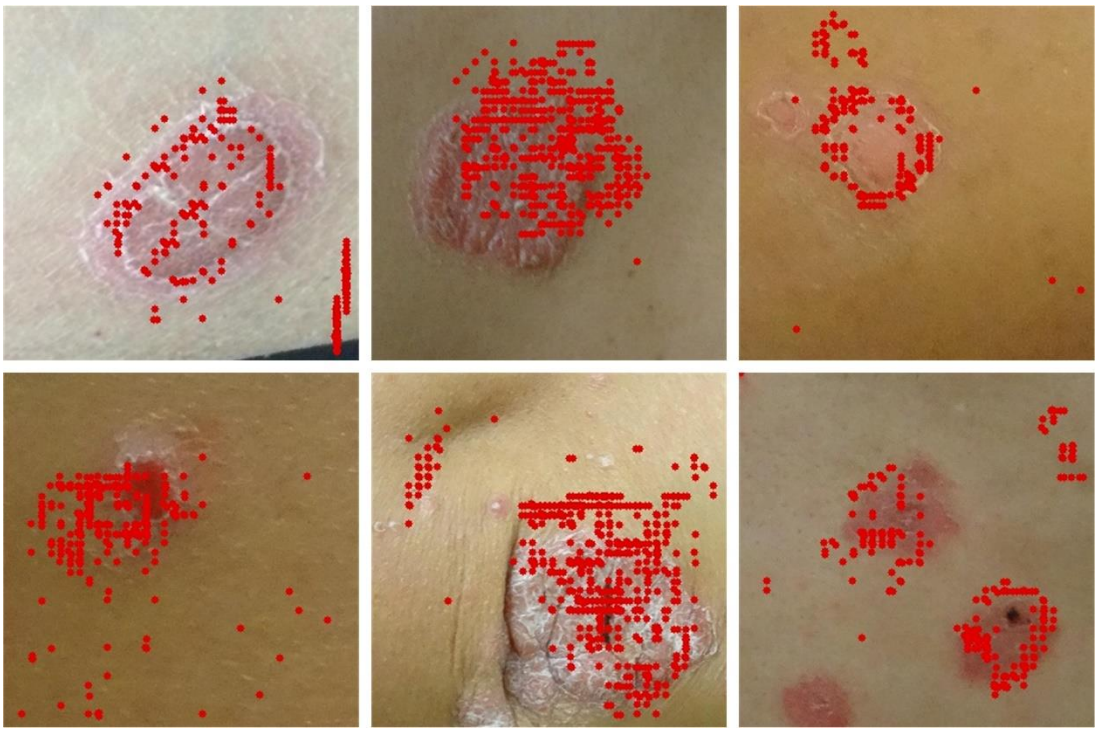


图 3-4 HOG 特征可视化展示

3.1.3 CH 特征

颜色直方图（Color Histograms, CH）是最常用的表达颜色特征的方法，最早被 Swain[24]等人提出可以将颜色直方图作为图像颜色特征的描述。其优点是计算简单，且不受图像旋转、缩放和平移变化的影响，对图像质量的变化(如模糊)也不甚敏感。颜色直方图可以是基于不同的颜色空间和坐标系，本文采取的是 HSV 空间的颜色直方图特征提取，其提取过程主要包括以下几个步骤：

（1）颜色空间转化：首先将 RGB 空间的图像转化到 HSV 空间，对于给定 RGB 彩色图像，通过公式（3-11）~（3-14）将其转换为 HSV 通道图像：

$$\begin{aligned}\max &= \max(R,G,B)\\ \min &= \min(R,G,B)\end{aligned} \tag{3-11}$$

$$H=\begin{cases}0^{\circ}, \max-\min=0\\ (G-B)/(\max-\min), R=\max\\ (B-R)/(\max-\min), G=\max\\ (R-G)/(\max-\min), B=\max\end{cases} \tag{3-12}$$

$$S=\begin{cases}0, \max=0\\ (\max-\min)/\max, \max\neq 0\end{cases} \tag{3-13}$$

$$V=\max \tag{3-14}$$

（2）颜色量化：颜色量化即将颜色空间划分为若干个等级。用于颜色量化的算法很多，如向量量化、聚类或神经网络。这里我们介绍一种简单的 HSV 空间的向量量化方法。如，我们把 (x,y) 处的 H，S 和 V 通道的像素值分成 L_h，L_s 和 L_v 三个等级，这样我们就得到了该点对应的一维特征向量，如公式（3-15）所示。

$$Hist_{(x,y)}=\omega_h * H_{(x,y)}+\omega_s * S_{(x,y)}+\omega_v * V_{(x,y)} \tag{3-15}$$

其中，ω_h，ω_s 和 ω_v 代表三个空间在特征向量中所占的权重。通过这种表示，特征向量 $Hist$ 的取值范围为：$[0,\ \omega_h(L_h-1)+\omega_h(L_h-1)+\omega_v(L_v-1)]$。

（3）颜色直方图统计：对每个 $Hist$ 等级的像素个数进行统计就可以绘制成直方图，$Hist$ 等级的总数为 $\omega_h(L_h-1)+\omega_h(L_h-1)+\omega_v(L_v-1)$，表征颜色空间划分的精细程度。一般情况下，颜色空间分的越细，小区间的数目越多，直方图对颜色的分辨能力越强。然而，小区间的数目过大会增加计算负担。一种有效减少小

区间个数的方法是只选取像素数目较多的直方图来构造图像特征，因为这些表征图像主要颜色的小区间能够近似表征图像的颜色特征。

本文针对皮肤疾病图像进行 CH 特征提取，我们将提取到的 CH 特征可视化如图 3-5 所示。其中，图中第一列为原图，第二列为其对应的颜色直方图，其中的 3 个子图从左至右依次为 H、S、V 的统计直方图。我们将 H、S、V 的取值范围都是[0, 255]，即 256 级。那么，最终对于任意图像提取到的颜色直方图特征的维度都是固定的 3*256=768 维。

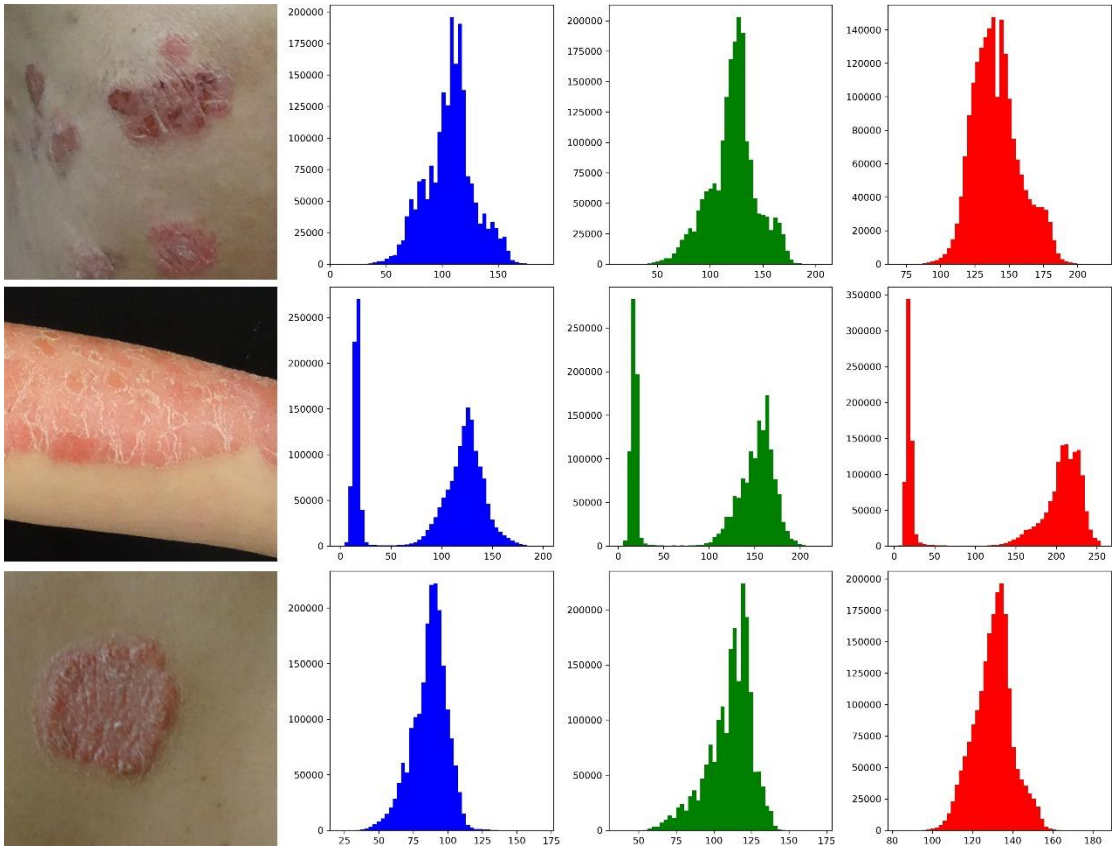


图 3-5 CH 特征可视化展示

3.1.4 CN 特征

颜色命名(Color Names, CN)[25]是一种基于颜色语义标签的特征描述子，具有较强的物体颜色和物体轮廓的描述能力，在图像分类领域中具有许多应用。在计算 CN 特征时，需要完成从 RGB 颜色空间到图像像素的 CN 值的概率映射。文献[26]通过研究认为语义学上的颜色包括 11 种基本颜色元素：黑、蓝、棕、灰、绿、橘黄、粉红、紫、红、白和黄。文献介绍了一个像素点的 RGB 值到颜色予以标签之间的映射模型。本文基于文献[26]和[27]对于 CN 特征的定义，介绍 CN 特征的提取过程如下：

（1）图像区域划分：为了建立一种局部特征，我们将待提取特征的图像 I 都被定义成由 m 个大小相等的带状区域，如式（3-16）：

$$I=[I_1,I_2,I_3,...,I_m] \tag{3-16}$$

（2）区域特征向量提取：为了消除背景噪声的影响，全局颜色命名描述符只计算皮肤疾病的病变区域。对每个条带都提取颜色特征。带状区域 j 的特征向量如式（3-17）：

$$I_j=[CN_{1j},...,CN_{sj}] \tag{3-17}$$

其中CN_s是指颜色为 s 的可能性。我们使用颜色分布模型计算每个带状区域的颜色分布。在这个模型中，我们只需要将 RGB 颜色空间（0-255）按照步长为 8 进行分割，那么一共分为256/8=32 个区间,因为有三个通道，那么一共给有 32*32*32 种组合。由此 RGB 被离散化为$32*32*32=32768$个索引，建立了一个32768*11个维度的查找表，模型的核心在与建立 RGB 与 11 种预先定义的颜色分布之间的对应关系。对于带状区域 j，颜色分布CN_s定义如下：

$$CN_{sj}=\frac{1}{N}\sum_{x_{RGB}\in R_j}p(CN_s|x_{RGB}) \tag{3-18}$$

其中，$p(CN_s|x_{RGB})$是一个确定的像素的 RGB 值被分配给特定颜色命名CN_s的可能性。R_j是带状区域 j 的前景区域，N 表示R_j中总共的像素值。显然，s 从 1 到 11 的CN_{sj}的总和为 1。然后将 m 个带状区域的的 CN 特征串联起来可以得到一个 11*m 维的特征向量。最后，我们采用 PCA 算法将特征维度降到 m。

本文针对皮肤疾病图像进行 CN 特征提取，我们将提取到的 CN 特征可视化如图 3-6 所示。其中，图中第一行为原图，第二行为其对应的颜色命名特征图。从图中我们可以看出，该颜色特征能够较好地捕捉颜色特征，如非皮肤区域、正常皮肤区域以及皮肤病变区域。另外，它对于病变区域的不同病变颜色也能够做出很好区分，如白色鳞屑、红斑和结痂区域等。

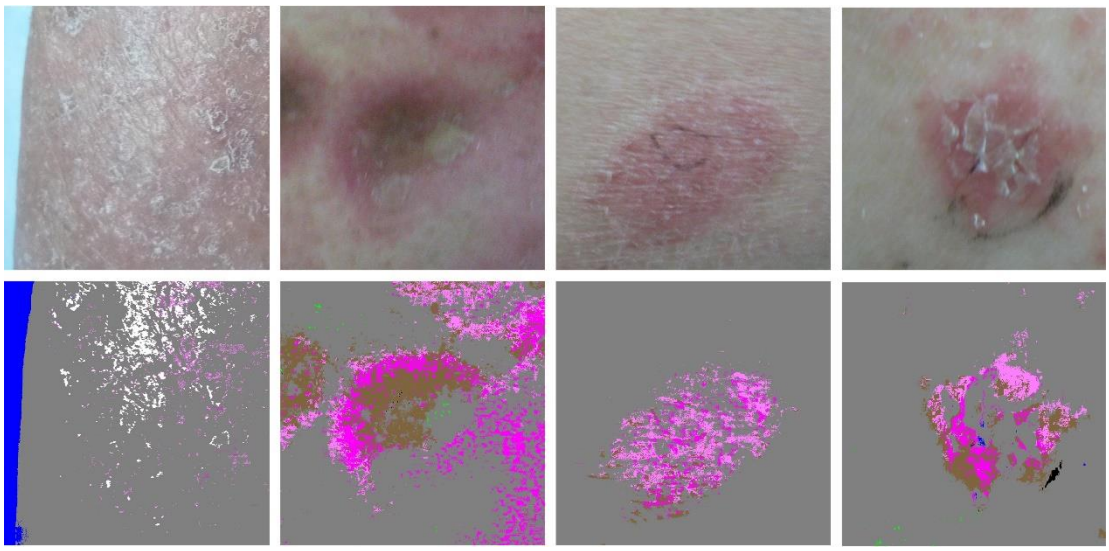


图 3-6 CN 特征可视化展示

3.2 实验方案设计

为了评估上述视觉特征对于红斑鳞屑类皮肤疾病的表征能力，本章采用支持向量机（Support Vector Machine, SVM）作为分类器，选取湘雅 ES-Patch 数据集作为本章的实验数据集。实验框图如图 3-7 所示。

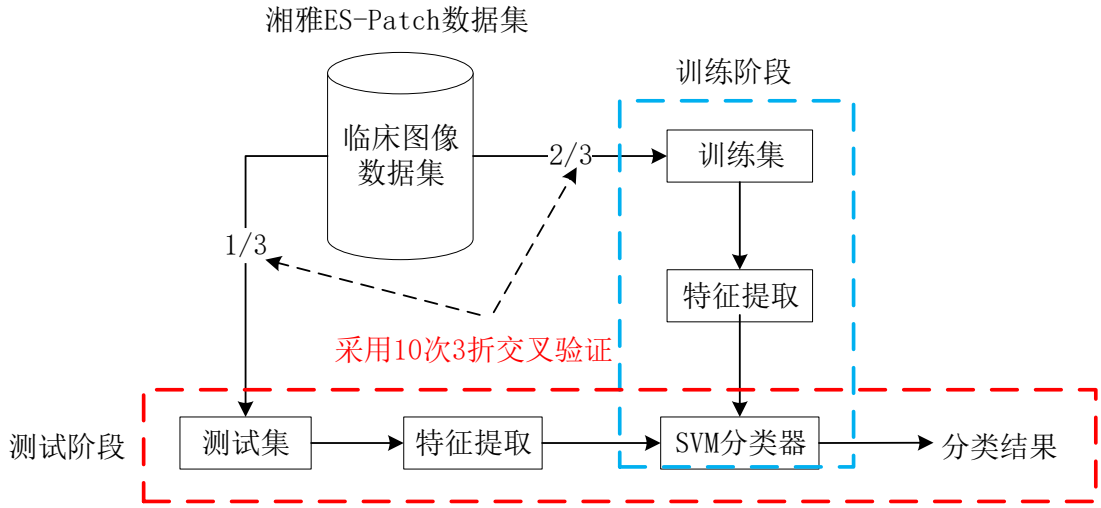


图 3-7 基于传统特征的红斑鳞屑类皮肤疾病分类实验框图

如图 3-7 所示，本章实验采用的是湘雅 ES-Patch 数据集，包含扁平苔藓、副银屑病、红斑狼疮、湿疹、银屑病这 5 类皮肤疾病，每类疾病各 900 张。采用 10 次 3 折交叉验证的实验方式，即进行 10 次实验，每次实验中训练集和测试集随机按照 2：1 的比例分配，结果取 10 次实验的平均值。之所以采用 Patch 数据集，是因为 Patch 数据集更加集中于病变区域，这样我们提取的疾病特征也就更

多与病变区域相关。

SVM 分类器采用的是 LIBSVM[28]工具包进行实现。LIBSVM 是台湾大学林智仁(Lin Chih-Jen)教授等人开发设计出来的一个简单、容易操作和快速有效的 SVM 模式识别与分类回归的软件包。它不但提供了编译好的可在 Windows 操作系统的运行文件，还提供了多种其他操作系统和语言的源代码，方便其他操作系统的用户的使用。基础的单个 SVM 是一个二分类器，LIBSVM 采用所有对所有（All-Versus-All）[29]的思想构建多个二分类器并结合投票策略实现了多分类的功能。

整个实验流程分为训练和测试两个阶段，在训练阶段，我们采用不同的特征提取方法对训练集的图像进行特征提取，之后训练得到不同特征提取下的 SVM 分类器。在测试阶段，我们对测试集提取相应的特征，输入训练好的 SVM 分类器，统计不同特征提取方法下的测试结果。

3.3 实验结果分析

基于湘雅 ES-Patch 数据集，我们对训练库和测试库中的图像提取上述的视觉特征并利用多分类 SVM 进行分类。准确率计算公式如式（3-19）所示：

$$准确率=\frac{测试分类正确的图像数量}{测试集图像总量}\times 100\% \qquad (3\text{-}19)$$

除了上面阐述的 4 种特征提取方法外，我们还采用了文献[30]的方法，文献[30]采用了 SIFT+CN 对鸟类图像进行分类，也是属于细粒度对象识别任务。通过交叉验证实验，我们最终的实验结果如表 3-1 所示：

表 3-1 视觉特征分类结果

序号	视觉特征	特征维度	分类器	准确率（%）
1	SIFT	2100	SVM	35.85
2	HOG	3780	SVM	22.78
4	CH	768	SVM	14.19
5	CN	2100	SVM	30.20
6	SIFT+CN[30]	2100	SVM	52.19

从表 3-1 的分类结果，我们可以总结出如下几点发现：（1）以上基于视觉特征的分类方法的效果在整体上都不够理想，最优的分类特征是 SIFT+CN 的组合特征，达到的识别准确率为 52.19%；（2）在该数据集上，纹理特征比颜色特征起着更重要的分类作用，在一些皮肤疾病的图像上前景和背景都比较相似，病

变往往呈现不同的纹理和形状，如环形、线形、凹形和凸形；（3）颜色特征也有适用的情况，存在着某些病变区域形状相似的病例，这个时候可以利用颜色特征进行区分，例如某些湿疹和扁平苔藓的图像。

考虑到前面提到的（2）和（3）的情况，文献[30]的 SIFT+CN 的方法由于结合了图像的纹理特征和颜色特征，在两种情况下都能提取到有效的疾病特征，因而识别效果更好。尽管如此，52.19%的识别效果仍然不能令人满意，无法承担起红斑鳞屑类疾病的辅助诊断任务。

3.4 本章小结

本章主要介绍了 5 类图像视觉特征，包括 SIFT 特征、HOG 特征、LBP 特征、CH 特征和 CN 特征。这些特征一定程度上描述了皮肤疾病的一些临床表现，如皮损部位的颜色、纹理和形状等。基于湘雅 ES-Patch 数据集，本章采用 10 次 3 折交叉验证的实验方式，利用多分类的 SVM 方法对上述视觉特征的分类效果进行了评估。实验结果表明，基于传统视觉特征的红斑鳞屑类疾病的鉴别准确率普遍都不理想，最高准确率为结合了纹理特征和颜色特征的 SIFT+CN 的方法，但是其准确率也只有 52.19%，这说明了依靠传统视觉特征的方法无法很好地承担红斑鳞屑类疾病的鉴别任务。

4 基于卷积神经网络的红斑鳞屑类皮肤疾病分类

深度学习是目前机器学习领域最前沿和热门的研究课题之一，它能建立模拟人脑的神经网络，模仿人脑的思维机制来解释数据。深度学习中的卷积神经网络是目前模式识别、计算机视觉领域用到的最为主流的方法之一。本章采用基于卷积神经网络的方法对皮肤疾病图像的特征进行自动提取，利用大量的带有标签信息的疾病图像数据对模型进行训练，构建起了基于卷积神经网络的皮肤疾病临床图像分类模型。

4.1 基于卷积神经网络的图像分类

4.1.1 卷积神经网络理论

卷积神经网络(Convolutional Neural Network, CNN)最早于 1998 年被 LeCun[31]等人提出。目前广泛应用于图像识别、语音处理和文本翻译等领域[32][33][34]。卷积神经网络（CNN）作为深度学习的代表算法之一，以其强大的特征自学习能力席卷了机器视觉领域最具权威的 ILSVRC 的各项任务[35][36][37]。最近，CNN 在许多医学图像分析任务中也取得了很好的应用，包括医学图像分类任务[38][39]、检测任务[40][41]和分割任务[42][43]。

在图像分类领域，CNN 的可以直接输入图像数据，模型利用梯度下降算法更新参数，能够自动学习到有利于图像分类的特征，而不关心各个图像特征间的差异性，省去了传统的特征提取方法中人工设计特征的繁琐操作。从结构上看，卷积神经网络是一种带有卷积结构的深度神经网络，我们以图 4-1 中的 LeNet5 为例介绍其基本结构。

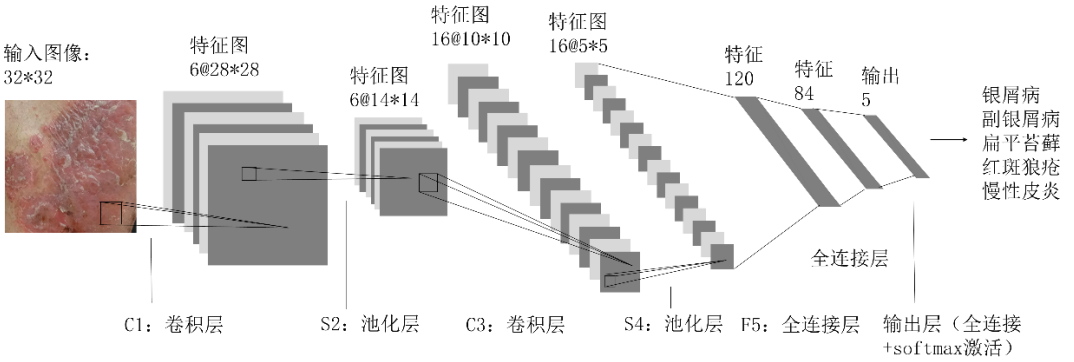


图 4-1 LeNet5 的基本结构

由图 4-1 所示，一个 CNN 主要包括卷积层、池化层和全连接层这 3 个结构。我们假设输入为 32*32 的皮肤疾病 RGB 图像，输出为 5 类皮肤疾病的概率值。

其中，C 层为卷积层，S 层为池化层，F 层为全连接层。下面分别对卷积层、池化层和全连接层做基本介绍。

卷积层是卷积神经网络的核心基石。在图像识别里我们提到的卷积是二维卷积，即离散二维滤波器（也称作卷积核）与二维图像做卷积操作，简单的讲是二维滤波器滑动到二维图像上所有位置，并在每个位置上与该像素点及其邻域像素点做内积。卷积操作被广泛应用与图像处理领域，不同大小的卷积核可以提取不同的特征，例如边沿、线性、角等特征。在深层的卷积神经网络中，通过不同层次的卷积操作可以提取出图像由低级到复杂的特征。卷积层输出的特征图的尺寸的计算公式如下：

$$O_w = (I_w - F_w + 2P)/S + 1 \tag{4-1}$$

$$O_h = (I_h - F_h + 2P)/S + 1 \tag{4-2}$$

其中，I_w, I_h 和 O_w, O_h 分别为输入图像和输出图像的宽和高，F_w 和 F_h 是卷积核的尺寸。P 和 S 是指卷积操作中的填充和移动步长。另外，卷积层输出的特征图的个数等于卷积核的个数。如图 x 中，以 C1 层为例，C1 层有 6 个 5*5 的卷积核，卷积操作的填充为 0，移动步长为 1，故 C1 层得到的特征图的宽和高都为：（32-5+2*0）/1+1=28。

池化层往往在卷积层后面，池化操作可以大大降低特征图的尺寸，同时也降低过模型拟合的可能性。池化层包含特定的池化函数，不同的池化函数可以将特征图转化为不同形式的池化层输出，表达不同的特征图统计量的含义。池化方法一般包括 L_p 池化和随机池化 2 种，计算公式如式（4-3）和（4-4）：

$$A_k^l(i,j) = \left[\sum_{x=1}^{f} \sum_{y=1}^{f} A_k^l(s_0 i + x, s_0 j + y)^p \right]^{\frac{1}{p}} \tag{4-3}$$

$$A_k^l = \lambda L_1(A_k^l) + \lambda L_\infty(A_k^l), \quad \lambda \in [0,1] \tag{4-4}$$

其中 k 为特征图的通道数，f 、s_0 和 p 对应卷积核大小、卷积步长和填充数，l 表示该层处于卷积神经网络的第 l 层。

卷积神经网络中的全连接层一般都是位于卷积神经网络中的最后部分，顾名思义，全连接层中的每个神经元都与前面一层的所有神经元都有连接。设置全连接层的目的是将特征图展开成特征向量，最终得到预测的输出层，其中，一般的分类卷积神经网络的输出层也是特殊的全连接层，神经元的个数即为分类的类别总数。

4.1.2 迁移学习理论

迁移学习顾名思义就是利用先前的任务中已经学习得到的知识来作为经验指导新的学习任务，加快新的学习任务的学习过程或者提升新任务的学习效果。在本文中，为了加快网络的学习效率，我们也采用了如图 4-2 所示的迁移学习的思想。②的方式属于迁移学习的方式，对于特定结构的 CNN 网络，我们均在 ImageNet 这一非常大规模的图像数据集进行预训练，然后将预训练的权重作为训练的初始参数加载于模型上，之后再通过我们的数据集进行再训练。此外，我们还对比了①的非迁移学习的方式，即完全在本任务上从头开始训练一个模型。迁移学习目前已经广泛地应用于深度学习网络，它非常适用于小数据集的情况下，即能够利用在不相关的大型数据集训练得到信息来辅助小数据集的训练任务，而本文的红斑鳞屑类皮肤疾病图像正具备这一特点，我们后文实验部分都将在各个 CNN 结构上比较这 2 种方法。

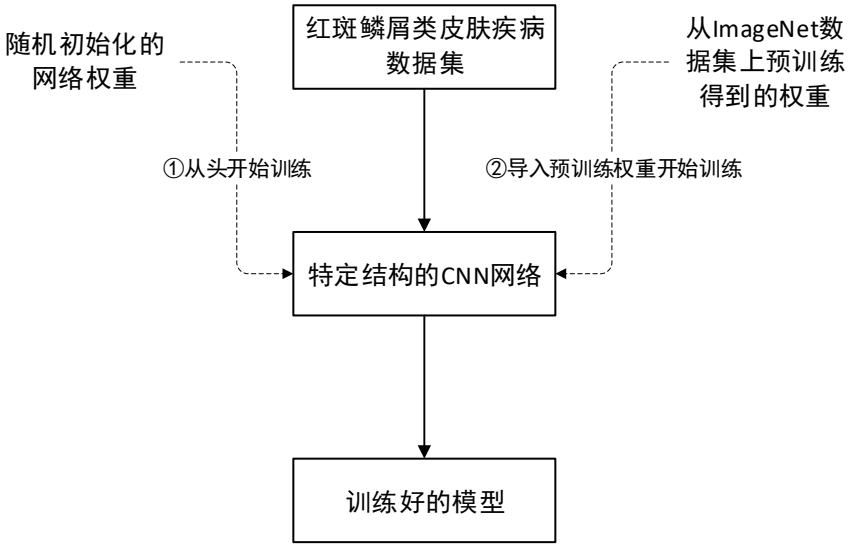


图 4-2 导入 ImageNet 上的预训练权重进行迁移学习

4.1.3 本文采用的几种主流的卷积神经网络介绍

下面主要介绍几种主流的卷积神经网络模型：

（1）InceptionV3

InceptionV3[44]是由谷歌研究组于 2015 年提出的 CNN 结构，属于 2014 年提出的 GoogleNet（即 InceptionV1）发展至今的第三个版本。该网络通过采用全局平均池化层取代全连接层，极大的降低了参数量，另外还引入了 Inception 模块组的结构，方便增添和修改，是非常实用的模型。随后的 InceptionV2 中，引入了 Batch Normalization 的方法，加快了训练的收敛速度。在 InceptionV3 模型中，

通过将二维卷积层拆分成两个一维卷积层，不仅降低了参数数量，同时减轻了模型训练过程中的过拟合现象。本文计划采用经典的 InceptionV3 模型构建基于红斑鳞屑类皮肤疾病的分类器，图 4-3 是其网络结构。

InceptionV3 网络结构		
类型	Kernel尺寸/步长（或注释）	输入尺寸
卷积	3*3/2	299*299*3
卷积	3*3/1	149*149*32
卷积	3*3/1	147*147*32
池化	3*3/2	147*147*64
卷积	3*3/1	73*73*64
卷积	3*3/2	71*71*80
卷积	3*3/1	35*35*192
Inception模块组	3个Inception Module	35*35*288
Inception模块组	5个Inception Module	17*17*768
Inception模块组	2个Inception Module	8*8*1280
池化	8*8	8*8*2048
线性	Logits	1*1*2048
Softmax	红斑鳞屑类皮肤疾病的分类输出	1*1*5

图 4-3 InceptionV3 网络结构

（2）InceptionResNetV2

InceptionResNetV2[45]是由谷歌研究组于 2016 年提出的 CNN 结构，作者借鉴了残差网络 ResNet[46]中残差学习的思想，将其应用到了 InceptionV3 模型当中，得到了 CNN 分类性能的进一步提升。InceptionResNetV2 简化了原有的 Inception 模块，网络结构中采取了残差连接的方式可以使网络层数堆叠得更多。2121212。图 4-4 是其网络结构。

InceptionResNetV2 网络结构		
类型	注释	输入尺寸
Stem模块	采用Inception设计的卷积组	299*299*3
Inception-ResNet-A	5个Inception-ResNet-A	35*35*256
Reduction-A		35*35*256
Inception-ResNet-B	10个Inception-ResNet-B	17*17*896
Reduction-B		17*17*896
Inception-ResNet-C	5个Inception-ResNet-C	8*8*1792
池化	平均值池化	8*8*1792
Dropout层	Keep=0.8，防止过拟合	1*1*1792
Softmax	红斑鳞屑类皮肤疾病分类输出	1*1*5

图 4-4 InceptionResNetV2 网络结构

（3）DenseNet121

DenseNet[47]是由康奈尔大学、清华大学、Facebook FAIR 实验室等于 2017 年提出的 CNN 结构，是基于 ResNet 结构的一个提升版本。相比原始版本的 ResNet，DenseNet 提出了一个更激进的密集连接机制，在 CNN 网络前向传播的过程中，每一层都与其它所有层相连。即对于一个有 L 层的 DenseNet 就有 $L(L+1)/2$ 个连接。这样的结构大大减少了参数的个数，增强了特征在各个层之间的流动，提高了模型的训练效率。图 4-5 为 DenseNet 原文中的 4 种网络结构的示意图。

DenseNet网络结构					
类型	输出大小	7*7 Conv,stride 2			
卷积	112*112	3*3 Max pool,stride 2			
池化	56*56	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Dense 模块(1)	56*56	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv
Transition层(1)	56*56	DenseNet-121			
	28*28	2*2 Average pool, stride 2			
Dense 模块(2)	28*28	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv
Transition层(2)	28*28	DenseNet-121			
	14*14	2*2 Average pool, stride 2			
Dense 模块(3)	14*14	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv	1*1 Conv * 6 3*3 conv
Transition层(3)	14*14	DenseNet-121			
	7*7	2*2 Average pool, stride 2			
Dense 模块(4)	7*7	1*1 Conv * 16 3*3 conv	1*1 Conv * 32 3*3 conv	1*1 Conv * 32 3*3 conv	1*1 Conv * 48 3*3 conv
Classification 层	1*1	7*7 全局平均值池化			
		5 维全连接，Softmax			

图 4-5 DenseNet 的 4 种网络结构示意图

（4）Xception

Xception[48]是谷歌研究组继 Inception 后提出的对 InceptionV3 的另一种改进.作者认为，通道之间的相关性与空间相关性最好要分开处理，Xception 采用深度可分离卷积模块替换原来 InceptionV3 中的卷积操作。Xception 在 ImageNet 比赛上比 InceptionV3 的准确率稍高，同时参数量有所下降，Xception 中加入的类似 ResNet 的残差连接机制也显著加快了 Xception 的收敛过程。潜在的问题是，虽然深度可分离卷积可以带来准确率的提升或是理论计算量的大幅下降，但由于其计算过程较为零散，现有的卷积神经网络实现中它的效率都不够高，例如本文中 Xception 的理论计算量是远小于 InceptionV3 的，但其训练时的迭代速度反而更慢一些。本文所采用的 Xception 结构如图 4-6 所示。

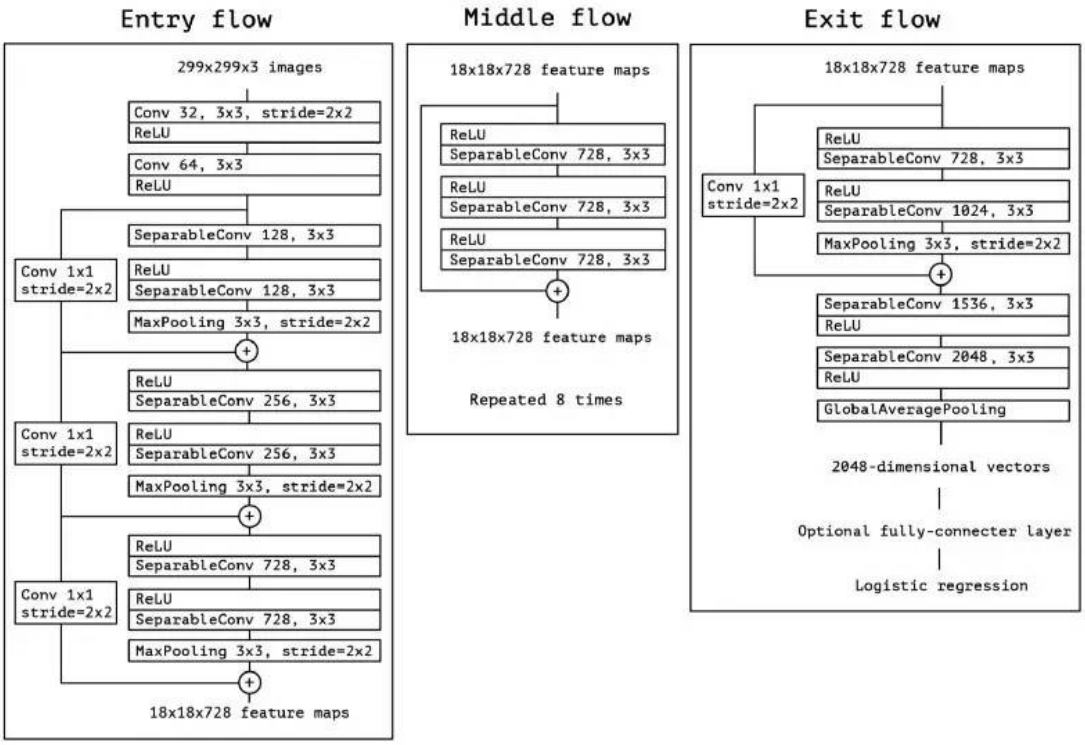


图 4-6 Xception 网络结构示意图

经过上面的介绍，本文选取 InceptionV3、InceptionResNetV2、DensNet121 和 Xception 这 4 种主流的卷积神经网络结构对红斑鳞屑类皮肤疾病进行分类。表 4-1 是本文实现的这 4 种卷积神经网络在 ImageNet 上的分类性能。采用 Tensorflow[49]作为深度学习框架，实验机器配置为 i7-8700K+3*TITAN Xp。本实验保存的各个 CNN 模型的预训练权重将用在后续皮肤疾病分类任务的迁移学习之中。

表 4-1 四种主流 CNN 模型在 ImageNet 数据集上的分类性能的比较

	Top-1Error	**Top-5Error**
InceptionV3	69.25	88.69
InceptionResNetV2	74.07	91.94
DensNet121	77.18	93.49
Xception	78.93	94.45

4.2 实验方案设计

为了评估 CNN 网络对于红斑鳞屑类皮肤疾病的分类能力，本章选取了上文介绍的 4 种 CNN 网络分别对湘雅 ES-Patch 数据集进行训练和测试，整体实验框图如图 4-7 所示。

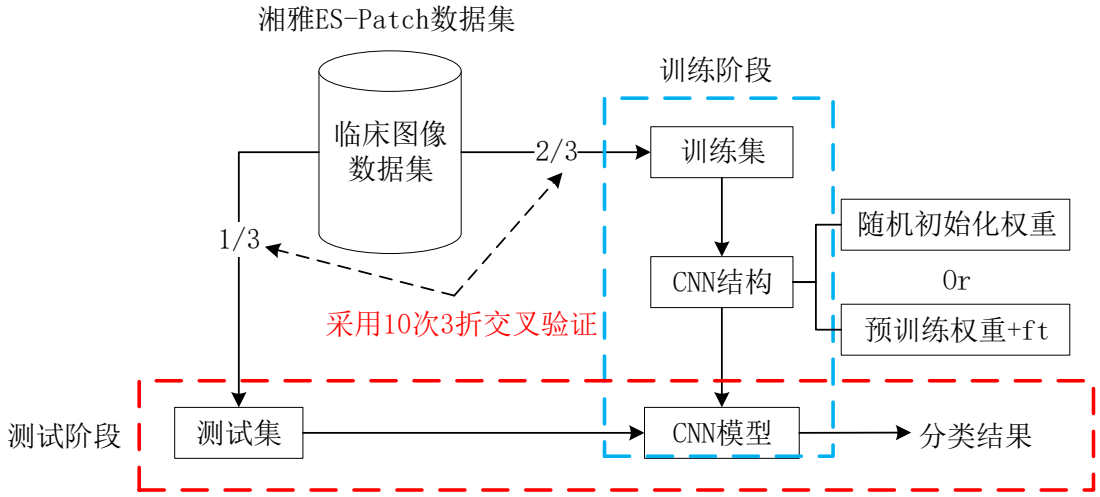


图 4-7 基于卷积神经网络的红斑鳞屑类皮肤疾病分类实验框图

如图 4-7 所示，为了保证实验的准确性，我们将图像按照不同的病人分开，训练集和测试集的分离同样采用 10 次 3 折交叉验证的方式，即训练集和测试集的比例为 2:1。为了保证对于 CNN 算法性能的公正性评估，4 种 CNN 训练均保持同样的硬件配置，迭代相同次数。为了对比迁移学习的作用，我们组织了从头开始训练和导入在 ImageNet 数据集的学习得到的预训练权重的对比实验。

此外，我们将对比实验中分类效果最优的模型与湘雅医院的 30 名专业医生进行了对比，对比的图像为经过病理金标准确诊的 100 例红斑鳞屑类皮肤疾病的临床图像。我们每个病例只挑选一幅图像，一共 100 幅，其中每类疾病各 20 幅。30 名医生都是拥有 5 年以上皮肤科临床工作经验的湘雅医院的皮肤科医生。对比的方式为：

①医生方面：我们制作了一个医生阅片的网站，如图 4-8 所示，医生可以自由的在线上对于这 100 幅临床图像进行判读，给出该病例是五类红斑鳞屑类皮肤疾病中的哪一类的答案，每例的判读时间不超过 30 秒。网站后台将对将判读结果与真实金标准结果对比，得到每位医师判读准确率；

②CNN 模型方面：直接将 100 例图像输入 CNN 模型，同样对模型给出的每例的判断结果与真实值进行比较得出 CNN 的判读准确率。

图 4-8 医生在线阅片的网站

4.3 实验结果分析

4.3.1 CNN 分类模型对比实验

经过实验，我们得到了 4 个 CNN 的 8 组分类实验的结果，其中带 ft 的是表示采用了该卷积神经网络在 ImageNet 上的预训练权重进行迁移学习，不带 ft 的表示是使用该卷积神经网络从头开始训练模型。分类实验结果统计如表 4-2 所示。

表 4-2 基于卷积神经网络的红斑鳞屑类皮肤疾病分类结果对比

	卷积神经网络模型	准确率
1	InceptionV3	73.53%
2	InceptionV3+ft	77.08%
3	InceptionResNetV2	75.32%
4	InceptionResNetV2+ft	78.53%
5	DenseNet121	76.87%
6	DenseNet121+ft	79.79%
7	Xception	78.73%
8	Xception+ft	81.60%

从表 4-2 的实验结果可以看出，基于迁移学习后的 CNN 分类准确率整体都上都比没有采用迁移学习的有所提升，这证明了迁移学习的策略确实对于小样本数据下的模式识别问题有较大的帮助。其中，基于 Xception+ft 的模型分类准确率可以达到 81.60%，是 8 组对比实验中准确率最高的，远优于传统视觉特征中的最优的分类特征 SIFT+CN 的 52.19%。

为了更加深入的分析 CNN 分类器对于每一类疾病的分类性能，我们统计了 Xception+ft 分类器的混淆矩阵，如表 4-3 所示

表 4-3 基于 Xception+ft 模型的分类混淆矩阵

		预测标签				
		扁平苔藓	副银屑病	红斑狼疮	湿疹	银屑病
实际标签	扁平苔藓	81%	3%	3%	12%	1%
	副银屑病	8%	69%	7%	5%	11%
	红斑狼疮	6%	5%	82%	7%	0
	湿疹	10%	4%	5%	80%	1%
	银屑病	0	2%	0	2%	96%

从混淆矩阵可以看出，每一行中准确率最高的值都集中在在对角线上，这说明了本模型对每一类疾病的分类准确率都还尚可。其中银屑病的分类效果最好，可以达到 96%，副银屑病的分类效果最差，只有 69%。从混淆矩阵，我们还可以对每类疾病的出错情况进行分析。在本实验中，副银屑病容易错分成银屑病，扁平苔藓和湿疹容易互相错分。

图 4-9 本数据集中临床特征相似的副银屑病和银屑病

我们对数据集中的一些错分的图像进行分析，如图 4-9 所示，第一行为副银屑病，第二行为银屑病，二者都具有红斑和白色鳞屑的特点，且有些图片病变区

域的轮廓都比较相似。扁平苔藓和湿疹也容易互相错分，我们同样将测试结果中部分出错（互相错分）的图像展示如图 4-10。这两种疾病在图像颜色、纹理和其病变区域分布上都有一定的相似度。

图 4-10 本数据集中临床特征相似的扁平苔藓和湿疹

4.3.2 CNN 模型与皮肤科专家医生对比实验

前文通过 CNN 分类模型的分类性能的对比实验，我们得出在本任务上最优的分类结果是由 Xception+ft 取得的。为了获取 CNN 模型相对于人类医生的鉴别水平，我们还组织了 CNN 模型与湘雅医院的 30 名专业医生的读图鉴别的对比实验，本实验设计的详情参见文中的 4.2 节。实验结果统计如表 4-4 所示。

表 4-4 CNN 诊断模型与 30 名专业皮肤科医生的对比

组别	扁平苔藓	副银屑病	红斑狼疮	湿疹	银屑病	总体
30 名医生平均	83%	75%	77%	78%	85%	79.60%
Xception+ft	81%	68%	80%	83%	93%	81.00%

由表 4-4 的结果可以看出，CNN 模型对于红斑鳞屑类皮肤疾病临床图像的鉴别能力与医生基本相当，甚至在总体的鉴别准确率略微优于医生。我们还可以发现，医生对于每种疾病的鉴别准确率更为平均，CNN 模型却在副银屑病这类疾病上准确率较低，而在银屑病这类上准确率很高。

我们认为造成 CNN 模型比皮肤科专家医生鉴别效果更好的原因是医生的诊断需要结合病人的病史或者部位信息，而我们的评判标准只考虑了临床图像。这种诊断场景对于医生来说并不习惯，而 CNN 模型以其强大的特征学习和图像分类能力在这个方面就凸显自己的优势了。值得强调的是，这一实验结果只是说明了 CNN 模型的“读图鉴别”能力比医生略优，不代表实际的诊断能力。

通过组织这一对比实验，我们也收到了各位专家医生对于组织“人机对战”的宝贵建议。比如对于医生方面需要给出明确皮损的部位、病人的基本病史（如发病的年龄、性别等）以及既往史等。通过医生反馈的问题，我们同时也认识到了本模型的缺点。本模型考虑的输入数据只有临床图像，没有考虑病人其他维度的临床信息。当本诊断模型需要诊断更多类别的疾病时，这种纯图像输入的模型的劣势就会突显出来。我们认为 CNN 的诊断模型也需要还原临床的真实诊断环境，我们下一步的工作是结合病史等维度的数据，去开发在临床环境中稳定使用的辅助诊断工具，服务广大的医生和患者。

4.4 本章小结

本章首先对卷积神经网络和迁移学习的理论基础进行了简要阐述，之后对后面实验用到的 4 种主流的卷积神经网络结构进行了介绍，包括 InceptionV3 模型、InceptionResNetV2 模型、DenseNet121 模型以及 Xception 模型。仍然是基于湘雅 ES-Patch 数据集，本章采用对 4 个 CNN 模型在是否采用迁移学习的实验条件下进行了 8 组对比实验。实验结果表明，基于迁移学习的策略在皮肤疾病这一小数据集上确实比从头开始训练的要具有优势。其中，Xception+ft 的模型分类准确率可以达到 8 组对比实验中最高的 81.60%，优于第 3 章中同样实验条件下的传统的视觉特征的分类效果。最后，我们还组织了 CNN 模型与专业皮肤科医生的“人机对战”实验。实验结果证明，CNN 模型在 100 张临床图像的识别任务上总体准确率略微优于医生，但是医生的识别准确率更为平均。由于本实验只考虑了临床图像识别，没有将更多病人信息提供医生，因此这一实验结果只是说明了 CNN 模型的“读图鉴别”能力比医生略优，并不代表实际的诊断能力。

5 基于多任务学习的红斑鳞屑类皮肤疾病病变区域检测

前面 2 章的研究对象都是基于湘雅 ES-Patch 数据库的图像，即是裁剪过后的以病变区域为主的 Patch 图像。本章将研究对象转移至带有病变区域标注的湘雅 ES-Raw 数据集，尝试利用多任务学习的方式，实现红斑鳞屑类皮肤疾病的病变区域的检测。其中，病变区域的检测包括 2 个子任务，即对于病变区域的定位和疾病类型的分类。本章将从多任务学习的理论介绍开始，介绍本文实现红斑鳞屑类皮肤疾病的病变区域检测的整个过程。

5.1 基于多任务学习的 Faster R-CNN

多任务学习（Multi-task learning，MTL）[50]是机器学习中一个很有前景的领域，其目标是利用多个学习任务中的相关信息来辅助单个任务的学习过程。一般而言，MTL 有两个基本因素[51]。第一个因素是任务之间的相关性，任务之间的相关性是基于对不同任务关联信息的理解，我们可以将这种关联信息融入到模型的设计之中。第二个因素是基于任务的定义。在机器学习中，学习任务主要包含分类和回归等监督学习任务、聚类等无监督学习任务和强化学习任务等，不同的学习任务对于多任务学习需要有不同的网络设置。

Faster R-CNN[52]是一个典型的基于多任务学习在深度学习中应用的例子，包含目标位置回归和目标物体分类两个任务，其中 RPN 模块的卷积层在两个任务之间共享。Faster R-CNN 是目标检测领域中的一个具有里程碑意义的网络，它第一次真正实现了目标检测网络端到端的训练，将图像基础特征提取网络、候选区域生成网络（RPN）、边界框（bbox）回归、目标分类整合到一个网络中，使目标检测速度有了很大提升。下面我们将以 ZFNet[53]作为特征提取网络为例，对 Faster R-CNN 的结构作简要介绍，其结构简图如图 5-1 所示。

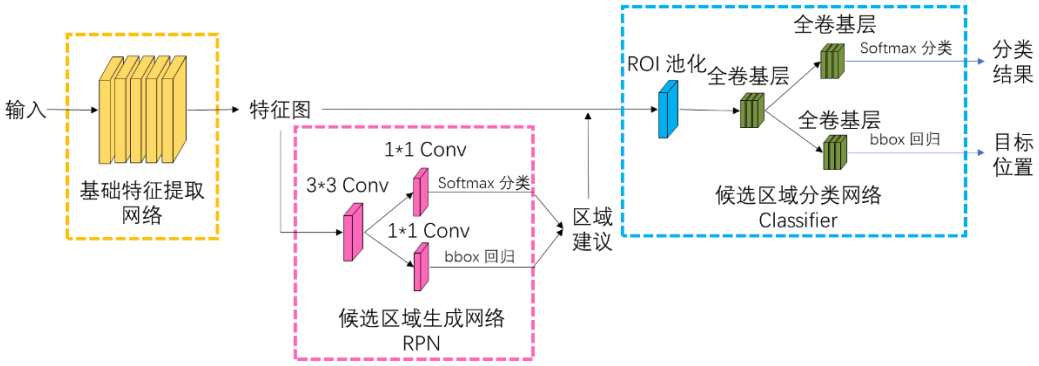


图 5-1 Faster R-CNN 结构简图

如上图所示，Faster R-CNN 的结构主要分为三大部分，第一部分是基础特征提取网络，第二部分是候选区域生成网络，第三部分是对候选区域进行分类的网络。下面分别对这 3 个部分做简要介绍：。

（1）基础特征提取网络：对于任意图像，我们可以利用特定的 CNN 结构（卷积+激活+池化层）提取候选图像的特征图。特征提取的网络可以是简单的 ZFNet 和 VGG16，也可以是复杂的 GoolgleNet 网络等。

（2）候选区域生成网络：候选区域生成的 RPN 网络可分为两个子网络，分类网络通过 softmax 分类 anchors（锚点）获得前景和背景，检测网络用于计算对于 anchors 的边界框回归的偏移量，以获得精确的目标候选区。图 5-2 为 RPN 网络的示意图，其输入为 ZFNet 最后的第五个卷积层的输出，维度是 256，假设特征图上的每个像素点上有 k 个 anchor，每个 anchor 判断是否是目标和背景的二分类，因此每个点的分类输出为 2k 维度的 scores（评分向量）。而对于位置回归，每个 anchor 都有[x, y, w, h]4 个偏移量，所以回归输出的维度为 4k。RPN 网络训练的损失函数的计算公式如下所示：

$$L_{reg}(t^u,v)=\sum_{i\in[x,y,w,h]} smooth_{L_1}\left(t_i^u-v_i\right),\quad smooth_{L_1}(x)=\begin{cases}0.5x^2, & |x|<1\\ |x|-0.5, & |x|\geq 1\end{cases} \tag{5-1}$$

$$L\left(\{p_i\},\{t_i\}\right)=\frac{1}{N_{cls}}\sum_i L_{cls}(p_i,p_i^*)+\lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i,t_i^*) \tag{5-2}$$

其中，式（5-1）表示 L_{reg} 的计算方式，采用 SmoothL1 损失函数进行训练。另外，在式（5-2）中，我们注意到回归误差中 L_{reg} 与 p_i^* 相乘，即 bbox 的回归只对包含分类目标的 anchor 计算误差。相应的，我们在标记数据的时候，也不用标记背景部分（正常皮肤部分、毛发部位或衣服等），只用标记各种皮肤疾病具有疾病特征的皮损部分。

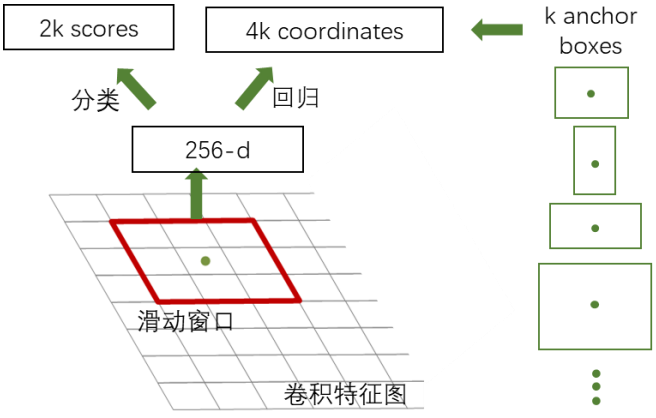


图 5-2 RPN 网络的示意图

（3）候选区域分类网络：候选区域分类网络 Classifier 利用 RPN 输出的区域建议提取 ROI（感兴趣区域）部分，接着送入全连接网络中作后续具体的目标分类和边界框的精细回归。与 RPN 网络类似，Classifier 网络也可分为目标分类和边界框回归两个部分。其中，分类器主要是对提取的 ROI 做最后的目标分类。由于 RPN 网络只是区分了图像中的前景和背景部分，本阶段的目标分类是对各类型的红斑鳞屑类皮肤疾病做具体分类。ROI 边框修正和 RPN 中的 anchor 边框修正原理一样，同样使用 SmoothL1 损失函数。值得注意的是，ROI 边框修正也是对于非背景的 ROI 进行修正，对于类别标签为背景的 ROI，则不进行 ROI 边框修正的参数训练。对于分类器和 ROI 边框修正的训练，可以公式描述如下：

$$L\left(p,u,t^{u},v\right)=L_{cls}(p,u)+\lambda\left[u\geq 1\right]L_{reg}(t^{u},v) \tag{5-3}$$

上式中 u>=1 表示 ROI 边框修正是对于非背景的 ROI 而言的，实验中，上式的 λ 取 1。总的来说，FasterR-CNN 的损失函数分两大块，第一大块是训练 RPN 的损失函数，第二大块是训练 FastR-CNN 中分类器损失函数，因此，总的 loss 函数可以描述如下：

$$L_{final}=L(\{p_i\},\{t_i\})+L(p,u,t^{u},v) \tag{5-4}$$

5.2 基于多尺度特征融合的 Faster R-CNN

Faster R-CNN 从提出以来一直是目标检测领域主流的框架之一，即使后面推出了 R-FCN[54]，Mask R-CNN[55]等改进版本，但基本结构变化不大。同时也有 SSD[56]，YOLO[57]等有开拓意义的新作，但精度上依然以 Faster R-CNN 为最好。对于一般行人，车辆，文字等通用目标检测问题，Faster R-CNN 都能得到非常好的效果。本文的目标检测对象为各种红斑鳞屑类皮肤疾病的临床图像，由于其拍摄距离的远近天然就存在不同尺度的大小，且由于其各分类目标（各种皮肤疾病）间差异较小属于细粒度分类。结合这两个特点，本文对 Faster R-CNN 的特征提取网络做出改进，提出了一种基于多尺度特征融合的改进方案，旨在通过不同层级的图像特征融合实现检测效果的提升。

5.2.1 改进动机

本文对 Faster R-CNN 改进的动机有 2 个：

（1）Faster R-CNN 的原网络没有对不同尺度的图像的输入做处理，尤其是本任务中输入的皮肤病图像通常会有不同尺度，即不同拍摄距离下皮损区域在图

像上的呈现尺度不同；

（2）红斑鳞屑类疾病属于一组临床表现类似的皮肤病，其分类性质属于细粒度分类。细粒度分类领域有研究报道过利用多尺度特征融合达到准确率提升的案例，因此，我们这里尝试 Faster R-CNN 的结构改进中加入这一思想。

下面将详细阐述这 2 个改进的动机：

1）对不同尺度的图像的输入做处理：由于卷积网络的平移不变性，CNN 对于预测目标的位置变化的适应性很强，但是对输入尺度变化的表现很差。目前，许多检测问题在处理多尺度图像输入的思路可以总结成图 5-3 中的 4 个方法，其中蓝色的线框表示特征图，边线越粗表示语义信息越强。

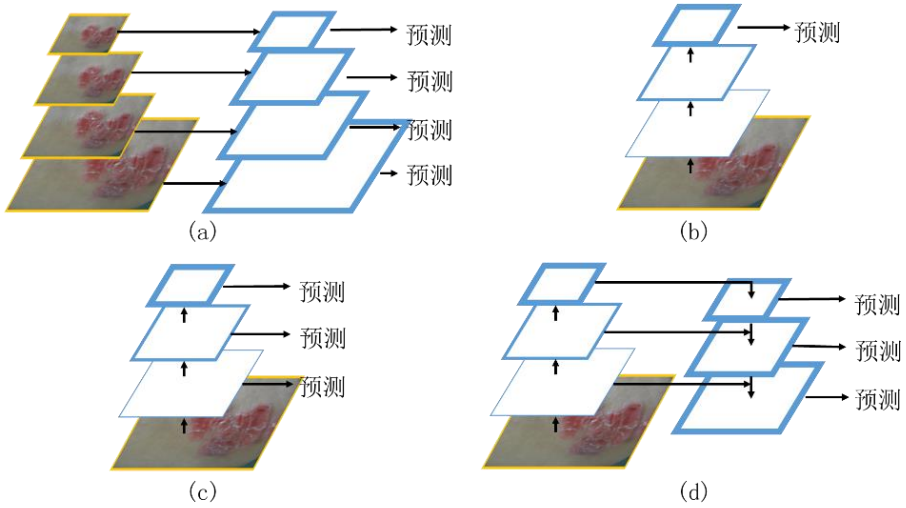


图 5-3 处理多尺度图像检测问题的思路

在图 5-3 中，(a)方法一般多用在一些非深度学习的算法中，通过将图像缩放到不同大小，构成一系列的图像金字塔网络，然后在每层提取特征并做出预测。这种方法在深度神经网络的设计中使用较为困难，主要就是内存和时间上的花销较大。(b)方法属于一般 CNN 做分类问题的模式，即利用 CNN 网络的顶层特征来进行预测，包括 Faster RCNN 也是这个思想。之所以利用顶层特征做预测，是因为 CNN 的卷积层越高，每个单元格的感受野就越大，语义化程度是越高，但是相对浅层特征就没有那样精细，分辨率也下降得比较厉害，这个问题在小目标上非常突出。也有研究人员尝试了类似图(c)的思路，即在顶层这种单元格感受野大的层上预测尺度比较大的物体，而在浅层这种感受野小、分辨率比较高的层上预测尺度比较小的物体。但是这样会导致的一个问题就是浅层虽然分辨率提高了，但是语义化程度还是不够高，去预测小目标效果还是不够好。(d)是一种基于特征融合的思路，即在顶层得到了语义化程度比较高的特征后，经过不断上采样，然后和 CNN 网络中那些浅层特征融合，这样得到的特征既有较高的语义性，又有较高的分辨率，这样再去分别预测不同尺度的物体就会有比较好的效果了。

2）细粒度图像分类领域结合多尺度特征融合的方法：细粒度图像分类问题是计算机视觉领域一项极具挑战的研究课题,其目标是对子类进行更加精细的识别，如区分不同种类的鸟。由于子类别间细微的类间差异和较大的类内差异,传统的分类算法不得不依赖于大量的人工标注信息。近年来,随着深度学习的发展，深度卷积神经网络为细粒度图像分类带来了新的机遇。大量基于深度卷积特征算法的提出,促进了该领域的快速发展。2017 年，微软亚洲研究院提出了一个基于多尺度特征融合的细粒度分类网络 RA-CNN[58]，在 CUB-200-2011（鸟类数据集）,StanfordDogs（狗类数据集）和 StanfordCars（车类数据集）上都取得了当时最优的分类效果。RA-CNN 通过设定特定的损失函数进行调优，并结合了三个不同尺度的特征进行融合，采用 ImageNet 上预训练的权重初始化三个尺度的分类网络，实验结果确实证明了这种特征融合的方式有助于提升图像细粒度分类的准确性。

5.2.2 网络结构

根据前面思路的介绍，我们可以知道 CNN 的高层特征具有丰富的语义信息，对检测网络的分类任务有提升作用。而低层特征感受野小，具有较高空间分辨率，对于检测任务可以帮助获得更好的精度。因此，为了结合多尺度的卷积神经网络特征，我们设计了一个基于多尺度特征融合的 Faster R-CNN 网络。其中，特征提取网络采用的是如图 5-4 所示 VGG16 模型，整体的检测模型如图 5-5 所示。

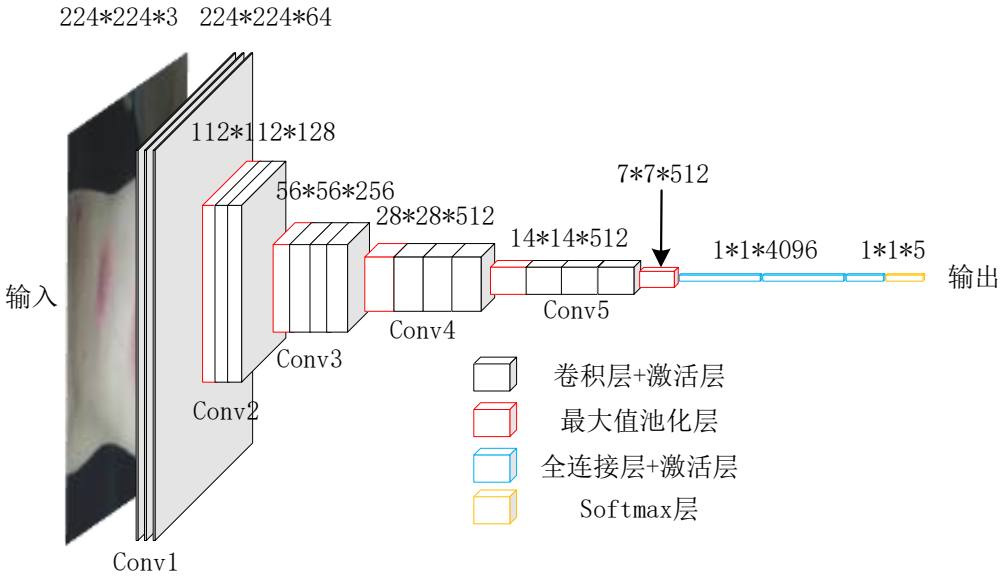


图 5-4 基于 VGG16 的特征提取网络

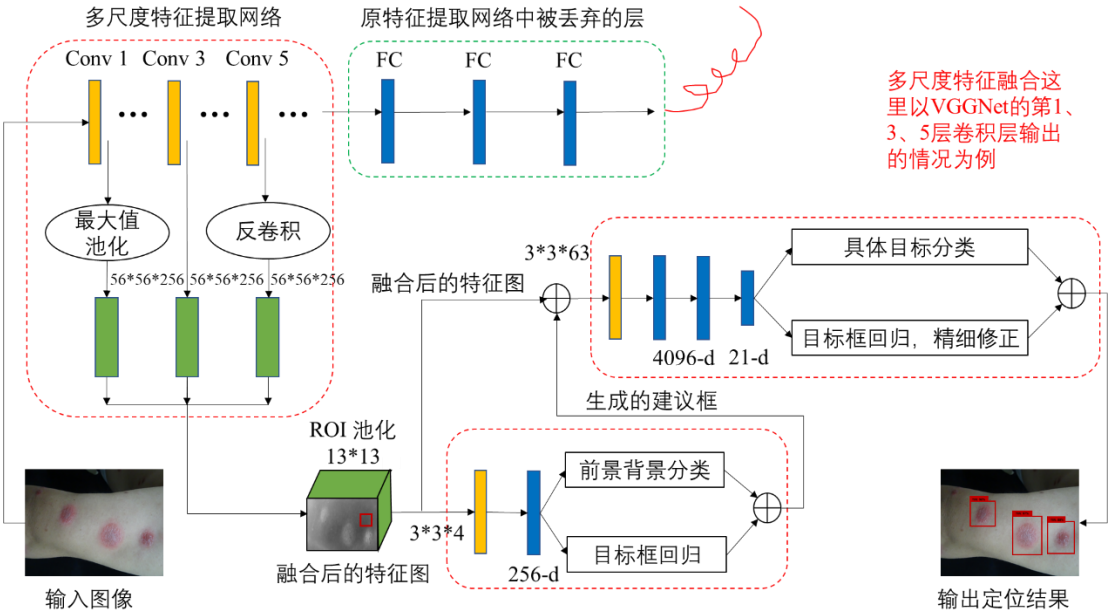


图 5-5 基于多尺度特征融合的 Faster R-CNN 网络

从图 5-5 中我们可以看到，对比原始的 Faster R-CNN 网络，我们主要对特征提取网络部分进行了修改。图示是以 VGG16 的 Conv1、Conv3、Conv5 三个卷积层的特征进行融合的情况为例，实际情况下，针对 VGG16 会有不同的特征融合组合，下文实验中我们将针对不同的特征融合组合进行对比实验，探讨最优的融合策略。

5.3 实验方案设计

本章实验主要是采用上文设计的基于多尺度特征融合的 Faster R-CNN 网络进行红斑鳞屑类皮肤疾病病变区域检测。实验数据集采用的是湘雅 ES-Raw 数据集，包含红斑鳞屑类皮肤疾病的原图和病变区域的边界框。为了评估多尺度特征融合的最优策略，我们设置了 9 组不同特征融合策略：

（1）3 组单层卷积特征：我们分别选择了特征提取网络 VGG16 的 Conv1，Conv3 和 Conv5 三个卷积层的输出作为特征提取结果，Conv1 层需要最大值池化，Conv5 需要做上采样，保证提取到的特征图大小都是 56*56*256。之所以选择这 3 个卷积层，是因为这三个卷积层处于 VGG16 的初级、中级和高级三个阶段，其特征代表了图像不同语义级别和空间分辨率，具有一定的代表性；

（2）3 组两层卷积层的融合特征：我们分别选择了 Conv1+Conv3，Conv3+Conv5 和 Conv1+Conv5 这 3 个特征融合的组合，融合特征同样会采取特征归一化操作，保证最终输出的特征图大小一致。可以看到，我们没有选择相邻的 2 个卷积层，是因为相邻的 2 个卷积层跨度不够大，特征图之间可能具有一定

的相似性。之所以选择这 3 个组合，也是分别选取了 VGG16 的不同阶段的特征图进行组合，融合的特征也分别代表了不同语义级别特征图的组合

（3）3 组三层卷积层的融合特征：我们分别选择 Conv1+Conv2+Conv3，Conv3+Conv4+Conv5 和 Conv1+Conv3+Conv5 这 3 个特征融合的组合，这里的 Conv1+Conv2+Conv3 和 Conv3+Conv4+Conv5 分别代表了相对初级和高级语义级别的特征图组合，而 Conv1+Conv3+Conv5 采用了间隔一层卷积层的形式进行特征融合，包含了高、中、低三个层级的特征。

对于对比实验结果的性能评价，我们选用均值平均精度 *mAP* 来对模型的定位和分类精度同时进行评价。*mAP* 的定义首先出现在 PASCAL Visual Objects Classes(VOC)竞赛中，按照定义其计算过程主要与以下 5 个式子相关。

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5-5}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5-6}$$

$$IoU = \frac{Box_{predict} \cap Box_{true}}{Box_{predict} \cup Box_{true}} \tag{5-7}$$

$$AP = \int_0^1 \text{Precision}(\text{Recall}) d(\text{Recall}) \tag{5-8}$$

$$mAP = \frac{1}{classes} \sum_{i=1}^{classes} \int_0^1 \text{Precision}(\text{Recall}) d(\text{Recall}) \tag{5-9}$$

在上述式子中，Precision 表示准确率，Recall 表示召回率，*IoU* 表示检测目标框与真值目标框的重合度，*AP* 为检测目标中的某类的平均精度。对于 *mAP* 的计算过程，我们可以以下面的方式来进行推导。

首先 *mAP* 是由对于检测目标的平均精度 *AP* 相加再求平均值计算得来的，即公式（5-9），其中对于本任务，classes 为 5。然后，由公式（5-8）可知平均精度 *AP* 为准确率-召回率曲线的曲线下面积，然而目标框的置信度阈值和 *IoU* 都会影响 Precision 和 Recall，这里的做法我们是参考 PASCAL VOC 比赛，将 *IoU* 设为固定值 0.5，即重合度为 0.5 以上视为该目标被检测出来。然后，根据设置不同的目标框的置信度阈值，由公式（5-5）和（5-6）就能计算出准确率和召回率，从而就能计算出该分类目标的平均精度 *AP* ，最终就能得出最终的评价指标均值平均进度 *mAP* 。

5.4 实验结果分析

对于每种特征融合策略，我们对于融合后的特征图都采用了归一化的操作，保证所有对比实验的配置一样。以 *IoU* 为 0.5 视为目标被检出，最终得到的检测实验结果如表 5-1 所示。

表 5-1 病变区域检测实验的结果对比

	特征融合组合	*mAP*
1	Conv1	62.58%
2	Conv3	63.00%
3	Conv5	65.62%
4	Conv1+Conv3	65.82%
5	Conv3+Conv5	66.89%
6	Conv1+Conv5	67.60%
7	Conv1+Conv2+Conv3	66.82%
8	Conv3+Conv4+Conv5	71.79%
9	Conv1+Conv3+Conv5	73.60%

从上面的实验结果，我们可以分析出以下几点发现：

（1）从整体实验结果上来看，对比特征融合组合 1、2、3，4、5、6 和 7、8、9，多尺度特征的融合确实能提升模型检测的效果；

（2）对比特征融合组合 1、2 和 3，我们发现卷积神经网络的高层特征比低层特征的对检测结果的贡献度更好，因此目前许多分类和检测系统都采用 CNN 的高层特征作为后续区域建议网络的输入；

（3）对比特征融合组合 7、8 和 9，最优的特征融合组合是为序号为 9 的 Conv1+Conv3+Conv5，可以达到 73.60%的 *mAP*。这表明，相邻的卷积层之间特征具有一定的相关性，利用卷积神经网络的更广范围的卷积特征的融合能更有效地提升目标检测任务的效果。

最优融合策略 Conv1+Conv3+Conv5 下的一些图像的病变区域检测结果如图 5-6 所示，图中其中从上到下每行各代表一类皮肤疾病，分别为扁平苔藓、副银屑病、红斑狼疮、湿疹和银屑病。图中不同颜色的矩形框表示模型预测的病变区域位置和疾病的类别，每个矩形框上百分比的数字表示该条预测的置信度。由于图像较小，我们将病变区域检测结果用文字的形式展现在对应的图像下方，例如，对于第一张图像，网络的检测结果为：“#部位 1 扁平苔藓：87%”。注意，这里展示的都是检测框与真值框的重合度 *IoU* 在 0.5 以上的检测结果。

图 5-6 红斑鳞屑类皮肤病变区域检测结果示意图

从图中我们可以看出，本实验结果整体上都能找到病变区域的大体位置，且置信度都比较高。不足之处是预测的目标框较真实病变区域偏大，这个是与数据集构建的时候的图像标签有关，即实际作为训练标签的人工框选的目标就框选得过大了一些。另外，副银屑病的识别效果较差，其置信度较低，从图中第二行给出检测结果就可以看出。究其原因，我们认为是本数据集中副银屑病的疾病特征本身就比较复杂，在类内差异就较大，因此 CNN 网络对该类识别效果差，在前

文的病变区域图像分类试验中，副银屑病的准确率就是五种疾病中最低的。

本实验中病变区域检测的方法最好的均值平均精度为 73.60%，之前图像分类实验中最高准确率为于 Xception+ft 的 81.60%。然而，这两个数字是不能简单进行比较的。首先，这两个实验采用的数据集就是不一样的。图像分类实验采用的是湘雅 ES-Patch 数据集，为裁剪保存的皮肤病变区域的图像。而目标检测实验采用的是 ES-Raw 数据集，背景更为复杂。然后，这两个实验的评价指标也是不一样的，目标检测实验的评价指标更为严格，它是考虑了病变区域定位任务和疾病分类任务的综合指标。

5.5 本章小结

本章的主要内容是对红斑鳞屑类皮肤疾病的原图进行病变区域检测。首先，对多任务学习的理论进行了简要介绍，重点介绍了多任务学习中的两种参数共享机制。接着，对基于多任务学习在深度学习中应用——Faster R-CNN 的整体框架进行了阐述。本文针对皮肤疾病数据集图像具有多种尺度和细粒度分类的特点，提出了基于 Faster R-CNN 的多尺度特征融合的改进策略。针对不同的特征融合组合，我们设计了 9 组对比实验。在湘雅 ES-Raw 数据集上对算法进行评估，最优的特征融合组合为 Conv1+Conv3+Conv5，可以达到 73.60%的均值平均精度。这表明，利用处于卷积神经网络的更广范围的由低级到高级的特征间的融合可以有效提升目标检测任务的效果。

6 红斑鳞屑类皮肤疾病辅助诊断系统的设计与实现

红斑鳞屑类皮肤疾病是一类临床表现极其相似的一类皮肤疾病，但治疗手段差异很大，其鉴别诊断历来是皮肤病科的一个难题。目前，临床上的红斑鳞屑类皮肤疾病的诊断需要结合临床表现、皮肤镜诊断以及组织病理学检查等，对医生诊断水平和医疗设备要求都较高。在中国的许多基层地区，由于缺乏专业的皮肤科医生和医学设备，这类疾病的误诊和漏诊现象时有发生。为了能缓解这一问题，本文设计并开发了一款红斑鳞屑类皮肤疾病的辅助诊断系统。本系统能辅助医生对于红斑鳞屑类皮肤疾病的诊断决策，为患者提供智能自诊，同时还能在实际应用中积累更加全面的数据来进行算法的更新和迭代。

6.1 系统功能

基于上文的介绍的红斑鳞屑类皮肤疾病的分类和检测算法，本系统接收的系统输入为皮肤疾病的临床图像，涉及疾种包括扁平苔藓、副银屑病、红斑狼疮、湿疹和银屑病，具有如下的系统功能：

（1）基于用户框选病变区域的皮肤疾病分类

由于本文训练的皮肤疾病分类模型是基于 Patch 数据集，所以本系统疾病分类功能接受的图像输入也是基于疾病病变区域的。如图 6-1 所示，用户拍照上传的图像可能包含一些复杂的背景干扰，本系统提供矩形框裁剪的功能，用户可以自行裁剪病变区域的大致位置。系统接受用户图像上传之后，利用上文基于卷积神经网络的疾病分类算法，将给出该图像最有可能的三个分类结果以及分别的置信度。如图 6-1，图中给出的诊断结果表示本系统对该上传图像的识别结果为：92.2%的可能性为湿疹，4.1%的可能性为扁平苔藓，3.7 的可能性为银屑病。

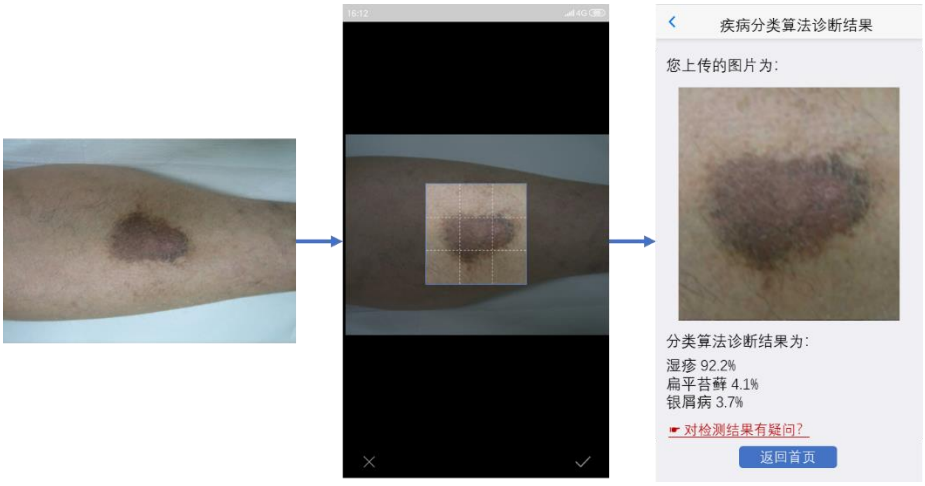


图 6-1 基于用户框选病变区域的皮肤疾病分类

（2）全自动皮肤疾病图像病变区域检测

由于目标检测算法的特性，本功能模块可以直接接受皮肤疾病的原始拍摄图像输入。对于传入图像，系统将利用训练好的病变区域检测模型给出病变区域检测结果，在返回的图像上将以矩形框的形式给出病变区域的位置、疾病的名称以及置信度。病变区域检测模型的参数设置为重合度 *IoU* 50%以上。如图 6-2 中，图中给出了检测后的图像，图中黄色的矩形框代表检测算法检测到了一个病变区域的位置，识别结果为该病变区域可能是银屑病，置信度为 85%。

图 6-2 全自动皮肤疾病图像病变区域检测

（3）诊断结果错误上报和诊断模型更新

由于本系统的分类和检测功能都有一定水平的限制，而且都是采用的基于卷积神经网络的深度学习算法，因此我们需要更多的皮肤疾病数据来对算法进行训练和迭代，尤其是当前算法模型识别错误的数据。因此，我们设置了错误结果上报的功能，给患者提供反馈意见的渠道的同时也可以帮助实现算法的不断更新。

图 6-3 诊断结果错误上报和诊断模型更新

如图 6-3 中的诊断结果页面，用户对于检测结果有疑问或质疑的可以点击“对检测结果有疑问”进入到“诊断结果错误上报”页面。在诊断结果上报页面，用户可以上传相关检测错误的图片，填写错误情况说明，然后点击“点击上报”按钮将该条情况上报。服务器后台将接收到该条上报记录并插入到数据库。

专家医生将会定期对上报上来的错误记录进行查阅，核对每条记录中检测图像对应的真实标签，核对完毕后这些用户上传的数据将被加入到系统的待训练数据集中。当待训练数据集累计到一定数量的时候，系统将自动启动模型的重新训练，实现诊断模型的更新。

6.2 系统设计与实现

本文采用 C/S 网络架构进行红斑鳞屑皮肤疾病辅助诊断系统的开发，系统分为客户端和服务器两个组成部分。如图 6-4 所示，客户端为可以安装在手机和平板等智能设备上的 APP 应用程序，用户可以通过该 APP 上传患病照片并接收来自于服务器端智能算法处理的结果。服务器端等待来自客户端的请求，发现任务即调用相应算法处理该任务并回传结果到客户端。二者的通信采用 HTTP 协议实现。下面分别对本系统客户端和服务器端的实现方法和工作流程做简要介绍。

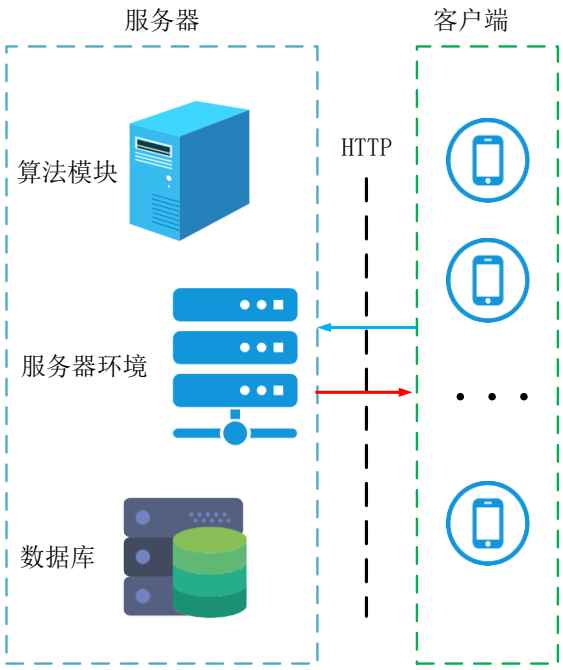


图 6-4 基于 C/S 网络架构的系统设计

（1）客户端

客户端采用 Html5 技术进行开发，编程语言为 HTML、JavaScript、CSS。基于 Html5 编写的 APP 可以实现与原生 APP 同样强大的功能和性能，具有快速开

发、低成本、易维护和跨平台等优点。本客户端支持 Android 4.4 及以上版本和 IOS 5.0 及以上版本。其工作流程图如图 6-5 所示，现按照 APP 的页面跳转方式将 APP 工作流程介绍如下：

1）任务选择页面：用户在 APP 内可以根据需求选择疾病分类任务或者病变区域检测任务；

2）图片上传页面：用户在该页面通过调用相机或者相册选择待上传的图片。其中，疾病分类任务需要用户对于病变区域进行自动框选，系统将保存用户框选后的图片区域进行上传；

3）等待结果页面：用户点击上传按钮，客户端将提交任务类型信息和待识别图片到服务器，之后，客户端将停留在等待页面，等待服务器返回识别后的结果；

4）结果展示页面：客户端得到服务器处理结果，呈现在诊断结果页面。用户可以查看和保存诊断结果，对于诊断结果有疑问的，用户还可以点击“对检测结果有疑问”进入诊断结果错误上报页面，上报相关错误信息。对于诊断结果没有疑问的用户则可以选择重新上传图片或直接退出客户端。

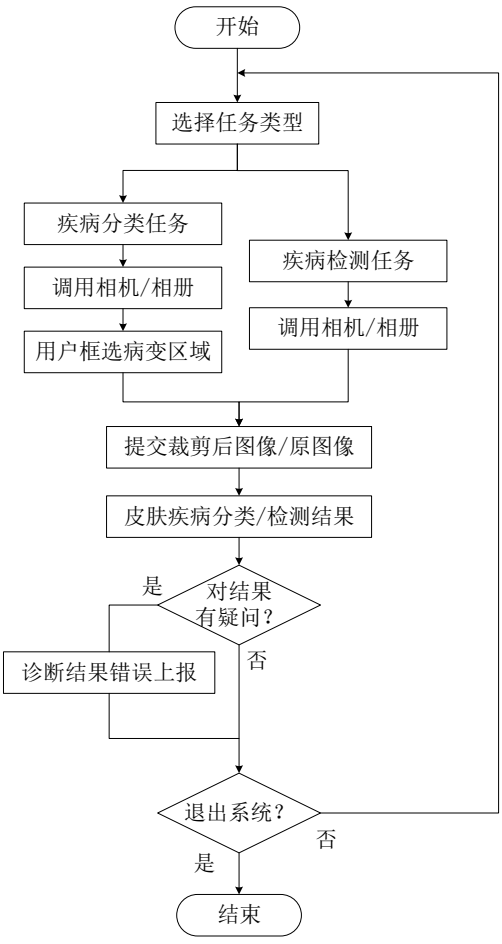


图 6-5 客户端工作流程

（2）服务器端

服务器端包括 WEB 服务模块和深度学习算法模块。WEB 服务基于 Django 框架进行实现，数据库采用 MySQL。深度学习算法模块采用 Keras 高级神经网络 API 进行搭建，并采用 NVIDIA 推出的通用并行计算架构 CUDA 和 CUDNN 进行运算加速。其工作流程图如图 6-6 所示，现简要介绍服务器端工作流程如下：

首先是服务端诊断任务处理流程：

a）WEB 服务一直都在启动状态，等待新的诊断任务请求的到来；

b）接收到新的诊断任务后，WEB 服务将根据传入的任务类型调用相应算法模型对传入图片进行识别；

c）识别结束后，算法模块将释放计算资源，并将结果返回给客户端，继续等待新的任务到来；

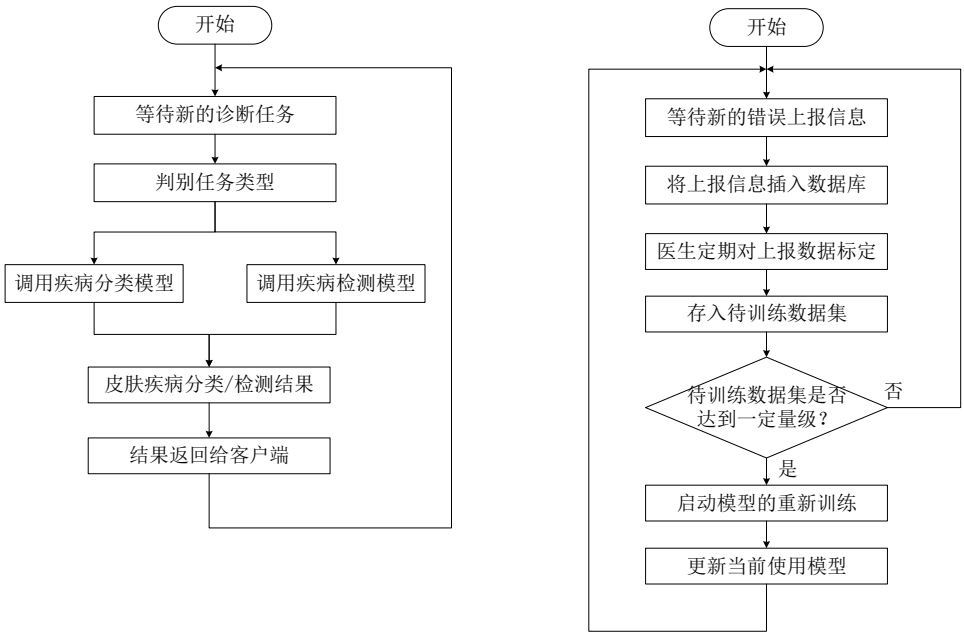
其次是服务端的诊断模型更新流程：

a）WEB 服务一直都在启动状态，等待新的诊断错误上报的信息；

b）接收到新的诊断错误上报信息后，WEB 服务将对应诊断错误信息解析后插入数据库表中；

c）专家医生将定期对上报的诊断错误记录进行查阅，核对每条记录中的真实标签，核对完毕的数据将被加入到系统的待训练数据集中；

d）判断训练数据集是否达到一定量级，如果是则启动模型的重新训练并将新训练的模型替换当前使用的诊断模型，如果未达到一定量级则继续等待新的错误上报信息。



(a) 诊断任务处理流程　　(b) 诊断模型更新流程

图 6-6 服务器端工作流程

6.3 本章小结

本章首先介绍了红斑鳞屑类皮肤疾病辅助诊断系统的主要功能，包括疾病分类，病变区域检测和基于用户反馈诊断错误信息的诊断模型更新。接着，本章对系统设计中客户端和服务器端的具体实现和工作流程做简要介绍，客户端为供患者医生使用的移动端的 APP，用户可以通过操作该 APP 上传图片分析任务，服务器端接收任务后调用相应算法模块进行处理并将结果返回给客户端。此外，本章还着重介绍了本系统诊断模型自我更新的机制，即我们在客户端设计了一个用户可以反馈本诊断模型识别可能有误的图像的渠道。系统将利用这些用户上报的数据进行算法的更新和迭代，使得模型诊断性能更加鲁棒。

7 总结与展望

7.1 论文总结

红斑鳞屑类皮肤疾病是一类临床表现极其相似的一类皮肤疾病，其鉴别诊断历来是皮肤病科的一个难题。目前，临床上的红斑鳞屑类皮肤疾病的诊断需要结合临床表现、皮肤镜诊断以及组织病理学检查等，对医生诊断水平和医疗设备要求都较高。本文提出一项针对红斑鳞屑类皮肤疾的研究计划，研究疾病的种类包括扁平苔藓、副银屑病、红斑狼疮、湿疹和银屑病，旨在通过计算机视觉的相关算法，开发一套对各类型红斑鳞屑类皮肤疾病进行准确分类和定位的病辅助诊断系统。具体工作如下：

（1）由于目前现有的皮肤疾病数据集来源主要以欧美人种为主，带有水印，且包含的红斑鳞屑类皮肤疾病的图片较少。因此，我们与合作单位湘雅医院一起构建了一个基于中国人种的皮肤疾病临床图像数据集，数据集中每张图像都带有病变区域的标记信息，且具有患者的全部医学资料作为支持。据我们了解，本数据集是目前世界范围内用于计算机辅助诊断（CAD）系统研究的最大的中国人种皮肤疾病临床图像数据集。

（2）基于视觉特征的红斑鳞屑类皮肤疾病分类。本文引入了 5 类图像视觉特征：包括 SIFT 特征、HOG 特征、LBP 特征、CH 特征和 CN 特征，这些特征一定程度上描述了皮肤疾病的一些临床表现。基于湘雅 ES-Patch 数据集，本章采用交叉验证的实验方式利用多分类的 SVM 方法对上述视觉特征的分类效果进行了评估。实验结果表明，基于传统视觉特征的红斑鳞屑类疾病的鉴别准确率普遍都不理想，最高准确率也只有 52.19%，这说明了依靠传统视觉特征的方法无法很好地承担红斑鳞屑类疾病的鉴别任务。

（3）基于卷积神经网络的红斑鳞屑类皮肤疾病的分类。卷积神经网络是目前深度学习领域最前沿和热门的方法之一。尤其在计算机视觉方面，卷积神经网络的应用开启了一个全新的时代。本文引入了 4 种主流的卷积神经网络：包括 InceptionV3、InceptionResNetV2、DenseNet121 以及 Xception。基于湘雅 ES-Patch 数据集，本章采用对 4 个 CNN 模型在是否采用迁移学习的实验条件下进行了 8 组对比实验。实验结果表明，基于迁移学习的策略在皮肤疾病这一小数据集上确实比从头开始训练的要具有优势。其中，Xception+ft 的模型分类准确率可以达到 8 组对比实验中最高的 81.60%，优于第 3 章中同样实验条件下的传统的视觉特

征的分类效果。此外，我们还组织了 CNN 模型与 30 名专业皮肤科医生的“人机对战”实验。实验结果表明，CNN 模型在总体准确率略微优于医生，但是医生的识别准确率更为平均。由于本实验只考虑了临床图像识别，没有将更多病人信息提供医生，因此这一实验结果只是说明了 CNN 模型的“读图鉴别”能力比医生略优，并不代表实际场景中的诊断能力。

（4）基于多任务学习的红斑鳞屑类皮肤疾病病变区域检测。不同于（2）和（3）中的疾病分类任务，本文针对皮肤疾病辅助诊断引入了多任务学习的概念，包括疾病区域定位和疾病类型分类两个任务。本文针对皮肤疾病数据集图像具有多种尺度和细粒度分类的特点，提出了基于 Faster R-CNN 的多尺度特征融合的改进策略。针对不同的特征融合组合，我们设计了 9 组对比实验。在湘雅 ES-Raw 数据集上，最优的特征融合组合为 Conv1+Conv3+Conv5 可以达到 73.60%的均值平均精度。实验结果表明，利用处于卷积神经网络的更广范围的由低级到高级的特征间的融合可以有效提升目标检测任务的效果。

（5）设计并实现了一个红斑鳞屑类皮肤疾病的辅助诊断系统。本文基于辅助诊断算法研究的内容，具有基于临床图像的疾病分类，病变区域检测和基于用户反馈诊断错误信息的模型更新迭代等功能。本系统能辅助医生对于红斑鳞屑类皮肤疾病的诊断决策，为患者提供智能自诊，同时还能在实际应用中不断积累更加全面的数据来进行算法的更新和迭代。

7.2 工作展望

本研究具有很大的应用前景和社会意义，有助于协助基层医院医生进行红斑鳞屑类皮肤疾病的初级诊断，辅助医生进行基本的诊疗决策，同时还能为患者实现智能自诊。但是，我们也清楚地认识到仅仅依靠临床图像来进行疾病辅助诊断还是不够的。接下来的工作，我们将会加入更多维度的医学诊断信息，扩充分类的疾病的种类，争取更大程度的还原临床真实的诊断环节，推动人工智能的疾病辅助诊断的落地应用。

参考文献

[1] Güvenir H A, Demiröz G, Ilter N. Learning differential diagnosis of erythemato-squamous diseases using voting feature intervals[J]. Artificial intelligence in medicine, 1998, 13(3): 147-165.

[2] Errichetti E, Stinco G. Dermoscopy in differential diagnosis of palmar psoriasis and chronic hand eczema[J]. The Journal of dermatology, 2016, 43(4): 423-425.

[3] Güvenir H A, Emeksiz N. An expert system for the differential diagnosis of erythemato-squamous diseases[J]. Expert Systems with Applications, 2000, 18(1): 43-49.

[4] Abbasi N R, Shaw H M, Rigel D S, et al. Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria[J]. JAMA, 2004, 292(22): 2771-2776.

[5] Argenziano G, Fabbrocini G, Carli P, et al. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis[J]. Archives of dermatology, 1998, 134(12): 1563-1570.

[6] Yang J, Sun X, Liang J, et al. Clinical skin lesion diagnosis using representations inspired by dermatologist criteria[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 1258-1266.

[7] Cortes C, Vapnik V. Support-vector networks[J]. Machine learning, 1995, 20(3): 273-297.

[8] Ho T K. Random decision forests[C]//Proceedings of 3rd international conference on document analysis and recognition. IEEE, 1995, 1: 278-282.

[9] Altman N S. An introduction to kernel and nearest-neighbor nonparametric regression[J]. The American Statistician, 1992, 46(3): 175-185.

[10] Kamiński B, Jakubczyk M, Szufel P. A framework for sensitivity analysis of decision trees[J]. Central European journal of operations research, 2018, 26(1): 135-159.

[11] Esteva A , Kuprel B , Novoa R A , et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639):115-118

[12] Han S S, Kim M S, Lim W, et al. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm.[J]. Journal of Investigative Dermatology, 2018.

[13] Sun X , Yang J , Sun M , et al. A Benchmark for Automatic Visual Classification of Clinical Skin Disease Images[C]// European Conference on Computer Vision. Springer International Publishing, 2016.

[14] Liao H. A deep learning approach to universal skin disease classification[R]. University of Rochester Department of Computer Science, CSC, 2016.

[15] Yu L, Chen H, Dou Q, et al. Automated melanoma recognition in dermoscopy images via very deep residual networks[J]. IEEE transactions on medical imaging, 2017, 36(4): 994-1004.

[16] Yang J, Xie F, Fan H, et al. Classification for Dermoscopy Images Using Convolutional Neural Networks Based on Region Average Pooling[J]. IEEE Access, 2018, 6: 65130-65138.
[17] Xie F, Fan H, Li Y, et al. Melanoma classification on dermoscopy images using a neural network ensemble model[J]. IEEE transactions on medical imaging, 2017, 36(3): 849-858.
[18] 王诗琪, 刘洁, 朱晨雨, et al. 皮肤科医师与深度卷积神经网络诊断色素痣和脂溢性角化病皮肤镜图像比较[J]. 中华皮肤科杂志, 2018, 51(7):486.
[19] Wang J , Mackenzie J D , Ramachandran R , et al. Neutrophils Identification by Deep Learning and Voronoi Diagram of Clusters[C]// Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, 2015.
[20] Pal A., Chaturvedi A., Garain U., et al. CapsDeMM: Capsule Network for Detection of Munro's Microabscess in Skin Biopsy Images. [C]// Medical Image Computing and Computer-Assisted Intervention – MICCAI 2018. Springer International Publishing, 2018.
[21] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
[22] Csurka G, Dance C, Fan L, et al. Visual categorization with bags of keypoints[C]//Workshop on statistical learning in computer vision, ECCV. 2004, 1(1-22): 1-2.
[23] Surhone L M, Tennoe M T, Henssonow S F, et al. Histogram of Oriented Gradients[J]. Betascript Publishing, 2016, 12(4):1368 - 1371.
[24] Swain M J, Ballard D H. Color indexing[J]. International journal of computer vision, 1991, 7(1): 11-32.
[25] Van De Weijer J, Schmid C, Verbeek J, et al. Learning color names for real-world applications[J]. IEEE Transactions on Image Processing, 2009, 18(7): 1512-1523.
[26] Berlin B, Kay P. Basic color terms: Their universality and evolution[M]. Univ of California Press, 1991.
[27] Benavente R, Vanrell M, Baldrich R. Parametric fuzzy sets for automatic color naming[J]. JOSA A, 2008, 25(10): 2582-2593.
[28] Chang C C, Lin C J. LIBSVM: a library for support vector machines[J]. ACM transactions on intelligent systems and technology (TIST), 2011, 2(3): 27.
[29] Ding C H Q, Dubchak I. Multi-class protein fold recognition using support vector machines and neural networks[J]. Bioinformatics, 2001, 17(4): 349-358.
[30] Goring C, Rodner E, Freytag A, et al. Nonparametric part transfer for fine-grained recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 2489-2496.
[31] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
[32] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and

pattern recognition. 2016: 779-788.
[33] Sainath T N, Weiss R J, Wilson K W, et al. Multichannel signal processing with deep neural networks for automatic speech recognition[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017, 25(5): 965-979.
[34] Hörnstein J, Gustavsson L, Lacerda F, et al. Multimodal word learning from infant directed speech[C]//2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2009: 2748-2754.
[35] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006, 18(7): 1527-1554.
[36] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint arXiv:1502.03167, 2015.
[37] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
[38] Setio A A A, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks[J]. IEEE transactions on medical imaging, 2016, 35(5): 1160-1169.
[39] Anthimopoulos M, Christodoulidis S, Ebner L, et al. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network[J]. IEEE transactions on medical imaging, 2016, 35(5): 1207-1216.
[40] Shin H C, Roth H R, Gao M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning[J]. IEEE transactions on medical imaging, 2016, 35(5): 1285-1298.
[41] Tajbakhsh N, Shin J Y, Gurudu S R, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning?[J]. IEEE transactions on medical imaging, 2016, 35(5): 1299-1312.
[42] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
[43] Chen H, Qi X, Yu L, et al. DCAN: deep contour-aware networks for accurate gland segmentation[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2016: 2487-2496.
[44] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2818-2826.
[45] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//Thirty-First AAAI Conference on Artificial Intelligence. 2017.
[46] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
[47] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern

recognition. 2017: 4700-4708.
[48] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1251-1258.
[49] Abadi M, Barham P, Chen J, et al. Tensorflow: A system for large-scale machine learning[C]//12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16). 2016: 265-283.
[50] Caruana R. Multitask learning[J]. Machine learning, 1997, 28(1): 41-75.
[51] Baxter J. A Bayesian/information theoretic model of learning to learn via multiple task sampling[J]. Machine learning, 1997, 28(1): 7-39.
[52] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[C]//Advances in neural information processing systems. 2015: 91-99.
[53] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[C]//European conference on computer vision. springer, Cham, 2014: 818-833.
[54] Dai J, Li Y, He K, et al. R-fcn: Object detection via region-based fully convolutional networks[C]//Advances in neural information processing systems. 2016: 379-387.
[55] He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
[56] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.
[57] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
[58] Fu J, Zheng H, 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. 2017: 4438-4446.

攻读学位期间主要的研究成果

论文成果

[1] Xie B, **He X**, Li Y. RGB-D static gesture recognition based on convolutional neural network[J]. The Journal of Engineering, 2018, 2018(16): 1515-1520.

[2] Wang Y, **He X**, Li F, et al. 553 A computer-aided diagnosis system for bullous disease based on deep learning[J]. Journal of Investigative Dermatology, 2019, 139(5): S95.

[3] Chen X, Zhao S, Huang K, **He X**, et al. 585 Application of artificial intelligence diagnosis in seborrheic keratosis and basal cell carcinoma in Chinese race[J]. Journal of Investigative Dermatology, 2019, 139(5): S100.

[4] Xie B, **He X**, Zhao S, Li Y, Su J, Zhao X, Kuang Y, Wang Y, Chen X. XiangyaDerm: A Clinical Image Dataset of Asian Race for Skin Disease Aided Diagnosis[C]// Medical Image Computing & Computer-assisted Intervention: Miccai International Conference on Medical Image Computing & Computer-assisted Intervention, 2019, 已投稿.

[5] 谢斌, **何小宇**, 黄伟红, 沈敏学, 李芳芳, 赵爽. 基于卷积神经网络的基底细胞癌和色素痣的临床图像鉴别[J]. 中南大学学报(医学版), 2019, 已投稿.

[6] Wu Z, Zhao S, Peng Y, **He X**, Zhao X, Huang K, Wu X, Fan W, Li F, Chen M, Li J, Huang W, Chen X, Li Y. Studies on different CNN algorithms for face skin disease classification based on clinical images [J]. IEEE Access, 2019, 已接收.

[7] Xie B, Zhao S, Wu Z, Huang K, Shen M, **He X**, Zhao X, Zhu W, Chen M, Su J, Chen X, Li Y. Doctor-level Multi-Classification of Skin Diseases and A dataset of Chinese population[J]. Journal of Medical Systems, 2019, 已投稿.

[8] Zhao S, Xie B, Li Y, Zhao X, Kuang Y, Su J, **He X**, Wu X, Fan W, Huang K, Su J, Peng Y, Huang W, Chen X. Smart Identification of Psoriasis by Images Using Convolutional Neural Networks: A Case Study in China. 2019. 已投稿.

[9] Zhao X, Wu X, Li F, Li Y, Huang W, Huang K, **He X**, Fan W, Wu Z, Chen M, Li J, Luo Z, Su J, Xie B, Zhao S. The application of deep learning in the risk grading of skin tumors for patients using clinical images[J]. Journal of Medical Systems, 2019, 已投稿.

[10] Zhao S, Xie B, Li Y, Zhao X, Kuang Y, Su J, **He X**, Wu X, Fan W, Huang K, Su J, Peng Y, Huang W, Chen X. Smart Identification of Psoriasis by Images Using Convolutional Neural Networks: A Case Study in China. 2019. 已投稿.

专利成果

[1] 谢斌, **何小宇**. 一种基于多维数据的红斑鳞屑类皮肤疾病辅助诊断方法. 中国:201910211342.0[P]，2019-03-20.

主要科研项目

[1] 智慧医疗辅助诊断系统及算法研究，2017.10-现在，项目组主要成员。

[2] 隧道全自动喷浆机器人，2017.10-现在，项目组主要成员。

[3] 无人车底层开放平台，2017.04-2017.11，项目组主要成员。

[4] 华为触觉互联网，2016.10-2017.01，项目组主要成员。

创新创业竞赛

[1] “建行杯”第四届湖南省“互联网+”大学生创新创业大赛一等奖，省级，2018 年 9 月，团队主要成员。

[2] 2018 第四届中国“互联网+”大学生创新创业大赛银奖，国家级，2018 年 10 月，团队主要成员。

硕士期间参加学术交流活动

[1] 2018 年 6 月，第二届亚洲人工智能技术大会（ACAIT2018），重庆。

致谢

时间总是走得匆匆，硕士三年一转眼就要结束。随着毕业日期的日趋临近，毕业论文的完成也随之进入了尾声。从开始进入课题到论文的顺利完成离不开老师、同学、朋友给我热情的帮忙。在这里请接受我诚挚的谢意！感谢这一路走来支持过和帮助过我的老师、同学和好朋友们！

首先最想要感谢我的研究生导师——谢斌老师。我在本科的时候就听闻过谢老师上课的风格引人入胜，十分有魅力，很幸运在硕士阶段能够进入谢门，接受谢老师的指导，做自己喜欢做的研究和项目。谢老师对于学生的指导不仅是学术方面，而是在学生成长的各方各面。记得我有段时间情绪不对感觉自己压力较大，谢老师就马上找到我谈话，及时为我开导，让我不要有压力。谢老师教会我最重要的一点就是一定要对自己答应下来的事情负责，不要因为最后完成不了而推脱责任，一定要尽自己的全力去做好。

感谢智能所的各位老师给与我的关怀与帮助！感谢蔡自兴老师、王勇老师、李仪老师、余伶俐老师、陈白帆老师、邹逸群老师、刘丽珏老师、雷琪老师、唐琎老师、高琰老师、肖晓明老师、郭幡老师及谭平老师。谢谢诸位老师，衷心祝愿老师们工作顺利，身体健康！

感谢尤作师兄，感谢作哥一路以来对我的指导。硕士的三年期间，如果自己个人技术上如果有那么一点提升的话，那肯定离不开作哥亲力亲为的指导。作哥严谨的态度，沉稳的风格，优秀的技术都让我崇拜不已。祝愿作哥未来的发展越来越好！

感谢我的同门宋迪，感谢你研究生三年以来对我的陪伴和照顾。我们一起经历的事情太多太多，之前一年多的时间我们两个人可以说是朝夕相处，形影不离，那段日子虽然很累但是很充实。祝愿你的未来一片光明，发挥你的优势，做自己想做的事情，未来一定如你所愿！

感谢同届的其他同学们，刘方、茹智阳、黄余、孔德成、吴谦、晋周南、余俭、王欣、刘懿、薛涛略、舒雷、周聪、夏旭梅、梁仰辉、周玲君、罗青、向达、余国勤、杨海洋等，我们一起参加智能所的活动一起欢乐，谢谢能遇见你们这群好同学。

感谢我们谢门的其他同学们。感谢带领我一起做项目的喻仲斌和梁照栋师兄，你们给了我兄长一般的关怀，祝愿你们在公司里面越做越好！感谢谢门大师姐李沁，感谢你在学术上对我的指导，将来肯定还有很多东西请你指教。感谢师妹张文婷、师弟林学斌、秦觅、赵新宇、程清、丘文杰、徐勇、彭哲、万思远、孙铭

蔚、吴迪等，是你们的存在让我感受到了集体的温暖和关爱！

最后，衷心感谢我的父母！感谢你们含辛茹苦供我上学，感谢你们二十多年来对我的培养和教育。未来的路，我一定会加倍努力，不辜负你们的期望！

2019 年 5 月 19 日星期日

于中南大学升华楼