



南京大學

研究生毕业论文 (申请硕士学位)

论文题目 基于深度学习的尘肺病医学影像分析与应用

作者姓名 何甲

学科、专业方向 计算机科学与技术

指导教师 商琳 副教授

研究方向 机器学习

2019年5月

学 号 : MG1533014
论文答辩日期 : 2019 年 5 月 20 日
指 导 教 师 : (签字)

Medical Image Analysis and Application of Pneumoconiosis Based on Deep Learning

by

Jia He

Directed by

Associate Professor Shang Lin

Department of Computer Science and Technology
Nanjing University

May 2019

*Submitted in partial fulfilment of the requirements
for the degree of Master in Computer Science and Technology*

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目： 基于深度学习的尘肺病医学影像分析与应用
计算机科学与技术 专业 2016 级硕士生姓名： 何甲
指导教师（姓名、职称）： 商琳 副教授

摘要

尘肺病是一种由于长期吸入生产性粉尘而引起的以肺组织弥漫性纤维化为主的全身性疾病，它目前是我国职业病中发病率第一的疾病。尘肺病的诊断依据主要来自于医生对肺部医学影像的分析判断，而在现代的医疗中，计算机辅助的医学影像分析技术已经被广泛使用。近些年来由于深度学习的重大突破，基于深度学习方法的医学影像分析研究逐渐成为主流。然而尘肺病作为我国高发的职业病，关于它的研究却并不多，还存在着待研究的问题：一是基于人工特征算法的尘肺病影像识别精度较低，二是算法分类结果缺乏可解释性。本文结合尘肺病本身的病理特点，围绕这两个问题展开了关于尘肺病X光片识别的研究与应用。

首先，针对人工特征方法对于尘肺病患病期别分类精度较低的问题，提出了基于深度卷积神经网络的层级分类方法。该方法分为两阶段，首先判断样本是否患病，然后再细分识别为“患病”样本的期别。在该方法的设计中，考虑了两个要点：一是单层级的基分类器要尽可能准确，为了设计好的基分类器，本文对比分析了多种深度卷积神经网络的分类结果，并基于代价敏感方法对选出深度卷积神经网络进行调优；二是所有层级的基分类器联合起来最终分类结果要尽量准确，为了调优最终效果，本文通过对比实验设计了截断点方法。经过实验证明，层级分类方法相较于普通的深度学习端到端式多分类方法提升效果明显，宏查全率由0.7723提升到0.9162，它能够更好的区分尘肺病第3期样本。

再者，针对尘肺病诊断需要可靠解释的问题，本文基于深度学习可解释性技术找到可疑患病区域给出了分类依据，并利用可疑区域的细粒度病理特征提升了分类精度。本文研究了两种特征可视化方法来探究尘肺病高维抽象特征。首先是通过Gradient-weighted Class Activation Mapping (Grad-CAM) 技术，获取到网络特征抽取层的分类权重与高维抽象特征，通过它们找到了每个样本分类过程中可能是结节、弥纤、尘斑等的图像病理区域，并以热力图的形式将这些区域高亮的标记出来，这些高亮区域即分类依据，称为“判别性区域”。再者在判别性区域的发现基础上，通过Selective Convolutional Descriptor Aggregation

(SCDA) 技术，在只有期别的标签情况下，将疑似尘斑等图像病理区域的高维特征从高维卷积特征中筛选出来，这种特征称为“细粒度病理特征”。本文对细粒度病理特征进行了进一步探究：降维可视化分析出了尘肺病医学影像各期别的特征区别，发现尘肺病的3期是瓶颈；使用该特征作为多种学习器的训练数据进行对比，在端到端的方法中尘肺病宏查全率由0.7723 提升到0.8857，层级分类方法中尘肺病宏查全率由0.9162提升到0.9292，经过实验验证了细粒度病理特征的有效性。

最后，依据硕士期间参与的尘肺病识项目，设计和实现了尘肺病识别系统。本系统提供尘肺病分类与可疑患病区域定位的服务，这两个服务的核心算法分别是层级分类方法与可解释性技术中的判别性区域发现方法。该系统能够辅助医疗从业人员更快的判断尘肺病，以及辅助他们二次检查模型的分类结果，从而减少漏诊误诊的情况发生。

关键词：深度学习，医学影像分析，尘肺病，可解释性，尘肺病识别系统

南京大学研究生毕业论文英文摘要首页用纸

THESIS: Medical Image Analysis and Application of Pneumoconiosis Based on Deep Learning

SPECIALIZATION: Computer Science and Technology

POSTGRADUATE: Jia He

MENTOR: Associate Professor Shang Lin

Abstract

Pneumoconiosis is systemic disease of diffuse pulmonary fibrosis caused by long-term inhalation of industrial dust, which currently the disease with the highest incidence rate among occupational diseases in China. The diagnosis of pneumoconiosis is mainly based on the doctor's analysis and judgment of the lung medical image, while computer-aided medical image analysis technology has been widely used in modern medical care. In recent years, due to the great breakthrough of deep learning, medical image analysis based on deep learning is becoming more and more popular. As pneumoconiosis is a high incidence of occupational disease in China, there are still some problems to be studied in the medical image analysis of pneumoconiosis: one is the low accuracy of pneumoconiosis recognition based on artificial feature algorithm, the other is the lack of interpretability of algorithm classification results. To solve these problems, identification of pneumoconiosis is studied in this paper the contributions are as follows:

Firstly, to solve that artificial feature method has low classification accuracy, a hierarchical classification method is proposed for the classification of pneumoconiosis. The hierarchical classification method is divided into two stages: our method first determines whether the sample is sick, and then subdivides the categories of "sick" samples. Two points are considered in the process of method design: one is that the single-level base classifier should be as accurate as possible, the other one is that the final classification result should be as accurate as possible when the all levels are combined. For solving the first problem, we compare the results of various deep convolutional neural network classifications, including the indicators before and after they added cost-sensitive methods, and select the best one. For solving the second problem,

a variety of cut points values are compared with the final classification index results in the experiment. The effectiveness of the hierarchical classification method is proved by our experiments. The experimental results show that the hierarchical classification method is more effective than the deep learning end-to-end multi-classification method. The macro-recall is increased from 0.7723 to 0.9162, and our method distinguish samples of stage 3 better.

Secondly, for solving that diagnosis of pneumoconiosis needs to be explained reliably, we base on the deep learning interpretability technology to find the suspected disease areas giving the classification basis, besides, using the fine-grained features of these areas to improve the classification accuracy. Two feature visualization methods are studied to explore the state of high-dimensional abstract features of pneumoconiosis. First of all, Gradient-weighted Class Activation Mapping (Grad-CAM) method is used to obtain the classification weights and high-dimensional features of the network feature extraction layer. The weights and features are used to find pathological areas that may be nodules, fibrils, dust spots and so on. These areas are highlighted by a heat map, which is the basis for classification, called the “discriminant area”. Then, based on the discovery of discriminative areas, Selective Convolutional Descriptor Aggregation method (SCDA) is used to screen the “fine-grained pathological features” from high dimensional convolution feature in the case of only the label of the period. Fine-grained pathological features refer to high-dimensional features of image pathological regions such as suspected dust spots. We further explore these features, including dimensionality reduction for visualization observation, and train several machine learning classifier with them. The results shows that macro-recall increased from 0.7723 to 0.8857 in the end-to-end method , and macro-recall raise from 0.9162 to 0.9292 in the hierarchical classification method. The effectiveness and rationality of the method is verified by our experiments.

Finally, the pneumoconiosis recognition system is designed and implemented according to the requirements of the pneumoconiosis recognition project I participated in during my master’s degree. The system provides services for classification of pneumoconiosis and location of suspected disease areas, and the core algorithms of these two services are hierarchical classification methods and discriminative region discovery methods in interpretable techniques. The system can assist medical practitioners to diagnose pneumoconiosis faster and reduce the occurrence of missed diagnosis and misdiagnosis by double check.

Keywords: Deep learning, Medical image analysis, Pneumoconiosis, Interpretability, Pneumoconiosis Recognition System

目 录

目录	vi
第一章 绪论	1
1.1 研究背景与意义	1
1.2 研究现状与相关工作	3
1.2.1 深度卷积神经网络	3
1.2.2 基于深度学习的医学影像分析	8
1.2.3 深度学习可解释性研究	12
1.3 待研究问题	15
1.4 本文工作	16
1.4.1 本文工作内容	16
1.4.2 本文组织	17
第二章 基于深度学习的尘肺病识别层级分类方法	19
2.1 概述	19
2.1.1 背景与动机	19
2.1.2 本章工作	19
2.2 相关知识	21
2.2.1 稠密连接网络	21
2.2.2 代价敏感方法	22
2.3 算法介绍	24
2.3.1 层级分类方法	24
2.3.2 基分类器选择与调优	25
2.3.3 截断点选取	26
2.4 实验结果与分析	27
2.4.1 实验数据集	27
2.4.2 实验设置	31

2.4.3 评价指标	31
2.4.4 结果与分析	32
2.5 小结	36
第三章 基于判别性区域发现的尘肺病细粒度病理特征提取	37
3.1 概述	37
3.1.1 背景与动机	37
3.1.2 本章工作	38
3.2 相关知识	38
3.2.1 CAM可解释性技术	38
3.2.2 主成分分析	39
3.3 算法介绍	40
3.3.1 符号描述	40
3.3.2 判别性区域发现	41
3.3.3 细粒度病理特征提取	42
3.4 实验结果与分析	45
3.4.1 实验数据集和实验设置	45
3.4.2 判别性区域发现结果	46
3.4.3 细粒度病理特征提取结果	47
3.5 小结	53
第四章 尘肺病识别系统设计与实现	54
4.1 概述	54
4.1.1 背景与动机	54
4.1.2 本章工作	54
4.2 系统设计	55
4.2.1 需求分析	55
4.2.2 流程分析	58
4.3 系统实现	59
4.3.1 系统架构	59
4.3.2 开发环境	60
4.3.3 关键技术	60

4.3.4	系统展示	62
4.4	小结	62
第五章 总结与展望		64
5.1	工作总结	64
5.2	工作展望	65
参考文献		66
简历与科研成果		72
致谢		73

表 格

2.1	尘肺病数据集各类样本数量	27
2.2	训练集与测试集划分	29
2.3	上采样以及平衡后训练集与测试集划分	30
2.4	尘肺病数据集分类结果	33
2.5	基分类器加入代价敏感分类结果	34
2.6	层级分类方法截断点取值与宏召回率对应表	34
3.1	尘肺病分类结果	52
4.1	开发软硬件环境	60

插 图

1.1 深度学习在医学影像分析领域论文分布 [25]	2
1.2 一层隐藏层的多层感知机	3
1.3 卷积操作	5
1.4 最大池化操作	6
1.5 LeNet模型结构 [23]	6
1.6 深度卷积神经网络模型准确率、运算量、参数量比较 [7]	8
1.7 ChestXray14 数据集样本图示 [44]	11
1.8 导向反向传播过程 [48]	13
1.9 反卷积（左）与反池化（右）	14
2.1 DenseNet示意图 [18]	20
2.2 DenseNet（左）与ResNet（右）与在跨层连接上的区别	20
2.3 层级分类方法流程	24
2.4 基分类器训练流程	26
2.5 各期尘肺病患病图例	28
2.6 数据增广过程	30
2.7 TenCrop过程	30
2.8 部分候选基分类器的混淆矩阵	32
2.9 基分类器加入代价敏感前后对比：加入前（左、中）与加入后（右）	33
2.10 截断点与层级分类方法宏召回率的关联性	35
2.11 端到端方法与层次分类方法最终分类混淆矩阵	35
3.1 全局平均池化（GAP）	38
3.2 尘肺病判别性区域发现流程	41
3.3 判别性特征展示	43
3.4 尘肺病细粒度病理特征提取流程	44
3.5 尘肺病判别性区域发现样例	47

3.6	判别性区域样例: DenseNet121 (上) 与Vgg16 (下)	48
3.7	细粒度病理特征降维可视化结果: PCA (左) 与t-SNE (右)	49
3.8	细粒度病理特征PCA降维成分与方差占比关系	50
3.9	降维细粒度病理特征5折交叉学习曲线	50
3.10	降维细粒度病理特征多种学习器分类结果混淆矩阵	51
4.1	尘肺病识别系统用例图	55
4.2	尘肺病识别系统主处理流程	57
4.3	尘肺病影像识别子流程	58
4.4	尘肺病识别系统架构图	59
4.5	层次分类方法基分类器训练曲线:第一阶段(左)和第二阶段(右)	62
4.6	尘肺病识别系统: 上传界面	63
4.7	尘肺病识别系统: 结果界面	63

第一章 绪论

1.1 研究背景与意义

在现代的医疗诊断当中，计算机辅助诊断（Computer Aided Diagnosis, CAD）已经被广泛使用，CAD中的医学影像能提供大量医学信息，已经成为了医生的重要参考依据。医学影像分析（Medical Image Analysis）是指综合利用数字影像处理、人工智能等计算机信息学科的技术分析医学影像，辅助医生诊断。在医学影像分析中，利用计算机识别算法能够有效的降低医生主观经验，认知能力，疲劳程度等诊断过程中的不稳定因素，同时加快诊断效率。随着我国医疗条件不断完善，医院每天都会产生大量医学影像，医学影像分析可以说是步入大数据时代。同时，主要得益于GPU应用带来的计算力的巨大提升与数据持续增长，深度学习近年来获得了突飞猛进的进步。深度学习实质上是多隐层的神经网络，能够从数据中学习到层次特征，十分适合从影像、文本、语音等高维数据中学习到更加精准的特征，从而提升分类或者预测的准确性。所以在自然语言处理、计算机视觉、语音识别等领域和相应的挑战赛中，它成为了主流方法并且屡次取得了破纪录的好成绩，并且深度学习模型的研究仍然在飞速发展当中。

医学影像分析领域积极拥抱深度学习这种技术。两者的结合，从医学影像数据中利用深度学习挖掘出有用知识，已经成为了学术界和工业界的研究热点。目前，经过了前几年相关研究的爆发式增长，其几乎已经渗透到医学影像分析的各个领域，包括了医学影像配准 [46]、影像增强 [6]、分类 [3, 29]、检测 [30, 43]、组织器官与病灶分割 [20, 49] 等。

目前已经有较多的研究集中在使用深度学习对肺部疾病的医学影像分析上，但是大部分工作主要是有结节或病灶的疾病的分类和检测。各大广泛有影响力的竞赛平台上涌现了大量优秀的深度学习肺结节的检测算法，如LUNA16（Lung Nodule Analysis 2016）肺结节检测竞赛中的算法 [37]，Kaggle平台上的Data Science Bowl 2017 肺癌检测比赛中的公布了前两名的算法，以及天池平台医疗AI肺部结节智能诊断比赛中的优秀算法。同时各大科技企业也纷纷推出了各自优秀的肺癌辅助诊断方案，如腾讯优图在2017年9月推出的觅影系统，对早期的肺癌和良性肺结节准确分别达到达到85%和84%以上；阿里巴巴2017年7月推出的Doctor You系统肺结节识别准确率到达90%以上。肺部疾病辅助治疗方法大都集中于肺结节/病灶检测的主要原因是该疾病的数据量满足

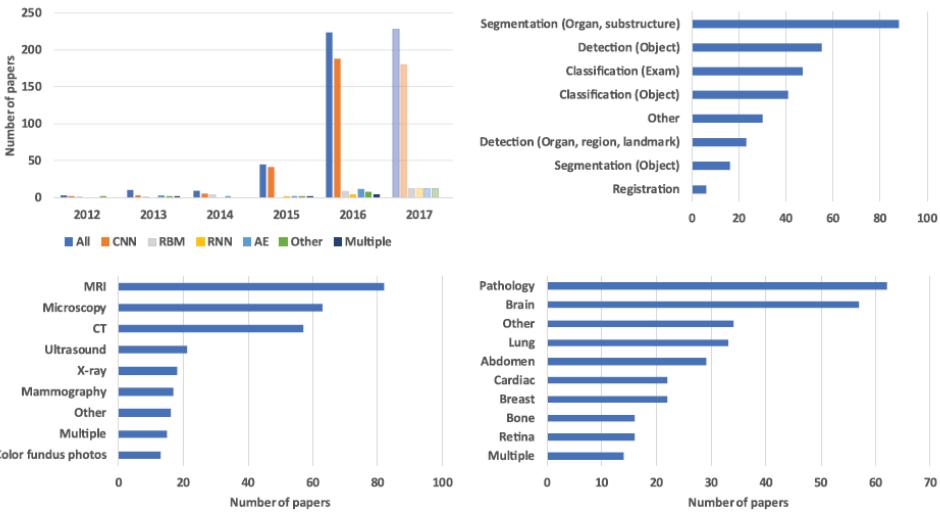


图 1.1: 深度学习在医学影像分析领域论文分布 [25]

深度学习的需求：一是已经有了公开的规模较大、标注质量高的肺结节数据集，通常是X光片或CT影像；二是因为肺结节疾病的特性，影像上肺结节作为样本可以被单独重复采样，经过专门技术处理使得数据集中每一张医学影像可获得几十上百的样本，极大的丰富训练数据。

而关于尘肺病的研究目前却非常少。尘肺病又名矽肺病，是我国确诊的职业病发病率排名第一的病症，根据国家卫计委的统计数据比例高达90%。它是由于在职业活动中长期吸入生产性粉尘（灰尘），并在肺内潴留而引起的以肺组织弥漫性纤维化（瘢痕）为主的全身性疾病。煤矿工人、粉尘作业工人、油漆工人、清洁工等都是尘肺病的高发患病人群，尘肺病会引起肺结核、肺炎、肺源性心脏病衰竭等并发症，患病后目前无法根治的方法，治疗手段以及时预防、及早发现延缓病情进展避免并发症为主。对于尘肺病因为涉及病人敏感信息或其他原因目前尚无公开数据集，同时因为自身病理特性每一张影像只能作为一个样本，所以目前使用深度学习方法进行尘肺病分类和检测研究较少。

尘肺病作为无法根治的慢性疾病“早发现早治疗”就显得尤为重要。但是目前我国存在问题是医疗资源地域分布不均匀，高水平的医生和高质量的医疗设备大都集中一二线城市、三甲医院，县城村镇医疗与之相比还有着巨大的差距，往往这种疾病必须到较高级别的医院才能及时发现。高危患病人群能够定期体检，并且有专业的医生和设备是发现病症的关键，但高发患病人群通常没有这样的经济条件和医疗条件。成熟而有效的CAD系统具有快速、廉价的特点。研究智能的尘肺病识别算法，对高患病人群而言，以低廉的代价定期体检并有效诊断显得意义重大。同时我国医患比严重失衡，各大医院每天产生大量

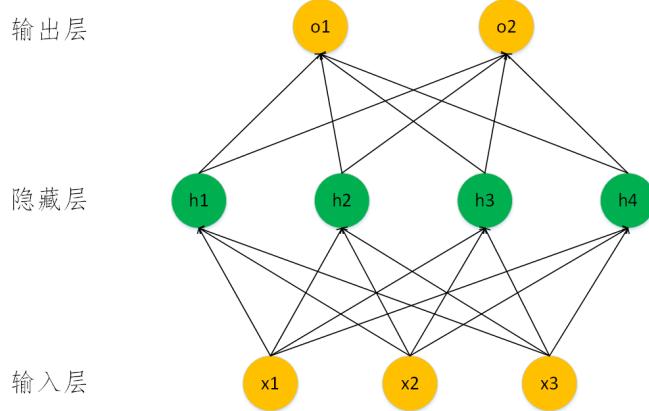


图 1.2: 一层隐藏层的多层感知机

的X光片或CT需要处理与分析而高素质的专业医生却较少，一位放射科或临床医生每天都要阅片数百张，工作压力巨大。研究智能的尘肺病识别算法，对医院医生而言，能够辅助医生对医学影像进行初步的诊断和筛选，加快医生检查效率降低医生疲劳带来的主观误差。因此研究基于深度学习的尘肺病X光片识别算法，有着广阔的应用空间和积极的意义。

1.2 研究现状与相关工作

1.2.1 深度卷积神经网络

机器学习是能从训练数据中学习，发掘出知识和类似的模式，然后对新的类似数据做出决策和预测，它在日常生活中各个领域内机器学习技术都已经得到了广泛运用。但是大部分机器学习任务有一个重要的步骤是特征工程，需要依赖人的对于该任务的认知通过先验知识从数据中构建一些特征，然后将构建的特征作为算法模型的样本去训练模型，这就使得机器学习的准确度很大程度时候依赖于技术人员的认知水平。深度学习作为机器学习的一个子领域，是能够自动的从原始数据中学习并抽取特征的算法模型。

深度学习实质上是多层感知机（Multi-Layer Perceptron, MLP），它拥有多层级的特征学习能力。具体来说，原始数据先经过前馈运算，通过层层神经网络的转化到最后与标签计算loss，结合优化算法和BP算法（Back Propagation, BP）将loss回传到层层网络并更新参数来达到学习的目的，MLP能够得到非常复杂的表达函数，理论上也可以逼近任何复杂任务抽象出来的数学表达函数。

MLP如图1.2示，给定一个小批量样本 $\mathbf{X} \in \mathbb{R}^{n \times d}$ ，其批量大小为 n ，输入个数为 d ，假设多层感知机只有一个隐藏层，其中隐藏单元个数为 h ，记隐藏层的

输出（也称为隐藏层变量或隐藏变量）为 \mathbf{H} ，有 $\mathbf{H} \in \mathbb{R}^{n \times h}$ 。隐藏层和输出层均是全连接层，可以设隐藏层的权重参数和偏差参数分别为 $\mathbf{W}_h \in \mathbb{R}^{d \times h}$ 和 $\mathbf{b}_h \in \mathbb{R}^{1 \times h}$ ，输出层的权重和偏差参数分别为 $\mathbf{W}_o \in \mathbb{R}^{h \times q}$ 和 $\mathbf{b}_o \in \mathbb{R}^{1 \times q}$ 。

含单隐藏层的多层感知机的计算公式如式1.1和1.2：

$$\mathbf{H} = \phi(\mathbf{X}\mathbf{W}_h + \mathbf{b}_h) \quad (1.1)$$

$$\mathbf{O} = \mathbf{H}\mathbf{W}_o + \mathbf{b}_o \quad (1.2)$$

其中，输出 $\mathbf{O} \in \mathbb{R}^{n \times q}$ ， ϕ 表示激活函数，激活函数（Activation Function）是非线性函数，用于非线性变换，常用的激活函数包括ReLU函数、sigmoid函数和tanh函数。在分类问题中，通常对输出 \mathbf{O} 做softmax运算，并使用softmax回归中的交叉熵损失函数；在回归问题中，对输出层的输出个数设为1，并将输出 \mathbf{O} 直接提供给线性回归中使用的平方损失函数。

事实证明，对比结构化数据深度学习特别适合从图像、文本、语音等高维数据中学习到更加精准的特征来提升模型识别准确度。在不同的领域研究者都提出了不同的优秀的深度学习模型，如计算机视觉领域常用的卷积神经网络（Convolutional Neural Networks, CNN），自然语言领域常用的循环神经网络（Recurrent Neural Network, RNN），无监督学习领域常用的自编码器（auto-encoder）。基于深度学习方法的医学影像分析常用的网络结构是卷积神经网络，下面主要是介绍卷积神经网络的相关知识。

• 卷积神经网络常见概念

卷积神经网络是深度学习网络结构中的一种，被广泛的用在计算机视觉相关的任务，如图像分类、物体检测、图像语意分割等。卷积神经网络相较于多层感知机具有更好的局部感知能力、更小的计算复杂度、更抽象更鲁棒的抽象能力。它主要是一种有着大量的卷积层的神经网络，除此以外还有池化层，全连接层。

1) 卷积层

卷积层是对输入的数据进行卷积操作，它作用是使得网络有着局部感知和权值共享的能力，使得网络能够学到图像的局部空间特征信息的同时减少了模型参数，降低计算量和减少过拟合风险。具体来说如图1.3所示，对输入数据 \mathbf{X} 使 K 个卷积核 $\mathcal{W} = \{\mathbf{W}_1; \mathbf{W}_2; \dots; \mathbf{W}_k\}$ 和偏置 $\mathcal{B} = \{\mathbf{B}_1; \mathbf{B}_2; \dots; \mathbf{B}_k\}$ 进行卷积，每个卷积核和偏置都可以得到一个新的特征映射 \mathbf{X}_k 。新的特征映射再经过激活函数，即一个非线性变换 $\sigma(\cdot)$ ，形象的说明如图1.3所示，卷积核会以指定

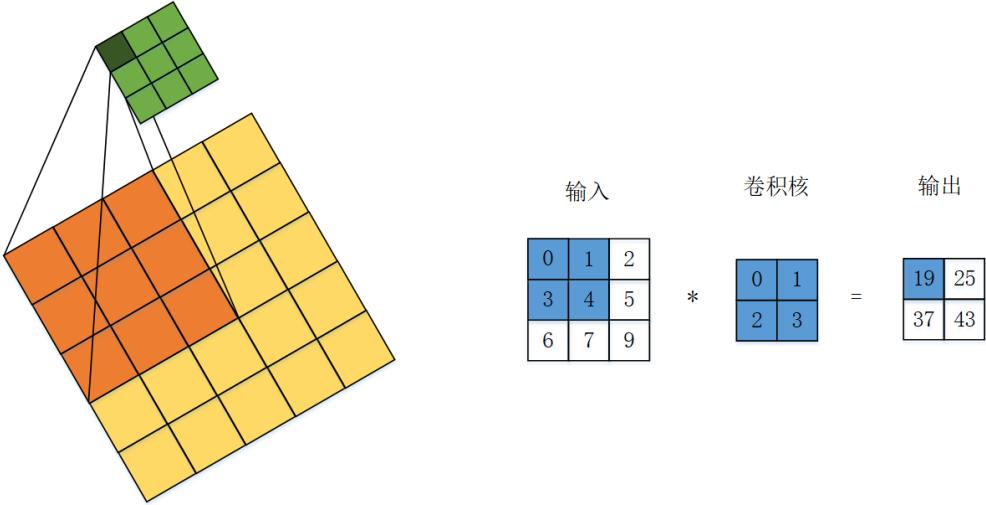


图 1.3: 卷积操作

的步幅在整个输入 \mathbf{X} 从左到右, 从上到下的顺序滑动并进行上述的互相关性计算, 并且每一层卷积层都重复这样的操作。卷积计算公式如公式1.3:

$$X_k^l = \sigma(W_k^{l-1} * X^{l-1} + b_k^{l-1}) \quad (1.3)$$

2) 池化层

同卷积层一样, 池化层每次对输入数据的一个固定尺寸 (池化窗口) 中的元素计算输出, 它的作用是使得网络具有了一定的空间不变性, 大量的减少了中间生成的特征的参数量, 减低过拟合风险和极大降低计算量。具体来说, 不同于卷积层里计算输入数据 \mathbf{X} 和卷积核 \mathbf{X} 的互相关性, 池化窗口也以从左到右、从上到下的顺序依次在输入数组上划过, 并每次计算窗口内所有元素的最大值或者平均值, 分别称为最大池化 (Max Pooling) 和平均池化 (Mean Pooling), 具体操作如图1.4示。

3) 全连接层

全连接层原理上与两层的感知机相同, 计算公式如式1.4和1.5, 与MLP公式(式1.1和式1.2)相比没有了激活函数 ϕ 。它的作用是把前一层所有神经元与下一层的所有神经元相连接, 做一个线性变换以期望汇总获得图片经过卷积后的全局信息。它一般为CNN的最后一层, 再之后经过损失函数的计算 $loss$, $loss$ 依具体任务而设定, 在分类任务中通常为交叉熵。

$$\mathbf{H} = \mathbf{X}\mathbf{W}_h + \mathbf{b}_h \quad (1.4)$$

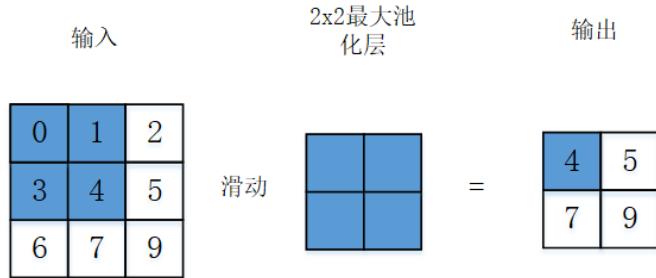


图 1.4: 最大池化操作

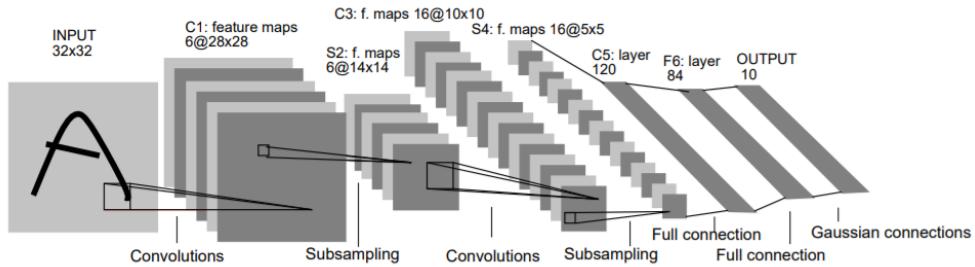


图 1.5: LeNet模型结构 [23]

$$O = HW_o + b_o, \quad (1.5)$$

- 卷积神经网络发展过程

图像分类是计算机视觉要解决的主要问题之一，尘肺病X光片识别分类亦属于此任务范畴。近些年来图像分类任务在深度卷积神经网络助力下取得了突破性的进展，优秀的深度学习模型大量出现。用于图像分类任务的CNN在这些年的发展主要沿着两个方向发展：首先是网络朝着更深、更宽的方向发展，不断追求着有着更好的特征表示能力，更高识别精度的网络模型；再者是网络朝着更灵活、更轻便的方向发展，不断的追求着有更小的内存消耗，更低的计算复杂度，更有效的训练方式的网络模型。

早在1989年就出现了CNN，LeCun提出LeNet [23]，LeNet结构如图1.5示，在当时就已经出现了卷积层，池化层，全连接层等CNN的基本网络结构。2012年随着AlexNet [22]在当年的ImageNet ILSVRC图像分类竞赛获得冠军，沉寂多年深度神经网络获得历史突破。AlexNet能够获得成功，不仅归功于使用GPU加速了网络训练以及ImageNet提供了大规模的标注图像数据，还归功于创新的网络结构——包括新的激活函数ReLU和解决过拟合策略Dropout等新技术的提出。

2014年提出了Vgg [38]网络再次刷新比赛成绩，它大体上与AlexNet有着相似的结构，但为了获取更充分图像特征，它采用了 3×3 的卷积核替代了 5×5 与 7×7 的卷积核，使用了较小卷积核来实现更深的网络。同时期，2013年底Lin等提出Network in Network（NIN）[24]结构，它主要是使用有更高的抽象，更强泛化能力的非线性滤波器Mlpconv代替传统卷积层；并且在最后分类层使用全局均值池化层取代全连接层，大量的减少模型参数并降低过拟合风险。在NIN的基础上，Google提出了Inception Block结构，将 1×1 , 3×3 , 5×5 的conv和 3×3 的pooling级联堆叠在一起，一方面增加了网络的宽度另一方面增加了网络对尺度的适应性；还使用了 1×1 的卷积来减少特征映射的维度和增加模型的非线性表达能力。将多个Inception Block串联起来搭建了深度网络GoogLeNet [40]。

2015年，Inception Block结构持续改进，将 5×5 卷积核拆成两层 3×3 的卷积核，并加入 $1 \times n$ 和 $n \times 1$ 的卷积核进一步降低参数。同时，He等发现，网络的深度增加到一定的层数分类精度不升反降出现退化，为此提出残差学习和残差网络ResNet [17]，该模型中存在跳转连接Shortcut，简单而有效地解决了非常深的网络梯度消失（或梯度弥散）的问题，成功训练了深达152层的网络并取得ImageNet ILSVRC 2015比赛多项冠军，至此深度神经网络的错误率已经低至3.57%，超过了人眼错误率。在此之后，2017年提出的密集连接卷积神经网络DenseNet [18]在Shortcut基础上，加入了稠密块和过渡层，网络任何两层之间都有直接的连接，加强了网络的特征传播和特征重用能力，进一步降低了错误率，在此之后ImageNet ILSVRC大赛已达到促进计算机视觉发展的目的，就此停办。

上述神经网络的发展过程中，显示出越深越宽通常拥有着更强的特征表达能力更高的识别精度。同时带来的问题是网络参数变得很多，训练数据集有限容易过拟合；网络越大计算复杂度越大，难以应用；网络越深梯度方向传播越往后越容易出现梯度消失（或梯度弥散），难以训练优化。

所以有大量工作又关注于更小的内存消耗，更低的计算复杂度的网络模型。Google将GoogLeNet改进到Inception V4 [39]，该网络将Inception Block结构与ResNet结合，使用了Inception-residual net结构。实验结果表明，Inception V4结构的网络性能与ResNet相当情况下训练时间只需要一半，并且训练过程收敛更快 [39, 41]。Zagoruyko等通过增加宽度的方法提高性能，他们提出了宽度残差网络 [47]（WRNs），在参数数量相当的情况下，WRNs比深度ResNet好训练，实验表明16层WRNs网络在精度和效率上达到、甚至优于上千层的ResNet网络。DenseNet加入了稠密块和过渡层，能大幅度降低网络参数数量，可以更有效地

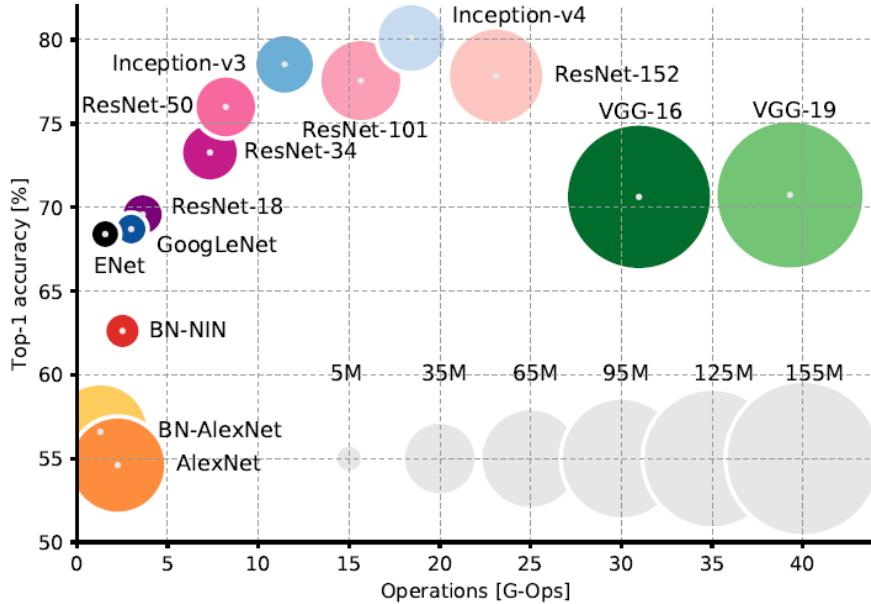


图 1.6: 深度卷积神经网络模型准确率、运算量、参数量比较 [7]

训练。Chen等提出了一种结合ResNet 和DenseNet 优势的新型卷积网络结构DPN (Dual Path Networks)，该网络通过残差旁支通路再利用特征和密集连接通路探索新特征 [8]。之后还有Iandola 等则致力于简化网络结构和减少网络参数，提出了网络结SqueezeNet [19]，该网络仅需只AlexNet1/50的参数量即可达到相同的精度。深度卷积神经网络发展到现在，模型结构逐渐朝着更加轻量化、更低运算量与更深更宽、更强的特征表达能力两个方向相互融合，部分经典网络模型比较如图1.6所示。

1.2.2 基于深度学习的医学影像分析

- 医学影像分析概述

医学成像设备已经在各级医院被广泛的使用，产生了各种医学影像。医学影像主要包括超声影像、X光片、计算机断层扫描（Computed Tomography, CT）、核磁共振（Magnetic Resonance Imaging, MRI）、数字血管剪影（Digital Subtraction Angiography, DSA）等。医学影像分析相关的研究与计算机视觉、数字影像处理相关领域有着紧密的联系，计算机视觉中的三大研究方向：分类、检测、分割研究中使用的方法经常被借鉴在医学影像处理上。类似的，主要的医学影像分析任务分为医学影像分类与识别、医学影像定位与检测、医学影像分割任务。

医学影像分割任务：影像分割是识别影像中感兴趣的目标区域（如肿瘤）内部体素及其外轮廓，它是临床手术影像导航和影像引导肿瘤放疗的关键任务。复杂的医学影像分析任务常常需要综合进行分类、检测与分割，例如为了进行诊断乳房X射线影像中病灶的良/恶性，要经过病灶检测、病灶分割、病灶分类。由病理学影像分析判断癌症严重程度时，需要首先检测、分割细胞核，然后基于分割结果进行特征和统计分析，最后分类得到分级结果。

医学影像分类任务：临床医生常需要借助医学影像来辅助诊断人体内是否有病灶，并对病灶的轻重程度进行量化分级，因此自动识别影像中的病灶区域和正常组织器官，对目标医学影像能够正确的进行分类，是医学影像分析的基本任务。

医学影像检测任务：人体组织器官解剖结构和病灶区域不仅能够被正确的识别，确定它们各自在图中的正确位置同样是医疗检查过程中非常重要的一步，定位的精度直接影响治疗的效果。影像目标检测任务则是不仅能够识别影像中的特定目标，并且还要确定任务目标的具体物理位置。

- 基于深度学习的医学影像分析

深度学习，尤其是卷积神经网络卷积神经网络，已经在医学影像分析中有着广泛的运用。

在医学影像分割任务中，此任务本质上是像素级别的分类。一般的流程分为数据处理、感兴趣区域（Region of Interest, ROI）提取、神经网络分割、分割结果后处理。Long等 [27]提出了全卷积神经网络（FCN），首次以端到端的形式实现了从输入到输出的图像像素级分类，保证对任意尺寸的图像都能进行处理并获得与原始输入图像大小一致的分割结果，但基于FCN的方法 [16, 42, 51]的研究，其结果还是比较不精确。Ronneberger等 [34]提出了U-Net结构更加适用于医学图像分割，并且针对数据量较小的医学影像数据集进行了专门的处理，相比于FCN它能取得更加精准的分割结果。U-Net的良好性能获得了医学影像分割研究者的青睐，并在此基础上提出多种改进。Drozdzal等 [13]引入了类似ResNet的shortcut 短连接结构，Poudel等 [31]将类似U-Net的网络结构与循环网络单元GRU结合，Alom等 [1]将U-Net、残差网络结构与RNN结构结合。这些最近的技术都取得了不错的效果。

在医学影像分类与识别任务中，主要包含了两类情况：一种是直接对输入影像进行疾病识别与分类，这种情况常常直接借用的深度卷积神经网络的模型包括了Vgg、ResNet、DenseNet等就能取得较好的结果，但与数据治理有着比较大的关系；另一种是对于体积较小不易察觉的病灶识别，深度学习方法很难

捕直接捕捉到病灶关键位置，因此外加其他的步骤。一种是通过分割手段，将病灶所在大体的局部区域提取出来，再对局部可疑区域进行识别，这种方法通常用到具有明确解剖位置的病灶上，如肝癌。这种方法通常会在病灶分类前，先训练一个检测模型，检测模型多借鉴了YOLO [32]、SSD [26]、Faster RCNN [33]。第二种是借助深度学习中的注意力机制（Attention），让模型自动捕捉到关键区域同时降低无关区域影响，在肺炎、糖尿病视网膜病变等应用中，这种方法可以学习到关键区域主要覆盖在病灶及相关位置附近，因此可以作为弱监督下病灶定位方法。一般情况下，通过以上方法，若在当前影像检测到了病灶，则该影像就能被判定为对应的疾病。后续还有研究者从特征区分能力、定位精度、效率等方面进行了改进。Arevalo [4] 等提出了乳腺癌诊断的特征学习框架，采用CNN自动学习区分性特征，对乳房X线照片病变分类。Kooi [21] 等比较了传统CAD中手动设计和自动CNN特征提取方法，两者都在约4.5万乳房X线影像的大数据集上训练，结果显示CNN在低灵敏度下优于传统CAD系统方法，且在高灵敏度下两者相当。Gao [15] 等探讨了影响CT脑影像分类任务的深度学习技术应用的重要性，特别是利用CNN，旨在为AD早期诊断提供补充信息。

医学影像分析中深度学习技术的广泛运用，源于其在计算机视觉领域的重大突破，这个突破很大部分归功于ImageNet [11]的出现，它是一个拥有超过1500万张影像，约2.2万类物体的带标签的大规模标注数据集。数据驱动的深度学习在具体疾病的分析任务中常常遇到数据集带来的以下问题：

首先尽管医院放射科积累大量影像数据，一是由于医学影像数据的隐私性很多二是具体到某一专业性更高的细分疾病领域，能获取的相关疾病的医学影像往往不多；再者医学影像标注代价高昂是更大的难点，因为深度学习目前主流方法还是有监督学习的方法，需要领域内专家来标注且标记大量数据时间代价高昂；最后，医学影像中疾病各个类别样本数量往往高度的不平衡，这是任务本身属性所决定的，医学检查中一种疾病的正常人的样本是远多于患者的样本，患病较轻的样本是多于患病较重的样本的。

相关疾病的数据集好坏是该疾病能否用医学影像分析取得好效果的关键，本文研究的是尘肺病的识别，接下来介绍关于肺部疾病的数据集的一些工作。

肺部疾病医学影像分析数据集主要有两种，一种是X光片，一种是CT。X光片辐射量小费用低能够显示特定的病理结构，但有投影部位重叠分辨率较低等缺点。CT是更先进，它能扫描肺部多个截面的状况提供更多的空间诊断信息，对隐藏的和较小的病变组织都有着更高的灵敏度，同时带来的假阳性率也会更高。LUNA16（Lung Nodule Analysis 16）是2016年推出的一个肺结节检测数据集，旨在作为评估各种肺结节CAD的benchmark，是目前世界上肺肿

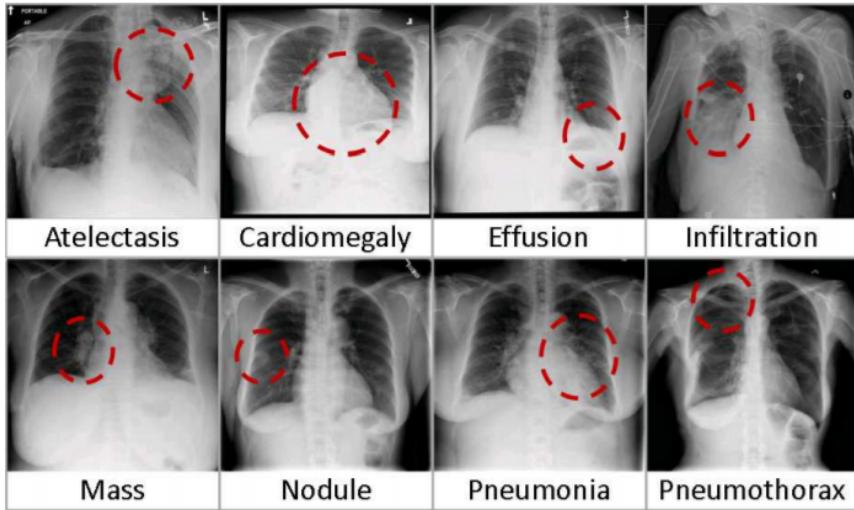


图 1.7: ChestX-ray14 数据集样本图示 [44]

瘤检测最常用最大的数据集，它包含了888个CT影像，1186个结节。LUNA16取自LIDC/IDRI [5]数据集，对比于LIDC/IDRI数据集，它的肺结节经过了三个以上放射科医生意见一致的标注，筛掉了直径小于3mm的结节，结节大小的范围更理想、影像质量更好、标注质量更高。胸部X光也有相应的数据集，2017年美国国立卫生研究院（NIH）公布了目前世界上最大规模的肺部X光数据集ChestX-ray14 [44]，包含了3万名患者的112,120张正面视图的X光片。该数据集包含了14种肺部疾病的标注（肺不张、变实、浸润、气胸、水肿、肺气肿、纤维变性、积液、肺炎、胸膜增厚、心脏肥大、结节、肿块和疝气），每张影像可以有多个疾病标签，标注准确率超过90%。除此之外，在各大数据挖掘竞赛平台上也有质量较高的肺部疾病数据集，如Kaggle Data Science Bowl 2017肺癌检测大赛的CT影像数据集，2017阿里天池医疗AI肺癌智能诊断比赛的CT影像数据集。

由于只有这些公布的数据集，目前使用深度学习方法辅助肺部疾病诊断主要集中在肺结节的识别，而关于尘肺病X光片的医学影像分析研究比较少。尽管病症的类别不同，但是这些研究的数据样本都是X光片或CT影像并使用了深度学习方法，与基于深度学习的尘肺病X光片识别任务有类似之处，对本文工作仍具有极大的借鉴意义。本小节将从肺结节识别将前人的工作进行总结。

各类肺部疾病如肺癌、结核、真菌感染等，都会导致肺结节的产生，因此准确的识别肺结节是判断各种肺部疾病重要依据。各类研究者从多种深度学习技术手段对肺结节进行识别。肺结节有多种类型，不同类型肺结节需要被正确识别，Gao 等 [14]由单张CT平片判断肺部的磨玻璃影、网状影、蜂巢影以

及肺气肿等间质性病变, 以及Anthimopoulos 等 [2]将CT 平片分割为多个局部子片, 并分别使用深度网络进行蜂巢及网状影等的病理检测。Ciompi 等 [10] 提出了一种多通道多尺度CNN 对不同类型的肺结节进行识别, 包括了分叶状结节、钙化结节、叶间裂周围结节、实性结节、非实性结节、部分实性结节, 它的算法平均识别精度为69.6%而该研究中的两名医生的平均识别准确度为72.9%, 诊断结果接近于医生水平, 并且研究显示出适度增大CNN滑窗尺度有利于分辨不同类型的肺结节和血管断面。多通道多尺度CNN是一种重要的手段, [36]提出了基于二维视角的CNN来解决抑制假阳性肺结节的问题, 该网络有着两个输入通道的并行结构, 每个输入通道对应输入候选结节影像块的一个视角截面, 这种并行的结构使得二维CNN能抽取更丰富的空间信息。在LUNA16的数据集上, 该方法能在平均每个CT容忍1 个假阳性结节的情况下达到89.2%的召回率。Dou等人 [12]又提出了三维CNN来解决假阳性肺结节抑制的问题, 该研究利用了三个输入尺度不同的三维CNN来抽取肺结节多尺度的纹理特征。对比 [36]三维CNN能更好地抽取CT影像中的三维空间信息, 该方法在LUNA16数据集上平均每个CT容忍1个假阳性结节的情况下取得了97.2%的召回率。由于数据量缺乏, 以及各类别数据样本经常不均衡, 迁移学习的思路也经常在各种研究中使用Christodoulidis 等人 [9]尝试借助多源迁移学习方法训练CNN, 将各数据源训练的不同CNN 的权重参数汇成一个权重参数集, 再训练集合各权重参数的CNN, 使CNN 分类肺间质性病变的平均准确率提升2%。

1.2.3 深度学习可解释性研究

深度学习是一个黑箱, 这主要是因为它的高度非线性性质, 每个神经元都是由上一层的线性组合再加上一个非线性函数的得到, 无法直观的像理解线性回归 (Linear Regression) 那样通过模型参数来确定哪些维度的特征比较重要。深度卷积神经网络的可解释性相关的研究主要是从“反向”的思路进行的, 即从输出特征通过各种CNN前向算法的逆过程逐层回溯到输入空间然后可视化, 下面进行一些简单的介绍。

CNN的可解释性一种技术是基于梯度回传的思想, Zeiler [48]等的研究就是这种思路, 他提出反卷积 (deconvolution) 与导向反向传播 (guided backpropagation), 该方法可以将网络学习到的特征或者神经网络的隐层信息“反向”的映射到输入空间, 从而转化成人类可以理解的有实际含义的图像。

CNN输出的特征图中数值有大有小, 其中数值较大的特征可以认为是很大程度上激活网络神经元, 可以认为这些数值较大的高维特征所对应的原始图像的区域是影响分类结果的区域, 找到这种对应关系即可给出每次分类一些解释。

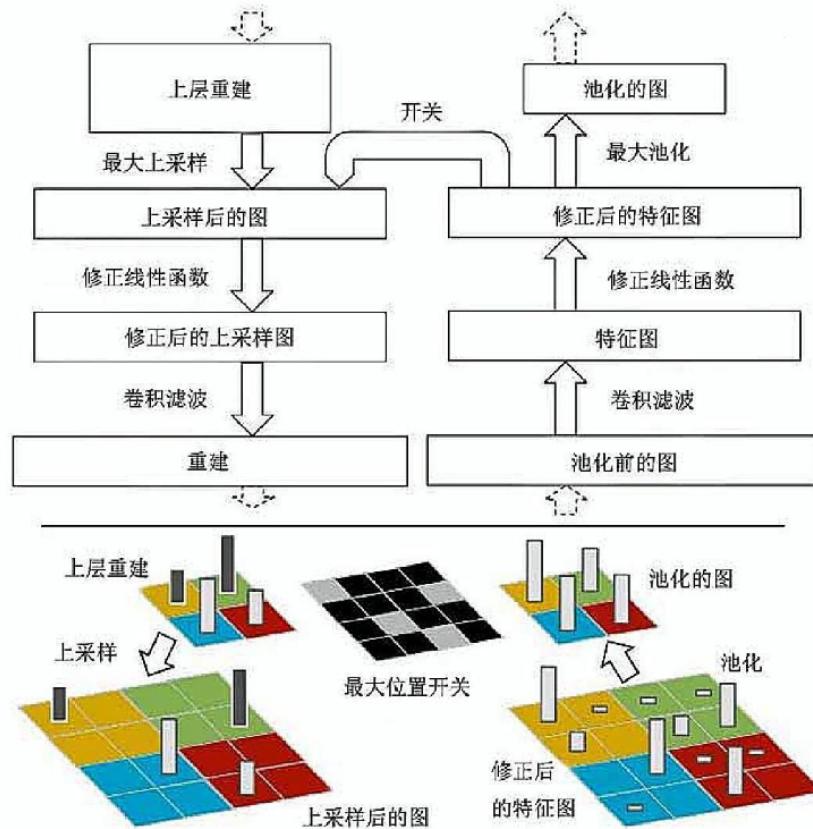


图 1.8: 导向反向传播过程 [48]

CNN一个典型的完整的正向卷积过程是由卷积—激活—池化三个步骤组成的，而这个“反向”的特征空间映射到输入空间主要通过反池化-反激活-反卷积的逆过程来实现，整个导向反向传播过程如图1.8示。

1) 反池化

反池化是池化的逆操作，因为池化的过程就只保留了主要信息，舍去部分信息，所以反池化无法通过池化的结果还原出全部的原始数据。如果想从池化后的这些主要信息恢复出全部信息，则存在信息缺失，反池化通过补位的方式实现最大程度的信息完整。池化分为最大池化和平均池化，其反池化也需要与其对应。

平均池化和反平均池化：反平均池化不需要记录特征图中各值得位置，它将池化结果中的每个值（平均池化每个窗口内的均值）都填入其对应原始数据区域中相应位置即可。

最大池化和反最大池化：反最大池化需要记录特征图中最大激活值的坐标位置，它只把池化过程中最大激活值所在位置坐标值激活，其他的值设置为0。

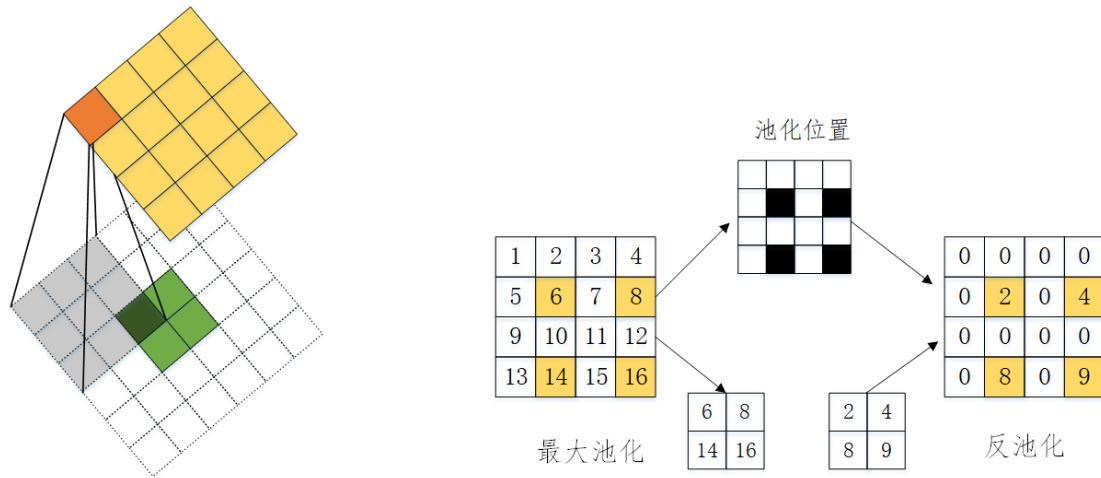


图 1.9: 反卷积 (左) 与反池化 (右)

这个过程也是一种近似，因为在池化过程中，除了最大值的位置，其他的值也是不为0，具体操作如图1.9右图示。

2) 反激活

激活函数的取反过程。在CNN中一般使用Relu作为激活函数，而反激活的值和实际的激活值没有任何区别，即只保留正数，其余值为0即可。

3) 反卷积

反卷积真实的含义应该是转置卷积 (transposed convolution)，CNN卷积过程实际上将一些共用参数组合成了一个滤波器的形式，都可以转变成一个矩阵乘法的操作，只不过对是一个参数重复度很高的稀疏矩阵。将这个稀疏矩阵表示为 C ，则后一层 X^{l+1} 和前一层 X^l 的关系就可以表示为式1.6，反卷积的操作就是要对这个矩阵运算过程进行逆运算，即通过和 C 得到 X^{l+1} ，能得到反卷积的计算的过程为式1.7。

$$CX^l = X^{l+1} \quad (1.6)$$

$$X^l = C^T X^{l+1} \quad (1.7)$$

C^T 往往可以表示为卷积层对输入层的梯度，也就是说通过对卷积层进行适当的补0操作，再用原始的卷积核转置之后的卷积核进行卷积操作，就可以得到相应的梯度矩阵与当前卷积层的乘积。可以看出其实卷积层的前向传播过程就是

反卷积层的反向传播过程，卷积层的反向传播过程就是反卷积层的前向传播过程。反卷积过程如图1.9左图示。

导向反向传播是基于反卷积的改进技术，两者的基础都是反向传播（BP），三者唯一的区别在于反向传播过程中经过Relu层时对梯度的不同处理策略：BP算法是前向时特征图输入Relu时大于零的值的所对应的梯度传递回去；deconvolution是反向时梯度值大于零的梯度传递回去；guided backpropagation既要前向的特征图输入大于零，又要反向梯度值大于零的对应的梯度回传回去。

guided backpropagation屏蔽掉了很多对当前层的激活值没有实际作用激活值将其归为0，通过反卷积操作就得到了仅对当前层有实际贡献的输入层的数值——将其作为该层的特征表示，表现效果更好。使用BP得到的图像噪声较多，无法辨认出有意义的图像结构。deconvolution可以大概看清楚物体的轮廓，但是有大量噪声在物体以外的位置上。guided backpropagation基本上没有噪声，特征很明显的集中在物体部位上。

1.3 待研究问题

尽管基于深度学习的医疗影像分析已经取得了很大进步，但是具体到尘肺病医学影像分析还有以下问题：

第一，基于人工特征方法尘肺病识别精度较低。众多研究表明在医学影像分析领域，深度学习方法能够取得比人工特征方法有更好的精度，目前深度学习方法在医学影像分析逐渐成为主流，基于深度学习方法的尘肺病应该能取得更高的精度但是相关研究却较少。算法研究的两个基础是：数据与算力，基于深度学习的尘肺病方法研究较少的一个主要原因是数据集的缺失。首先，尘肺病影像总量少，虽然各大医院有着各种医学影像，但是具体到该种疾病上图片数量往往不多。而且医学影像由于涉及病人隐私有些不便公开，能够获得数据就更少，这种数据量通常难以满足数据驱动的深度学习的需求。再者，尘肺病影像各期别数量相差悬殊，通常患病样本与健康样本比例相差极为悬殊，或者不同期别的类别数量差异巨大。这是由于疾病本身原因所决定的，因为尘肺病检查中健康人的数据是远多于患者的数据，较严重的期别是远少于较轻期别的数量，数据的类别不平衡会导致算法模型不够鲁棒，精度不够高。最后，数据标签质量参差不齐且标记成本高昂，尘肺病影像识别目前主要还是基于监督学习的方法，识别算法的精度与标签的质量高低有着极大的联系。医院由于历史条件数据整理保存不当或者其他原因，很多积累的数据并没有相应的说明等各原因，数据集标签往往有误差，但是由于是专业性的影像，算法设计者没有领域知识进行修正。一个样本影像通常需要多名领域内的专家再次确认标签是

否有误和重新标记，代价高昂，而且数据量对专家而言较多时间代价太大。

第二，尘肺病识别结果缺乏解释依据。医学诊断结果需要严谨的解释，基于深度学习的医学影像分析中常用的获得解释性的方法是：将一张样本影像划分成许多尺寸相同的子图（它们原图上可以重叠也可以不重叠），然后将子图像输入模型，依据子图像的模型评分确定该子图像患病程度，将所有子图还原到原始样本影像的位置，就可以得到整个样本影像所有区域的评分，评分越高的区域，就越有可能是导致样本最终分类结果的区域，即以一种“正向”的思路，获得样本被分类的原因。但是尘肺病的医学判定标准是，“依据对全肺区小阴影密集度的判定的综合结果”，即一张尘肺病影像只能作为一个样本，随意分割可能会破坏标签属性。由于深度学习模型是个黑箱模型，以及尘肺病本身的病理特性，基于深度学习方法进行尘肺病图像分析时可以得到较高精度但却缺乏解释：为什么模型将某个样本判别为某一类。

由于上述的问题，基于深度学习的尘肺病医学影像分析的相关研究缺乏。以肺部疾病为例，目前大多数研究都集中在肺部结节的识别与检测中，而类似于尘肺病，气胸等其他疾病的研究却少有。因此，结合尘肺病数据集本身样本分布与疾病本身的病理特性等特点，具体任务具体分析，研究能智能识别尘肺病的有效的算法模型与应用，不仅能够填补相关领域的一些空白，同时对于缓解医疗资源短缺让更多的人享受更好的医疗有着积极意义。

1.4 本文工作

本文基于深度学习方法，围绕尘肺病医学影像分析中人工特征方法识别算法精度较低与算法分类结果缺乏可解释性的问题，尘肺病X光片识别的研究与应用开展了以下工作：

1.4.1 本文工作内容

- 针对人工特征方法对于尘肺病患病期别分类精度较低的问题，提出了基于深度卷积神经网络的层级分类方法。层级分类方法分为两阶段，首先判断样本是否患病，如果超过第一阶段模型的截断点则认为患病，否则健康；然后再细分认为“患病”样本的类别。其设计出发点是，在诊断中医生对于患病期别的重视程度不同，即健康与否的漏诊代价比患病期别的误诊代价大得多。层级分类方法考虑了两个要点：一是单层级的基分类器要尽可能准确，为了选择好的基分类器，实验中研究分析了多种深度卷积神经网络的分类结果，包括它们加入代价敏感方法调优前后的差别；二是所有层级的基分类器联合起来最终分类结果要尽量准确，为了调优最终效果，在

实验中进行了多种截断点取值与最终分类指标结果的对比，选择了最好的截断点。同时在训练基分类器的过程中，针对数据不平衡的问题在数据层面使用了多种扩充方法与模型层面使用了代价敏感方法对基分类器进行调优。层级分类方法通过先“粗筛”再“细筛”的方式有效的提升了识别精度。

- 针对尘肺病诊断需要可靠解释的问题，本文基于深度学习可解释性技术找到可疑患病区域给出了分类依据，并利用这些区域的高维特征提升了分类精度。深度学习是一个黑箱模型，本文研究了两种特征可视化方法来探究尘肺病高维抽象特征的状态。首先是通过Grad-CAM技术，获取到网络特征抽取层的分类权重与高维抽象特征，通过它们找到了每个样本分类过程中可能是结节、弥纤、尘斑等的图像病理区域，并以热力图的方法将这些区域高亮的标记出来，这些高亮区域即分类依据，称为“判别性区域”。判别性区域发现是尘肺病医学影像分类之外的一个很好的补充，以利于医生在模型诊断过程中进行二次检查确保更少的漏诊误诊出现。再者，在判别性区域的发现基础上，通过SCDA技术，在只有期别的标签情况下，将判别性特征中超过一定阈值的部分作为最有可能是病理症状而非噪声的高维特征筛选出来，这种特征称为“细粒度病理特征”。本文对细粒度病理特征进行了进一步探究：降维可视化分析出了尘肺病医学影像各期别的特征区别；使用该特征作为多种学习器的训练数据进行进一步分类对比，验证了筛选出的细粒度病理特征的有效性和合理性。
- 依据硕士期间参与的课题组项目的需求，设计和实现了尘肺病识别系统。该系统的核心算法是层级分类方法与可解释性技术中的判别性区域发现方法，系统基于B/S 架构，使用的技术是包括了Flask, React, Pytorch等的Python 技术栈的流行的技术，实现了用户登录，影像上传，影像分类，结果获取等功能。本系统提供尘肺病分类与可疑患病区域定位的服务，期望能够辅助医疗从业人员更快的判断尘肺病以及能够二次检查减少漏诊误诊的情况发生。

1.4.2 本文组织

围绕本文工作，将以五个章节展开介绍，结构如下：

第一章为绪论。包括了本文的研究背景与研究意义、本文涉及的技术背景与相关研究、相关研究仍存在的问题、本文主要工作与本文组织结构。

第二章介绍了尘肺病层级分类方法。首先介绍了层级分类方法提出的动机

与针对的问题，再是层级分类算法涉及的具体技术细节，接着介绍了层级分类方法的整体流程与层级分类方法如此设计的出发点，最后介绍的是实验部分，包括了实验数据集的划分与处理，实验参数设置，实验结果对比与结论，证明了层级分类方法的有效性。

第三章介绍了尘肺病分类的可解释性工作。首先介绍了可解释性工作在尘肺病分类中的必要性，再是可解释性技术的涉及的具体技术细节，接着依次介绍了可解释性技术的具体构成：“判别性区域发现”与“细粒度病理特征提取”的具体实验流程，最后分别介绍了两者的实验结果，包括实验参数设置，实验结果对比与结论，证明了尘肺病可解释性技术的有效性。

第四章介绍了尘肺病识别系统。首先介绍了项目需求分析，再介绍了系统整体流程分析与核心算法流程分析，最后介绍了系统实现过程，包括了整体架构设计，涉及的工程技术，最终结果展现。

第五章对本文进行了总结，并对以后尘肺病识别相关工作做出了展望。

第二章 基于深度学习的尘肺病识别层级分类方法

2.1 概述

2.1.1 背景与动机

在绪论中已提到深度学习方法在图像分类任务上已经取得了很好的识别准确度，它们使用的数据集是诸如ImageNet图像主体多为日常物体的通用图像数据集，其中各类别样本分布均衡、数量较为充足。它们所面临的场景是各类物体对于模型而言重要性是一样的，基于这种场景，通过缩减人工预处理和后续处理，尽可能使模型从原始输入到最终输出“端到端”的思想已经被广泛应用。

然而基于深度学习的医学影像分类并不能简单的将计算机视觉中图像分类方法套用过来。具体到尘肺病分类而言，在实际医疗过程中有一个很重要的区别是，不同类别的样本对于一个模型的重要性是不一样的。即医生对是否患病和患病到了第几期的关注程度是不同的，患病者的漏诊误诊的代价比患病期别的误诊代价要大的多。因此有必要在尘肺病多分类的算法流程中，对于这种情况进行区分。

除此之外，还有诸多问题：相比于常用的通用数据集上万级别的数量，本尘肺病数据集样本数量少，约2500张胸部X光片；图片质量参差不齐，X光片出片时间跨度超过20年，由于医疗条件的变迁有些影像清晰有些非常模糊；类间样本数量差距巨大，尘肺病分为四期，各类数据大致比例2:20:2:1；还有尘肺病本身的病理特性，如判定标准主要是依据全肺区的阴影密集度的综合情况决定，制约了X光片的预处理手段。由于这些问题的存在，不同的细分领域的疾病的医学影像分类任务需要具体任务具体分析，结合数据集本身的样本分布与疾病本身的标签特性等特性设计合适的识别算法。尘肺病作为我国职业病患病率第一的疾病，而目前关于使用深度学习的尘肺病识别研究又很少，围绕解决尘肺病数据存在的这些问题，丰富相关研究正是本章工作的原因。

2.1.2 本章工作

围绕基于深度学习的尘肺病X光片识别的目标，为了解决上节提到的问题，本章具体工作如下：

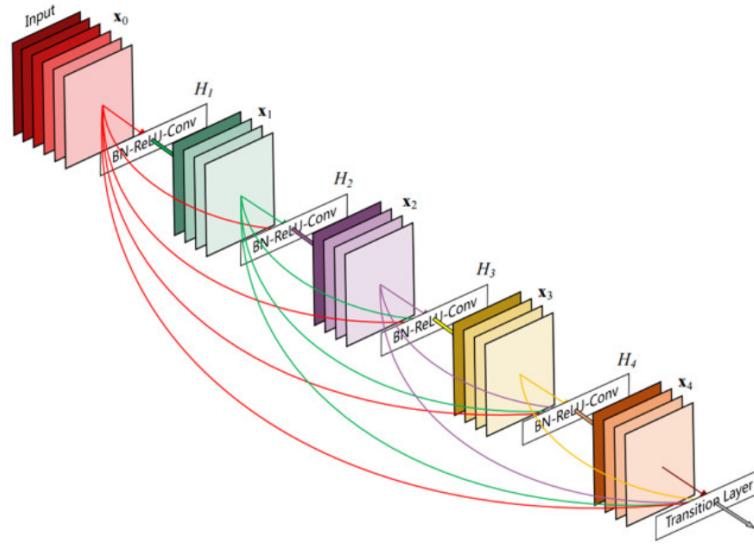


图 2.1: DenseNet示意图 [18]

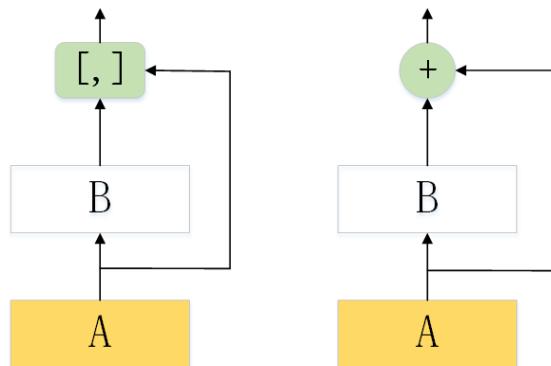


图 2.2: DenseNet（左）与ResNet（右）与在跨层连接上的区别

- 结合尘肺病医疗流程，设计了基于深度学习的层级方法，即首先筛选出样本患病与否，再细分患病期别的流程。
- 层级分类方法中，对于单阶段的基分类器选择，对比了多种深度卷积神经网络效果与加入代价敏感方法前后效果，选出了最优的单阶段基分类器；对于所有阶段最终分类精度优化，设计了截断点，对比多组截断点选择出最合适的值提升了最终分类精度。

2.2 相关知识

2.2.1 稠密连接网络

本章在对比实验中使用了诸如Vgg [38], ResNet [17], DenseNet [18]等优秀的图像分类的深度卷积神经网络。DenseNet作为实验中主要使用卷积神经网络，本小节进行一些简单的介绍。

DenseNet是一种具有密集连接的卷积神经网络。在该网络中，任何两层之间都有直接的连接，即网络每一层的输入都是前面所有层输出的并集，而该层所学习的特征图（feature map）也会被直接传给其后面所有层作为输入。图2.1是DenseNet的示意图，可以看见它的主要结构包括：稠密的跨层连接（shortcut connections）、稠密块（dense block）与过渡层（transition layer）。

1) 跨层连接

DenseNet的优点之一是网络中的后面的层都能直接与其前面的任意一层相连，实现特征的重复利用，这得益于稠密的跨层连接结构。跨层连接最初在ResNet的残差结构中提出，主要作用是解决网络添加过多的层后训练误差往不降反升的现象，即CNN中越深的网络梯度弥散越严重的问题。残差结构数学表达如式2.1式2.2，其中 x_l 和 x_{l+1} 分别表示的是第 l 个残差单元的输入和输出，而每个残差块一般包含多层结构。 F 是残差函数，表示学习到的残差， $h(x_l) = x_l$ 表示恒等映射， f 是激活函数，一般为Relu 函数。

$$y_l = h(x_l) + F(x_l, W_l) \quad (2.1)$$

$$x_{l+1} = f(y_l) \quad (2.2)$$

这种跨层连接对后续的CNN产生了深刻的影响，DenseNet在此基础上进行改进。相比ResNet的跨层连接，主要区别如图2.2示，其中A和B表示前后相邻的抽象的运算模块：DenseNet模块B的输出与模块A通过跨层连接在通道上连接，直接将A的输出和B的输出一起传入后面的层；而ResNet是模块B与模块A是通过跨层连接进行了加法运算再输入后面的层。另外一点区别是，ResNet的跨层连接只连接相邻的两个残差块，是一个递进的层级结构；DenseNet网络中的每一层都直接与其后面的层相连，如图2.1示，这使得网络中的后面的层都能直接与其前面的任意一层相连，实现特征的重复利用，这是称为“稠密连接”的原因。

DenseNet通过稠密的跨层连接，公式由式2.3 变为2.4，其中 $H_l(\cdot)$ 表示第 l 层的非线性变换。

$$x_l = H_l(x_{l-1}) + x_{l-1} \quad (2.3)$$

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (2.4)$$

2) 稠密块与过渡块

DenseNet的另一个优点是网络学习非常少的特征图即可达到非常高的识别精度，参数量和计算量都显著减少达到了降低CNN冗余性的目的。这归功于它的主要构建模块—稠密块和过度块。稠密块定义了输入和输出是如何连结的，而过度块用来控制通道数，使之不过大。

DenseNet因为大量使用了稠密连接，这样显然增加了参数量，网络参数显得比较冗余。DenseNet通过将每一层都设计的特别窄很好的解决了这个问题。首先是设计了稠密块，稠密块中每一个单元实际上都是一个瓶颈层（bottleneck layer），它包括了一个1x1卷积核（conv）和一个3x3卷积核，bottleneck layer可以改变特征图的通道数，这样既能降维减少计算量，又能融合各个通道的特征。其次在稠密块之间使用了过度块，他包括了一个批归一化层（batch normalization），一个1x1的卷积核和一个池化层（pooling），这种结构同样起到了降低冗余的作用，使得一个稠密块输出的特征图，能够被一个[0,1]之间的参数来限制输出的数量。

因为以上的模型设计，DenseNet有着省参数，省计算，抗过拟合的优点。DenseNet达到与ResNet相当的精度所需的计算量与参数量大约都只有ResNet的一半左右，这是非常好的优点，深度学习模型的轻量化和高效化是实际应用中非常强的需求。同时DenseNet可以综合利用浅层复杂度低的特征，能够更容易得到一个光滑的具有更好泛化性能的决策函数，所以有着很好的抗过拟合性能。在原论文中已证明，DenseNet 在不做数据增广（data augmentation）的CIFAR数据集上比以前的网络在同样的情况下的错误率都更低。这种表现尤其适合于训练数据相对匮乏的应用，适合于尘肺病这种医学影像分类问题。

2.2.2 代价敏感方法

代价敏感方法主要出发点是，在分类中当不同的分类错误会导致不同的惩罚力度时如何训练学习器。如在金融信用卡盗用检测中，“将盗用误认为正常使用”的代价”与“正常使用误认为盗用的代价”不同；在医学健康检查中“将患病样本误诊为健康的代价”与“将健康样本误诊为患病样本的代价”不同。为权衡不同类型错误所造成的不同损失，将错误赋予非均等代价，一般的学习

算法通过最小化分类损失使分类错误率最小化, 而代价敏感学习则以最小化分类代价为目标。

在医学影像分类任务中数据不平衡的情况下, 样本数目较少的类别会欠学习, 我么可以通过增加小样本错分的惩罚代价并将此惩罚代价直接体现在目标函数中, 这样通过优化目标函数就可以调整模型更加关注到小样本的数据。

代价敏感方法通常有两种, 一种是基于代价敏感矩阵的代价敏感方法, 另一种是基于代价敏感向量的代价敏感方法。

$$C = \begin{bmatrix} C(1,1) & C(1,2) & \dots & C(1,K) \\ C(2,1) & C(2,2) & \dots & C(2,K) \\ \vdots & \vdots & \ddots & \vdots \\ C(K,1) & C(K,2) & \dots & C(K,K) \end{bmatrix} \quad (2.5)$$

- **基于代价敏感矩阵的方法:** 代价敏感矩阵反映了类别级别的错分惩罚。例如在分类问题中, 训练集 $\{x_n, y_n\}_{n=1}^N$ 共 N 个样本, 其中样本标记 y 属于 K 类。不同的代价被表示成为一个 $K \times K$ 的代价矩阵 C 中, 基于代价敏感矩阵方法就是利用矩阵 C 对不同样本类别施加错分惩罚, 如式2.5所示, $C(y_i, y_i) \in [0, \infty)$ 表示类 y_i 错分为 y_j 类的代价, 其取值不小于0, 且 $C(y_i, y_i) = 0$ 。可以看到, 式2.5中的代价敏感矩阵反映了类别级别的错分惩罚, 施加了代价惩罚以后训练目标变成训练得到某分类器 g 使得期望代价之和 $\sum_{n=1}^N C(x_n, g(x_n))$ 最小
- **基于代价敏感向量的方法:** 代价敏感向量反映了在样本级别错分惩罚。例如对某样本 (x_n, y_n) 有对应的一个 K 维代价敏感向量 $c_n \in [0, \infty)^K$, 其中 c_n 的第 k 维表示该样本被错分为第 k 类的惩罚, 其第 k 维应为0。基于代价敏感向量的方法在模型训练阶段是将样本级别的代价敏感向量与样本以 (x_n, y_n, c_n) 三元组形式一同作为输入数据送入学习算法。不难发现, 代价敏感矩阵法实际上是代价敏感向量法的一种特殊形式, 即对于某类的所有样本其错分惩罚向量为同一向量。

代价敏感方法处理不平衡样本问题的前提是需事先指定代价敏感矩阵或向量, 其中关键是错分惩罚或错分权重的设定。在具体操作中可以根据样本比例、分类结果的混淆矩阵等信息指定代价敏感矩阵或向量中错分权重的具体取值。

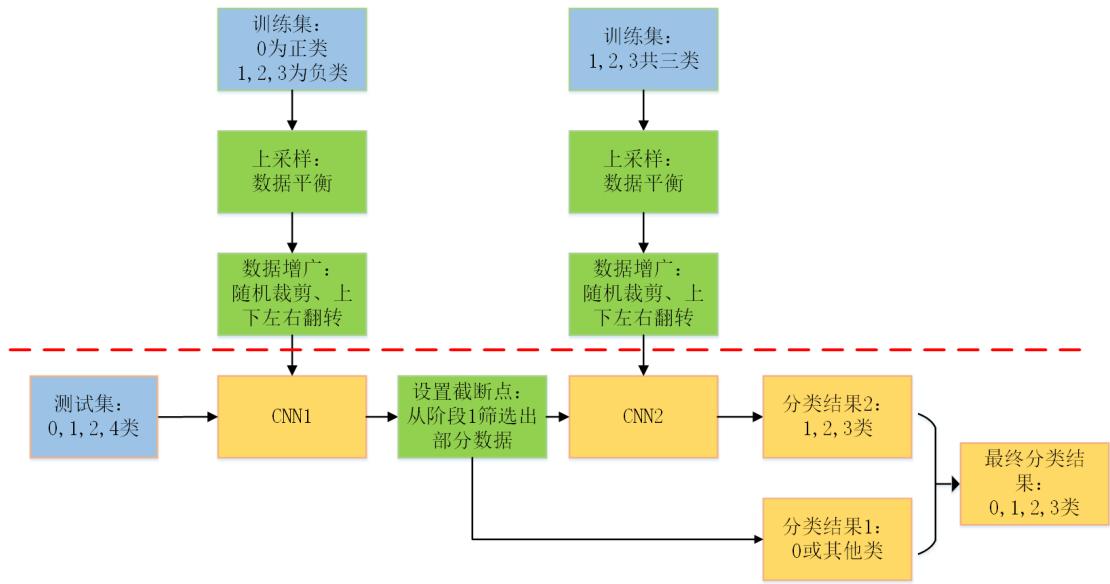


图 2.3: 层级分类方法流程

2.3 算法介绍

在本章中，结合尘肺病诊断流程中对于是否患病以及患病具体期别的关注程度不同，设计了分为两阶段的层级分类方法，即首先判断样本是否患病，再判断患病到第几期的流程。并且在层级分类过程中，针对不同层级的深度学习模型加入了代价敏感方法进行调优，同时在两个层级之间加入了截断点筛选数据进行联合调优最终分类精度。

2.3.1 层级分类方法

将深度学习中常用的从所用数据输入到输出结果中途没有人工干预的方法，在本章中即一次性的4分类深度学习模型，称为“端到端”方法；将识别过程分为两阶段，中间每个模型都经过人工调优的方式称为“层级分类方法”。相比基于深度学习的多分类任务中仅用一个深度神经网络的端到端的方法，层级分类方法的主要考量因素是医生对于患病/健康和患病到哪种程度的敏感性不一样：对于患病病人的漏诊比患病期别的误诊代价要大的多。具体而言，基于基分类器，首先训练一个2分类器用于区分患病（1, 2, 3类，作为正样本）和健康（0类，作为负样本）样本，然后再训练一个3分类器用于区分不同患病程度的样本（1, 2, 3类），层级分类方法流程如图2.3示，算法如2.1示。

层级方法中，要使的算法的最终分类指标比较高，主要涉及两个问题：1)

算法 2.1 层级分类方法

输入: 尘肺病数据集 $D = \{(x_k, y_k)\}_{k=1}^n$; 二分类模型 F_1 , 三分类模型 F_2 ; 阈值 t ;

输出: 尘肺病数据集预测结果 $\{(x_k, \hat{y}_k)\}_{k=1}^n$

```

1: for all  $(x_k, y_k) \in D$  do
2:   计算当前样本通过模型  $F_1$  的输出评分  $s_1$ ;
3:   if  $s_1 \geq t$  then
4:     计算当前样本通过模型  $F_2$  输出结果  $\hat{y}$ ;
5:   else
6:     由评分  $s_1$  计算当前样本通过模型  $F_1$  输出结果  $\hat{y}$ ;
7:   end if
8: end for

```

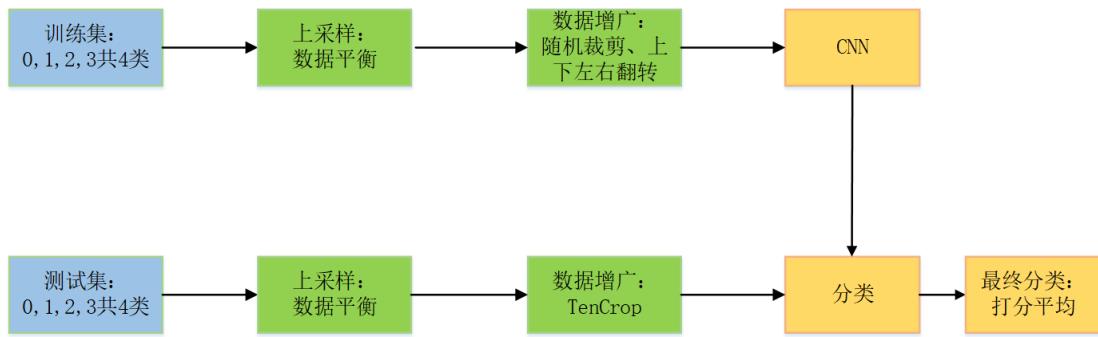
对单层级而言，每个流程基分类器精度要尽量的高，实验中测试了多种经典深度卷积神经网络并选择了其中最好的作为基分类器，并且针对数据类别不平衡的情况使用了代价敏感的方法对模型进行了优化；2) 对整个流程而言，需要联系起来通盘考虑，实验中观察了不同的截断点值对最终四分类指标的影响，进行了多组实验并选取了合适的截断点。

2.3.2 基分类器选择与调优

基分类器选择阶段，即使用端到端的方法直接训练出4分类模型，选择其中最优的。实验所使用了AlexNet, Vgg16, ResNet34, DenseNet121共四种经典的深度卷积神经网络，并且使用了深度学习中常用的数据增广方法（data augmentation）：训练阶段使用的是随机裁剪，上下左右翻转；在测试阶段使用的是TenCrop方法。流程如图2.4示

已知本文使用的尘肺病数据类别极度不平衡，可以使用代价敏感方法进一步调优模型的表现。已知样本空间 X , 类别空间 Y , 给定样本 (x, y) , $x \in X$, $y \in Y$ 代价敏感损失函数 $L_C(y, F(x))$ 度量 F 对 (x, y) 的预测损失，其中 C 是代价矩阵见式2.5。学习算法优化整个样本空间的代价敏感损失 $E_{X,Y}[L_C(y, F(x))]$ (代价敏感风险)，求取代价敏感决策函数 F_C^* 如式2.6。给定 x , 式2.6 可转化为优化条件代价敏感风险 $E_Y[L_C(y, F(x))|x]$, 即采用 F 对 x 类别进行预测产生的风险，如式2.7示。

$$F_C^* = \arg \min_F E_{X,Y} [L_C(y, F(x))] \quad (2.6)$$



$$F_C^* = \arg \min_F E_Y [L_C (y, F(x)) | x] \quad (2.7)$$

具体方法是，通过对比实验选出在尘肺病数据集上指标最好的基分类器，再用分别用数据集1（1, 2, 3类作为正和0类作为负）和数据集2（1, 2, 3类）再去训练基分类器，这样分别获得到第一阶段与第二阶段基分类器的分类混淆矩阵 M_1 与 M_2 ，基于 M_1 与 M_2 设定各自代价矩阵 C_1 和 C_2 ，再依据式2.6用代价敏感学习方法重新训练基分类器，这样最终得到层级分类算法中最优的基分类器。

2.3.3 截断点选取

设置好最优的基分类器后，层级分类方法中在第一阶段和第二阶段之间有设置置信截断点。在使用神经网络作为学习器的预测过程中，模型会把每个样本属于每个类的打分给出，这些打分是属于[0.0, 1.0]之间的实值并且和为1，实际上这些打分是CNN的softmax层输出的样本属于每个类的概率。将这个概率值与一个阈值比较，若大于阈值则分为正类，否则为反类。这个阈值即截断点（cut point），易见截断点的选取能够影响学习器的泛化能力。实际上，算法可以根据输出的概率值将测试样本排序，最可能是正例的样本排在最前面，最不可能是正例的样本排在最后面。分类的决策过程就相当于在这个排序中选取截断点，选取位置不同，最终的分类决策结果不同。如果结果更重视查准率，截断点位置越靠前，如果结果更重视查全率，截断点位置越靠后。

层次分类方法第一阶段是个二分类器（判断患病与否），在不设置截断点是会取两者中分数较大的所代表的类别作为分类结果，即默认0.5为阈值，选择超过0.5的打分所代表的类别作为预测结果。设置截断点以后，只有预测打分超

表 2.1: 尘肺病数据集各类样本数量

诊断结果	标签	数量
健康	0	316
尘肺病第一期	1	2533
尘肺病第二期	2	389
尘肺病第三期	3	79

过这个阈值的样本才被认为是正类，接着这部分被认为是“正（患病）”数据被送入第二阶段的分类器进行细分类，即先判断是否是尘肺病，再判断是第几期尘肺病。这是结合实际尘肺病检查的，健康/患病的误诊代价比患病到第几期的误诊代价更大，所以第一阶段目的是降低患病样本被漏诊几率，第二阶段目的是降低样本患病类别被误诊几率。

具体方法是，实验在(0.0, 1.0)之间每隔0.05选取一个值作为截断点，共19组值。当样本在第一阶段打分超过截断点后，则认为是正类否则是负类(0类)，接着再将认为是正类的数据送入阶段2给出分类结果，这样联合计算最后的分类指标进行对比选取最合适的截断点。

2.4 实验结果与分析

本节首先介绍实验中所使用的数据集与实验设置，包括了原始数据集划分，针对尘肺病数据集存在问题进行数据平衡后的情况，预处理中的数据增广方法；然后介绍网络训练中的参数与设置；最后汇报层级分类算法的实验结果，包括了与端到端4分类算法对比，添加敏感权重前后指标对比，添加截断点前后指标对比，通过实验验证了层级分类算法的有效性。

2.4.1 实验数据集

实验尘肺病X光片数据集由课题组项目提供，出于隐私保护的原因并未公开。该数据集样本共分为4类，标签分别是“健康”，“尘肺病第一期”，“尘肺病第二期”，“尘肺病第三期”，标签分别用0, 1, 2, 3类表示，其中1, 2, 3类表示的患病程度依次变严重，数据集数量具体如表2.1示及各类影像样本如图2.5示。可以看到，0, 1, 2, 3类样本分布不均，尤其是3类，只有几十张与其他类别的数量差距很大。并且患病越严重的样本，X光片中肺部成聚集状的阴影就越明显和越多。

数据集预处理

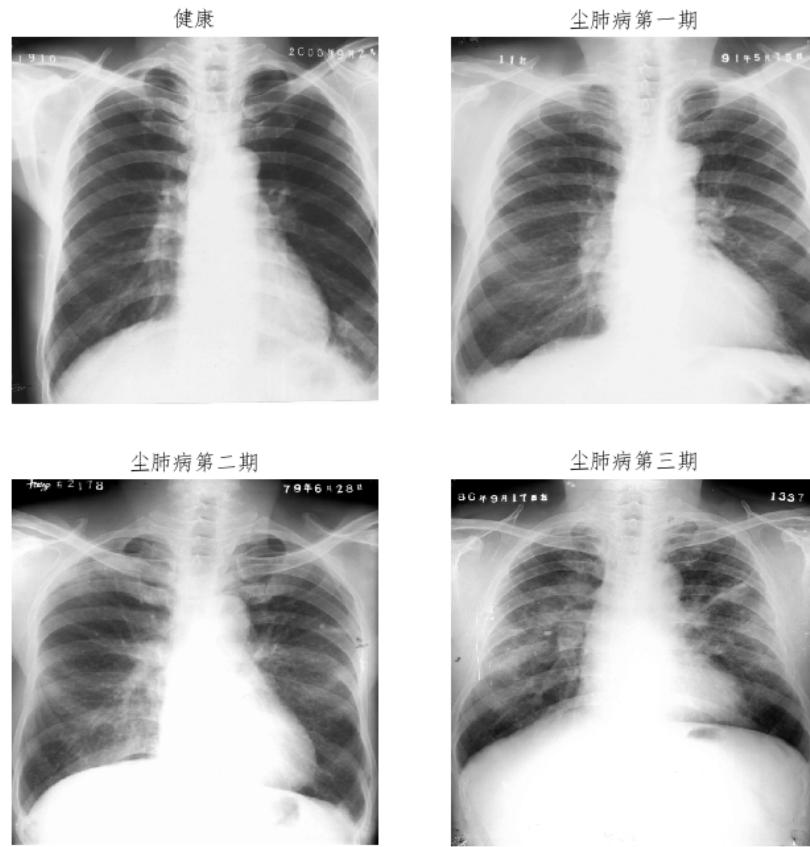


图 2.5: 各期尘肺病患病图例

原始的数据集不论是从样本数量上还是各类样本比例上都不能满足深度学习的需求，除此之外，这批数据还有一些处理上的难点，具体如下：

- **影像标签具有整体性：**根据医学知识，在X光片上，尘肺病判定标准主要是依据对全肺区小阴影密集度的一个综合结果，即X光片的标签是基于整体肺区阴影综合结果给出的。一张X光片在后续的数据预处理步骤中也应该作为一整张图片输入，如果随意分割可能会破坏标签属性是，例如一张标签为车的图片主体为车但包含天空，若数据预处理过程中将图片分割了，天空的部分若标签仍然是车，是不合理的。组织器官，对目标医学影像能够正确的进行分类，是医学影像分析的基本任务。
- **影像噪声大：**如上一点所述，尘肺病判定标准主要是依据对全肺区的阴影密集度。但是，如图2.5示，X光片上胸腔骨骼占据了较大比例，而肺区的气管，阴影等画面信息相对较少。根据先验知识，模型分类依据应该聚焦于肺区，而依据第一点图片又不可以随意分割必须以整体信息被模型学

表 2.2: 训练集与测试集划分

数据集	基分类器	层次分类方法	
		第一阶段	第二阶段
	0,1,2,3期 共四类	1,2,3 期(正) 0期(负)两类	1,2,3期 共三类
训练集	2843	2843	2572
测试集	474	474	429
总计	3317	3317	3001
类间比例	4:25:4:1	15:2	25:4:1

习，可见对于分类问题，这种数据引入了较多噪声。

针对数据集存在的这些问题，结合整个算法流程，简单的介绍针对数据集的划分与预处理的过程：

1) 数据集划分

本小节将原始数据集重新按需求组成三个数据集，一个数据集为端到端方法使用，两个数据集为层级分类方法使用。每个小数据集采取不放回抽样的方法得到训练集与测试集，保证训练集与测试集的独立同分布，训练集与测试集的比例约为6:1数据集划分集划分后样本数量如表2.2示。

2) 影像压缩与数据平衡

原始数据集每张X光图片文件大小约5M，数据较大不仅不适合于深度神经网络入口的规格，并且会在模型的后续计算中带来过大的计算量。本文中将每张图片先进行压缩到300k 左右，并通过resize操作将每张影像变为 600×600 像素。

同时，数据集4类大致比例是4:25:4:1，3类数据非常少。类间数据不平衡对CNN有着较大的影响，而且数据量太小会导致模型过拟合。本文中对较少的种类，采取上采样的策略，被采样的图片每一次都是随机的裁取0.8的部分并resize到 600×600 像素作为一个新样本。采样后四类间的比例约为16:25:16:10，具体数据如表2.3示：

3) 数据增广

尘肺病数据集在上采样后数据仍然较少，需要用到数据增广。数据增广是深度学习中常用的技巧之一，主要针对一个样本做出各种变换，让数据集尽可能的多样化，间接扩大数据集使得训练的模型具有更强的泛化能力。

表 2.3: 上采样以及平衡后训练集与测试集划分

数据集	基分类器 0,1,2,3期 共四类	层次分类方法	
		第一阶段	第二阶段
		1,2,3 期(正) 0期(负)两类	1,2,3期 共三类
训练集	5819	5819	4429
测试集	969	969	737
总计	6788	6788	5166
类间比例	16:25:16:10	3:1	25:16:10

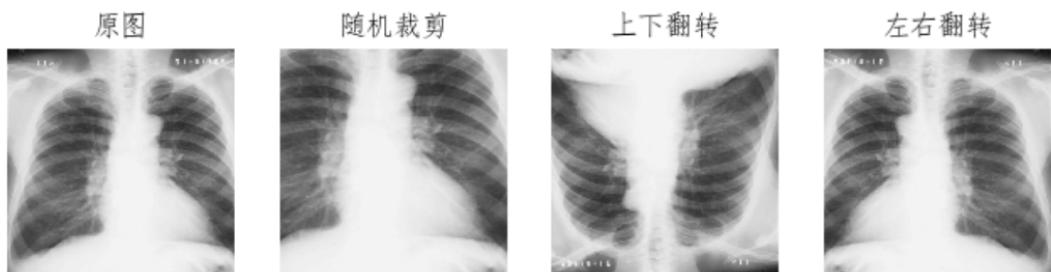


图 2.6: 数据增广过程

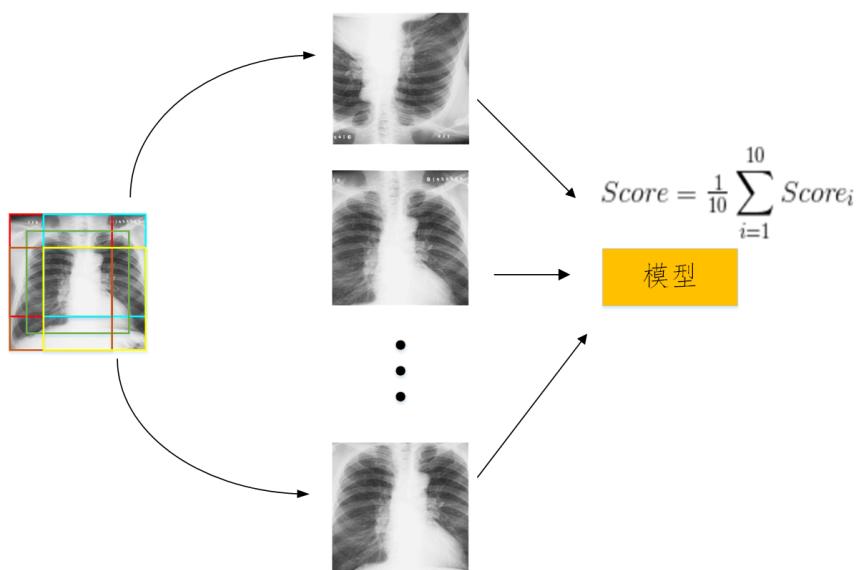


图 2.7: TenCrop过程

在训练过程，本节实验使用的方法包括3种：以0.8画幅的比例随机裁剪，随机水平翻转，随机垂直翻转，过程如图2.6示。

在测试过程中，本节实验使用的是TenCrop方法。TenCrop 方法是指对一张图片在上下左右以及中心裁剪，然后全部翻转（水平或者垂直），获得十张图片，然后对十张图片的模型打分求平均，最后综合判定这一张影像的最终分类，是一个类似bagging 集成思路用于提升模型准确率，过程如图2.7示。

2.4.2 实验设置

基分类器选取实验中，数据集设置如表2.3 中“基分类器”一栏所示，多种网络训练使用了1块GTX1080Ti 显卡，优化策略为SGD (stochastic gradient descent)，基于ImageNet 的预训练模型进行fine-tuning，训练参数为：学习率 (lr) 为0.001，动量 (momentum) 为0.9，每经过40 个epoch 设置lr衰减到原来的0.1 (衰减率)，共训练了80个epoch。

基分类器使用代价敏感方法实验中，数据集设置如表2.3中“层次分类方法”一栏所示，依2.4结果网络选择DenseNet121，训练使用了1块GTX1080Ti 显卡，优化策略为SGD (stochastic gradient descent)，基于ImageNet 的预训练模型进行fine-tuning，训练参数为：学习率 (lr) 为0.001，动量 (momentum) 为0.9，每经过40 个epoch 设置lr衰减到原来的0.1 (衰减率)，共训练了80个epoch。

2.4.3 评价指标

实验使用宏查全率 ($macro - R$) 为分类器的性能评价指标，下面进行简单的介绍。

本文使用的尘肺病数据集共4类，类间样本数量相差非常悬殊，查全率 (recall) 和查准率 (precision) 相比于准确率 (accuracy) 是在样本不平衡数据集中更加科学全面的指标。查全率又叫召回率，代表着预测为正且真实为正的样本占真实为正样本的比例。在二分类问题中，样本有正负之分，预测结果也有正负之分，根据真实类别与模型预测类别组合可划分为真正例 (TP)，假正例 (FP)，真反例 (TN) 和假反例 (FN)，查全率 R 与查准率 P 分别定义为2.8和2.9所示。

$$R = \frac{TP}{TP + FN} \quad (2.8)$$

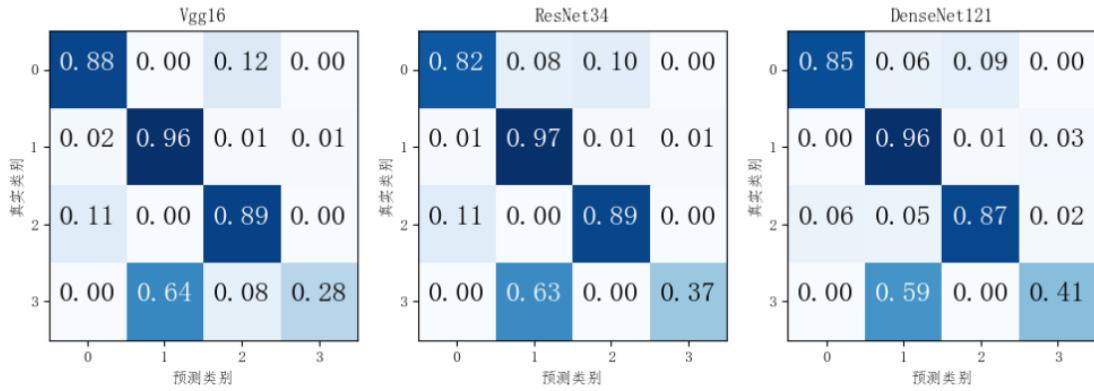


图 2.8: 部分候选基分类器的混淆矩阵

$$P = \frac{TP}{TP + FP} \quad (2.9)$$

具体而言， P 表示预测为正且真实为正的样本占真实为正样本的比例， R 表示预测为正且真实为正的样本占预测为正样本的比例。在尘肺病X光片分类任务中我们更加关注查全率 R ，即要尽可能全的将各类患病样本找出来，尽可能的减少患病样本的漏诊。

当遇到多分类问题，本文需要用到宏查全率作为评价指标。本文将 n 类分别作为正类其余类作为负类，可以构建 n 个二分类的混淆矩阵，再计算出 n 个混淆矩阵的查全率，记为 R_1, R_2, \dots, R_n ，再计算平均值即得到宏查全率($macro - R$)，公式如式2.10。

$$macro - R = \frac{1}{n} \sum_{i=1}^n R_i \quad (2.10)$$

2.4.4 结果与分析

1) 基分类器选择结果

各深度卷积神经网络在尘肺病数据集上的分类结果如表2.4 “基分类器”栏所示，结果表表明端到端4 分类的单模型宏召回率到0.77就到达极限，最好的是DenseNet121，宏召回率为0.7723，本文选了其中指标较好且接近的Vgg16，ResNet34与DenseNet121观察它们的分类混淆矩阵，如图2.8示，表明3类被误分为1类是主要瓶颈，这跟数据集中3类样本极其缺少有直接关系。同时DenseNet121是在这个问题中 $macro - R$ 最高的单个模型，所以本文在层次分类方法中的CNN1与CNN2均使用DenseNet121作为基分类器。

表 2.4: 尘肺病数据集分类结果

模型	说明	$macro - R$
AlexNet	基分类器	0.5503
Vgg16		0.7586
ResNet34		0.7651
DenseNet121		0.7723
基分类器	层级分类方法	0.9016
基分类器-代价敏感		0.9050
基分类器-截断点		0.9121
基分类器-截断点-代价敏感		0.9162

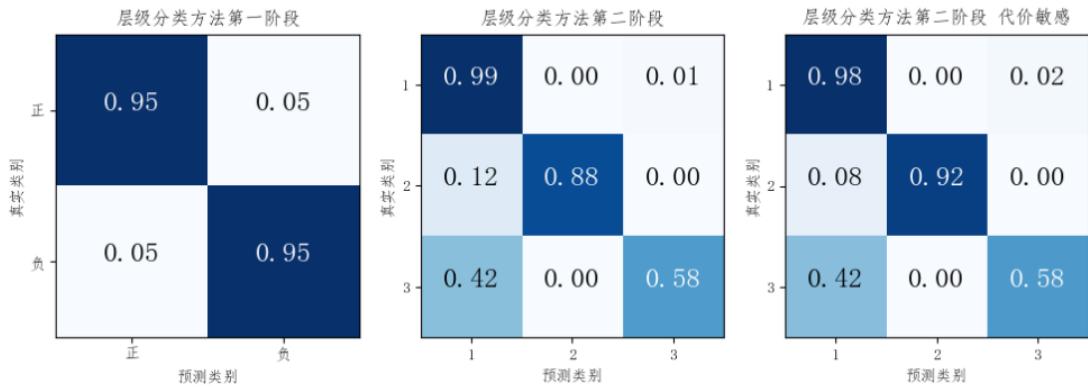


图 2.9: 基分类器加入代价敏感前后对比：加入前（左、中）与加入后（右）

2) 基分类器代价敏感优化结果

选定DenseNet121作为基分类器后用代价敏感方法进行优化。进行代价敏感方法需要知道原学习的分类结果，图2.9左、中图为层级分类方法两个阶段各自基分类器的分类结果的混淆矩阵。可以看到第一阶段的基分类器性能较好，优化空间不大；而第二阶段的基分类器表现较差，易将3和2类误分到1类，还有改进空间。依据图2.9中间的混淆矩阵，1:2:3的召回率比约99:88:58，本文使用基于分类结果代价敏感向量的方法进行调节，按召回率比例换算为罚项向量[5:8:9]加入损失函数，并重新训练第二阶段基分类器。

加入代价敏感方法后，分类结果混淆矩阵如图2.9 的右图所示，第一阶段基分类器和第二阶段基分类器加入代价敏感方法前后 $macro - R$ 如表2.5示，可以看到第二阶段基分类器加入代价敏感方法后，2类的误分情况得到一定的缓解，

表 2.5: 基分类器加入代价敏感分类结果

模型	说明	$macro - R$
基分类器	层次分类方法 第一阶段	0.9483
基分类器	层次分类方法	0.8173
基分类器-代价敏感	第二阶段	0.8242

表 2.6: 层级分类方法截断点取值与宏召回率对应表

基分类器 -截断点 -代价敏感	截断点	0.05	0.10	0.15	0.20	0.25	0.30	0.35
	$macro - R$	0.9098	0.9110	0.9107	0.9154	0.9140	0.9133	0.9157
基分类器 -截断点 -代价敏感	截断点	0.40	0.45	0.50	0.55	0.60	0.65	0.70
	$macro - R$	0.9162	0.9089	0.9050	0.9099	0.9078	0.9072	0.9045
基分类器 -截断点 -代价敏感	截断点	0.75	0.80	0.85	0.90	0.95		
	$macro - R$	0.9018	0.8965	0.8851	0.8697	0.8497		
基分类器 -截断点	截断点	0.05	0.10	0.15	0.20	0.25	0.30	0.35
	$macro - R$	0.9050	0.9061	0.9065	0.9112	0.9098	0.9091	0.9116
基分类器 -截断点	截断点	0.40	0.45	0.50	0.55	0.60	0.65	0.70
	$macro - R$	0.9121	0.9049	0.9016	0.9065	0.9045	0.9038	0.9011
基分类器 -截断点	截断点	0.75	0.80	0.85	0.90	0.95		
	$macro - R$	0.8985	0.8931	0.8819	0.8666	0.8467		

$macro - R$ 提升了约1个百分点，验证了加入代价敏感方法能够调优模型性能的有效性。

3) 截断点调优结果

如图2.10示，表2.4中基分类器-截断点比DenseNet121 的宏查全率高1.1个百分点，实验证明截断点的选择可以有效提升分类宏查全率。

在层次分类方法两阶段加入截断点，不同的截断点值与层级分类方法最终 $macro - R$ 值如表2.6示。在表2.4的“层次分类方法栏”是各种调优手段的最高 $macro - R$ ，其中各项的表意是：在层级分类方法中，“基分类器”表示基分类器没加代价敏感方法也没设置截断点（默认为0.50）；“基分类器-代价敏感”表示仅基分类器加入代价敏感方法；“基分类器-截断点”表示仅设置了截断点；“基分类器-截断点-代价敏感”表示既有基分类器加了代价敏感方法也设置了截断点。

可以看到，表2.4中“基分类器-截断点”的最好结果比“基分类器”高出

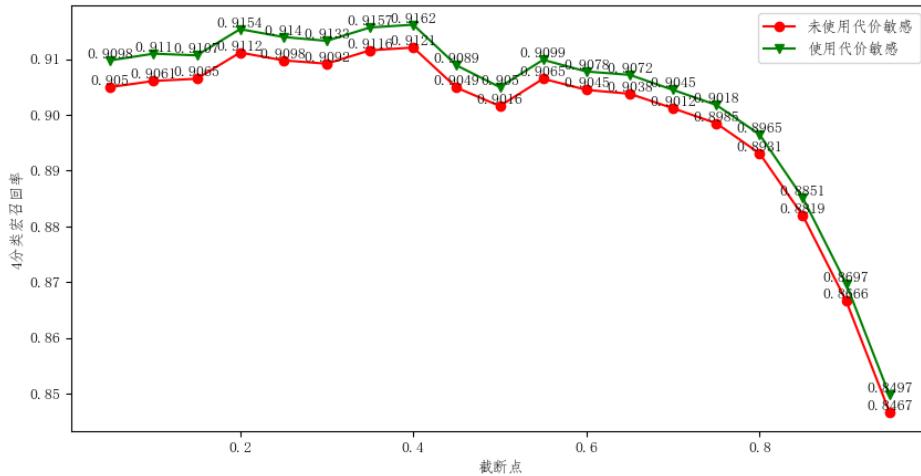


图 2.10: 截断点与层级分类方法宏召回率的关联性

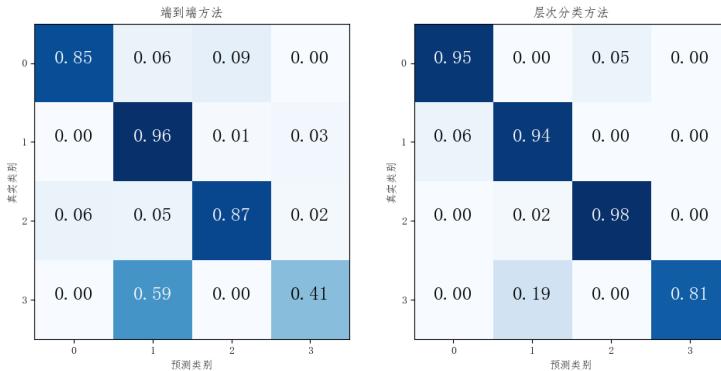


图 2.11: 端到端方法与层次分类方法最终分类混淆矩阵

约1个百分点，证明了截断点方法的有效性；“基分类器-代价敏感”比“基分类器”高出虽然不多，但是也证明了代价敏感方法的有效性，同时图2.10展示了“基分类器-截断点”与“基分类器- 截断点-代价敏感”在不同的截断点取值时层级分类方法整体的*marco - R*变化趋势，同样证明了代价敏感方法的有效性。表2.6是“基分类器-截断点”与“基分类器-截断点-代价敏感”的截断点的19组取值 ((0.0,1.0) 间每隔0.05 取一个值) 与最终*marco - R*具体对应关系。最后，层级分类方法“基分类器-截断点-代价敏感”在截断点为0.4时取得了尘肺病数据集上最好的结果，宏召回率0.9162。

同样是使用DenseNet121，在表2.4的层级分类方法“基分类器-截断点-代价敏感”，相比端到端的直接4分类方法“DenseNet121”宏查全率高出约14个百分

点，提升效果相当明显。图3.2两者分类结果的混淆矩阵，结果显示层次分类方法能够更好的区分0类和3类。从层次分类方法的流程上分析，一定程度上可以说明尘肺病X光片患病与不患病的样本也许存在着非常大的差距，所以层次分类方法第一阶段能够很好的将它们区分开，在这个基础上，患病期别样本的差距也许并不大，但是由于在层次算法第二阶段已经排除了健康样本的，使得模型能够足够聚焦到患病期别之间的差距，使得3分类的效果也不错，从而整个层次分类方法*marco - R* 达到了较高的数值。实验充分证明，层次分类方法在尘肺病分类这个问题的有效性。

2.5 小结

在本章中，针对尘肺病诊断中对于患病期别的重视程度不同提出了层级分类方法，即健康与否的漏诊代价比患病期别的误诊代价大得多。层次分类方法分为两阶段，首先判断样本是否患病，如果超过第一阶段模型的截断点则认为患病，否则健康；然后再细分认为“患病”样本的类别。层级分类方法考虑了两个要点：一是单层级的基分类器要尽可能准确，为此实验中对多种主流的深度卷积神经网络以及加入代价敏感方法调优前后进行了对比，为每一层级选择了最好的基分类器；二是两个层级通盘考虑起来最终分类结果要尽量准确，为此实验中进行了多种截断点取值与最终分类指标结果的对比，选择了最好的截断点。同时在训练基分类器的过程中，针对数据不平衡的问题在数据层面使用了多种扩充方法与模型层面使用了代价敏感方法对基分类器进行调优。层级分类方法通过先“粗筛”再“细筛”的方式，相较于普通的深度学习端到端式多分类方法有效的提升了识别精度，宏查全率由0.7723提升到0.9162，并且它能够更好的区分尘肺病第3期样本。

第三章 基于判别性区域发现的尘肺病细粒度病理特征提取

3.1 概述

3.1.1 背景与动机

医学影像分析作为辅助医生诊断的重要手段，识别算法的主要作用是对样本给出准确的判断结果，辅助医生更快做出诊断判断并降低医生负担。同时，如果算法还能给出每个样本被分类的原因供医生二次确认，这样就能更好的避免检查过程中误诊和漏诊。在尘肺病层级分类方法的实验中，实验证明了深度神经网络在尘肺病X光片分类中的有效性，然而还不能对每一次的分类结果给出一定的解释提供更强的诊断依据。

深度学习却本身是一个黑箱模型。深度神经网络是一个具有很多隐层的模型，它通过逐层的特征变换，将样本在原输入空间的特征表达变换到一个新的特征空间而拥有强大的表征能力。对比于人工构造的特征，它的层层变换所得到特征能够更好的刻画数据丰富的内在信息，同样因为它的层层变换使得模型缺乏直观解释性，并且随着网络结构加深就越难以理解。

为了解释尘肺病模型每一次分类的原因，探究究竟是影像的哪一部分特征（本章称为判别性区域）决定了最终的分类结果，本章将深度学习中可解释性技术——这种可解释性技术主要依据深度卷积神经网络的高维特征图与梯度的结合获得——运用到尘肺病X光片的分类中，从深度卷积神经网络的内部状态和机制来理解尘肺病分类模型。

同时，在上一章的实验中，本文基于监督学习的方式，将尘肺病整张影像以及对应的标签的输入网络去学习。在这个过程中，整张的尘肺病影像相较于其中的尘斑、结节等病理区域对于尘肺病识别任务而言是更粗粒度的特征，细粒度特征相比于粗粒度特征有着更少的噪声，作为训练数据可能会有更好的识别精度。但在基于监督学习的方式中，要使用细粒度病理特征作为训练数据，需要将这些病理区域先标记出来，但这种标签是本数据集所不具备的，本数据只有“期别”层面的标签。可解释性技术可以利用训练好的网络与判别性区域，由尾到头“反向”的捕捉到尘肺病影像中更细粒度的病理特征，并且这种“反向”方式不需要数据有更细粒度的病理区域标签，网络从“期别”层面的训练过程中就具有了这种能力。为了更充分利用深度卷积神经网络强大的特征表达能力，本章对更细粒度的病理特征进行了更多的探究。

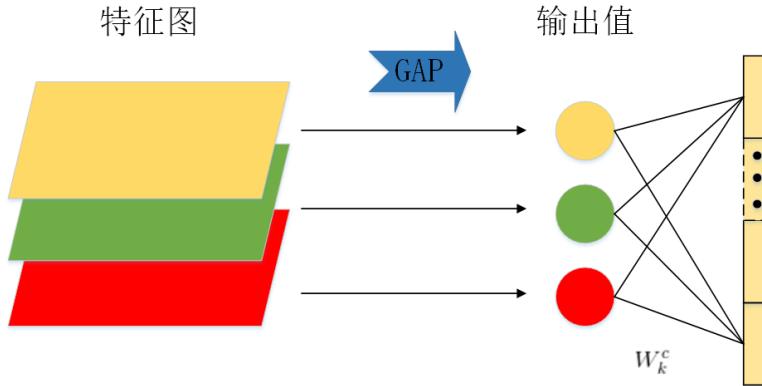


图 3.1: 全局平均池化 (GAP)

3.1.2 本章工作

基于深度学习可解释性技术，围绕尘肺病X光片识别中的判别性区域发现与细粒度病理特征的发掘利用，本章具体工作如下：

- 利用深度学习中基于高维特征图与梯度信息的可解释性技术，发现了尘肺病在分类过程中的判别性区域，为分类结果给出一定程度上的解释。
- 基于发现的判别性区域，从深度神经网络中提取出的尘肺病更细粒度卷积特征，并利用这种细粒度的病理特征进行再融合处理进行了分类效果的提升。

3.2 相关知识

3.2.1 CAM可解释性技术

CNN的可解释性技术有两个很重要的方法，一种是基于梯度的思想的利用反卷积的导向反向传播法（guided-backpropagation）方法 [48]，一种是基于高维特征图与分类权重结合的CAM（Class Activation Mapping, CAM）方法 [50]。本章实验主要可解释性技术是基于CAM方法的，以下进行一些简单介绍。

CNN的所有网络层作用大体分为两种，一部分用于特征提取（conv等），一部分用于分类（softmax等），在较为后期的CNN中，特征提取的最后一层已经使用全局平均池化层（Global Average Pooling, GAP）替代全连接层（fc），GAP结构如图3.1示。GAP相比于fc层有着减低模型参数，减少计算量与降低过拟合风险的优点。它最主要的特点是针对将前面层提取到的feature map，直接输出每个feature map的平均值，而非像fc对所有feature map进行线性变换，这样

保存了feature map的空间信息。正是这个特点，每个feature map通过GAP都可以对应得到一个均值，这些均值再经过softmax层得到最后的输出，在均值输入softmax层过程中有一个分类权重，可以对网络最后一层所有通道的feature map利用分类层的权值进行加权叠加。2016年Zhou [50]等人正是基于这种思路提出了CAM，以热力图（saliency map）的形式高亮的标出图像那些决定分类结果的区域。最后的class activation map是网络最后一层所有featur map的线性组合。然后为了方便观察，直接使用上采样，将class activation map还原到与原图一样的大小，通过叠加，就可以知道哪些区域是与最后分类结果最相关的部分。

Grad-CAM [35]是CAM的改进技术，区别就在于如何获取每个特征图的权重。CAM的特征图加权系数利用了GAP，要求深度神经网络必须有GAP层，但是有部分网络结构并没有GAP；而Grad-CAM的加权系数是通过反向传播得到的，因此它不需要CNN必须要有GAP层。因此Grad-CAM 比CAM适用的范围更广，对各类网络结构，各种任务都可以使用。

3.2.2 主成分分析

图像经过深度神经网络层层转换后，在靠后的层往往能抽取出图像非常抽象的特征，在深度神经网络越来越深的今天这种抽象特征往往有着非常高的维度。本章涉及到尘肺病X光片的细粒度卷积特征的进一步的处理和探索，在实验中主要使用了主成分分析技术对复杂特征进行降维，下面进行一些简单的介绍。

复杂的高维数据会有一些问题：原始的高维空间中，包含了冗余信息以及噪音信息；高维空间样本具有稀疏性，导致模型比较难找到数据特征；过多的变量会妨碍模型查找规律等。通过数据降维，可以减少数据维度，便于计算和可视化，更有利于有效信息的提取综合及无用信息的摈弃寻找到数据内部的本质结构特征。降维的本质是学习映射函数 $f : x \rightarrow y$ ，将原高维空间中的数据点映射到低维度的空间中，其中 x 是原始数据点的表达， y 是数据点映射后的低维向量表达，通常 y 的维度小于 x 的维度。降维主要有两种：线性映射和非线性映射方法。

主成分分析（Principal Component Analysis, PCA）是一种常用的线性降维方法，它的目标是通过某种线性投影，将高维的数据映射到低维的空间中表示，并期望在所投影的维度上数据的方差最大，以此使用较少的数据维度，同时保留住较多的原数据点的特性。PCA算法具体步骤如算法3.1示。

直观来说，如果把所有的点都映射到一起，那么几乎所有的信息（如点和

算法 3.1 PCA 算法

输入: 样本集 $D = \{x_1, x_2, \dots, x_m\}$; 低维空间维数 d'

输出: 投影矩阵 $\mathbf{W} = \{w_1, w_2, \dots, w_{d'}\}$

- 1: 对所有样本进行中心化: $x_i \leftarrow x_i - \frac{1}{m} \sum_{i=1}^m x_i$;
- 2: 计算样本的协方差矩阵 $\mathbf{X}^T \mathbf{X}$;
- 3: 对协方差矩阵 $\mathbf{X}^T \mathbf{X}$ 进行特征值分解;
- 4: 取最大的 d' 特征值对应的特征向量 $w_1, w_2, \dots, w_{d'}$;

点之间的距离关系) 都会丢失。而如果映射后方差尽可能的大, 那么数据点则会分散开来, 以此来保留更多的信息。数学上用协方差矩阵表征变量之间的相关程度 (维度之间关系), 对协方差矩阵的特征值分解, 然后将原始数据向这些特征值方向上投影。有 n 维向量 \mathbf{W} 为目标子空间的一个坐标轴方向 (称为映射向量), 最大化数据映射后的方差有式3.1。其中 m 是数据实例个数, \mathbf{x}_i 是第 i 个数据实例向量表达, $\bar{\mathbf{x}}$ 是所有数据实例的平均向量。定义 \mathbf{W} 为包含所有映射向量为列向量的矩阵, 经过线性代数变换, 可以得到目标函数式3.2, 其中 tr 是矩阵的迹, \mathbf{A} 是数据协方差矩阵。容易得到最优的 \mathbf{W} 是由数据协方差矩阵前 k 个最大的特征值对应的特征向量作为列向量构成的, 这些特征向量形成一组正交基并且最好地保留了数据中的信息。

$$\max_{\mathbf{W}} \frac{1}{m-1} \sum_{i=1}^m (\mathbf{W}^T (\mathbf{x}_i - \bar{\mathbf{x}}))^2 \quad (3.1)$$

$$\begin{aligned} \min_{\mathbf{W}} & tr(\mathbf{W}^T \mathbf{A} \mathbf{W}), \text{ s.t. } \mathbf{W}^T \mathbf{W} = \mathbf{I} \\ & \mathbf{A} = \frac{1}{m-1} \sum_{i=1}^m (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \end{aligned} \quad (3.2)$$

PCA优点是在误差最小的情形下提取了主要信息; 缺点是将所有的样本作为一个整体对待, 去寻找一个均方误差最小意义下的最优线性映射投影, 而忽略了类别属性, 而它所忽略的投影方向有可能刚好包含了重要的可分性信息。

3.3 算法介绍

3.3.1 符号描述

本章将通过Grad-CAM [35]方法找到的那些决定尘肺病影像分类结果的影像部分称为“判别性区域”; 将判别性区域对应的CNN高维特征图, 经过SCDA (Selective Convolutional Descriptor Aggregation) [45]方法筛选后的特征称为“细

粒度病理特征”；将深度卷积神经网络最后一层卷积层输出的高维特征称为“*convfeature*”，有特征图（feature map） $\{A_i\}_{i=1}^m$ ；将*convfeature*乘以分类权重后的判别性区域对应的高维卷积特征称为“判别性特征”，有特征图 $\{F_i\}_{i=1}^m$ 。

3.3.2 判别性区域发现

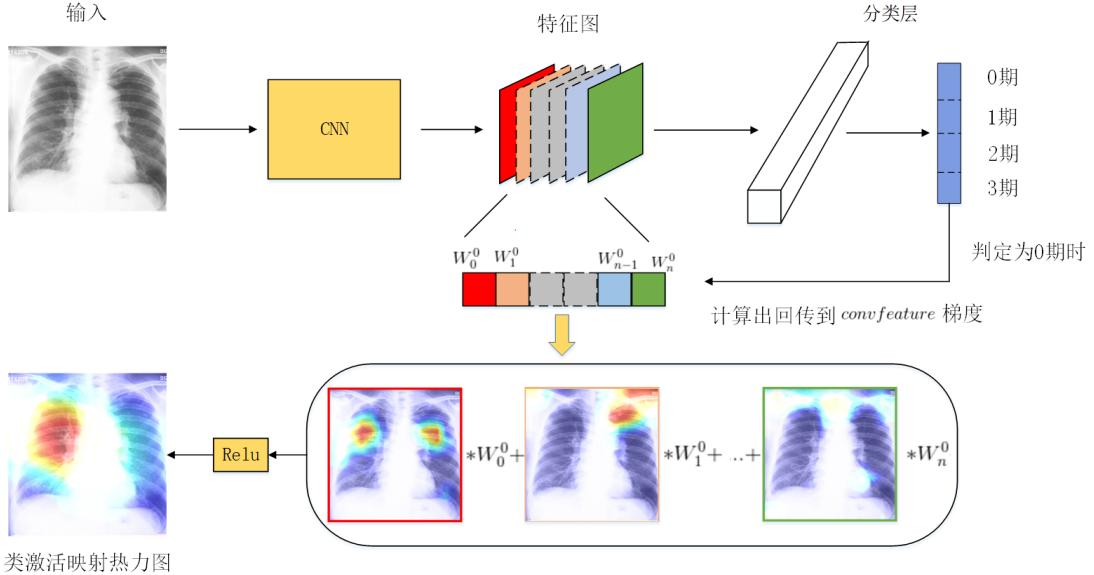


图 3.2: 尘肺病判别性区域发现流程

判别性区域发现流程如图3.2示，实验需要首先训练完毕一个尘肺病4分类模型，训练好的模型作为特征探测器不再被训练。接着利用Grad-CAM方法获取数据集每张图片的判别性区域。在CAM技术中，CNN中作为特征提取层的最后一层网络的feature map经过GAP后，每一个feature map包含了不同类别的信息，只需要提取出分类时所有的权重，往回找到对应的特征图，然后对所有打的feature map进行加权求和即可获得判别性区域。由于DenseNet121 原始结构并没有包含GAP层，本文采用了梯度的全局平均来计算权重。定义Grad-CAM 中第 n 个特征图对应类别 k 的权重为 W_n^k ，可以通过下面的公式3.3计算得到。

$$W_n^k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^k}{\partial A_{ij}^n} \quad (3.3)$$

其中 Z 为feature map的像素个数， y^k 为第 k 类得分的梯度， A_{ij}^n 为第 n 个特征图中， (ij) 位置处的像素值，然后再求得所有的特征图对应的类别的权重大后进行

加权求和，这样便可以得到最后的热力图，求和公式如式3.4。实验算法如算法3.2示。

$$L_{Crad-CAM}^k = \text{ReLU} \left(\sum_n W_n^k A^n \right) \quad (3.4)$$

算法 3.2 尘肺病判别性区域发现

输入：尘肺病数据集 $D = \{(x_k, y_k)\}_{k=1}^n$ ；

输出：原始样本与其对应判别性区域图 $\{(x_k, \hat{x}_k)\}_{k=1}^n$

- 1: 训练完毕的尘肺病分类深度卷积神经网络 G ;
 - 2: **for all** $(x_k, y_k) \in D$ **do**
 - 3: 计算当前样本通过网络 G 的输出 \hat{y}_k ;
 - 4: 计算当前样本通过网络 G 的最后一层卷积层的输出 $convfeature$, 包含特征图 $\{A_i\}_{i=1}^m$;
 - 5: 依据式3.3计算输出 \hat{y}_k 对应的特征图 $\{A_i\}_{i=1}^m$ 权重 W ;
 - 6: 依据式3.4计算权重 W 与特征图 $\{A_i\}_{i=1}^m$ 得到判别性特征 $\{F_i\}_{i=1}^m$;
 - 7: 特征图 $\{F_i\}_{i=1}^m$ 加和均值 $\frac{1}{m} \sum_{i=1}^m F_i$ 成二维矩阵并通过线性插值的方法变换为样本 x_k 的尺寸并通过相加重叠, 得到高亮判别性区域影像 \hat{x}_k ;
 - 8: **end for**
-

3.3.3 细粒度病理特征提取

在判断性区域发现的实验中，本文能够找到样本影像中那些极大程度影响尘肺病分类结果的影像区域。在实验分析中，如图3.5示，随机的抽取了一些样本图片进行展示，可以观察到在没有患病区域这种细粒度的标签的情况下，判别性区域的 $convfeature$ 自动的集中在肺区；也会集中噪声或背景区域，比如在脊椎、肋骨区域；也会有大小不一的多块区域区域，有些在肺区而有些在噪声区域。尘肺病的医学判定是依据X光片上观察肺部纤维化、尘斑等症状的严重程度来诊断的，判别性区域的 $convfeature$ 并不能完全反应到这些细粒度的病理特征上，因此在本小节实验中使用SCDA方法在判别性特征提取出细粒度病理特征，并用更多的算法对提取出来的细粒度特征进行进一步的探究。

一张图片经过CNN层层变换以后得到 $convfeature$ ，这个中间特征是一个通道数为 c 高度为 h 宽度为 w 的三维张量 T ，在本实验中使用的DenseNet121的 $convfeature$ 尺寸为 $1024 \times 7 \times 7$ 。 T 的每个通道的特征图与分类权重 W 相乘即是判别性特征。为了便于直观观察，将判别性特征通过线性插值转换到原始影像的尺寸并重叠，展示结果如图3.3所示。

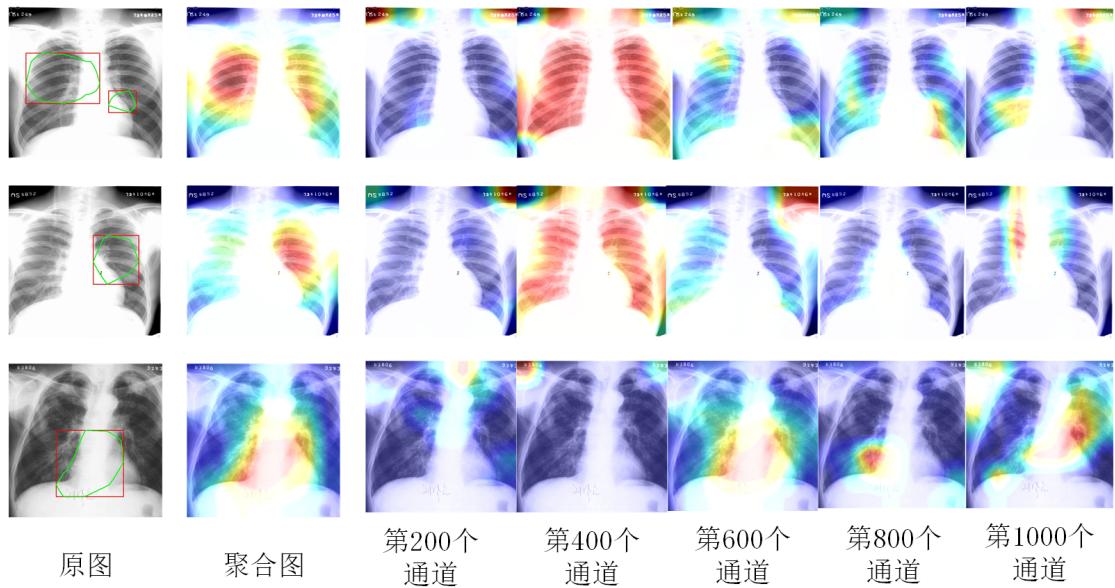


图 3.3: 判别性特征展示

图3.3中越亮的区域代表对应区域的响应值越大，可以观察到：对于一个样本而言，判别性特征的 c 个通道上每个通道之间的特征图上的响应值不同；对于判别性特征的 c 个特征图而言，不同的样本在同一个通道上特征图响应值也不一样。这些响应值，值越大对分类结果影响越大，就越有可能是高维特征中的细粒度的病理特征所对应的值，实验中将张量 T 中这些较大的响应值选取出来，即视为尘肺病的细粒度病理特征。

尘肺病细粒度病理特征提取总体流程如图3.4所示。具体步骤是，首先样本输入模型，按照标签与算法3.2获得判别性特征，再沿着判别性特征的深度方向（通道维度）将所有的特征图相加，将 $c \times h \times w$ 三维张量 T 变为 $h \times w$ 二维矩阵 B ，称为聚合图即 $B = \sum_{n=1}^m S_n$ (S_n 为判别性特征中的第 n 通道上的特征图)。

对于加合矩阵 B ，它代表了判别性特征的各个响应值的综合响应强度， B 有 $h \times w$ 个加合起来的网络响应值，因为此时CNN提取出来的特征尚未破坏原始影像的空间位置信息，它就能对应着原始影像 $h \times w$ 个位置。 B 中某个值 (i, j) 越大那么就越可能是病理特征。因为没有细粒度的病理标记，为此实验设计了 B 中所有位置响应值得平均值 \bar{b} 作为阈值，若 B 中某个位置 (i, j) 超过 \bar{b} ，则表示该值是病理特征，该值对应的原图区域是纤维化、尘斑等病理影像。将 B 中大于 \bar{b} 的部分提取出来，作为掩码矩阵 M 如式3.5示，其中 i, j 表示矩阵 B

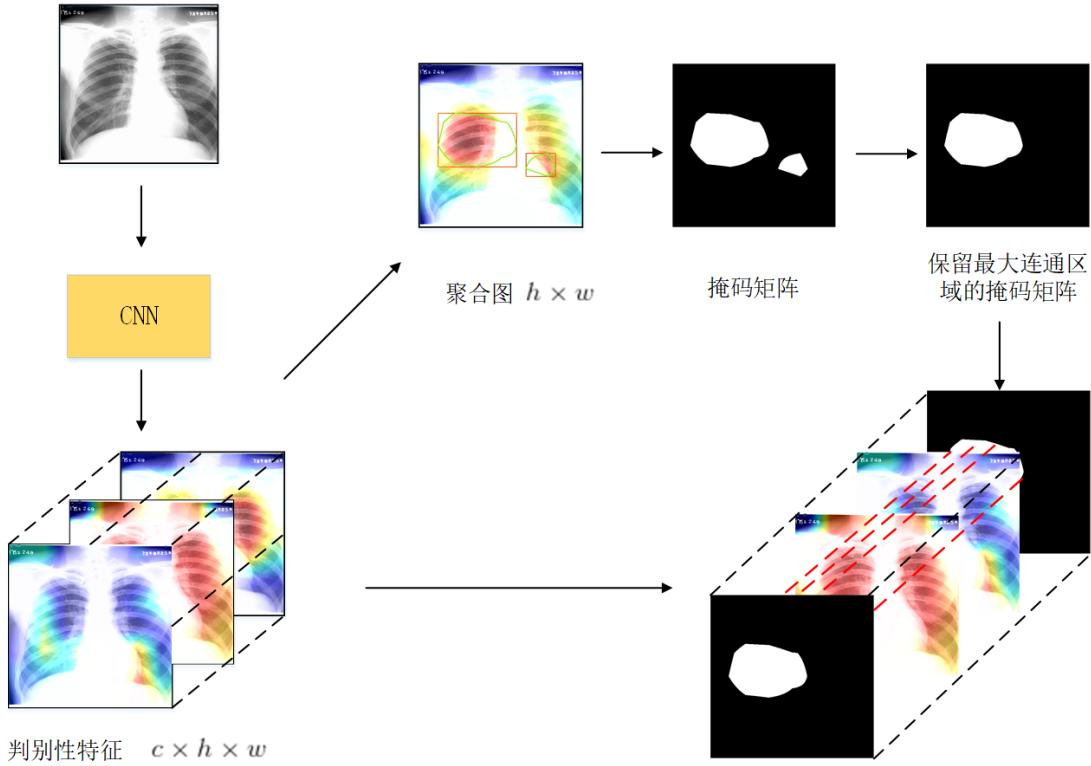


图 3.4: 尘肺病细粒度病理特征提取流程

的第 i 行第 j 列的响应值。

$$M = \begin{cases} 1 & \text{if } B_{i,j} > \bar{b}, \\ 0 & \text{otherwise} \end{cases} \quad (3.5)$$

然后将掩码矩阵 M 分别与张量 T 的每个特征图进行点乘，即可筛选出判别性特征细粒度病理特征 F ，如式3.6示，其中 $F_{(i,j)}$ 为判别性特征的所有特征图。

$$Z = \{F_{(i,j)} | M_{i,j} = 1\} \quad (3.6)$$

掩码矩阵 M 实质上是对影像高维卷积特征在判别性特征基础上的进一步筛选，在判别性区域实验结果中我们还观察到一些图片有些多块判别性区域，比如图3.5的第三列右图有两块高亮区域，其中左边的较小的高亮区域集中到肺部边缘的肋骨，是明显的噪声。这种情况对应到掩码矩阵 M 即代表着会有多块连通区域。对于这种情况实验在 M 基础上进一步筛选，本文只保留 M 中最大连通区得到新的掩码矩阵 \tilde{M} 。将掩码矩阵 \tilde{M} 分别与张量 T 的每个特征图进行点乘，

即可筛选出更加精细的判别性特征细粒度病理特征 \tilde{Z} , 如式3.7所示。

$$\tilde{Z} = \left\{ F_{(i,j)} | \tilde{M}_{i,j} = 1 \right\} \quad (3.7)$$

综上所述, 尘肺病细粒度病理特征提取算法如算法3.3所示。

算法 3.3 尘肺病细粒度病理特征提取

输入: 尘肺病数据集 $D = \{(x_k, y_k)\}_{k=1}^n$;

输出: 原始样本与其对应细粒度病理特征 $\{(x_k, \tilde{Z}_k)\}_{k=1}^n$

- 1: 训练完毕的尘肺病分类深度卷积神经网络 G ;
 - 2: **for all** $(x_k, y_k) \in D$ **do**
 - 3: 计算当前样本通过网络 G 的输出 \hat{y}_k ;
 - 4: 计算当前样本通过网络 G 的最后一层卷积层的输出*convfeature*, 包含特征图 $\{A_i\}_{i=1}^m$;
 - 5: 依据式3.3计算输出 \hat{y}_k 对应的特征图 $\{A_i\}_{i=1}^m$ 权重 W ;
 - 6: 依据式3.4计算权重 W 与特征图 $\{A_i\}_{i=1}^m$ 得到判别性特征 $\{F_i\}_{i=1}^m$;
 - 7: 依据 $\{A_i\}_{i=1}^m$ 计算加合矩阵 B , 并计算 B 的均值 \bar{b}
 - 8: 依据式3.5计算掩码矩阵 M
 - 9: 计算并只保留 M 最大连通图得到 \tilde{M}
 - 10: 用 \tilde{M} 与 $\{F_i\}_{i=1}^m$ 依据式3.7计算细粒度病理特征 \tilde{Z}
 - 11: **end for**
-

3.4 实验结果与分析

本节首先介绍实验数据集的划分和网络训练的详细参数, 然后分别展示判别性区域发现的实验结果和细粒度高维特征在各分类器中的分类精度。

3.4.1 实验数据集和实验设置

本章判别性区域发现与细粒度病理特征提取实验中, 都需要先训练CNN用作特征提取器。训练CNN使用的训练数据集仍为尘肺病X光片数据集, 采用的数据预处理方式与第二章层级分类方法所使用的方法一样, 具体的数据集划分方式如表2.3中的“基分类器”列所示, 其中数据集中仍有4类, 0:1:2:3 类别的比例为16:25:16:10。网络训练完毕后, 尘肺病判别性区域发现与细粒度病理特征提取使用的未预处理的尘肺病数据集, 具体划分与如表2.2的“基分类器”列所示, 0:1:2:3 类别的比例为4:25:4:1。

同时在第二章层级分类方法中已经验证，DenseNet121是端到端的尘肺病4分类中效果最好的CNN，因此在判别性区域发现的实验中以DenseNet121作为基分类器，训练好以后用于探测判别性区域。网络训练使用了1块GTX 1080Ti显卡，优化策略为SGD，基于ImageNet的预训练模型进行fine-tuning，训练参数为：学习率（lr）为0.001，动量（momentum）为0.9，每经过40个epoch设置lr衰减到原来的0.1（衰减率），共训练了80个epoch。

3.4.2 判别性区域发现结果

实验数据集只有图片“期别”级别的标签，并没有更精细的针对于影像患病区域的轮廓或者bounding-box的标注。因此判别性区域发现的实验中，是没有定量的量化指标，只能定性的从判别性结果中观察。同时，判别性区域实验的出发点就是辅助医生对于模型分类结果进行人为的二次检查，因此在没有量化指标的这种情况下，它也足够作为尘肺病分类过程中一个很好的补充。

基于DenseNet121的Grad-CAM判别性区域样例如图3.5示。随机的抽取了8个样例，每一列右边为判别性区域结果，均是CNN判定正确时的样本各自的可疑区域，左边为判别性区域在原影像上的对应区域，实验用bounding-box将其轮廓圈出。可以看到，判别性区域自动的集中到了肺区，显示出Grad-CAM在只有“类别”级别的标签时而没有具体“患病区域”的标签时，能够发现尘肺病可疑的患病区域，它具有一定的“弱检测”的能力。同时，这种可解释性不一定准确，比如图3.5的第一行右图，虽然预测准确，但判别性区域显示出CNN决定的原因却是中间的骨头区域，显然不合医学判断逻辑；第三排右图显示，判别性区域虽然在肺区，但左边判别性区域在原图上对应到了肋骨边缘，并没有器官结节等组织，也不符合医学判断逻辑。

作为对比，用同样的方法又训练了一个基于Vgg16的特征提取器，两者对于同一个样本判别性区域发现结果如图3.6示，该样本标签为0期，DenseNet121判断正确而Vgg16判定为1期，证明不同的CNN在针对同一张图片关注的区域是不同的；同时DenseNet121的所覆盖的判别性区域比Vgg16的一般更大，有理由推测DenseNet121 学习到了更多的信息所以能够判断更为准确。

判别性区域发现的实验结果说明了的本章方法能够发现到患病区域，但是判别性区域并不十分准确，会定位到脊椎肋骨或者空白部分。但因为在判别性区域发现的任务中本文的目的并不是准确的检测出患病区域，而是针对每次分类结果给出解释。在数据标签受限的弱监督学习的情况下，尘肺病判别性区域发现能够大概给出可疑的患病区域，能够作为辅助医生二次检查的一个很好的补充。

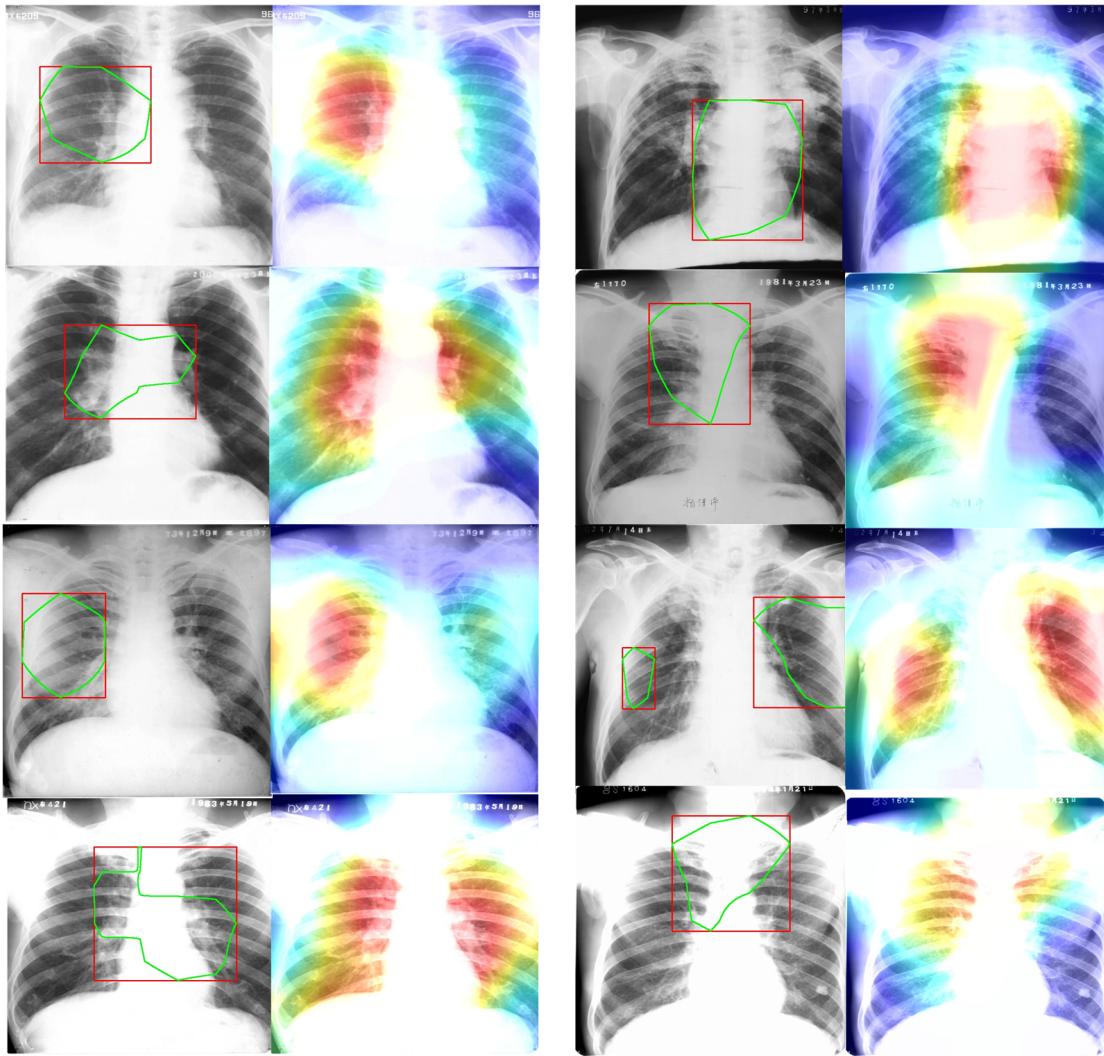


图 3.5: 尘肺病判别性区域发现样例

3.4.3 细粒度病理特征提取结果

图3.3是细粒度特征提取过程中，为了便于观察，将判别性特征在各个通道的值通过线性插值转换到原图的大小重叠而成的热力图，可以观察到对于一个样本而言，判别性特征的 c 个通道上每个通道之间的特征图上的响应值不同；对于判别性特征的 c 个特征图而言，不同的样本在同一个通道上特征图响应值也不一样。高亮区域代表响应值较大大，响应值越大则对分类结果影响越大。

实验在提取到尘肺病细粒度病理特征后进行进一步的探究实验：包括了可视化观察细粒度病理特征；以及用多种传统机器学习学习器进行分类，来验证细粒度病理特征相比于深度神经网络端到端的方式提取的高维特征是否更好。

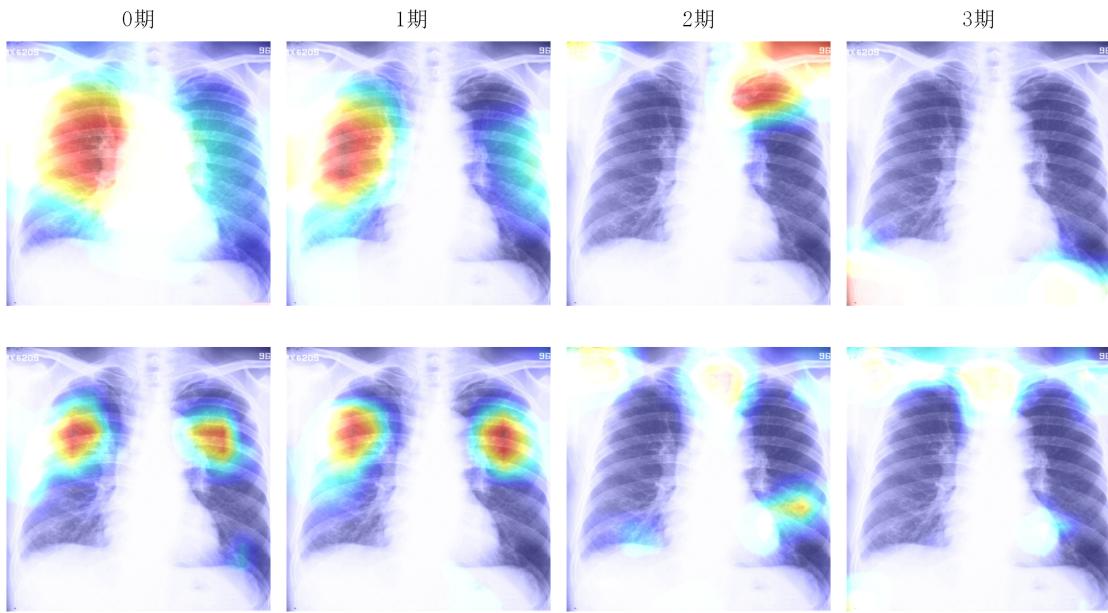


图 3.6: 判别性区域样例: DenseNet121 (上) 与Vgg16 (下)

降维可视化

使用DenseNet121提取到的细粒度病理特征为 $1024 \times 7 \times 7$ 约50000维的特征向量，为了观察这些特征的分布，实验使用了主成分分析方法（PCA）与t-SNE [28]方法将数据降维到2维和3维进行观察。同时，50000维高维特征相较于数据集约2500的数据量而言，极大可能存在大量的冗余信息，这种情况在用传统机器学习分类器对细粒度病理特征分类过程中，极有可能出现过拟合的情况。为了避免后续分类器比较过程中出现过拟合，实验中使用了主成分分析方法（PCA）将细粒度病理特征降到多个低维空间进行了对比实验。

t-SNE (t-distributed stochastic neighbor embedding) 是流形学习的一种，属于非线性降维方法，t-SNE算法对每个数据点近邻的分布进行建模，其中近邻是指相互靠近数据点的集合。在原始高维空间中，将高维空间建模为高斯分布，而在二维输出空间中，可以将其建模为t分布。该过程的目标是找到将高维空间映射到二维空间的变换，并且最小化所有点在这两个分布之间的差距。与高斯分布相比t 分布有较长的尾部，这有助于数据点在二维空间中更均匀地分布。

PCA也是流行的降维算法，它的目标是寻找能尽可能保留数据方差的数据维度，数据某个维度上有较大的方差，意味着数据在这个维度上就能保留越多的信息，然后PCA 将原始数据降维到这些维度上。相比于PCA，t-SNE一般有着更好的可视化效果。t-SNE更加适用于将高维数据降维到2维或者3维进行可视

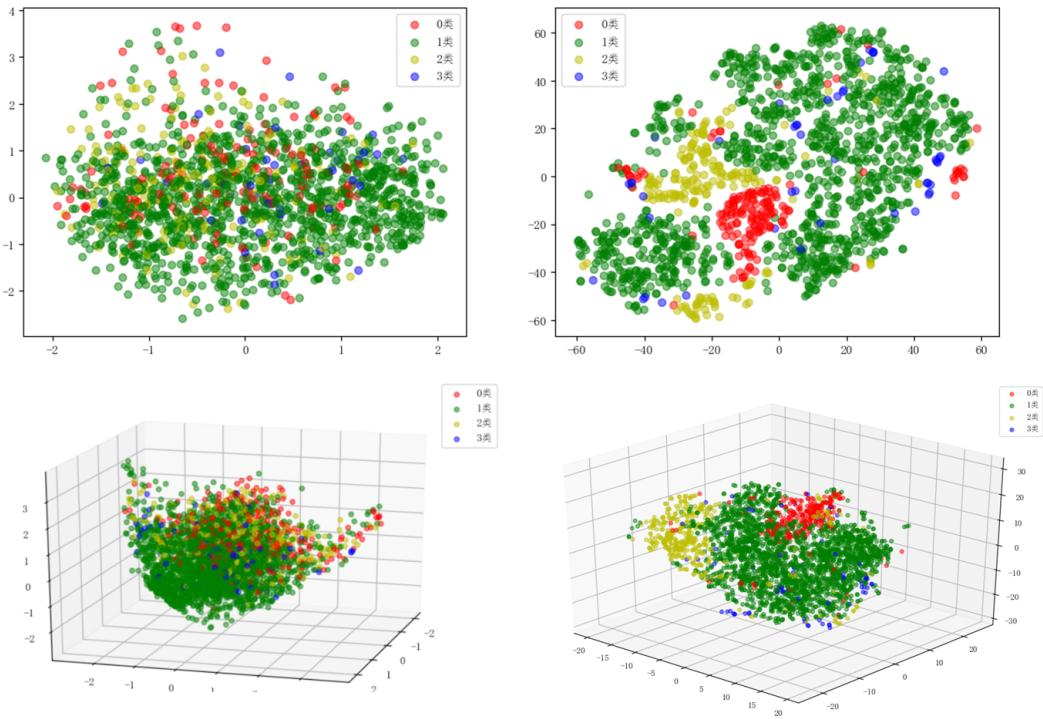


图 3.7: 细粒度病理特征降维可视化结果: PCA (左) 与t-SNE (右)

化观察, 方便使用者更加直观地了解数据的内在结构与验证算法有效性。细粒度病理特征降维可视化结果如图3.7示, 分别用PCA与t-SNE降维到2维与3维进行观察。

在降维后可视化图中, 如果数据点能够被很好的分离开来, 那么学习器同样也能找到一种将未知新数据投影到相应类别的好方法, 对于给定的学习器, 就能达到较为理想的准确率: 如果可视化结果每一个类别都是孤立开来, 那么简单的学习器就能将该类别与其他类区分开; 但如果类别重叠, 那么可能就要构建更精细复杂的学习器去学习数据, 分类预测结果可能更加离散、错误率更高。从图3.7观察可得到, PCA可视化结果(图左侧)中所有类别的点都混杂在一起, 在2维和3维空间中显然细粒度特征不具有很好的可分性; t-SNE可视化(图右侧)相较而言有更好的区分性, 其中0, 1, 2类明显的各自聚集在一起, 而3类零散的分布在整个空间中。可以认为, 学习器很难的区分尘肺病第3类样本对应的细粒度病理特征, 而0, 1, 2类对应的细粒度特征容易被学习器区分。

分析

将细粒度病理特征降维可视化后本文进一步分析, 图3.8为细粒度病理特

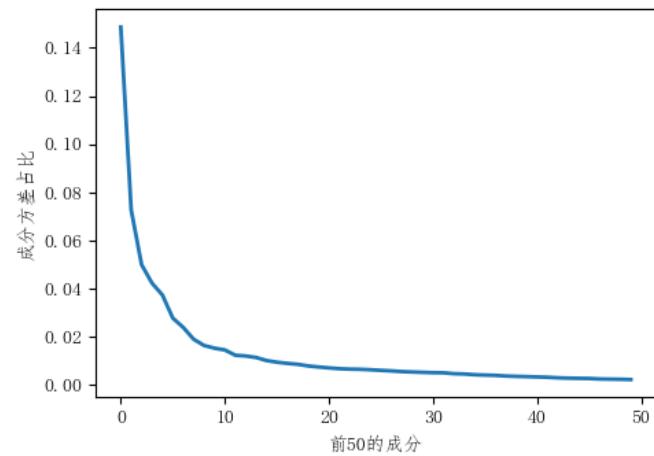


图 3.8: 细粒度病理特征PCA降维成分与方差占比关系

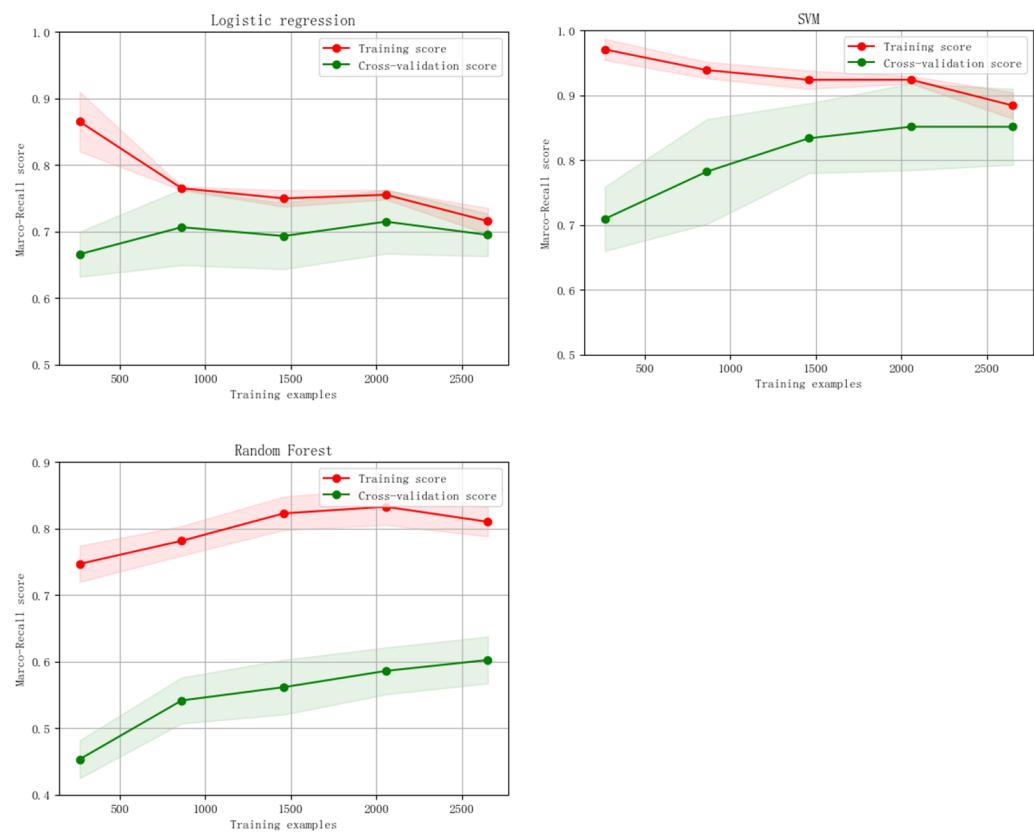


图 3.9: 降维细粒度病理特征5折交叉学习曲线

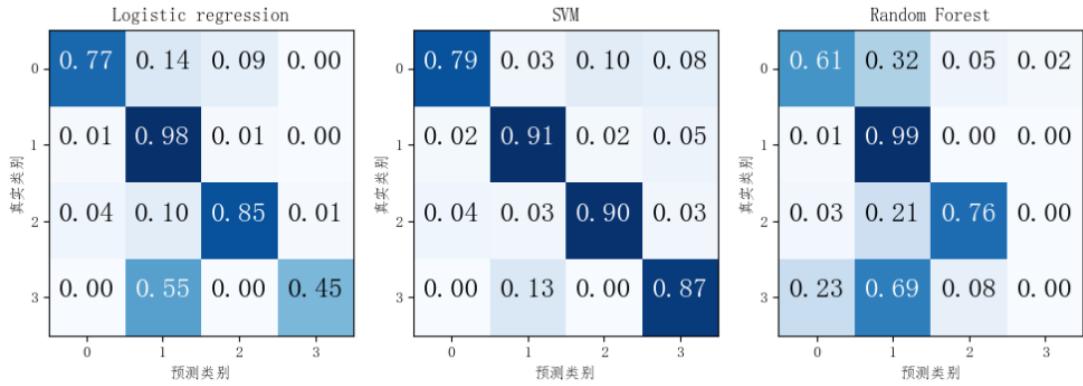


图 3.10: 降维细粒度病理特征多种学习器分类结果混淆矩阵

征进行降维后各主成分的方差占比，图3.8选取了占比前50的成分进行观察，可以看见前20个成分就几乎可以涵盖细粒度病理特征的所有信息。因此实验中选取3组参数：4, 8, 16即分别将细粒度病理特征从50000维降到4维，8维，16维进行后续分类实验。

将降维后的细粒度病理特征分别用几种常用的学习器（Logistic Regression, SVM, Random Forest）重新进行训练分类，三个学习器训练过程中学习曲线的变化如图3.9示。在每个学习器训练过程使用了5折交叉验证，涉及SVM二分类学习器采用了OvO策略，经过多次实验中本文选择了一些主要参数如表3.1中说明：pca_n代表降维维数，C为L2正则项系数，gamma为SVM的RBF核函数的参数，rf_n为Random Forest 弱学习器个数，rf_split为划分节点时每个划分最少的样本数，rf_leaf 为叶子节点最少的样本数。评价指标仍然为宏查全率marco - R最终结果如表3.1示。

从结果表3.1中可以看见，细粒度病理特征分类效果中最好的是SVM，而Logistic Regression与Random Forest 方法都效果不太好且marco - R都以较大的差距低于端到端的深度学习方法（比AlexNet高）。这三个传统机器学习器分类结果的混淆矩阵如图3.10示，可以发现三者性能瓶颈都在于难以区分3类，特别是Random Forest在3类上的区分度表现特别差，这可能与它缺乏罚项有一定关系，尤其在3类样本数量非常少时极容易过拟合。这种分类情况与上一小节t-SNE可视化结果中，0, 1, 2类明显聚集容易被区分而3类杂乱分布的情况相吻合。同时三个传统机器学习器都是在细粒度特征降维到16维时取到各自最好的效果，虽然在可视化结果中细粒度病理特征并不能很好的被线性可分，16维相比于8维和4维更高，这也符合数据在低维空间线性不可分时在越高维空间可能存在线性可分的情况。

表 3.1: 尘肺病分类结果

模型	参数	方法	说明	<i>marco - R</i>
Logistic regression	pca_n=16 C=0.01	细粒度病理特征提取	端到端 DenseNet121	0.6929
SVM	pca_n=16 C=0.0001 gamma=0.001	细粒度病理特征提取	端到端 DenseNet121	0.8857
Random Forest	pca_n=16 rf_n=20 rf_split=3 rf_leaf=3	细粒度病理特征提取	端到端 DenseNet121	0.6198
AlexNet	-	端到端	-	0.5503
Vgg16	-	端到端	-	0.7568
ResNet34	-	端到端	-	0.7651
DenseNet121	-	端到端	-	0.7723
基分类器-截断点-代价敏感	-	层级分类方法	-	0.9162
SVM	pca_n=16 C=0.0001 gamma=0.001	细粒度病理特征提取	层级分类方法 基分类器-截断点-代价敏感	0.9292

同时可以看见，对细粒度病理特征分类效果最好的SVM的结果，比单个端到端深度学习方法最好的DenseNet121的*marco - R*由0.7723提升到0.8857，高出约11个百分点。并且，它非常接近第二章层级的深度学习方法的结果(*marco - R*为0.9162)，仅相差约3个百分点。因为细粒度病理特征是在一个4分类的CNN的基础上进行的，它相比于CNN在特征层提取特征后直接输入分类层进行分类的方式，在中间可以看作是多了特征筛选的步骤。

除此之外，实验中还比较了使用“基分类器-截断点-代价敏感”层级分类方法的提取的细粒度病理特征的效果。实验过程类似以上说明的基于端到端4分类方法的过程：样本经过层级分类算法第一个基分类器时，若输出预测值低于截断点，则认为是0期，接着在第一个基分类器的基础上提取该样本的细粒度病理特征；否则样本进入第二个基分类器，基于第二个基分类器与输出结果提取细粒度病理特征。在端到端4分类方法中已经验证了SVM为细粒度病理特征最好的分类器的基础上，将细粒度最终降到16维用SVM分类

的 $marco - R$ 为0.9292，相比于原始层级分类方法“基分类器-截断点-代价敏感”的0.9162的 $marco - R$ 结果又高了约1.2个百分点。实验结果验证了细粒度病理特征筛选的方式的有效性。

3.5 小结

本文基于深度学习可解释性技术找到可疑患病区域并给出了分类依据，利用这些区域的高维特征提升了分类精度。本文研究了两种特征可视化方法来探究尘肺病高维抽象特征的状态。首先是通过Grad-CAM技术，获取到网络特征抽取层的分类权重与高维抽象特征，通过它们找到了每个样本分类过程中可能是结节、弥纤、尘斑等的图像病理区域，并以热力图的方法将这些区域高亮的标记出来，这些高亮区域即分类依据，称为“判别性区域”。再者在判别性区域的发现基础上，通过SCDA技术，在只有期别的标签情况下，将疑似尘斑等图像病理区域的高维特征从高维卷积特征中筛选出来，这种特征称为“细粒度病理特征”。本文对细粒度病理特征进行了进一步探究：降维可视化分析出了尘肺病医学影像各期别的特征区别，发现尘肺病的3期样本是识别算法的瓶颈；使用该特征作为多种学习器的训练数据进行对比，在端到端的方法中尘肺病宏查全率由0.7723 提升到0.8857，层级分类方法中尘肺病宏查全率由0.9162提升到0.9292，经过实验验证了细粒度病理特征的有效性。

第四章 尘肺病识别系统设计与实现

4.1 概述

4.1.1 背景与动机

在课题组与其他科研机构合作的项目中，合作方提供了实际的尘肺病人数数据，并提出了搭建尘肺病智能识别系统的需求。本章基于项目需求，将前两章中分别介绍的尘肺病层级分类方法以及判别性区域发现的方法作为项目核心算法，将识别算法落地为运用在医学诊断中的实际可用的系统。

将尘肺病识别算法的研究落地，不仅仅是完成本项目；也是作为算法研究转化到工程项目，从理论到实际运用的一次尝试。目前我国医疗数据中超过90%的数据来自于医学影像，但这些数据几乎全部需要人工分析。人工分析有着不够精确，依赖经验，缺乏量化标准造成误诊的缺点；也有着因人眼视力误差及以及精神疲劳造成漏诊的不足。同时，数据表明我国医学影像数据年增长率约为30%，而放射科医师数量年增长率约为4%，放射科医师数量增长与影像数据增长速度不匹配，专业的医疗从业人员短缺，放射科医生处理影像数据的负荷变大是可以预见的。因此，不仅是研究可靠的医学影像智能化算法有着积极意义，将算法转化为生产生活中可用的实际系统的一部分，也有着实际意义。

4.1.2 本章工作

围绕尘肺病识别项目，为了解决医院实际尘肺病辅助诊断中的问题，本章在前两章提及的算法研究基础上实现了尘肺病识别系统，包括：

- 围绕项目的实际需求，进行了总体的需求分析与系统流程分析，并对尘肺病识别系统进行了总体设计。
- 依据系统设计，基于B/S架构实现了尘肺病识别系统。包含了网络服务搭建，网页界面搭建，深度学习模块搭建，以及数据库的搭建，系统提供了尘肺病分类与可疑患病区域定位的服务。

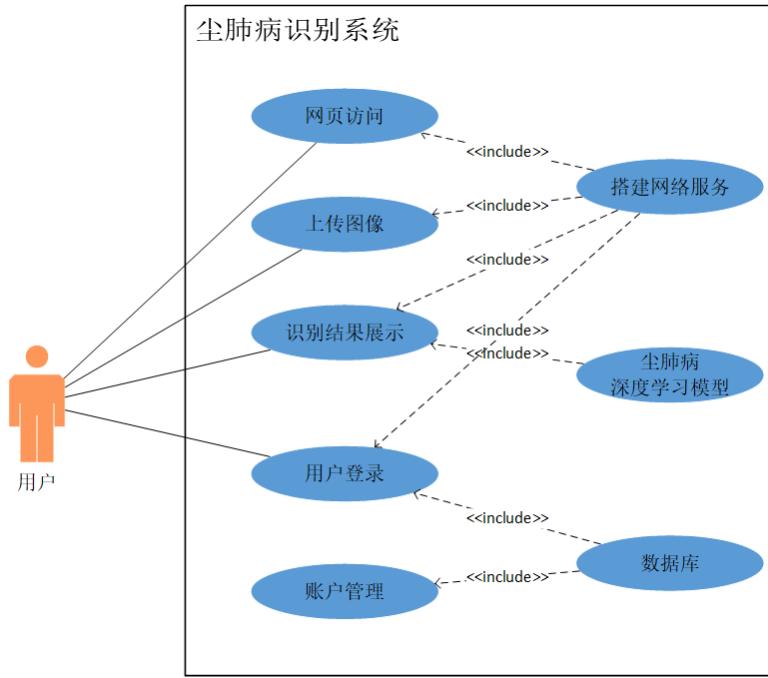


图 4.1: 尘肺病识别系统用例图

4.2 系统设计

4.2.1 需求分析

在系统设计与开发之前，先要分析清楚实际项目的需求。本系统是为专业的医疗人员提供尘肺病识别识别服务，依据B/S架构，通过搭建网站的方式，为用户提供尘肺病的期别的分类与可疑的患病区域的定位的服务。用户通过网络服务链接到指定网页，通过账户密码登录本系统，然后上传尘肺病影像，系统接收到用户传过来的影像，调用深度学习功能模块判断影像类别，并将判断结果和生成可疑区域结果图返回给用户。通过本系统提供的服务，辅助用户加快诊断尘肺病的速度，并提供二次检查的依据从而降低误诊漏诊的几率。尘肺病系统的用例图如图4.1示。

网络服务

用户需要利用互联网，通过浏览器访问到本系统的界面，然后才能上传尘肺病X光影像，使用尘肺病识别与可疑区域定位的服务，接收最后的结果。网络服务需求如下：

- **搭建网站：** 将深度学习模型部署到服务器上，并且在服务器上搭建网络服务，绑定域名后即可对外使用。

- 网页页面：服务器上搭建的网络服务需要有对应的前端网页页面，图形化的网页页面方便用户操作。
- 用户登录：网页页面需要支持用户登录，用户使用本系统的识别服务必须要先有账号密码。
- 上传影像：网页页面需要支持上传影像的功能，上传的影像保存在服务器的临时文件夹中，供尘肺病识别程序使用。
- 执行影像识别程序：载入部署在服务器上的深度学习模型，调用单张的尘肺病识别程序，调用可疑区域发现程序，并将分类结果与可疑区域结果保存到临时的识别结果文件中。
- 结果显示：将结果文件返回到用户的浏览器界面，进行展示。

账号管理服务

只有拥有授权的账号的用户才能通过账户密码登陆本系统使用尘肺病识别服务，账号信息通过数据库存储并管理，在数据库中的账号才是合法账号。需求如下：

- 增加账号：在数据库中为用户新建一个账号，包括用户名与密码。
- 删除账号：在数据库中删除一个账号，包括用户名与密码。
- 修改账号：修改数据库中一个账号的信息，只能修改账号的密码。

影像识别服务 尘肺病识别与可疑区域定位的服务，是本系统的核心任务，需求如下：

- 载入文件系统的影像：用户上传的影像保存在服务器临时文件夹，需要支持依路径载入影像、将影像转换为系统指定的尺寸（ $600 \times 600 \times 3$ ）、将转换影像转化为指定数据格式以供深度学习模型使用的功能。
- 影像预处理：需要支持深度学习中的数据增广方法：TenCrop。
- 执行分类程序：调用部署在服务器的深度学习模型，将预处理后的数据输入深度学习学习模型，获取最终分类结果与各类评分概率。
- 执行可疑区域发现程序：调用部署在服务器的深度学习模型，利用分类程序的执行结果，调用判别性区域发现方法，获取影像判别性区域结果。

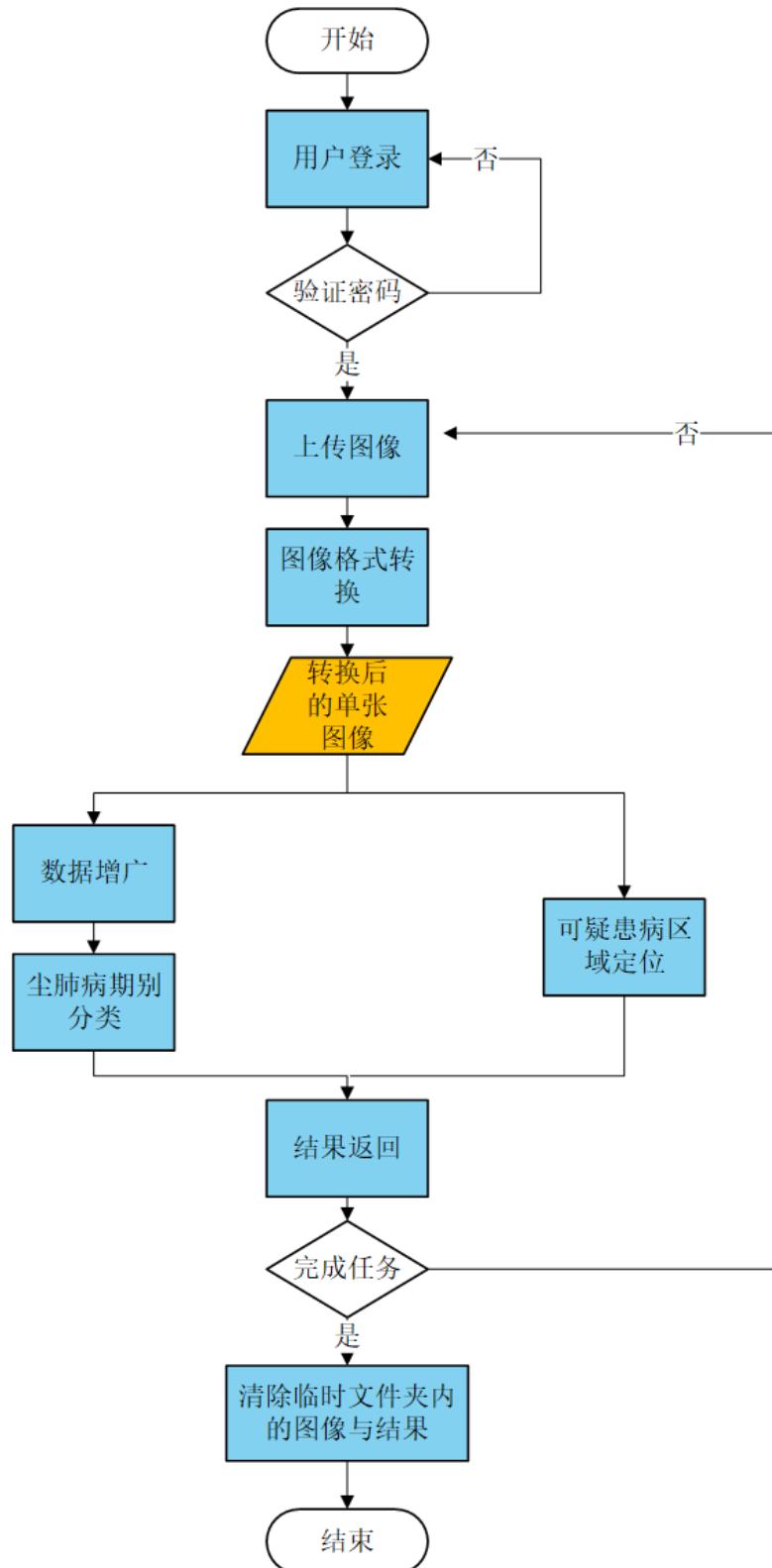


图 4.2: 尘肺病识别系统主处理流程

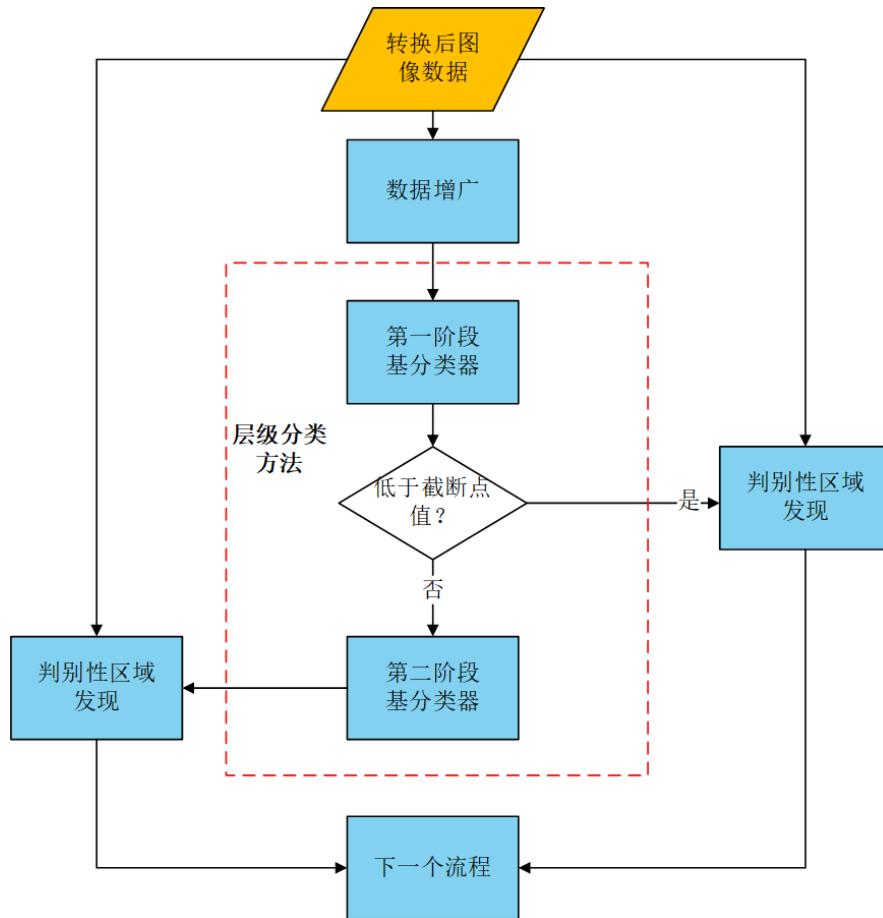


图 4.3: 尘肺病影像识别子流程

4.2.2 流程分析

主流程分析

系统主要是由两部分构成：首先用户经过浏览器网页，输入账号密码验证身份后，通过验证登录系统；然后上传需要识别的尘肺病图片，尘肺病识别系统接收到图片进行分类和可疑患病区域定位，将结果再返回到用户网页界面。尘肺病识别系统主处理流程图如图4.2示。

尘肺病识别子流程分析

本系统核心是提供尘肺病识别与可疑区域定位的服务，服务中的尘肺病识别的算法是由本文第二章提出的“层级分类方法”；可疑患病区域定位的算法是第三章提到的“判别性区域发现”。尘肺病识别流程与可疑区域定位流程属于图4.2所示的主流程的子流程，具体如图4.3示。当影像数据经由“图像格式转化”步骤转换为深度学习模型能够使用的数据格式后，经由层级分类算法：首先经过第一阶段的基分类器，得到打分如果低于截断点数值，则判定为0期，并

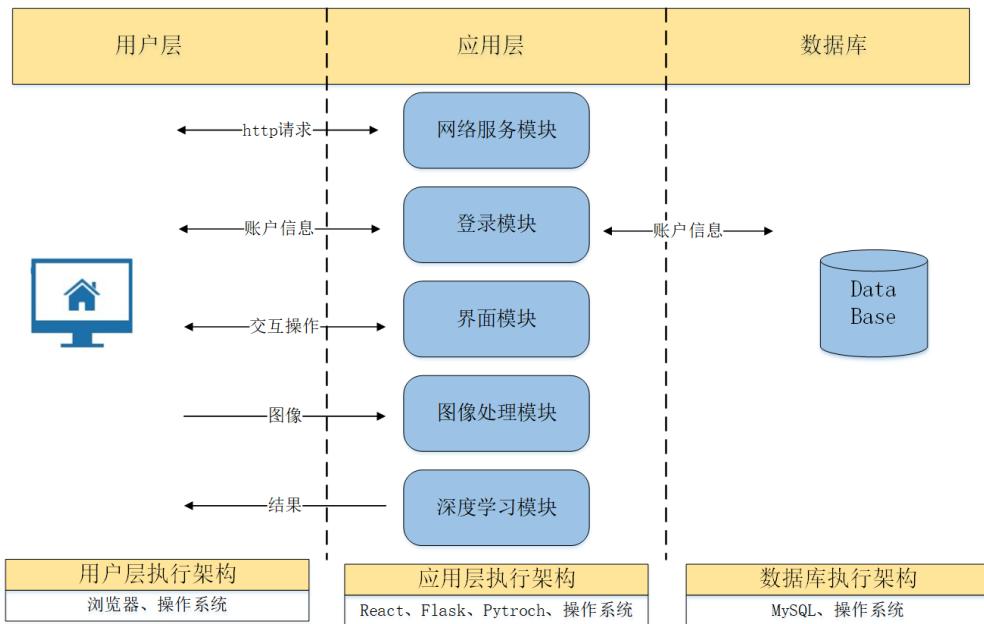


图 4.4: 尘肺病识别系统架构图

再次利用第一阶段基分类器、基分类器判定结果、转换后影像经由判别性区域发现方法得到可疑患病区域结果影像，然后进入下个流程；否则影像接着经过深第二阶段基分类器，得到判定结果，并再次用第二阶段基分类器、基分类器判定结果、转换后影像经由判别性区域发现方法得到可疑患病区域结果影像，然后进入下个流程。

4.3 系统实现

4.3.1 系统架构

经过上一节的需求分析与流程分析，总体架构如图4.4所示。系统由三层构成，分别是用户层，应用层和数据库。用户在最上面用户层，包含了发送http请求访问页面，与系统的前端页面进行交互，账号登录请求，上传影像，获取结果展示等服务。中间的应用层是系统的提供服务的核心，包含了提供网络服务的网络服务模块，提供前端交互页面的界面模块，提供账号验证服务的登录模块，提供对上传影像进行格式转换与预处理的影像处理模块，以及提供利用深度卷积神经网络进行识别与可疑患病区域定位服务的深度学习模块。最后一层是数据库，用户的账号信息储存在数据库中，它给中间层登录模块提供账号信息，以进行校验。数据库对于用户是不可操作的，只能由系统管理员进行账号的增加、修改、删除操作。

表 4.1: 开发软硬件环境

硬件环境	CPU	Intel(R)Core(TM)i7-8750H CPU @ 2.20GHz
	GPU	NVIDIA GeForce GTX 1080 Ti
	内存	32.0 GB
软件环境	操作系统	Ubuntu 16.04
	第三方框架	Pytorch, Flask, React, Anaconda-Python3.6, OpenCV
	深度学习模型	DenseNet121
	数据库	MySQL

4.3.2 开发环境

本系统除了网页前端页面涉及到JavaScript，主要还是基于Python技术栈进行开发，开发和运行环境如表4.1示。

4.3.3 关键技术

Flask

本系统基于B/S架构提供网络服务，由应用层的网络服务模块承担，它的开发基于Flask框架。Flask是一个使用Python 编写的轻量级Web应用框架，依赖于Werkzeug WSGI （Python Web Server Gateway Interface，WSGI）工具箱与Jinja2模板引擎。Flask被称为“microframework”，它主要作用是Python应用程序或框架和Web 服务器之间的一种接口。Flask被称为“微框架”，因为它有着保持核心的简单，但同时又易于扩展的特点。Flask作为Web框架，它仅保留了核心功能：请求响应处理和模板渲染，这两类功能分别由Werkzeug WSGI 工具箱和Jinja2完成。数据库抽象层、表单验证，或是其它任何已有多种库可以胜任的功能Flask都没有，但是它支持用扩展来给应用添加这些功能。Flask的这些特性，使得它在小型Web开发方面变得非常流行。

React

本系统的网络服务中，需要提供网页界面，网页前端服务由应用层的界面模块承担，它的开发基于React框架。React又称React.js或ReactJS，狭义上讲是由Facebook 开源的一个为数据提供渲染为HTML 视图的开源JavaScript库，广义上来讲它不仅是一个用于前端用户界面（User Interface，UI）开发的JavaScript框架，更是一套完整的前端开发生态体系。React 侧重于MVC 开发模式中的View 层，它可以用于Web端，IOS、Android和服务器端的界面开发，多用于Web 界

面开发。

React相比于其他前端框架最重要的特点是虚拟DOM（文档对象模型）和组件化。在Web界面开发中，常常需要将变化的数据实时反应到UI上，这时就需要对DOM进行操作。复杂或频繁的DOM操作通常是性能瓶颈产生的原因。React为此引入了虚拟DOM的机制：在浏览器端用Javascript实现了一套DOM API。基于React进行开发时所有的DOM构造都是通过虚拟DOM进行，每当数据变化时，React都会重新构建整个DOM树，然后React将当前整个DOM树和上一次的DOM树进行对比，得到DOM结构的区别，然后仅仅将需要变化的部分进行实际的浏览器DOM更新，而且React能够批处理虚拟DOM的刷新，有着高效的性能。虚拟DOM不仅带来了简单的UI开发逻辑，同时也带来了组件化开发的思想，所谓组件，即封装起来的具有独立功能的UI部件。React推荐以组件的方式去重新思考UI构成，将UI上每一个功能相对独立的模块定义成组件，然后将小的组件通过组合或者嵌套的方式构成大的组件，最终完成整体UI的构建。因此React有着高效的性能、分离的设计等优点。

Pytorch

本系统提供尘肺病影像识别和可疑患病区域定位服务，由应用层的深度学习模块承担，它的开发基于Pytorch框架。Pytorch是一个由Facebook开源的深度学习框架，专门针对GPU 加速的深度神经网络开发及应用。PyTorch提供了运行在GPU/CPU之上的基础张量操作库，提供了包含各种前沿的神经网络模型的库，支持模型训练功能，还提供了支持共享内存的多进程并发库等。

相比于其他的深度学习框架，Pytorch最重要的特点是：Python 优先和动态计算图（dynamic computational graph）结构。Pytorch的Python优先的设计思想是方便开发者能直接的使用广大的Python库和软件（如NumPy、SciPy和Cython），使之处于机器学习第一大语言Python的生态圈之中。目前主流的深度学习框架分为静态图框架和动态图框架两种，TensorFlow、Caffe、Theano等采用的是静态图框架，Pytorch是动态图框架。静态图框架的网络定义都是声明式的，通过先定义后运行的方式：即需要先搭建好神经网络整个结构，然后再一次又一次使用同样的结构进行计算，不需要再重新构建计算图，所以计算速度会更快；然而静态图架构一旦需要修改网络结构时需要从头重新构建整个网络。动态计算图每次计算都会重新构建一个新的计算图，速度可能会慢，带来的好处是可以随时改变神经网络的结构，而不影响其计算：可以随意的调用函数，可以随意修改结构，也可以使用简单易懂的搭建方法来创建自己的神经网络。因此采用动态图架构的Pytorch有着简单易用、直观清晰、更加利于开发者调试的优点，这些优点使得它变得越来越流行。

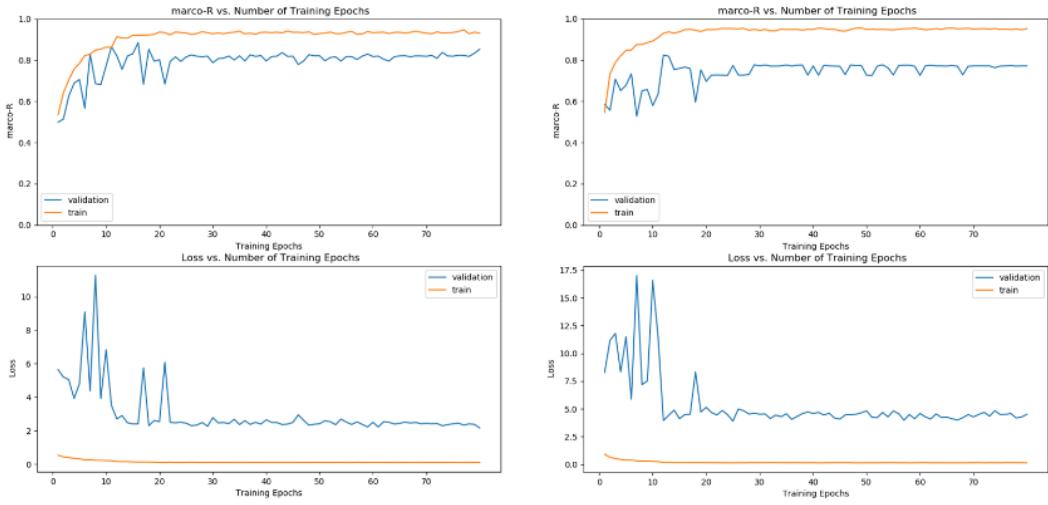


图 4.5: 层次分类方法基分类器训练曲线:第一阶段(左)和第二阶段(右)

4.3.4 系统展示

在搭建完本系统之前，需要已经训练好的深度学习模型，按照第二章层次分类方法训练两个深度卷积神经网络分别作第一阶段基分类器和第二阶段基分类器（分别作为正负分类和1,2,3期分类），具体训练设置参照第二章层次分类方法的实验数据集和实验设置小节。图4.5为第一阶段基分类器与第二阶段基分类器的训练过程曲线。

在模型训练完毕后与尘肺病识别系统一起部署到服务器，绑定域名即可对外提供服务，以下是尘肺病识别系统核心功能的部分展示截图：上传影像服务界面如图4.6示，尘肺病判别以及可疑区域判定结果获取服务界面如图4.7示。

4.4 小结

在本章中，依据合作项目的需求，设计和实现尘肺病识别系统。本系统的核心算法是前两章介绍的层级分类方法与判别性区域发现方法，基于B/S架构，使用的技术是包括了Flask、React、Pytorch等的Python技术栈的流行的技术，实现了用户登录、影像上传、影像分类、结果获取等功能。本系统提供尘肺病分类与可疑患病区域定位的服务，期望能够辅助医疗从业人员更快的判断尘肺病以及能够二次检查减少漏诊误诊的情况发生。

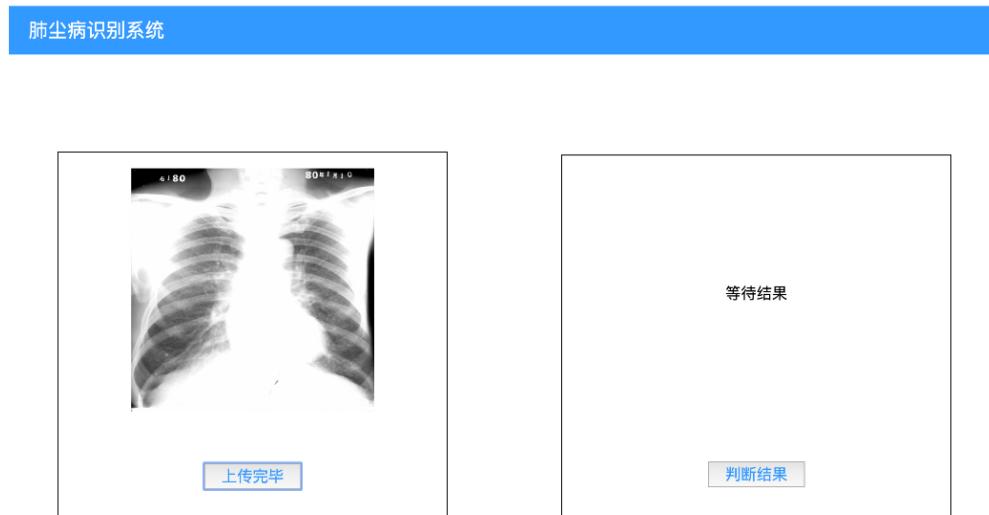


图 4.6: 尘肺病识别系统: 上传界面

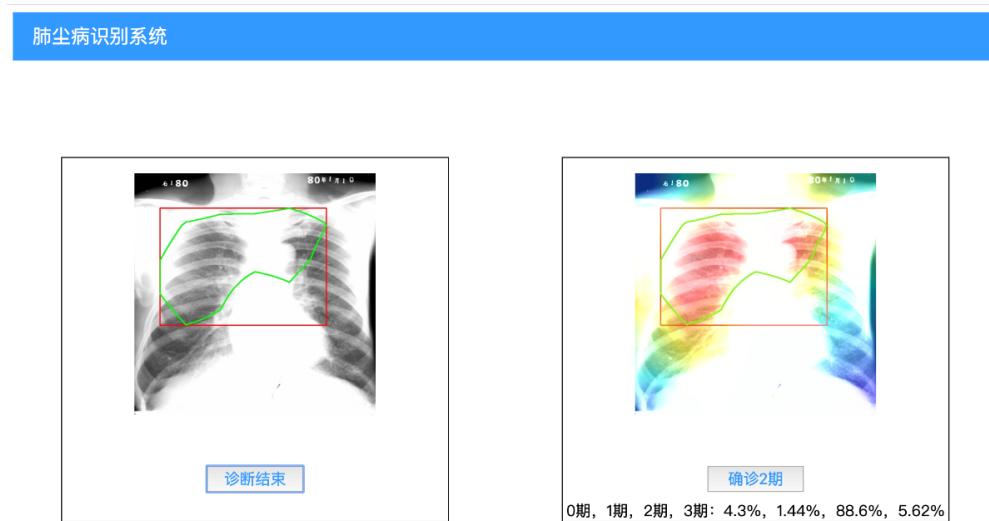


图 4.7: 尘肺病识别系统: 结果界面

第五章 总结与展望

5.1 工作总结

伴随着深度学习技术近些年的突破性进展，在现代的医疗中，基于深度学习方法的计算机辅助诊断技术已经是主流发展趋势。在肺部疾病领域，目前已经有较多的研究集中在使用深度学习对肺结节或肺癌的诊断当中，但是具体到另外一个细分领域，尘肺病，作为我国职业病发病率第一的肺部疾病相关的研究却比较少。研究能智能识别尘肺病有效的算法模型与应用，填补相关细分领域的一些空白，丰富相关研究，是一个十分有意义的课题。本文围绕深度学习方法，进行了尘肺病识别与识别的可解释性研究与应用，本文工作内容如下。

首先，针对人工特征方法对于尘肺病患病期别分类精度较低的问题，提出了基于深度卷积神经网络的层级分类方法。该方法大体分为两阶段，首先判断样本是否患病，然后再细分认为“患病”样本的期别。在该方法的设计中，考虑了两个要点：一是单层级的基分类器要尽可能准确，为了选择好的基分类器，实验中研究分析了多种深度卷积神经网络的分类结果，包括它们加入代价敏感方法调优前后的差别；二是所有层级的基分类器联合起来最终分类结果要尽量准确，为了调优最终效果，在实验中进行了多种截断点取值与最终分类指标结果的对比。经过实验证明，层级分类方法相较于普通的深度学习端到端式多分类方法提升效果明显，宏查全率由0.7723提升到0.9162，层级分类方法能够更好的区分尘肺病第3期样本。

再者，针对尘肺病诊断需要可靠解释的问题，本文基于深度学习可解释性技术找到可疑患病区域给出了分类依据，并利用这些区域的高维特征提升了分类精度。本文研究了两种特征可视化方法来探究尘肺病高维抽象特征。首先是通过Grad-CAM技术，获取到网络特征抽取层的分类权重与高维抽象特征，通过它们找到了每个样本分类过程中可能是结节、弥纤、尘斑等的图像病理区域，并以热力图的方法将这些区域高亮的标记出来，这些高亮区域即分类依据，称为“判别性区域”。再者在判别性区域的发现基础上，通过SCDA技术，在只有期别的标签情况下，将疑似尘斑等图像病理区域的高维特征从高维卷积特征中筛选出来，这种特征称为“细粒度病理特征”。本文对细粒度病理特征进行了进一步探究：降维可视化分析出了尘肺病医学影像各期别的特征区别，发现尘肺病的3期是的瓶颈；使用该特征作为多种学习器的训练数据进行对比，在端到端的方法中尘肺病宏查全率由0.7723提升到0.8857，层级分类方法中尘肺病宏查

全率由0.9162提升到0.9292，经过实验验证了细粒度病理特征的有效性。

最后，依据硕士期间参与的尘肺病识项目，设计和实现了尘肺病识别系统。本系统提供尘肺病分类与可疑患病区域定位的服务，这两个服务的核心算法是层级分类方法与可解释性技术中的判别性区域发现方法。该系统能够辅助医疗从业人员更快的判断尘肺病以及能够二次检查减少漏诊误诊的情况发生。

5.2 工作展望

虽然本文中的工作取得了一定的成果，但是仍然有一些可以改进与提高的地方，以下列举了一些可以继续探究的方面：

设计更加适用于X光片的卷积神经网络。实验中使用的是主流的计算机视觉中图像分类的深度卷积神经网络，但是尘肺病病X光片有本身的特性：例如数据曝光度较高，分辨率较高，胸腔处骨骼（噪声）较多等问题。设计能够适用于这些特点的网络结构或数据预处理流程值得被进一步探究。

利用迁移学习提升算法分类精度。在尘肺病层级分类算法中，实验中多次观察到模型性能瓶颈是第三期分类不够准确，这个原因归结于第三期样本较少。目前世界上已经有更好的开源的肺部X光片数据集，如ChestXray-14，尽管它不包含尘肺病但包括了其他14种肺部病理特征，如何设计合理的迁移学习指标，将其中共通的领域知识迁移过来，增强目标领域内模型的鲁棒性值得我们研究。

更加精确的可解释性指标。在判别性区域发现与细粒度尘肺病特征提取的章节中，本文的可解释性是通过高亮标出判别性区域实现的，这只是一个“弱检测”的方法，依靠人的肉眼来确认，能够量化的可解释性指标有待进一步研究。

参考文献

- [1] Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T.M., Asari, V.K., 2018. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. arXiv preprint arXiv:1802.06955 .
- [2] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S., 2016. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE transactions on medical imaging 35, 1207–1216.
- [3] Antony, J., McGuinness, K., O'Connor, N.E., Moran, K., 2016. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks, in: 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE. pp. 1195–1200.
- [4] Arevalo, J., González, F.A., Ramos-Pollán, R., Oliveira, J.L., Lopez, M.A.G., 2016. Representation learning for mammography mass lesion classification with convolutional neural networks. Computer methods and programs in biomedicine 127, 248–257.
- [5] Armato, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., 2011. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Medical physics 38, 915–931.
- [6] Benou, A., Veksler, R., Friedman, A., Raviv, T.R., 2016. De-noising of contrast-enhanced mri sequences by an ensemble of expert deep neural networks, in: Deep Learning and Data Labeling for Medical Applications. Springer, pp. 95–110.
- [7] Canziani, A., Paszke, A., Culurciello, E., 2016. An analysis of deep neural network models for practical applications. arXiv preprint arXiv:1605.07678 .
- [8] Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S., Feng, J., 2017. Dual path networks, in: Advances in Neural Information Processing Systems, pp. 4467–4475.

- [9] Christodoulidis, S., Anthimopoulos, M., Ebner, L., Christe, A., Mougiakakou, S., 2017. Multisource transfer learning with convolutional neural networks for lung pattern analysis. *IEEE journal of biomedical and health informatics* 21, 76–84.
- [10] Ciompi, F., Chung, K., Van Riel, S.J., Setio, A.A.A., Gerke, P.K., Jacobs, C., Scholten, E.T., Schaefer-Prokop, C., Wille, M.M., Marchiano, A., 2017. Towards automatic pulmonary nodule management in lung cancer screening with deep learning. *Scientific reports* 7, 46479.
- [11] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database, in: *2009 IEEE conference on computer vision and pattern recognition*, Ieee. pp. 248–255.
- [12] Dou, Q., Chen, H., Yu, L., Qin, J., Heng, P.A., 2017. Multilevel contextual 3-d cnns for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering* 64, 1558–1567.
- [13] Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The importance of skip connections in biomedical image segmentation, in: *Deep Learning and Data Labeling for Medical Applications*. Springer, pp. 179–187.
- [14] Gao, M., Xu, Z., Lu, L., Harrison, A.P., Summers, R.M., Mollura, D.J., 2016. Multi-label deep regression and unordered pooling for holistic interstitial lung disease pattern detection, in: *International Workshop on Machine Learning in Medical Imaging*, Springer. pp. 147–155.
- [15] Gao, X.W., Hui, R., Tian, Z., 2017. Classification of ct brain images based on deep learning networks. *Computer methods and programs in biomedicine* 138, 49–56.
- [16] Goyal, M., Yap, M.H., 2017. Multi-class semantic segmentation of skin lesions via fully convolutional networks. *arXiv preprint arXiv:1711.10449* .
- [17] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

- [18] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708.
- [19] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K., 2016. SqueezeNet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size. arXiv preprint arXiv:1602.07360 .
- [20] Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., Biller, A., 2016. Deep mri brain extraction: a 3d convolutional neural network for skull stripping. NeuroImage 129, 460–469.
- [21] Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C.I., Mann, R., den Heeten, A., Karssemeijer, N., 2017. Large scale deep learning for computer aided detection of mammographic lesions. Medical image analysis 35, 303–312.
- [22] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.
- [23] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. Neural computation 1, 541–551.
- [24] Lin, M., Chen, Q., Yan, S., 2013. Network in network. Computer Science .
- [25] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis 42, 60–88.
- [26] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C., 2016. Ssd: Single shot multibox detector, in: European conference on computer vision, Springer. pp. 21–37.
- [27] Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440.

- [28] Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-sne. *Journal of machine learning research* 9, 2579–2605.
- [29] Menegola, A., Fornaciali, M., Pires, R., Avila, S., Valle, E., 2016. Towards automated melanoma screening: Exploring transfer learning schemes. *arXiv preprint arXiv:1609.01228* .
- [30] Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016. Brain tumor segmentation using convolutional neural networks in mri images. *IEEE transactions on medical imaging* 35, 1240–1251.
- [31] Poudel, R.P., Lamata, P., Montana, G., 2016. Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation, in: *Reconstruction, segmentation, and analysis of medical images*. Springer, pp. 83–94.
- [32] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788.
- [33] Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks, in: *Advances in neural information processing systems*, pp. 91–99.
- [34] Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer. pp. 234–241.
- [35] Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization, in: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 618–626.
- [36] Setio, A.A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S.J., Wille, M.M.W., Naqibullah, M., Sánchez, C.I., van Ginneken, B., 2016. Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging* 35, 1160–1169.
- [37] Setio, A.A.A., Traverso, A., De Bel, T., Berens, M.S., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., 2017. Validation, compar-

- son, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. *Medical image analysis* 42, 1–13.
- [38] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *Computer Science* .
- [39] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning, in: Thirty-First AAAI Conference on Artificial Intelligence.
- [40] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2014. Going deeper with convolutions .
- [41] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826.
- [42] Tran, P.V., 2016. A fully convolutional neural network for cardiac segmentation in short-axis mri. *arXiv preprint arXiv:1604.00494* .
- [43] van Tulder, G., de Bruijne, M., 2016. Combining generative and discriminative representation learning for lung ct analysis with convolutional restricted boltzmann machines. *IEEE transactions on medical imaging* 35, 1262–1272.
- [44] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M., 2017. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2097–2106.
- [45] Wei, X.S., Luo, J.H., Wu, J., Zhou, Z.H., 2017. Selective convolutional descriptor aggregation for fine-grained image retrieval. *IEEE Transactions on Image Processing* 26, 2868–2881.
- [46] Wu, G., Kim, M., Wang, Q., Munsell, B.C., Shen, D., 2016. Scalable high-performance image registration framework by unsupervised deep feature representations learning. *IEEE Transactions on Biomedical Engineering* 63, 1505–1516.

- [47] Zagoruyko, S., Komodakis, N., 2016. Wide residual networks. arXiv preprint arXiv:1605.07146 .
- [48] Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks, in: European conference on computer vision, Springer. pp. 818–833.
- [49] Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., Shen, D., 2015. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage 108, 214–224.
- [50] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A., 2016a. Learning deep features for discriminative localization, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2921–2929.
- [51] Zhou, X., Ito, T., Takayama, R., Wang, S., Hara, T., Fujita, H., 2016b. Three-dimensional ct image segmentation by combining 2d fully convolutional network with 3d majority voting, in: Deep Learning and Data Labeling for Medical Applications. Springer, pp. 111–120.

简历与科研成果

基本情况

何甲，男，苗族，1993年6月出生，湖北省来凤县人。

教育背景

2016.9~2019.6	南京大学计算机科学与技术系	硕士
2011.9~2015.6	南京邮电大学物联网学院	本科

攻读硕士学位期间完成的学术成果

- [1] He J, Shang L, Ji H, et al. Deep learning features for lung adenocarcinoma classification with tissue pathology images[C]//International Conference on Neural Information Processing(ICONIP). Springer, Cham, 2017: 742-751.

攻读硕士学位期间参与的科研课题

1. 江苏省自然科学基金面上项目“融合语境的视频行为分析不确定性方法建模及应用研究”(No.BK20161406)
2. 横向项目 基于并行运算架构的尘肺病期别判定

攻读硕士学位期间获奖情况

1. 2017年度南京大学光华奖学金
2. 2017年度优秀研究生

致 谢

入学的时刻仿佛还在眼前，然而时光飞逝，在南京大学硕士求学转眼而至末尾，此刻百感交集，在这里我要向那些帮助过我的人表达诚挚的感谢。

感谢我的导师商琳副教授。三年时间，承蒙商老师的悉心指导，一步步帮助我确定研究方向，为我指明研究中关注的重点，教育我科学的研究中应该秉持的态度，在我遇到难点时支持我、鼓励我，正是她的辛勤付出，我才能顺利快乐的度过硕士的科研生活并完成此篇论文。商老师不仅教学当中尽职尽责，在生活当中也是和蔼可亲，她作为老师，也作为我的长辈，教育我做人做事道理，遇到困难她积极乐观的态度深深的影响了我。三年前因为一个偶然的机会进入商老师的办公室与她面谈，从此能成为她的学生是怎样一种幸运与荣幸。

感谢我的同组的各位师兄师弟师姐师妹，在学习与工作当中给予我的支持与帮助；感谢计算机系的各位老师，正是他们的教导才让我掌握了宝贵的专业知识。

感谢三年交到的新朋友宾义、郇兆鑫、胡瑞、钱浩哲、宁天昊、安蔚、胥棲梧，与他们每一次谈天说地、打的每一场球、吃的每一顿饭都是难以忘记的美好回忆；也感谢我的老朋友邓雪，在本文成文过程中给予我技术上的帮助与指导。

最后，感谢我的家人。父母、爷爷奶奶、姐姐，正是他们的养育与无私的付出，我才能从一个懵懂的孩子经过二十六载的岁月成长为如今的年轻人。

虽然求学生涯已近末尾，但它使我明白，师长、好友、家人正是支撑一个年轻人对这个世界产生疑惑时继续保有期待与热情的内在精神力量。感恩在南大经历的一切，那些学生时代的光，是永远开在心里的花。