



西北工业大学

本科毕业设计论文

题 目 基于深度学习的胸腔 X 光图像的病变差异性识别

专业名称 电子科学与技术

学生姓名 胡小婉

指导教师 王毅

毕业时间 2019.06

摘要

医学 X 光图像是一种根据不同组织对 X 射线的不同吸收程度来反映人体内部结构的成像方式，可以显示多种复杂的病理特征，是目前最常用的放射学检查手段之一。本文研究的胸部 X 光图像主要用于检测多种常见胸肺部疾病，但与传统的图像分类相比，X 光图像在灰度值上的对比度很小，小的病理区域很难被发现，而成像过程中所引起的伪影和畸变很容易淹没临床特征，这给胸部 X 光图像的疾病检测带来了许多困难。

在胸部X光图像分类上，目前仅使用某种单独的学习策略通常不足以训练参数量大、层数多的复杂深度网络，限制了网络训练速度以及分类效果的提升。因此，本文利用深度学习技术，提出了一种**基于双通道分离网络（Dual Channel Split Network, DCSN）**的胸部病变的差异性识别方法。该方法在特征提取过程中，舍弃了单网络结构，而是将输入图像进行多个网络的并行处理，在通道方向上将特征图分离，使其分别通过上下分支的不同网络进行特征学习，并最终将双通道的输出进行融合，用于分类和定位任务。本文提出的网络模型不仅能有效缓解深层模型的负担，同时证实了不同分支的信息共享和集成策略有助于改善网络信息流，加速网络收敛。此外，本文对该网络的性能与多种胸部疾病分类算法进行比较，验证了本文所提出的双通道分离网络框架在十四种胸部疾病的识别任务中，得到的AUC（Area Under Curve）分数更高，表明该网络对胸部病变的识别准确率更高，同时在可视化任务中可以对疾病位置完成较为精准的定位，可作为一种高效的医疗辅助手段，为放射科医生提供可靠的诊断参考。

关键词：X 光图像分类，双通道分离，残差学习，密集连接，特征融合

ABSTRACT

Medical X-ray image is an imaging method that reflects the internal structure of the human body according to the different absorption degree of X-rays of different tissues. It can display a variety of complex pathological features and is one of the most commonly used radiological examination methods. The chest X-ray images studied in this paper are mainly used to detect a variety of common chest and lung diseases, but compared with the traditional image classification, the contrast of the X-ray image on the gray value is small, and the small pathological area is difficult to be found. The artifacts and distortions caused by the imaging process can easily overwhelm the clinical features, which bring many difficulties to the detection of chest X-ray images.

In the classification of chest X-ray images, the use of only a single learning strategy is usually not enough to train complex deep networks with large parameters and many layers, thus limiting the improvement of classification effect. Therefore, this paper proposes an image classification method based on Dual Channel Split Network (DCSN) by using deep learning technology. In the feature extraction process, the method discards the single network structure, and performs the parallel processing of the input image on multiple networks, and separates the features in the channel direction, so that the feature learning is performed through different networks of the upper and lower branches, respectively. The two-channel output is finally fused for classification and positioning tasks. The network model proposed in this paper can not only effectively alleviate the burden of deep models, but also prove that information sharing and integration strategies of different branches can improve network information flow and accelerate network convergence. This paper compares the performance of the network with a variety of chest disease classification algorithms, and verifies that the proposed network's AUC (Area Under Curve) score is higher, which shows that the network has higher recognition accuracy for chest lesions, and the position of the disease can be accurately positioned. We believe that this algorithm model can be used as an effective medical aid to provide a reliable diagnostic reference for radiologists.

KEY WORDS: X-ray image classification, DCSN, residual learning, dense

connection, feature fusion

目 录

第一章 绪论	7
1.1 研究背景	7
1.2 研究目的和意义	8
1.2.1 研究目的	8
1.2.2 研究意义	8
1.3 研究现状分析	9
1.4 本文章节安排	11
第二章 神经网络概述	13
2.1 卷积神经网络	13
2.2 网络结构	13
2.3 分层解读	14
2.3.1 卷积层	14
2.3.2 激活层	14
2.3.3 池化层	15
2.3.4 全连接层	15
2.4 反向传播算法	16
2.5 本章小结	16
第三章 网络模型设计	17
3.1 残差学习网络 (ResNet)	17
3.1.1 残差学习	17
3.1.2 网络结构	18
3.2 密集学习网络 (DenseNet)	19
3.2.1 密集连接	19
3.2.2 网络结构	20
3.3 两种网络的比较	21
3.4 深度双通道分离网络 (Dual-Channel Splitting Network, DCSN)	20
3.4.1 双通道分离网络结构	20
3.4.2 双通道分离模块	22
3.5 病变检测可视化	24
3.6 本章小结	25
第四章 实验过程及结果	27

4.1 实验环境.....	29
4.2 数据集.....	27
4.2.1 数据集介绍.....	27
4.2.2 数据详解.....	29
4.2.3 数据预处理.....	29
4.3 评价指标.....	32
4.4 实验 I：基于残差学习网络的胸部病变差异性识别	31
4.4.1 实验内容.....	31
4.4.2 实验结果.....	31
4.5 实验 II：基于密集连接网络的胸部病变差异性识别	37
4.5.1 实验内容.....	37
4.5.2 实验结果.....	37
4.6 实验 III：基于通道分离双网络（CSDN）的胸部病变差异性识别.....	41
4.6.1 实验内容.....	41
4.6.2 实验结果.....	41
4.7 实验结果可视化.....	41
4.8 本章小结.....	43
第五章 全文总结	47
5.1 结论.....	47
5.2 未来展望.....	47
参考文献.....	49
致 谢.....	51
毕业设计小结	53

第一章 绪论

1.1 研究背景

近年来,医学影像作为医学技术中的重要领域之一,作为一种医疗辅助手段用于诊断和治疗,已经成为目前医学检测项目不可缺少的步骤,因而发展十分迅速。而随着现代科学技术的进步,跨学科交叉研究与整合已成为现代科学发展的一个突出特征和重要途径。随着新的成像技术和设备的出现和计算机技术的发展,医学图像处理技术在医学研究和临床实践中发挥着越来越重要的作用^[1]。

胸部 X 射线 (Chest X-ray, CXR) 成像是目前用于筛查和临床诊断的最广泛可用的放射学检查之一。然而,由于图像上不同类型的肺部病变的复杂病理,因此对 CXR 图像进行自动检测并理解已经成为目前在技术上具有挑战性的任务。因为胸部 X 光射线成像技术不仅快捷、简便,而且价格经济,所以人们在检查胸廓(包括肋骨,胸椎,软组织等)、胸腔、肺组织、纵隔、心脏等的胸部疾病时,通常都会选择 X 光。CXR 目前是诊断肺炎等疾病的最佳检测项目^[2],在临床诊断及护理^{错误!未找到引用源。}和流行病理研究^{错误!未找到引用源。}中发挥了关键作用。而治疗效果与发现的早晚密切相关,由于个体差异,这些症状在发病前并不一定明显。传统的医院放射诊断方法可能无法及时准确地发现相应的症状,为医生提供决策依据。导致患者无法及早得到治疗,错过了最佳治疗时间。据研究调查,仅在美国就有超过 100 万成年人因肺炎住院,每年约有 50,000 人因为患有该类疾病而死亡^[5]。所以,利用先进的计算机技术对胸部 X 光图像进行检测与分析,提高胸部疾病的确诊率是目前医学图像处理领域需要解决的一大问题。

计算机辅助诊断 (Computer Aided Diagnosis, CAD) 方法始于 60 年代^[6] 但在那段时间没有取得很大成功。大规模的 CAD 使用在 80 年代到来,采用了不取代医学专业人员的新方法,但只能帮助他们进行诊断^[7]。最近计算能力的增长使得卷积神经网络 (Convolutional Neural Networks, CNN) 应用在图像的识别和检测上得以扩展^[8],特别是在 2012 年引入 AlexNet 之后^[9]。这种增长也适用于使用卷积神经网络对患者进行分类的 CAD 系统。过去十年以来,作为一项有前途的技术,医学人工智能 (Medical Artificial Intelligence) 在全球范围内引起了人们的兴趣。特别是在深度学习的出现和急速发展之后,计算机辅助诊断技术的革命正式开始并在许多生物医学应用中受到影响。

在临床实践中,基于 CXR 图像的分析 and 诊断在很大程度上取决于具有至少多年专业经验的放射科医师的专业知识。因此,如果开发一种能够仅通过读取患

者的 CXR 图像来自动检测不同类型的胸部疾病的计算机辅助系统，对医学诊断领域来说影响重大。这一切都建立在精心设计的人类知识转移到机器智能的基础之上。但目前利用计算机检测 CXR 图像上的胸部疾病仍然是机器智能的一项具有挑战性的任务，主要难点如下：

1) 缺乏训练数据集以及图像的相关注释/标签，专业知识限制导致标注难度高。此外，就算有专家标注的数据，存在的标签噪声也会产生干扰。

2) 从不同类型的胸部疾病样本中提取的视觉模式通常在外观、大小和位置上具有高度多样性，传统图像处理方法对嘈杂标签监督不力而无法学习有意义的表征。

3) 图像类别不平衡，数据库中通常正常图像比异常图像多很多，导致临床特征容易被淹没。

4) CXR 图像的空间尺寸通常为 2000×3000 像素。局部病理图像区域可以显示出不同的大小或范围，但是与全图像比例相比通常非常小，病变区域难以检测，在 X 射线图像上训练分类器比通用图像更难。

1.2 研究目的和意义

1.2.1 研究目的

在放射学研究中，CXR 作为在胸部疾病诊断的临床实践中最常用和最经济的筛查技术，在放射学检查中得到广泛应用。数以千计的 CXR 图像被医院捕获，CXR 图像的自动分析将有效地帮助临床诊断和病理学发现。因此，利用当下流行的机器智能及深度学习算法开发一种高效率高准确性的胸部疾病分类检测网络至关重要。

本文的研究目的为：

1) 利用深度学习，构建一种新的深度神经网络模型，完成对肺扩张、心脏肿大、肺积液、浸润性肺炎、肿块、肺结核、肺炎、气胸、巩固性肺炎、浮肿、肺气肿、纤维化、胸膜增厚和疝气十四种疾病的检测。

2) 对于输入的胸部 X 光片图像，模型可以对其存在的疾病类型进行分类，同时将图像中最能预示患有疾病的区域可视化，对诊断出的疾病位置进行标注，便于医生进行二次诊断。

3) 探索一种完整的计算机辅助系统模型来检测并诊断出常见的胸部疾病类型，提高诊断准确率。

1.2.2 研究意义

由于 X 光图像灰度值差异小且病变区域难以肉眼辨别，导致现有的基于人工来识别胸片存在效率低下、精度不高以及误诊、漏诊的问题，故 X 光图像的自动分析与检测成为支持临床诊断和治疗计划的一项越来越重要的技术。利用人工智能算法和医学图像处理技术对胸部 X 光图像的诊断提供一种可靠的分类和

定位方法，不仅能够高精度地提供分类结果，还可以标注出相关的视觉区域，方便医生二次诊断。在提高疾病检测准确率的同时大大缩短了诊断时间，通过计算机辅助手段，提高检测的速度和精度，降低胸部疾病的漏诊和误诊的几率，改善医疗服务质量，实现早诊断、早发现、早治疗，使相关胸部疾病的死亡率大大降低，这项研究在常见胸部疾病的医学诊断和分析上具有极其深远的意义。

1.3 研究现状分析

CXR 是肺和心脏病诊断中最常见的放射学检查之一。目前，阅读胸部 X 光图像主要依靠专业知识和仔细的人工观察。由于图像中不同肺部病变的复杂病理和细微纹理变化，即使经历了长期临床培训和专业指导，放射科医师也可能犯错误。因此，开发一种自动的医学图像分类方法以支持临床实践者是十分必要的。

基于 CXR 扫描产生的各种数据，已经进行了许多研究以自动检测胸部疾病。已经有不少利用处理传统自然图像的方法进行的探索，查普曼等人讨论了手工制定规则、贝叶斯网络和决策树三种专家系统在识别 CXR 报告时的表现^[10]，使用概率输出时，专家系统的表现与放射科医生相似，但在采用分类器的标准评价指标时，对贝叶斯网络和决策树的输出进行的二进制转换明显降低了系统的性能。在此之后，Ye 等人通过模糊阈值方法并引入加权支持向量机（Support Vector Machine, SVM）分类器减少 CXR 图像的肺结节分类中的假阳性^[11]，提高了部分结节检测的准确率，但初步的粗糙分割大大降低了 SVM 分类的效果。由于传统方法提取特征的表达能力不足，在医学图像的细节检测任务中有一定局限性，效果上还有很大的提升空间。

深度学习技术自从在 2006 年提出以来^[12]，其超高准确率和高效率已经在多个图像识别任务的效果上完成了质的提升，在多个自然图像数据集上已经表现出超越人类的能力。在医学图像领域，也提出了许多关于深度学习和大型数据集的最新进展算法，深度学习在医学图像分析领域中针对对象（通常是放射学成像中的人体解剖学或病理学结构）检测和分割任务产生明显的性能提升。2016 年谷歌利用深度卷积神经网络在自动识别糖尿病性视网膜病变上取得重大进展^[13]，其算法性能与专业医生齐平，2017 斯坦福大学在 Nature 上发表的关于皮肤癌检测^[14]的文章，证实了利用 Inception V3 的方法在皮肤癌分类任务上表现优异，甚至超过人类的识别能力，同时，胸部 X 片的自动诊断也引起了相关肺结核分类算法^[15]和肺结节检测^[16]的研究人员的注意。深度卷积神经网络(Convolutional Neural Networks, CNN)在细粒度对象分类的各种通用和高变任务中显示出巨大的潜力。表明基于深度学习的机器学习算法对医学图像的检测具有高灵敏度和特异性。

深度学习在医学图像领域的快速发展导致相关研究人员对医学数据集的需求日益增加，数据集容量的大小及质量的好坏将直接影响训练出的深度学习网络

在各个任务上的性能。在胸部医学图像方面，第一个大型数据集是 Shiraishi 在 2000 年发布的 JSRT 数据集^[17]，仅包含 247 张 X 光图像，每张图像只有是否包含肺结节的单标签。之后，谷歌在 2015 年发布了含有 7470 张的多标签胸部疾病数据集，即 OpenI^[18]，包含 8 种胸部疾病。美国国家癌症研究所（National Cancer Institute）同期发布了 LIDC-IDRI 肺结节公开数据集^[19]。在 CXR 图像分类中，美国国家卫生研究院（National Institutes of Health, NIH）2017 年公开的医院级数据库 ChestX-ray14^[20]，包含 14 类图像标签，每个图像标签对应一种或多种胸部疾病。此外，NIH 临床中心在 2018 年公开发布的大型 CT 图像数据集 DeepLesion^{错误!未找到引用源。}为 CT 图像的临床检测做出了巨大贡献。值得注意的是，大部分数据集中的图像标签并非直接来自病理学家的手动注释，而是由相关的放射学报告中的自然语言处理技术开采。因此，训练集中的类标签是嘈杂的，这不可避免的给疾病分类任务带来一些额外的挑战。

有了相关数据集的支持，目前深度学习的一些特定算法在各种医学成像任务中的实验结果表现优异，甚至已经能够超越医学专业人员，比如糖尿病视网膜病变检测^[13]、皮肤癌分类^[14]、心律失常检测^{错误!未找到引用源。}和出血鉴定^[23]。在胸部 X 光图像的深度学习研究中，Wang 等人^[20]将分类器连接到完全卷积网络作为 ChestX-ray14 数据集图像上胸部疾病分类的特征提取器，其中他们对比了四种经典的 CNN 架构，即 AlexNet^[24]、VGGNet^[25]、GoogLeNet^[26]、ResNet^[27]，以使用全局图像来表示存在多种病理。此外，使用相同的网络，疾病病变区域以弱监督的方式定位。将胸部 X 光图像分类视为多标签识别问题，后来，Yao 等人^[28]调查了 ChestX-ray14 数据集中 14 个病理类标签之间的隐藏相关性。使用 DenseNet^[29] 的变体作为图像编码器，他们采用长短期记忆网络（Long Short-Term Memory, LSTM）^[30] 来捕获图像数据之间的依赖性。库马尔等人^[31]研究哪种损失函数更适合从头开始训练 CNN 并呈现用于全局图像分类的增强级联 CNN。最近，Rajpurkar 等人提出了新的框架 CheXNet^[32]，它对 ChestX-ray14 图像上修订后的 121 层 DenseNet 进行了微调，并在胸部疾病检测方面取得了当前最先进的性能。此后，又提出了大量新型的网络结构，包括引入注意力机制的类别残差学习网络（Category-wise Residual Attention Learning, CRAL）^[33]，以及李飞飞团队提出的一种利用有限监督方法进行疾病识别和定位的方法，只使用少量注释就能得到很好的结果^[34]，还有一些对图像处理方式的探索，结合全局和局部特征，使网络能像医生一样聚焦于病变区域^[35]，以上方法均取得了较好的效果。

虽然现有的很多深度学习方法在胸部 X 光图像分类上已经取得了比较好的成果，但上述大部分方法在对图像进行特征提取时，均选择只采用单个网络及分类层的方法，不可避免地降低了网络对不同通道之间层次特征的提取及特征表示能力，同时也限制了网络的加深以及训练的效率。因此，本文探索了一种全新的

网络模型来对输入的图像特征进行建模，将单网络变成多通道的网络并行分支，采用多种学习策略融合训练的方法，实现更精确的分类检测效果。

1.4 本文章节安排

论文全文共分为六章。

第一章是绪论，简单的介绍了胸部 X 光图像的疾病检测技术研究课题的研究背景、目的与意义，概述了此类研究在国内外的的发展近况，以及具体内容的章节安排。

第二章对深度学习中的卷积神经网络及其基础理论进行了介绍，概述了网络结构组成及模型训练方法，并对其具体过程进行了阐述。

第三章介绍了实验中网络模型的设计过程，解释了残差学习和密集连接两种单独网络的工作原理及过程，之后对双通道的新型网络结构进行图示及说明。

第四章介绍了具体的实验过程和结果，包括实验的前期准备以及原始图像的预处理，通过使用第三章所介绍的方法，验证了三种不同的网络结构在具体 X 光图像分类任务上的表现，并使用 Python 工具包绘制多种曲线图进行结果对比和展示。

第五章为总结章节，对全文的工作进行了概括，阐述了本文实验存在的一些不足与该领域仍存在的主要难点，并提出了未来的研究方向与展望。

第二章 神经网络概述

2.1 卷积神经网络

卷积神经网络是深度学习的代表算法之一，它是一种包含卷积计算并且具有深度结构的前馈神经网络，在计算机视觉领域，卷积神经网络允许从输入数据的多层抽象中提取有用信息，这种多层学习架构对图像的细粒度特征表现出强大的提取和学习能力，使得该网络在图像分类和目标检测方面表现出优越的结果，逐渐成为当下最常用的图像处理方法。

2.2 网络结构

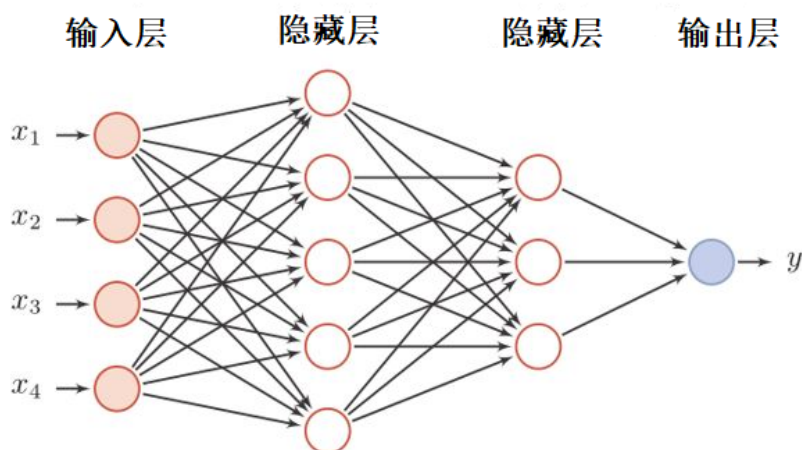


图2-1 典型卷积神经网络结构

典型的神经网络由输入层，输出层和多个隐藏层构成，其中每一层都由多个神经单元组成，如图 2-1 所示。

输入层接受数据输入，包括一维数据或多维数据，根据神经网络的维度及数据特征决定，比如处理一张彩色图像，输入的就是包含 RGB 通道的三维数据，通常在输入层之前对数据进行预处理，标准化的输入特征处理不仅有利于提高算法的运行效率，在增强学习表现方面也有一定作用。

隐藏层中包含卷积层、激活层、池化层和全连接层 4 类，分层提取不同级别的特征并进行线性整合和非线性激活，将输入数据进行变换并提取内在联系，对图像信息进行抽象及建模，为分类和检测提供依据。

输出层通常是一个分类层，对隐藏层提取出的抽象表征进行变换，并根据不

同的任务目的，有针对性地进行不同数学处理和计算，学习出不同模型，最终输出代表图像种类的概率或目标位置。

2.3 分层解读

2.3.1 卷积层

卷积层中一般包含多个卷积核，他们的作用一般是对输入数据进行特征提取和变换，其中，每个卷积核一般由权重参数和偏差参数组成，与前向神经网络的内部神经元结构有一定相似之处，每一个神经元都与前层的多个神经元建立连接通路，卷积计算的“感受野”尺寸大小对输出数据的维度有重要影响，一般由卷积核的尺寸大小来决定，通常按照式 2-1 在单个神经元的感受野内对输入特征做矩阵元素计算：

$$I^{l+1}(i, j) = [\mathbf{I}^l \otimes \mathbf{w}^l](i, j) + b = \sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [I_k^l(s_0 i + x, s_0 j + y) w_k^{l+1}(x, y)] + b \quad (2-1)$$

$$(i, j) \in \{0, 1, 2, \dots, L_{l+1}\}, L_{l+1} = \frac{L_l + 2p - f}{s_0} + 1$$

其中， b 为偏差量， I^l 和 I^{l+1} 表示第 $l+1$ 层的卷积输入和输出， L_{l+1} 为 I^{l+1} 的尺寸， $I(i, j)$ 对应特征图的像素， K 为特征图的通道数， f 、 s_0 和 p 是卷积层参数，对应卷积核的大小、卷积步长和填充步数。

小的卷积核可抓取小区域的局部细节特征，大的卷积核可抓取较大区域的全局特征，之后进行线性整合，更深层次的特征图可通过使用多级卷积操作来得到，同时，利用“权值共享”原可有效减少权重的参数量，便于学习和训练。

2.3.2 激活层

激活层通常由激活函数组成，主要作用是用来模拟输入的细微变化，在这里引入的非线性因素，可以大大提高网络的表示和可学习能力，他引入了多层网络模型的第一个非线性结构，理想的激活函数通常需要具备非线性、连续可微、函数单调、范围内不饱和以及原点处近似线性等数学性质，但现在常用的激活函数大多只满足上述部分要求。

$$A_{i,j,k}^l = f(\mathbf{I}_{i,j,k}^l) \quad (2-2)$$

其中，函数 f 代表激活层函数，通常使用 ReLU 函数^[36] 和 Sigmoid 函数^[37]。

图 2-2 为常用激活函数的曲线特征。

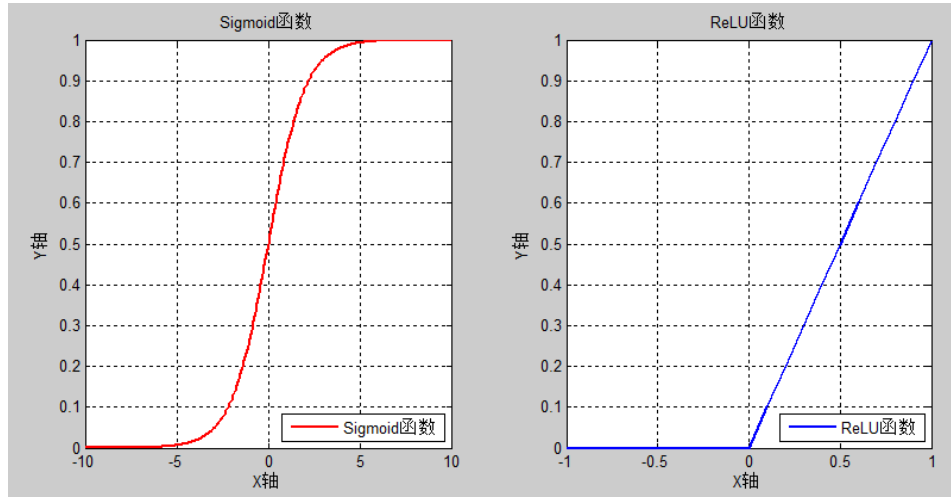


图2-2 常用激活函数

2.3.3 池化层

输入数据先在卷积层进行特征提取和激活之后，将输出特征转移到池化层，进行下一步特征选择和信息过滤，这种处理可以在为位置和尺寸改变之后仍然保持一定程度的不变性，同时不同层次的特征可以在特征图内部进行高度聚合响应，一方面，使特征图的尺寸减少，网络计算复杂度被大大降低；另一方面将大量特征进行压缩来实现主要特征提取。目前使用最多的池化操作包括平均池化（Average pooling）和最大池化（Max pooling）。

$$A_k^l(i, j) = \left[\sum_{x=1}^f \sum_{y=1}^f A_k^l(s_0 i + x, s_0 j + y)^p \right]^{\frac{1}{p}} \quad (2-3)$$

式 2-3 中步长 s_0 、像素 (i, j) 的含义与卷积层相同， p 为指定参数，当 $p=1$ 时，为平均池化，当 $p \rightarrow \infty$ 时，为最大池化。所有的卷积神经网络中基本都会包含池化操作，因为它不仅可以输出特征向量在维度上进行缩小，减少计算量，同时还可以改善整个网络的训练结构，使不易出现过拟合并表现出较好的泛化性能。

2.3.4 全连接层

卷积神经网络的最后一层一般为全连接层，“全连接”指前层全部的神经元都与下一层所有神经元有连接通路，如图 2-3 所示，连接前面几层处理得到的所有的特征，作为分类层，将最终输出值送给指定种类的分类器函数。全连接层的作用为：将学到的“分布式特征表示”映射到样本空间中，减少不同的特征位置对最终分类结果带来的影响，特征图的三维结构在该层的操作中丧失，被展开成向量传给不同分类器，如逻辑函数（Logic）或归一化指数函数（Softmax），然后利用损失函数来控制传播误差，不断改进网络表现，最终输出目标的分类预测结果。

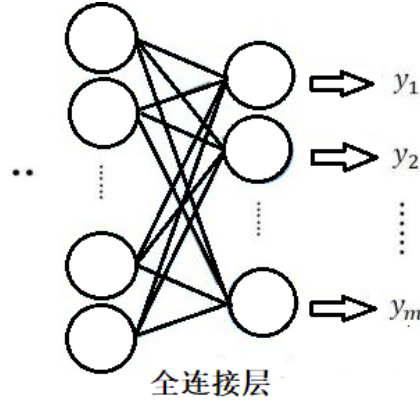


图2-3 全连接层图示

2.4 反向传播算法

卷积神经网络在监督学习中通常使用反向传播框架（Back Propagation, BP）进行学习，反向传播算法以误差为主导。它的最终目的是为了得到最优的全局参数矩阵，从而在多种特定的分类或回归任务中应用多层神经网络。

该算法主要包含两个步骤：激励传播和权重更新。在算法实现中，这两步进行反复迭代，输入信号的前向传输直到输出产生误差，然后误差信息的反向传播更新权重矩阵，最后在双向信息流中优化权重。直到网络对输入的响应达到预定的目标范围，这种双向的传播算法可以说是梯度下降在链式法则中的应用。

全连接层的 BP 计算与传统的前馈神经网络相同，卷积层的反向传播是一个与前向传播类似的交叉相关计算：

$$\begin{aligned} \left(\frac{\partial E}{\partial A}\right)_{i,j}^l &= \sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [w_k^{l+1}(x, y) \left(\frac{\partial E}{\partial A}\right)_{s_0^l+x, s_0^l+y, k}^{l+1}] f'(A_{i,j}^l) \\ w^l &= w^{l+1} - \alpha \left(\frac{\partial E}{\partial w}\right)_k = w^{l+1} - \alpha [A^{l+1} \left(\frac{\partial E}{\partial A}\right)_k^{l+1}] \end{aligned} \quad (2-4)$$

式中， E 为代价函数（cost function）计算的误差、 f' 为激励函数的导数、 α 是学习速率（learning rate），卷积神经网络的误差函数可以有多种选择，常见的包括 Softmax 损失函数（Softmax Loss）、铰链损失函数（Hinge Loss）、三重损失函数（Triplet Loss）等^[38]。

2.5 本章小结

本章主要对深度学习中经典的神经网络算法进行概述，介绍了卷积层、池化层和全连接层的具体实现过程和作用，以及解释了最基础的反向传播算法，了解 CNN 中每层的细节以及各层之间的联系，明确在反向传播中网络的信息如何流动，特征如何提取，以及各层的权重参数如何迭代更新，是设计并训练一个新的深度神经网络的开始。

第三章 网络模型设计

前两章主要介绍了深度学习发展背景以及构建 CNN 的基础算法，本章将详细介绍如何构建新的神经网络模型。在 3.1 和 3.2 节中分别介绍了目前流行的两种网络模型，这两种网络在现有的视觉任务中均表现出良好的性能。之后，在 3.3 节中，将通过融合上述两种网络的特征，设计一种新的并适用于 CXR 图像的深度网络模型。

3.1 残差学习网络（ResNet）

3.1.1 残差学习

深度卷积网络自然地融合了低、中、高层次的不同特点，通过深化网络层次，丰富了其特征内容。因此，在构造卷积网络时，提取的特征层次随着网络的深度的增加而变得更加丰富。然而，梯度消失、梯度爆炸和网络退化等问题在使用深层网络结构训练中出现的更加频繁，极大地影响了网络的性能和训练速度，如图 3-1(a)所示，随着层数的增加，在训练过程中误差随之增加，而(b)图中测试误差的增加显示了网络性能的退化。网络深度的增加所引起的梯度消失/爆炸导致无法收敛，即使通过标准的初始化和正则化层解决了收敛问题，网络的退化却引入了更大的误差。

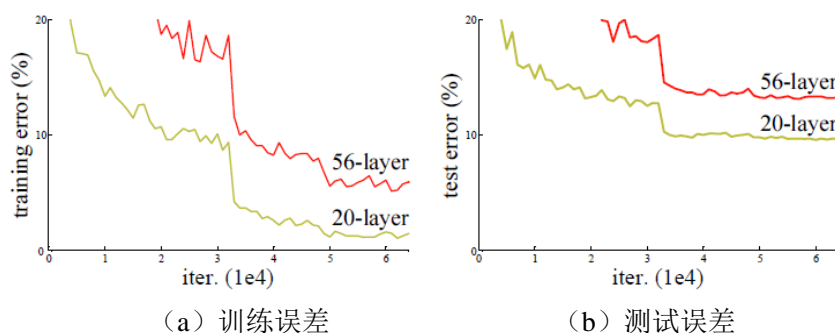


图3-1 深度网络的性能退化

2015 年提出的深度残差学习网络^[39]（ResNet）就是为了解决网络深度变深以后的性能退化问题，该结构的重点在于在浅层网络的基础上叠加 $y = x$ 的层（称 Identity Mappings，恒等映射）来建立层间的短路连接，可以让网络避免因深度增加而带来的退化现象，多层非线性网络存在无法逼近恒等映射网络的缺点，而残差学习的方法解决了上述问题，使网络在训练中表现出更好的性能。

3.1.2 网络结构

ResNet 的思想是假设一个网络层的结构层次为最优，而对于深层次网络中的冗余层，希望它们能够拥有恒等映射的性质，使输入和输出保持一致，ResNet 学习的是残差函数：

$$F(\mathbf{x}) = H(\mathbf{x}) - \mathbf{x} \quad (3-1)$$

这里如果 $F(\mathbf{x}) = 0$ ，则为恒等映射，事实上 ResNet 是“短路连接”的恒等映射下的特例，它不引入多余的参数和计算复杂度。如果优化目标函数的输出是实现一个恒等映射，而不是一个零映射，那么选择学习输入的扰动对该恒等映射的变化量输出，比重新学习新的映射函数要容易更多。

网络由多个残差块（Building）组成，结构如图 3-2，一个残差块有两层，定义为：

$$F = \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{x}) \quad (3-2)$$

其中， σ 代表非线性函数 ReLU， \mathbf{W}_1 和 \mathbf{W}_2 代表每层的参数矩阵。

然后通过一个 shortcut 连接和第二个 ReLU 函数，获得输出：

$$\mathbf{y} = F(\mathbf{x}, \{\mathbf{W}_i\}) + \mathbf{x} \quad (3-3)$$

这里 \mathbf{x} 和 \mathbf{y} 分别表示所计算层的输入向量和输出向量， \mathbf{W}_i 代表每层要学习的参数，函数 $F(\mathbf{x}, \{\mathbf{W}_i\})$ 表示要学习的残差映射的响应函数。操作 $F + \mathbf{x}$ 通过 shortcut 方式连接并逐元素相加。

公式 3-3 中 \mathbf{x} 和 F 的维数必须保持相同，如果情况并非如此（例如输入或输出通道发生改变），我们可以通过引入短路连接操作来解决，如此可执行线性投影 \mathbf{W}_s 来使尺寸得到匹配：

$$\mathbf{y} = F(\mathbf{x}, \{\mathbf{W}_i\}) + \mathbf{W}_s \mathbf{x} \quad (3-4)$$

在公式 3-3 中也可以引入方阵 \mathbf{W}_s ，但实验表明，恒等映射在退化问题的解决上表现良好，并且可减少大量参数计算，所以 \mathbf{W}_s 主要被用于匹配维度。

实验证明，要想在在结果上有明显提升，则通常需要两层以上的残差块（Building），如图 3-2 所示，仅一层的残差块（形式为 $\mathbf{y} = \mathbf{W}_1 \mathbf{x} + \mathbf{x}$ ）并没有在性能上表现出明显提升。

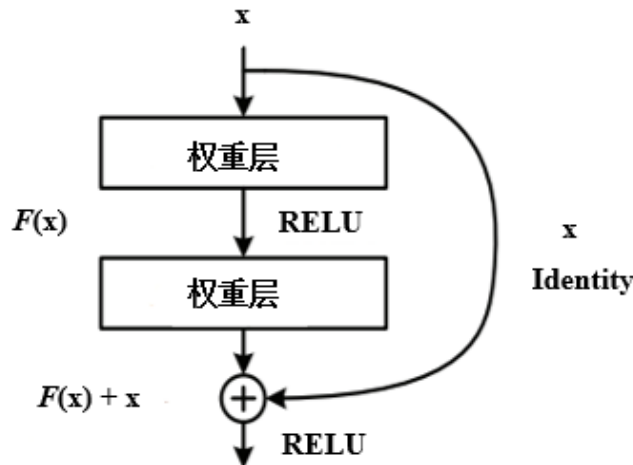


图3-2 残差学习：一个 Building 块

ResNet 的优势，是它用经济高效的算法解决了当今训练深层网络的过程中广泛存在的网络退化问题和梯度消失问题。使用残差网络结构 $H(\mathbf{x}) = F(\mathbf{x}) + \mathbf{x}$ 代替 $H(\mathbf{x}) = \mathbf{x}$ ，增加一个捷径连接，可知对于学习 $H(\mathbf{x}) = \mathbf{x}$ ，选择学习 $F(\mathbf{x}) = 0$ 更加容易，更新冗余层的参数时也更加方便。在反向传播更新参数时，这种残差结构也避免了梯度变为零的情况，不会导致梯度消失问题。

3.2 密集学习网络（DenseNet）

3.2.1 密集连接

随着训练的卷积神经网络的层数和深度不断增加，在网络训练过程中，传递的前向信号和梯度信息经过多层后可能逐渐消失。为了解决这类问题，使信号在输入和输出层之间高速流动的一种旁路(skip-layer)技术被提出，并且在 Highway 和 Resnet 体系结构中最先得到应用。该技术的核心思想是通过一个跨层连接来连接网络的前后层。因此，在此核心概念的基础上，提出了一种新的连接模型。为了使网络各层之间的信息流最大化，对网络中的所有层都执行成对连接操作，使得网络中的每一层都能接受在他之前所有层的特征作为输入，即一种密集连接网络^[29]（DenseNet）。如图 3-3 所示，其中 \mathbf{x}_i 为各层特征图， H_i 为中间层。

密集学习主要拥有以下两个特性：

1) 在一定程度上解决网络训练过程中存在的梯度消失的问题。因为从图 3-3 可以看出，在反向传播时每一层都会结合之后所有层的梯度信息，所以前层的梯度不会随着网络深度的增加而逐渐消失。

2) 由于该结构可以进行特征复用，所以大量的特征的获取只需要较少的卷积单元，最终得到的网络模型也有较小的尺寸，参数量较少，节省计算。

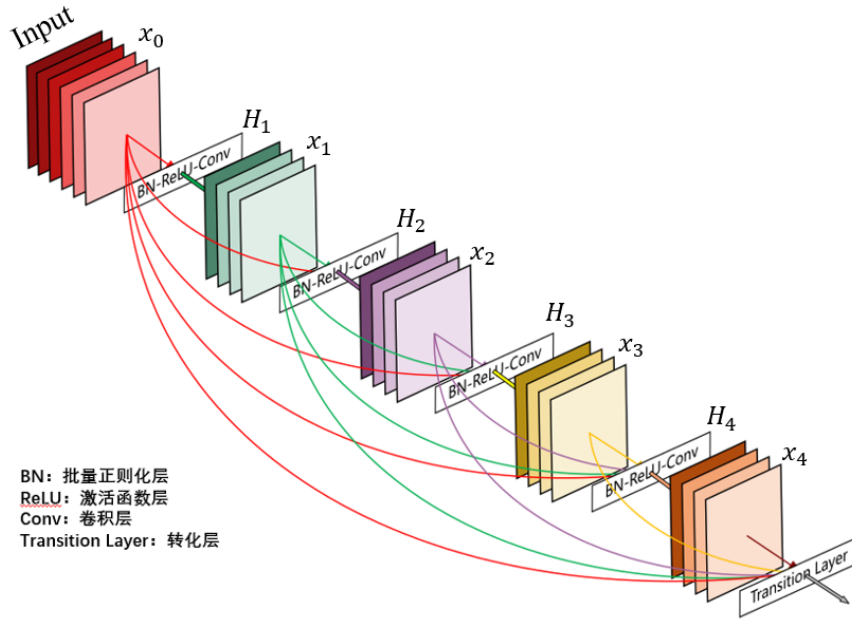


图3-3 5层的 DenseBlock 结构示意图

3.2.2 网络结构

DenseNet 的网络结构如图 3-4 所示，其中 DenseBlock 的内部细节见图 3-3，转换层包括一个卷积层和一个池化层以调整通道数量，可根据不同的网络深度进行多级操作。

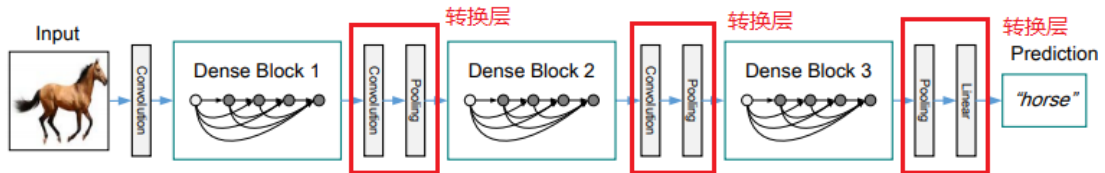


图3-4 DenseNet 网络结构图

图 3-4 中的 DenseBlock 输出：

$$\mathbf{X}_l = H_l([\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_{l-1}]) \quad (3-5)$$

式 3-5 中， \mathbf{X}_l 指代 DenseBlock 中的第 l 层输出， $\mathbf{X}_0 \sim \mathbf{X}_{l-1}$ 分别代表在 l 层之前所有层的输出，因此，第 l 层的输出主要由 $\mathbf{X}_0 \sim \mathbf{X}_{l-1}$ 层的输出组成，所以第 l 层的输入特征通道数量可表示为：

$$D_l = k_0 + k(l-1) \quad (3-6)$$

其中， k 表示增长率，它的含义为：前 $l-1$ 层中有 k 个特征通道进入 l 层。

$H_l(\cdot)$ 函数主要由以下层级组成：

$$BN + ReLU + Conv(1 \times 1) + BN + ReLU + Conv(3 \times 3) \quad (3-7)$$

其中， $Conv(1 \times 1)$ 层级的引入称为 Bottleneck Layers，其作用是为了减少输入的通道数量。

图 3-4 中的转换层（Transition Layers）在每两个 DenseBlock 结构之间插入，主要由 1×1 的卷积层和 3×3 的池化层组成，降低 DenseBlock 的输出通道数量，进一步减少参数数量。

密集学习网络（DenseNet）以其紧凑的内部表示和减少的特征冗余在更深的网络中作为特征提取器表现出良好的表征性能和优化特性，同时在参数数量增加时没有表现出性能下降和过拟合现象，同时在小型数据集上有非常好的抗过拟合性能，这种密集连接的网络结构在医学图像处理中得到广泛的应用。

3.3 两种网络的比较

残差学习网络（ResNet）：主要思想是在层间中加入代表恒等映射的快捷连接，如图 3-5 所示，通过在前层和后层之间建立的“短路连接”，在训练过程中更方便梯度信息的反向传播，因此可以训练出更深的卷积神经网络。

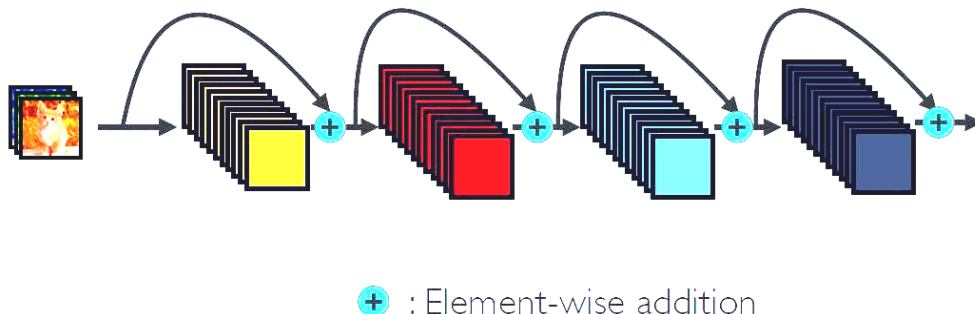


图3-5 ResNet 网络的残差学习结构（其中+代表的是元素级加法）

第 l 层的输出可表示为：

$$\mathbf{x}_l = H_l(\mathbf{x}_{l-1}) + \mathbf{x}_{l-1} \quad (3-8)$$

该结构的优点是后面层中的梯度可以通过恒等函数直接传输到前面的层。但是 ResNet 中恒等响应函数的输出与残差函数的输出的连接方式为加法操作，可能会对网络中的信息量流动造成一定影响。

密集学习网络（DenseNet）：其思路与残差学习网络有很多相似之处，但是它建立的是一种密集连接结构，每一层都与其他层有紧密联系。如图 3-6 所示，特征图在通道上的密集连接可以有效实现特征重用，这也是 DenseNet 一大优势。

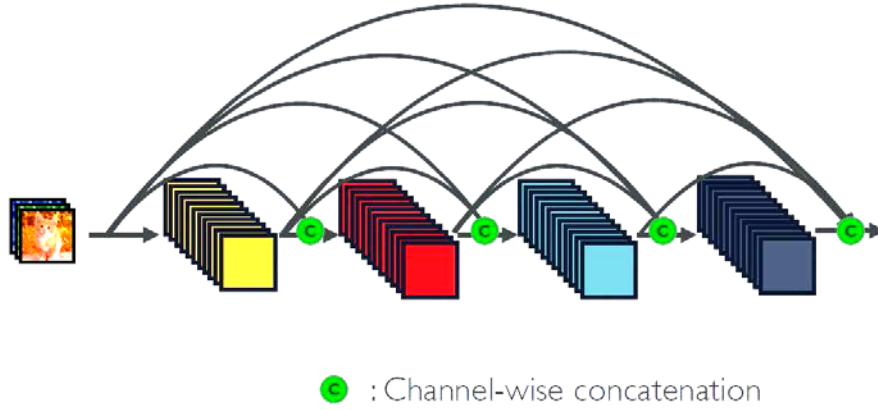


图3-6 DenseNet 网络的密集连接结构（其中 c 代表的是通道级连接）

第 l 层的输出可表示为：

$$\mathbf{x}_l = H_l([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}]) \quad (3-9)$$

ResNet 是一种加法连接。而 DenseNet 是直接聚合来自不同网络中不同层的多类特征图，这可以实现特征重用，提升效率，这一特点是 DenseNet 与 ResNet 最主要的区别。

3.4 深度双通道分离网络（Dual-Channel Splitting Network, DCSN）

3.4.1 双通道分离网络结构

通过 3.3 节对残差学习网络（ResNet）和密集学习网络（DenseNet）的比较可知，这两种模型在提取深层特征和图像分类中均表现出良好的特性，然而，许多当前的模型等价的处理不同通道上的分层特征，这对模型有区别和有针对性地处理分层特征没有帮助。因此，本文探索了一种新颖的双通道分离网络（DCSN），如图 3-7 所示。

该网络将分层特征划分成两个分支，即残差分支和密集分支，该网络不是将前一层的特征映射完全转移到下一层，而是将特征映射分成具有不同信息传输的两个不同部分，残差学习使特征重用，密集学习能够实现新的特征探索，这两者对于学习良好的表示均十分重要，之后再两个分支的信息融合，这种结构可以促进分层特征之间的信息共享和集成，有助于改善整个网络的信息流，使网络更容易训练。

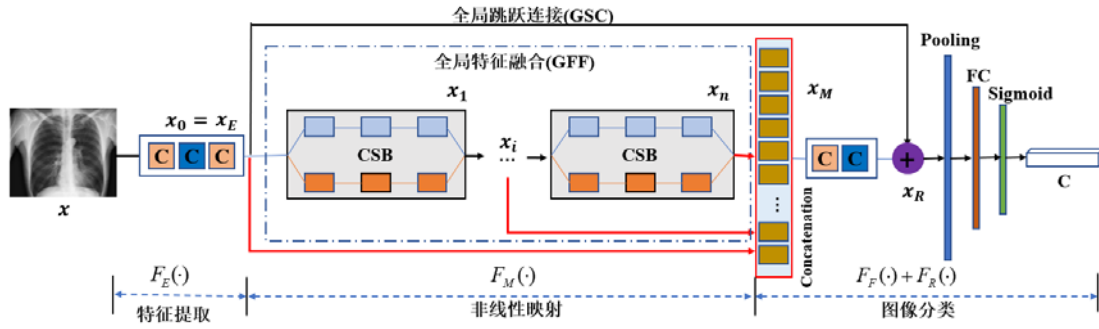
双通道分离网络的整体结构可分为三部分：

1) 特征提取（Feature Extract）

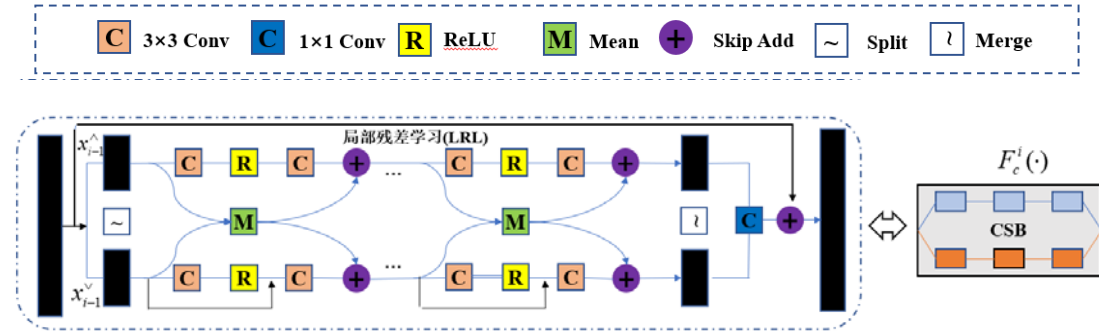
特征提取部分由 2 个 3×3 的卷积层和 1 个 1×1 的卷积层组成，定义 $F_E(\cdot)$ 为该处理过程的响应函数，所以提取的浅层特征为：

$$\mathbf{x}_E = F_E(\mathbf{x}) \quad (3-10)$$

其中， x 为输入的图像原始数据。



(a) 双通道分离网络 (DCSN) 的整体结构



(b) 通道分离模块 (Channel Split Block, CSB) 的内部结构

图3-7 双通道分离网络结构。(a) DCSN 的整体结构, 包含三个组成部分: 特征提取、非线性映射和图像分类, GFF 模块用来融合分层特征, GSC 模块用来稳定模型训练。(b) CSB 模块内的中间特征映射被分成两个分支, 一个构建为残差学习分支 (ResNet), 另一个被构建为密集学习分支 (DenseNet)

2) 非线性映射 (Nonlinear Mapping)

非线性映射部分主要由通道分离模块 (CSB) 组成, 将提取的特征分成两个分支进行处理, 这里我们定义 CSB 模块的响应函数为 $F_c(\cdot)$, 该函数的细节将在 3.4.2 节中详细描述, 假设该部分有 n 个 CSB 模块, 则第 i 个 CSB 模块的输出为:

$$\mathbf{x}_i = F_c^i(\mathbf{x}_{i-1}), i = 1, 2, \dots, n \quad (3-11)$$

可知该部分的原始输入 \mathbf{x}_0 为特征提取部分的输出 \mathbf{x}_E , 故最终输出为:

$$\mathbf{x}_n = F_c^n(\mathbf{x}_{n-1}) = F_c^n(F_c^{n-1}(\dots(F_c^1(\mathbf{x}_0))\dots)) \quad (3-12)$$

3) 图像分类 (Image Classification)

图像分类部分包括两个步骤: 全局特征融合和特征图分类。

(a) 全局特征融合

全局特征由前面每一个 CSB 模块的输出聚合而成, 如图 3-7(a)所示, 进行全

局特征融合（Global Feature Fusion, GFF）、假设有 n 个 CSB 模块，表达为一个聚合向量 \mathbf{x}_M ：

$$\mathbf{x}_M = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n] \quad (3-13)$$

然后，此聚合向量经过两个卷积层处理（ $3 \times 3 + 1 \times 1$ ）后得到融合特征，这里将其定义为函数 $F_F(\cdot)$ 。此外，还添加了一个全局跳跃连接模块（Global Skip Connection, GSC），将浅层特征直接添加到聚合特征中进行一次局部残差学习，从而得到最终的全局特征 \mathbf{x}_R ：

$$\mathbf{x}_R = F_F(\mathbf{x}_M) + \mathbf{x}_0 = F_F([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n]) + \mathbf{x}_0 \quad (3-14)$$

(b) 特征图分类

将得到的全局特征 \mathbf{x}_R 分别经过池化层（Pooling）、全连接层（FC）和 Sigmoid 归一化层，最终输出为 14 维向量。代表每种疾病的预测概率：

$$\tilde{p}(c|I) = 1 / (1 + \exp(-p(c|I))) \quad (3-15)$$

其中， I 代表输入图像， $\tilde{p}(c|I)$ 表示图像 I 属于类别 c 的概率分数， $c \in \{1, 2, \dots, C\}, C = 14$ ，之后通过最小化二进制交叉熵（Binary Cross-Entropy, BCE）损失函数来优化模型参数：

$$L(I, y) = -\frac{1}{C} \sum_{c=1}^C l_c \log(\tilde{p}(c|I)) + (1 - l_c) \log(1 - \tilde{p}(c|I)) \quad (3-16)$$

l_c 表示类别 c 的训练标签。

由图 3-7 可知该网络的深度 D 可表示为：

$$D = n(2m + 1) + 8 \quad (3-17)$$

其中， n 为 CSB 模块的数量， m 为每个 CSB 模块中操作的级数。

3.4.2 双通道分离模块

DCSN 网络中的通道分离模块用来合并不同传播机制的不同信息。如图 3-7(b) 所示，所提出的 CSB 模块的主要作用是进行通道分裂和合并，以及残差学习和密集学习两个分支的融合。在模块中还应用局部残差学习（Local Residual Learning, LRL）来进一步改善信息流，稳定训练过程。

通道分离模块（CSB）的实现可分为三个部分：

1) 通道分离和融合

对于第 i 个 CSB 模块，输入特征 \mathbf{x}_{i-1} 沿着通道方向被分成上下两个分支，分别经过残差学习网络（ResNet）和密集学习网络（DenseNet）处理后得到上分支特征 $\mathbf{x}_{i-1}^{\dot{U}}$ 和下分支特征 $\mathbf{x}_{i-1}^{\dot{D}}$ ，定义为：

$$\mathbf{x}_{i-1}^{\dot{U}}, \mathbf{x}_{i-1}^{\dot{D}} = S_c(\mathbf{x}_{i-1}), i = 1, 2, \dots, n \quad (3-18)$$

同理，在 CSB 模块的输出也存在一个融合操作 $M_c(\cdot)$ ，将上下分支的结果沿通道方向重新聚合，重新作为下一个模块的输入：

$$\mathbf{x}_i = M_c(\mathbf{x}_{i-1}^{\dot{U}}, \mathbf{x}_{i-1}^{\dot{D}}), i = 1, 2, \dots, n \quad (3-19)$$

通道分离和融合操作可以使不同层次的特征通过不同的网络处理，得到更优的特征表示，同时也有利于网络的加深。

2) 特征重用及新特征探索

在 CSB 模块的内部存在多级操作，每一级都对分离特征进行一次分支处理，并且在每级网络处理之后还包含一个类似恒等映射的均值相加模块，可以促进网络内信息流动，有利于反向传播训练，这里我们定义 $H_{i,j}^R(\cdot)$ ， $H_{i,j}^D(\cdot)$ 分别为残差分支和密集分支的响应函数， $M_{i,j}(\cdot)$ 为求均值函数，假设每个 CSB 模块共有 m 级操作，则第 i 个 CSB 模块中的第 j 级操作后的输出可表达为：

$$\begin{bmatrix} \mathbf{x}_{i-1,j}^{\dot{U}} \\ \mathbf{x}_{i-1,j}^{\dot{D}} \end{bmatrix} = \begin{bmatrix} H_{i,j}^R(\mathbf{x}_{i-1,j-1}^{\dot{U}}) \\ H_{i,j}^D(\mathbf{x}_{i-1,j-1}^{\dot{D}}) \end{bmatrix} + \begin{bmatrix} M_{i,j}(\mathbf{x}_{i-1,j-1}^{\dot{U}}) \\ M_{i,j}(\mathbf{x}_{i-1,j-1}^{\dot{D}}) \end{bmatrix}, i = 1, 2, \dots, n \ \& \ j = 1, 2, \dots, m \quad (3-20)$$

该部分的上下双分支单独处理不同属性的层次特征，可以有效整合两种网络的优势，得到更好的特征表示，同时也能加深网络，稳定训练。

3) 局部残差学习（LRL）

在单个 CSB 模块的最后，将上下分支得到的输出级联并通过一个 1×1 的卷积层（用 $C_1(\cdot)$ 表示）进行融合并调整通道数量，为了保留原始信息，防止梯度消失，我们在每个 CSB 模块的最后增加局部残差学习分支（如图 3-7(b)所示），将原始输入与输出进行相加连接，故经过第 i 个 CSB 模块的最终输出特征可表示为：

$$\mathbf{x}_i = C_1(M_c(\mathbf{x}_{i-1,m}^{\dot{U}}, \mathbf{x}_{i-1,m}^{\dot{D}})) + \mathbf{x}_{i-1} \quad (3-21)$$

此后， \mathbf{x}_i 将作为下一模块的输入以及最终聚合特征的一部分，直到 n 个 CSB 模块全部结束。

3.5 病变检测可视化

为了能更好的观察病变的显著区域,也为了为医生的二次诊断提供更直观的诊断依据,本文将测试图像检测出的胸部疾病进行定位可视化,这里使用由卷积神经网络的全局池化层(Global Average Pooling, GAP)生成的类激活图(Class Activation Mapping, CAM)^[40]来表示,具体过程如图 3-8 所示。

将 DCSN 网络输出的卷积特征图后面加入 GAP 层,然后将 X 光图像中的重要区域用输出层权重映射回卷积层特征,将类激活图与输入图像的尺寸保持一致,就能清楚地显示出与特定类最相关的区域。之后在原图像上进行标记,得到胸部疾病分类显著性区域的可视化定位图。用公式表示如下:

$$S_c = \sum_k w_k^c \sum_{x,y} f_k(x,y) = \sum_{x,y} \sum_k w_k^c f_k(x,y) \quad (3-22)$$

其中,用 $f_k(x,y)$ 代表最后一个卷积层在空间坐标 (x,y) 中单元 k 的激活值。然后,对于每个单元 k ,通过 GAP 后的结果为 $\sum_{x,y} f_k(x,y)$ 。 w_k^c 代表单元 k 对应的类 c 的权重。实际上, w_k^c 就是该单元对类 c 的重要性,最终得到每一类的得分,实现特定类的区别性区域的可视化。CAM 方法本质上是不同空间区域的线性加权可视化。

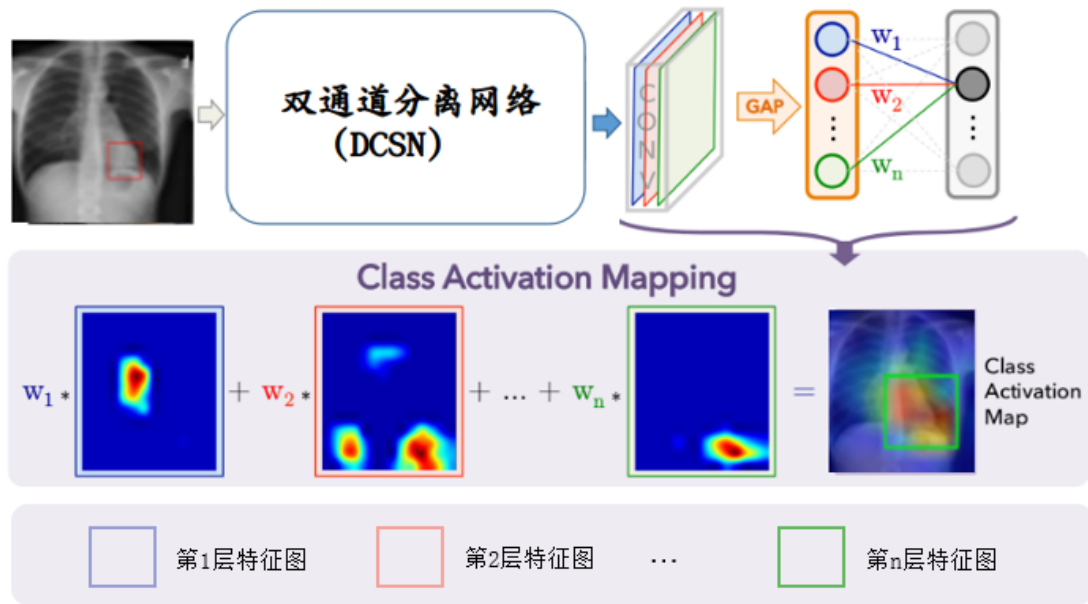


图3-8 预测的类别分数被映射回先前的卷积层以生成类别激活图

3.6 本章小结

本章主要介绍了网络模型设计的详细过程，前半部分阐述了目前流行的 ResNet 和 DenseNet 两种深度网络的内部结构及处理过程，并对两种网络的优缺点进行对比分析，进而在后半部分提出结合两种网络特点的 DCSN 网络，采用通道分离的双分支结构，详细介绍网络的组成和具体处理流程。并在本章最后介绍了一种基于 CAM 的定位可视化方法，可对图像分类结果进行显著性分析。

第四章 实验过程及结果

本章将详细介绍实验的具体过程，包括实验的前期准备和多个对比实验，阐述基于多个不同类型网络的实验细节。在 4.2 和 4.3 节中介绍本实验采用的数据集及结果评价指标。在实验阶段，首先用两个实验分别分析仅采用单个网络的模型性能，该部分见 4.4 和 4.5 节。在 4.6 节中，使用新设计的双网络 DCSN 模型，并将结果与之前的实验进行对比。

4.1 实验环境

本文在实验中尝试了多种深度学习框架，在实验初期，分别使用 TensorFlow 和 PyTorch 进行了完整代码的编写，在后续测试过程中，经过对代码简洁性和模型训练速度的比较，最终选择 PyTorch 框架作为本实验的主要代码工具。

在本实验中使用的代码环境中，系统版本为 Ubuntu16.04 x64，Python 版本为 3.6，CUDA 版本为 9.0，Pytorch 版本为 0.2.0。

4.2 数据集

4.2.1 数据集介绍

本实验使用的数据集是由 NIH 临床中心于 2017 年开源的 ChestX-ray14 数据集，是目前最大的临床医疗胸部 X 光图像数据集，该数据集包含 30,805 名患者的 112,120 张正面视图的 X 射线图像，以及利用自然语言处理技术（NLP）从相关放射学报告中挖掘的 14 类疾病的图像标签。数据集中含有 14 类常见胸部疾病的 X 光图像，其中包括肺不张、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，每张图像可以有多个疾病标签，图像对应的原始 CXR 的放射学报告并未公开。据称，该数据集是利用现有的诊断报告，通过使用文本挖掘方法来对每张图像进行标注，准确率达到 90% 以上。

4.2.2 数据详解

1) 数据集结构

该数据集包括：

Image_001.zip 至 Image_0012.zip：存储原始胸部 X 光图像，共 112120 张。

Data_Entry_2017.csv：存储每张 X 光图像的分类标签和病人数据。

test_list.txt：划分的测试集图像，约占总数据集的 20%。

train_val_list.txt：划分的训练集图像，约占总数据集的 70%。

BBox_List_2017.csv：标注疾病的边界框坐标，包含 984 条图像信息。

2) 图像

每张胸部 X 光图像的存储格式为 PNG 格式，图像分辨率为 1024*1024。数据集包括十四种胸部疾病 X 光图像（一张图像可能含有多种疾病，故存在重叠）以及正常 X 光图像，具体分布如表 4-1 所示。

表4-1ChestX-ray14 数据集中的正常和含疾病图像的数量分布情况

疾病名称	英文	数量	重叠数量
肺不张	Atelectasis	11535	7323
变实	Consolidation	4467	3353
浸润	Infiltration	19871	10319
气胸	Pneumothorax	5289	3099
水肿	Edema	2303	1669
肺气肿	Emphysema	2516	1621
纤维变性	Fibrosis	1686	959
积液	Effusion	13307	9348
肺炎	Pneumonia	1353	1046
胸膜增厚	Pleural thickening	3385	2258
心脏肥大	Cardiomegaly	2772	1678
结节	Nodule	6323	3617
肿块	Mass	5746	2138
疝气	Hernia	227	117
正常	No findings	60421	0

3) 注释

Data_Entry_2017.csv 文件中存储图像注释信息，每行描述一张图像，每行保存的信息包括：

- Image Index: 图像文件名称，即索引。
- Finding Labels: 疾病类型（类别标签），共十五种。
- Follow-up : 定期复查情况。
- Patient ID: 病人编号
- Patient Age: 病人年龄
- Patient Gender: 病人性别
- View Position: X 射线的方向，AP 表示正面，PA 表示反面。
- Original Image Width: 原始图像宽度。

- Original Image Height: 原始图像高度。
- Original Image Pixel Spacing x: x 方向的原始图像像素间距
- Original Image Pixel Spacing y: y 方向的原始图像像素间距

BBox_List_2017.csv 文件中存储疾病边界框信息，每行描述一张图像，每行保存的信息包括：

- Image Index: 图像文件名称，即索引。
- Finding Label: 疾病类型（类别标签），每张图像仅标注一个。
- Bbox x: 边界框的 x 轴坐标。
- Bbox y: 边界框的 y 轴坐标。
- Bbox w: 边界框的宽度。
- Bbox h: 边界框的高度。

4.2.3 数据预处理

数据预处理可分为数据增强和图像预处理两部分。

1) 数据增强

数据增强主要是为了减少网络的过拟合现象，通过对训练图片进行多种方式的变换，从已有的训练数据集中生成一些新的数据，来扩大训练数据。得到泛化能力更强的网络。

数据增强的方法包括：

- 旋转/反射变换：将图像随机旋转一定角度或改变图像方向。
- 翻转变换：沿不同方向对图像随机翻转。
- 缩放变换：对图像尺寸进行放大或缩小。
- 平移变换：在平面上将图像进行水平移动。

2) 图像预处理

由于数据集中图像是一种平稳的数据分布，所以需要进行一些预处理操作消除公共部分，以凸显个体之间的特征和差异，同时有利于在反向传播中加快网络中每一层权重参数的收敛。

图像预处理的方法包括：

- min-max 标准化（归一化）：将像素灰度值映射到[0, 1]之间。

公式为：新数据 = (原数据 - 最小值) / (最大值 - 最小值)

- z-score 标准化（规范化）：基于每张图像的像素的均值和标准差进行标准化处理。

公式为：新数据 = (原数据 - 均值) / 标准差

- 正则化：在样本集中的每个样本均需进行计算，得到 p-范数，然后使用该范数做每个元素的除数，对数据进行转换以缩放到单位范数。

● 稀疏编码：是一种重新编码过程，找到对应样本集中每个样本的一组基向量，使用这些基向量对输入向量进行线性表示，从而获得样本的稀疏表达。该方式仅激活数据中最有区分度的视觉特征向量，目的是可以更好的表示图像并简化计算量。

4.3 评价指标

本实验使用了多种评价指标，对实验结果进行对比评估，主要种类如表 4-2 所示。

表4-2分类器的多种评价指标及意义

类型	指标意义
真阳性 (TP)	被模型预测为正的正样本
真阴性 (TN)	被模型预测为负的负样本
假阳性 (FP)	被模型预测为正的负样本，可以称作误报率
假阴性 (FN)	被模型预测为负的正样本，可以称作漏报率
灵敏度 (TPR)	$TPR = TP / (TP + FN)$
特指度 (TNR)	$TNR = TN / (TN + FP)$
假阳率 (FPR)	$FPR = FP / (FP + TN)$
假阴率 (FNR)	$FNR = FN / (TP + FN)$
接受者操作特性曲线 (Receiver Operating characteristic Curve, ROC)	曲线上每个点反映着对同一信号刺激的感受性。
AUC (Area under Curve)	ROC 曲线下的面积，介于 0.1 和 1 之间，可以评价分类器的好坏，值越大越好

ROC 曲线的平面的横坐标是假阳率 (FPR)，纵坐标是灵敏度(TPR)。假阳率指在所有实际为阴性的样本中，被错误地判断为阳性之比率。灵敏度指在所有实际为阳性的样本中，被正确地判断为阳性之比率。

在对分类器的测试过程中，TPR 和 FPR 的值是根据该分类器在测试样本上的表现得到。每得到一组数据，即可映射成 ROC 平面上的一组坐标进行描点绘图。设定多个分类阈值进行遍历，就画出得到一条曲线，该曲线经过点(0, 0)和(1, 1)。曲线距离左上角越近，证明分类器效果越好。

AUC 值为 ROC 曲线所覆盖的区域面积，假设将一张图像判断为正类的分数作为分类器的输出，则 AUC 的物理意义为，任意取一正一负样本，属于正样本的分数大于属于负样本的分数的概率。显然，AUC 越大，分类器分类效果越好。

4.4 实验 I：基于残差学习网络的胸部病变差异性识别

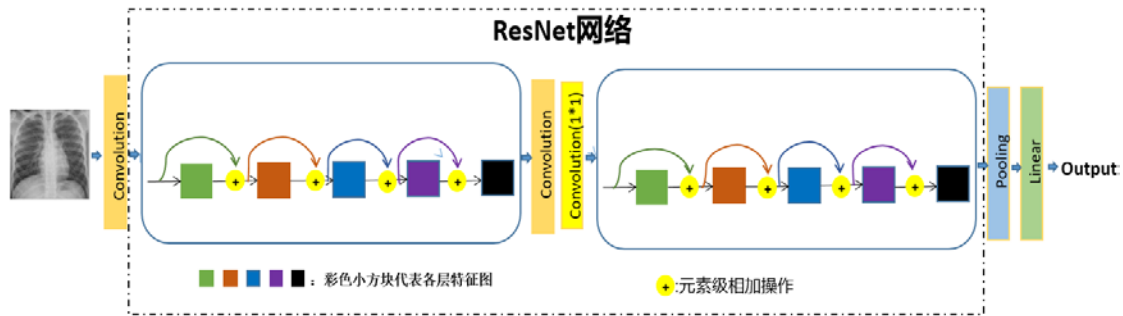


图4-1 使用残差学习网络(ResNet)的网络结构

4.4.1 实验内容

该实验使用单分支的残差学习网络（ResNet）作为特征提取器，ResNet 是通过每一层之间的元素级的相加操作，建立前后层之前的“短路连接”，在每层之间引入恒等映射，将后面层的梯度更容易传输到前面的层，网络结构如图 4-1 所示。具体实验的网络结构种类为 ResNet50，ResNet101 和 ResNet152，其网络组成由表 4-3 所示，其中将网络的最后一个全连接层替换为一个池化层和一个输出为 14 维的线性分类层，之后连接一个 Sigmoid 非线性层，输出每张图像对应的 14 种胸部疾病的概率值。

具体网络处理流程如图 4-1，其中的小矩形框代表一次 ResBlock 操作，之后连接两个卷积层，卷积核大小分别为 3×3 和 1×1 ，该层作为转换层调整特征图的大小和通道数量，同时在每个卷积层和激活层之间加入批量正则化层来防止过拟合，网络中 ResBlock 模块的数量可决定网络的深度，其表现出不同的分类特性。

在模型训练过程中，采用论文^[41]提到的权重初始化^[42]，梯度优化使用随机梯度下降（Stochastic Gradient Descent, SGD），mini-batch 的大小为 256，学习速率初始化为 10^{-4} ，在每 10 次迭代进行指数衰减，动量参数为 0.9，迭代次数为 1000，选择验证损失最低的模型作为最终训练好的网络模型。

在网络模型测试阶段，使用数据集集中的 20% 作为测试集图像，与图像的真实标签进行对比，并绘制模型结果的 ROC 曲线。

4.4.2 实验结果

实验中分别使用 ResNet-50、ResNet-101 和 ResNet-152 三种网络作为基础，训练其对应的分类器网络，并利用矩阵计算得到其各个评价指标的价格，绘制其 ROC 曲线如图 4-2、图 4-3、图 4-4 所示。将 AUC 分数绘图对比于图 4-5，并显示其具体数字于表 4-4 中。

表4-3 残差学习网络（ResNet）的结构^[39]

Layer	Output Size	ResNet-50	ResNet-101	ResNet-152
Conv1	112×112	7×7 conv, stride 2		
Conv2_x	56×56	3×3 max pool, stride 2		
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
Conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
Conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	1000D 全连接层, Softmax 分类		

表4-4 三种残差学习网络在分类任务上的 AUC 分数

	Resnet-50	Resnet-101	Resnet-152
Atelectasis	0.4490	0.4890	0.5686
Cardiomegaly	0.5337	0.5455	0.5859
Effusion	0.4491	0.5361	0.5762
Infiltration	0.4290	0.5247	0.5068
Mass	0.4725	0.5639	0.4616
Nodule	0.4909	0.4997	0.5130
Pneumonia	0.5415	0.4955	0.4692
Pneumothorax	0.3635	0.4007	0.5897
Consolidation	0.4614	0.4014	0.5948
Edema	0.5895	0.4288	0.4560
Emphysema	0.4780	0.5100	0.4011
Fibrosis	0.5435	0.5970	0.5261
PT	0.4524	0.5278	0.4707
Hernia	0.3881	0.4549	0.4738

ResNet 随着网络深度的加深，在特定疾病上的分类性能有一定提升，部分类别的 AUC 分数有所上升，但整体分数偏低，说明使用单个 ResNet 的网络结构没有在 CXR 图像的分类任务上表现出较好的效果。

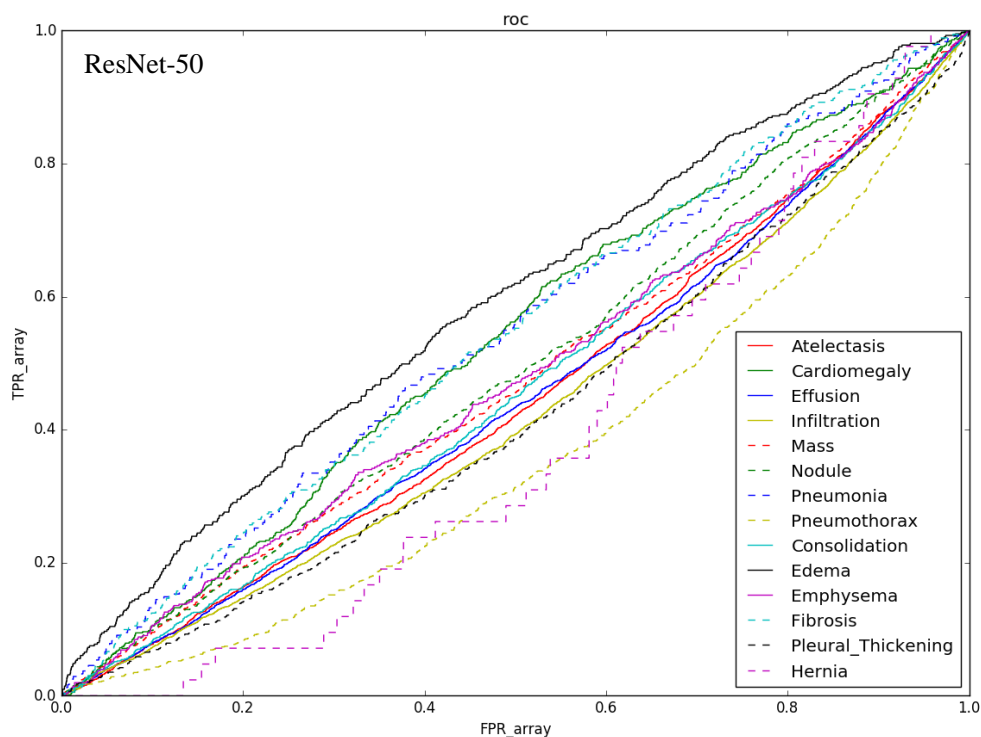


图4-2 ResNet-50 网络的 ROC 曲线图

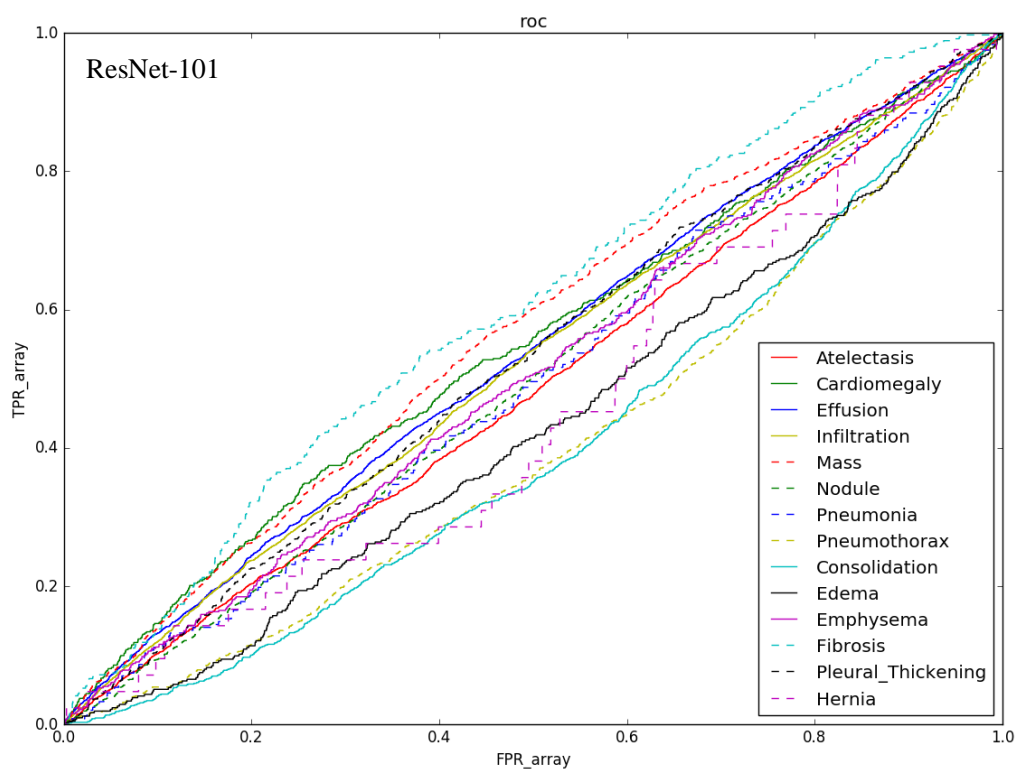


图4-3 ResNet-101 网络的 ROC 曲线图

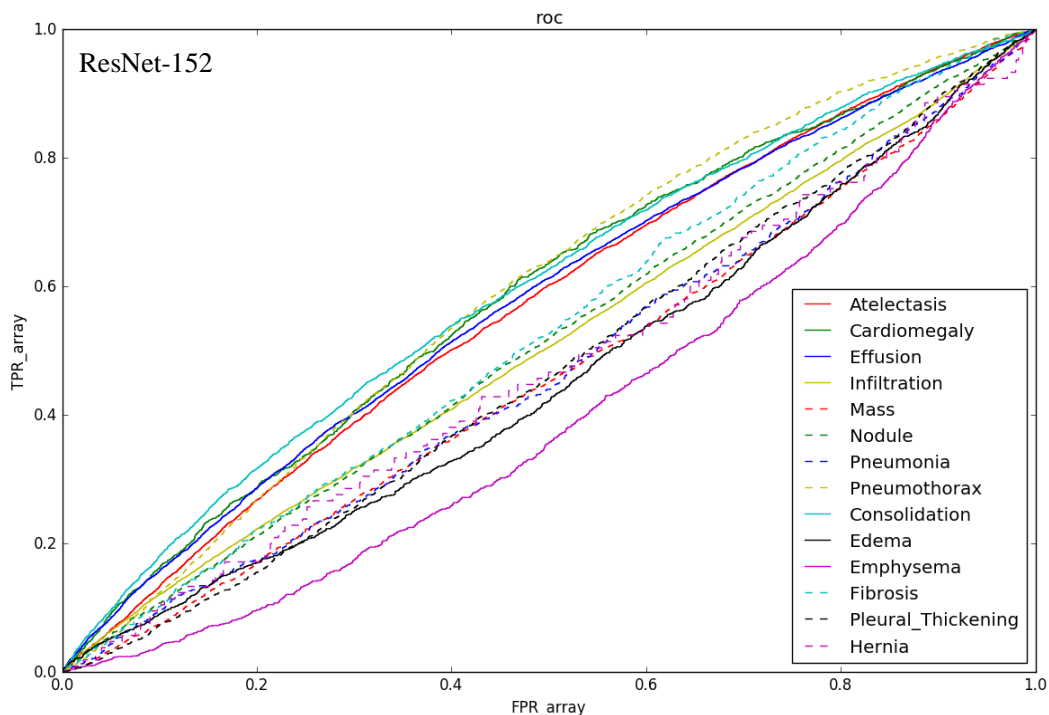


图4-4 ResNet-152 网络的 ROC 曲线图

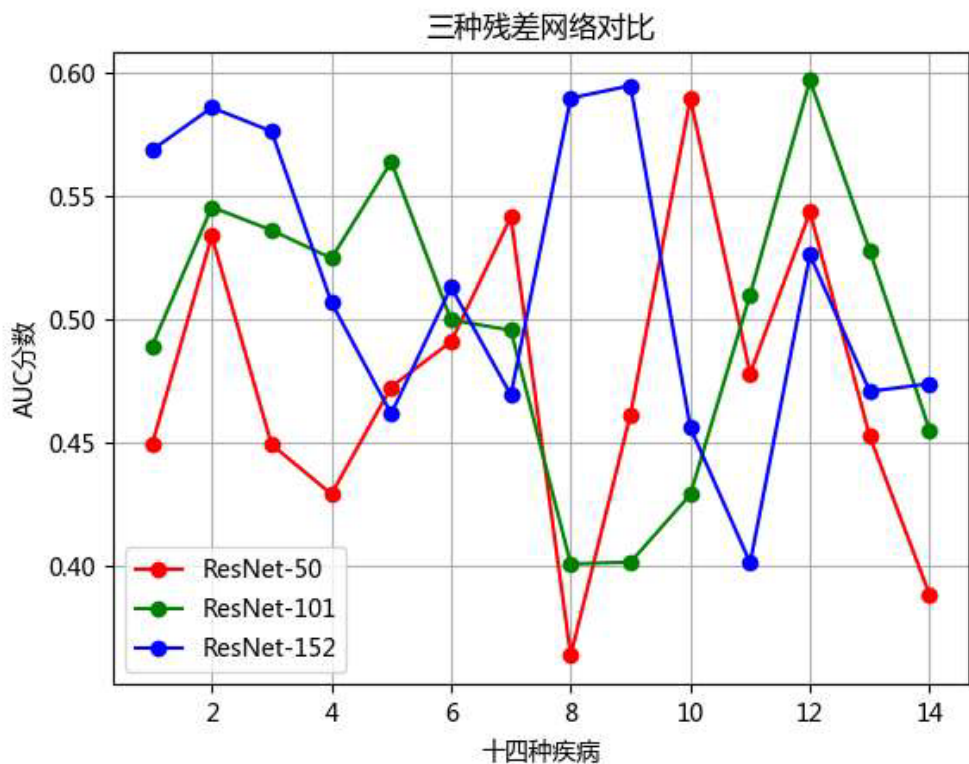


图4-5 三种残差学习网络的 AUC 分数对比图

由图 4-5 的曲线图对比可知，152 层的残差网络由于有更多的残差块而获得一定的 AUC 分数增益，与 50 层网络相比，在一些疾病上的表现较好。101 层网络在参数较少的情况下，获得了性价比更高的分类结果。

4.5 实验 II：基于密集连接网络的胸部病变差异性识别

4.5.1 实验内容

该实验使用单分支的密集连接网络（DenseNet）作为特征提取器，密集网络是将网络中的所有层都进行通道级的连接操作，每一层都与后面的层两两连接，接受前面的所有输入，最大化网络中的信息流，可以综合浅层的简单特征，从而有更好的泛化性能。具体实验的网络结构种类为 DenseNet121，DenseNet169，DenseNet201，其中将网络的最后一个全连接层删掉，增加一个池化层和一个输出为 14 维的线性分类层，之后通过一个 Sigmoid 函数层，输出每张图像对应的 14 种胸部疾病的概率值。

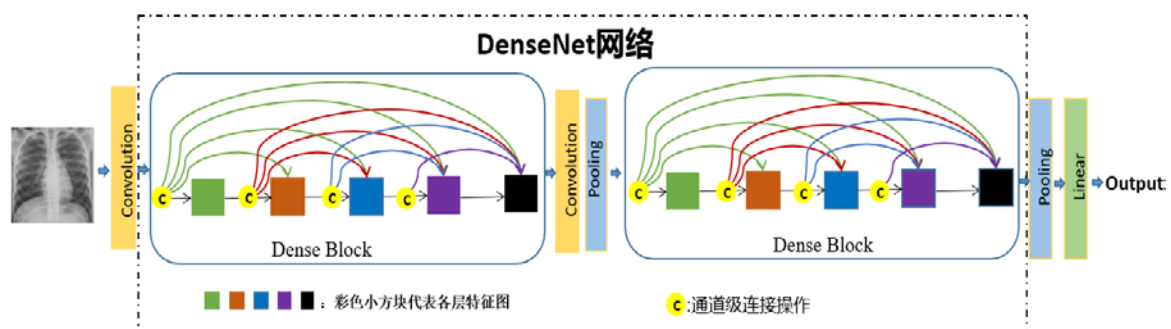


图4-6 使用密集连接网络（DenseNet）的网络结构

具体网络处理流程如图 4-6。其中的小矩形框代表一次 DenseBlock 操作，之后连接一个卷积层和一个池化层作为转化层，这个操作可以在网络中重复，以增加网络的深度。DenseNet 的具体网络组成如表 4-5 所示。

在模型训练阶段，网络权重使用 ImageNet 的预训练模型进行初始化，梯度优化使用自适应学习率的 Adam 算法（使用标准参数： $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$ ），小批量 mini-batches 为 16，损失函数使用 BCE 损失，初始学习率设置为 10^{-3} ，并在每 10 次迭代进行指数衰减，迭代次数为 100，选择验证损失最低的模型作为最终训练好的网络模型。

在网络模型测试阶段，使用数据集中的 20% 作为测试集图像，与图像的真实标签进行对比，并绘制模型结果的 ROC 曲线。

4.5.2 实验结果

实验中分别使用 DenseNet-121、DenseNet-169 和 DenseNet-201 三种网络作为基础，训练其对应的分类器网络，并利用矩阵计算得到其各个评价指标的结果，绘制三种网络的 ROC 曲线如图 4-7、图 4-8、图 4-9，并将最终得到的 AUC 分数在图 4-10 中进行对比，显示具体数字于表 4-6 中。

表4-5 密集连接网络（DenseNet）的结构^[29]

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201
Convolution	112×112	7×7 conv, stride 2		
Pooling	56×56	3×3 max pool, stride 2		
Dense Block (1)	56×56	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv		
	28×28	2×2 average pool, stride 2		
Dense Block (2)	28×28	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 12$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 12$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv		
	14×14	2×2 average pool, stride 2		
Dense Block (3)	14×14	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 24$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 32$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 48$
Transition Layer (3)	14×14	1×1 conv		
	7×7	2×2 average pool, stride 2		
Dense Block (4)	7×7	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 16$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 32$	$\begin{bmatrix} 1\times1, conv \\ 3\times3, conv \end{bmatrix} \times 32$
Classification Layer	1×1	7×7 global average pool		
		1000D 全连接层, Softmax 分类		

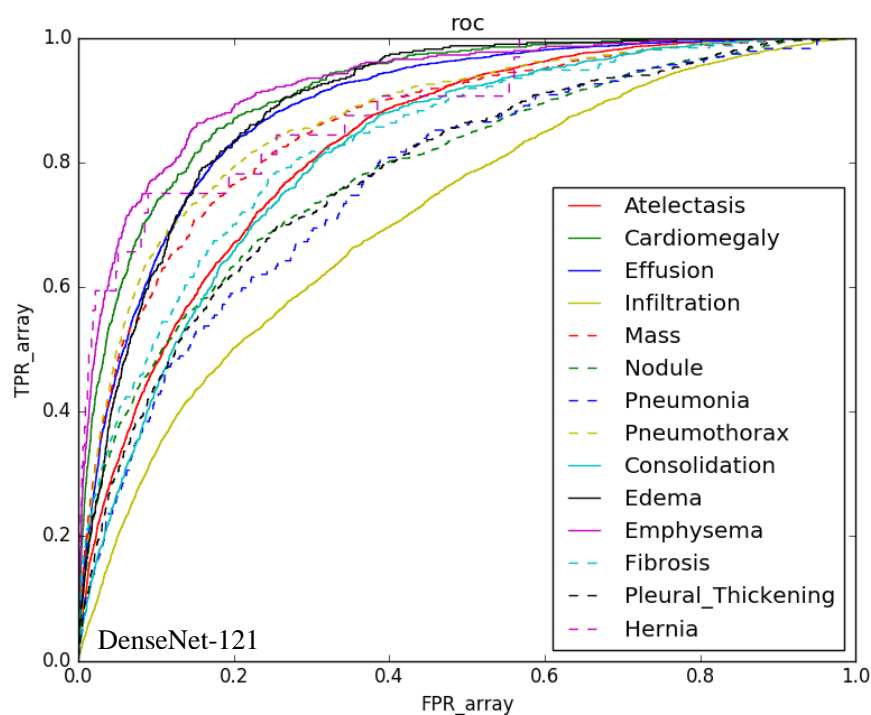


图4-7 DenseNet-121 的 ROC 曲线图

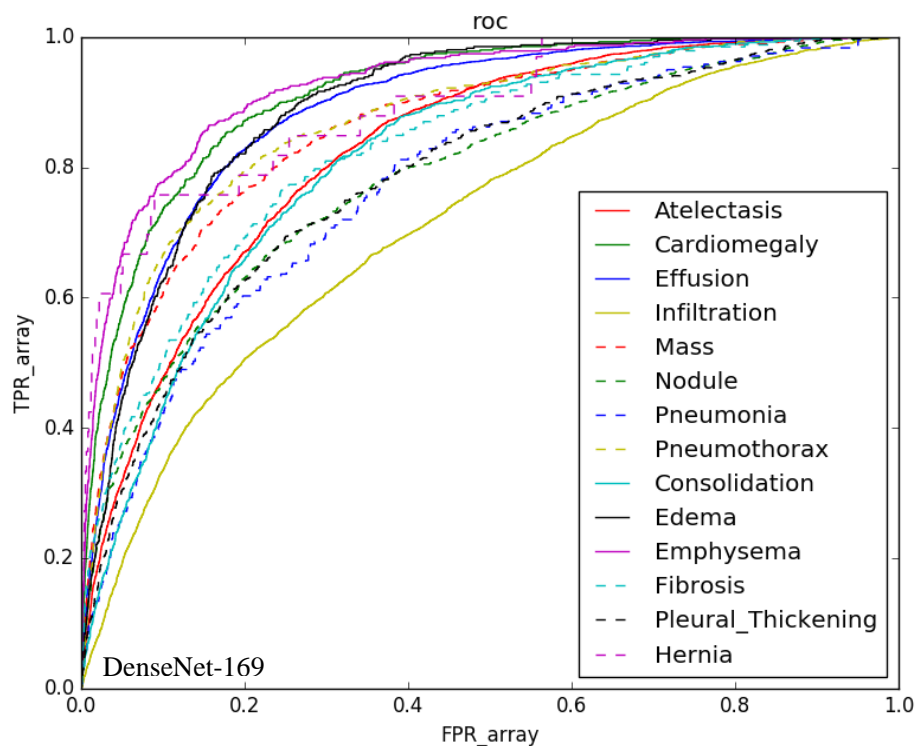


图4-8 DenseNet-169 的 ROC 曲线图

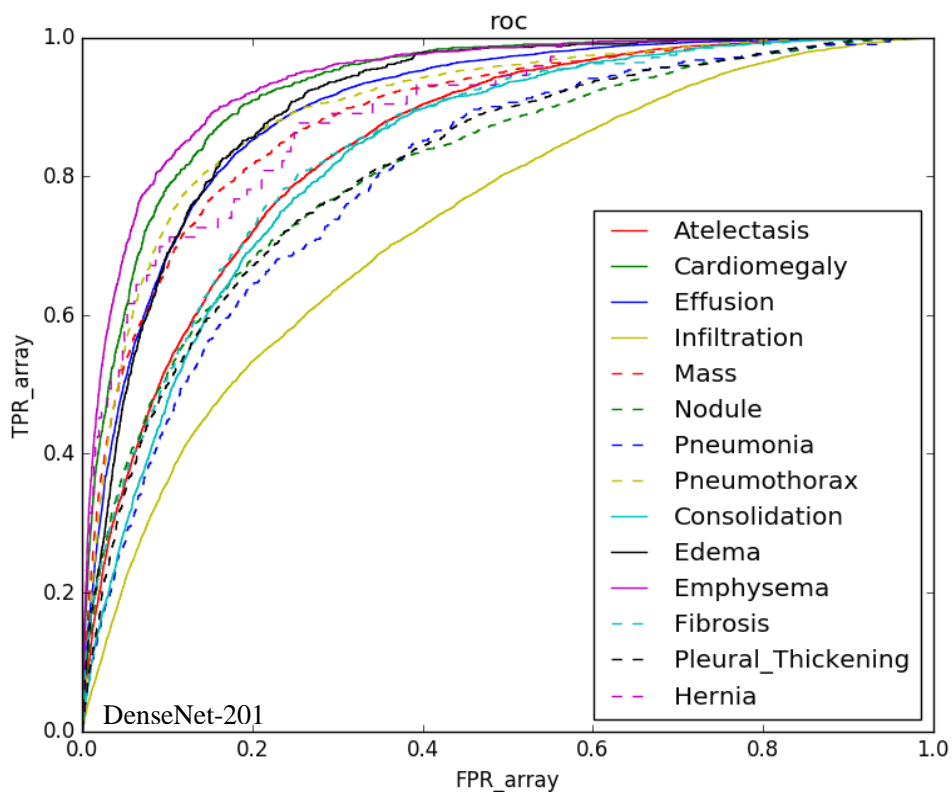


图4-9 DenseNet-201 的 ROC 曲线图

表4-6 三种密集连接网络在分类任务上的 AUC 分数

	DenseNet-121	DenseNet-169	DenseNet-201
Atelectasis	0.8285	0.8279	0.8467
Cardiomegaly	0.9132	0.9143	0.9287
Effusion	0.8878	0.8872	0.8996
Infiltration	0.7144	0.7146	0.7348
Mass	0.8604	0.8608	0.8867
Nodule	0.7882	0.7862	0.8130
Pneumonia	0.7695	0.7742	0.8017
Pneumothorax	0.8710	0.8702	0.9006
Consolidation	0.8147	0.8149	0.8321
Edema	0.8931	0.8918	0.9045
Emphysema	0.9233	0.9247	0.9390
Fibrosis	0.8320	0.8277	0.8449
PT	0.7825	0.7851	0.8134
Hernia	0.8876	0.8917	0.8921

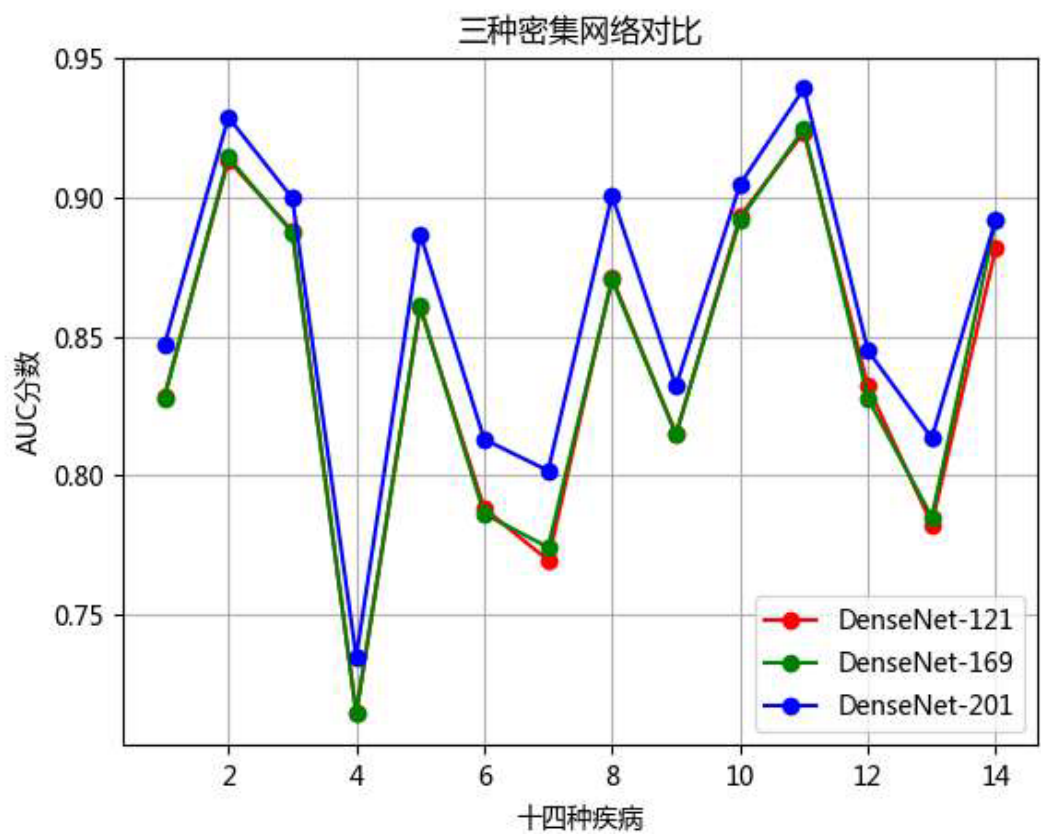


图4-10 三种密集连接网络的 AUC 分数对比图

根据图 4-10 的曲线对比图可知 DenseNet 随着网络深度的增加,在该分类任务上的性能表现有一定的提升,并由表 4-6 中的具体 AUC 分数可以说明 DenseNet 相比于 ResNet 网络, AUC 分数有相当明显的提升,也验证了章节 3.2 中提到的密集连接结构在小数据集上有更好的泛化性能。同时在使用较少参数网络和降低计算量的原则下, 121 层密集连接网络有较好的性价比。

4.6 实验 III: 基于通道分离双网络 (CSDN) 的胸部病变差异性识别

4.6.1 实验内容

该实验采用双通道分离网络结构 (DCSN), 如图 3-7 所示。将输入图像的浅层特征图分离到两个分支中同时进行处理, 其中上分支为残差学习分支 (ResNet), 下分支为密集学习分支 (DenseNet), 之后融合上下分支的输出得到全局特征进行最终分类。

在模型训练阶段, 网络的深度参数设置为 $n = m = 4$, 对于模型的初始参数我们使用 Xavier 进行初始化^[43], minibatch 的大小为 16, 特征图的数量为 256, 对于上分支的残差学习, 使用 ResNet50 网络结构, 对于下分支的密集学习, 使用 DenseNet121 的网络结构, 梯度优化使用自适应学习率的 Adam 算法 (使用标准参数: $\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8}$), 学习率为 10^{-4} , 同时在每 10 次迭代后进行指数衰减, 迭代次数为 2000, 选择验证损失最低的模型作为最终训练好的网络模型。

在网络模型测试阶段, 使用数据集中的 20% 作为测试集图像, 与图像的真实标签进行对比, 并绘制模型结果的 ROC 曲线。

4.6.2 实验结果

实验中使用双通道分离网络 (DCSN) 网络作为基础, 训练其对应的分类器网络, 并利用矩阵计算得到其各个评价指标的结果, 绘制其 ROC 曲线如图 4-11, 将三种网络最终得到的最优 AUC 分数画图进行对比, 具体分数可见表 4-7。

由图 4-12 中结果可知, 融合两种分支网络的双通道分离网络 (DCSN) 跟单独使用残差或密集网络的结构相比, AUC 分数有较为明显的提升, 证明该分类器对胸部疾病的识别准确率更高, 同时训练的收敛过程更快, 网络的泛化性能也更好。

由表 4-7 中与其他相关算法对比的结果可知, 本文的提出的双通道分离网络 (DCSN) 结构在与现有的胸部疾病分类的其他最优算法相比, 几乎每一类的 AUC 分数都有明显提升, 可说明本实验的算法有优越的性能表现, 已经成为目前常见胸部疾病的 X 光图像分类任务中表现最好的网络。

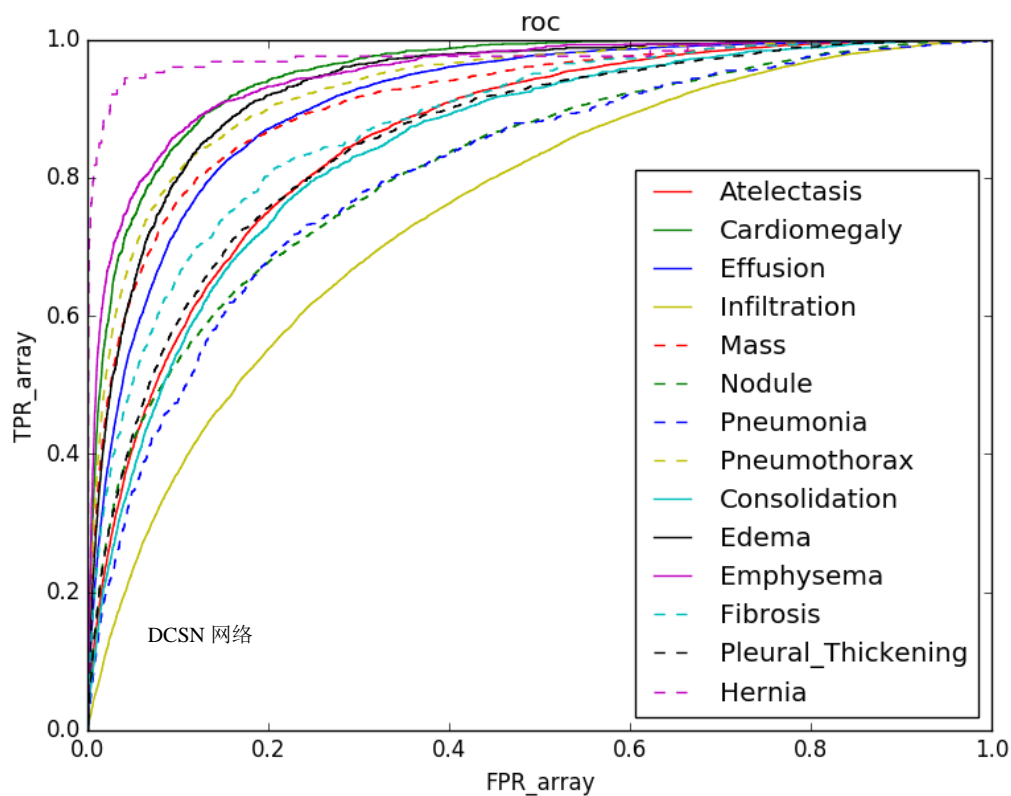


图4-11双通道分离网络（DCSN）的 ROC 曲线图

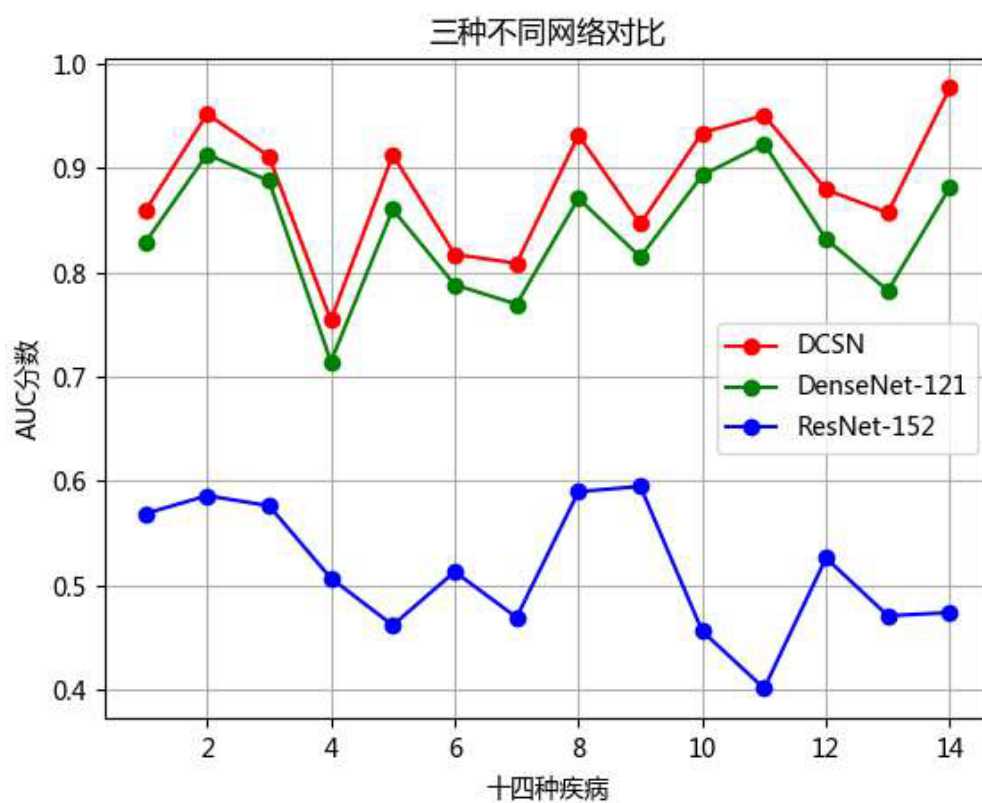


图4-12双通道分离网络（DCSN）与其他两种网络的 AUC 分数对比图

表4-7 本文的三种实验网络与现有其他算法在 AUC 分数上的对比

	ChestX [20]	Yao. [44]	DNetLoc [45]	CheX- Net ^[32]	AG-C NN ^[46]	ResNet (Best)	DenseNet (Best)	Ours
Ate.	0.7160	0.7330	0.7670	0.7795	0.8530	0.5686	0.8467	0.8587
Car.	0.8070	0.8580	0.8830	0.8816	0.9390	0.5859	0.9287	0.9520
Eff.	0.7840	0.8060	0.8280	0.8268	0.9030	0.5762	0.8996	0.9113
Inf.	0.6090	0.6750	0.7090	0.6894	0.7540	0.5068	0.7348	0.7541
Mas.	0.7060	0.7270	0.8210	0.8307	0.9020	0.4616	0.8867	0.9126
Nod.	0.6710	0.7780	0.7580	0.7814	0.8280	0.5130	0.8130	0.8174
Pna.	0.6330	0.6900	0.7310	0.7354	0.7740	0.4692	0.8017	0.8086
Pnx.	0.8060	0.8050	0.8460	0.8513	0.9210	0.5897	0.9006	0.9317
Con.	0.7080	0.7170	0.7450	0.7542	0.8420	0.5948	0.8321	0.8467
Ede.	0.8350	0.8060	0.8350	0.8496	0.9240	0.4560	0.9045	0.9340
Em.	0.8150	0.8420	0.8950	0.9249	0.9320	0.4011	0.9390	0.9506
Fib.	0.7690	0.7570	0.8180	0.8219	0.8640	0.5261	0.8449	0.8793
PT.	0.7080	0.7240	0.7610	0.7925	0.8370	0.4707	0.8134	0.8572
Hen.	0.7670	0.8240	0.8960	0.9323	0.9210	0.4738	0.8921	0.9771
Ave.	0.7380	0.7673	0.8066	0.8180	0.8710	0.5138	0.8598	0.8850

4.7 实验结果可视化

将测试集图像分别输入 DCSN 分类网络中，处理之后将网络的最后一个卷积层特征提取出来，并通过全局池化操作得到权重，之后根据公式(3-22)，得到类别分数生成类激活图（CAM）^[40]，并使用 OpenCV 中的 applyColorMap 函数进行可视化显示。

使用此方法标记胸部 X 光片的结果如图 4-13 所示，图中以特殊高亮形式表示区域的显著性程度，其中红色表示很重要，橙色、绿色和蓝色按顺序表示重要性逐级减弱。

由图 4-13 可知，本文提出的双通道分离网络（DCSN）在测试集图像的可视化定位任务中所显示的重要区域与预标注文件中的边界框区域有很高的重叠度，证明该网络所提取的特征拥有卓越的定位能力，在热力图中可准确显示特定病理的所在区域，故可通过该网络的可视化功能在临床诊断中为放射科医生提供分类依据。

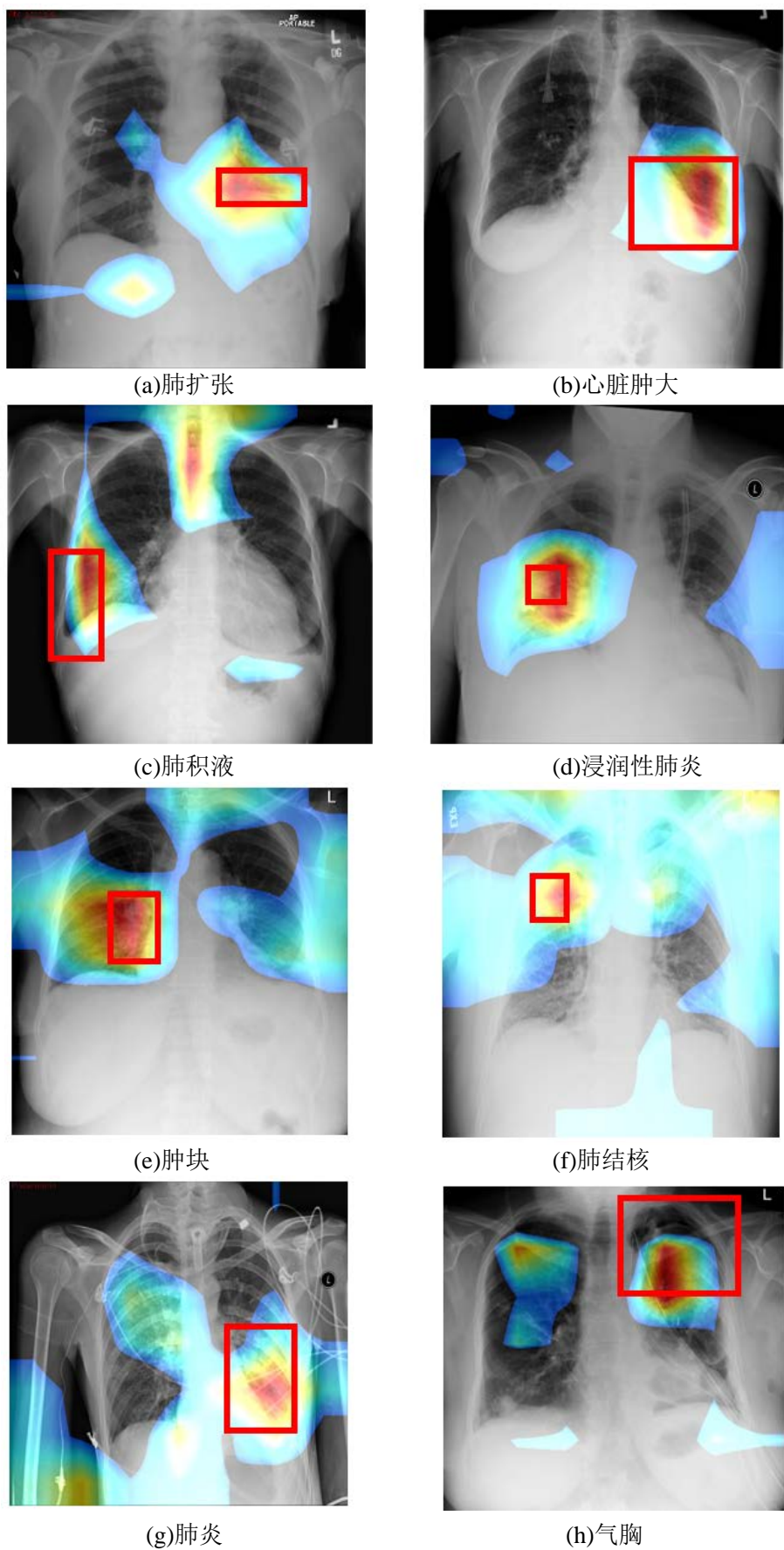


图4-13对于特定病理分类的显著区域进行可视化定位，图中红色框为标注文件中的边界框

4.8 本章小结

本章通过三组对比实验，在相同数据集上分别使用 ResNet、DenseNet 和 DCSN 网络进行训练，并将实验结果分别绘制 ROC 曲线并对比 AUC 分数。测试了仅使用一种学习网络与使用提出的双分支的通道分离网络在 CXR 图像分类任务性能上的差异。实验结果表明，使用 CSDN 网络训练的模型比单独使用 ResNet 或 DenseNet 有更优越的性能表现。同时与现有的部分最优算法相比，CSDN 在十四种胸部疾病的分类任务中，AUC 分数几乎均有明显提升，表明双通道分离的网络结构可有效实现特征重用和新特征的探索，获得更好的层次特征表示，同时在定位任务中也表现出了较高的准确率。该实验结果验证了本文提出的 CSDN 网络结构的确可获得分类和定位性能上的改进。

第五章 全文总结

5.1 结论

本文提出了一种基于双通道分离网络（Dual Channel Split Network, DCSN）的胸部病变的差异性识别方法，将输入数据提取的特征图分离输入到网络的上下两个不同分支中进行处理，最终得到全局融合特征进行分类，这种融合不同网络的层次特征的方法可以有效实现原始图像的特征复用和新特征的探索，可在分类任务中获得更好的特征表示，得到了较高的准确率，尤其是在胸部 X 光图像这种显著特征区域较少且像素间差异较小的图像类型的分类任务上表现出了更加优异的结果，同时实验中对病变区域的可视化标注不仅使分类依据更加直观，也大大缩短了放射科医生的二次诊断时间，所以，本文在实验中训练出的网络模型可作为一种新型的计算机辅助系统工具，在医疗图像的胸部 X 光片诊断领域发挥出一定作用，使胸部疾病的诊断过程更方便快捷，确诊率也更高。

5.2 未来展望

深度学习在医学图像处理领域的探索虽然还处于初期阶段，但是已经在许多诊断任务上获得了非常优异的效果，比如疾病分类，病变检测，器官分割，报告检索等，有些深度学习算法的表现甚至已经超越了一些专业医师，这些成果都充分证明了将深度学习技术迁移到医学领域的可行性，相关算法及专用深度网络模型的快速发展也显示出其巨大的潜力，例如本文的实验证明了通道分离操作有利于增加网络深度和多层分类特征的多样性，所以未来可以通过探索新的分支策略或不同信息融合方案来进一步提高模型表现结果，同时该思想也可以迁移到除 X 光图像之外的其他类型的医学图像处理上，如 CT 图像，超声波图像或核磁共振图像。相信之后关于常见胸部疾病的相关探索也一定可以取得更优秀的成果，为提高临床诊断准确率，缩短诊断周期提供更方便高效的计算机医疗辅助系统。

深度学习技术在医学图像领域飞速的同时也面临着挑战，一方面由于医疗数据的保密性和复杂性通常导致大型的医疗数据集难以获得，同时医学图像的标注往往需要专业的医师来完成，过程中需要耗费大量的人力和物力。另一方面，由于医疗领域属于跟生命紧密相关的特殊领域，对相关辅助诊断系统的错误容忍度极低，而深度学习算法中存在的一些不可解释性使得该类方法的社会接受度较低，人们不能像相信专业医生一样相信计算机视觉方法。所以，未来如何快速高效地获得可靠的医疗数据集来提高算法效果以及如何让大众相信并接受相关计算机技术的确诊率和可靠性，是我们之后仍然需要努力去解决的问题。

参考文献

- [1] 张贵英. 计算机图像处理技术在医学中的应用[D], 2012.
- [2] World Health Organization. Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children[R]. Geneva: World Health Organization, 2001.
- [3] Franquet T. Imaging of pneumonia: trends and algorithms[J]. European Respiratory Journal, 2001, 18(1): 196-208.
- [4] Cherian T, Mulholland E K, Carlin J B, et al. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies[J]. Bulletin of the World Health Organization, 2005, 83: 353-359.
- [5] CDC, 2017. URL <https://www.cdc.gov/features/pneumonia/index.html>
- [6] Lodwick G S, Haun C L, Smith W E, et al. Computer diagnosis of primary bone tumors: A preliminary report[J]. Radiology, 1963, 80(2): 273-275.
- [7] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential[J]. Computerized medical imaging and graphics, 2007, 31(4-5): 198-211.
- [8] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436.
- [9] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.
- [10] Chapman W W, Fizman M, Chapman B E, et al. A comparison of classification algorithms to automatically identify chest X-ray reports that support pneumonia[J]. Journal of biomedical informatics, 2001, 34(1): 4-14.
- [11] Ye X, Lin X, Dehmeshki J, et al. Shape-based computer-aided detection of lung nodules in thoracic CT images[J]. IEEE Transactions on Biomedical Engineering, 2009, 56(7): 1810-1820.
- [12] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. science, 2006, 313(5786): 504-507.
- [13] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. Jama, 2016, 316(22): 2402-2410.
- [14] 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.
- [15] Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks[J]. Radiology, 2017, 284(2): 574-582.
- [16] Islam M T, Aowal M A, Minhaz A T, et al. Abnormality detection and localization in chest x-rays using deep convolutional neural networks[J]. arXiv preprint arXiv:1705.09850, 2017.
- [17] Shiraishi J , Katsuragawa S , Ikezoe J , et al. Development of a Digital Image Database for Chest Radiographs With and Without a Lung Nodule[J]. American Journal of Roentgenology, 2000, 174(1):71-74.
- [18] Demner-Fushman D , Kohli M D , Rosenman M B , et al. Preparing a collection of radiology examinations for distribution and retrieval[J]. Journal of the American Medical Informatics Association, 2015:ocv080.
- [19] Armato S , Mclellan G , M McNitt - Gray, et al. WE - B - 201B - 02: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Public Database of CT Scans for Lung Nodule Analysis[J]. Medical Physics, 2010, 37(6Part6):3416-3417.
- [20] Wang X, Peng Y, Lu L, et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2097-2106.
- [21] Yan K, Wang X, Lu L, et al. DeepLesion: Automated Deep Mining, Categorization and Detection of Significant Radiology Image Findings using Large-Scale Clinical Lesion Annotations[J]. 2017.
- [22] Rajpurkar P, Hannun A Y, Haghpanahi M, et al. Cardiologist-level arrhythmia detection with convolutional neural networks[J].arXivpreprint arXiv:1707.01836, 2017.
- [23] Grewal M, Srivastava M M, Kumar P, et al. Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans[C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, 2018: 281-284.
- [24] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.
- [25] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale

- image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [26] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [27] 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.
- [28] Yao L, Pohlenz E, Dagunts D, et al. Learning to diagnose from scratch by exploiting dependencies among labels[J]. arXiv preprint arXiv:1710.10501, 2017.
- [29] 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.
- [30] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [31] Kumar P , Grewal M , Srivastava M M . Boosted Cascaded Convnets for Multilabel Classification of Thoracic Diseases in Chest Radiographs[J]. 2017.
- [32] Rajpurkar P , Irvin J , Zhu K , et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning[J]. 2017.
- [33] Guan Q, Huang Y. Multi-label chest X-ray image classification via category-wise residual attention learning[J]. Pattern Recognition Letters, 2018.
- [34] Li Z, Wang C, Han M, et al. Thoracic disease identification and localization with limited supervision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8290-8299.
- [35] Guan Q, Huang Y, Zhong Z, et al. Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification[J]. arXiv preprint arXiv:1801.09927, 2018.
- [36] Hahnloser R H , Sarpeshkar R , Mahowald M A , et al. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit.[J]. Nature, 2000, 405(6789):947-951.
- [37] Han J , Moraga C . The influence of the sigmoid function parameters on the speed of backpropagation learning[C]// International Workshop on Artificial Neural Networks. Springer, Berlin, Heidelberg, 1995.
- [38] 张小威. 基于深度学习的图像美学质量评价[D]. 杭州电子科技大学, 2017.
- [39] 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.

- [40] Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2921-2929.
- [41] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. arXiv preprint arXiv:1207.0580, 2012.
- [42] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1026-1034.
- [43] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[C]//Proceedings of the thirteenth international conference on artificial intelligence and statistics. 2010: 249-256.
- [44] Yao L, Poblens E, Dagunts D, et al. Learning to diagnose from scratch by exploiting dependencies among labels[J]. arXiv preprint arXiv:1710.10501, 2017.
- [45] Guendel S, Grbic S, Georgescu B, et al. Learning to recognize abnormalities in chest x-rays with location-aware dense networks[C]//Iberoamerican Congress on Pattern Recognition. Springer, Cham, 2018: 757-765.
- [46] Guan Q, Huang Y, Zhong Z, et al. Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification[J]. arXiv preprint arXiv:1801.09927, 2018.

致 谢

在毕业设计的这段时间内，首先要感谢给予我很多指导和帮助的两位导师：清华大学的王好谦老师和本校的王毅老师，从开题到最终定稿，两位老师都花费了很多时间和精力，王毅老师更是不辞辛劳的多次帮我改论文，指出我论文的每一处失误，让我能为本科交出一份满意的答卷，在这里，请允许我向你们表达我最衷心的感谢，感谢老师的付出，祝愿您们工作顺利，身体健康。

此外，还要感谢实验室的师兄师姐们，在我初来乍到还不熟悉科研环境时为我排忧解难，当我实验上遇到问题，宋磊师兄和李萌师兄都会耐心帮我解答，也要感谢张宇伦师兄在我实验中期遇到瓶颈时建议我尝试另一种方法，为我开拓了新思路，我才能得到满意的实验结果，我毕业设计的很多成果与你们的无私帮助是分不开的，你们努力和严谨的科研态度也深深感染了我，在你们的身上我学到了科研工作者的刻苦与踏实，从不轻言放弃。相信当我踏入研究生，这将成为我科研道路上的第一笔财富，谢谢你们。

我还要感谢我的室友们，从大一到大四，一路上有你们的陪伴真的很幸运。与你们一起学习，一起逛街吃饭，一起熬夜备考的日子是我大学中最纯真最无忧无虑的时光，很开心能成为云B238宿舍中的一员，我们是来自东南西北的四个北方人，我们是永远的张小越。因为有了你们的包容和鼓励，我才能有这么美好的大学时光，也许我们即将毕业，去向不同的地方，但这并不是分离，愿我们是一辈子的好朋友，愿你们所求皆如意，所行化坦途。

最后，感谢我的父母，无论我成功或失败，积极或颓废，你们都在我身后支持我，是你们的开明与鼓励才能让我有了宠辱不惊的资本，让我有对抗全世界的勇气，让我难过时不彷徨，兴奋时不自负，有了你们才有了我，我会继续成长，成长到足以保护你们，希望那一天明天就到来。

愿我心怀梦想，不丢信仰，当一个苦读勤学的赶路者。

愿我看透疾苦，深爱人间，做一个有爱有光的善良人。

毕业设计小结

不知不觉，从开题到准备答辩，已经过去了快 6 个月的时间，毕业设计也接近尾声，不禁感慨万千，不敢相信自己真的独立地完成了一个课题，也有些怀疑自己的结果能不能得到老师们的认可，总之喜忧参半，收获颇多。

最开始拿到选题时，自己是茫然的，学了三年的电子科学，突然换到医学图像处理，一个自己之前从来没有接触过的领域，开始是怀疑和动摇的，不知道自己能不能把这个选题完成。开始也确实走了很多弯路，文献看来很多，实际收获却很少，甚至一开始没有找到解决问题的正确思路，实验室的师兄们看出我的偏离，建议我先从基础知识开始学起，先掌握图像处理的大框架。这一方法果然奏效，我渐渐入了门，了解了图像处理方面的很多原理和方法，然后接触深度学习，被卷积神经网络的神奇所吸引，当我独立训练出第一个网络模型时，兴奋之情溢于言表。从那以后，我开始有选择性的阅读文献，寻找提高任务性能的新方法，期间经历了很多次失败，甚至花费一个多星期训练出的模型，到最后根本不能用。这里要再次感谢两位导师和同实验室的同学们，他们为我提供了很多新思路，让我在后续实验过程中才能得到满意的结果。我也从中学到了很多。一次毕业设计不像一次考试，只需要掌握重点，也不像我本科参加的很多机器人比赛，因为毕设要求学生对一个领域有完整的知识体系，要条理清晰，思路明确，知道自己要干什么，怎么去做才能在正确的科研道路上得到一个好的成果。这是学校在我们踏入研究生生活之前给我们上的重要的一课。

虽然自己的结果还有很多不足之处，未来还有很多需要改进的地方，但我相信只要保持紧张，一直努力就不会后悔。答辩还未开始，希望自己能把自己这半年来做的努力给老师们都展现出来，为我的本科生涯画上一个圆满的句号。