

# 基于深度学习的肺部多疾病自动诊断

苟鑫

(西北民族大学 甘肃兰州 730030)

**摘要:** 肺部疾病是常见的高发性疾病, X线检查是早期发现和诊断肺部疾病的首选影像学方法, 现行医学影像人工阅片方式受医生水平和经验限制, 并且容易受主观因素影响出现误诊和漏诊的现象。因此本文研究了基于深度学习的肺部疾病自动智能分类, 目的在于辅助医师高效准确地进行肺部疾病诊断。首先利用对比度图像自适应阈值均衡化算法对数据进行了预处理, 通过几何变换方法对数据集进行了扩充, 从而构建了适合于深度学习模型的肺部疾病 X 线图像数据集。然后选用 ResNet-50 作为肺部疾病分类模型, 利用 GOOGLE 公司的开源深度学习框架 TensorFlow 进行了实验验证。实验结果表明 ResNet-50 模型在正常和肺炎两类肺部 X 光图像的特征提取和识别上效果最佳准确率分别达到了 93% 和 86%。

**关键词:** ResNet-50; X 线医学影像; 肺部疾病分类

## 引言

肺是人体呼吸系统最重要的器官, 是人体与自然之间的进行物质交换的通道, 对维持人体生命健康有着重要作用。随着社会的发展, 生活环境的改变, 肺部疾病也受到越来越多人的重视。肺部病变种类繁多, 肺炎、气胸、肺结节等都是目前比较严重的肺部疾病, 其中肺结核在全球呼吸道传染性疾病中最为常见, 因为耐多药性患者比例的不断增多, 该病在传染性疾病中死亡率排名第二<sup>[1]</sup>。在我国肺炎是比较常见的高发性肺部疾病, 尤其儿童、老人和机体免疫较差的人, 极易受到肺炎支原体的攻击而导致患病<sup>[2]</sup>。

利用从历史数据中提取的疾病特征, 构建基于深度学习技术的计算机辅助诊断系统, 不仅能有效的提高工作效率, 还能大大减少出现漏诊、误诊的概率<sup>[3]</sup>。深度学习技术在医学领域得到了广泛的运用。在眼底视网膜图像识别中, 夏婷<sup>[4]</sup>使用多模型融合技术将 SE-ResNet 和 ResNet 进行加权模型融合对眼底视网膜彩图进行分类研究, 在疾病分类中 AUC 达到了 0.9458。在肺部图像识别中, 张建波<sup>[5]</sup>使用改进的 CheXnet 模型对 14 种肺部疾病智能诊断的平均 AUROC 达到了 0.8505。

目前, 在深度学习的肺部图像诊断方面, 大多数的工作者所进行的研究都是基于肺部 CT 图像疾病的诊断, 对于利用 X 光图像进行多种疾病的综合诊断研究较少, 且生活中人们进行肺部检查时更常见的选择是 X 光检查, 故本文研究基于深度学习的肺部疾病自动诊断, 通过 ResNet-50 对 X 线医学图像进行肺结节、肺不张、肺炎和正常 4 种类别分类。

## 1 数据及预处理

### 1.1 实验数据

本次实验选择 NIH 研究员吕乐博士等整理的 ChestX-ray 数据集作为实验数据, 该数据集包括 30000 多名患者超过 100000 张正面视图的 X 光射线图像, 以及 14 类疾病的图像标签<sup>[6]</sup>, 每一张图像中可能存在多个标签。数据集标签包括肺不张、肺炎、肺气肿、实变、气胸、肺结节、心脏肥大等。

ChestX-ray 数据集的图像中有的只包含 14 类疾病中的一种, 有的包含其中多种疾病, 为了提高模型对于疾病识别的准确率, 避免无关因素的干扰, 本文对该数据集进行了筛选, 只选取了每张图像

中仅包含有肺结节、肺不张、肺炎的肺部 X 光图像和正常的肺部 X 光图像作为本文的数据集。最后得到的实验数据集如表 1 所示。

表 1 实验数据集

序号	类型	样本数量
1	肺结节	2539
2	肺不张	3951
3	肺炎	4576
4	正常	1583

### 1.2 数据预处理

图像预处理是图像分类任务中进行模型训练前必不可少的一项工作。首先需要图像进行缩放处理, 使得输入图像的大小满足模型的需要, 从而保证模型训练的准确性以及效率, 然后利用平移、旋转和水平翻转对缩放后的图像进行线性变换, 提高模型的鲁棒性。最后, 图像在经过以上一系列操作之后图像对比度发生细微下降, 因此本文使用限制对比度图像自适应阈值均衡化算法对变换后的图像进行处理使图像的特征更加清晰<sup>[7]</sup>。

模型训练的过程中需要将所有的图像数据读入内存中, 以便进行训练, 而原数据集中图像的分辨率大小为 1024 × 1024, 尺寸太大占用过多的内存, 可能导致程序崩溃或无法运行, 同时模型需要处理的数据量也会随之增加, 从而影响模型训练的速度, 且大尺寸的图像中包含着大量的细节信息, 而模型可能无法有效处理这些信息, 将对模型的准确率造成影响, 因此需要对图像进行缩放处理使其变换为 224 × 224 大小的图像。同时考虑到数据集中存在着类别不均衡的问题, 本文利用图像的几何变换: 平移、旋转和水平翻转对图像进行了扩展。

图像经过预处理之后对比度下降, 使得图像中的细节信息难以区分, 从而影响了模型的识别能力。因此本文使用自适应直方图均衡化算法对图像特征进行增强, 提高模型的识别能力<sup>[8]</sup>。自适应直方图均衡化算法是基于直方图进行的, 对于改善细节丢失具有很好的效果。该算法是将图像分为大小为 N × N 的若干个小块, 分别对这些小块进行直方图均衡化处理, 然后对计算每个小块的全局对比度, 以此为阈值将其进行二值化处理, 最后将所有小块的二值图拼接到一起, 得到最终的输出图像。

1.3 数据集构建

经过上述处理方法对实验数据集处理之后，最终得到数据样本 39545 张，其中肺不张 9950 张，肺结节 10049 张，肺炎 10048 张和正常 9498 张。在划分数据集时采用 8:1:1 的比例划分训练集、验证集和测试集，具体划分情况如表 2 所示。

表 2 扩展后数据情况

集合	肺不张	肺结节	肺炎	正常	合计
训练集	7961	8040	8039	7599	31639
验证集	995	1005	1005	950	3955
测试集	994	1004	1004	949	3951
总计	9950	10049	10048	9498	39545

2 肺部疾病分类模型

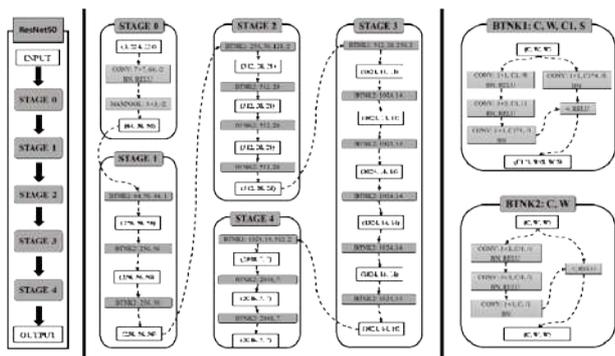


图 1 ResNet-50 结构

以深度学习为基础的图像分类任务的性能与模型的深度有关，但随着深度学习模型层数的不断增加可能会造成梯度消失和网络模型退化的现象<sup>[9]</sup>。为了保证分类任务的准确性同时避免出现上面的这些问题，本文选择 ResNet-50 进行研究。

ResNet-50 模型结构

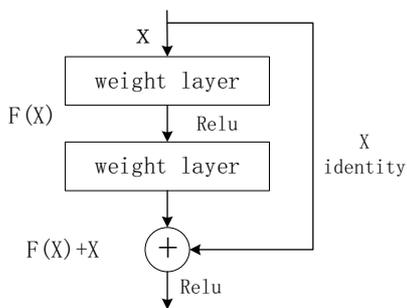


图 2 残差结构

ResNet-50 是由 Microsoft Research Asia 的 Kaiming He 等人提出的，其结构图如图 1 所示，若输入 ResNet-50 的图像大小为  $224 \times 224$  经过 STAGE0 中的卷积层处理之后输出图像的大小为  $112 \times 112$ ，再经过全局池化层处理得到的输出图像大小为  $56 \times 56$ ，依次进行 STAGE1、STAGE2、STAGE3 和 STAGE4 五个阶段处理，每一个阶段图像的大小缩减为原来的二分之一，而通道数扩展为原来的两倍。与之前提出的神经网络模型相比较，ResNet 模型具有更深的网络结构，并且有效的解决了由于网络层数加深所产生的梯度消失或梯度爆炸的问题，使得模型性能更加优良。而这都得益于残差块的引入，使得网络模型的深度可以达到更高，残差块中的输入可以跨层直接

传递到后续的层中，从而保留了更多的信息，减少了信息的丢失。残差块的结构如图 2 所示，其定义公式如公式 (6) 所示。

$$X_{j+1} = f[X_j + F(X_j W_j)] \quad (6)$$

式中  $X_j$  和  $X_{j+1}$  分别为第  $j$  个残差单元的输入和输出，而  $F(X_j W_j)$  是待学习的残差映射，其中  $W$  为卷积核， $f$  为激活函数，使用的 ReLU 函数，该函数能够将模型输出的结果中的负数全部转化为 0，而正数和零保存不变。

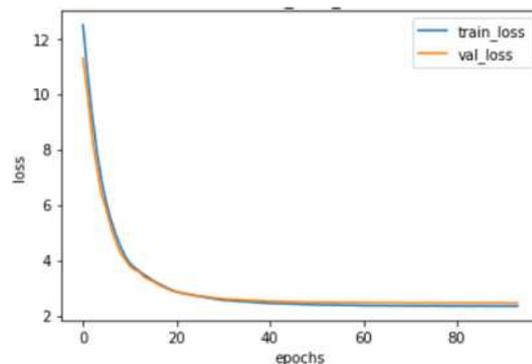
3 实验及结果分析

本文使用 Python 作为编程语言配合 Google 的深度学习开源框架 TensorFlow 搭建图像分类模型。硬件设备包括处理器和显卡，处理器型号为 Intel(R) Xeon(R) Platinum 8255C，包含 12 个内核，运行内存为 43G；而显卡为 RTX 3090。软件环境主要有操作系统 ubuntu 20.04、深度学习框架 TensorFlow2.9.0、显卡驱动库 CUDA 和 GPU 加速库 cudnn。

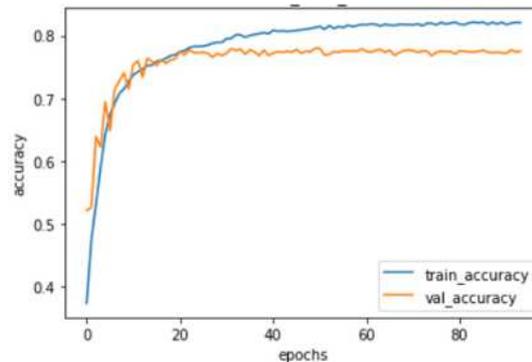
表 3 模型参数

损失函数	交叉熵
优化器	自适应学习率优化算法
激活函数	ReLU
学习率	0.00005
最大迭代次数	300

在模型的选择方面，本文选择 ResNet-50 作为肺部疾病分类的模型，且模型参数设置如表 3 所示，为了避免训练过程中出现过拟合现象，本文还设置了早退机制和学习率衰减策略，当模型在验证集上的准确率在 10 次迭代内依然没有提升，则停止模型的训练，并且模型每训练 9880 步学习率就衰减为原来的一半。



(a) Loss-Epochs 曲线图



(b) Accuracy-Epochs 曲线图

图 3 ResNet-50 训练结果

利用处理完毕的实验数据集对模型进行训练，模型的参数随着训练次数的迭代不断进行更新，loss 也在训练过程中降低，最终稳定在 2.34 左右，如图 3 中 (a) 所示，而模型训练的准确率最终也稳定在 0.82 左右，如图 3 中 (b) 所示。通过对模型训练结果的分析可以看出，模型在迭代到 30 次时，其验证集准确率趋于稳定，但训练集上的准确率仍然具有轻微的上升趋势，这可以从模型的损失值 loss 分析得出，模型在迭代到 30 次时，其损失值 loss 依然没有趋于稳定，还具有下降的趋势，这时随着模型 loss 的降低，模型的预测值和真实值之间的误差减小，模型的准确率随之增加，而验证集的准确率之所以没有随着 loss 的下降而出现如训练集上的提升，是因为在迭代 30 次之后，相邻两次迭代之间，被正确识别并分类的样本数量基本上相同，但图像在其对应标签上的预测概率增加，因此导致了验证集上 loss 下降，但准确率并未随之增加。

表 4 ResNet-50 模型实验结果

模型评估指标	指标数值大小
精确度	79.03%
召回率	76.18%
F1 评分	77.57%
准确率	77.7%

ResNet-50 模型的实验结果如表 3 所示，从实验结果可以看出模型在测试集上的准确率为 77.7%，精确度、召回率和 F1 评分都较高，其中召回率达到了 76.18%，证明该模型有 76.18% 的概率将肺部图像分类正确；F1 评分为 77.57%，反映了模型的性能较为不稳定，原因是肺不张和肺结节的医学特征极为相似，而模型对于细节的处理存在不足，导致模型在这两类图像的识别上存在着较大的误差。

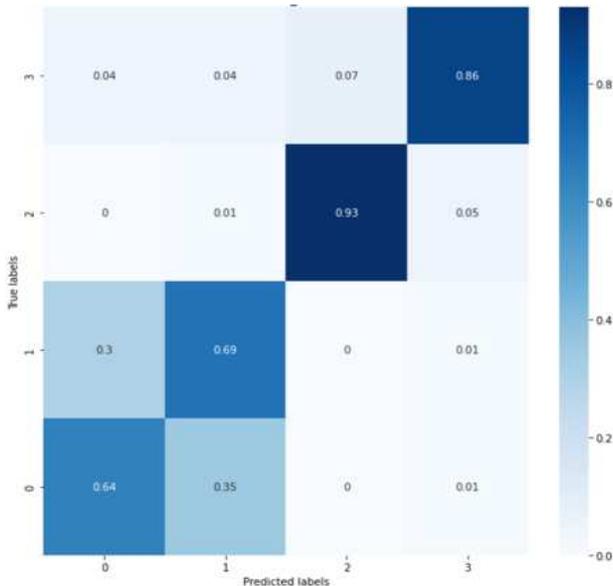


图 4 混淆矩阵图

模型在测试集上各类别的预测结果如图 4 所示，其中横坐标表示预测值，纵坐标表示实际值，0 表示肺不张，1 表示肺结节，2 表示正常，3 表示肺炎，从图中可以看出模型对肺炎和正常两类数据集的预测结果相较于其他类别较好，但在肺结节和肺不张的分类识别上效果较差，主要原因是两类疾病在 X 光图像上的医学特征都可能表现为肺组织密度的结节或斑点状突起，导致了模型在对两类疾病进行分类时造成了误判的现象。

## 4 总结与展望

### 4.1 总结

随着我国生活水平的提升，人们对于肺部疾病的关注度越来越高，X 检查是目前最常见的肺部检查方式，但易受医师水平和主观因素的影响。基于深度学习的肺部疾病诊断系统的出现将很好的帮助医师进行肺部疾病的诊断，不仅提高了诊断速度，而且还能减少误诊和漏诊的概率。本文的主要工作是：

(1) 利用限制对比度图像自适应阈值均衡算法对形态学变换后图像出现的对比度下降，造成的细节信息难以区分进行处理，从而提高模型的识别能力。

(2) 基于 Tensorflow 框架搭建 ResNet-50 模型对肺部图像进行四分类，其中模型在肺炎和正常两类图像上的分类结果分别达到 86% 和 93%，并且它们的准确度和 F1 评分分别达到了 93.93%、92.57% 和 92.31%、89.07%，证明模型在这两类图像的分类上效果较好且稳定。

### 4.2 展望

本文主要对肺结节、肺不张、肺炎和正常四类 X 光图像进行分类研究，模型在肺炎和正常图像的识别和分类上效果较好，但在肺结节和肺不张的分类上存在较为严重的误诊现象，这将是后期研究工作的重点。接下来的研究过着如下：

在模型中加入注意力机制，增加模型对于细节的处理能力，从而准确区分肺结节和肺不张在 X 光图像上的一些细微区别，从而正确识别两类疾病。

### 参考文献：

- [1]郭晓红,李一鸣.肺结核一线抗结核药物治疗的研究进展[J].医学综述,2021,27(05):948-951+956.
- [2]刘玉霞,邓宽国,尹石华,史健,王卫平,孙征,曲艳文,刘利.肺结核与支原体肺炎患者外周血淋巴细胞亚群及细胞因子分析[J].临床肺科杂志,2022,27(04):581-585.
- [3]王震,田军委.计算机视觉辅助训练系统研究——以射击类项目为例[J].西北大学学报(自然科学版),2013,43(04):563-567.DOI:10.16152/j.cnki.xdxbzr.2013.04.010.
- [4]夏秋婷.基于深度学习的眼底视网膜图像多疾病分类研究[D].杭州电子科技大学,2020.DOI:10.27075/d.cnki.ghzdc.2020.000261.
- [5]张建波.基于深度学习的肺部病识别与诊断研究[D].天津理工大学,2021.DOI:10.27360/d.cnki.gtdgy.2021.000493.
- [6] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmark on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017
- [7]唐文学.基于多特征融合与机器学习的散斑缺陷精确识别[D].电子科技大学,2021.DOI:10.27005/d.cnki.gdzku.2021.000472.
- [8]何迪.基于深度学习的肺部 X 光图像分类与目标检测[D].西安电子科技大学,2021.DOI:10.27389/d.cnki.gxadu.2021.002733.
- [9]倪锦园,张建勋,张馨月.深度宽残差网络注意力机制的人脸表情识别[J].重庆理工大学学报(自然科学),2023,37(01):177-185.

作者简介：苟鑫（2000 年-），男，云南昭通人。大学本科学历，现就读于西北民族大学。研究方向：深度学习，计算机视觉。

基金项目：西北民族大学 2022 年校级大学生创新训练计划项目（项目编号：X202210742166）