

# 《人工神经网络》大作业开题报告

李诗懿

计算机科学与技术系  
清华大学

shiyi-li15@mails.tsinghua.edu.cn

陈旭鹏

生命科学学院  
清华大学

xp-chen14@mails.tsinghua.edu.cn

## 1 任务定义

利用深度学习模型对X光（胸片）数据进行分类，用于帮助诊断心脏疾病。通过尝试用于图像处理领域的深度学习模型，试图加入专业医生的先验经验，对VSD、ASD、TOF等（先天性）心脏疾病进行分类诊断。胸片被广泛地应用于肺部和心脏的诊断，尤其是用于预诊断工作，让医生快速地从胸片中获取大量信息是具有挑战性的，如果能以较高的准确率进行自动诊断，就可以很大地帮助医院医生的诊断，尤其是对于偏远地区，先天性心脏病的诊断具有较大的意义。但是缺少标准的公开数据集与基准使得很难比较和建立一套所谓最好的预测模型。目前已有的基于深度学习模型的研究相较于传统的基于特征的模型可以有大约20%的分类准确率上的提升，但是目前的研究大多基于几种比较经典的深度学习模型的使用，在可靠性、多分类问题和小样本问题上还有很多未解决的问题。

## 2 数据集

数据集分为个人整理与公开来源两部分。通过从两家医院以及openi的公开数据集中收集到了11230张胸片。包括VSD、ASD、TOF等心脏病类型的病人胸片以及正常人胸片。另外包括病人的性别，年龄和患病类型注释。

Table 1: 数据集描述

来源	患病样本	正常样本	备注
南阳市第二人民医院	2826	500	DICOM格式
南阳市中心医院	-	-	DICOM格式，共3949张，未分类
openi数据(Indiana Hospital)	469	3486	jpg格式

以其中一个病人的胸片示例(figure1)，包括正面与侧面两张胸片（部分病人只有其中一张胸片）

## 3 挑战和基线

### 3.1 挑战

我们遇到的挑战主要有数据处理、标注、样本量与均衡性、单张样本信息量及可靠性等问题。

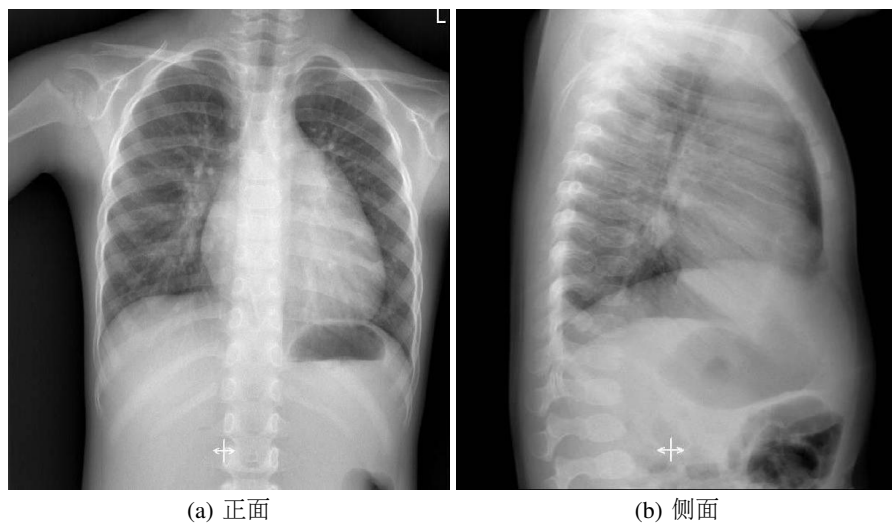


Figure 1: 胸片样例

- **数据处理与清洗**  
目前收集到的数据需要进行一系列的预处理才可以使用。需要去除隐私信息，筛选有标签的样本，调整到统一的分辨率等。
- **数据标注**  
经过向医生的请教，我们可能需要对图片进行一些具体的标注，这需要耗费一定的时间。
- **样本量与均衡性**  
目前来看可用的样本数量大约只占总样本的约百分之三十，因此整体数据集并不是很大，且其中ASD、VSD等疾病占比较大，一些比较罕见的心脏病亚型样本量较少。
- **单张样本信息量问题**  
胸片作为心脏病的诊断方法，最为方便和便宜，但是相比于CT和MRI数据，信息含量较低，部分图片比较模糊，轮廓边缘不清晰，比较难以使用。
- **可靠性问题**  
作为医疗图像问题，一个很大的挑战在于诊断的可靠性，我们需要思考如何分析并且保证模型正确地学习到了心脏的特征并进行分类。

### 3.2 基线

我们选取了几个比较有代表性的研究，展示其进行正常与否分类以及多分类的分类结果。

#### 3.2.1 研究1

其中一个研究选取了包括Alexnet、VGG以及ResNet在内的多种模型进行正常与否的分类。采取accuracy、AUC、sensitivity和specificity等指标进行评价。ROC曲线是true positive rate (TPR) 和false positive rate (FPR) 所画的图，其面积AUC值在0-1之间。正确分类的得病（阳性）样本为true positive (TP)，错误分类的得病样本为false negative (FN)，正确分类的正常样本为true negative (TN)，错误分类的正常样本为false positive (FP)

$$\begin{aligned}
 \text{TPR} = \text{sensitivity} &= \frac{TP}{TP + FN} \\
 \text{FPR} = 1 - \text{specificity} &= \frac{FP}{FP + TN}
 \end{aligned} \tag{1}$$

文章展示了使用dropout的几种网络的分类正常与否的结果（figure2）、使用ensemble方法的结果以及ROC曲线（figure3）

	ACCURACY (%)	AUC	SENSITIVITY (%)	SPECIFICITY (%)
ALEX NET	88.00%	<b>0.94</b>	88.00%	88.00%
VGG-16	<b>89.00%</b>	0.90	90.00%	88.00%
VGG-19	88.00%	0.88	86.00%	90.00%
RESNET-50	88.00%	0.90	88.00%	88.00%
RESNET-101	87.00%	0.91	82.00%	<b>92.00%</b>
RESNET-152	87.00%	0.88	<b>92.00%</b>	82.00%

Figure 2: 研究1结果图1

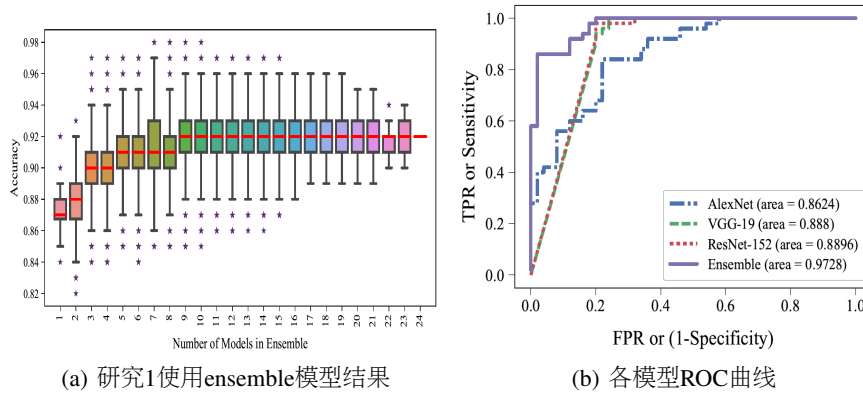


Figure 3: 研究1结果图2

### 3.2.2 研究2

第二个研究的特色是尝试了VGG-16和ResNet-101于多种心脏病亚型的诊断（判断是否得有该亚型心脏病）问题，结果如下（figure5）：

SINGLE DISEASE CLASSIFICATION RESULTS							
Disease Label	Specificity	VGG-16			Specificity	ResNet-101	
		Sensitivity	Top-3 Sensitivity			Sensitivity	Top-3 Sensitivity
normal	71.5%	78.5%	99.7%		70.5%	80.4%	99.5%
increased lung marking	79.4%	59.6%	100%		86.3%	43.2%	99.6%
aortosclerosis	97.0%	37.5%	92.4%		90.4%	61.1%	99.3%
increased heart shadow	99.8%	12.5%	50.0%		99.4%	25.0%	62.5%
pleural thickening	99.8%	16.7%	25.0%		99.6%	41.7%	58.3%
pulmonary interstitial hyperplasia	98.7%	57.1%	57.1%		99.9%	28.6%	42.9%
costophrenic angle blunting	99.9%	40.0%	60.0%		100%	60.0%	80.0%
pleural effusion	100%	60.0%	80.0%		99.9%	60.0%	100%
emphysema	99.8%	20.0%	20.0%		99.8%	20.0%	40.0%
bronchitis	99.6%	14.3%	28.6%		99.7%	28.6%	42.9%

Figure 4: 不同心脏病亚型的诊断效果

### 3.2.3 已有研究的问题:

- 只有针对正常与否的分类，或者是针对是否患有某种特定疾病的分类，缺乏多分类的结果。
- 可靠性不好，没有说明模型是否学习了正确的特征。

## 4 研究计划

我们希望在保证模型的准确度的同时，在模型的可靠性方面多进行一些探索。我们计划进行以下几个方法的工作：

- 了解利用胸片进行心脏病诊断的经验，帮助我们更好的设计模型。
- 对图片进行一些前期的预处理，通过常用的方式进行图像的augmentation，增加一定的数据量。
- 解决图像缺少标注的问题，目前的图像只有患有某种疾病的标注，没有对区域的细致标注，因此无法做区域的分割和检测，我们需要探索解决标注问题的方法。
- 对使用常用深度学习模型的结果进行分析，分析模型的可靠性。
- 探索如何解决模型的可靠性问题，使得模型可以重点学习图像的特定区域（与医生的经验结合），使得模型在保证准确率的同时更具有可靠性。

为此我们准备对图像数据进行基本的预处理，使用常用的深度学习模型进行分类，使用一些可视化方法分析学习结果，并且尝试让模型学习特定的区域（如心脏区域）

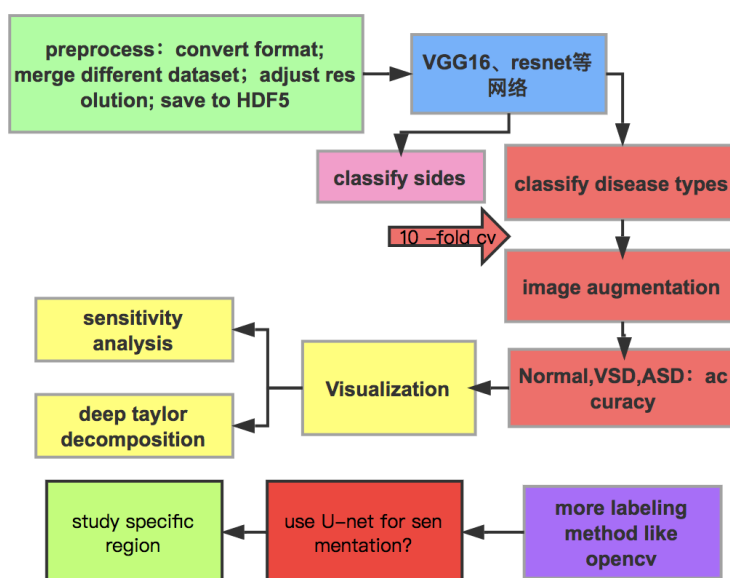


Figure 5: 计划流程图

## 5 可行性

我们认为我们的工作可能存在的一些困难，包括有用图像的筛选和预处理阶段的耗时、训练模型的耗时，如何解决标注问题与可靠性问题的探索上。尤其是标注过程可能会十分耗时，不能完全由人工标注，如何让模型学习特定区域也是一个可能的挑战。

我们认为使用人工加上opencv等工具进行标注的方法可能会节约一定的时间，可以尝试使用U-net对区域进行分割来实现对特定区域的学习，我们可以在带有GPU的服务器上训练我们的模型。细节的问题还有待具体实施过程中的逐步发现与解决。

## 参考文献

- [1]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
- [2]J. M. Carrillo-de Gea, G. Garcia-Mateos, J. L. Fernandez-Aleman, and J. L. Hernandez-Hernandez, "A computer-aided detection system for digital chest radiographs," Journal of Healthcare Engineering, vol.

- [3]S. Candemir, S. Jaeger,W. Lin, Z. Xue, S. Antani, and G. Thoma, "Automatic heart localization and radiographic index computation in chest x-rays," in SPIE Medical Imaging. International Society for Optics and Photonics, 2016, pp. 978 517–978 517. 2016, 2016.
- [4]A. Kumar, Y.-Y. Wang, K.-C. Liu, I.-C. Tsai, C.- C. Huang, and N. Hung, "Distinguishing normal and pulmonary edema chest x-ray using gabor filter and svm," in Bioelectronics and Bioinformatics (ISBB), 2014 IEEE International Symposium on. IEEE, 2014, pp. 1–4.
- [5]G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," arXiv preprint arXiv:1702.05747, 2017.
- [6]Mohammad Tariqul Islam, Md Abdul Aowal, Ahmed Tahseen Minhaz, Khalid Ashraf, "Abnormality Detection and Localization in Chest X-Rays using Deep Convolutional Neural Networks", arXiv preprint arXiv:1705.09850v3, 2017
- [7]Yuxi Dong, Yuchao Pan, Jun Zhang and Wei Xu, "Learning to Read Chest X-Ray Images from 16000+ Examples Using CNN", Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2017 IEEE/ACM International Conference on.