

医学图象分类报告

赵旺 无57 2015011157

王复英 无57 2015011175

孙铭 无56 2015011127

2018 年 6 月 15 日

Contents

1	背景调研	3
2	相关工作总结	3
2.1	基于非图像数据的ResNet50医学图像分类	3
2.2	基于位置感知密集网络的胸部X射线异常检测	3
2.3	基于Attention CNN的X射线胸片分类	4
3	方法描述	4
3.1	全局-局部信息融合	5
3.1.1	FPN结构	5
3.1.2	PSP结构	6
3.1.3	多尺度输入	6
3.1.4	两阶段结构	7
3.2	Loss函数改进	7
3.3	网络结构改进	8
4	实验结果与展示	8
5	思考讨论	9
6	小组分工	10
7	致谢	10

1 背景调研

据统计, 仅在美国, 每年就有超过100万人因为肺炎住院, 5万人因为该病而死亡 (CDC, 2017)。因此, 肺炎的及时诊断和治疗是当前的医学领域一大课题, 而当前诊断肺炎的最好方法是X光胸片检查。但是对于医生来讲, 从胸部X光检查照片中检测肺炎不是一个简单的任务, 肺炎在照片上的表现经常难以识别——它可能会与其他病症重叠, 也可能与其他很多良性异常类似, 这都导致放射科医师在诊断肺炎时表现差异很大, 而这些诊断上的差异往往是代价巨大的。

随着卷积神经网络在图象分类领域的广泛应用, 医学图象领域也出现了一大批深度学习的成功应用。肺炎诊断、以及更广泛的X光胸片诊断作为重要的医学课题, 也受到了很多的关注。Islam使用开源的OpenI数据集研究了多种卷积神经网络的差异; Yao等人[2]在此基础上给出了利用标签统计依赖关系的更好模型。Wang等人[1]发布了ChestX-ray14数据集, 其中包含了来自30805个患者的112120张正面扫描X光胸片, 每张图片都含有14种疾病的标注信息。Rajpurkar等人[3]提出CheXNet, 以densenet[4]为基础, 在ChestX-ray14数据集和实际胸片图像诊断中都取得了非常好的效果, 甚至已经超过了人类医师的判断。数据的丰富和研究人员的努力大大推动了X光胸片诊断的进步, 本次大作业即是在CheXNet的基础上进行改进, 试图得到更精确的分类结果。

2 相关工作总结

我们调研Chexnet之后的对ChestX-ray14数据集分类的工作, 然后以此为启发在Baseline的基础上做一些改进。

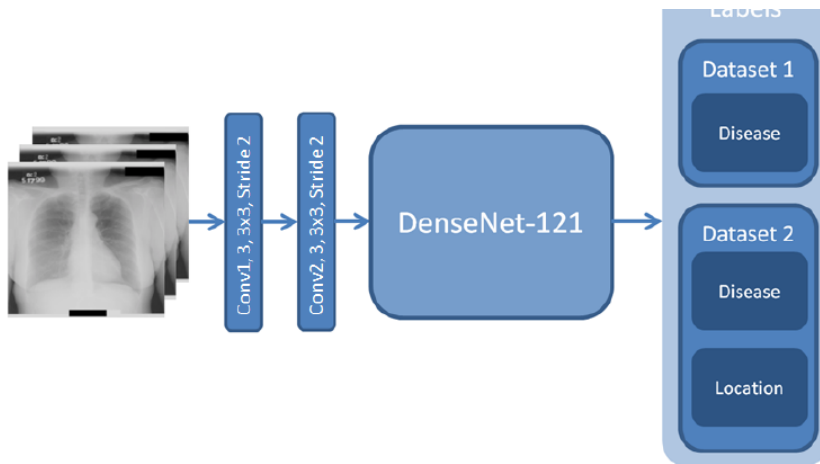
2.1 基于非图像数据的ResNet50医学图像分类

本文[10]作者详细调研了ResNet50在X射线胸片分类问题中的有效性。在此领域先前工作的基础上, 作者考虑了有无Finetune的迁移学习, 并从头开始训练了专门的X射线网络。为了充分利用X射线数据的高空间分辨率, 作者还加入了扩展的ResNet-50架构和一个在分类过程中集成非图像数据 (患者年龄, 性别和采集类型) 的网络。在系统性评估中, 作者使用5倍重采样和多标签损失函数, 评估ROC统计的不同病理分类方法的性能, 并使用秩相关方法分析分类器之间的差异。我们观察到已取得的性能有相当大的差距, 并得出结论: X射线专用ResNet-50集成了非图像数据, 可获得最佳的整体结果。

2.2 基于位置感知密集网络的胸部X射线异常检测

随着大规模数据集ChestX-ray14的发布, 人们已经提出了很多用于对胸部X射线上的病理进行分类的方法。然而, 大多数方法对数据集的分割是随机的, 忽略了很多同一患者出现在训

练和测试集合中的情况。另外，大多数方法不能明确地利用异常的空间信息或利用高分辨率图像。作者提出了一种基于位置感知密集网络（DNetLoc）的新方法[11]，将高分辨率图像数据和空间信息结合起来用于异常分类。此模型实现了在单个基准数据集（ChestX-Ray14）上公布的训练集和测试集分割方法的最佳平均AUC得分，当明确使用病理位置信息时AUC得分会得到提高。



作者采用了两个数据集：ChestX-Ray14和PLCO[12]。其中PLCO数据集包含了患病位置的信息，因此作者增加了空间位置标签来利用空间信息，最终实现了很好的分类效果。

2.3 基于Attention CNN的X射线胸片分类

本文[13]考虑了X线胸片疾病分类的任务。现有方法通常使用全局图像作为网络学习的输入。这种方法在两个方面受到限制：

1) 胸部疾病通常发生在具有疾病特异性的局部区域中，使用全局图像训练CNN可能会受不相关的嘈杂区域的影响。

2) 由于某些CXR图像的对齐不良，边界不规则的存在限制了网络的性能。

在本文中，作者通过提出了具有三个分支的AG-CNN来解决上述问题。作者在ChestX-ray14数据集上做了综合实验。首先作者以ResNet50为Baseline，得到的平均AUC为0.841。通过将局部信息与全局信息相结合后，AG-CNN将平均AUC提高到0.868。在使用DenseNet-121作为AG-CNN的基本结构时，平均AUC达到0.871，这是ChestX-Ray14上最高的分类准确率。

3 方法描述

在CheXNet获得较好的分类效果背后，我们也看到了其可能存在的一些不足之处，并基于

此提出了我们的改进方案，归纳起来有如下几点：

- 网络中增加更多的局部信息，实现更好的全局-局部信息融合
- 实现多尺度输入和处理，减小输入下采样带来的关键细节丢失
- 改进分类输出的Loss，增大类间差异，减小类内差异
- 换用其他的基础网络结构

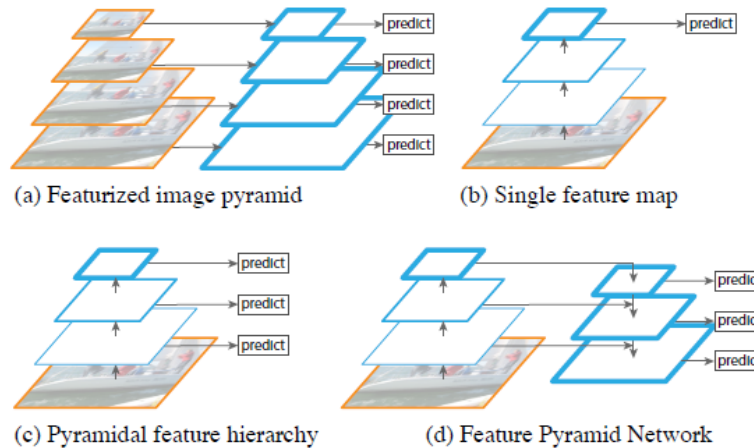
具体描述如下：

3.1 全局-局部信息融合

在疾病分类的时候，整体图像的信息和局部病变的特征同等重要。全局信息有助于帮助我们迅速找到大致正确的结果，同时整体层面上的决策不至于被局部噪声等影响；而局部信息有助于得到更加精细和准确的分类结果，在区分比较相似的病变时尤其如此。因此我们认为，在densenet或者resnet的基础上改进，增加更好的全局-局部信息融合是有必要的。

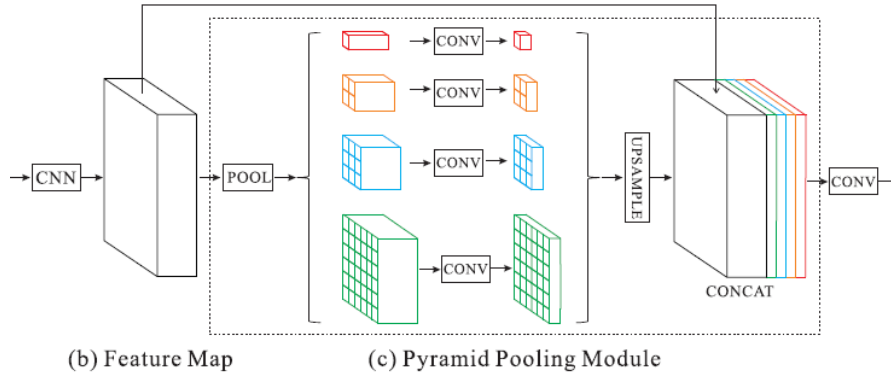
3.1.1 FPN结构

在[5]中Lin等人提出了FPN的结构来更好得到空间信息，如图所示，以减小由于特征图分辨率减小、语义增强带来的空间信息损失问题；实质上这可以看作是一种全局-局部信息的融合：语义信息更多存在于高层和全局中，而局部空间细节在不断的pooling和采样中逐渐损失；而FPN通过这样top-down的方式融合了空间细节和全局语义，在detection任务上取得了非常好的结果；我们认为这样的方法同样适用于既需要全局语义信息、又需要空间细节的胸片疾病诊断任务；



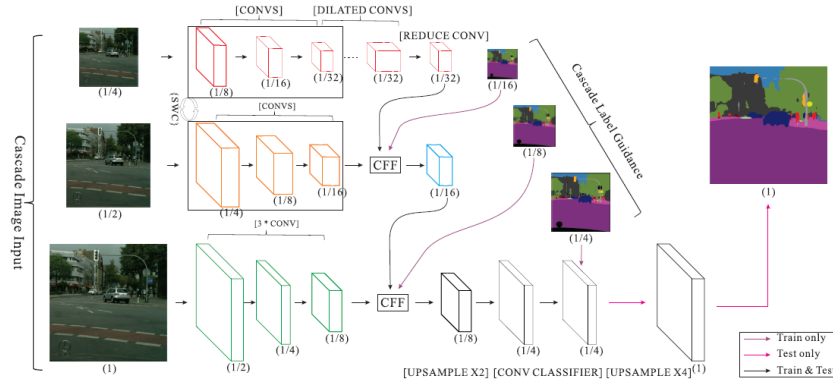
3.1.2 PSP结构

在[6]中Zhao等人提出了一种局部-全局Pooling的方式，以求在pooling操作的过程中最大程度地保留局部信息，以及融合局部和全局信息，如图6所示。通过不同尺度的pooling然后特征图插值concat的操作，细节信息和语义信息同时被保留；事实上其做法和FPN的初衷是类似的，但是更加精细，也更加结构化，也是一种不错的局部-全局信息融合方式。



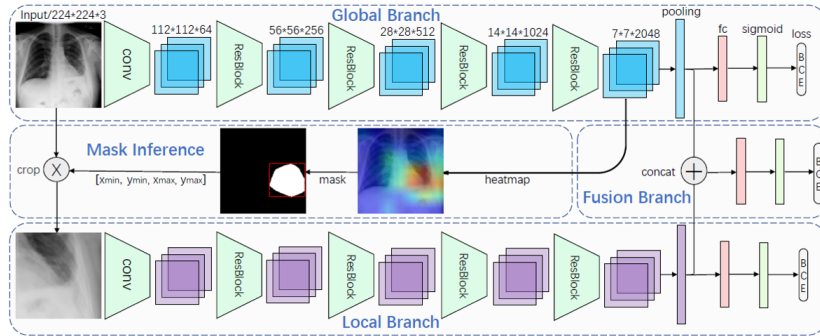
3.1.3 多尺度输入

由于Resnet-18等特征提取网络的最后全连接层是固定输入大小的，因此输入网络的图片必须都是固定大小，如 224×224 ，但是原图是 1024×1024 大小，直接降采样到 224×224 会损失很多的细节信息，有可能对分类结果是有影响的；受[8]中网络结构的启发，我们实现了多尺度输入的网络，原图分辨率的图片经过较浅的网络，保留丰富的空间信息并且降低复杂度，而 224×224 分辨率的图片经过较深网络，提取较强的语义信息；最后多尺度的输出进行插值concat，之后经过pooling输入全连接分类层，得到最终的分类结果。



3.1.4 两阶段结构

[9]中基于生成的热图给出一种two-stage的方法，第一阶段用传统方法提取特征并生成热图，在热图上定位出对结果影响最大的区域并crop出来，再将这部分图片送入第二阶段的网络提取特征，将两个阶段的特征融合后输出分类结果。如下图所示

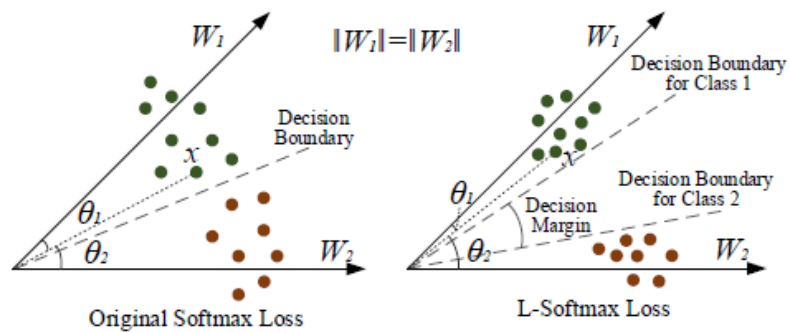


论文中作者给出的结果显示，该方法明显提升了分类的精度，证明了局部信息的重要性。由于时间精力的限制，我们没有复现该结果。

3.2 Loss函数改进

在分类问题中，已经有不少的文章注意到loss的重要性。传统的softmax或者sigmoid激活加交叉熵的Loss函数虽然简单方便，但是这样训练出的网络对类内和类间差异的区分却不是很好，即如果最后输出有两类可能性都比较高，理论上是类间差异过小或者类内差异较大，此时应该有较为强的惩罚，但是经过softmax或者sigmoid激活后，几乎只保留了最大项，这样就难以达到惩罚类间差异过小的目的。因此Liu等人[7]提出了Large-Margin Softmax Loss，如图所示。同

时还有众多的类似做法，试图达到同样的目的。我们认为这样的方法在分类问题上比较本质和有效的，因此我们试图应用或者尝试新的loss函数来更好解决这一问题，在尝试过程中发现的问题是由于现有标签中存在一张图同时患多种疾病和一种疾病也不患的可能，因此这些loss无法直接迁移使用。



3.3 网络结构改进

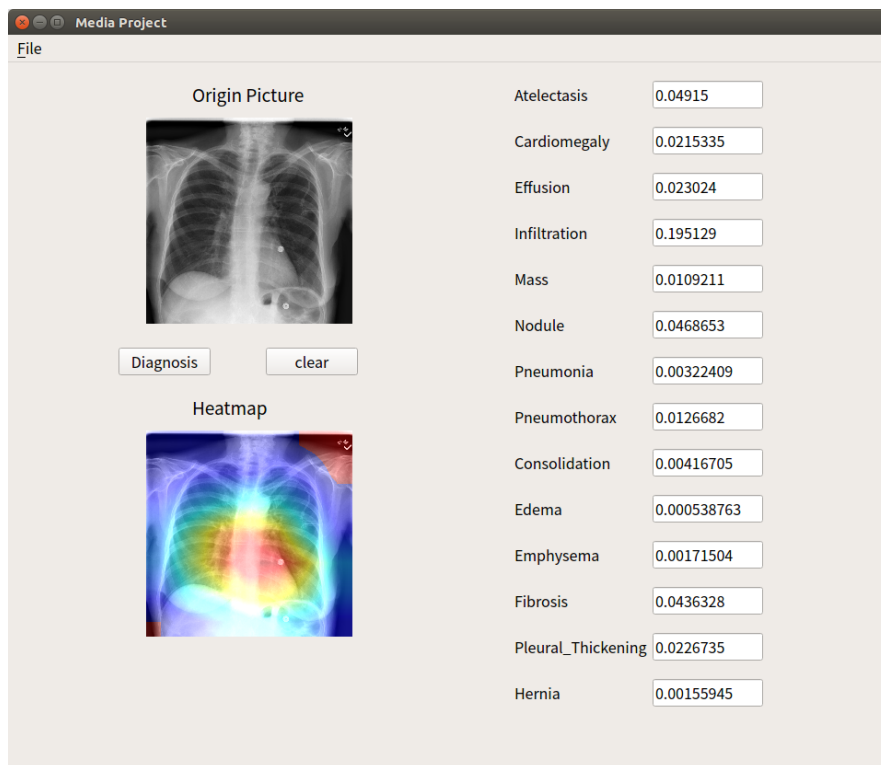
CheXNet中使用了DenseNet-121作为基础网络，我们尝试了其他几种不同的网络结构，如ResNet-18, ResNet-50, NasNet, Inception-v3等等，也取得了不错的效果。

4 实验结果与GUI展示

下表显示了实验结果

实验设置	AUROC Mean
ResNet-18 baseline	0.815944
ResNet-50	0.816823
Inception-v3	0.817215
DenseNet-121	0.841235
ResNet-50 + 15labels	0.806576
ResNet-50 + FPN	0.817053
ResNet-50 + PSP	0.816901
ResNet-50 + Multi-scale Input	0.817048

从结果中可以看出，大多数的尝试在baseline的基础上都有一些提升，但效果都比较有限；原因有可能是这些算法更适用于目标检测等任务，在分类问题上没有非常明显的作用；我们基于PyQt搭建了一个简洁的用户界面，可以实现用户选择一张X光胸片图像输入，输出热图以及患每种疾病的可能性，运行结果如下图所示：



5 思考讨论

本次大作业以医学图像分类为主题，以目前流行的深度学习和卷积神经网络为工具，在熟悉基础原理和算法的同时尝试着对算法进行改进，极大地丰富了组内同学的知识，提高了解决问题的能力和小组协作能力；小组同学认为，图像分类问题，本质上是图像特征的提取和比对，因此自编码器、支持向量机、神经网络等基于特征子空间的分类方法都可以使用，分类的好坏，一方面取决于特征提取的好坏，一方面也取决于特征比对的好坏；因此从特征提取网络和损失函数两方面来尝试改进是有道理的，而具体如何提升，什么方法比较适用，在前面已经详细描述；医学图像分类是一种特殊的任务，X光胸片图像更是如此，因此可以设计一些比较专用的算法来更好解决该问题，不需要要求可以泛化到任意图片。总的来说，医学图像分类前景广阔，在学术界和工业界都有重要的意义，同时当前的方法也还存在着不少欠缺的地方，值得我们进

一步去探究。

6 小组分工

小组中，赵旺和王复英主要负责算法实现和实验，孙铭主要负责算法调研和用户界面的搭建；三人贡献基本相同。

7 致谢

这次大作业的完成离不开助教和老师的辛勤付出，感谢助教提供的帮助和详细的答疑，感谢其他组同学的分享和启发，感谢老师的付出和指导！

8 参考文献

- [1] Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315, 2017.
- [2] Yao, Li, Poblens, Eric, Dagunts, Dmitry, Covington, Ben, Bernard, Devon, and Lyman, Kevin. Learning to diagnose from scratch by exploiting dependencies among labels. arXiv preprint arXiv:1710.10501, 2017.
- [3] Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning[J]. arXiv preprint arXiv:1711.05225, 2017.
- [4] Huang G, Liu Z, Weinberger K Q, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, 1(2): 3.
- [5] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//CVPR. 2017, 1(2): 4.
- [6] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2017: 2881-2890.
- [7] Liu W, Wen Y, Yu Z, et al. Large-Margin Softmax Loss for Convolutional Neural Networks[C]//ICML. 2016: 507-516.
- [8] Zhao H, Qi X, Shen X, et al. Icnet for real-time semantic segmentation on high-resolution images[J]. arXiv preprint arXiv:1704.08545, 2017.
- [9] 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.

- [10] Wen Y, Zhang K, Li Z, et al. A discriminative feature learning approach for deep face recognition[C]//European Conference on Computer Vision. Springer, Cham, 2016: 499-515.
- [11] Zhang M L, Zhou Z H. A review on multi-label learning algorithms[J]. IEEE transactions on knowledge and data engineering, 2014, 26(8): 1819-1837.
- [12] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//CVPR. 2017, 1(2): 4.
- [13] Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on. IEEE, 2016: 2921-2929.