

声明：我已知悉学校对于考试纪律的严肃规定，将秉持诚实守信宗旨，严守考试纪律，不作弊，不剽窃；若有违反学校考试纪律的行为，自愿接受学校严肃处理。

---

2019-2020 学年春季学期 COMP130032.01

## 《数字图像处理》期末报告

### 基于深度学习的新冠肺炎分类算法

---

学号：17300200036,

姓名：吴茜茜,

邮箱：17300200036@fudan.edu.cn

#### 摘要

新型冠状病毒 (Covid-19) 已在世界迅速传播, 由于每日病例的增加, 建立新冠肺炎的自动检测系统对跟踪疫情、诊断患者大有帮助。在本次课程项目中, 我们分别以新冠肺炎患者的胸部 X 光片和正常人胸部 X 光片为正负样本构建数据集, 使用预训练的肺部分割模型对原始 X 光图像处理, 之后分别使用三种基于卷积神经网络的模型 (ResNet18, MobileNet v2 和 DenseNet121) 来进行分类, 以识别出新冠肺炎患者。实验中 DenseNet 模型效果最佳, 获得了准确率 88.75% 和 F1 得分 88.9%, 证明了我们的模型具有较好的性能和鲁棒性。

#### 1 引言

新型冠状病毒 (Covid-19) 现成为袭击我们星球的最具传染性的疾病之一。自该病毒首次发现以来, 仅三个多月便已传播到 90 多个国家和地区, 感染了 3500000 多人。尽管当前的人工智能技术远远不能复制人类的智力, 但它们对跟踪疫情、诊断患者、判断病灶以及加快 Covid-19 的治疗过程大有帮助。

随着深度学习技术的发展, 卷积神经网络 (Convolutional Neural Network, 简称 CNN) 在医疗图像分类上应用广泛。它能够通过卷积运算, 自动地学习到图像域的特征, 有更强的表征能力, 因此提升了网络的性能。此外, 当前流行的 CNN 架构的训练策略是微调 (fine-tuning) 预训练的网络。这种方法无需庞大的标注训练数据便可直接应用于端到端的医学图像分类任务。目前, 基于 CNN 的分类网络模型已经应用到新冠肺炎的识别中 [3, 6]。

现有的分类方法主要基于病人的胸部 X 光成像或计算机断层扫描 (CT), 数量较少, 且对负样本的选择各有差异; 数据集本身图像质量不同, 不利于开发鲁棒的识别算法。很多研究者采用了数据增广的办法弥补样本不足, 但难以避免过拟合的现象, 导致模型的泛化能力下降, 难以实际应用。因此, 我们的工作就着力于研究一种鲁棒的基于分类任务来定位病灶区域的方法, 以此达到快速、低成本新冠肺炎诊断。

目前新冠肺炎的医学图像集有 CT 和胸部 X 光两类, 考虑到我们的目标是构建低成本的自动诊断系统, 我们以胸部 X 光作为数据集。我们首先使用了图像分割技术提取出胸部 X 光图像中的肺部; 之后, 我们在处理后的数据集中训练了不同 CNN 分类网络, 着重考虑提高模型的精准率 (precision) 和召回率 (recall)。实验证明, DenseNet 模型具有最高的性能, 达到了 88.75% 的准确率和 88.9% 的 F1 得分。

本次课程项目贡献如下:

- 将任务定义为新冠肺炎和非新冠肺炎患者的二分类任务。通过图像分割技术处理原始医学图像，再训练多个 CNN 分类网络 (Resnet18, MobileNet v2, DenseNet121)，比较各个模型的准确率和 F1 得分，选出最优模型。
- 在病灶区域的映射上，我们分别将训练得到的最优模型 (DenseNet121) 通过 Grad-CAM 和 Grad-CAM++，直接显示了新冠肺炎患者病变区域。

## 2 相关工作

### 图像分类

在有监督数据的条件下，新冠肺炎的检测可以被规约为图像分类问题。图像分类即从给定的类别集合中给图像分配标签的任务。2012 年，AlexNet 在 ImageNet 上大放异彩，拉开了 CNN 成为图像分类任务最主流方法的序幕。随后涌现出 VGG-Net，GoogleNet，ResNet 等等，都是更深的 CNN 网络，分类性能更好。另一方面，这些网络模型复杂，需要消耗更大的存储和计算开销，严重限制了在低功耗领域的应用。轻量级分类网络如 ShuffleNet 和 MobileNet 应运而生，虽然网络结构略简单，但实现了算力资源约束下的性能突破。

### Covid-19 分类

目前已有的新冠肺炎分类工作主要有两大类。一类是直接医学图像数据集中训练分类网络，比如 [6] 中在病人的 CT 图像上微调了 ResNet50 网络进行分类，[4] 使用 Inception 模型训练新冠肺炎分类网络。另一类则是首先进行图像的处理，如图像分割、数据增强等，继而在处理后的图像中训练分类网络，比如 [2] 中使用条件增强和生成对抗网络的方式扩大训练样本集，之后再运行 CNN 分类模型。

此外，各个分类网络中对负样本的选择不同。在 [6] 中，选择了新冠肺炎患者、健康人群和细菌性肺炎三类群体的胸部 CT 构成样本集合。在 [1] 中，样本集合由正常人群、新冠肺炎患者和非典患者三类人群的肺部医学图像构成。

### 医学图像分割

医学图像分割是将医学图像中具有某些特殊含义的部分分割出来，并提取相关特征，为临床诊疗和病理学研究提供可靠的依据。该领域的开山之作是 2015 年 Long Jonathan 等人提出的全卷积网络 (FullyConvolutional Networks, 简称 FCN) 结构，对图像进行像素级的分类，从而解决了语义级别的图像分割问题。在此基础上提出的 U-Net 模型效果更好，目前已广泛应用于医学图像的分割。

## 3 问题定义

由于负样本的选择各异，肺部疾病种类多，做多分类任务难以涵盖完全、对各类疾病的特征学习能力可能不足，模型效果差。因此，在本次课程项目中，我们所完成的是新冠肺炎患者和健康人群的胸部 X 射线图像的二分类任务。

我们约定在有监督的条件下处理新冠肺炎分类问题。监督数据集由  $\{(x_i, y_i)\}_{i=1}^N$  给定，其中  $(x_i, y_i)$  为一组配对数据， $x_i$  为胸部 X 射线图像， $y_i \in \{1, 0\}$  (0 代表正常人群 normal, 1 代表新冠肺炎患者)。我们的目标就是从监督数据集  $D$  中学习一个判别函数  $f_\theta: x \rightarrow y$ ，使得对任意一张胸部 X 射线图像，能够正确区分出新冠肺炎患者和正常人。

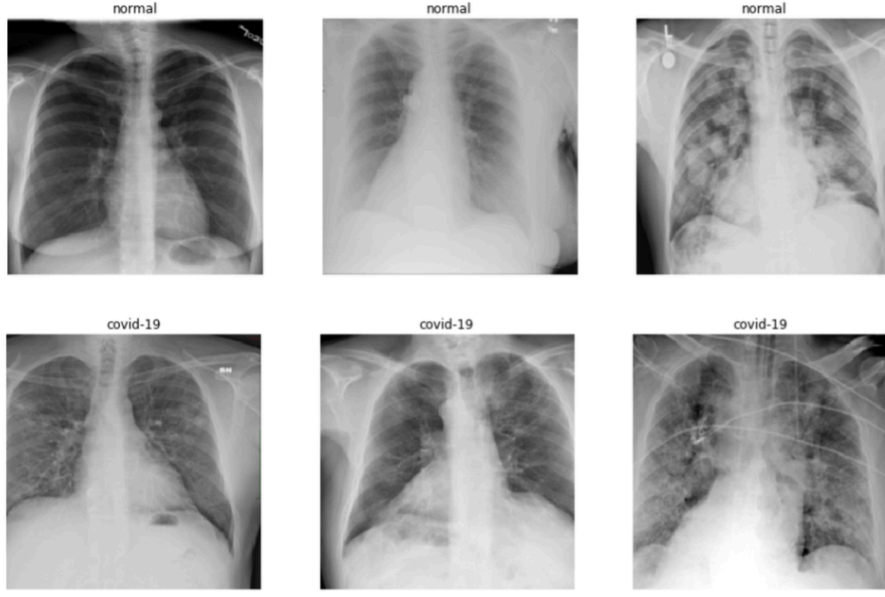


图 1: 样本数据集

## 4 算法

在本节中，我们将首先介绍分类系统的结构，然后介绍整个系统的训练方法。我们整个模型分为两个阶段 (2-stage)，如图 2 所示。

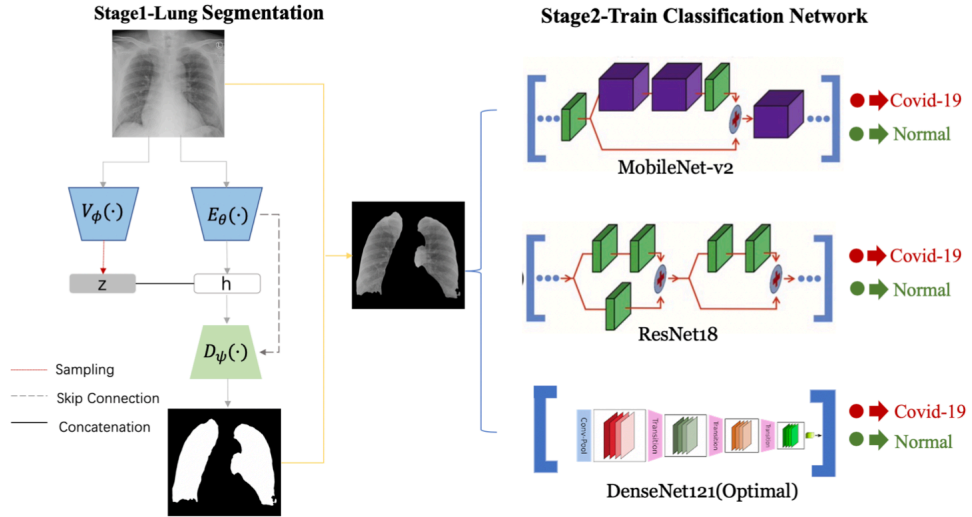


图 2: 算法整体架构

### 4.1 模型细节

我们的整个模型由两个阶段构成。首先进行肺部图像的分割。接着训练基于 CNN 的分类网络完成分类任务。对于图像分割网络，我们使用了 Raghavendra 等人提出的分割模型 lungVAE [5]。对于分类模型，我们分别实验了近几年在图像分类领域具有里程碑意义的三个网络，ResNet18，DenseNet121 以及 MobileNet-v2。

### 图像分割

有监督的图像分割问题，即给定输入图片  $x \in X$  和相应的分割  $s \in S$ ，通过大量的数据集  $\{(x_i, s_i)\}_{i=1}^N$ ，学习到映射函数  $f(\cdot) : X \rightarrow S$ 。在这一领域具有里程碑式的工作是 2015 年由 Ronneberger 等人提出的 U-net 模型。在该模型中，映射函数由编码器和解码器组成，因此  $f(\cdot) = E_\theta(D_\phi(\cdot))$ ，这里  $E_\theta$  和  $D_\phi$  分别对应参数为  $\theta$  和  $\phi$  的编码器和解码器。

在处理该课程项目的图像分割问题上，我们使用了 lungVAE 模型，该模型基于变分自动编码器 (Variational Autoencoder)。由于变分自动编码器能够获得更为丰富的隐含表示，因此网络的学习能力强，效果好。

该模型也是典型的编码器-解码器 (encoder-decoder) 模型，其中蓝色的部分是编码器，绿色是解码器。编码阶段，分别学习到两个部分均值  $E_\theta$  和偏差  $V_\phi$ ，其中学习均值的部分依然采用了 U-net 的网络模型。模型的特殊之处是将均值和偏差进行特征拼接，再之后通过 U-net 解码模型学习解码器  $D_\phi$ 。lungVAE 模型由大量胸部 X 射线图像训练得到，戴森系数达到 84.0%，效果优越。

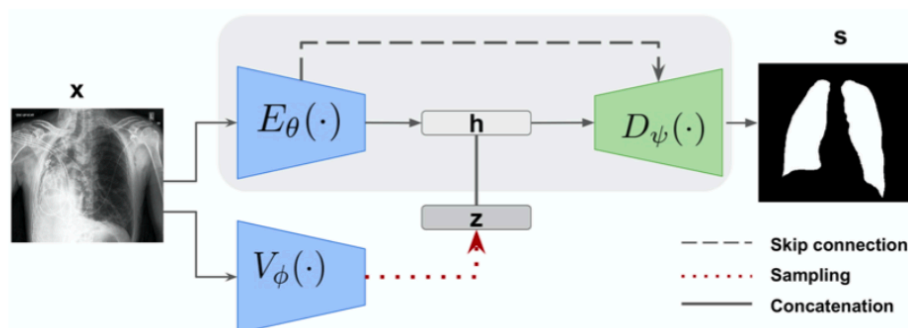


图 3: 图像分割模型

## 分类网络

当我们完成了图像分割，将肺部区域提取出来之后，就可以训练分类网络来完成分类任务。在这里，我们分别实验了 CNN 图像分类史上具有里程碑式的网络 ResNet, DenseNet 和 MobileNet。

### • ResNet

ResNet 由何恺明等人于 2015 年提出，是当前应用广泛的 CNN 特征提取网络。在此之前，CNN 的发展一直朝着学习更深的网络模型方向前进，从 AlexNet 的 7 层发展到了 VGG 的 16 乃至 19 层，后来更有了 GoogLeNet 的 22 层。因为网络越深，能够提取到不同的特征越丰富，越具有语义信息。但是，只简单地增加网络深度，会导致网络收敛性下降和网络退化等弊端。受制于此，ResNet 网络的作者想到了常规计算机视觉领域常用的 Residual Representation 的概念，并进一步将它应用在 CNN 模型的构建中。它通过建立前面层与后层间的短路连接，有助于训练过程中的梯度反向传播，从而训练更深更好的 CNN 网络。

### • DenseNet

相比于 ResNet，DenseNet 提出了更激进的密集连接机制：即互相连接所有的层。具体来说，每个层都会接受前面所有层作为其额外的输入。通过这种更密集的连接方式可以实现特重用和效率提升。下图是 ResNet 和 DenseNet 基本块的比较，可以直观地感受到 DenseNet 相较于 ResNet 在连接上更为密集紧凑。

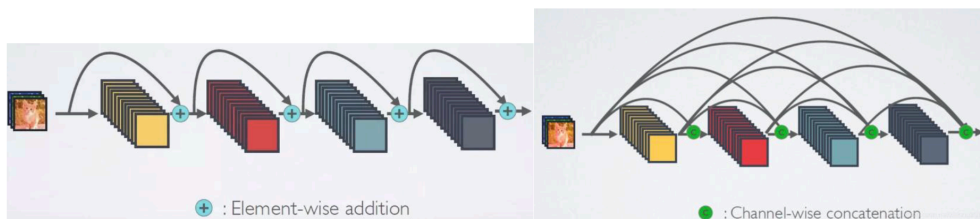


图 4: ResBlock 与 DenseBlock 结构对比

- MobileNet

在实际应用中 CNN 受限于硬件运算能力与存储开销, 因此必须有一种能在算法层面的有效压缩存储和计算量的方法。MobileNet 正在为我们打开这扇窗。

MobileNet v1 是 Google 于 2017 年发布的网络架构, 旨在充分利用移动设备和嵌入式应用的有限资源, 有效地最大化模型的准确性。它的核心是把卷积拆分成 Depthwise 和 Pointwise 两部分。Depthwise 将输入分组, 每组做  $3 \times 3$  的卷积, 相当于收集每个通道的空间特征, 但是通道间信息是相互隔离的。Pointwise 对输入做  $1 \times 1$  的卷积, 收集了每个点的特征, 实现了对通道间信息的交换。因此这两者结合即把普通的卷积分为两个步骤来做, 但是模型的参数和计算量显著下降。MobileNet v2 是 Google 继 v1 之后提出的下一代轻量化网络, 解决了 v1 在训练过程中的特征退化问题, 效果进一步提升。

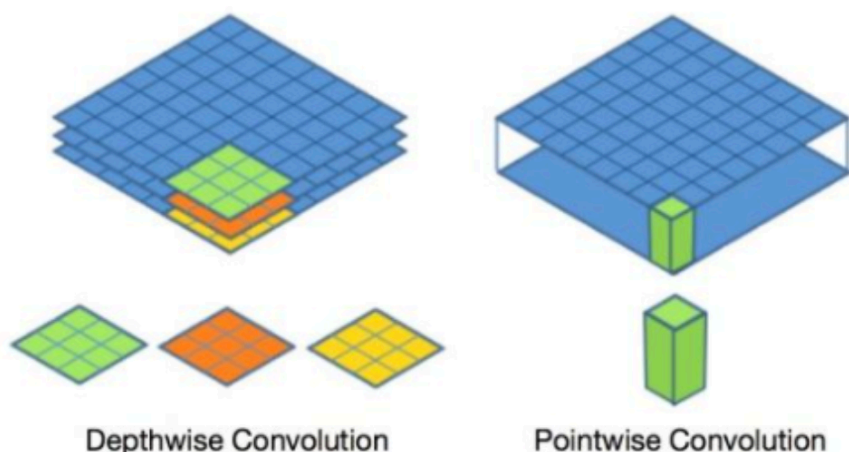


图 5: MobileNet 的操作:DepthWise,Pointwise

## 4.2 训练

我们的模型分为两个阶段。在第一阶段, 我们直接调用了预训练好的 lungVAE<sup>1</sup>, 该模型在肺部图像分割中的测试集上达到了 84.0% 的戴森系数和 82.4% 的准确率, 性能良好。

对于分类任务, 我们使用交叉熵 (cross-entropy) 作为损失函数。公式为:  $L = -\sum_{i=1}^N (y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i))$ . 其中  $y_i$  为图像标注类别 (ground truth),  $\hat{y}_i$  为图像经过分类网络得到的计算结果。

## 5 实验比较

### 5.1 数据集

<sup>1</sup><https://github.com/raghavian/lungVAE>

我们的样本数据集有新冠肺炎患者的胸部 X 射线图像<sup>2</sup> 和正常人的 X 射线图像<sup>3</sup> 共 392 张。其中正样本新冠肺炎患者的有 196 张，负样本 196 张，正负样本均衡。我们按照 7:2:1 的比例划分为训练集 (trainset)、测试集 (testset) 和验证集 (validationset)。

## 5.2 实验

### 实验步骤

我们首先调用预训练的肺部分割网络对原始图像进行处理，得到 mask。之后将 mask 与原始图像进行与运算，得到分割后的图像。进而在分割后的图像上进行分类网络的训练，包括 Resnet18, Densenet121, Mobilenetv2，训练的损失函数已在 4.2 部分说明。

### 实验结果

在实验参数的设定上，我们将 Mask 后的图像进行归一化，之后设定学习率为 0.002，采用 Adam 优化器进行权重更新。在三个网络上训练分类模型过程中的准确率如图 7 所示。可以看到 DenseNet 模型收敛较快，在验证集和测试集上均表现良好。MobileNet 效果不稳定，当 Epoch 增加，会出现严重的过拟合现象，模型性能大大下降。



图 6: 实验样本 (从左至右依次为原图，分割模型得到的 mask 和原图与 mask 作用的最终图)

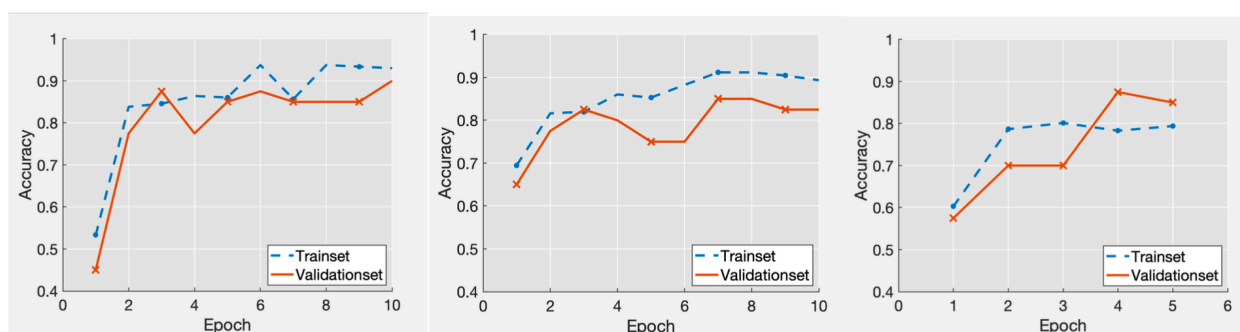


图 7: 在测试集和验证集准确率 (从左到右依次为 DenseNet, ResNet, MobileNet)

为验证分割的效果更优，我们设计了三个 Baseline。即分别在原图像进行 DenseNet、ResNet 和 MobileNet 的训练，我们比较了这三个模型作用于原图和经过肺部分割后图像在测试集上的准确率、召回率、精度以及 F1 得分。

<sup>2</sup><https://github.com/ieee8023/covid-chestxray-dataset>

<sup>3</sup>[http://openaccess.thecvf.com/content\\_cvpr2017](http://openaccess.thecvf.com/content_cvpr2017)



表 1: 三个分类网络作用于原图和经过分割处理后图像的性能比较

Model	Accuracy	Precision	Recall	F1-score
DenseNet	87.50%	87.50%	87.50%	87.50%
MobileNet	86.25%	93.94%	77.50%	84.93%
ResNet	83.75%	89.66%	65.00%	75.36%
Segmentation+DenseNet	88.75%	87.80%	90.00%	88.89%
Segmentation+MobileNet	86.25%	87.18%	85.00%	86.08%
Segmentation+ResNet	87.50%	89.47%	85.00%	87.18%

基于实验结果可以得出以下结论：

- 无论图像是否经过分割，对于这个二分类任务，DenseNet 模型的效果在三个网络中表现最好，在准确率和 F1 得分上表现最佳。作用于经过分割后的图像的 DenseNet 可以获得 88.75% 的准确率和 88.89% 的 F1-得分，性能优越。
- 在我们实验的三个网络上，作用于经过分割的图像的模型效果都好于直接作用于原图像。比较直接的反映是准确率、召回率以及 F1-得分。在召回率的提高上尤其显著，比如 ResNet 模型在未经分割的 X-ray 图像上仅有 65% 的召回率，但是作用在分割后的图像，达到了 85% 的召回率，提升效果明显。对于新冠肺炎的分类任务，提高模型的召回率意义重大。低召回率意味着许多新冠病人被误判为健康人，这一点是十分危险的。因此，一个好的分类模型必须在召回率上表现良好。可以看到 DenseNet 作用于处理后的图像，召回率高达 90%，这一结果证明了我们先分割处理后分类的 2-Stage 算法的合理性。

将三个分类网络直接作用于原图像和经过分割处理的图像在测试集上的 F1-得分，准确率，召回率比较如下图。可以更直观地感受到图像分割对模型性能提升的显著作用，也由此感受到我们算法的合理性。

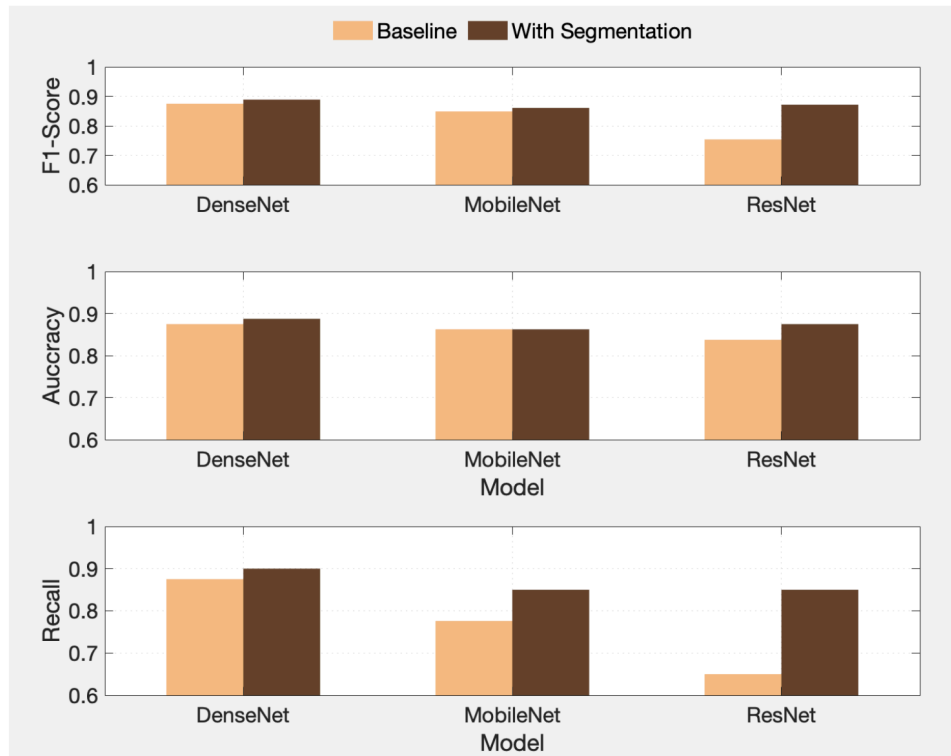


图 8: 比较三个分类网络作用于原图像和经过分割处理的图像的性能

## 案例分析

我们分别使用 Grad-CAM 和 Grad-CAM++ 模型作用于图像，进行病灶的定位。

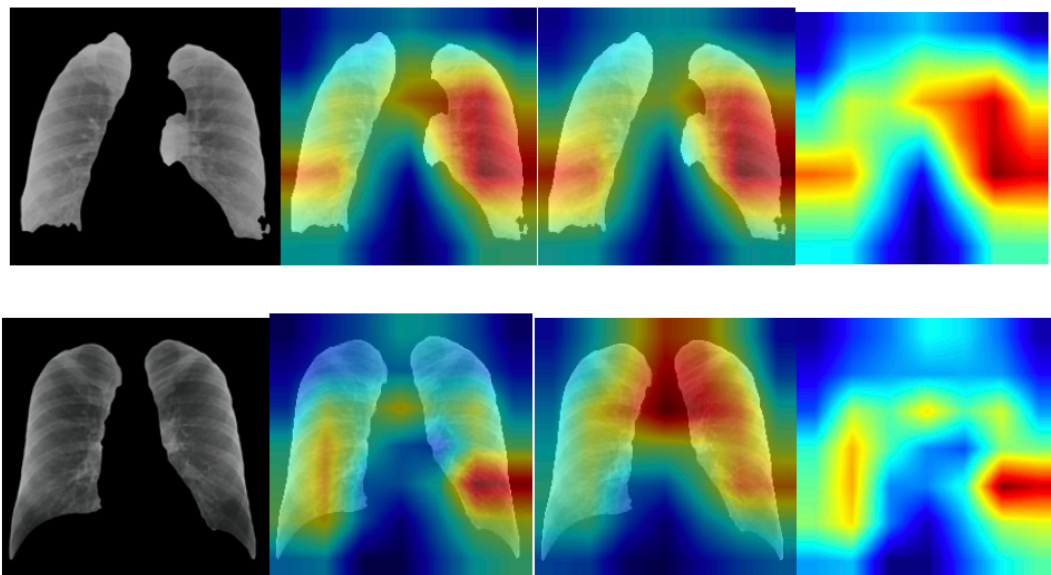


图 9: Covid-19 患者胸部 X 射线图、CAM 图、CAM++ 图、热力图

我们所使用的网络是训练得到的最优网络 DenseNet121，将它作用于患者胸部 X 射线图，使用 Grad-CAM（以及 Grad-CAM++）方法<sup>4</sup>。从图 11 中可以看到，我们的模型能够进行病灶区域的预测和判断。

## 6 总结与展望

本次课程项目，我们立足于新冠肺炎的分类，自主选择构建了合理的样本集，提出了一个 2-stage 算法，包括先进行肺部的图像分割，之后再经过 CNN 进行分类。通过在未经分割的原图像训练分类网络的比较试验，证明了分割模型的有效性。另外，我们比较了 DenseNet，ResNet 和 MobileNet 这三类网络在本任务上的性能，确定了 DenseNet 效果最优，达到了 88.75% 的准确率和 88.9% 的 F1-得分。我们的算法较直接在原图像训练分类网络，在召回率和 F1 得分上提升显著，证明了我们算法的有效性和可靠性。最后，我们随机选择了样本，在其上运行了 Grad-CAM 算法进行病灶区域的定位来直观展示我们模型的结果。

当然，本次处理也有不足。比如数据集规模略小，不利于网络学习更全面广泛的特征。对于分类网络，仅考虑了比较著名的三个网络，对其他的 CNN 分类网络缺乏更多的实验，或者加上 Attention 机制构建鲁棒性更高的分类网络；也缺少选择不同的分割模型的效果的对比实验，而仅使用了一个预训练效果好的肺部分割模型。

## 参考文献

- [1] Asmaa Abbas, Mohammed M. Abdelsamea, and Mohamed Medhat Gaber. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. 2020.
- [2] M. Loey and N.E. Smarandache, F.and M. Khalifa. A deep transfer learning model with classical data augmentation and cgan to detect covid-19 from chest ct radiography digital images. 2020.
- [3] Sethy PK and Behera SK. Detection of coronavirus disease (covid-19) based on deep features. 2020. 2020.

<sup>4</sup><https://github.com/yizt/Grad-CAM.pytorch>



- [4] Wang S, Kang B, Ma J, Zeng X, Xiao M, and Guo J. A deep learning algorithm using ct images to screen for corona virus disease (covid-19). 2020.
- [5] Raghavendra Selvan, Erik B. Dam, Sofus Rischel, Kaining Sheng, Mads Nielsen, and Akshay Pai. Lung segmentation from chest x-rays using variational data imputation. 2020.
- [6] Song Y, Zheng S, Li L, and Zhang X. Deep learning enables accurate diagnosis of novel coronavirus(covid-19) with ct images. 2020.