1. **数据集介绍**
2. **COVID-19 CT Lung and Infection Segmentation Dataset** **（**<https://zenodo.org/record/3757476>**）**

该数据集是目前最大的公开的具有像素级标签的Covid-19数据集，适用于图像分割任务。其中包含了20个确诊患者的CT数据和对应的mask标签，对于准确分割出Covid-19的病灶区域很有帮助。

1. **COVID-Chest-Xray-Dataset**

（<https://github.com/ieee8023/covid-chestxray-dataset>）

该数据集是目前最大的公开的X-Ray数据集，其中包含了Covid-19、SARS等肺部疾病的X-Ray图像和对应的图像级类别标签。有助于在基于X-Ray图像的肺部疾病诊断中识别出Covid-19图像。

1. **代表性文章／算法**
2. **Towards Efficient COVID-19 CT Annotation: A Benchmark for Lung and Infection Segmentation** (<https://arxiv.org/pdf/2004.12537.pdf>)

这篇文章提出了一个关于Covid-19的Benchmark，包含3个task：i.) 小样本分割任务，ii.) 从非Covid-19数据中学习如何分割Covid-19图像，iii.) 从Covid-19数据和非Covid-19数据中进行多病灶的分割（如Covid-19和整个肺叶）。文章利用已有框架NN-Unet给出了三个任务的Baseline。

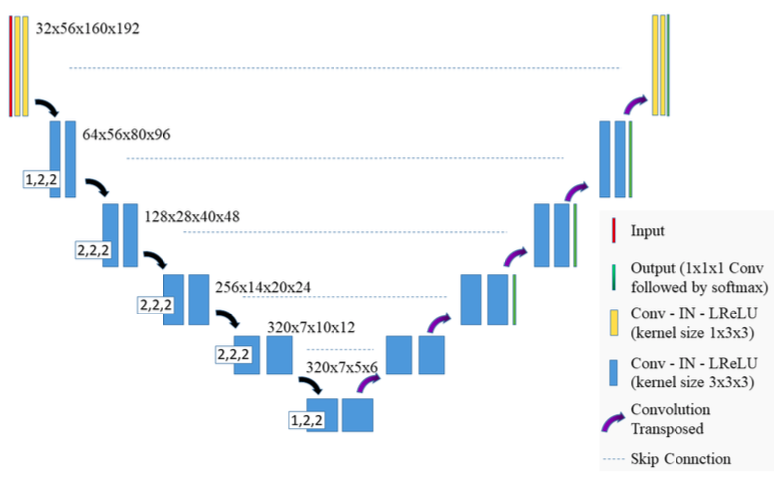
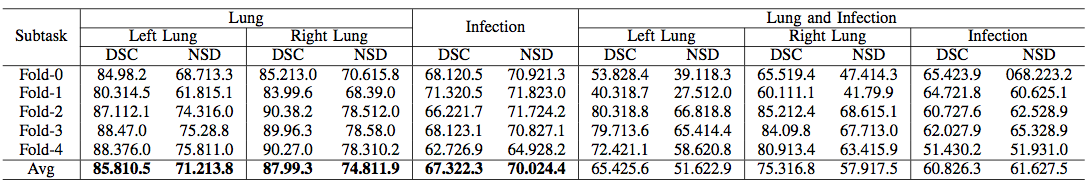


图 1 nnUnet方法

表1 关于左右肺叶分割和Covid-19病灶分割的5折交叉验证结果



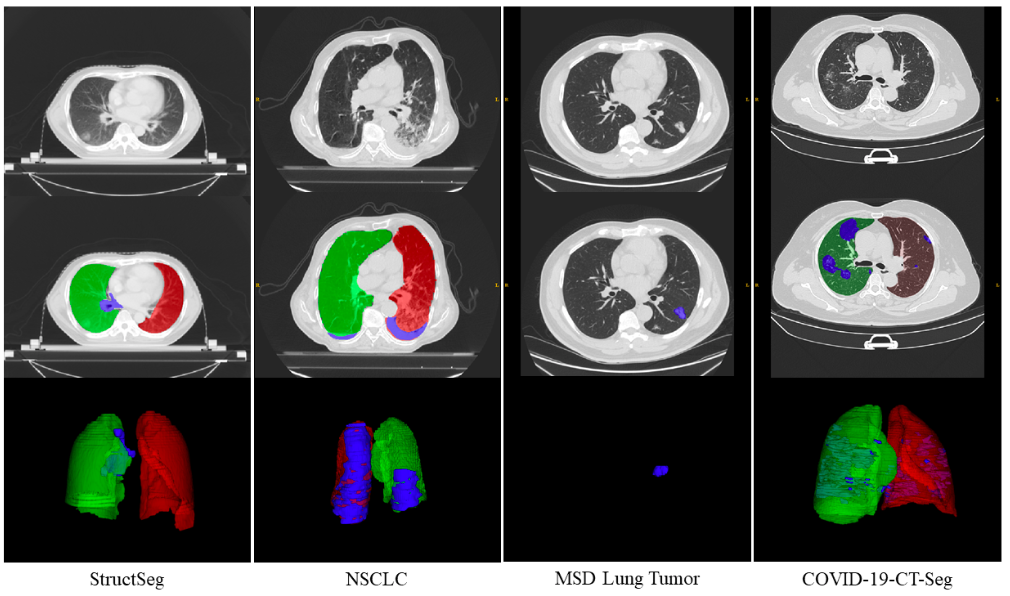


图 2 现有数据集以及提出的数据集COVID-19-CT-Seg中图像和Ground-Truth

**(2) Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network** （<https://www.medrxiv.org/content/10.1101/2020.03.30.20047456v1.full.pdf>）

这篇文章在chest X-Ray数据做分类任务，提出DeTraC模型，利用基于欧氏距离的标签分解方法，将数据集中包含的Covid-19和SARS标签分别分解为更细的子标签，最后再反向集成，得到真正的分类标签。用于计算欧式距离的特征通过深度神经网络(ResNet-18)得到，模型框架如下：

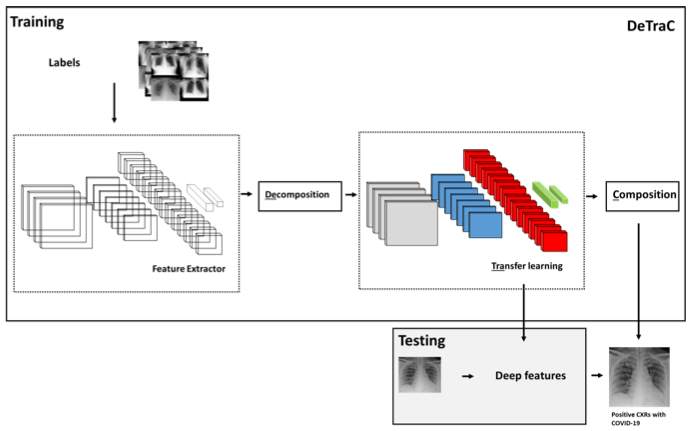
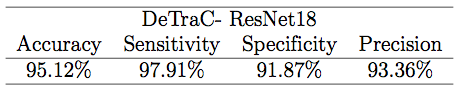


图 3 DeTraC框架

表2 X-Ray数据集分类结果



分类结果如表2所示，虽然的到了较好的结果，但是其训练样本存在严重的类别不平衡问题，有很大的改进空间。

**(3) A Deep Transfer Learning Model with Classical Data Augmentation and CGAN to Detect COVID-19 from Chest CT Radiography Digital Images.** (<https://www.preprints.org/manuscript/202004.0252/v1>)

这篇文章从多种途径收集了目前互联网上的Covid-19数据，利用数据增强和条件生成对抗网络(CGAN)的方式扩大训练样本集。其算法框架(DTL)如图4所示，利用多种神经网络的测试结果如表3所示。

表3 DTL测试结果

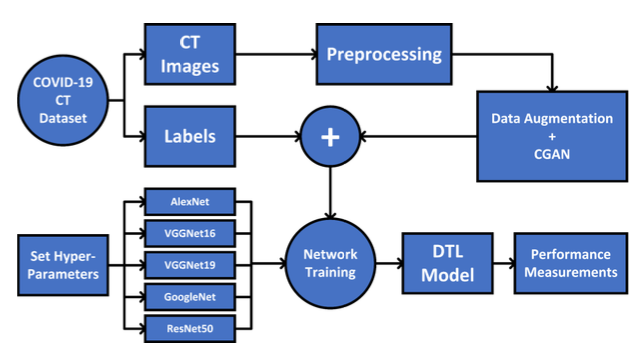
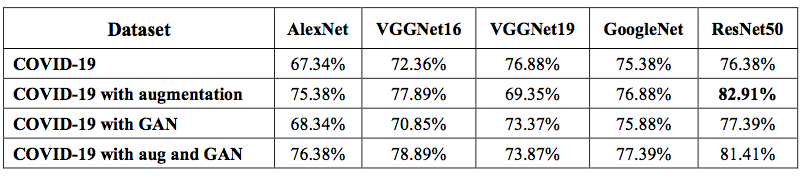


图4 DTL框架

1. **拟研究方向／问题**

* **现有方法使用的基于X-Ray的数据集中包含的正样本（Covid-19样本）数量较少，负样本的选择各异，且数据集本身的图像质量不同，不利于开发一个鲁棒的Covid-19识别算法，即分类任务难以定位到病灶区域。很多研究者采用了数据增强的方法，但是难以避免过拟合的现象。因此，一种鲁棒的基于分类任务来定位病灶区域的方法将有助于快速、低成本（CT的成本低于chest X-Ray）的Covid-19诊断。**
* **利用肺部CT图像的基于像素级标签的图像分割方法能够更准确地定位到Covid-19相关的病灶区域。但是目前公开的该类型数据集比较少且其中包含的样本也较少，直接采用监督学习的方式很容易过拟合，模型泛化能力较差。因此，在该类数据上通过小样本学习（few-shot learning）和元学习（meta-learning）的方法来提高模型的泛化能力更有效。**
* **目前关于Covid-19的算法模型大都是现有的方法和模型（如UNet框架），更准确的分类／分割效果依赖于对该任务有针对性的方法的研究。**