

# 生物大数据分析大作业报告

组员(按拼音首字母): 鲁周语, 佟昊炜, 朱俊超, 朱骏杰

## 生物大数据分析大作业报告

1. 设计背景
2. 模型框架
3. 实验设置
4. 实验结果与讨论
5. 不足之处
6. 总结
7. 致谢
8. 任务分工 (待加)

## 1. 设计背景

胸腔X光 (Chest X-ray) 是一种常见的医学影像检查方法, 用于评估和诊断胸部疾病。它是一种快速、非侵入性且相对低成本的检查技术, 广泛应用于临床实践中。在胸腔X光检查中, 患者站立或坐在X光机前, 将胸部暴露在X射线辐射下。然后, X射线机会通过患者的胸部, 生成一个二维的黑白影像, 显示了内部肺组织、肺血管和其他结构的信息。通过胸腔X光图像, 医生可以观察和评估以下方面:

1. 肺部解剖结构: 胸腔X光图像提供了肺部解剖结构的概貌, 包括肺野、肺叶、肺段和支气管。这有助于医生判断是否存在异常结构或变异。
2. 肺组织密度: 正常肺组织在X光图像上呈现为黑色, 表示透明度高, 透射X射线较多。异常肺组织, 如病变、浸润或肿块, 可能表现为白色或灰色阴影, 表示密度增加。
3. 胸腔膜和胸廓: 通过胸腔X光图像, 可以评估胸腔膜的状态以及胸廓的形状和对称性。异常的胸腔膜可能表现为积液、肿块或胸腔积气。
4. 心脏和大血管: 胸腔X光图像还提供了对心脏和大血管的初步评估。心脏大小、形状和位置的变化以及大血管的扩张或狭窄等异常都可以在图像中观察到。

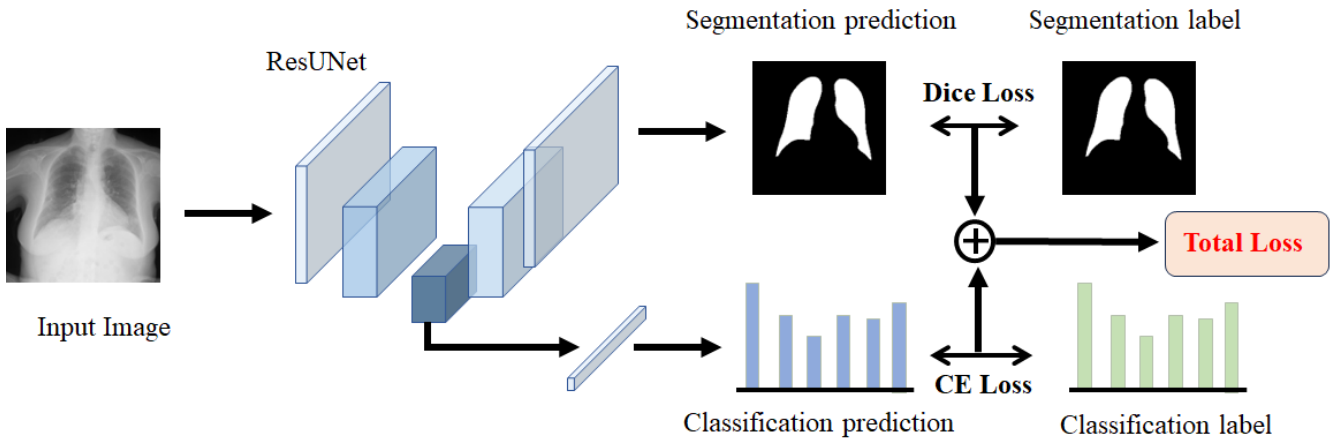
胸腔X光图像在临床上广泛应用于肺部疾病的筛查、诊断和监测。然而, 由于其二维特性和较低的分辨率, 胸腔X光图像的解释存在一定的主观性和限制性。因此, 借助深度学习方法, 我们可以探索如何利用计算机算法和人工智能来自动分析和解读胸腔X光图像, 以提高肺部疾病的诊断准确性和效率。

我们的目标是通过图像深度学习算法实现对患病情况的分类, 共分为四类: 正常 (Normal)、病毒性肺炎 (Viral Pneumonia)、新冠肺炎 (COVID) 和其他肺部感染 (Lung Opacity)。病毒性肺炎是由感染病毒引起的肺部炎症。常见的病毒包括流感病毒、腺病毒、呼吸道合胞病毒等。病毒性肺炎的X光图像通常显示为肺部浸润, 即肺组织的浑浊或模糊阴影。这种肺炎可能导致咳嗽、呼吸急促、胸痛等症状。新冠肺炎是由严重急性呼吸综合征冠状病毒2型 (SARS-CoV-2) 引起的呼吸系统疾病。COVID-19的病情可以从轻微的症状到严重的肺炎不等。肺部X光图像中, COVID-19患者常常表现为双侧肺部浸润性病变, 包括斑片状阴影、浸润和斑点状阴影。而细菌性肺炎、真菌感染、结核病等感染性疾病同样会引起肺部炎症和病变, 导致肺部X光图像中出现浸润、阴影和不规则的肺部纹理。

通过采用深度学习方法，我们可以利用大规模的标注胸腔X光图像数据集进行模型训练，并利用模型的自动特征学习和表示能力来区分不同的疾病类别。这种基于数据驱动的方法有望提高肺部疾病的诊断准确性和效率，并为临床医生提供可靠的辅助工具。由于本次任务给定的数据集为具有分类标签与分割标签的多标签肺部疾病数据集，考虑到单通道的CT图像在分类标签层级上的信息量较为有限，无法高效的实现肺部疾病分类任务。为充分利用已有的标注信息，考虑到不同肺部疾病类型与其肺部形态病变之间存在的天然相关性，设计了一个双分支的多任务模型，同时进行肺部疾病分类与分割任务。通过在分类任务中引入执行分割任务，模型可以对图像中的目标进行感知，并在生成过程中引入目标感知的约束。这可以提高生成模型对目标的理解和处理能力，使生成结果更加准确和语义合理，帮助模型更好的理解图像中的多样性。

## 2. 模型框架

本模型是一个多任务、U形网络，能够同时实现肺部疾病分类与分割任务。我们采用ResNet50作为编码网络用于提取图像的特征信息，通过一个分类预测头来预测当前图像的疾病分类结果。此外，我们采用UNet模型的解码器用于恢复与生成像素级的肺部分割结果。



)

Figure 1. 模型框架

对于分类任务，我们采用交叉熵损失来评估模型在分类过程中的性能，估计模型预测与真实标签之间的差异。交叉熵损失的计算公式如下，其中， $y$ 表示真实标签的概率分布， $\log$ 表示自然对数， $p$ 表示模型的预测概率分布。

$$CE = - \sum (y * \log(p)) \quad (1)$$

为帮助模型从像素级的分割标注中学习不同肺部疾病的独特属性，我们将分割任务是做一种正则化技术，帮助模型在多任务学习中的任务互相约束，让模型学习到更一般化的特征表示，提高分类任务的泛化能力，使模型对新样本的预测更准确。在实现的过程中，我们采用Dice loss来评估模型在分割过程中的性能，估计模型预测与真实标签之间的差异。Dice loss的计算公式如下，其中， $A$ 表示预测分割结果的二值化掩码（0表示背景，1表示目标）， $B$ 表示真实分割结果的二值化掩码。

$$DICE = 1 - (2 * |A \cap B|) / (|A| + |B|) \quad (2)$$

从而，模型的全部损失可以被表示为：

$$L_{total} = -\lambda_1 \sum (y * \log(p)) + \lambda_2 (1 - (2 * |A \cap B|) / (|A| + |B|)) \quad (3)$$

其中， $\lambda_1$ 和 $\lambda_2$ 为平衡两个损失的超参数，在实验设置中，为突出分类任务的重要性，设置， $\lambda_1 = 1$ 和 $\lambda_2 = 0.01$ 。

### 3. 实验设置

我们随机抽样的放大，采用8：2的比例设置训练集与测试集进行训练。本模型超参数设置包括：学习率为0.1，epoch=25，batchsize=4， $\lambda_1 = 1$ 和 $\lambda_2 = 0.01$ 。本模型用一块Nvidia RTX2060显卡训练8小时。

### 4. 实验结果与讨论

以下是对比模型部分的文字介绍：

1. 对比试验中选取的模型介绍： 在对比实验中，我们选择了三种模型进行比较：ResNet50、ConvNetXt和Conformer。
- ResNet50: ResNet50是一种经典的深度残差网络，具有50层的网络结构。它采用了残差连接（residual connections）的技术，通过跨层直接连接来解决梯度消失和网络退化的问题。ResNet50在图像分类任务中表现出色，并被广泛用于各种计算机视觉应用中。
  - ConvNetXt: ConvNetXt是一种基于卷积神经网络的模型，它采用了密集连接（dense connections）的思想。该模型在每个层之间建立密集的连接，使得每个层的输入都包含了前面所有层的信息，从而增强了特征传递和梯度流动。ConvNetXt在图像分类和分割任务中取得了良好的性能。
  - Conformer: Conformer是一种基于Transformer的序列建模网络，它在自然语言处理任务中取得了巨大成功。最近，Conformer也被应用于计算机视觉领域，用于图像分类和目标检测等任务。Conformer通过自注意力机制和多层感知机等模块来捕捉输入序列的长程依赖关系和特征表示，从而提高了模型的性能。
1. 模型性能展示： 我们对比了这三种模型在性能指标上的表现，包括AUC、准确率、精确率、召回率和F1值。根据实验结果，我们得到了每个模型在这些指标上的具体数值。

Model	Accuracy	Precision	Recall	F1 score	AUC
ConvNeXt	57.51%	0.6314	0.6158	0.6678	0.7876
ResNet50	64.22%	0.7011	0.7834	0.6580	0.8450
Conformer	60.32%	0.6805	0.7802	0.6245	0.8662
Our Model	64.73%	0.7074	0.7965	0.6631	0.8461

根据提供的对比试验数据，我们可以对不同模型的性能进行比较。下面是各个模型在不同评估指标上的表现：

ConvNeXt:

- 准确率（Accuracy）： 57.51%
- 宏平均精确率（Macro Precision）： 0.6314
- 宏平均召回率（Macro Recall）： 0.6158
- F1值： 0.6678
- AUC（Area Under the Curve）： 0.7876

ResNet50:

- 准确率（Accuracy）： 64.22%
- 宏平均精确率（Macro Precision）： 0.7011
- 宏平均召回率（Macro Recall）： 0.7834

- F1值: 0.6580
- AUC (Area Under the Curve) : 0.8450

Conformer:

- 准确率 (Accuracy) : 60.32%
- 宏平均精确率 (Macro Precision) : 0.6805
- 宏平均召回率 (Macro Recall) : 0.7802
- F1值: 0.6245
- AUC (Area Under the Curve) : 0.8662

我们的模型:

- 准确率 (Accuracy) : 64.73%
- 宏平均精确率 (Macro Precision) : 0.7074
- 宏平均召回率 (Macro Recall) : 0.7965
- F1值: 0.6631
- AUC (Area Under the Curve) : 0.8461

通过对比这些指标, 我们可以观察到不同模型在分类任务中的性能差异。从准确率来看, 我们的模型和ResNet50模型表现相似, 略优于ConvNeXt和Conformer模型。在宏平均精确率、宏平均召回率和F1值方面, 我们的模型和Conformer模型相比表现较好, 但仍低于ResNet50模型。在AUC指标上, 我们的模型和ResNet50模型处于相似水平, 高于ConvNeXt和Conformer模型。

这些结果表明我们的模型在COVID疾病分类任务中具有一定的优势, 相较于ConvNeXt和Conformer模型, 我们的模型在综合性能上表现更好。然而, 与ResNet50模型相比, 我们的模型在某些指标上仍有改进的空间。

## 2. ROC曲线和混淆矩阵

我们绘制了每个模型的ROC曲线, 并生成了相应的混淆矩阵。ROC曲线是一种用于评估多分类器性能的常用工具, 它展示了灵敏度和特异性之间的关系。混淆矩阵则提供了模型在不同类别上的分类结果, 包括真阳性、假阳性、真阴性和假阴性。

- ConvNeXt:

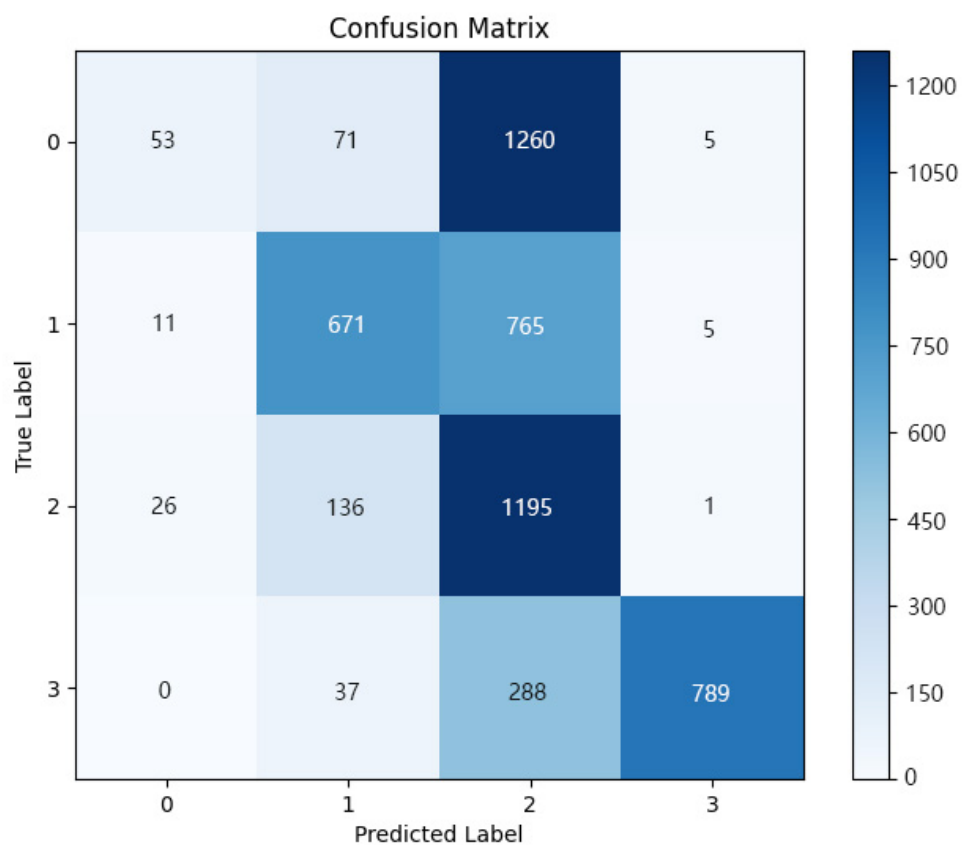
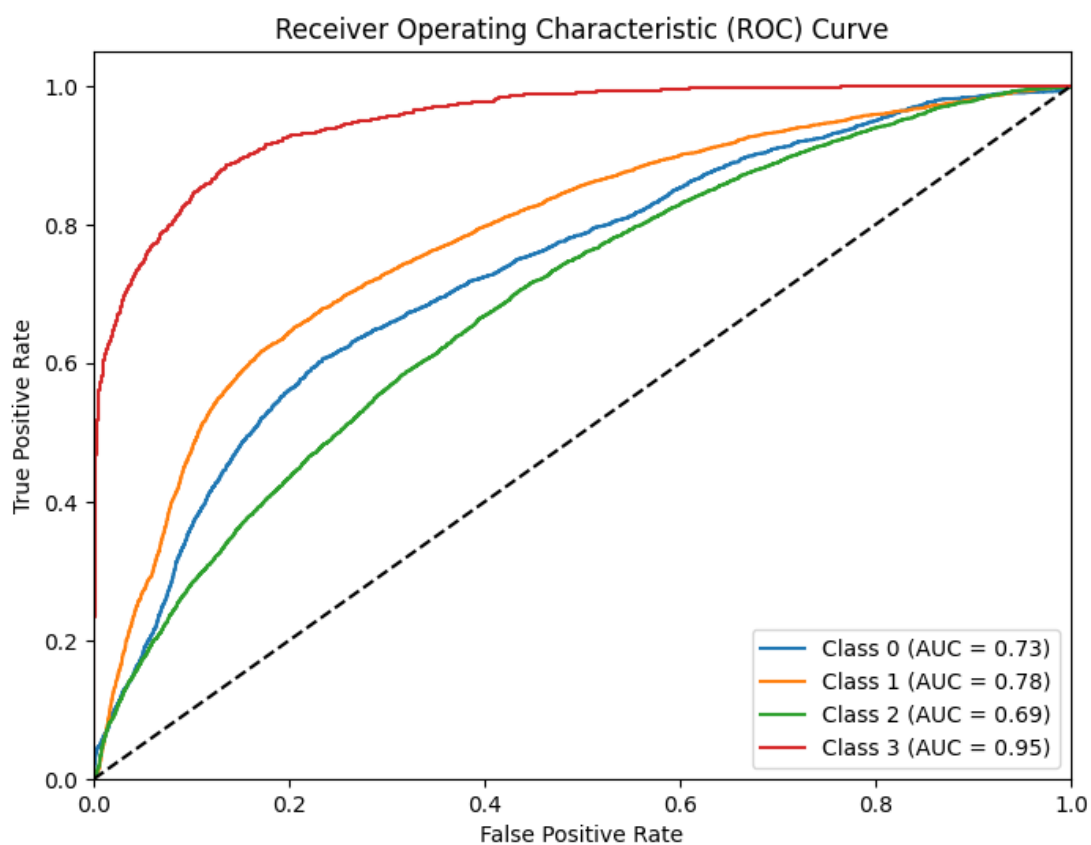


Figure 2. ConvNeXt的ROC与Confusion Matrix

- ResNet50

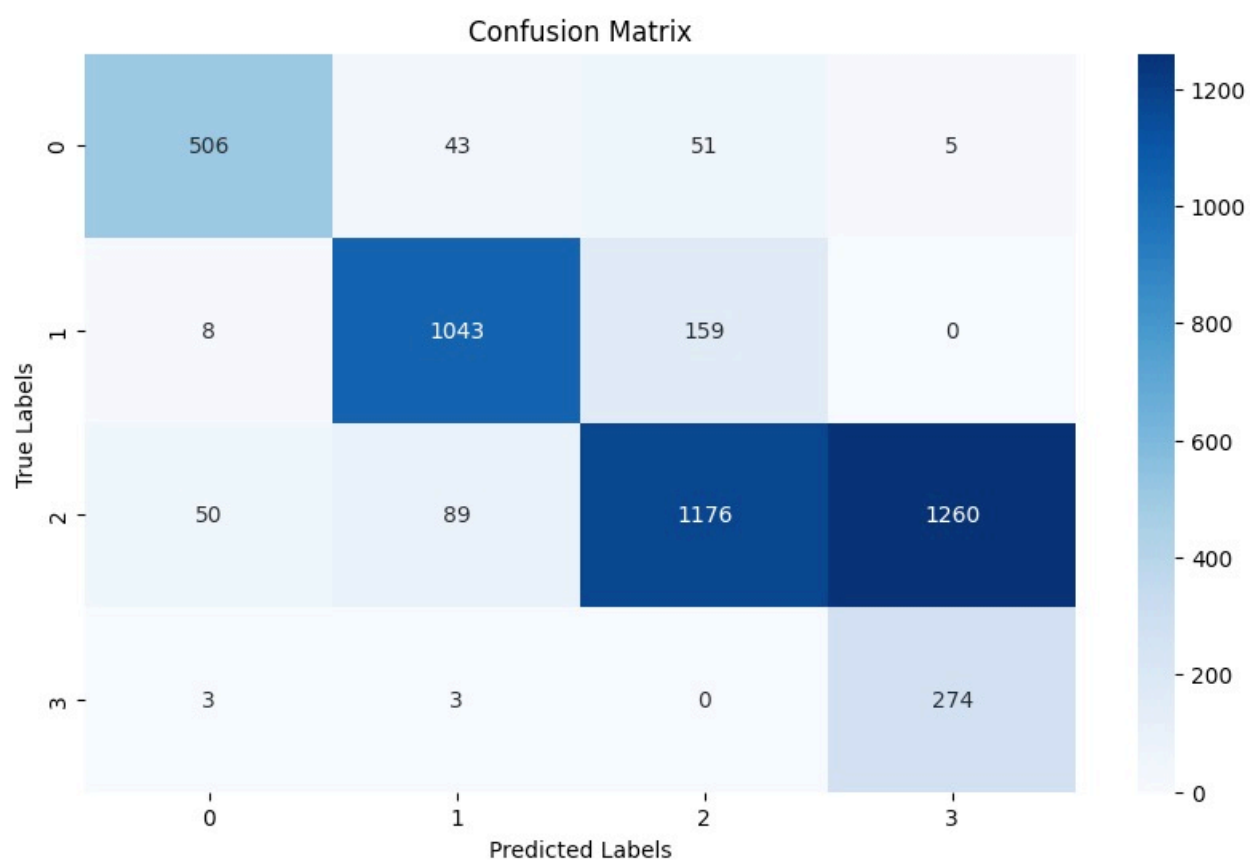
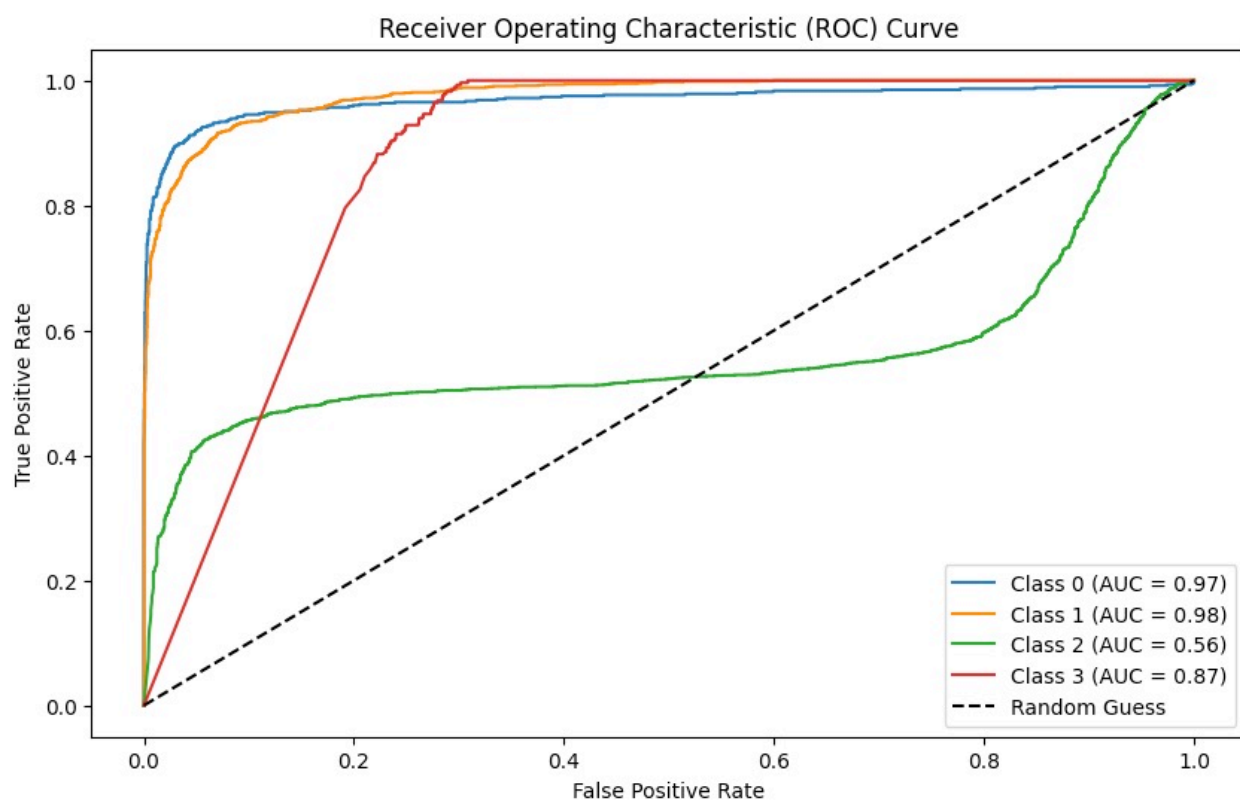


Figure 3. ResNet50的ROC与Confusion Matrix

- Conformer

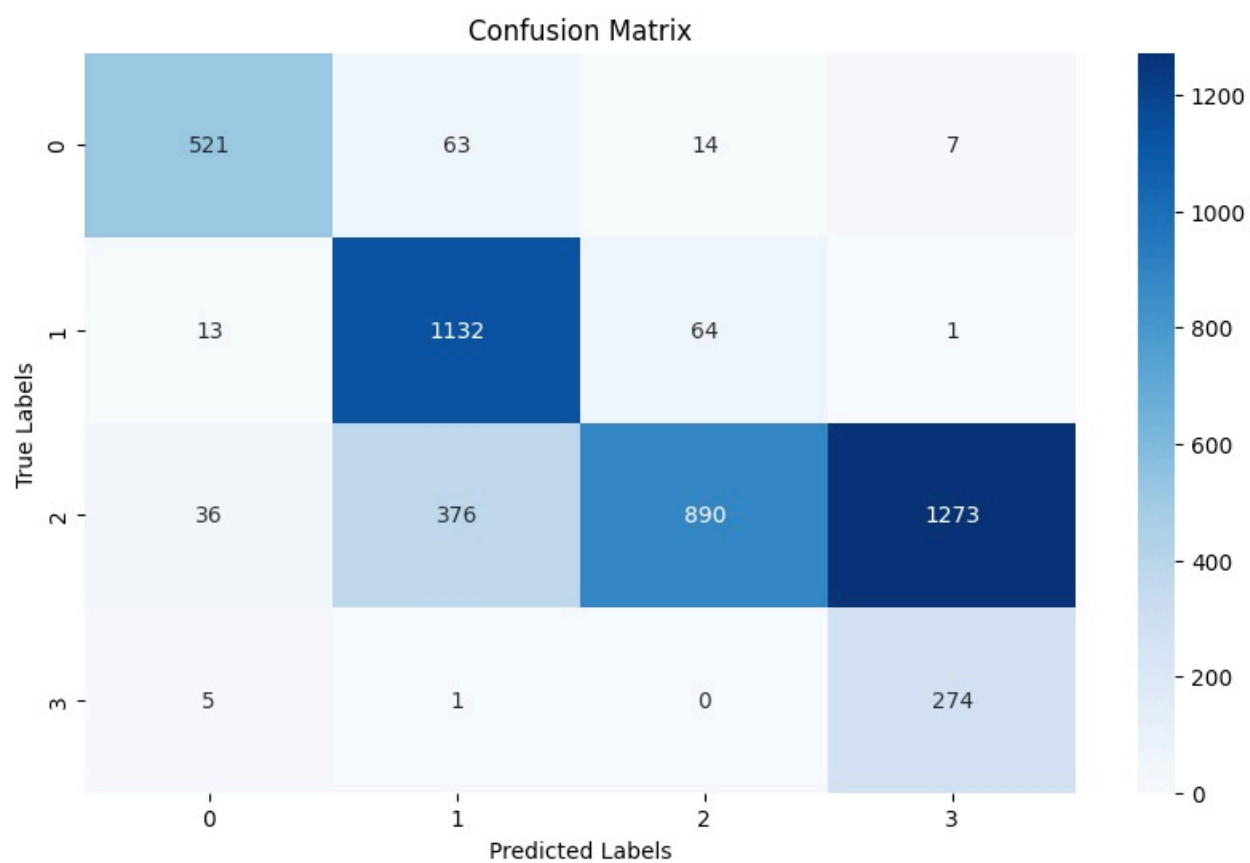
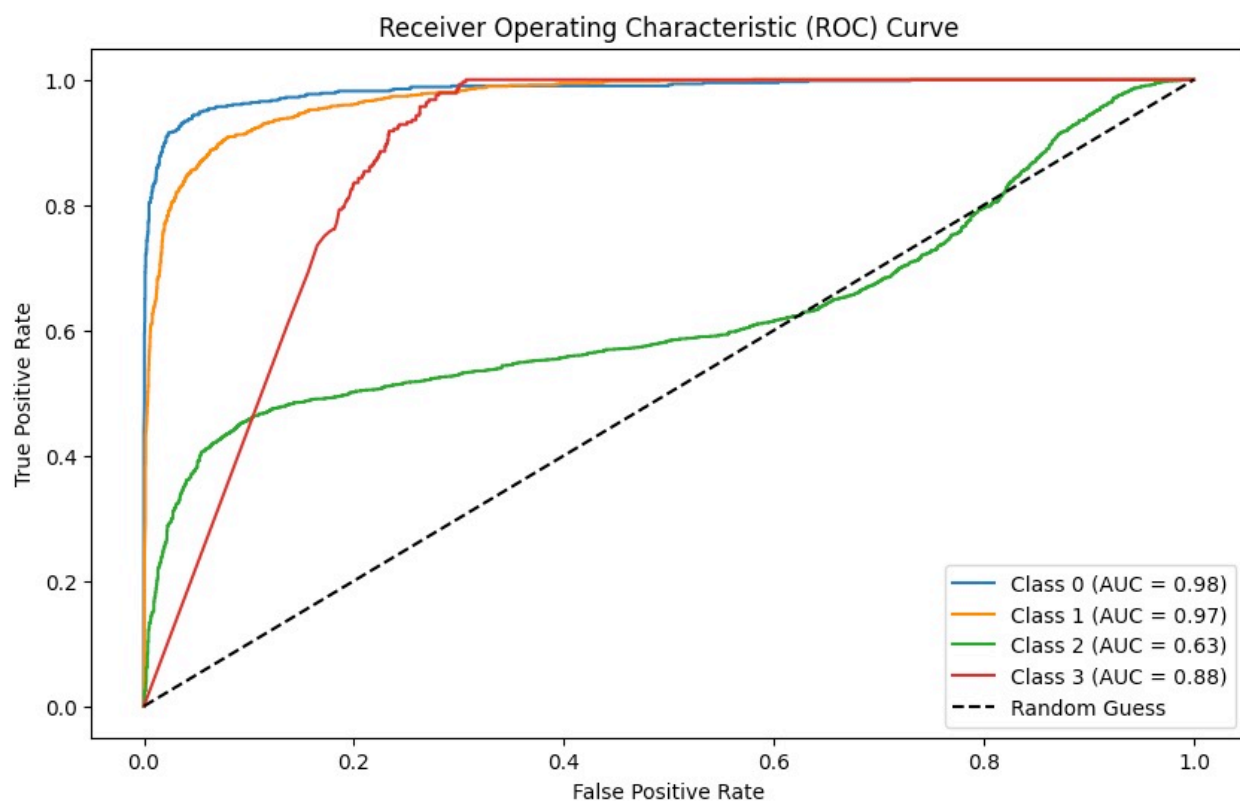


Figure 4. Conformer的ROC与Confusion Matrix

- Our Model

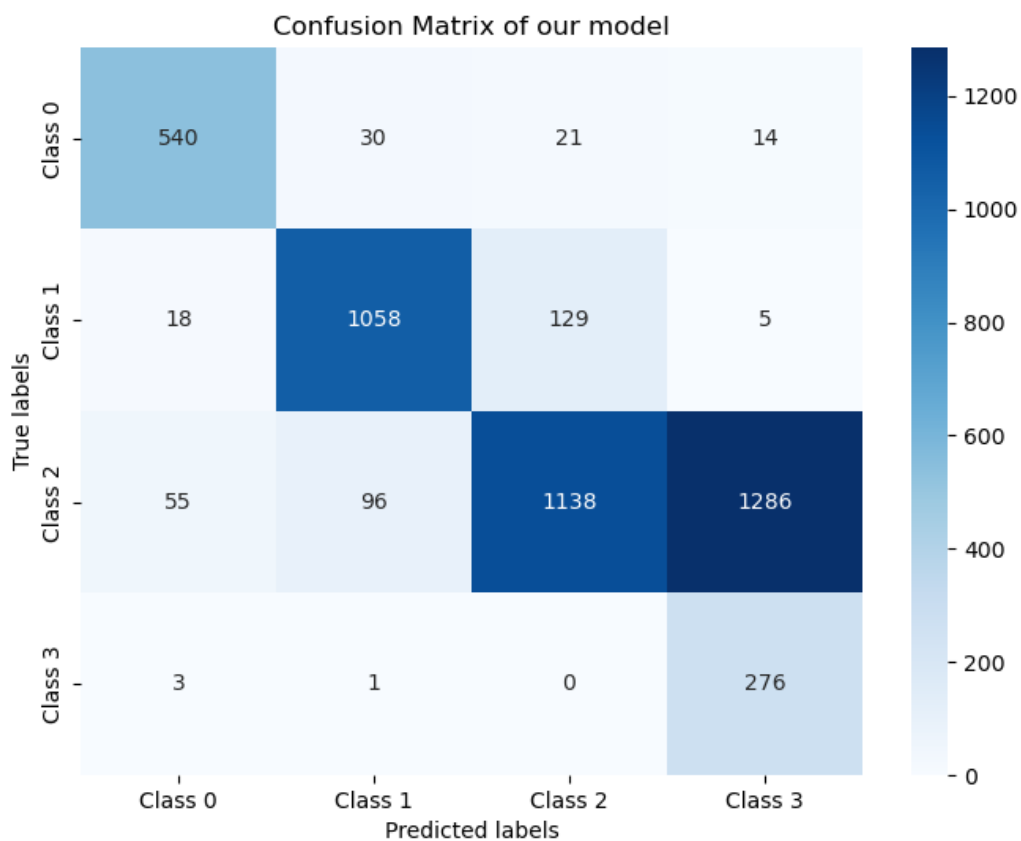
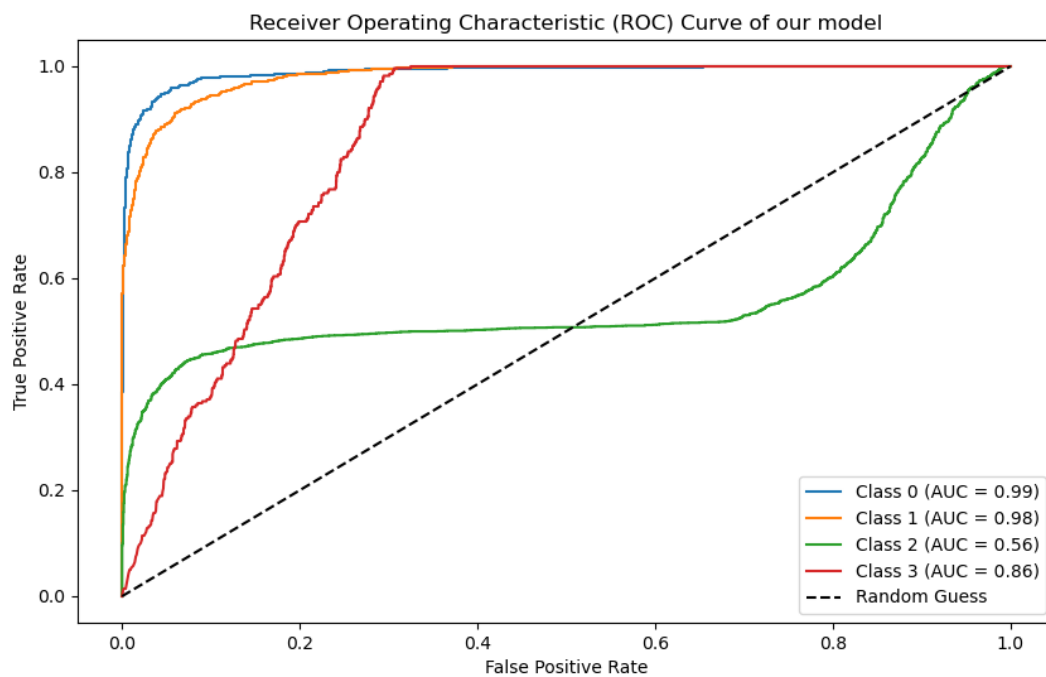


Figure 5. Our Model的ROC与Confusion Matrix

3. 对比结果讨论：



通过对比这些指标，我们可以观察到不同模型在分类任务中的性能差异。从准确率来看，我们的模型和ResNet50模型表现相似，略优于ConvNeXt和Conformer模型。在宏平均精确率、宏平均召回率和F1值方面，我们的模型和Conformer模型相比表现较好，但仍低于ResNet50模型。在AUC指标上，我们的模型和ResNet50模型处于相似水平，高于ConvNeXt和Conformer模型。

从ROC曲线的特征上，我们注意到除ConvNeXt外，class 2（即normal 类的covid）AUC值相对较低，排除其样本量少的可能性，可能具有较低的特征区分度，即其与其他类别之间的特征差异较小。这可能是由于该类别与其他类别在特征空间上存在较大的重叠，导致模型难以准确地区分这些类别。

这些结果表明我们的模型在COVID疾病分类任务中具有一定的优势，相较于ConvNeXt和Conformer模型，我们的模型在综合性能上表现更好。然而，与ResNet50模型相比，我们的模型在某些指标上仍有改进的空间。

## 5. 不足之处

由于时间与个人能力不足，对于模型还有更多的提升空间：

（1）对于网络结构方面，如对编码器与解码器的选取方面可以采用更为先进的模型，从而达到更为理想的实验效果。

（2）由于数据集的规模有限，为提升实验效果，可以在进一步增加数据增强手段，如随机旋转、翻折等方式，用以进一步的提升模型分类性能。

（3）对于分割任务对分类任务的辅助而言，除了正则项约束之外，还可以采用更多的方式，对分类与分割使用不同的网络，在网络之间通过参数共享等方式实现任务。

（4）分割任务通常需要更多的计算资源和内存来处理像素级别的预测。这可能会导致训练和推理的速度变慢，并对模型的可扩展性产生限制。

（5）在多任务学习中，不同任务之间可能存在相互干扰的问题。如果分割任务和分类任务之间的关联不明确或冲突，可能会导致模型性能下降。

（6）多任务学习中，不同任务之间共享特征表示的有效性是一个关键问题。某些任务可能需要特定于其自身的特征表示，而与其他任务的特征表示不同。在这种情况下，特征共享可能会受到限制，影响了模型的性能。

## 6. 总结

本研究中，我们使用一个多任务的U形网络模型，Res-UNet，用于同时实现肺部疾病分类和分割任务。我们的模型在与其它网络模型如Conformer和ResNet50的对比中，取得了非常理想的结果。

Res-UNet模型的设计灵感来自U形结构，通过编码器-解码器的架构，实现了有效的特征提取和重建。我们的模型能够自动学习和提取肺部X光图像中的关键特征，用于分类不同的肺部疾病，并且同时能够将识别图像中的病变区域。通过训练大规模的标注数据集，我们的Res-UNet模型展现了出色的性能。在肺部疾病分类任务中，我们的模型能够准确地将正常、病毒性肺炎、新冠肺炎和其他肺部感染等类别进行区分。与Conformer和ResNet50等其他网络模型相比，Res-UNet模型在肺部疾病分类任务上表现出更好的性能。通过将分割任务视作一种正则化技术，帮助模型在多任务学习中的任务互相约束，让模型学习到更一般化的特征表示，提高分类任务的泛化能力，使模型对新样本的预测更准确。

综上所述，我们的Res-UNet模型在肺部疾病分类任务上表现出优秀的性能，为医生和临床工作者提供了一种可靠且高效的辅助工具。未来，我们将进一步改进和优化我们的模型，以实现更准确和鲁棒的肺部疾病分析，并将其应用于临床实践中，提高患者的健康状况和医疗质量。

## 7. 致谢

感谢老师在课堂上的悉心讲述与助教的指导，帮助我们从零开始学习与了解生物大数据领域的经典算法。希望能够在后续课程的学习之中，进一步深入地了解更多生物大数据领域的算法与知识，丰富自身的科学素养与培养自身的学习能力，增强多学科交叉的知识学习。

## 8. 任务分工（待加）

### 1. 朱骏杰

- （1）Conformer对照实验
- （2）完成文章的介绍部分
- （3）ppt以及视频制作

### 2. 朱俊超

- （1）构建我们的模型
- （2）ppt及文章中有关我们模型的部分

### 3. 佟昊炜

- （1）ConvNeXt对照实验
- （2）完成模型评价指标部分的代码（ROC/AUC等）
- （3）ppt中有关结果讨论的部分

### 4. 鲁周语

- （1）ResNet50对照试验
- （2）完成模型评价指标部分的代码（ROC/AUC等）
- （3）文章中有关结果讨论的部分