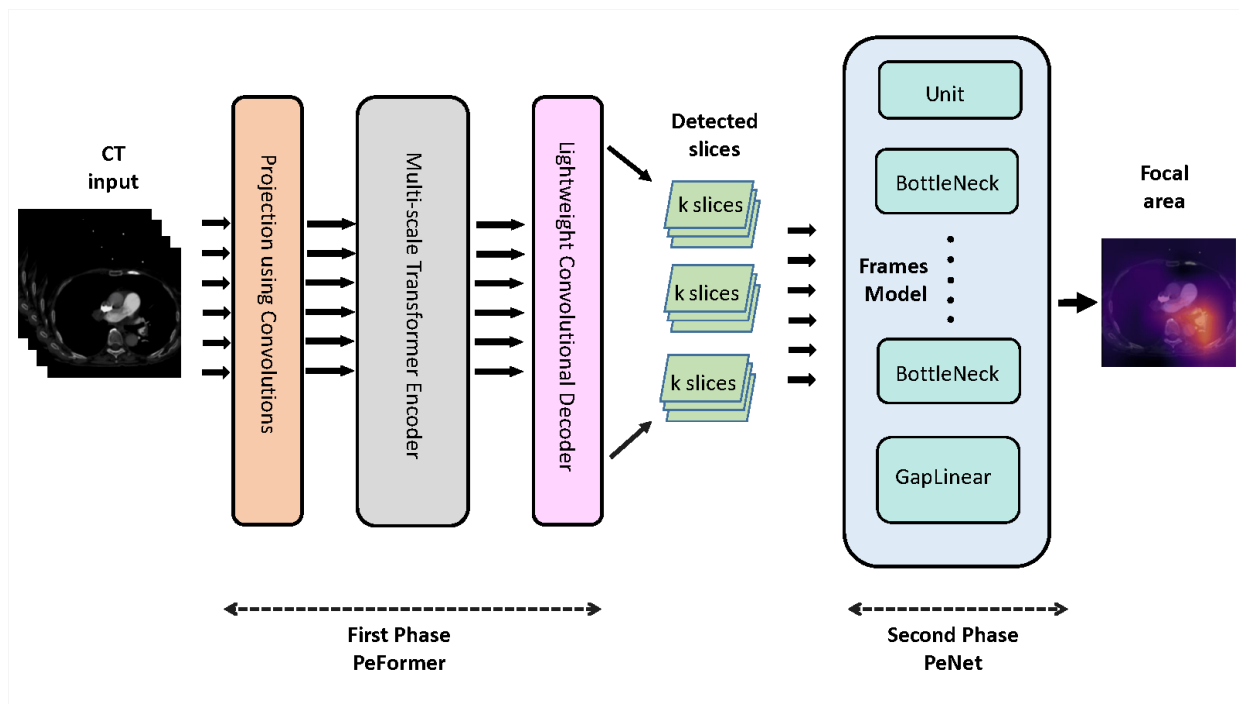


网络架构



特征提取

为了从CT图像中提取特征，我们采用具有时间维度建模能力的TSP方法，例如R(2+1)D-34。该方法能够有效地捕捉视频中的运动和时序特征。

注意力机制与ActionFormer

为了确定阳性或阴性，并定位可能含有病灶的视频帧序列位置，我们引入了注意力机制和ActionFormer。ActionFormer是一种模型，能够输出动作、onset和offset信息，用于对视频序列进行建模，并定位病灶关键帧。

病灶定位和PENet

基于输出的动作、onset和offset信息，我们能够确定可能包含病灶的视频帧序列位置。然后，我们提取每个可能的视频帧序列，并将其作为一组输入到经过微调的预训练PENet中。PENet是一个用于病灶检测的网络，能够计算每个视频帧序列的病灶概率。

我们对PENet的最后一个分类器进行改进，为每个视频帧划分不同的区域，如左上、左下、右上、右下、右中、主干等，然后计算每个视频帧的每个区域的病灶概率。这样以来每个区域都具有相应的概率，概率反映了存在病灶的可能性。

阳性判断和病灶区域获取

在获得每个区域的病灶概率后，我们对每个组的视频的的区域进行加权平均，我们可以设置一个阈值来确定阳性区域。超过这个阈值，则说明存在病灶，如果存在阳性区域，则将该CT图像包判定为阳性；否则，判定为阴性。

总体概述

通过以上网络架构，您可以输入CT图像，并获得以下结果：

- 阳性/阴性判定
- 若为阳性，获得相应的阳性视频帧序列

训练方案

数据集准备

首先，需要准备一个包含CT图像和相应标签的数据集。该数据集异常（阳性）样本。每个样本应具有对应的阳性视频帧序列位置和相应的病灶区域标注。

ActionFormer的训练

训练过程可以采用监督学习的方式，使用标注好的阳性视频帧序列位置标注作为目标。并且我们可以针对CT影像出现病灶的动作信息来微调损失函数，使其能够准确地输出动作、onset和offset信息。

PENet的微调和训练

在完成ActionFormer的训练后，可以进行PENet的微调。将之前预训练好的PENet加载到模型中，并使用阳性视频帧序列提取的数据作为训练样本。可以使用其他适合的损失函数来训练PENet，使其能够准确地预测每个视频帧区域序列的病灶概率。

网络整体训练和优化

在分别训练完特征提取网络、ActionFormer和PENet之后，可以将它们整合在一起进行联合训练。联合训练时，可以将所有网络的损失函数综合考虑，使用适当的权重进行加权组合。可以使用反向传播算法和优化器（如Adam）来最小化联合损失函数，从而优化整个网络。

模型评估与调优

训练完成后，需要对模型进行评估和调优。可以使用一个独立的测试集来评估模型的性能，计算准确率、召回率、精确度等指标。根据评估结果，可以进行模型的调整 and 参数调优，以进一步提高性能。