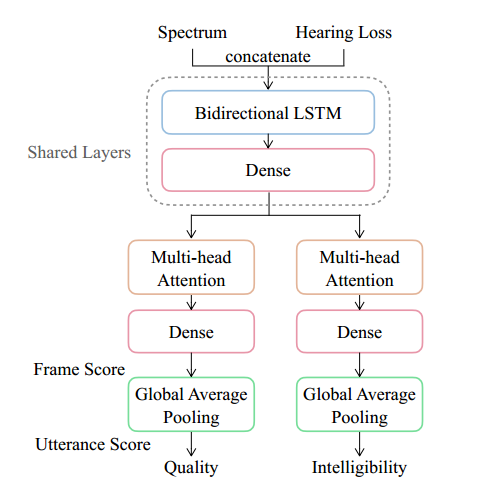
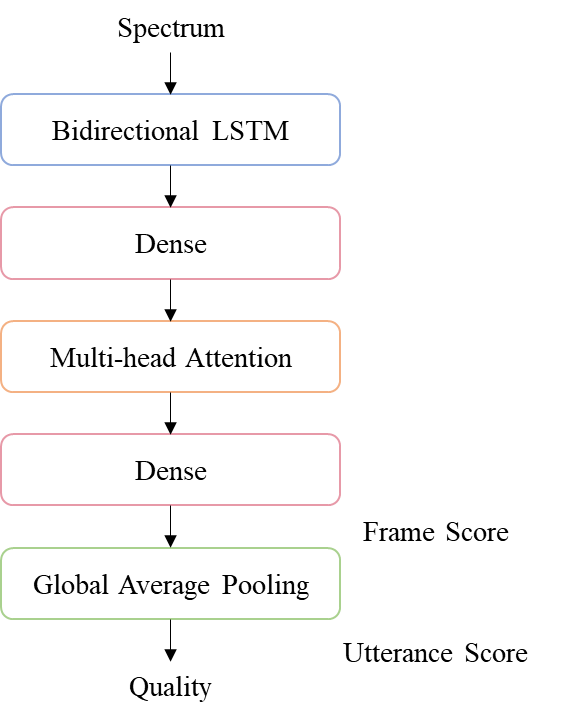
Module板块的覆写：

原网络结构：



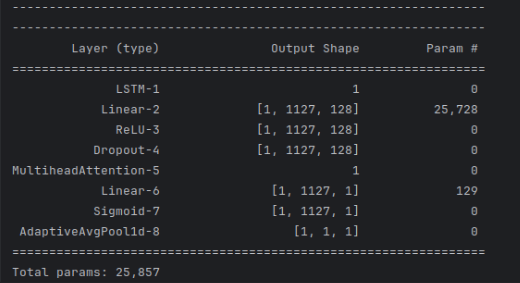
重新调整后的网络结构：



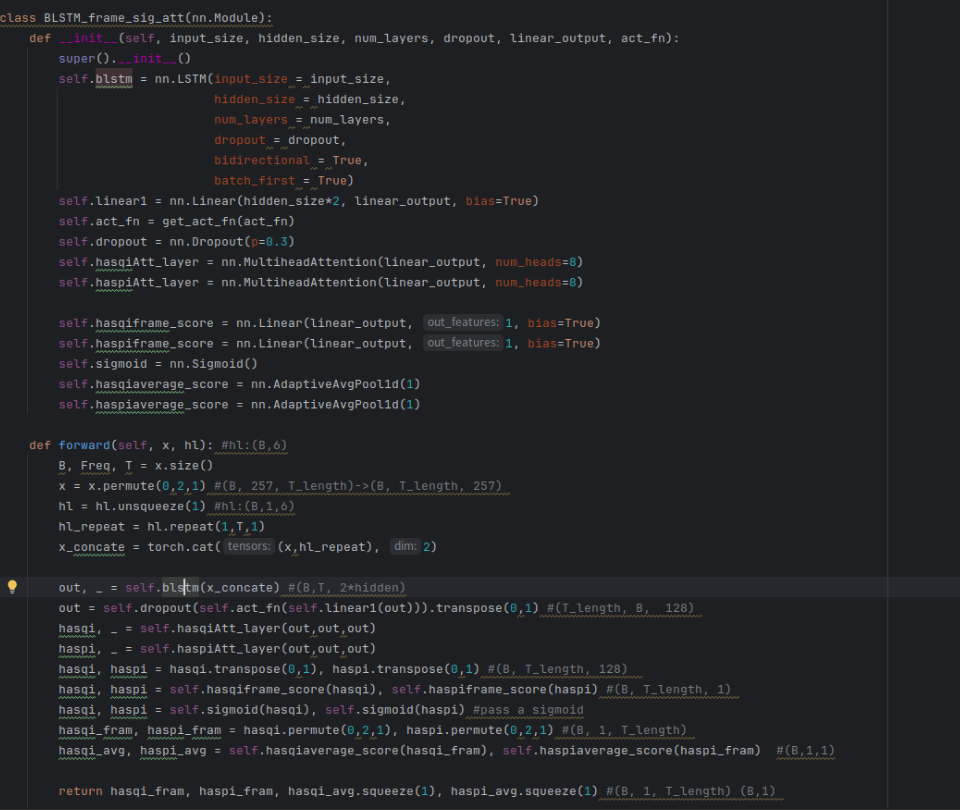
在网络修改过程中，幅度和听损双参数输入，利用torch.cat()拼接后输入网络进行预测。修改后不需要听损输入，直接输入幅度序列进行预测。

forward()函数的调整，删去原先的haspi和hl部分，保留hasqi部分，将hasqi作为POLQA，输入音频和truePOLQA进行训练。

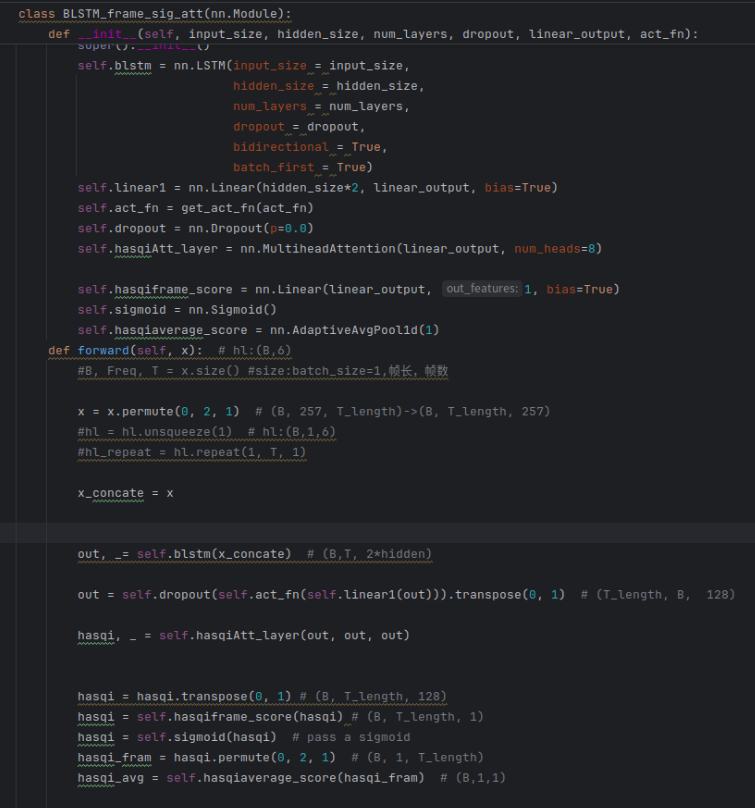
网络层的输入输出尺寸变化如图：

  
修改前后的module模板：

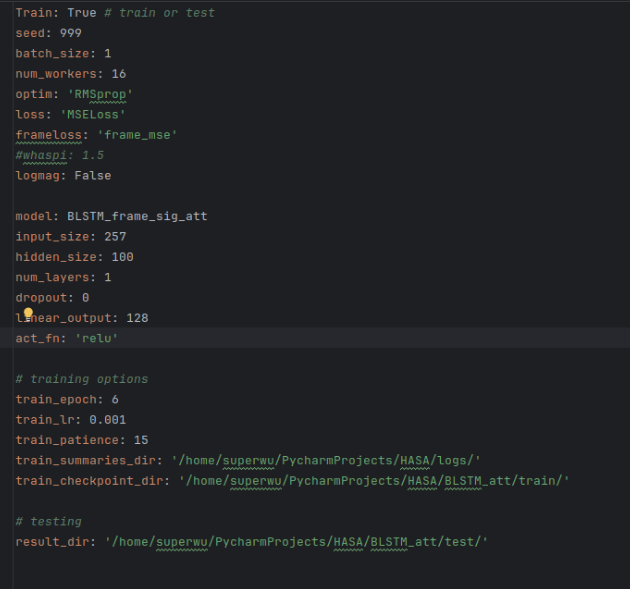
修改前：



修改后：



修改完module后，将trainer.py调整，去除haspi部分，调整输出，调整总损失输出。



上图为训练模型前的参数预设，学习率设为0.001，用约80000个48K劣化音频作为训练集，epoch=2，用1000个48K语音数据集作为测试集。

经过训练后得到训练结果如图：

LCC:0.949896，SRCC:0.920563，MSE：0.008879，

从预测图以及测试输出可见HASA-Net网络应用于48K语音的质量评估，相较于Quality-Net表现出更为优异的性能，后续可尝试调整网络的参数以及网络层，或者加入更多语音指标的预测，如STOI、DNSMOS等。

