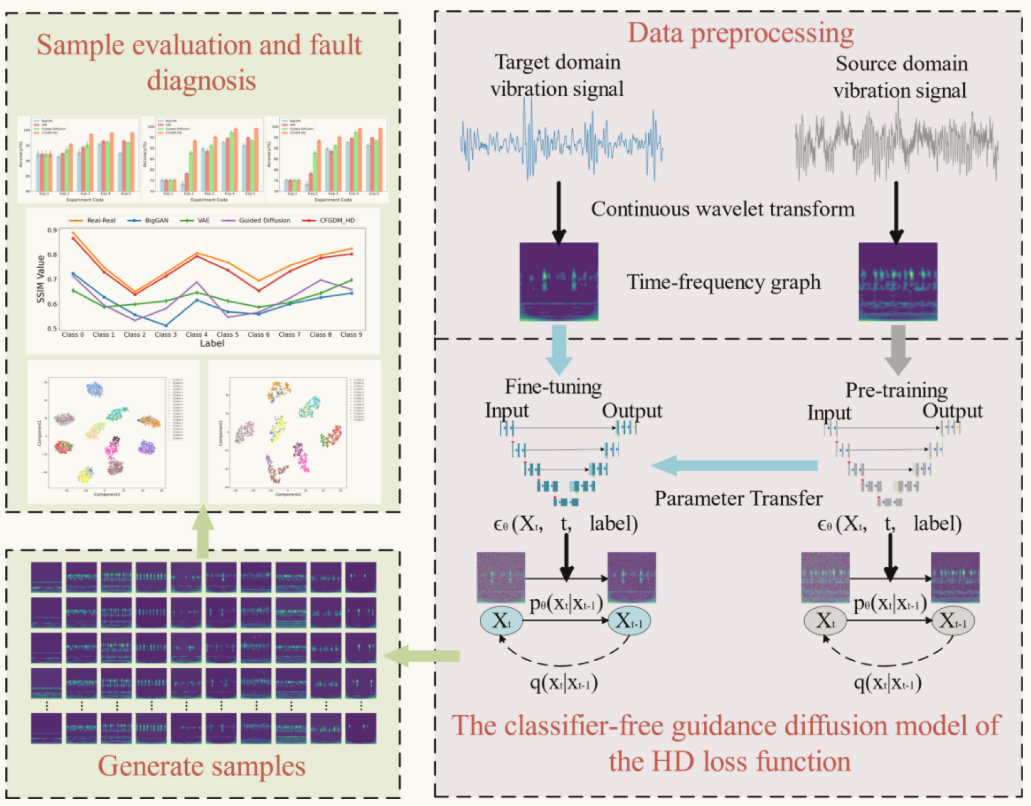
# 第十八次周报

## 这两周工作内容

### 阅读的文献

* 《A bearing fault data augmentation method based on hybrid-diversity loss diffusion model and parameter transfer》

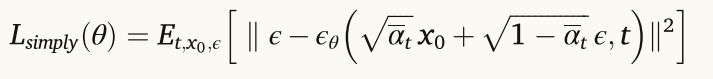
本文针对机械故障诊断中数据稀缺导致模型训练困难的问题，提出一种基于混合多样性损失扩散模型和参数转移的轴承故障数据增强方法（CFGDMHD）。该方法通过扩散过程从随机噪声中生成与真实样本具有相同数据分布的样本，能够同时生成多类别样本，无需额外的分类器指导。CFGDMHD通过引入多样性损失来提高生成样本的多样性。其模型结构如图：

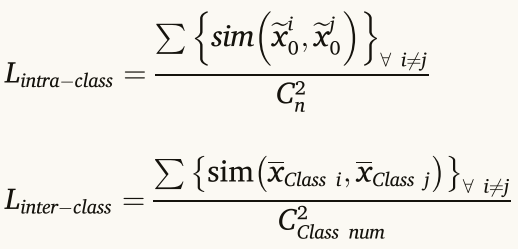


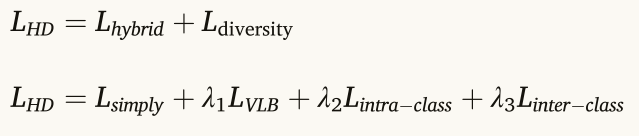
与以往大多模型不同，扩散模型提出时是应用于图像生成的，因此为了减小修改模型以适应一维振动信号的复杂操作，本文直接采用将原始振动信号使用连续小波变换转换到时频域进行处理，将时频图作为扩散模型的输入，并采用U-Net学习扩散模型的逆向过程，并通过联合无条件与有条件扩散模型的训练，实现无需外部分类器即可同时生成多类别故障样本，避免了额外分类器带来的对抗干扰和训练负担。这样它最终可以生成特定类别的噪声预测如下。



另外在原有混合损失（Hybrid Loss）的基础上引入多样性损失（Diversity Loss），由类内损失（Lintra-class）与类间损失（Linter-class）共同构成，采用余弦距离计算样本间的相似度，既保证了生成样本与真实样本的高相似度，又提升了样本间的差异性，避免模式坍塌。







此外，本文也采用迁移学习策略，先在大规模轴承数据集（如Paderborn）上进行预训练，再在小样本数据集（如CWRU、DIRG）上微调，有效缓解了小样本条件下模型训练不足的问题。

实验部分采用了 SSIM、IS、FID 三个指标对生成样本的质量和多样性进行量化评价，同时通过VGG16、ResNet50、DenseNet121等主流分类网络验证生成数据对故障诊断性能的提升。实验设置中将生成的数据与真实数据混合，不同的实验组逐步增加生成样本在训练集中的比例（如20:50、20:200、20:400、甚至仅生成样本无真实数据参与），对比不同方法在不同条件下的诊断准确率。结果表明，CFGDMHD在不同数据集上的生成质量和诊断准确率均优于ACGAN、VAE和Guided Diffusion等对比方法。

## 遇到的问题

### 2.1 扩散模型的数学推导较为复杂，特别是前向扩散与逆向扩散过程的概率建模、变分下界损失（VLB Loss）的优化目标，以及联合无条件与有条件扩散模型的设计，需要花费大量时间反复推敲，理解上仍存在一定难度。

## 收获与启发

### 3.1 对扩散模型在智能故障诊断中的应用有了更全面的认识，尤其是其相较于GAN和VAE的稳定性与可解释性优势，并且可以使用时频图作为研究对象，不过这会在一定程度上丢失原始一维振动信号的某些特征。

### 3.2 学习了在小样本问题下如何通过参数迁移和损失函数设计提升生成效果，有在生成过程中添加引导信息的可能，可以提升模型的可解释性。

## 下两周计划

### 4.1 继续阅读和整理扩散模型及故障预测相关文献。

### 4.2继续复现扩散模型的代码，实现简单的图像生成或信号去噪实验，加深对扩散过程、采样方法的实际理解。