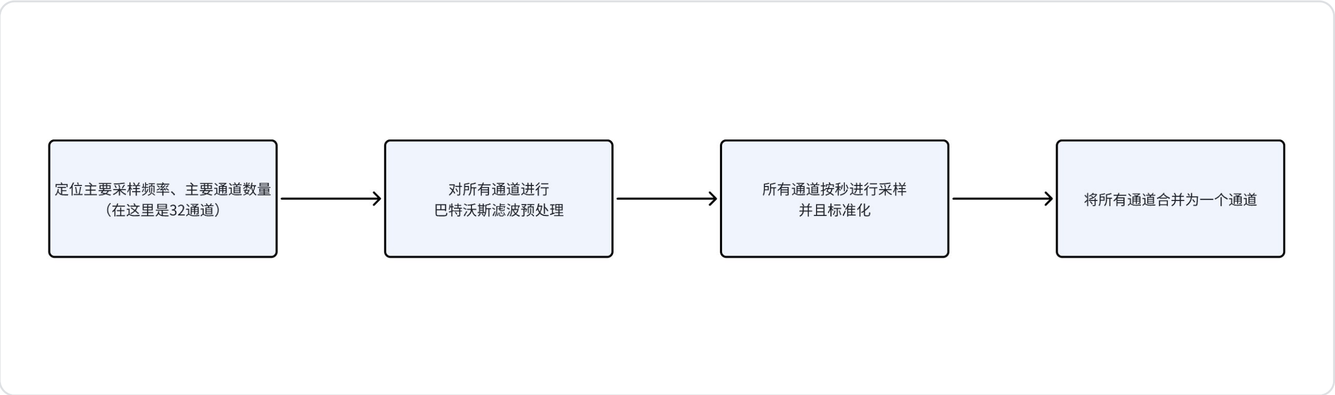
2.19 Studies on MFR-SAE

阶段总结：

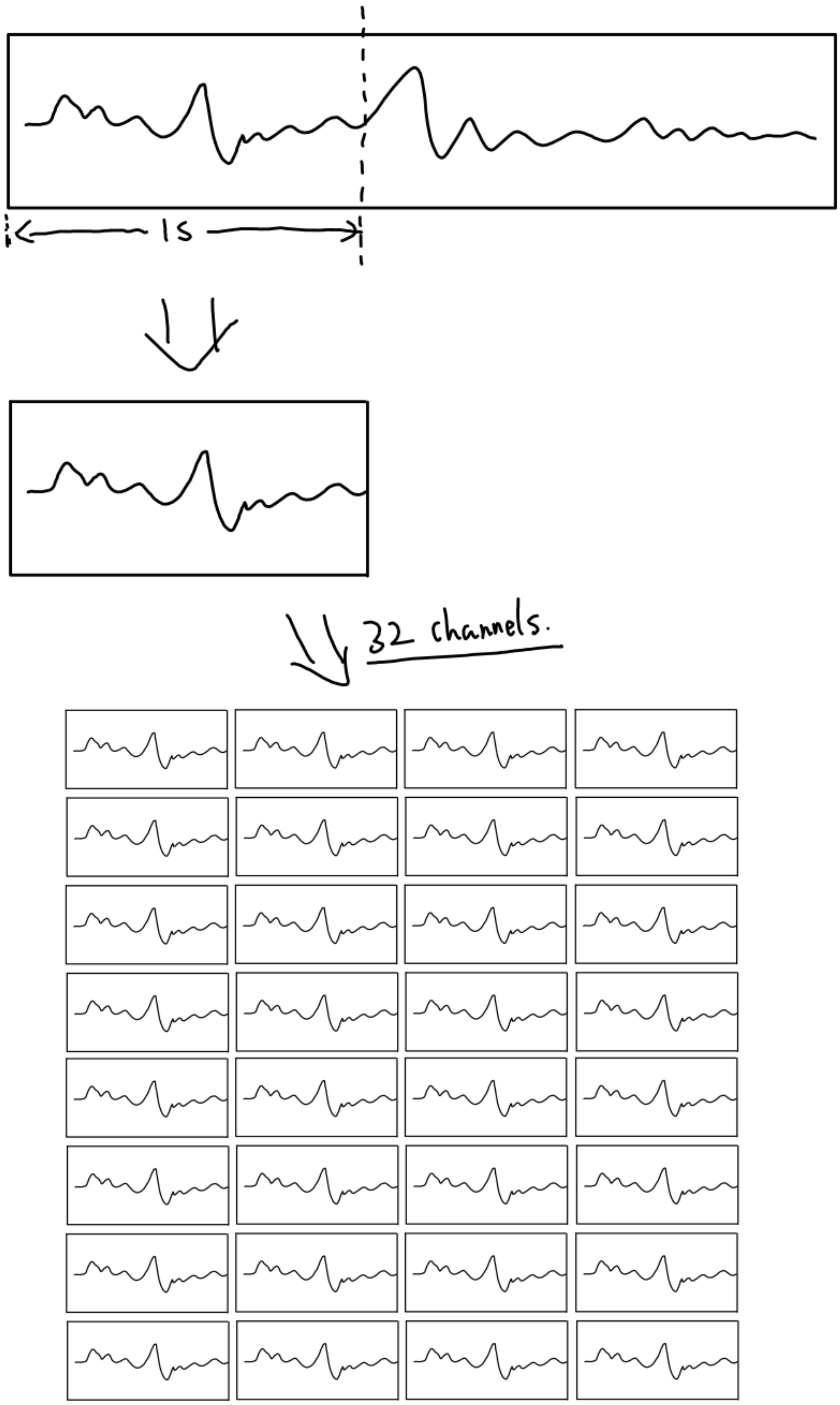
1. 掌握了 MFR-SAE
2. 掌握了使用 MFR-SAE 进行高逼真度 EEG 信号重建的方法
3. 对稀疏编码进行可视化（PCA）
4. 通过观察，验证了 SAE 与 PCA 的本质区别

MFR-SAE 计算过程

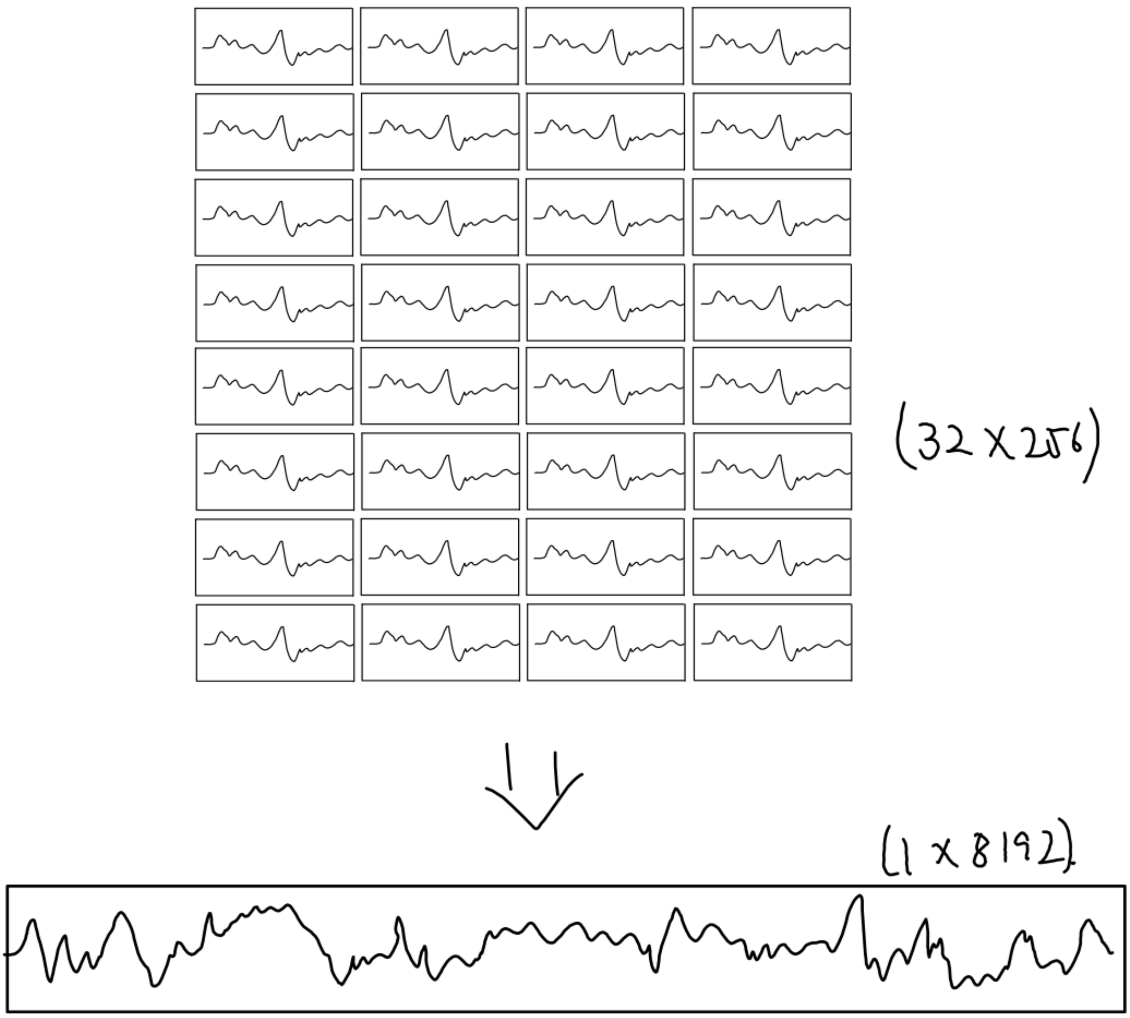
1. MFR-SAE 中对于 TUH-EEG 原始数据的预处理过程
    a. 预处理流程（EEG 数据具有32个通道）

    b. 数据形状预处理



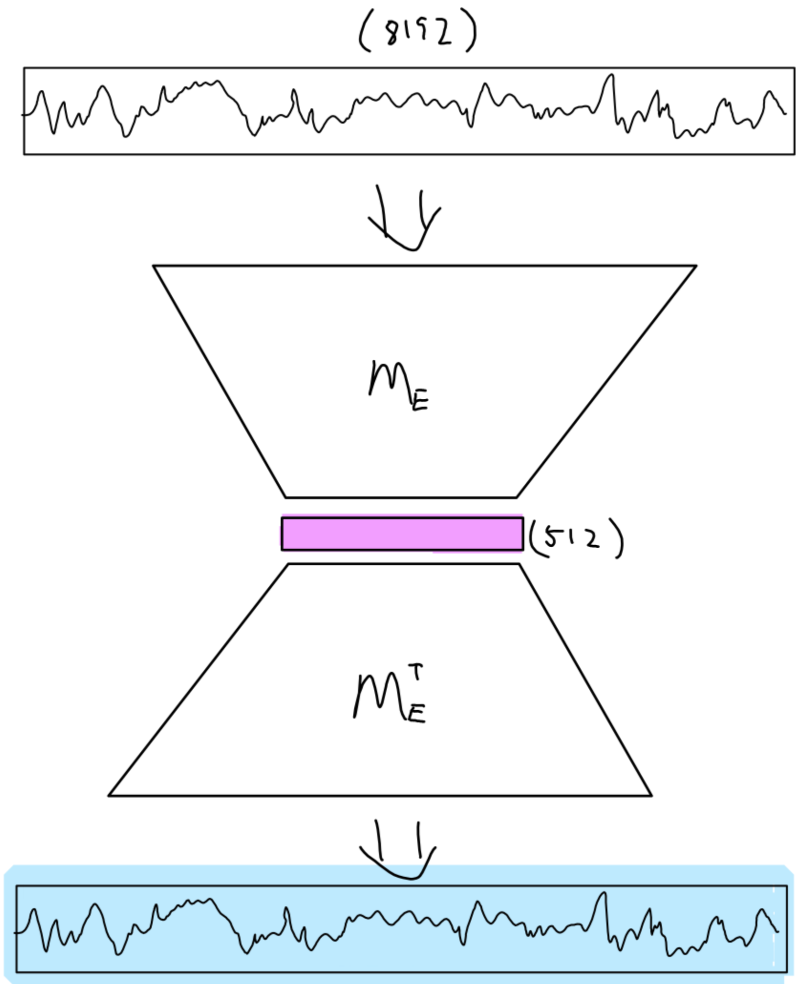
将所有通道的数据进行滤波，然后切割出 1 秒钟，得到32个通道的数据片段



将32个数据片段拼接成一个一维（超长）向量，作为 SAE 的输入

2. SAE 的处理流程

    a. 处理流程



在中间的编码层，进行 K_Sparse 的激活限制，然后用解码器输出

b. 权重捆绑技巧（tied weights）

在 SAE 中，采用 encoder 矩阵 M_E **的转置** M_E^T **作为 decoder .**

这种设计与字典学习（Dictionary Learning）有密切联系。在字典学习中，我们希望找到一个“字典” D，使得输入 x 可以近似表示为字典元素的线性组合，即

$$x \approx D\alpha$$

其中 α 是稀疏系数。这里，编码器的权重实际上就扮演了字典的角色，而稀疏激活（通过 TopK 策略得到的编码）则相当于稀疏系数。采用 tied weights 则相当于在编码和解码两个过程共享同一个字典，这样不仅符合字典学习的思想，也有助于提升特征的可解释性。

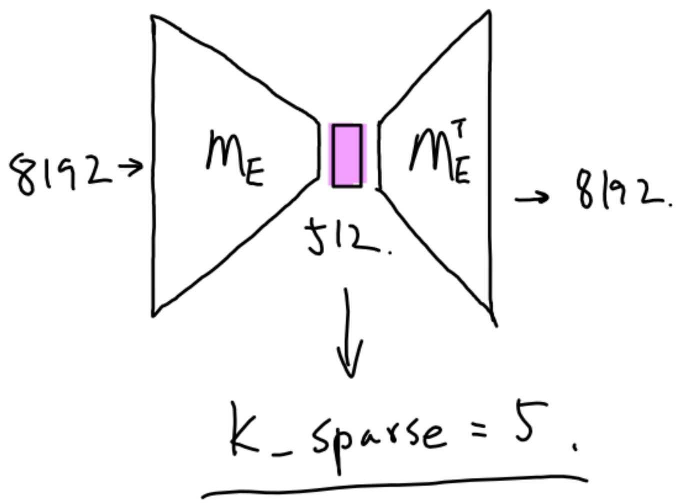
MFR-SAE 训练结果 & 效果改进

1. **原论文中的训练结果**

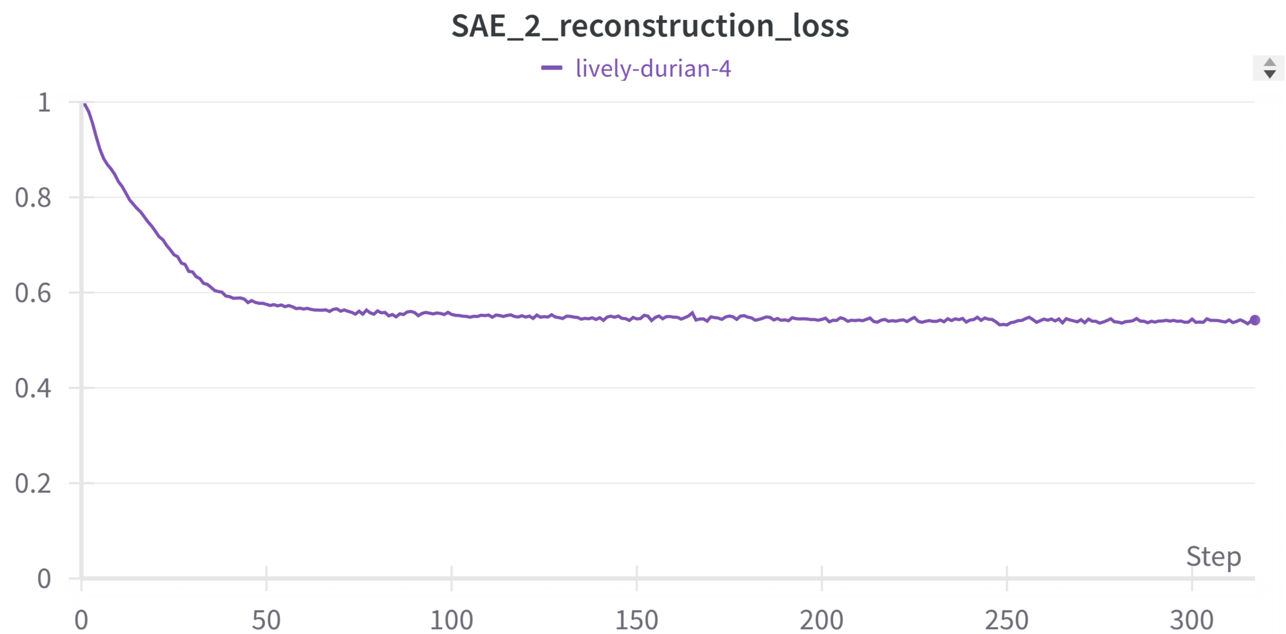
   a. hyperparameters（网络结构）

i. 编码层神经元数量：512

ii. K_Sparse: 5

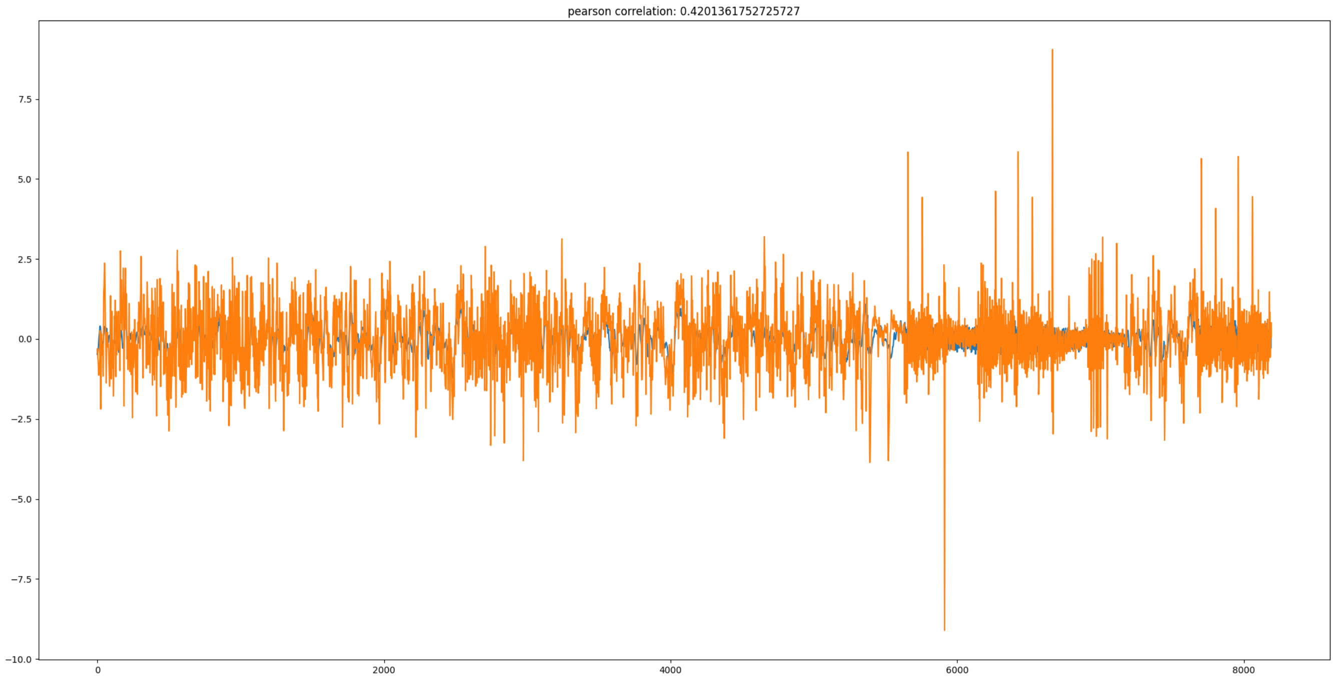


b. 训练 Loss

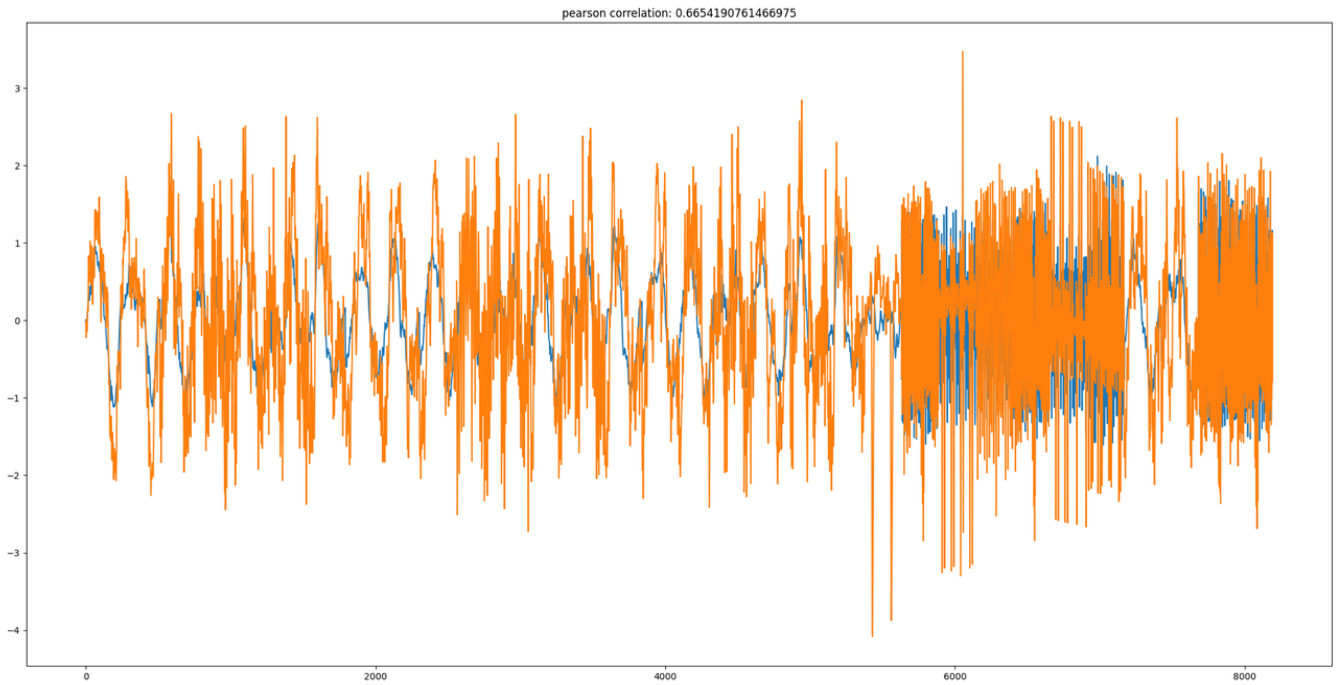


c. EEG 重建效果 & 相关系数

此时的 SAE 勉强能够重建 EEG 信号，但是相关性极低

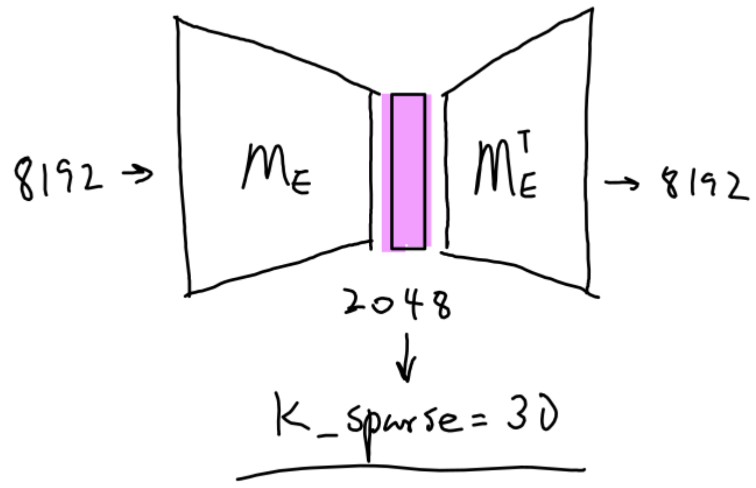


pearson-R = 0.42

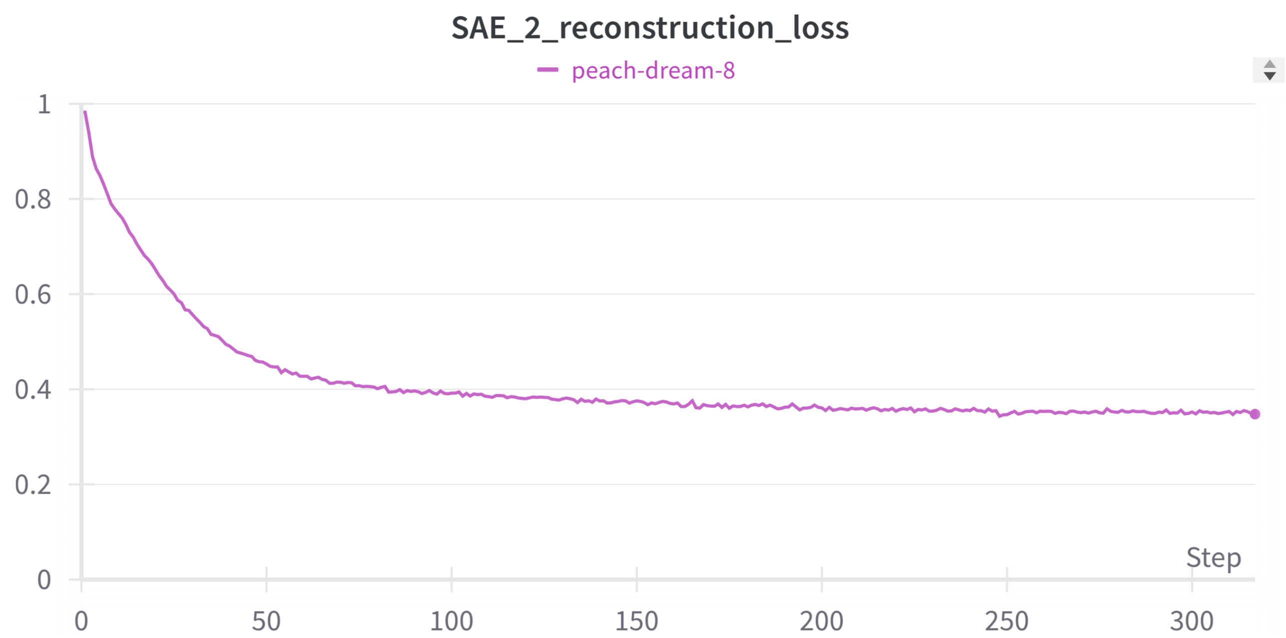


pearson-R = 0.66

2. 改进配置 1
   a. hyperparameters（网络结构）
      i. 编码层神经元数量：2048
      ii. K_Sparse: 30

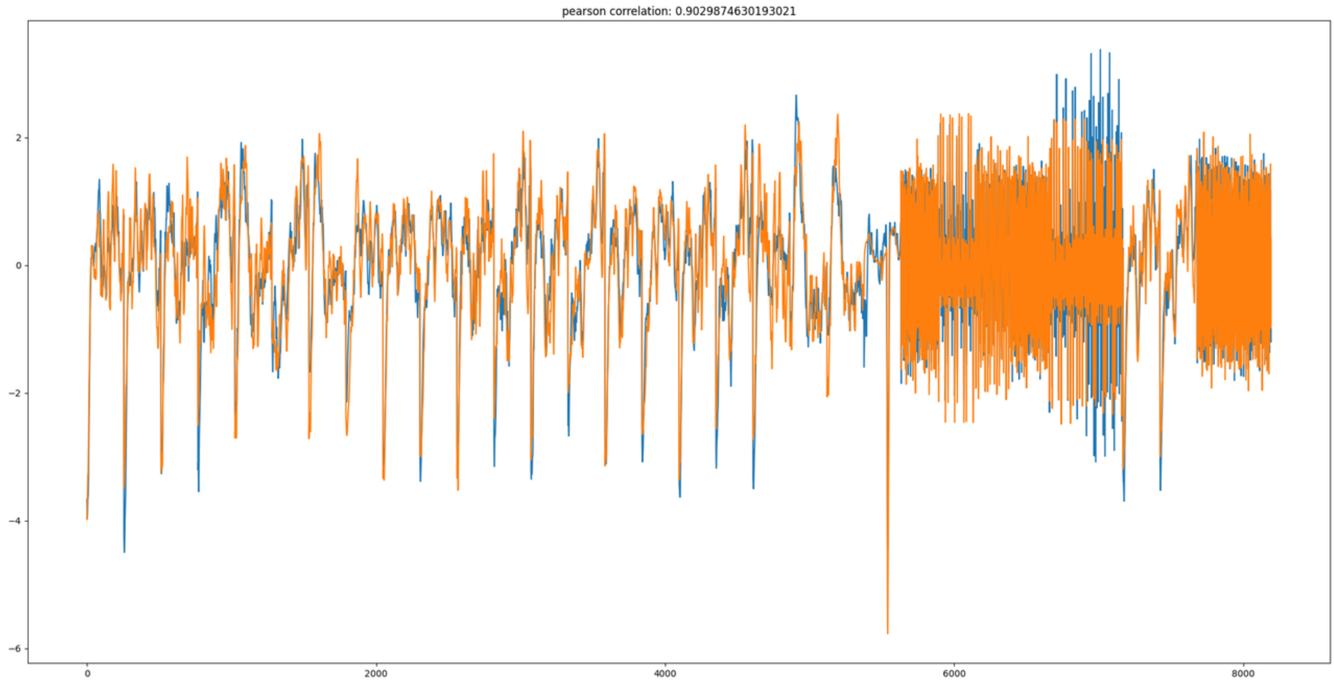


b. 训练 Loss

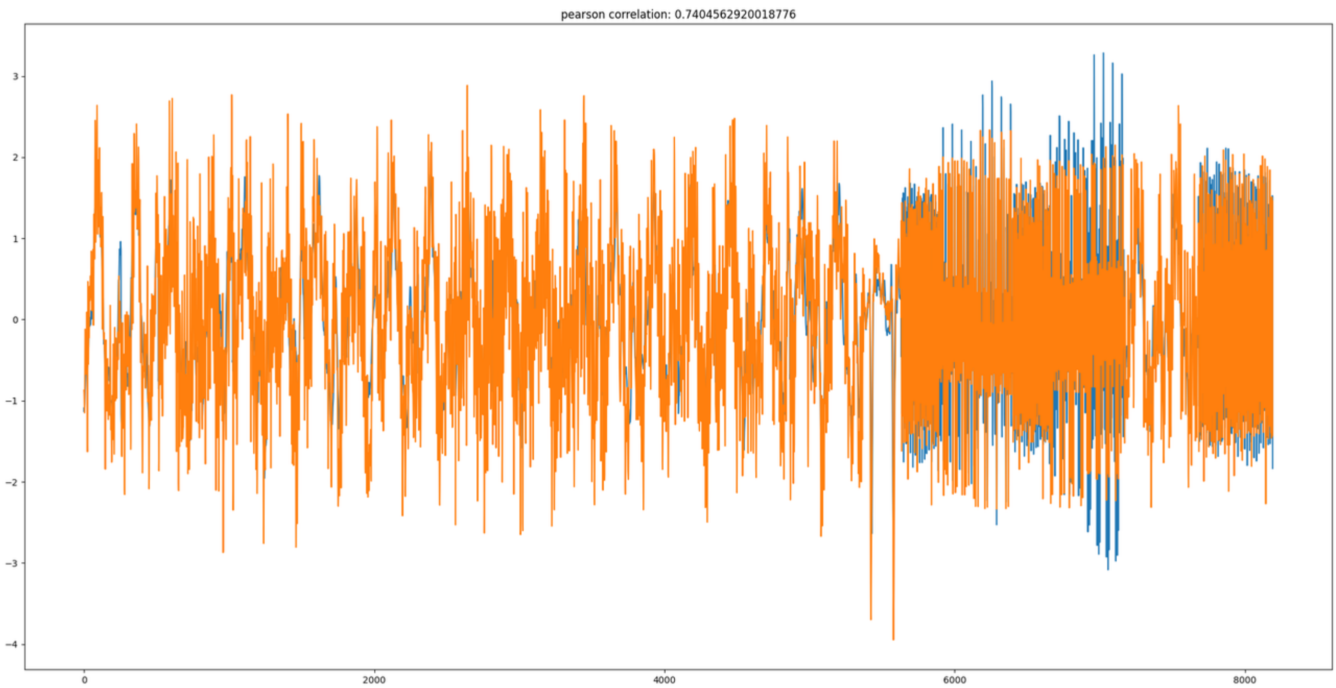


c. 重建效果 & 相关系数

此时的 SAE 已经基本能够重建 EEG 原始信号。

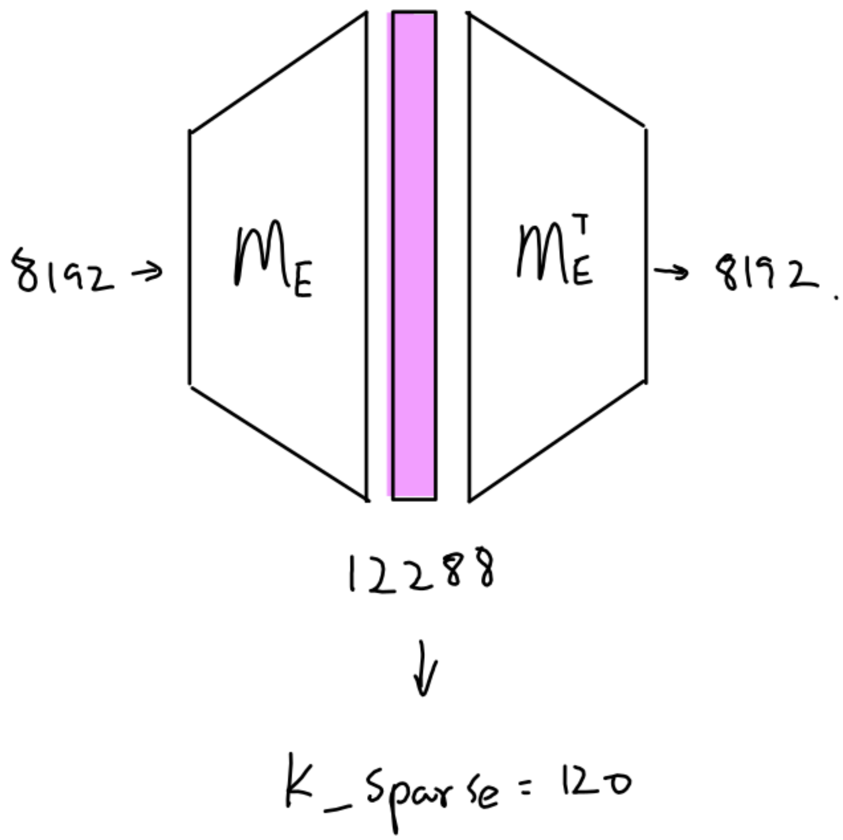


pearson-R = 0.90

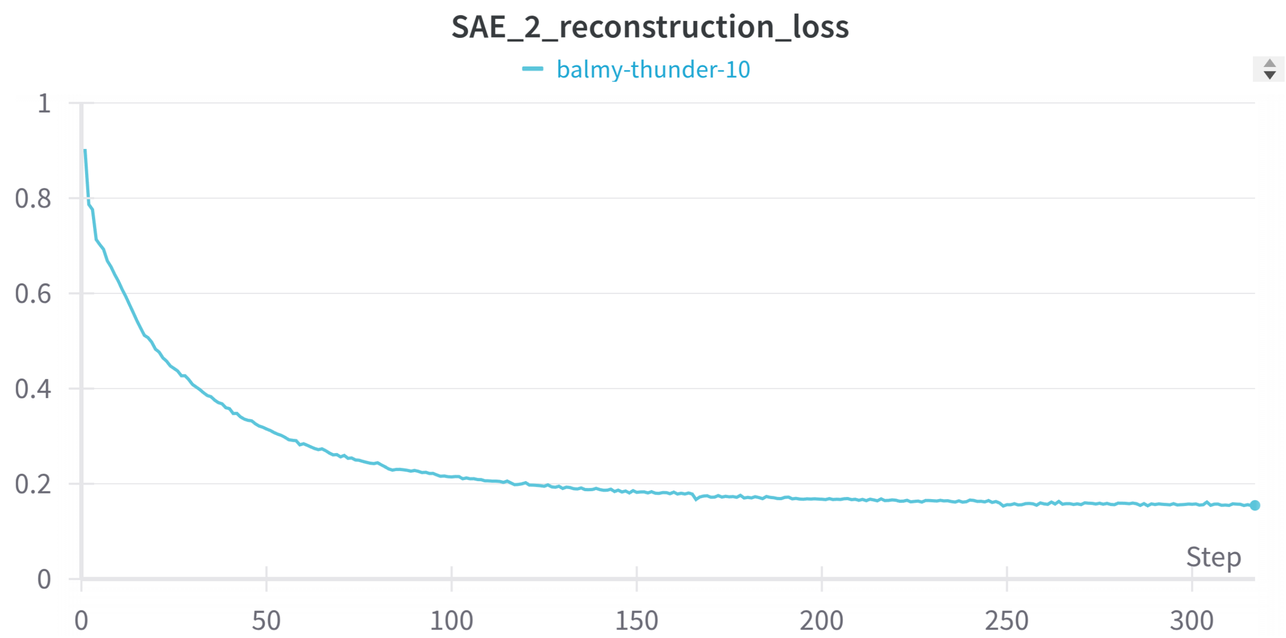


pearson-R = 0.74

3. **改进配置 2**
   a. hyperparameters（网络结构）
      i. 编码层神经元数量：12288
      ii. K_Sparse: 120

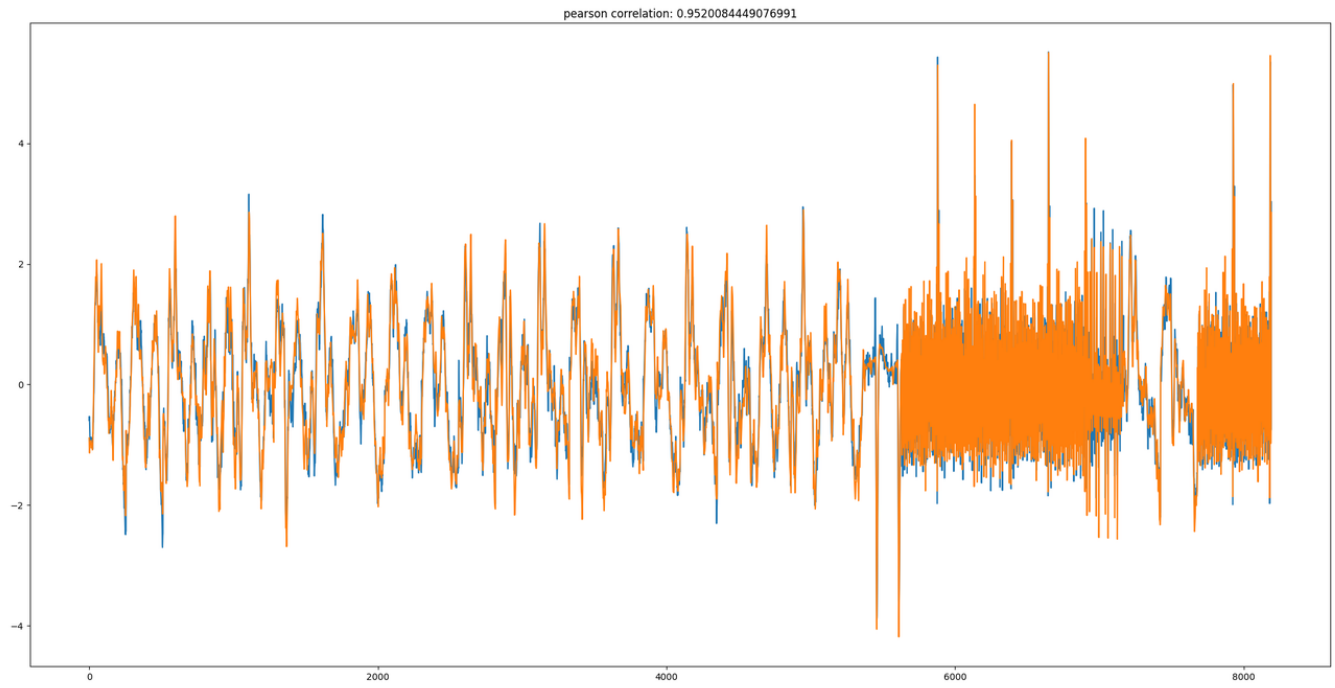


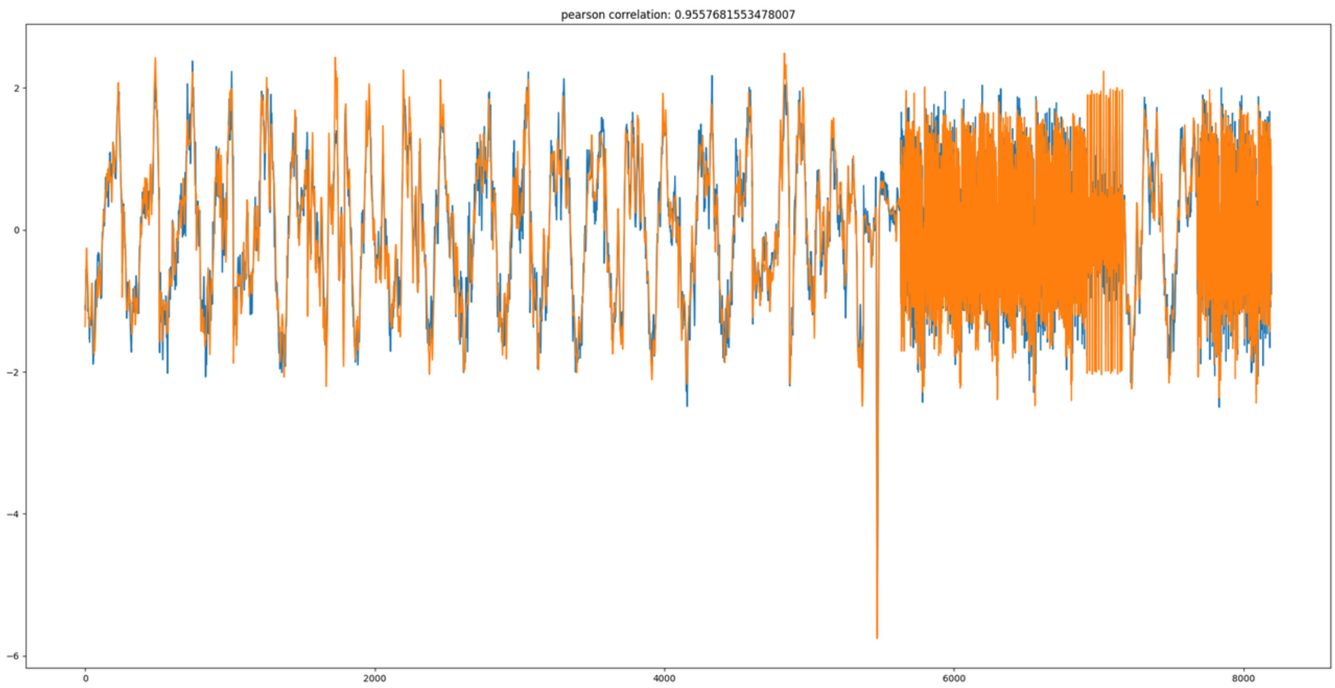
b. 训练 Loss

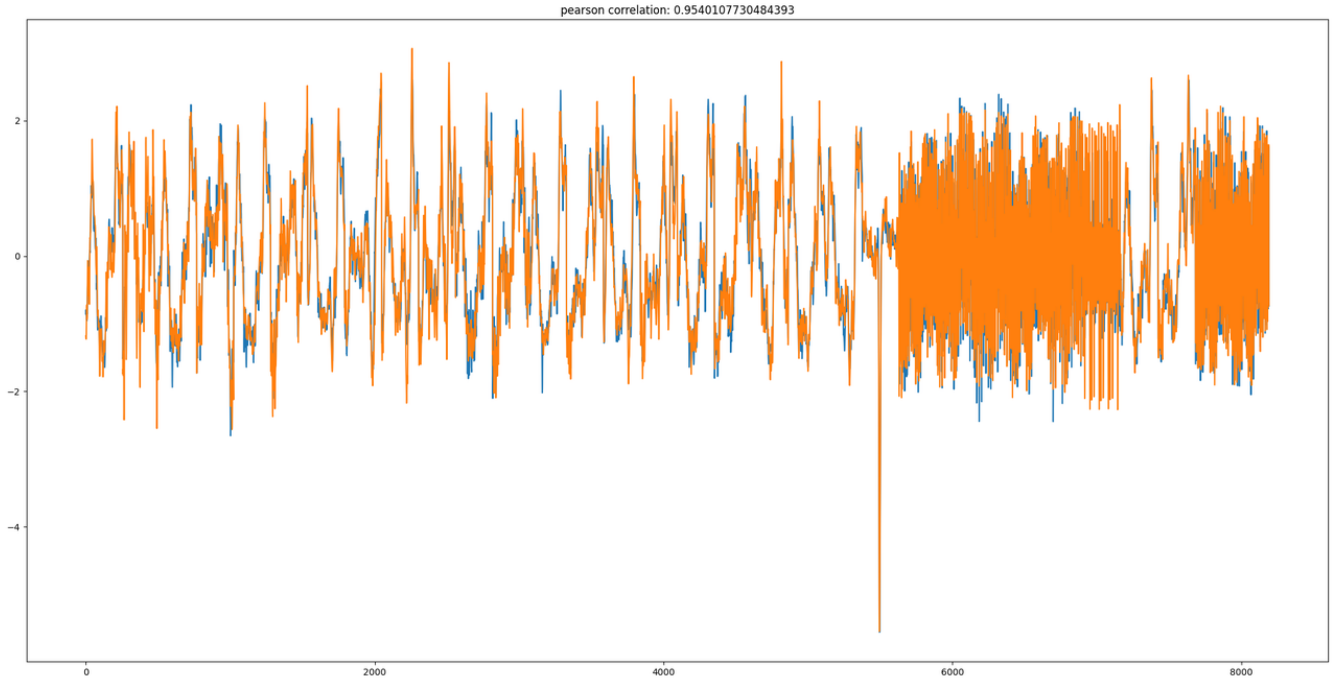


c. **重建效果 & 相关系数**

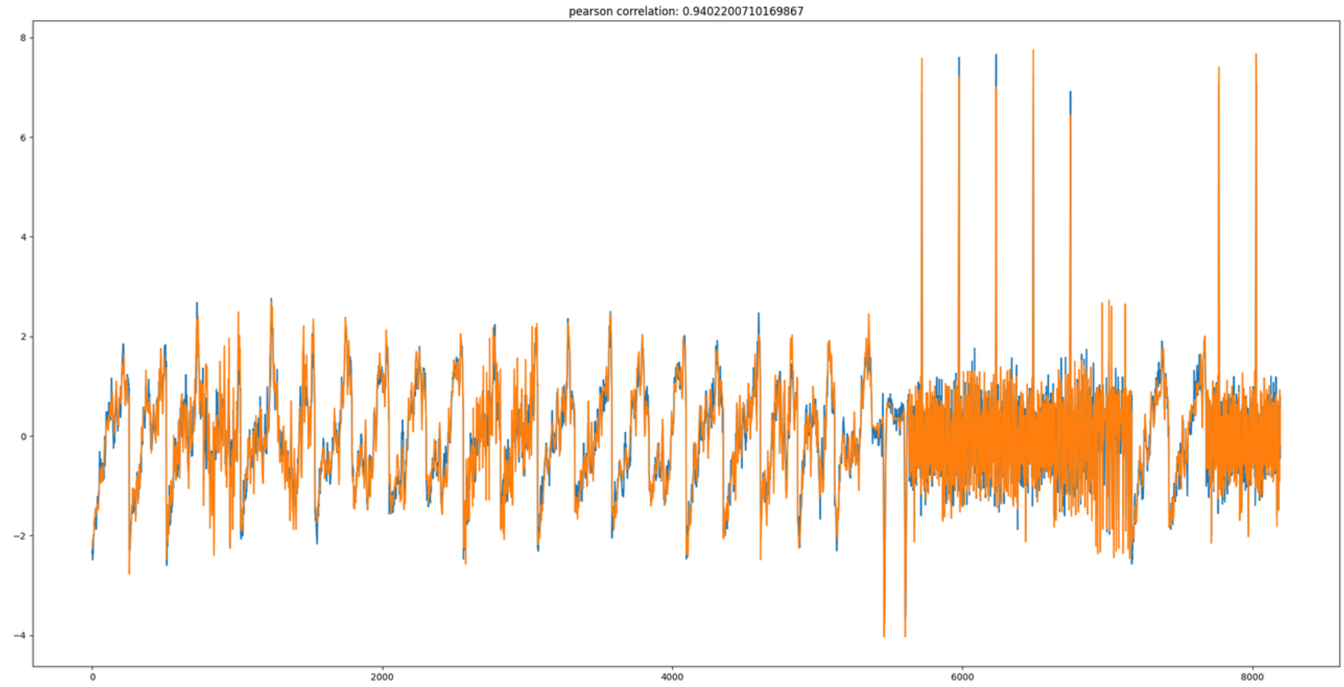
此版本为效果最好的一个模型，pearson-R 高达 93% 以上

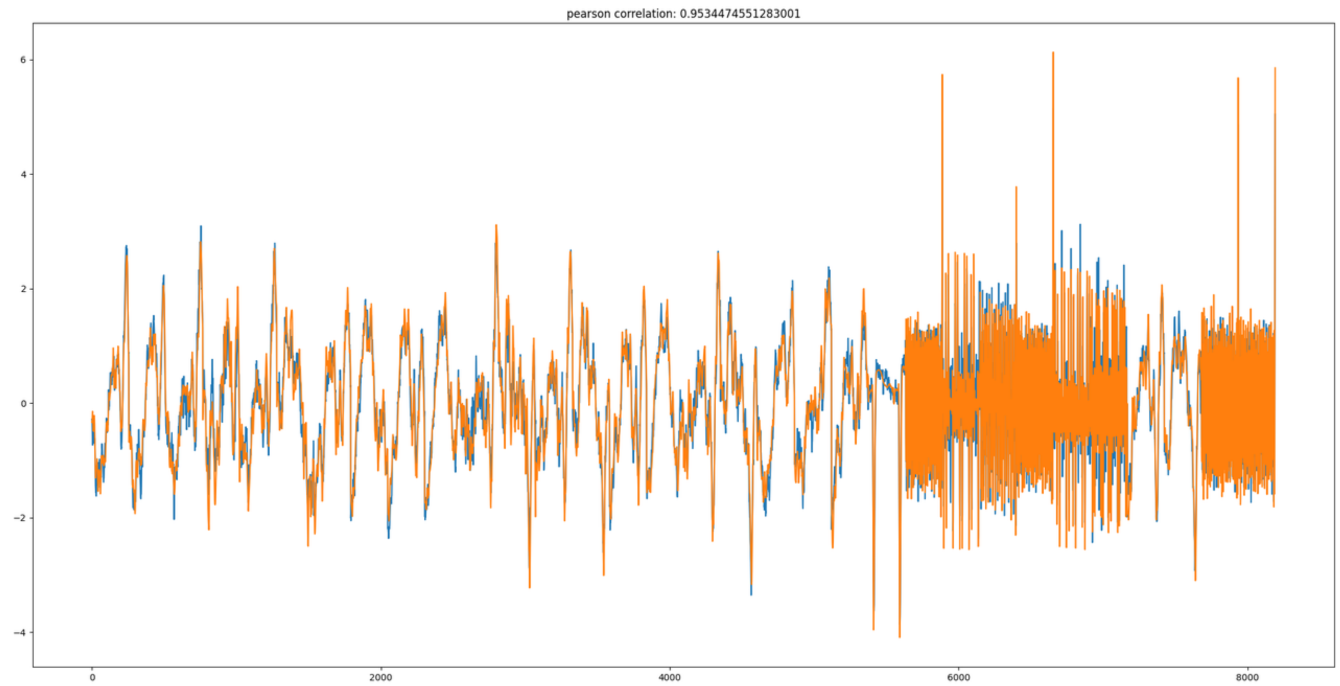

pearson correlation: 0.9520084449076991
4
2
0
−2
−4
0
2000
4000
6000
8000


pearson correlation: 0.9557681553478007
2
0
−2
−4
−6
0
2000
4000
6000
8000

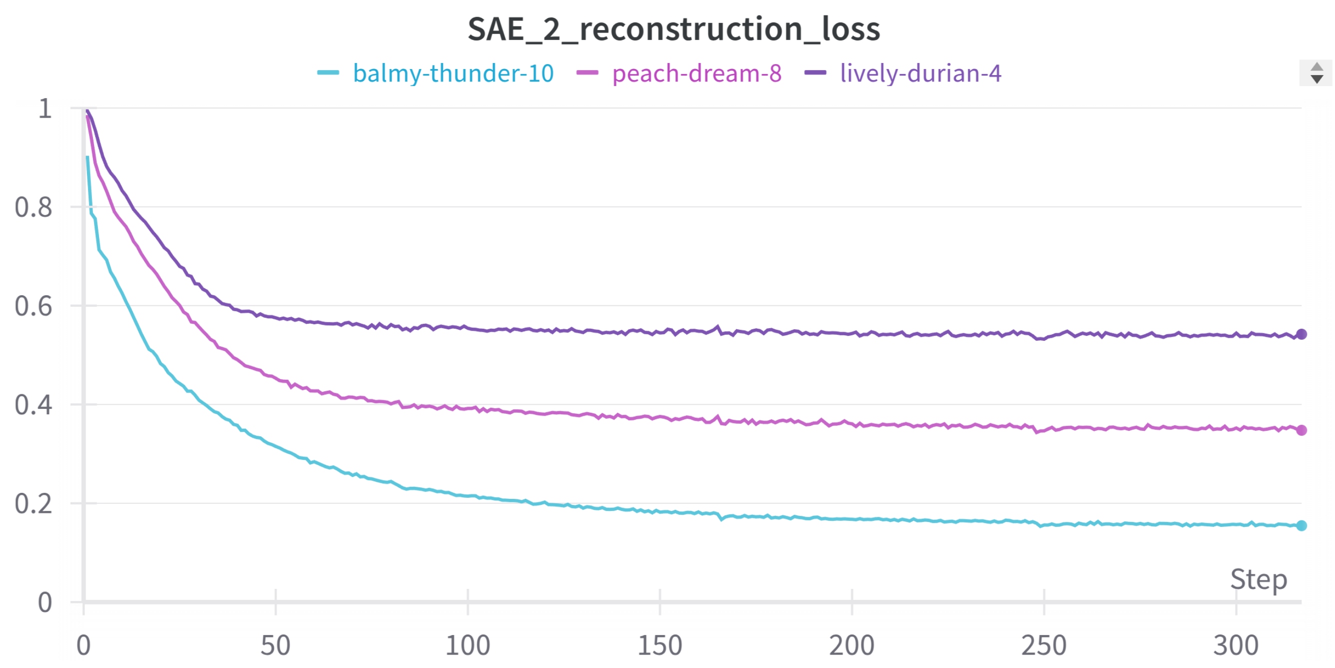

pearson correlation: 0.9540107730484393
2
0
−2
−4
0
2000
4000
6000
8000


pearson correlation: 0.9373873626684381
6
4
2
0
−2
−4
−6
−8
0
2000
4000
6000
8000


pearson correlation: 0.9402200710169867
8
6
4
2
0
−2
−4
0
2000
4000
6000
8000


pearson correlation: 0.9534474551283001
6
4
2
0
−2
−4
0
2000
4000
6000
8000

4. 三种参数的对比

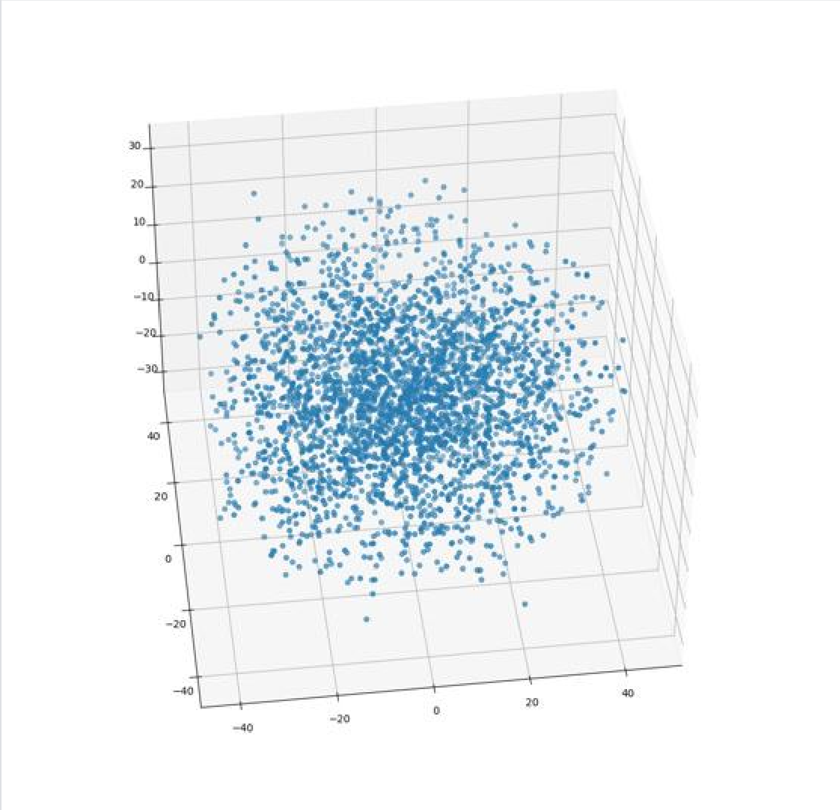


可以看出：

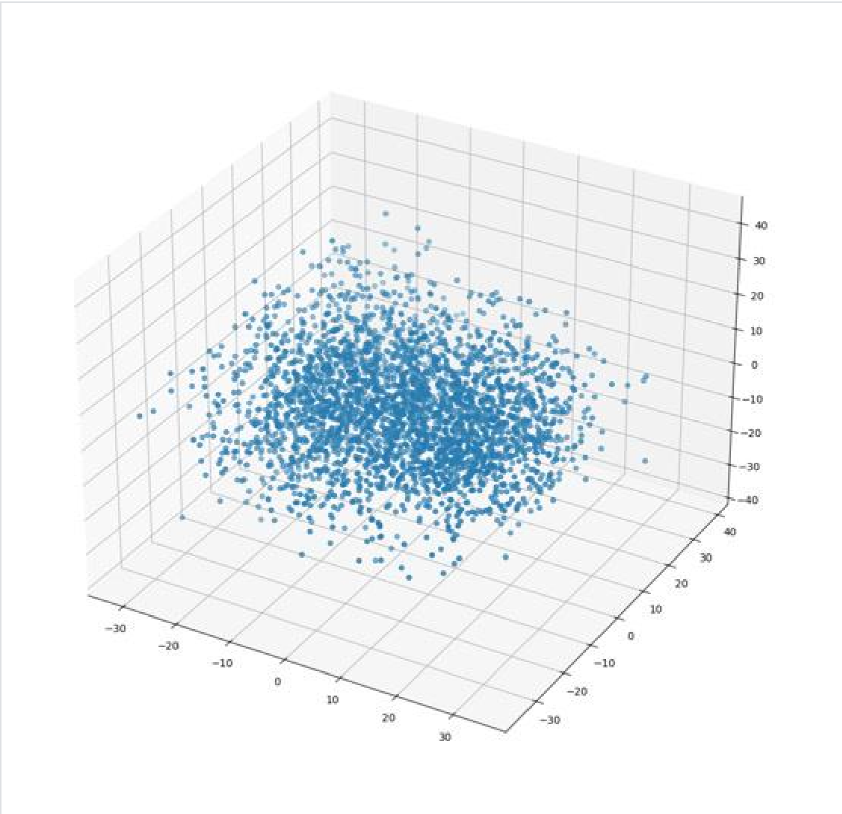
a. hidden_size 与 k_sparse 越大，重建效果越好
b. 尽管重建效果最好的是 k_sparse=120 配置，但是其同时激活神经元数量也远小于 input_size=8192 ，属于稀疏编码

SAE 中稀疏编码的结构（PCA）

1. PC0-2



2. PC3-5

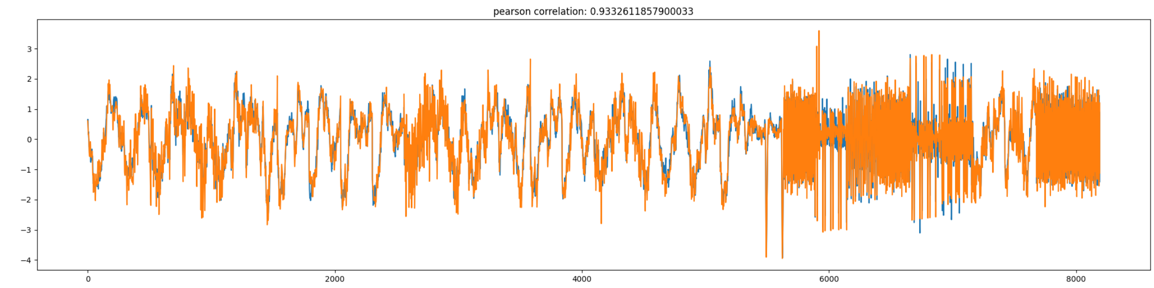


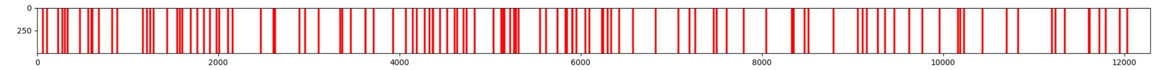
SAE 与 PCA 的本质区别

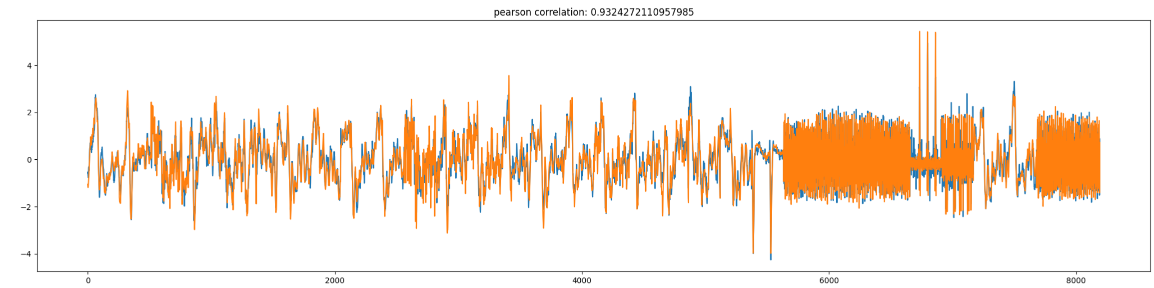
实际模型运行时，观察到了随数据变化而变化的 top-K 激活：

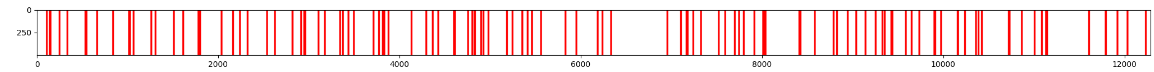

pearson correlation: 0.9578425008929172

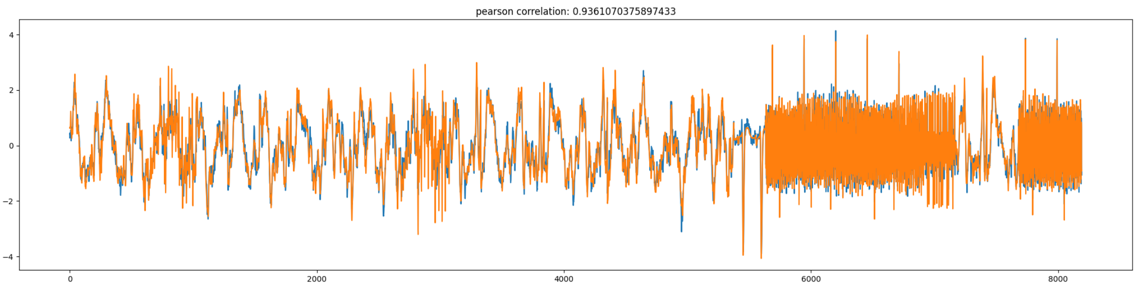

pearson correlation: 0.9311924798745964

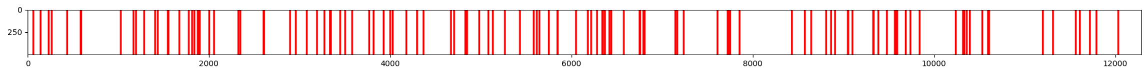

pearson correlation: 0.9332611857900033
3
2
1
0
-1
-2
-3
-4
0
2000
4000
6000
8000

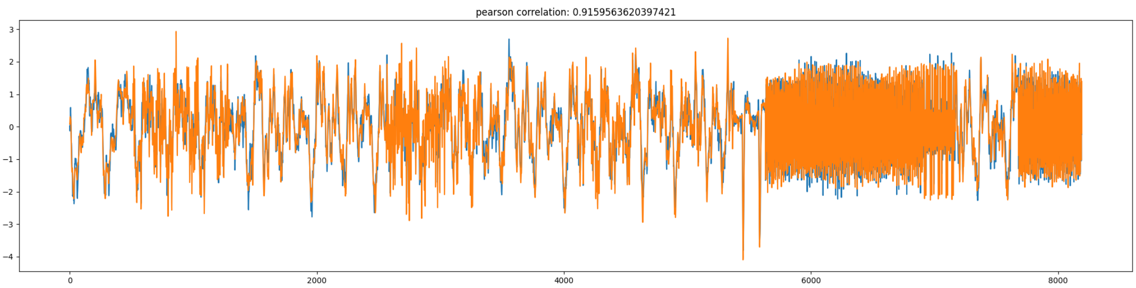

0
250
0
2000
4000
6000
8000
10000
12000

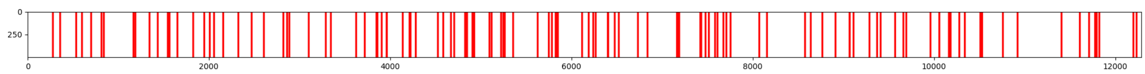

pearson correlation: 0.9324272110957985
4
2
0
-2
-4
0
2000
4000
6000
8000

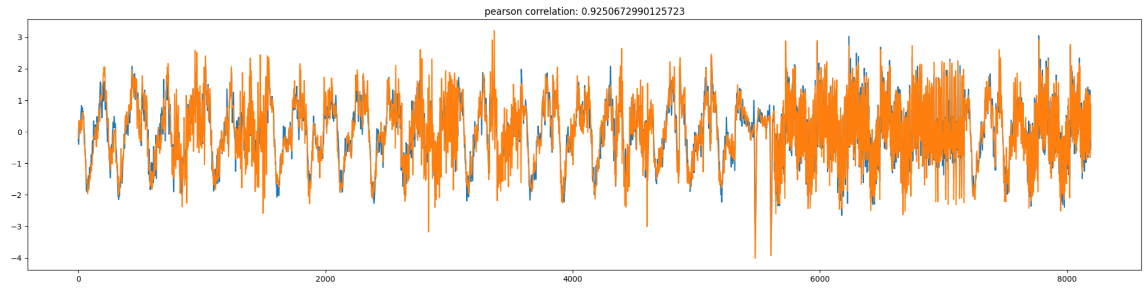

0
250
0
2000
4000
6000
8000
10000
12000

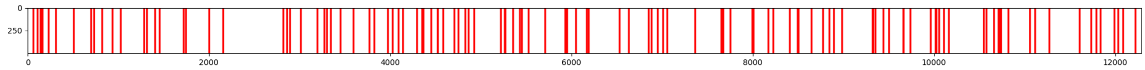

pearson correlation: 0.9361070375897433

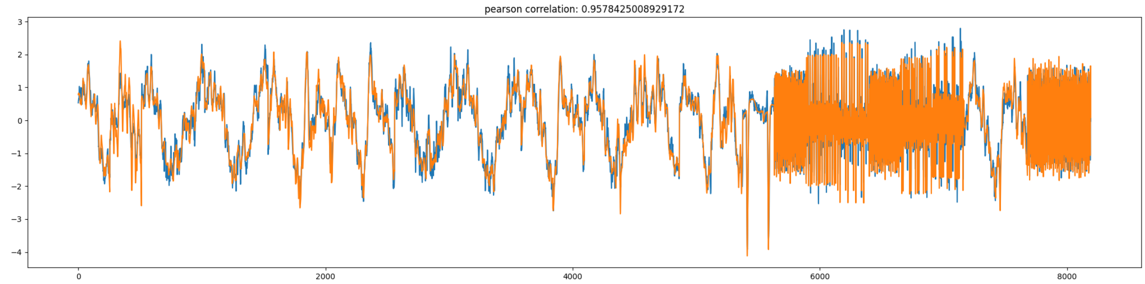


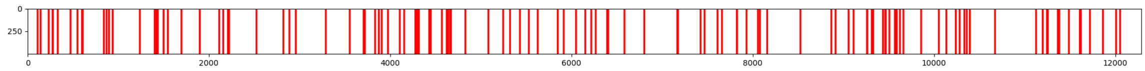

pearson correlation: 0.9159563620397421











本周 TODO

1. 目前的 SAE 将所有通道合并成了一个通道，这种做法是否适合接下来的分析，还需要重新考虑，可能要根据我们的需求对数据处理方法和SAE算法作出修改
2. 读入所有标签，将标签信号对齐到 (8192) 长度的 input 数据
3. 尝试用CCA分析：标签结构 vs Embedding结构