# Supplementary Materials:
Diff3Dformer: Leveraging Slice Sequence Diffusion for Enhanced 3D CT Classification with Transformer Networks

Zihao Jin, Yingying Fang, Jiahao Huang, Caiwen Xu, Simon Walsh, Guang Yang

Table 1: The experiment setting of the methods for comparison in the paper. z is the number of slices and p is the number of patches cropped from the whole CT scan.

Model Name	Learning Rate	Batch Size	Optimizer	Hardware	Input Size
WS-DenseNet121	1e-3	32	Adam Optimizer	One RTX3090	$64 \times 128 \times 128$
WS-ResNet101	1e-3	32	Adam Optimizer	One RTX3090	$64 \times 128 \times 128$
WS-Contrastive 3d	1e-4	4	Adam Optimizer	Two RTX3090	$64 \times 256 \times 256$
2.5D-ResNet101	1e-4	8	Adam Optimizer	Two RTX3090	$8 \times 256 \times 256$
AG-Swin Transformer	1e-4	2	Adam Optimizer	Two RTX3090	$z \times 224 \times 224$
ViT-patch	1e-5	4	Adam Optimizer	Two RTX3090	$p \times 64 \times 64$
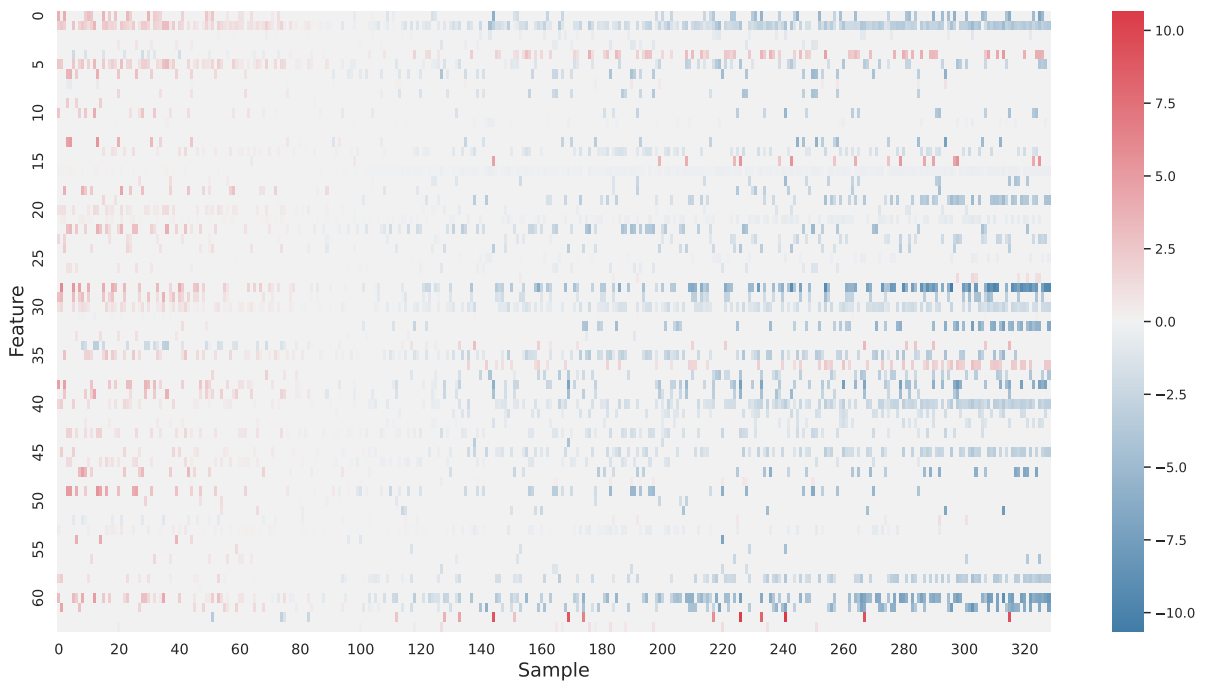
Figure 1: The heatmap represents the contribution of the cluster to the final patient-level risk score R on the FLD dataset. Patients ranked from highest to lowest risk score R on the horizontal axis from the left to right and 64 clusters on the vertical axis.
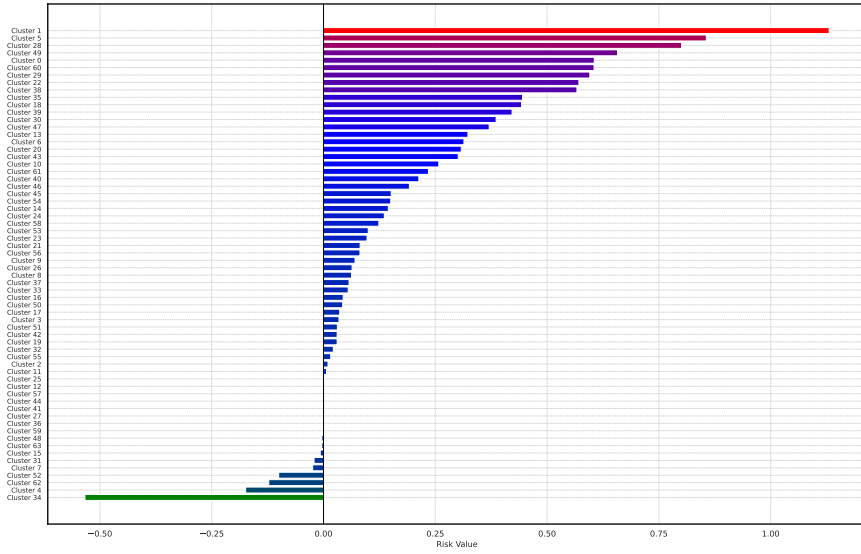
Figure 2: Cluster ranking by contribution to the 'mortality in one year' class on the FLD dataset.
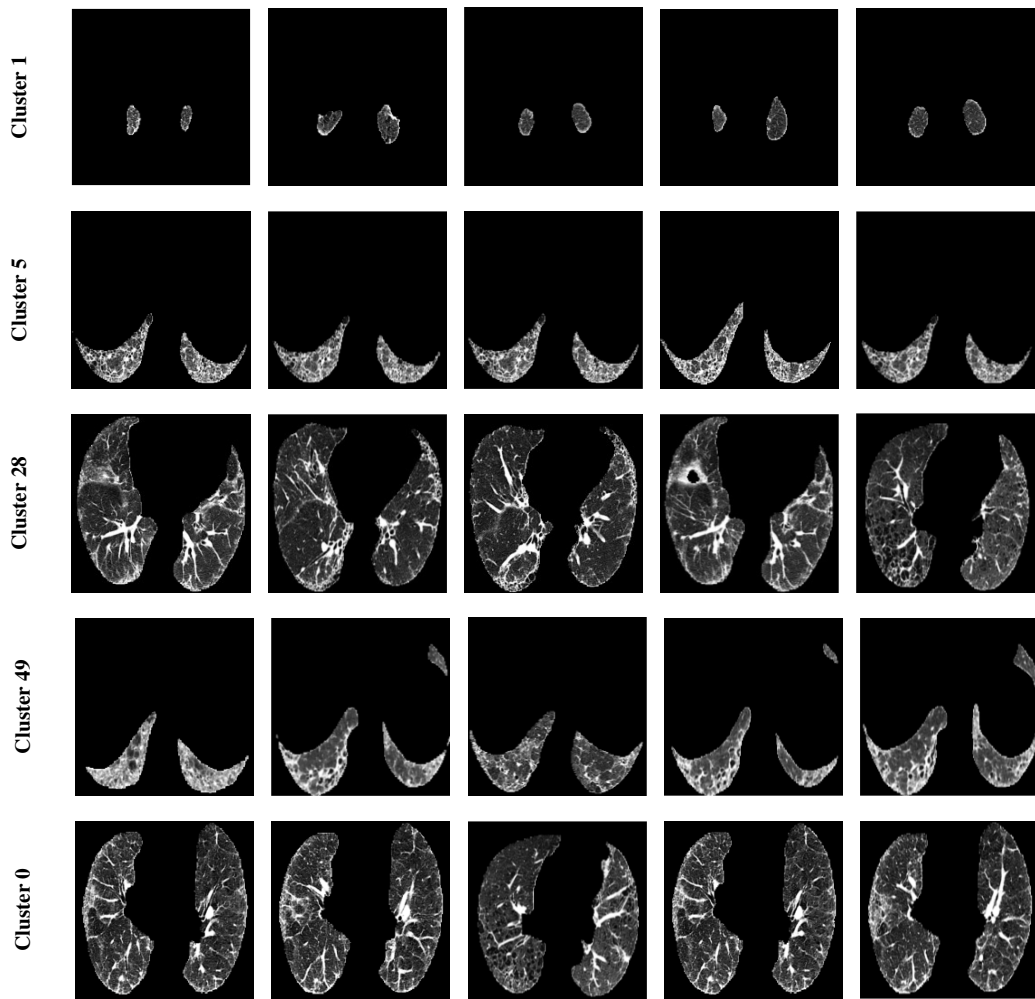
Figure 3: Visualization of the representative slices of high-risk clusters on the FLD dataset. The representative slices are those closest to the centroids of the cluster.