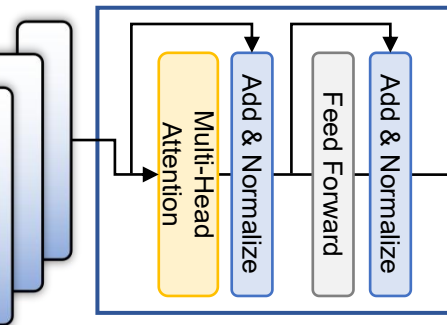


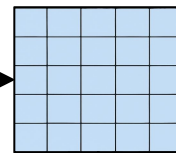
SMILES dataset



Chemberta



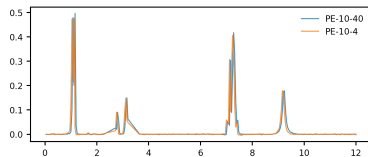
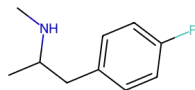
SMILES
representation



SMILES:
CC(CC1=CC=CC(C=C1)F)NC

Example data

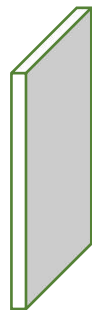
4-FMA



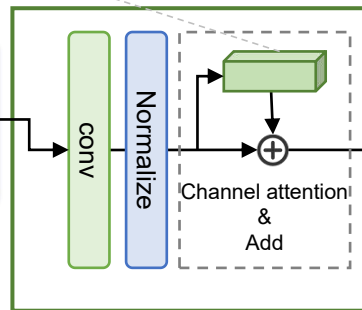
NMR dataset



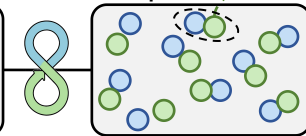
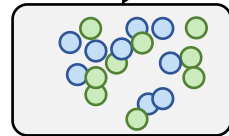
Data
preprocessing



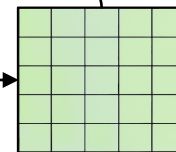
Residual block



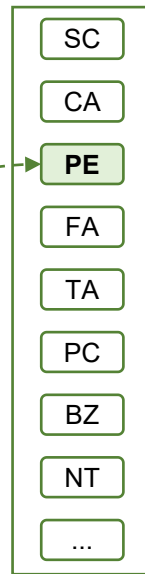
Vector space



Contrastive learning



NMR
representation



NPS labels