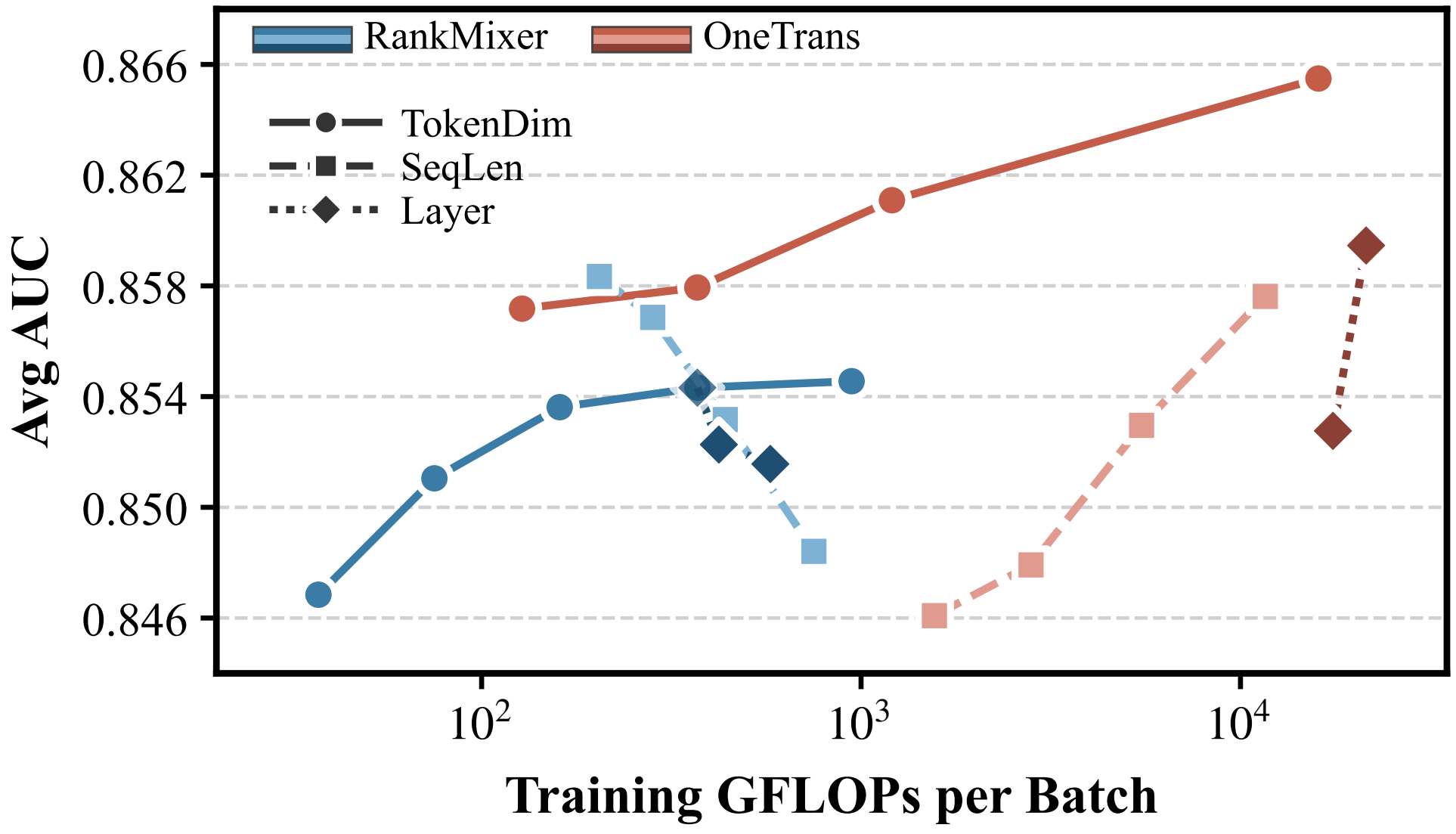

RankMixer
OneTrans
TokenDim
SeqLen
Layer
0.866
0.862
0.858
0.854
0.850
0.846
Avg AUC
10^2
10^3
10^4
Training GFLOPs per Batch