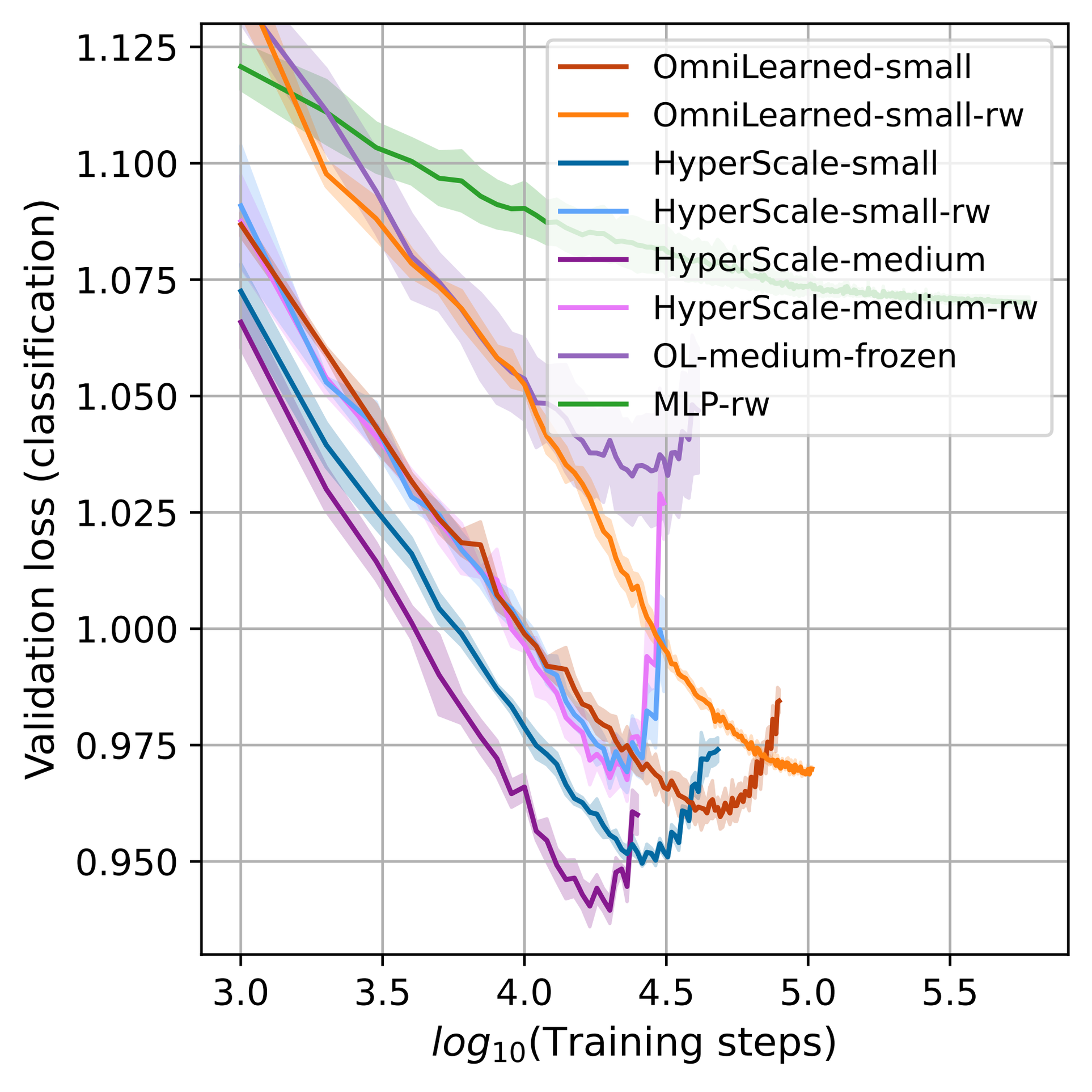

OmniLearned-small
OmniLearned-small-rw
HyperScale-small
HyperScale-small-rw
HyperScale-medium
HyperScale-medium-rw
OL-medium-frozen
MLP-rw
Validation loss (classification)
1.125
1.100
1.075
1.050
1.025
1.000
0.975
0.950
3.0
3.5
4.0
4.5
5.0
5.5
log10(Training steps)