

Pairwise training improves best loss on both datasets

