cos similarity histograms
(albert-xlarge-v2_wikitext_(x,y)_trained-params_dense-off_24-layers)