cos similarity heatmaps
(albert-xlarge-v2_uniform-tokens_(x,y)_random-params_dense-off_24-layers)