cos similarity histogram - layer 7 (albert-xlarge-v2 uniform-tokens (x,y) trained-params dense-off 24-layers 

