cos similarity histogram - layer 6 (albert-xlarge-v2 uniform-tokens (x,y) trained-params dense-on 24-layers 

