cos similarity histogram - layer 18 (albert-xlarge-v2 wikitext (Qx,Ky) trained-params dense-on 24-layers) 

