cosine T-SNE visualisation - layer 0 (albert-xlarge-v2_uniform-tokens_(Qx,Ky)_trained-params_dense-on_24-layer 75 50 25 0 -25-50-75-100 -50 -100-75-50-2525 75 100