cosine T-SNE visualisation - layer 3 ert-large-uncased_uniform-tokens_(Qx,Ky)_random-params_dense-on_24-lay 75 50 25 0 -25-50-75-100 --50 50 -100-75-2525 75 100