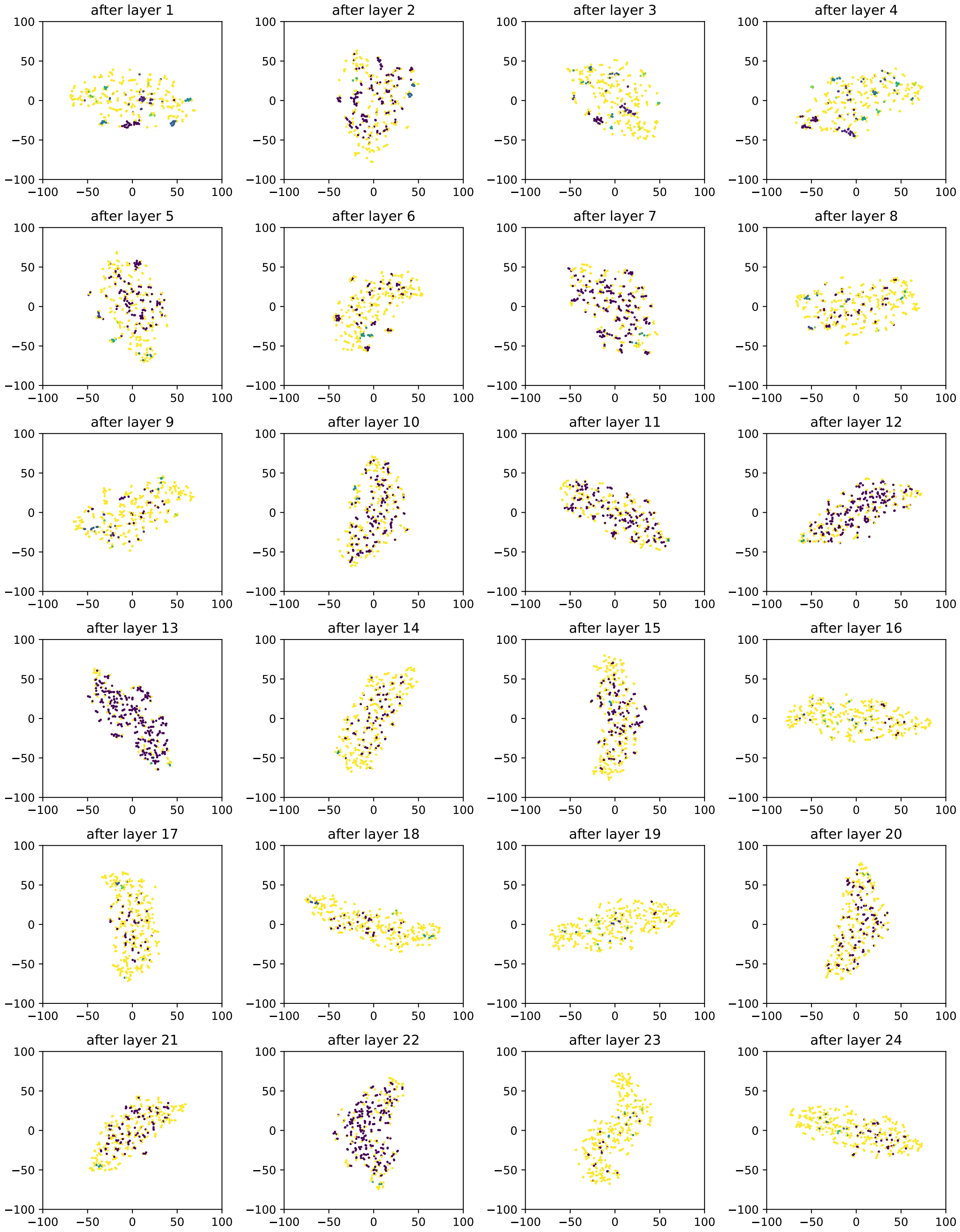attention_logit t-SNE embeddings
(wikitext_bert-large-uncased_params-random_dense-off_layers-24)