

Image-to-Text Recall@1 per Layer, Mask After Attention Softmax

