

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

