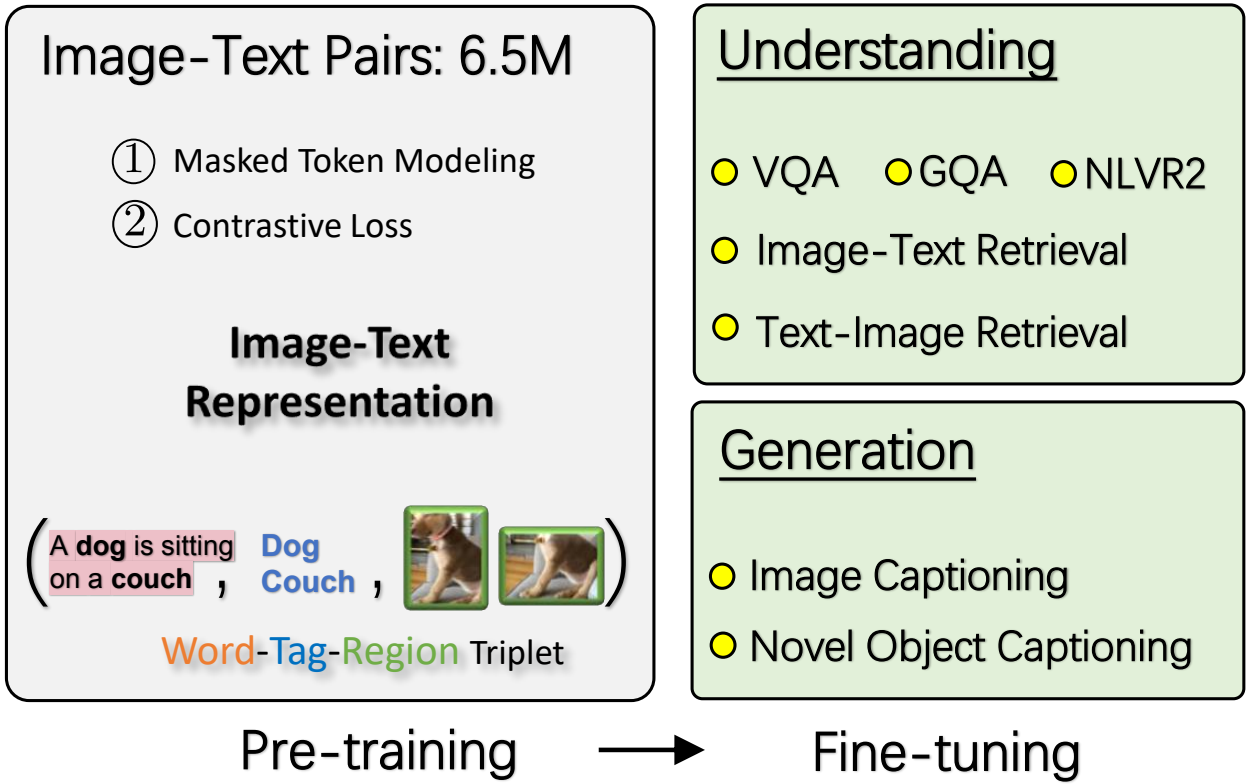

Fig. 1. OSCAR pipeline. The model takes a triplet as input sequence, is pre-trained with two objectives (a masked token loss over words & tags, and a contrastive loss between tags and others), and fine-tuned for 5 understanding and 2 generation tasks (detailed in Section 4).