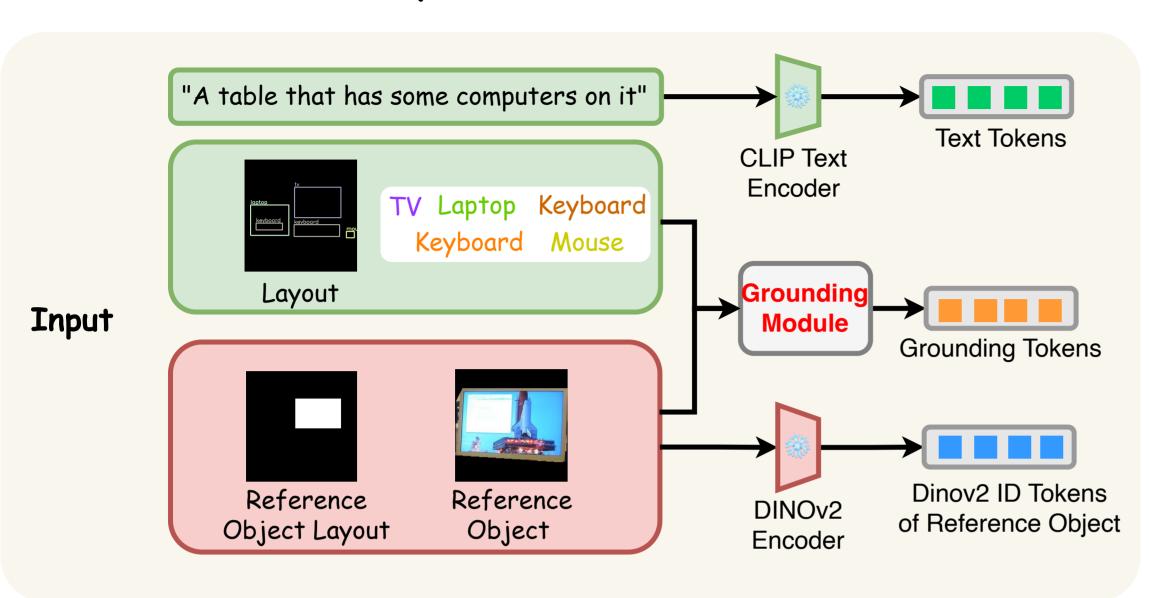
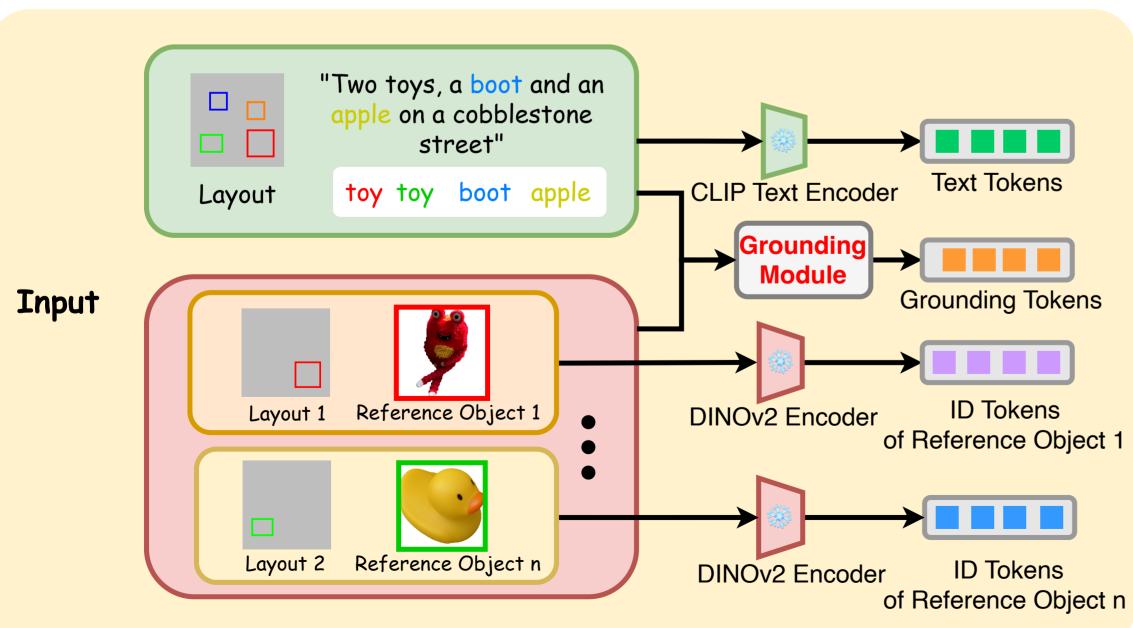
Step 1: Feature Extraction





Train



Step 2: Foreground-Background Cross-Attention Control

