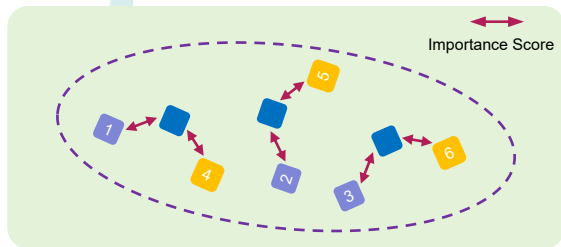




threshold



token importance rank and random mask %n  $(\text{rank} < \text{threshold})$



Visual Tokens

  
Pooled Text Tokens

Text Encoder

Prompt