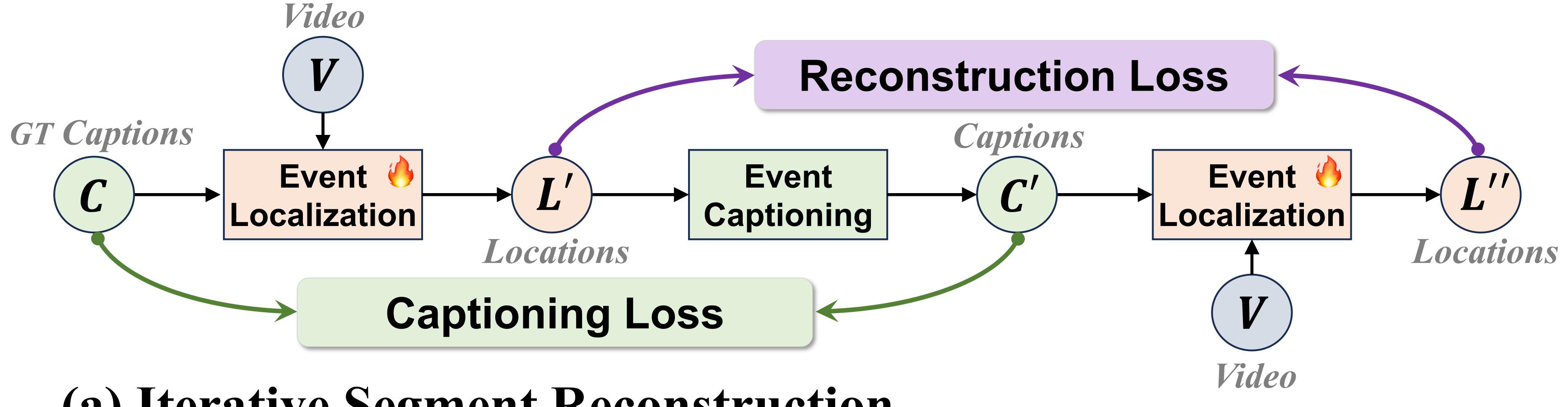
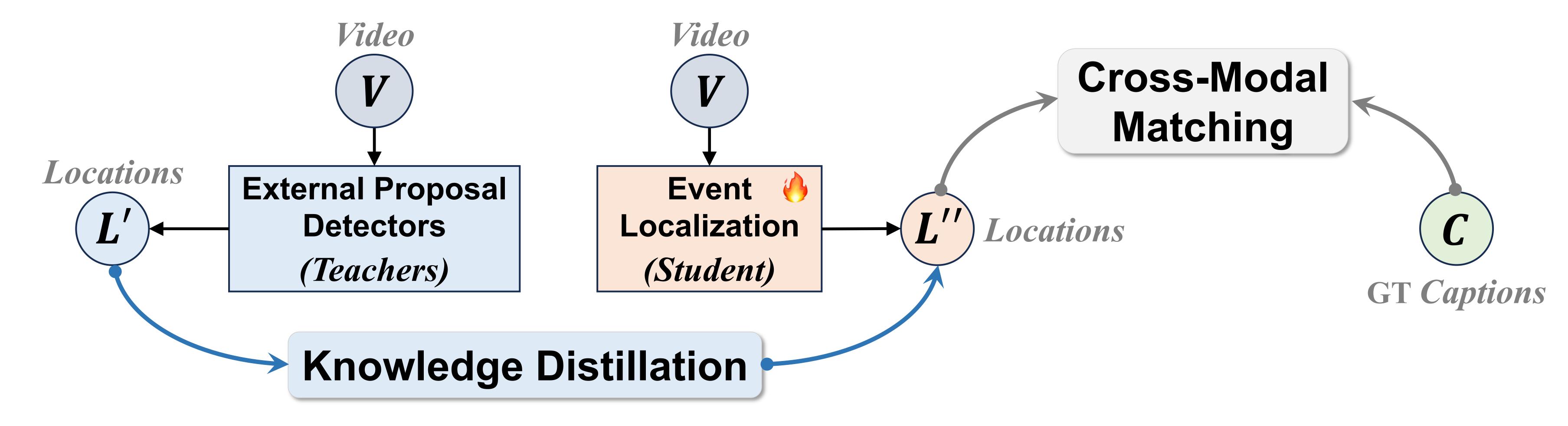


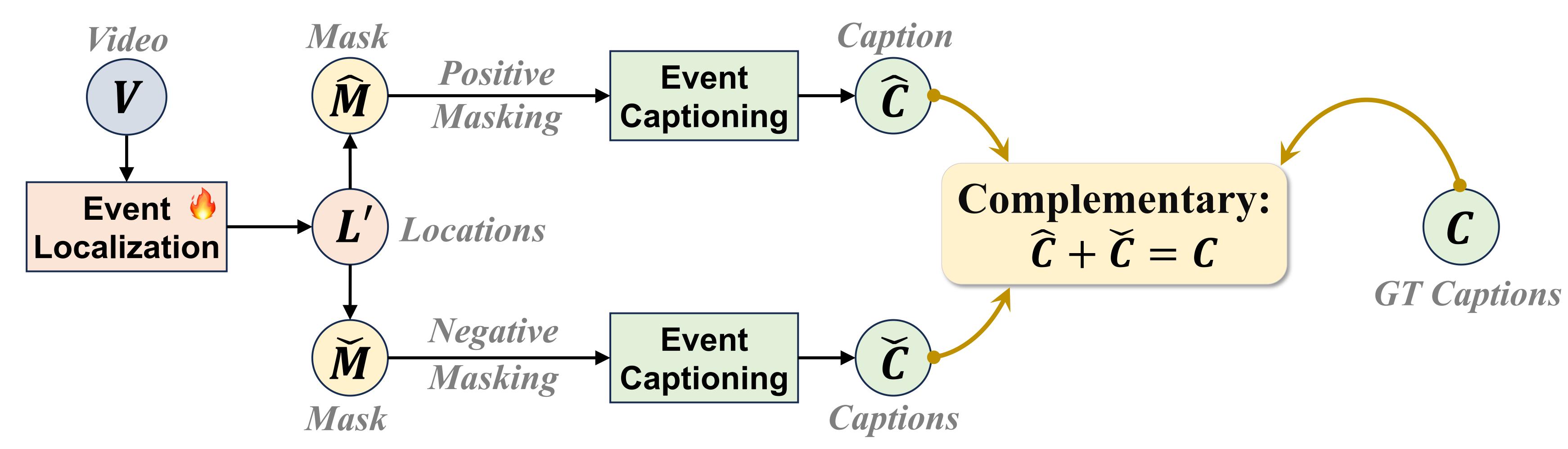
Event Localization via Differentiable Complementary Masking



(a) Iterative Segment Reconstruction



(b) Distillation and Matching



(c) Complementary Masked Captioning (Ours)