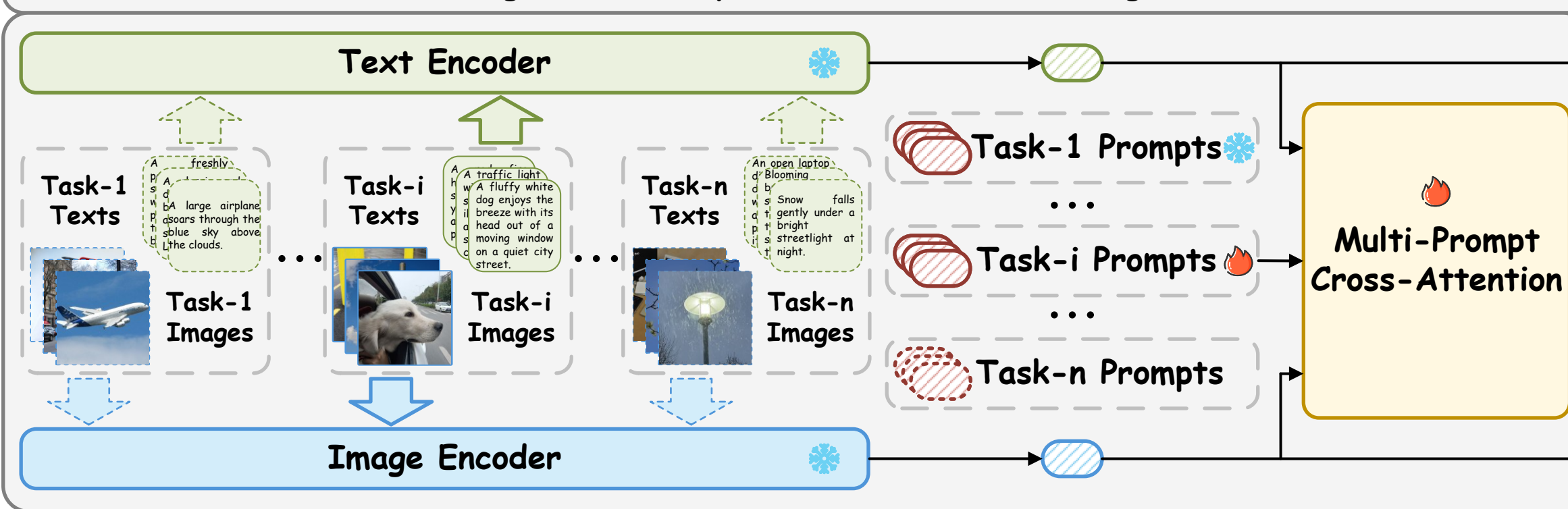
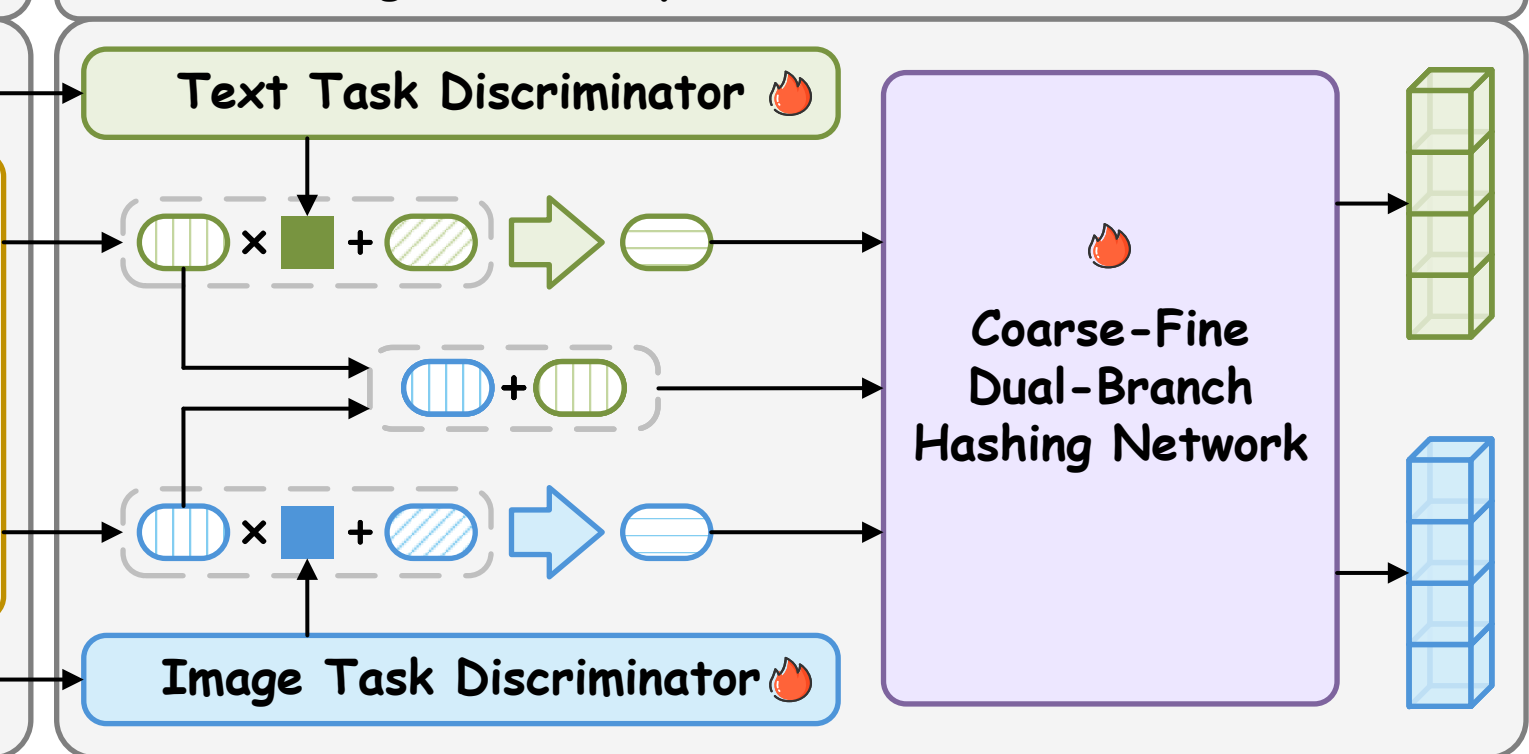


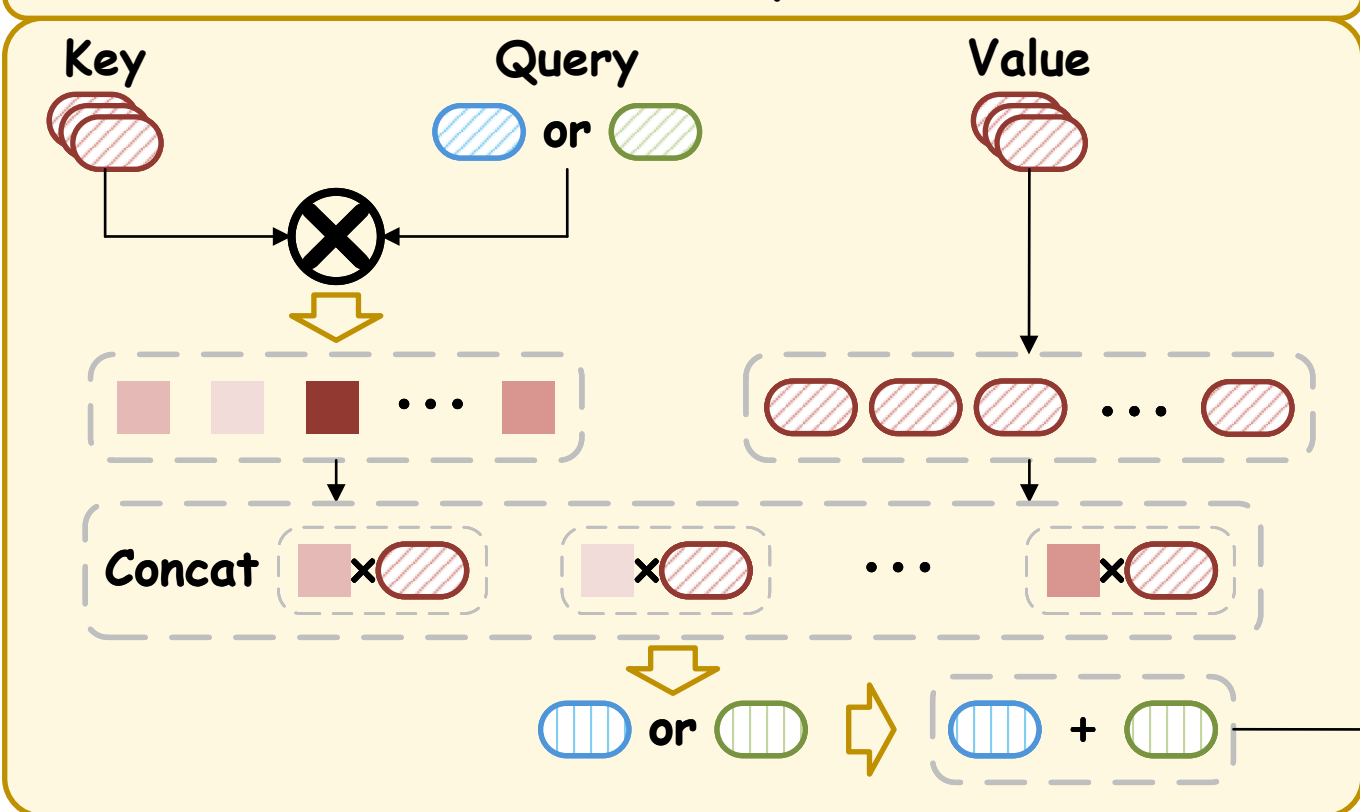
Stage I: Task-Specific Multimodal Learning



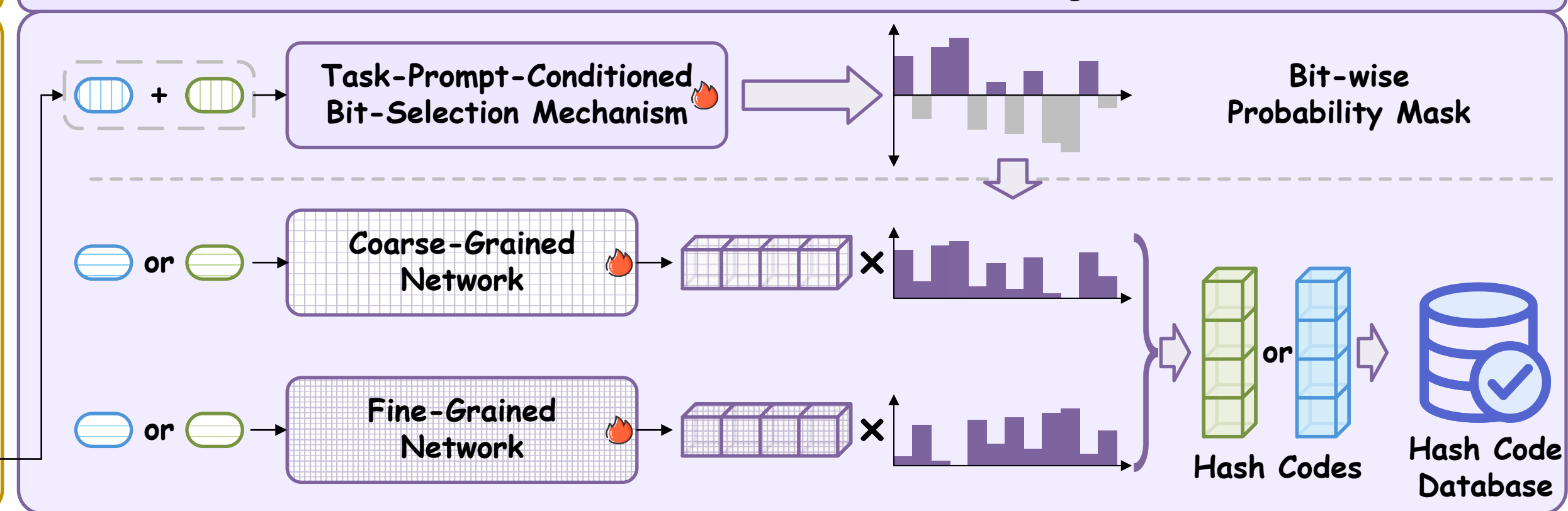
Stage II: Prompt-Masked Bit Localization



Detailed View: Multi-Prompt Cross-Attention



Detailed View: Coarse-Fine Dual-Branch Hashing Network



❄️ Frozen
 🔥 Learnable
 Base Features
 Prompt-Aggregated Features
 Enhanced Features
 Prompts
 Coarse/Fine Hash Code Candidates
 Hash Codes