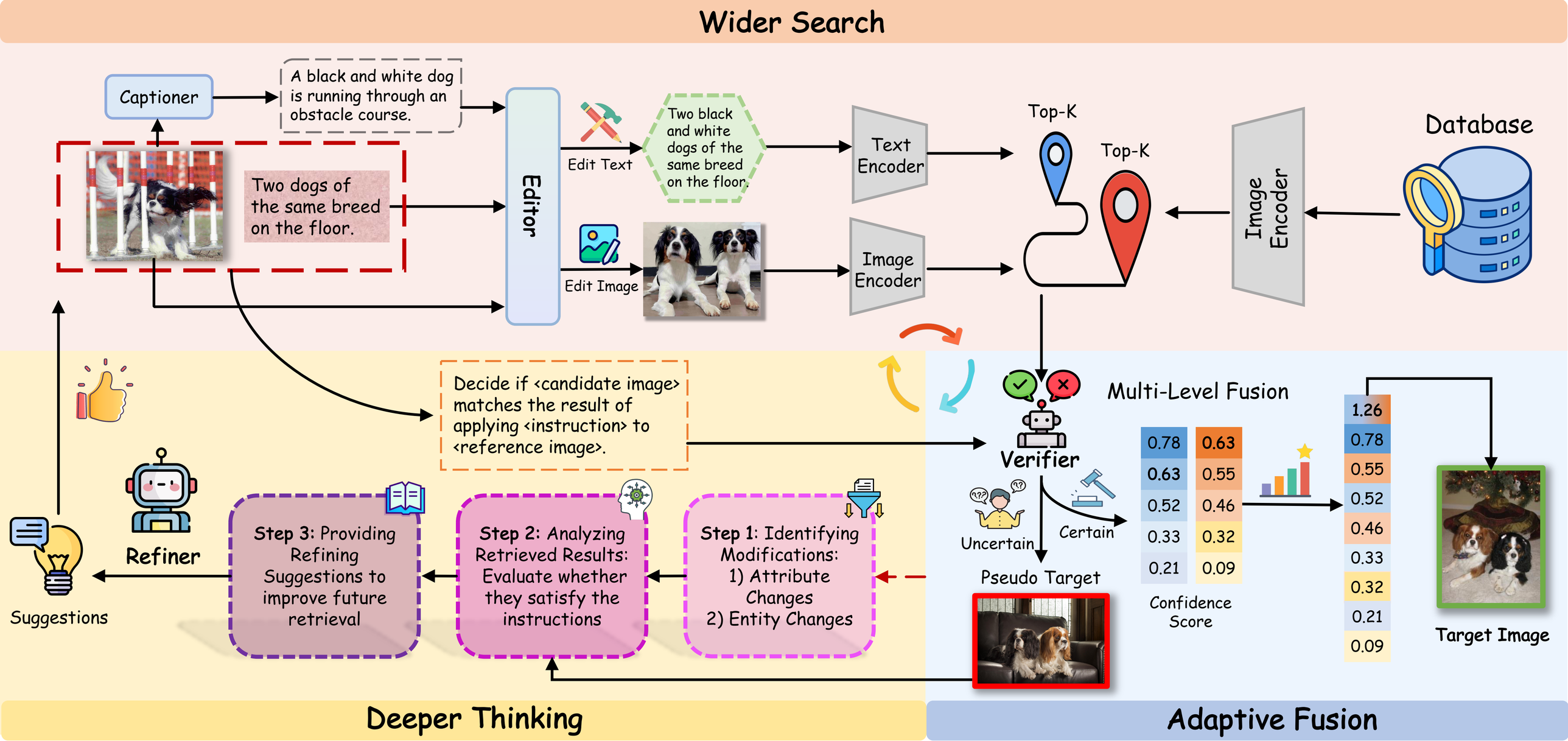
Wider Search
Captioner
A black and white dog is running through an obstacle course.
Two dogs of the same breed on the floor.
Editor
Edit Text
Two black and white dogs of the same breed on the floor.
Edit Image
Text Encoder
Image Encoder
Top-K
Top-K
Image Encoder
Database
Decide if <candidate image> matches the result of applying <instruction> to <reference image>.
Verifier
Uncertain
Certain
Pseudo Target
Multi-Level Fusion
0.78
0.63
0.63
0.55
0.52
0.46
0.33
0.32
0.21
0.09
Confidence Score
1.26
0.78
0.55
0.52
0.46
0.33
0.32
0.21
0.09
Target Image
Step 1: Identifying Modifications: 1) Attribute Changes 2) Entity Changes
Step 2: Analyzing Retrieved Results: Evaluate whether they satisfy the instructions
Step 3: Providing Refining Suggestions to improve future retrieval
Refiner
Suggestions
Deeper Thinking
Adaptive Fusion