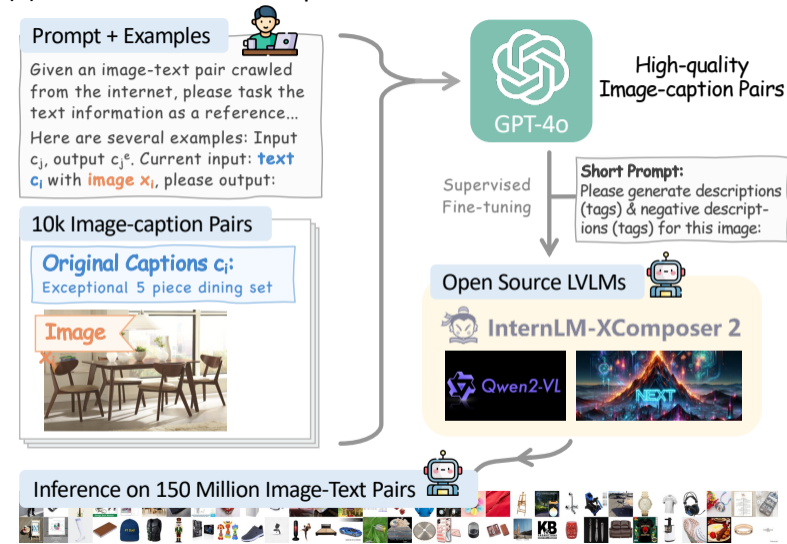


(a) Dataset Refinement Pipeline



(b) HQ-CLIP Training Framework

Original Captions:

Seaside on the island of Noirmoutiers in France

Description $\{d^+\}$:

The image shows a **beach** on Noirmoutier Island in France, where **a woman and a child** are **walking** barefoot on the **sandy shore**, the scene is captured on a **sunny day** with a clear blue sky and a few clouds...

Negative Description $\{d^-\}$:

The image shows a **mountainous area** in Switzerland. **A man and a child** are **hiking** through a dense **forest**. The scene is captured on a **rainy day** with overcast skies...

Negative Tags $\{t_i^-\}$:

mountains, forest, hiking, father and child, rainy day...

Tags $\{t_i^+\}$:

beach, seaside, mother and child, sunny day, seascape...

