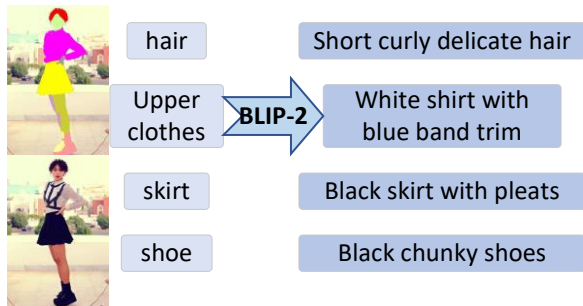


### Step: 01 ATR Phrase Generation

"Question: What kind of hair is she? Answer: It is"



### Step: 03 TBPR Sentence Slicing

The woman has shoulder length dark, straight hair and is wearing a short sleeve chambray blouse and light cropped pants. She is carrying a large, beige shoulder bag

LLMs

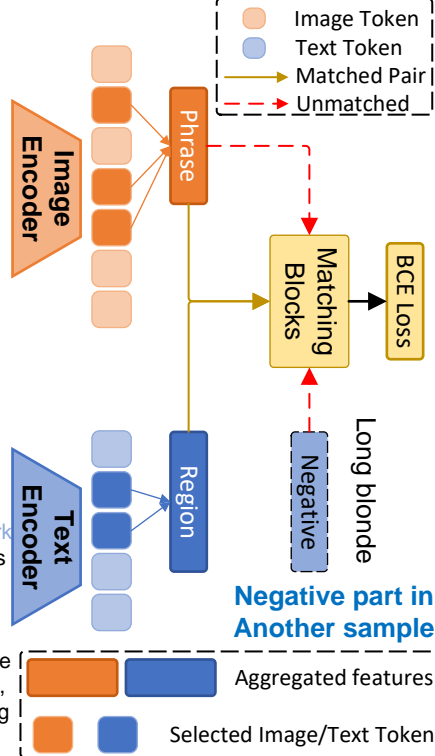
shoulder length dark, straight hair

a short sleeve chambray blouse

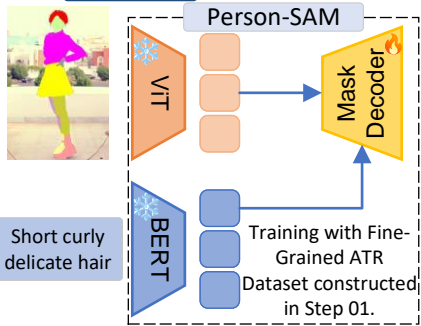
light cropped pants

a large, beige shoulder bag

### Step: 05 Fine-Grained Semantic Alignment

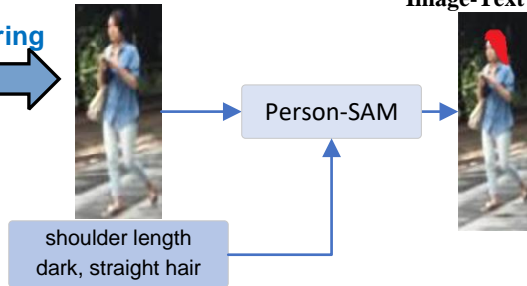


### Step: 02 Person-SAM Training



### Step: 04 Person-SAM Transferring

Transferring



Fine-grained Image-Text Pairs

Used

The woman has shoulder length dark straight hair and is wearing a short sleeve chambray blouse and light cropped pants. She is carrying a large, beige shoulder bag