

[Domain-**invariant** prompts]

Text  
Encoder

positive

pull

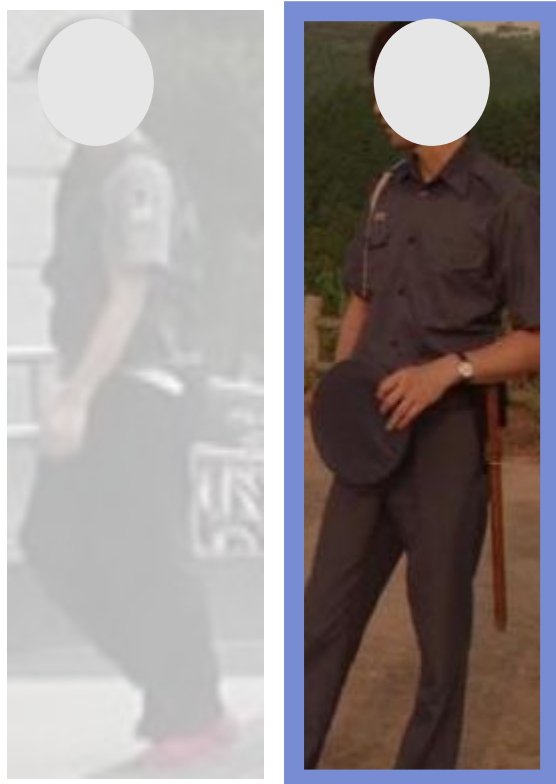


Image  
Encoder

anchor

push

[Domain-**relevant** prompts]

Text  
Encoder

negative

fine tuning

Bidirectional Guiding Method