Step1: Extract IDs and descriptions for each frame
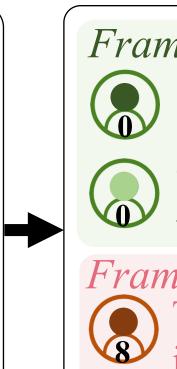
Step2: Extract ID-Matching sequence

Video frames

Video Caption

The scene begins with the man walking against the flow of traffic on a road … As the man continues walking along … The situation escalates when the white car moves … The scene concludes with the man walking away…

Frame 0
- (0) The man walks against the flow …
- (0) The driver isdriving the white…

Frame 4
- (4) The man continues walking along …
- (4) The motorcyclist almost hit a man…

Frame 8
- (8) The man go straight into the flow…
- (8) The pedestrian is hit and seems …

Frame 20
- (20) The man walks away with …
- (20) The driver stops the car and …

ID 1: (0) (4) (8) (20)
ID 2: (0)
... (8)
... (4)
ID n: (20)