Action Speed
Event Order Direction
Attribute Change

People Buildings Vehicles
Animals Artifacts Abstracts
Natural Objects Food Plants



Meta Information

Subject: ice cream

Attribute Change: melting

Multi-Choice QA

What is happening to the ice cream?

A. melting B. freezing

C. evaporating D. Solidifying

Answer: A. melting

Yes/No QA

Is the ice cream freezing?

Answer: no

Is the ice cream melting?

Answer: yes

Caption Matching

Which caption matches the video better? Caption A: The ice cream is melting.

Caption B: The ice cream is freezing.

Answer: Caption A: The ice cream is melting.

Caption Generation

Identify the information that consistent with the video and generate a video caption.

Information A: {'subject': 'ice cream', 'attribute_change': 'melting'}
Information B: {'subject': 'ice cream', 'attribute_change': 'freezing'}
Information C: {'subject': 'ice cream', 'attribute_change': 'evaporation'}

Information C: {'subject': 'ice cream', 'attribute_change': 'evaporating'}
Information D: {'subject': 'ice cream', 'attribute change': 'solidifying'}

Answer: The video shows an ice cream is melting.

- (1) Temporal Aspect & Static Content Selection
- (2) Video Collection
- (3) Meta-Information Collection
- (4) Instruction Collection

(4)