

KnowIT VQA: Answering Knowledge-Based Questions about Videos

Video QA aims to answer **temporal** questions about videos.



Q: How do they finish the conversation?

A: Shaking hands.

Knowledge-Based VQA answers questions requiring **external knowledge**.



Q: How much do these animals need to drink each day?

A: 50 gallon.

How about answering **knowledge-based questions about videos?**

KnowIT VQA: Answering Knowledge-Based Questions about Videos

Contributions:

1. **KnowIT VQA dataset:**

With 4 types of questions and knowledge annotations.

2. **ROCK model:**

Leverages external knowledge and video content to answer questions about videos.



Penny: What are you doing at work these days?

Sheldon: Oh. I'm working on time-dependent backgrounds in string theory. Specifically, quantum field theory in D-dimensional de Sitter space. in D-dimensional de Sitter space.

What night is it?

Wednesday

Monday

Friday

Saturday 

Retrieved Knowledge

Saturday is Sheldon's laundry night.

**Come to our poster!!
Paper ID 1192**

Dataset download

