



Does the person cook any meals throughout the day?

Long ADL Video (Up to 1 or more hours)

Existing Approaches



Long ADL Video



User Query



Cloud Vision
Language Model

↑ High Cost \$\$\$

↑ High Latency

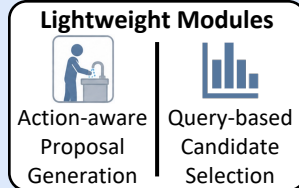
TIMEPROVE (Ours)



Long ADL Video



User Query



Lightweight Modules



Action-aware
Proposal
Generation



Query-based
Candidate
Selection

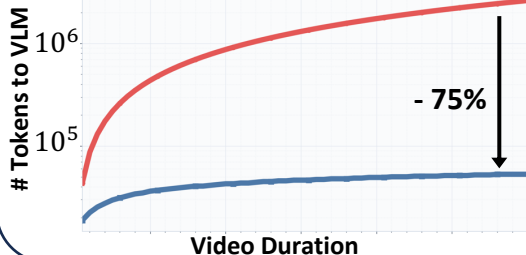


Cloud
VLM

↓ Low Cost \$

↓ Low Latency

— Baseline — TIMEPROVE (Ours)



— Baseline — TIMEPROVE (Ours)

