Slow-Fast Weights for Long-Term Video Understanding

Mentors: Rajat Modi, Dr. Rawat (University of Central Florida)
REU Student: David Gardner (Boston University)

July 30, 2025



Project Overview

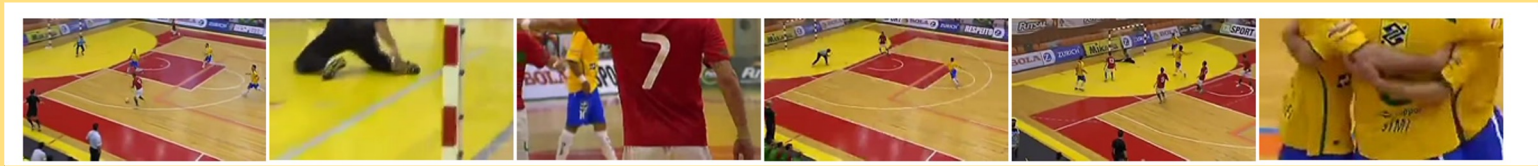
Motivation:

- Many long video applications (CCTV, sports)

Limitations:

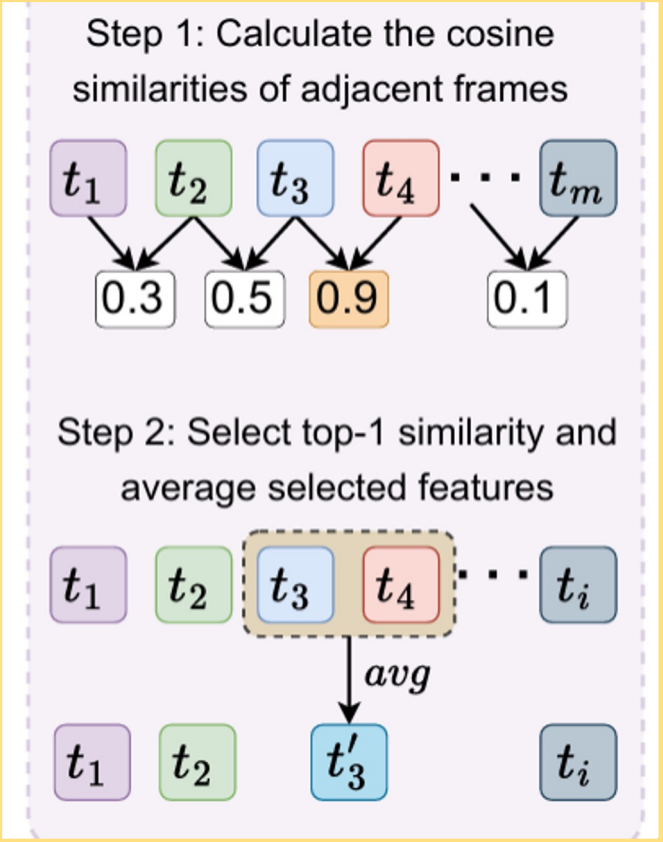
- High computational cost
- Memory bank compression loss



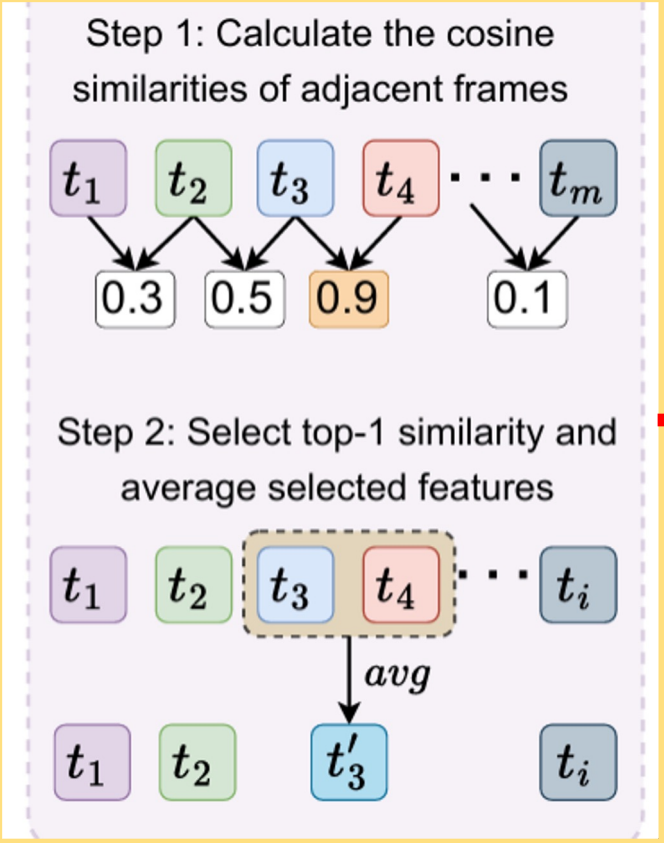


Question: What color is man with the number 7 in the video?

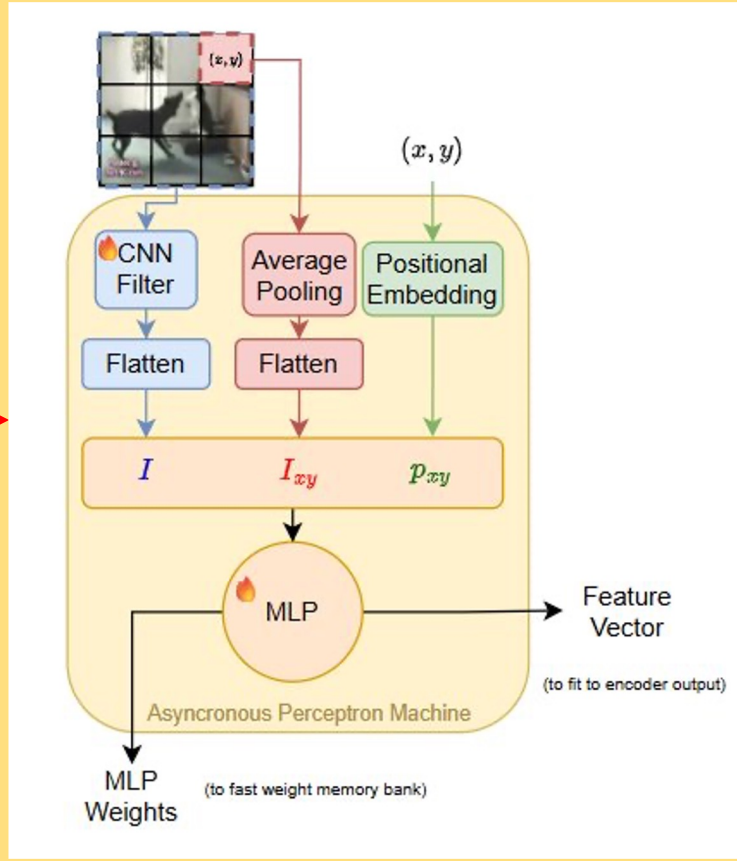
Answer: Red



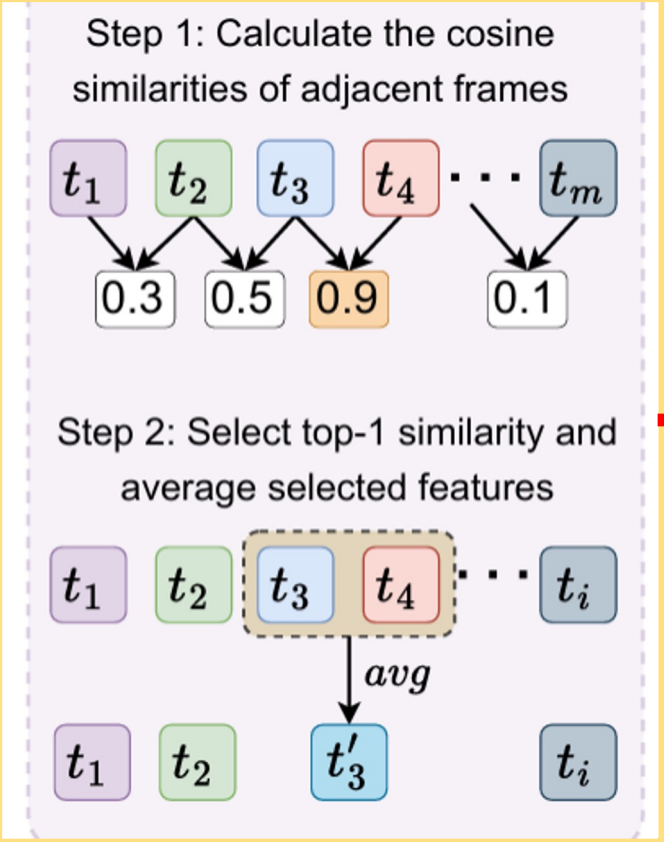
Existing Memory Compression



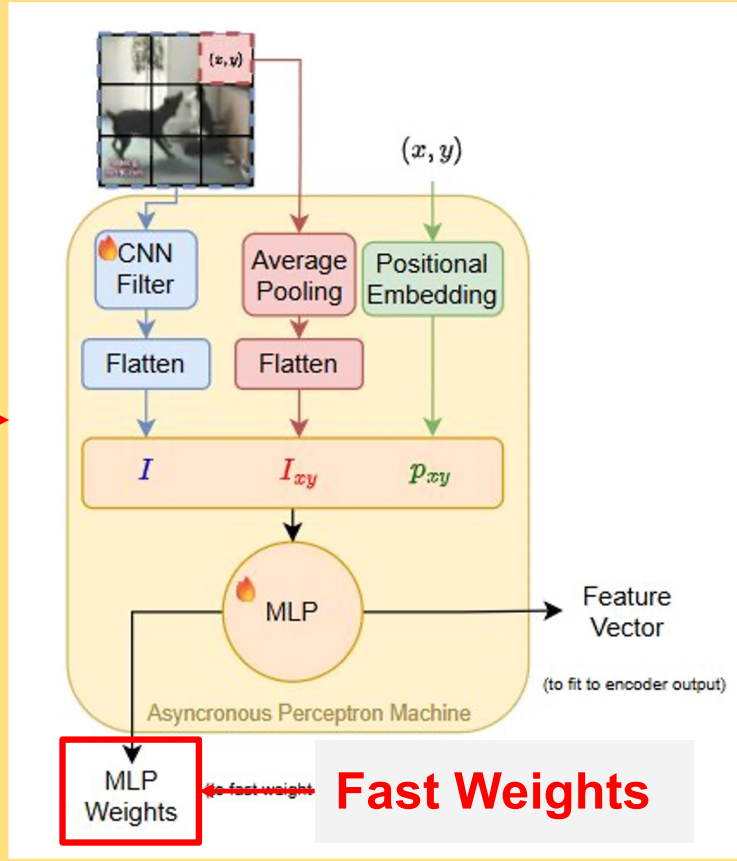
Existing Memory Compression



Our Memory Compression



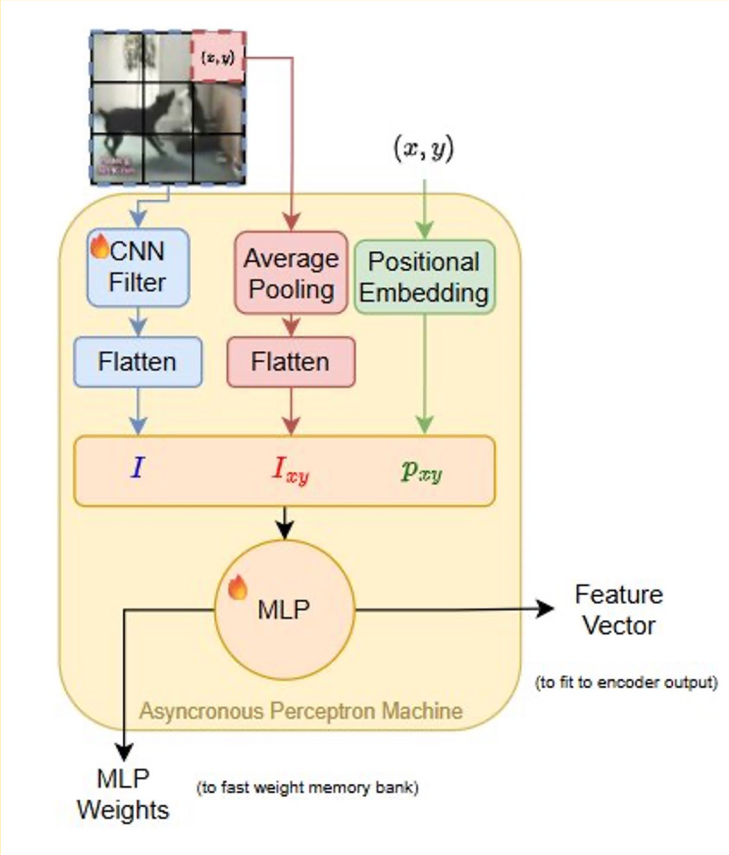
Existing Memory Compression



Our Memory Compression

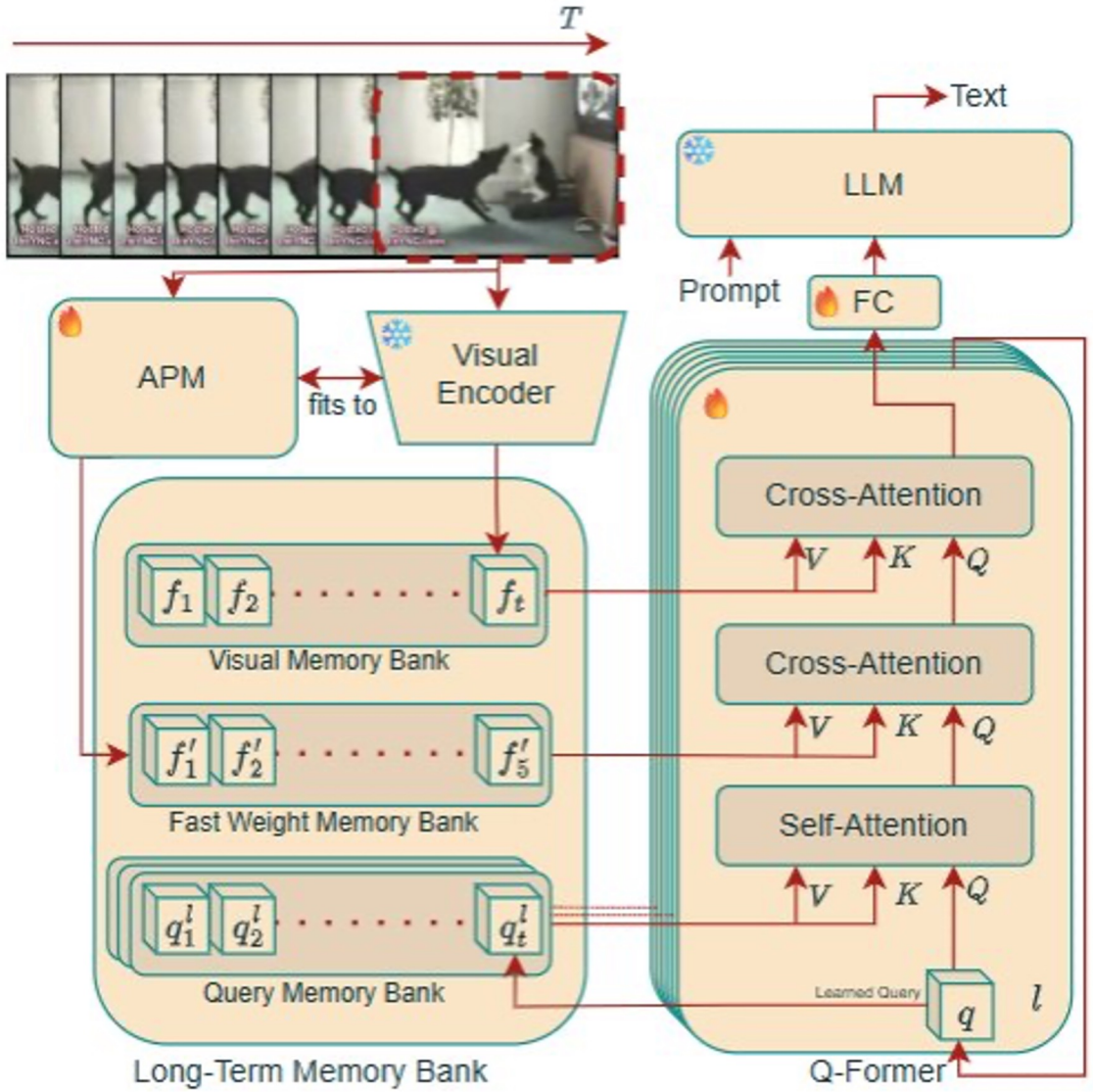
Why weights?

- Associative memory (pattern storage)
- Similar to brain
- Large capacity (n vs n^2)

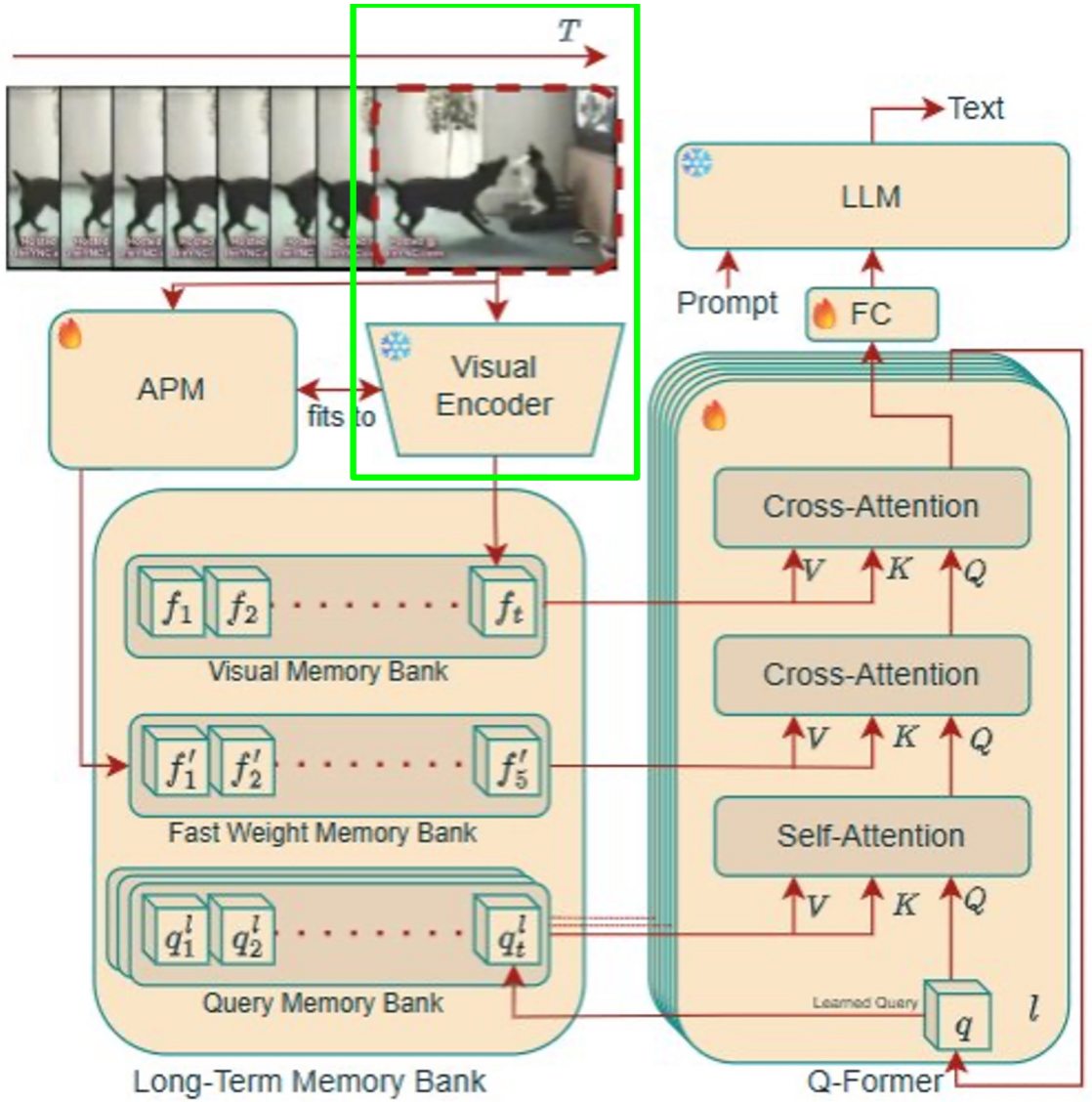


Our Memory Compression

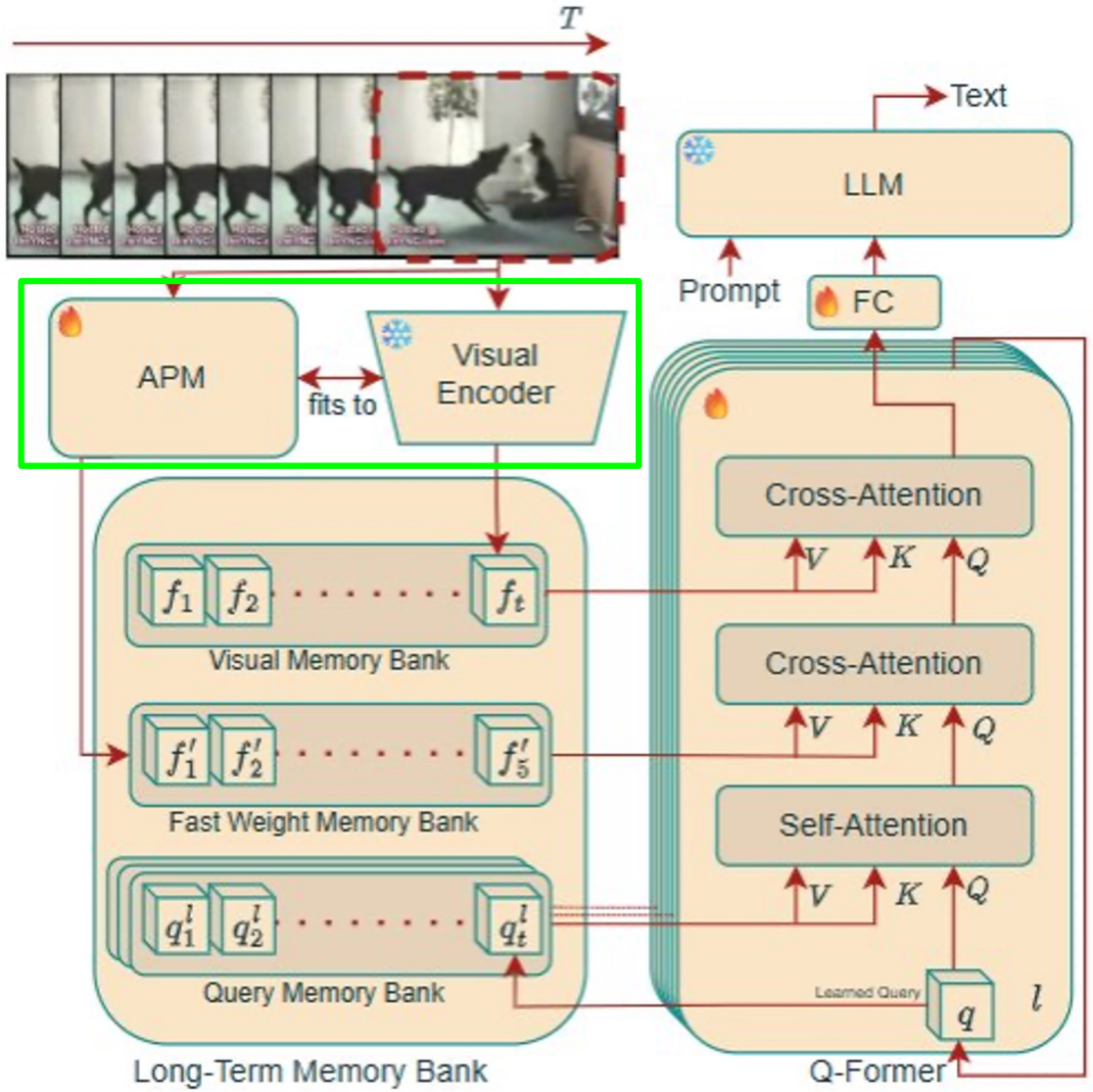
Our Architecture Flow



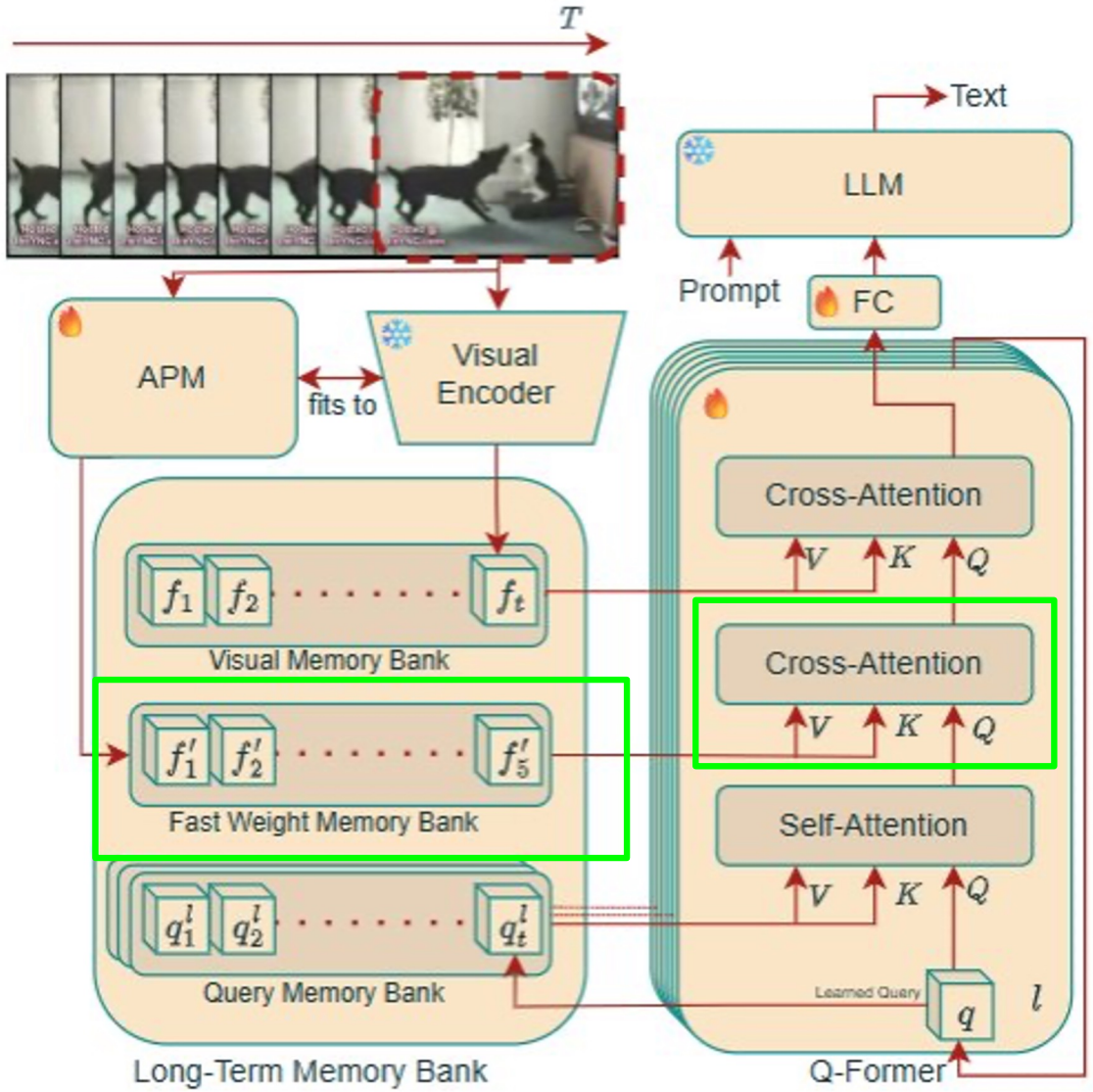
Our Architecture Flow



Our Architecture Flow



Our Architecture Flow



Results



Question: how many boys splashed in the little pool?

❌ **Ours:** 2

✅ **MA-LMM:** 3



Question: what did the officer pull a gun on?

✅ **Ours:** car

❌ **MA-LMM:** motorcyclist

Future Work

- Modify architecture to best interpret memory bank weights
- Experiment on longer video datasets

Thank You