Activity prediction pipeline

One batch : sequence of 10 frames

- ❄️ Image encoder
- ❄️ VLM
- ❄️ VLM

Image feature

Action anotations

❄️ Object detector

Activity annotations

Graph constructor ← Bounding box features

Action scene graph 0 Action scene graph 1 ... Action scene graph 9

Sequence of 10 graphs

Graph attention encoder

Activity prediction → Cross Entropy loss