



Prior work:

Single clip step prediction

Ours: Masked step prediction over all clips in video