

Short tubelet results in each window (e.g., 7 frames)

pick-and-roll
defense

dribble

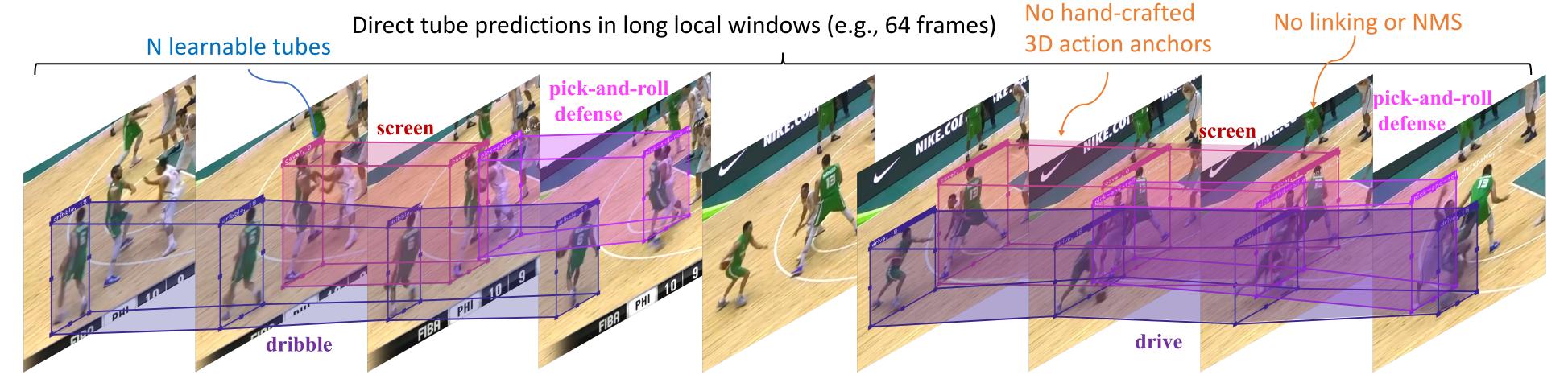
Linking Linking Linking

(sit, hold, eat)

Key frame

(b) Previous tubelet-level detectors in short local windows.

(c) AVA-style frame-level action detectors



(d) Our end-to-end tube-level action detector