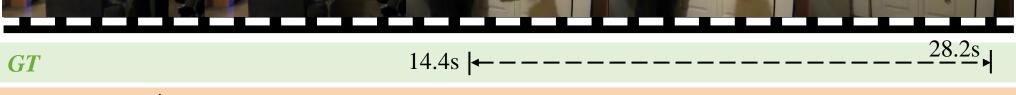
Query: person holds up a broom.





GT *VLSNet** 2.5s ← - - - - -**-**►|11.5s

14.9s BSSARD-VLSNet*