

提示

注释

标签

a video representing the {ComeHere} gesture,
a type of {DivingSignals1}

"DivingSignals1/ComeHere"

$i: 129$

语义分支

CLIP
EncoderText
AdapterFilter
Learner

SF

 $filter_s^i$ FC_s L_s^i Video
Encoder

视觉分支

 f_v $filter_v$ FC_v L_v

SLS

Prompts

CLIP
Encoder f_s S