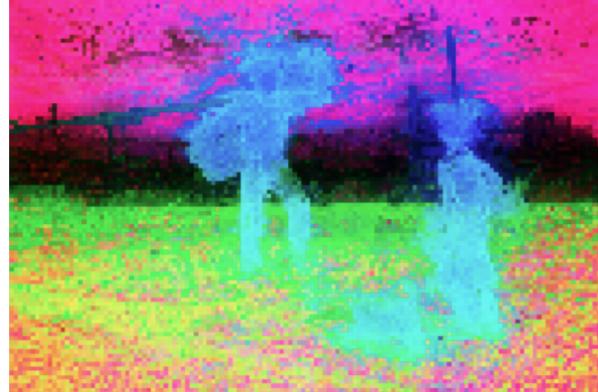


Early-fusion Grounding VLM

PCA maps: SigLIP2



DINOv3



Ours



Bounding box



Segmentation mask

