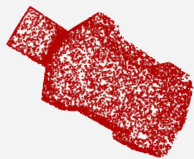
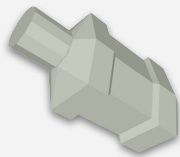


Object  
Point Cloud image



Object Render



New Hand Point Cloud



"Grasp the  
bottle around  
its body using  
a wrap grasp"

Pointnet++

Dino-v2

Hand Parameters

Mesh bashing

New Params

Tokenizer

Attention Pooling

Pointnet++

$\mathcal{X}(T, R, \theta)$

$t$

Learned  
Query

FLASH (Qwen 2.5-0.5B)

$\hat{v}$

