Geometry-Driven Diffusion Model

Rendered condition N-Encoder Encode conditions Use pretrained LLM **Text Embedding** to extract text prompt "A hand is holding dispenser on a table **LLaVA** covered with green cloth"

Appearance Regularization

