Rendering synthetic data (e.g., 3D CAD-rendered images) to generate annotations for learning deep models in vision tasks has attracted increasing attention in