Despite significant recent progress on generative models, controlled generation of images depicting multiple and complex object layouts is still a difficult problem