

In this paper, we focus on generating realistic images from text descriptions. Current methods first generate an initial image with rough shape and color, and the