

Generating an image from a given text description has two goals: visual realism and semantic consistency. Although significant progress has been made in gen