

Text Prompt for
Realistic Scenes

Fine-tuned
Diffusion Model



Monocular
Depth Estimator

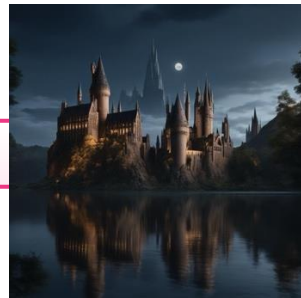


Monocular
Depth Estimator



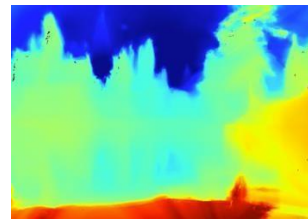
Text Prompt for
Unrealistic Scenes

Pre-trained
Diffusion Model



Pre-trained Image
Diffusion Model

Monocular
Depth Estimator



Align Scenes

(a) Fine-tuned Image Diffusion Model

(b) Pre-trained Image Diffusion Model