Monocular depth estimation model trained on synthetic data produces sharp and plausible depth when applied to real-world images transformed to the style of synthetic images.

**Motivation:**

Synthetic images captured from a graphically-rendered virtual environment primarily designed for gaming can be employed to train a monocular depth estimation model. However, this will not generalize well to real-world images as the supervised model easily overfits to local features present within the training domain.

**Proposed Approach:**

1) train a primary model to estimate monocular depth based on synthetic images. 2) use a secondary model to transform real-world images to the synthetic style before their depth is estimated.

**Run-time:** two forward passes required during inference – once through the style transfer network and once through the depth estimation model.

**Results:**

Our approach produces superior qualitative (sharper) and quantitative (lower error) results compared to the contemporary state-of-the-art.