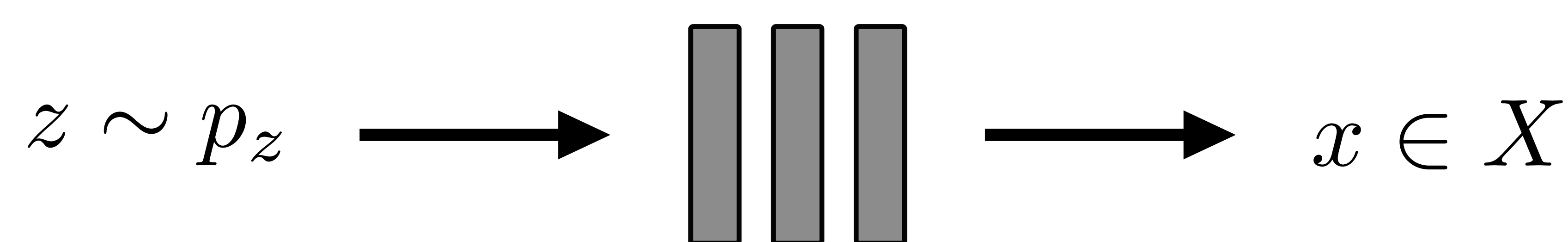


Motivation

State-of-the-art GANs achieve impressive **photorealistic image synthesis**

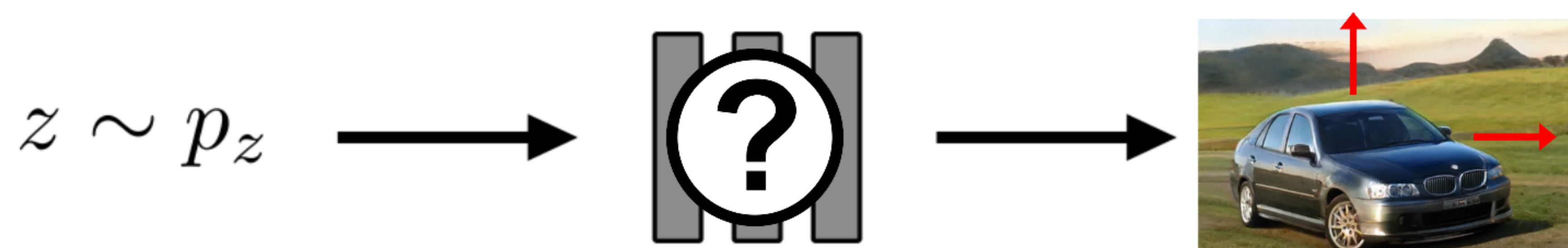
At test time, we can sample new latent codes and get new outputs in the training domain:



For example, when we train on images of cars, we can generate new images of cars:



But how can we manipulate the scene?

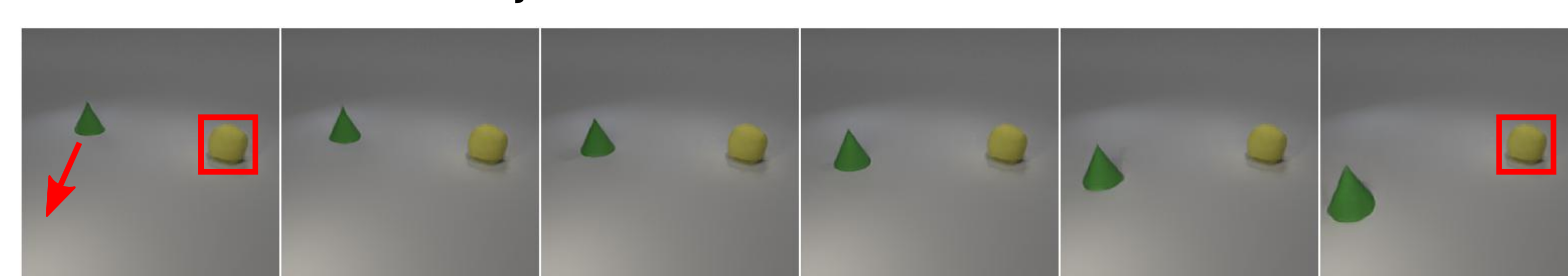
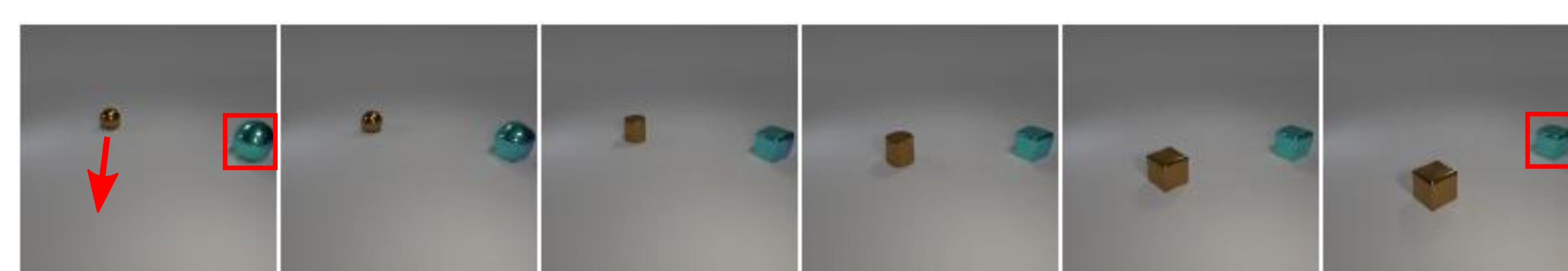


The Challenge

Most state-of-the-art GANs operate in the **two-dimensional image domain**

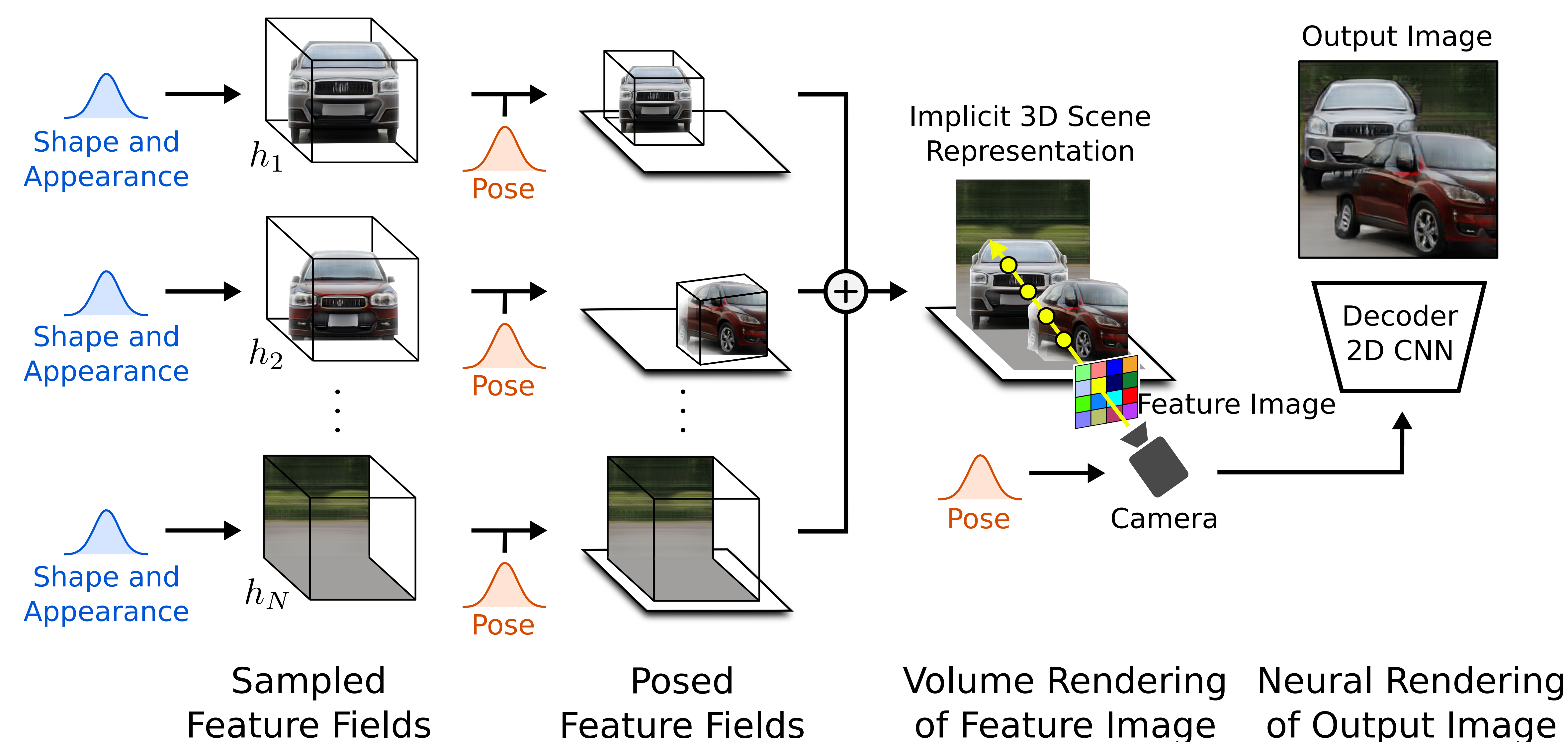
However, our world is **three-dimensional**

Disentangling factors of variation very challenging without **reasoning in 3D**:



Our Method

Our key idea is to incorporate a **compositional 3D scene representation** into the generator model:

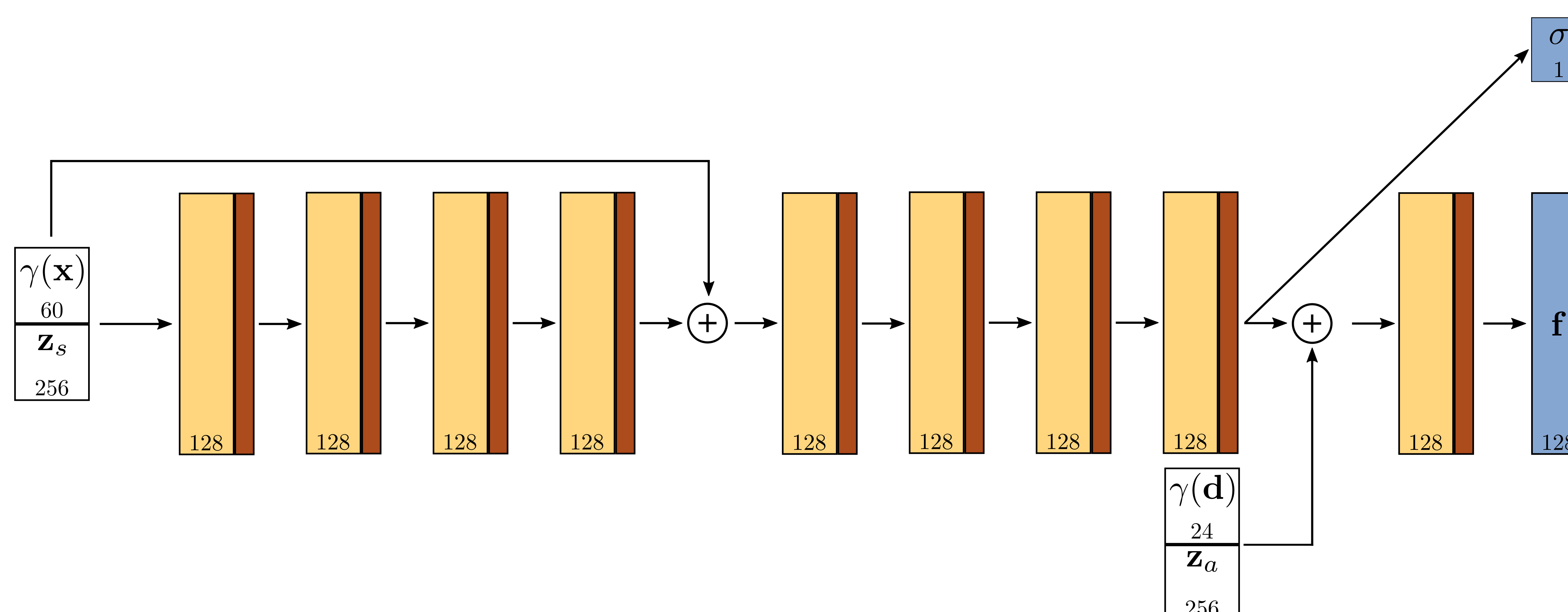


This architecture allows us to

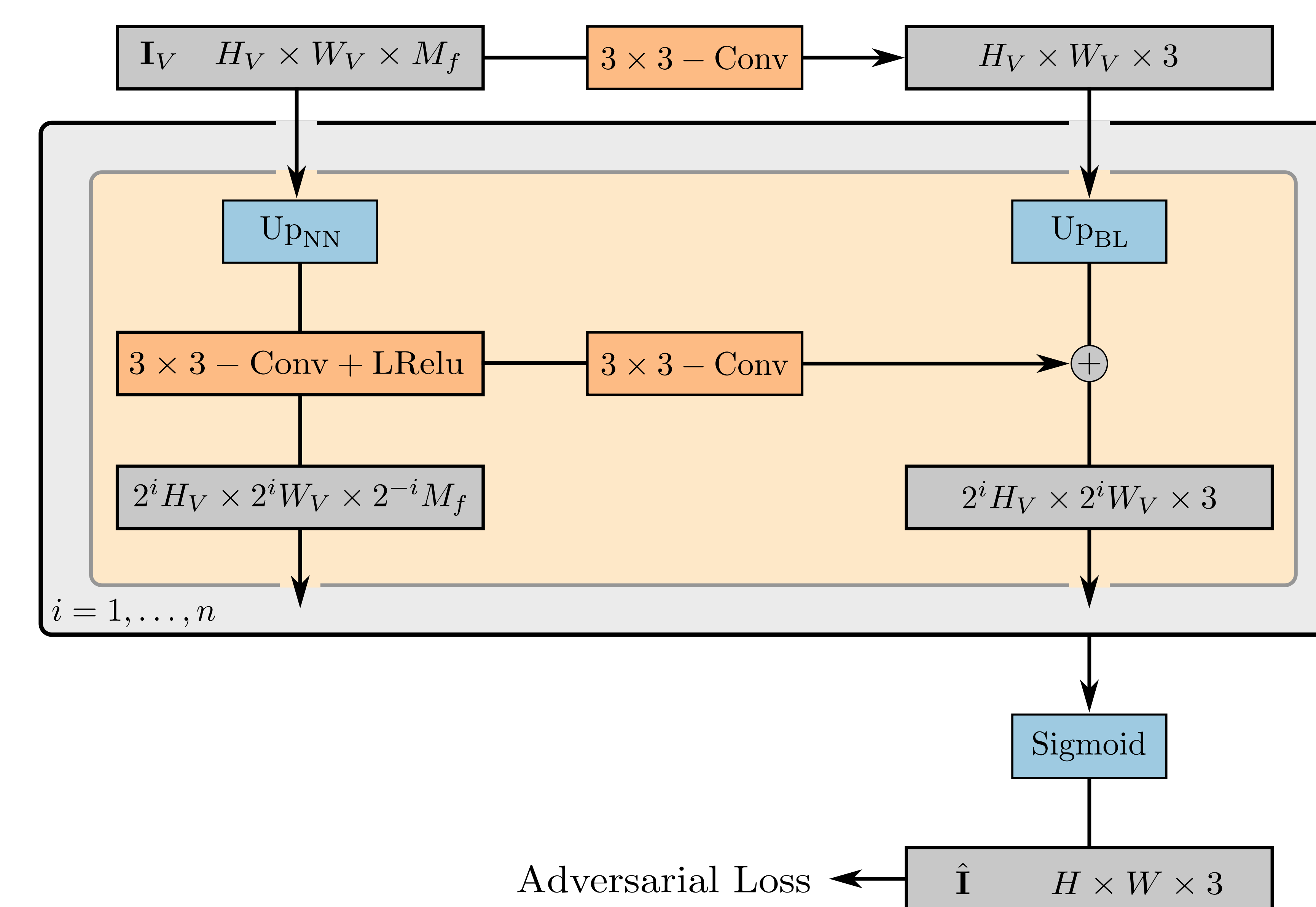
- + Train from **raw, unposed image collections**
- + Model objects in **3D**
- + **Disentangle** objects and the background

At test time, we can generate new images of scenes with **explicit control over the object and camera poses**

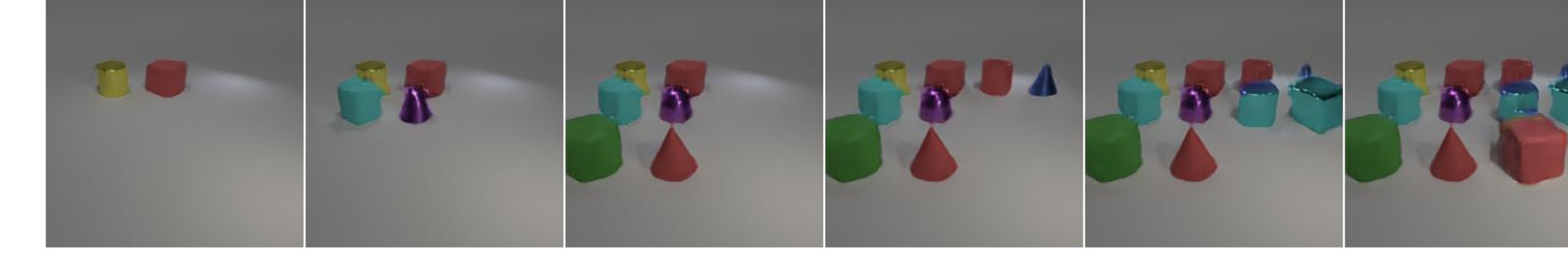
Feature Field Architecture



Neural Renderer Architecture



Controllable Scene Generation



Add more Objects at Test Time (Trained on One/Two Object Scenes)

