We propose Neural Scene Decoration (NSD) - a novel image generation problem which aims to generate images of decorated indoor scenes, given a background image and a collection of foreground objects to be inserted into the scene.

An ideal NSD system should:
- produce realistic results
- preserve the regions of background image not occluded by objects
- ensure the generated image follows the input layout

Many GAN networks exist for generating images from semantic maps or layouts. However, these methods typically lack background input. The background information is typically synthesized from scratch.

In this work, we propose a novel method for NSD, which:
- Encodes the object layout in a simple format parametrized by each object’s position and size, ensuring easiness for manipulation
- Explicitly accepts a background image as input

Network architecture
The generator employs a lightweight architecture trainable from a single GPU. The discriminator has separate branches for adversarial / object layout losses that can be mixed together.

Object labels - object layouts are encoded as K-channel spatial maps, where K is the no. of object classes in dataset.
- Box labels - each object represented as its bounding box
- Point labels - each object represented as unnormalized Gaussian function

Ground truth labels are extracted from the Structured3D dataset.

An ideal NSD system should:
- produce realistic results
- preserve the regions of background image not occluded by objects
- ensure the generated image follows the input layout

In general, images generated with point labels have objects that are less distorted and contains less artifacts.

Scene generation with default object sizes
Our method also works when using a single default size for all object instances of the same class, allowing scenes be generated by only specifying the object positions.