Instituto Universitario de Investigación en Ingeniería de Aragón
Universidad Zaragoza

We estimate the **volume** of objects **in the wild** from just a **single RGB image**.

We can use it to edit their appearance.

Normal Map Estimation in the Wild

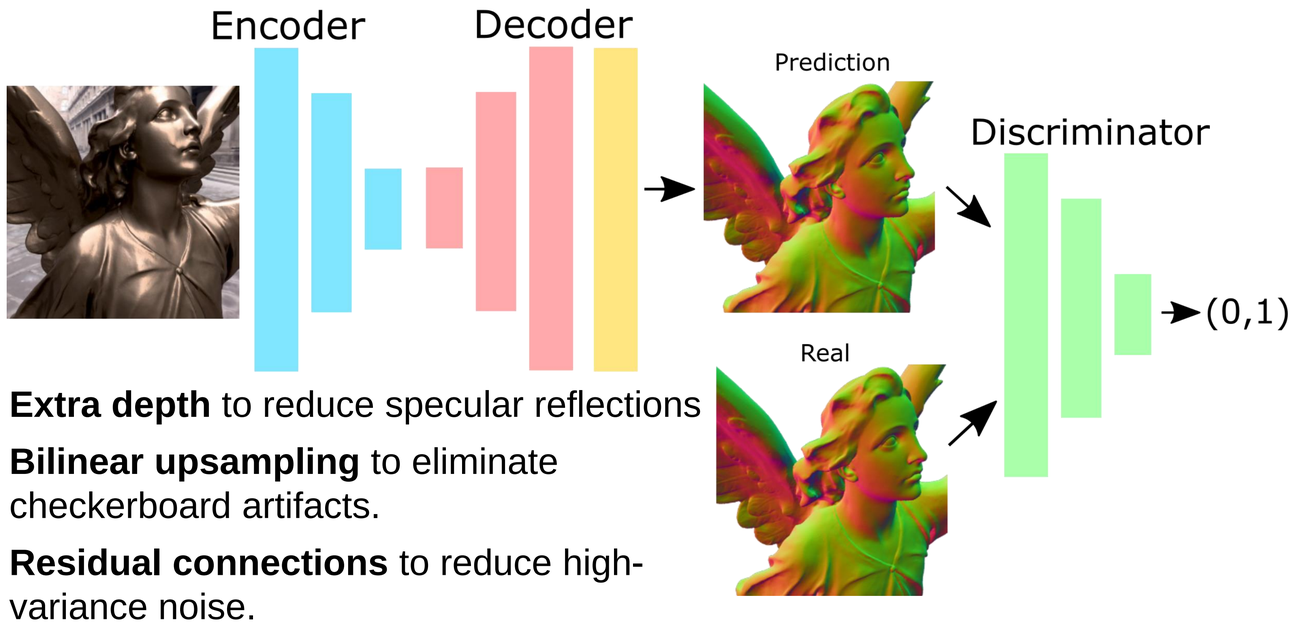
Jorge Condor, Manuel Lagunas, Johanna Delanoy, Belén Masiá, Diego Gutiérrez

MOTIVATION

- **Normal maps provide information about an object's 3D geometry** without access to the 3D model itself.
- In order **to edit** the **appearance** of objects **from images**, **we need** to estimate **geometry**, **illumination** and **material properties**.
- **Current methods** to compute volume from RGB images **require several viewpoints or control over the illumination**.

Our method estimates normal maps (volume) in the wild (completely uncontrolled environment, under any lighting condition), **requiring just a single image.**

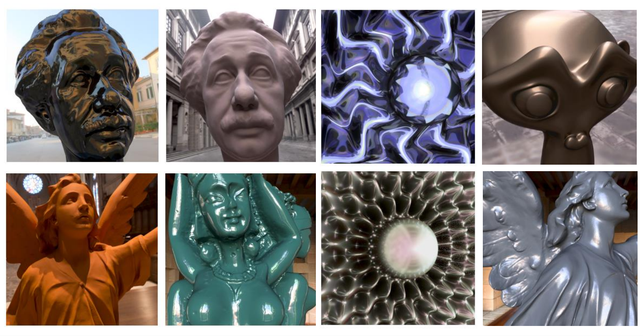
ARCHITECTURE: MODIFIED PIX2PIX



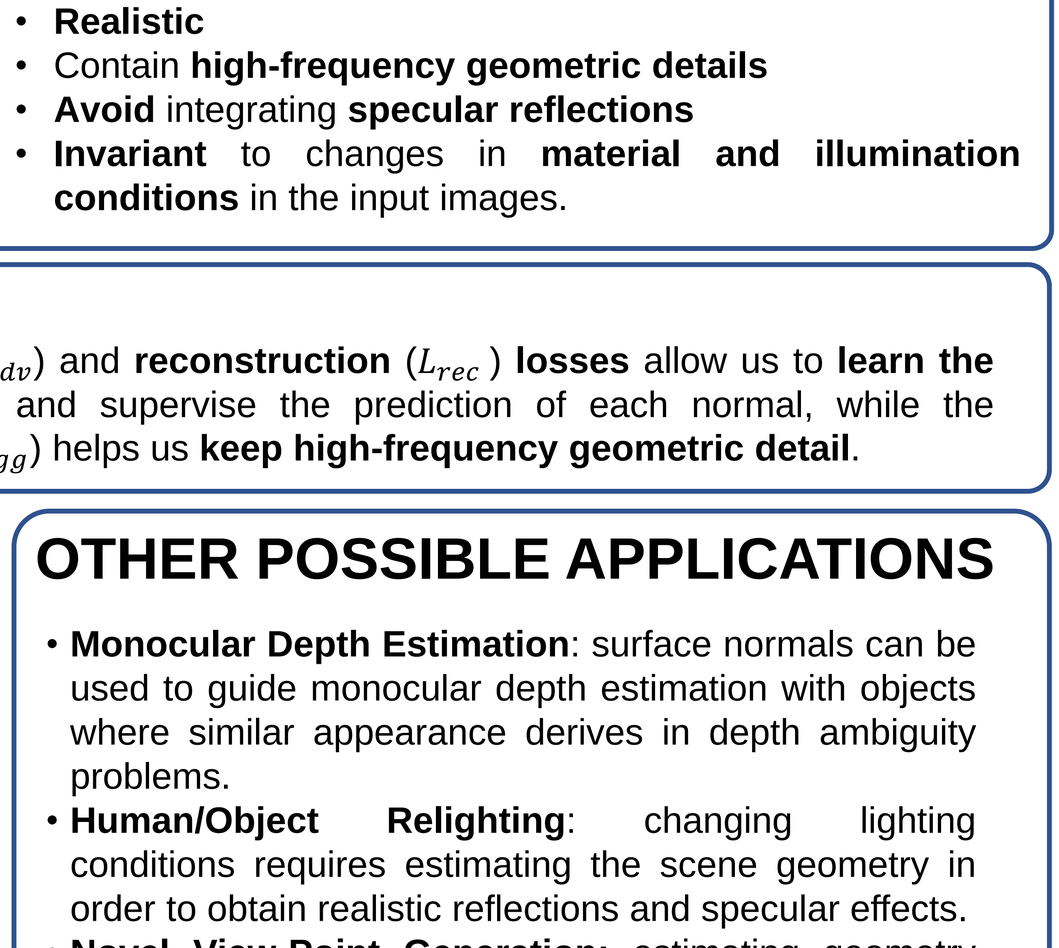
- **Extra depth** to reduce specular reflections
- **Bilinear upsampling** to eliminate checkerboard artifacts.
- **Residual connections** to reduce high-variance noise.

DATASET

We trained on **synthetic data**. The dataset was composed of different geometries, viewpoints, illumination conditions and materials, for a total of **42000 images**.



RESULTS



- **Realistic**
- Contain **high-frequency geometric details**
- **Avoid** integrating **specular reflections**
- **Invariant** to changes in **material and illumination conditions** in the input images.

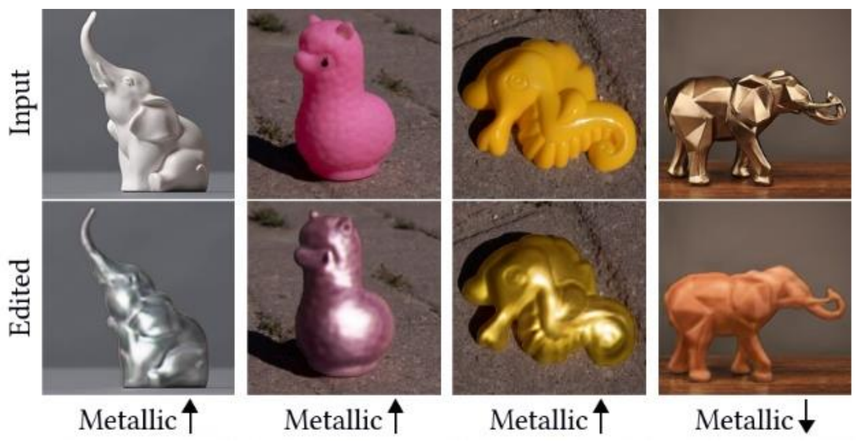
TRAINING

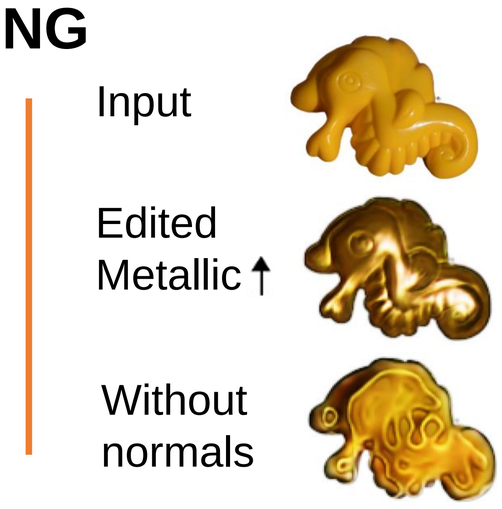
We train our model using the following **custom loss function**:

$$Loss = 0.25L_{adv} + L_{vgg} + 10L_{rec}$$

The **adversarial** (L_{adv}) and **reconstruction** (L_{rec}) **losses** allow us to **learn the target distribution** and supervise the prediction of each normal, while the **perceptual loss** (L_{vgg}) helps us **keep high-frequency geometric detail**.

APPLICATIONS OF OUR METHOD: IMAGE-BASED MATERIAL APPEARANCE EDITING





- Our volume estimations drive a framework that **edits the perceived material properties** of objects **using a single image** of them.

OTHER POSSIBLE APPLICATIONS

- **Monocular Depth Estimation**: surface normals can be used to guide monocular depth estimation with objects where similar appearance derives in depth ambiguity problems.
- **Human/Object Relighting**: changing lighting conditions requires estimating the scene geometry in order to obtain realistic reflections and specular effects.
- **Novel View-Point Generation:** estimating geometry from a few images can help reduce the number of images required to learn a 3D representation of an object and generate new images of it from novel points of view.

Graphics and Imaging Lab
LIRIS
Escuela de Ingeniería y Arquitectura
Universidad Zaragoza
Contact:
jorge.condorlacambra@gmail.com

Take a picture to **download** the **full paper** →