Instituto Universitario de Investigación
**de Ingeniería de Aragón**
**Universidad** Zaragoza

# In-the-wild Material Appearance Editing using Perceptual Attributes

**FULL PAPER**

**J. Daniel Subias**[1], Manuel Lagunas[2]

[1] Universidad de Zaragoza - I3A, Zaragoza, Spain
[2] Amazon

## PROBLEM

Intuitively **editing the appearance** of materials, **just from a single image, is a challenging task** given the complexity and ambiguity of the interactions between light and matter.

This problem has been **traditionally solved by estimating additional factors** of the scene like geometry or illumination, thus **solving an inverse rendering problem** where the interaction of light and matter needs to be modelled.

**We present a single-image appearance editing framework** that allows to intuitively modify the material appearance of an object by **increasing or decreasing high-level perceptual attributes** describing appearance (e.g., glossy or metallic).

## OUR FRAMEWORK

We rely on an **encoder-decoder architecture G that encodes the image x**, and manipulates the **latent space z** together with the **target attribute $att_t$** to generate the **edited image y**. We use skip connections with STU cells, a variant of the GRU [1,2], which allow encoder-decoder architectures to keep the relevant information of the input image in the edited output when manipulating the **latent space z**.
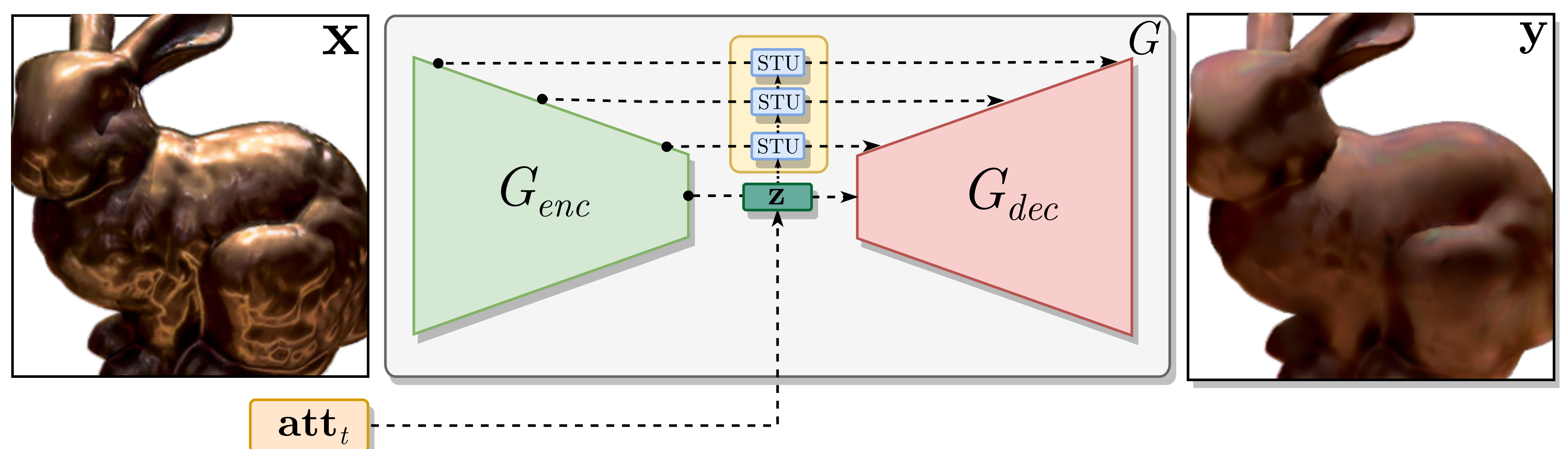


*High-level overview of our framework. Our generator G is composed of an encoder $G_{enc}$ and a decoder $G_{dec}$. It is capable of editing the input image x according to the target attribute $att_t$ to generate the edited image y.*

## TRAINING SCHEME

We adopt the adversarial training proposed by He et al. [3] and introduce a GAN model where **a discriminator D predicts whether an image is fake (edited) or real** and the high-level attribute value $att_t$. We leverage the dataset of Delanoy et. al [4] composed of **45,500 single-object images paired** with crowd-sourced ratings of high-level attributes.
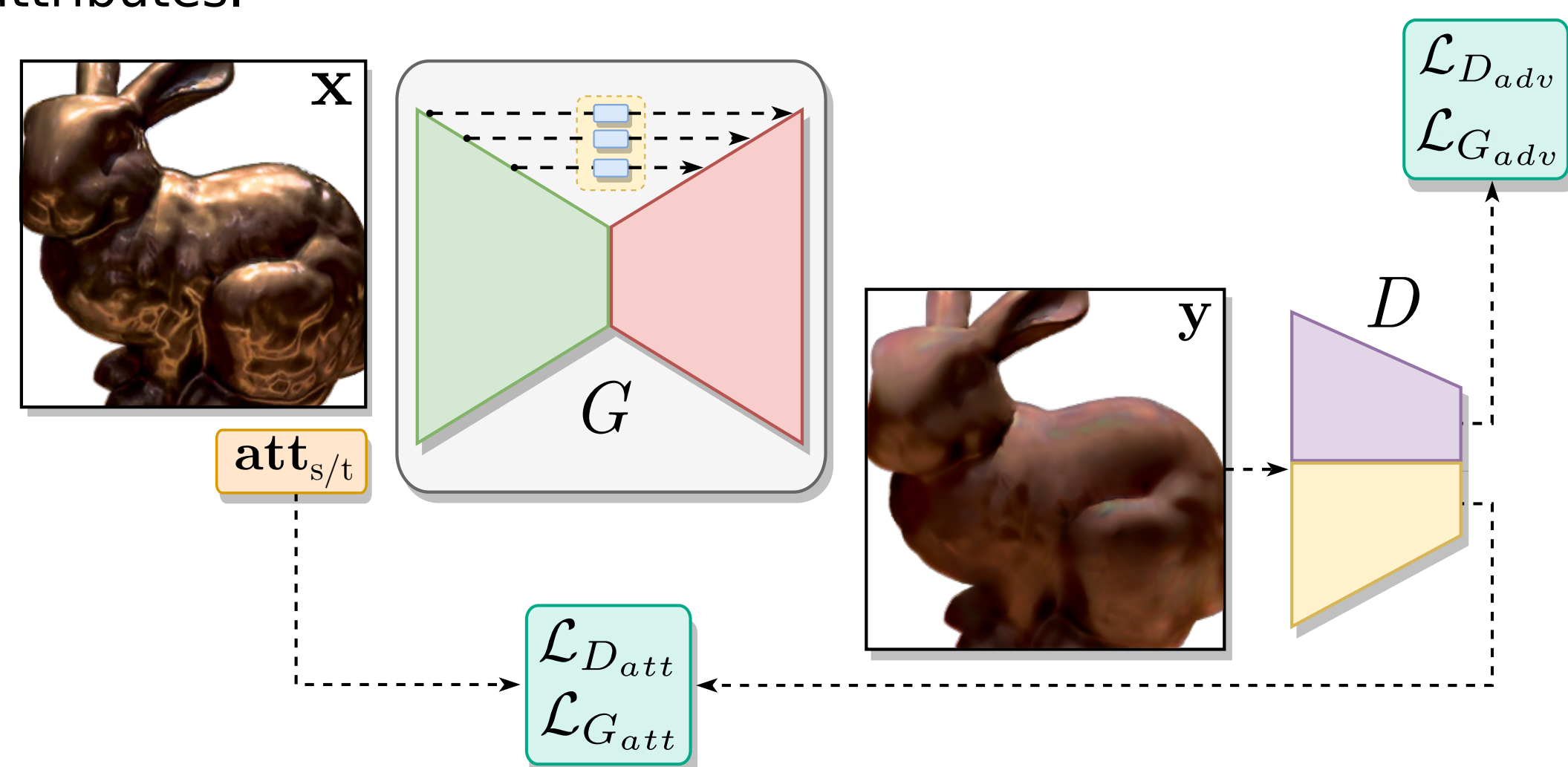


*Training scheme of our framework. The gray block represents our generator G, and the discriminator branches are illustrated by the purple and yellow blocks.*



*Five scenes of the geometries present in the training dataset.*

## RESULTS AND DISCUSSION

Our approach learns to edit perceptual cues properly while **objects' shape remains unchanged**. We test our framework's performance with real-world photographs downloaded from online catalogs of decorative items and with in-the-wild mobile photos taken by us.

**The current state-of-the-art method of Delanoy et al. also needs the normal map** as the input. The material appearance edits from **Delanoy et al. [4] strongly depend on** the shape of their **estimated normal** map [4]. An inaccurate estimation of the normal map may deform the original shape, especially in in-the-wild photographs where geometries are usually highly complex and geometric details not present in the normal map are lost.

**Our method does not need the normal map** as the input. Our results show that the presented method can achieve **realistic results**, almost on par with real photographs, on a **wide variety of different inputs**.
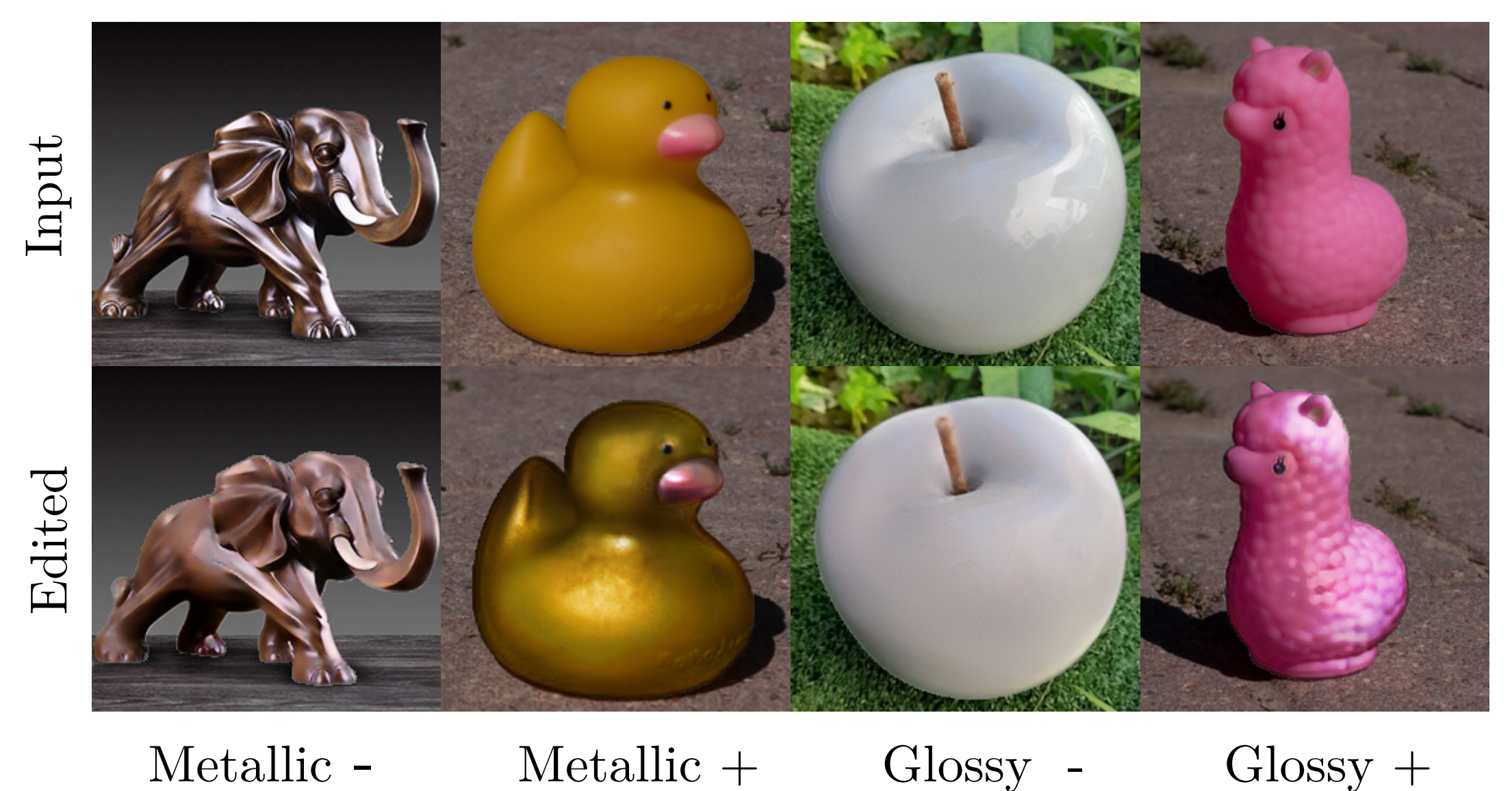


*A sample of the real photographs edited by our framework without suplementary information of the scene.*
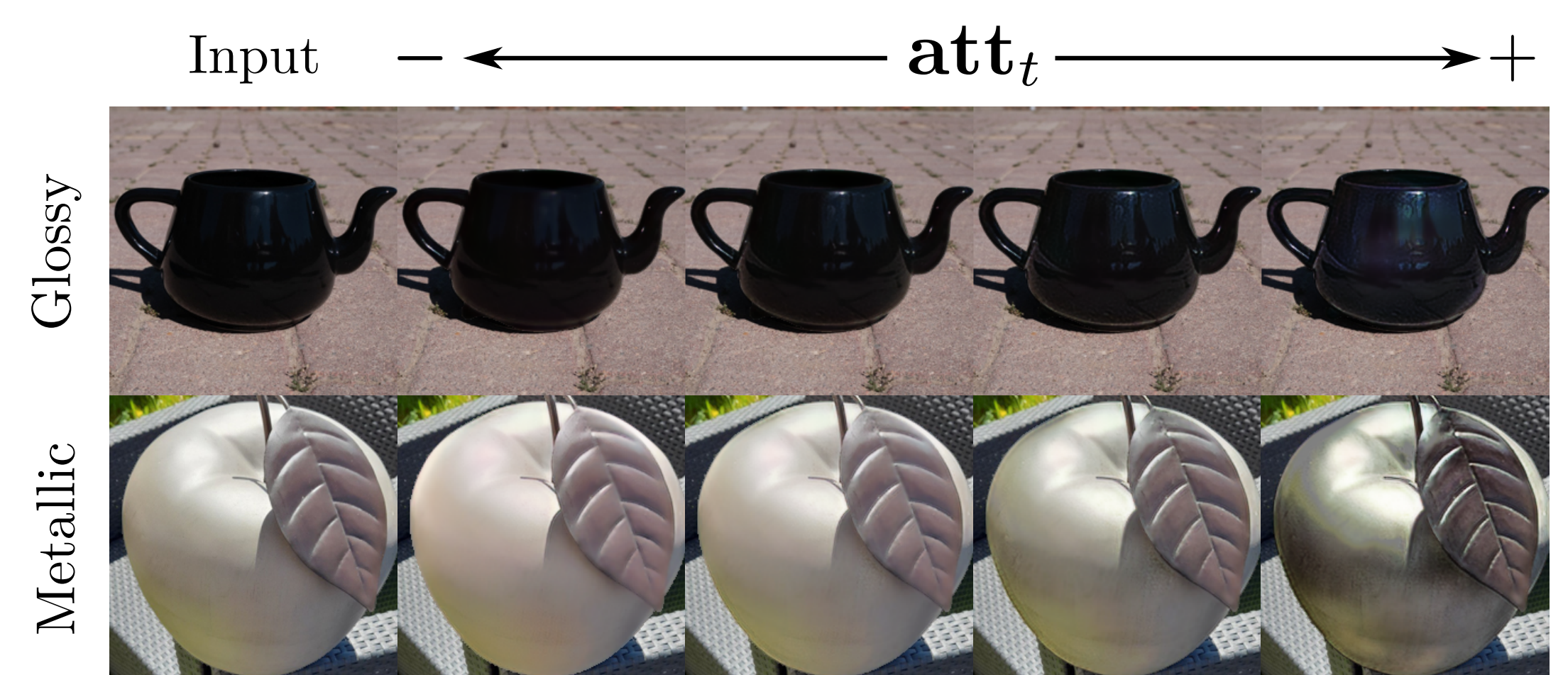


*Editing results by varying the metallic or glossy attributes. The "+" and "-" indicate whether the target high-level perceptual attribute is increased or decreased.*

## REFERENCES

[1] - CHO K., VAN MERRIËNBOER B., GULCEHRE C., BAH-DANAU D., BOUGARES F., SCHWENK H., B ENGIO Y.: Learning phrase representations using RNN encoder–decoder for statistical machine translation. In Proc. Empirical Methods in Natural Language Processing (EMNLP) (Doha, Qatar, Oct. 2014), Association for Computational Linguistics, pp. 1724–1734.

[2] - CHUNG J., GÜLÇEHRE Ç., CHO K., BENGIO Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling.CoRR abs/1412.3555 (2014).

[3] - HE Z., ZUO W., KAN M., SHAN S., CHEN X.: Attgan: Facial attribute editing by only changing what you want. IEEE Transactions on Image Processing 28, 11 (Nov 2019), 5464–5478.

[4] - DELANOY J., LAGUNAS M., CONDOR J., GUTIERREZ D., MASIA B.: A generative framework for image-based editing of material appearance using perceptual attributes. Computer Graphics Forum 41, 1(2022), 453–464.

**Graphics** and **Imaging Lab**
Universidad de Zaragoza

**amazon**

Escuela de Ingeniería y Arquitectura
**Universidad** Zaragoza

**Contact:** dsubias@unizar.es