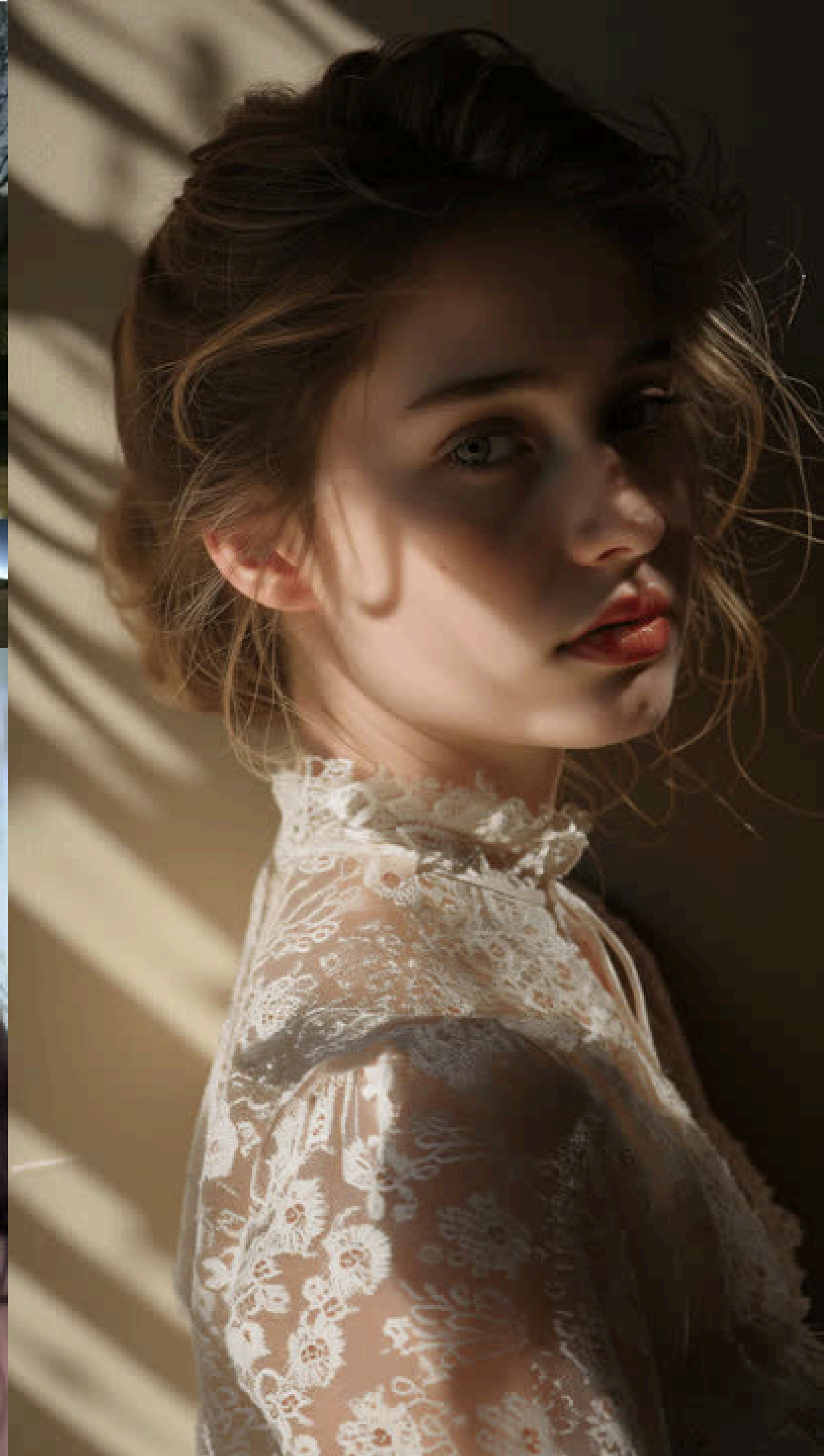


Towards controllable relighting

Maria Pilligua

Supervisor: Javier Vazquez

What is relighting?



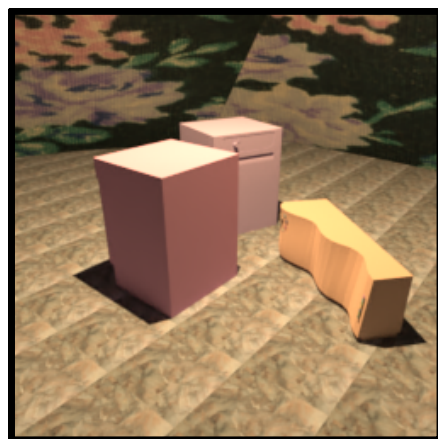
— Objectives

Achieve: Controllable image relighting

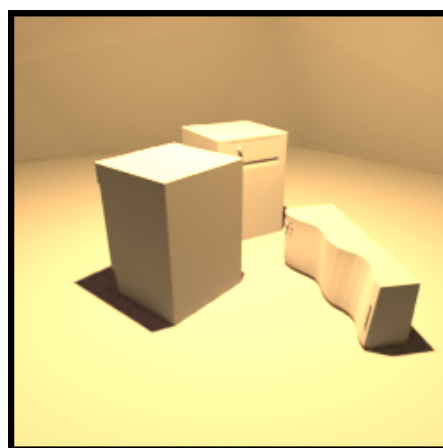
1. Study approaches
2. Design a pipeline with:
 - a. Model able to **relight** images
 - b. **Control** mechanisms
3. Train on a diverse dataset
4. Evaluate on unseen images

Context

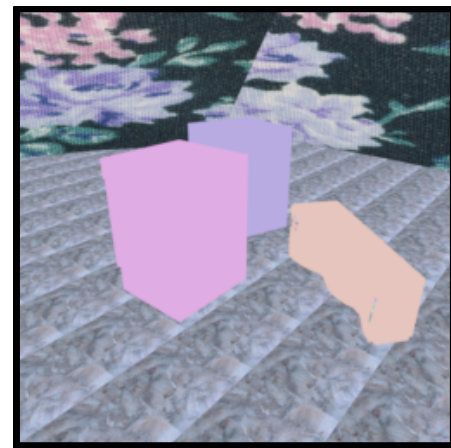
Intrinsic decomposition models



Input image



shading



albedo

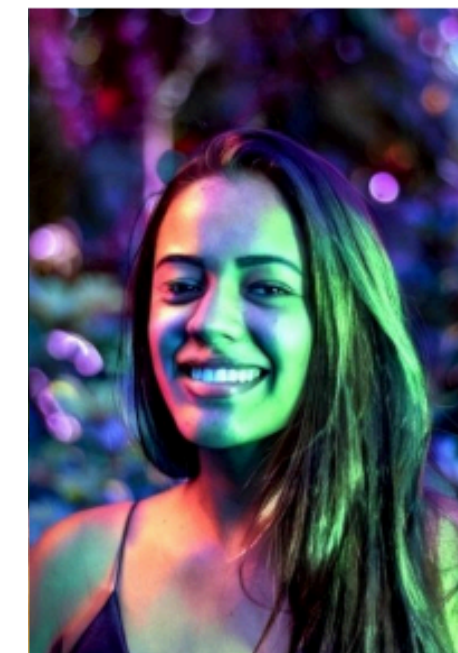
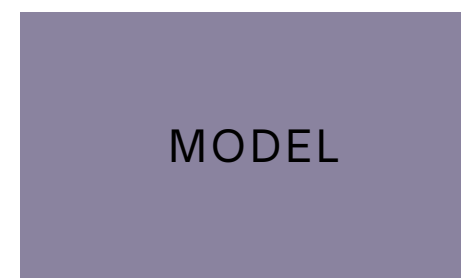
$$Image = Albedo \times Shading$$

Data-driven / Generative models

Image
+
Condition



Relit Output

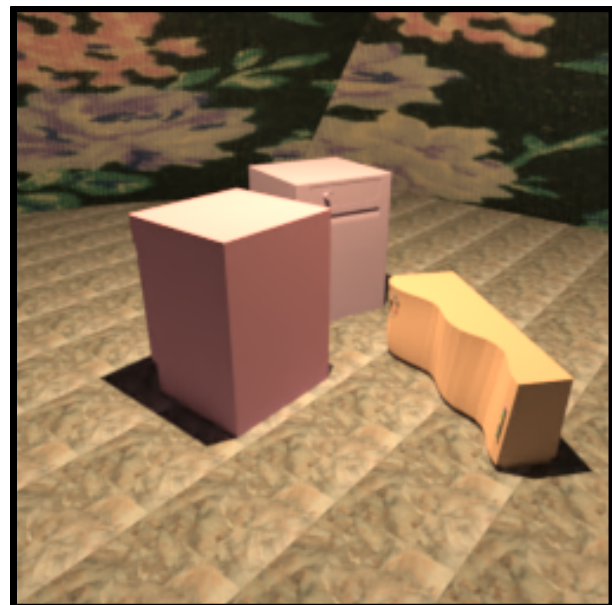


Conditioning: "Neon signs
on the street"

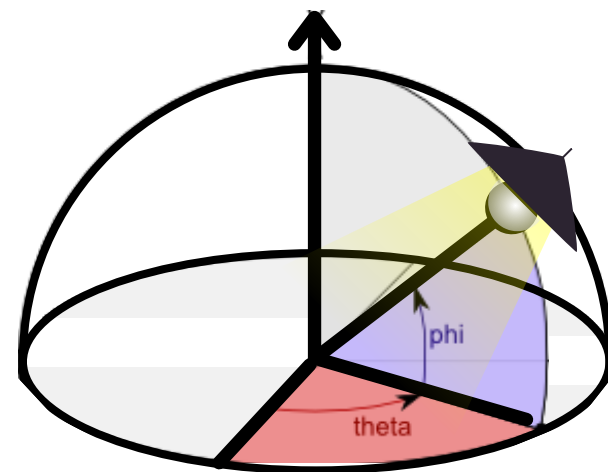
Datasets

ISR: Intrinsic Scene Relighting Dataset

- Synthetic → physically based graphics engine
- 78010 images + annotations



+

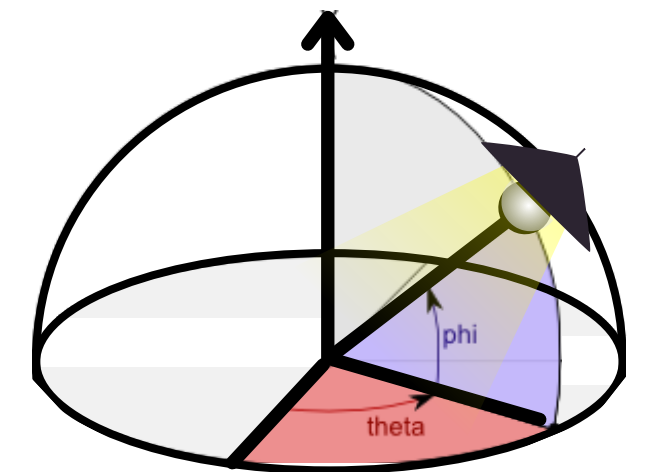


RSR: Real Scene Relighting Dataset

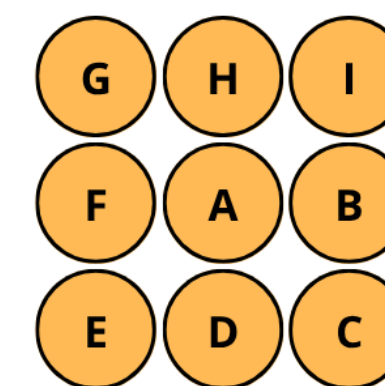
- Real → captured in a controlled laboratory
- 5K images + annotations



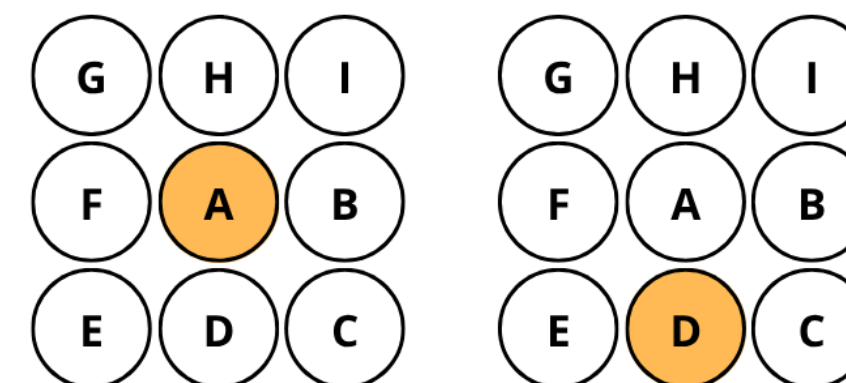
+



Datasets



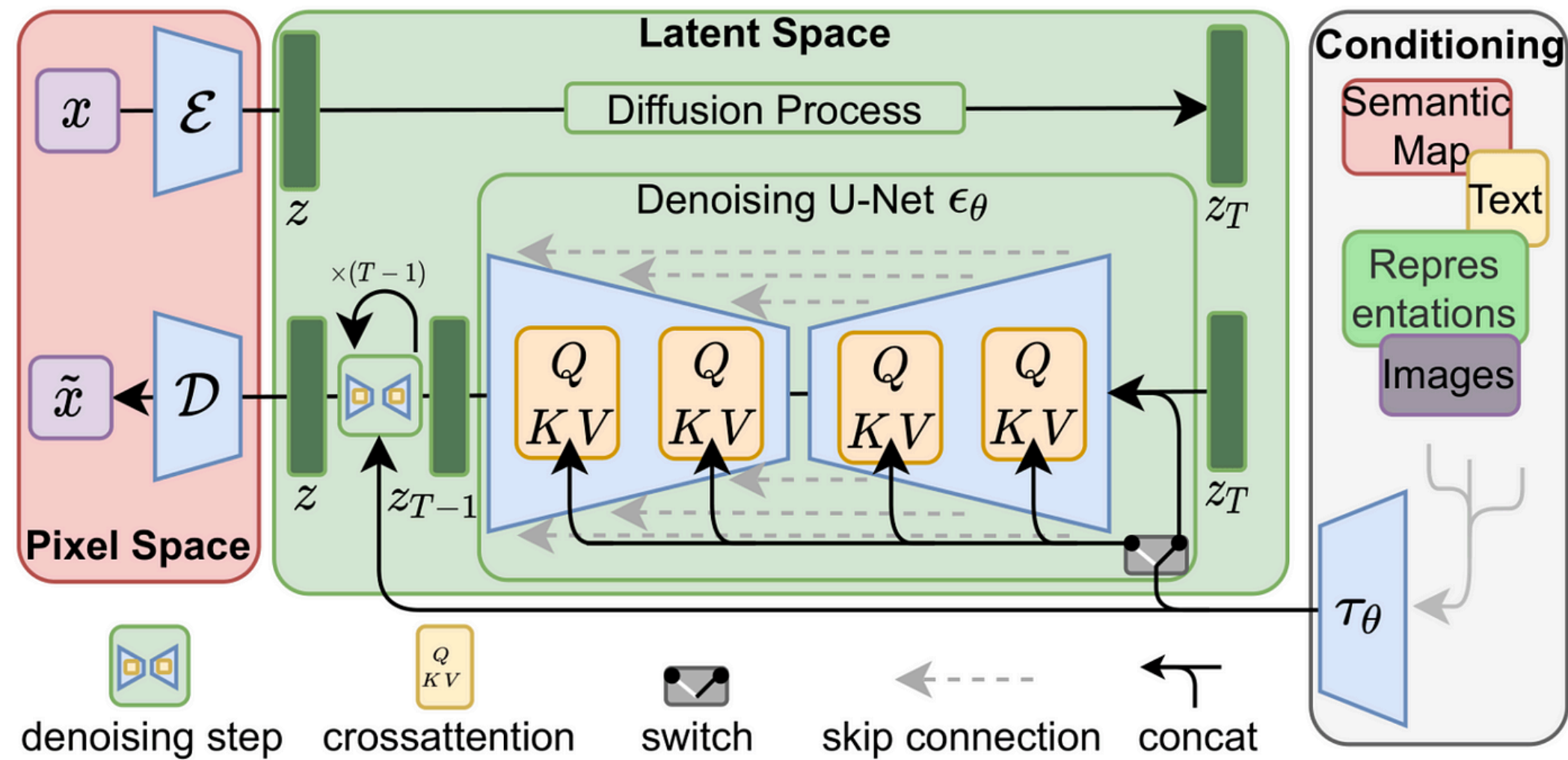
Lights setup



Work done

Diffusion-based relighting

Conditioning Stable Diffusion:



Stable Diffusion architecture

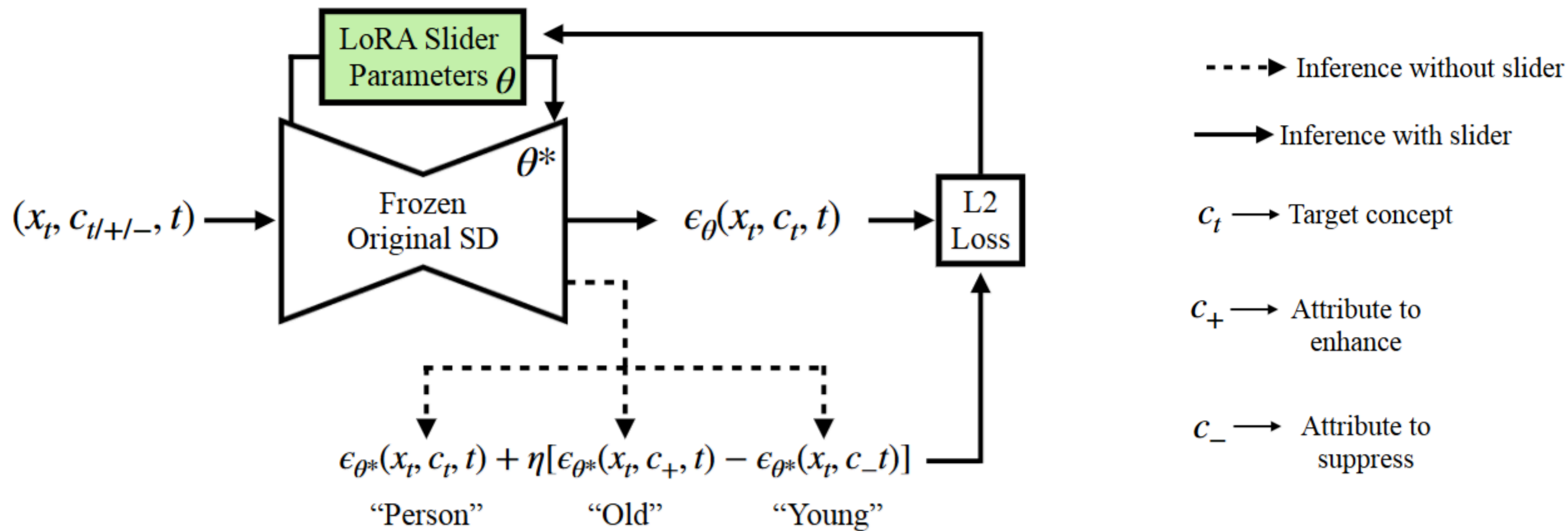


Concept Sliders



Be able to
change a single
attribute of an
image

Concept Sliders

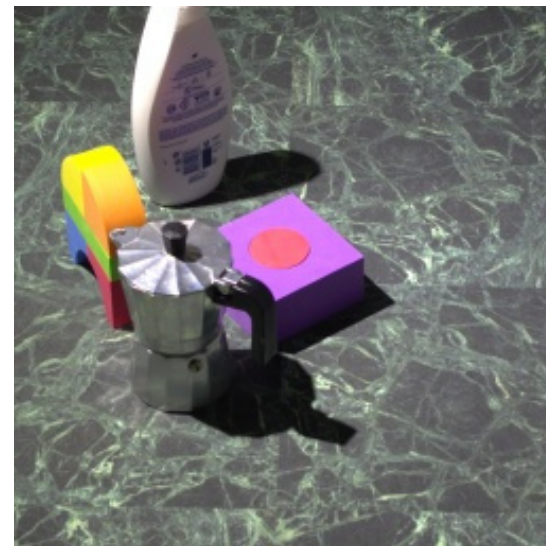


Concept Sliders

A

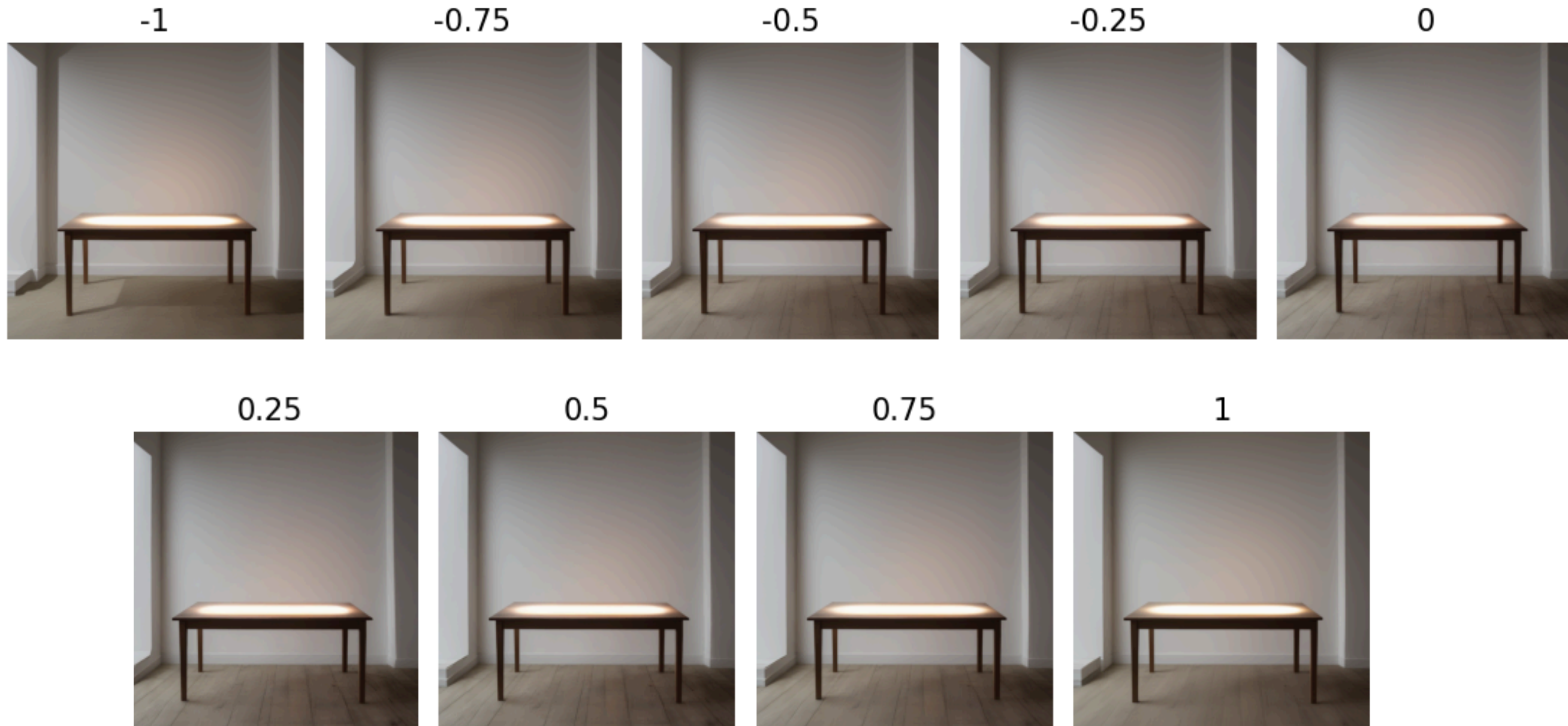


B

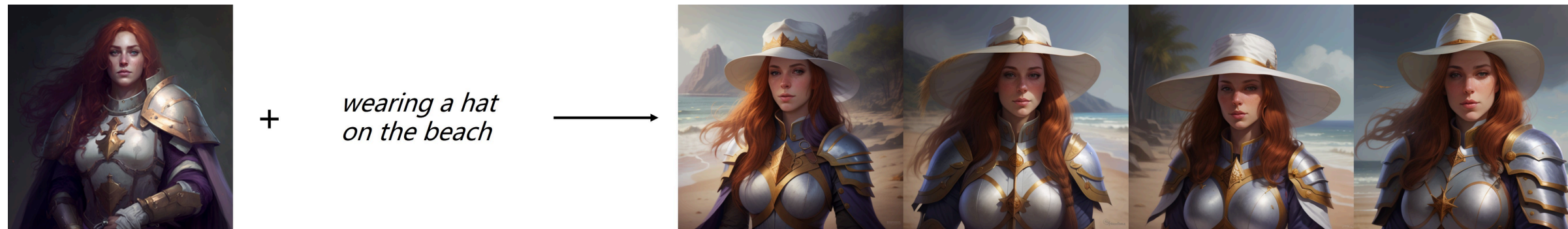
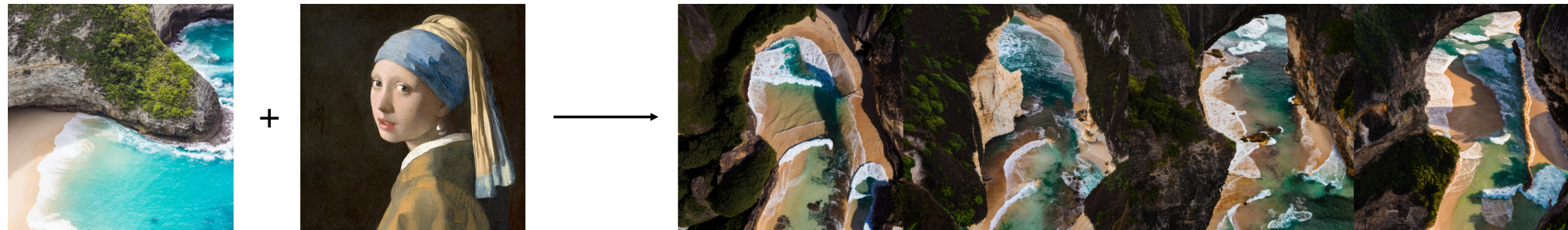


Concept Sliders

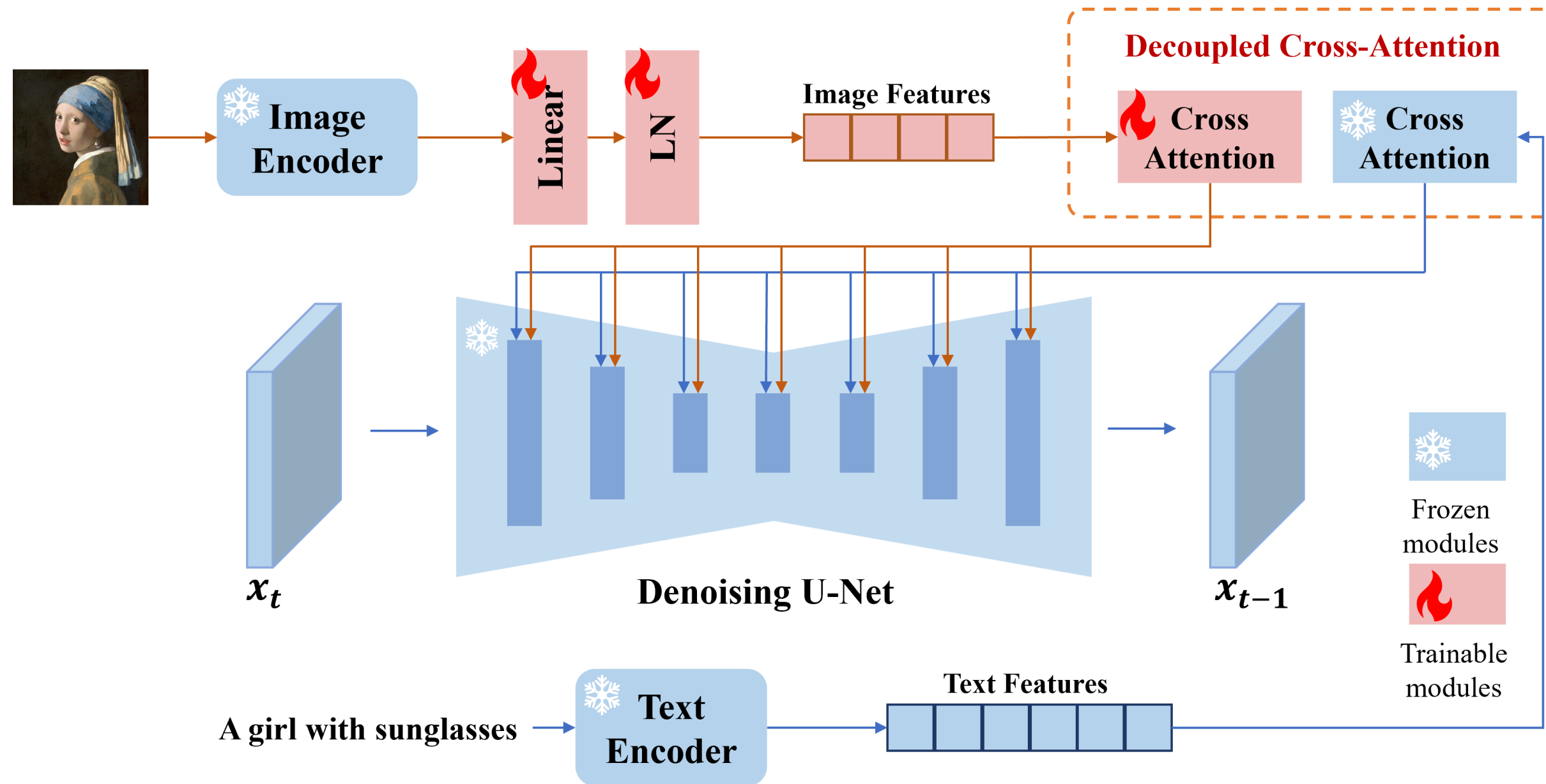
Qualitative results



IP-Adapter



IP-Adapter

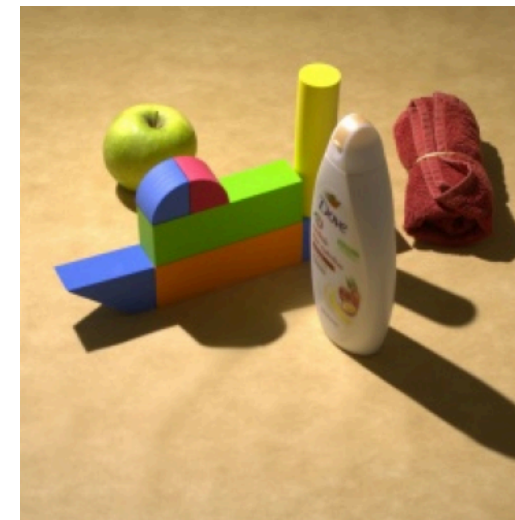


— IP-Adapter



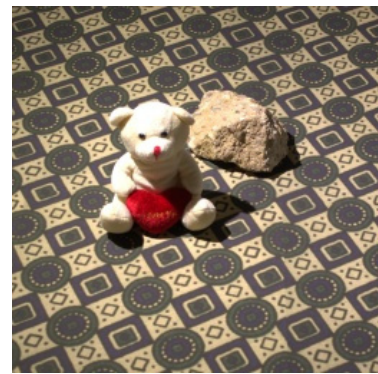
+

*A scene with
pan 0.62 and
tilt 0.10*



IP-Adapter

Stable diffusion with IP-Adapter

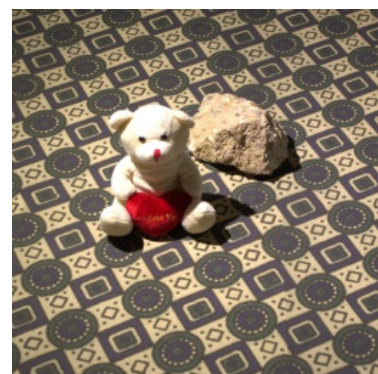


+

*A scene with
tilt 0.75, and
pan 0.19*



Stable diffusion **XL** with IP-Adapter

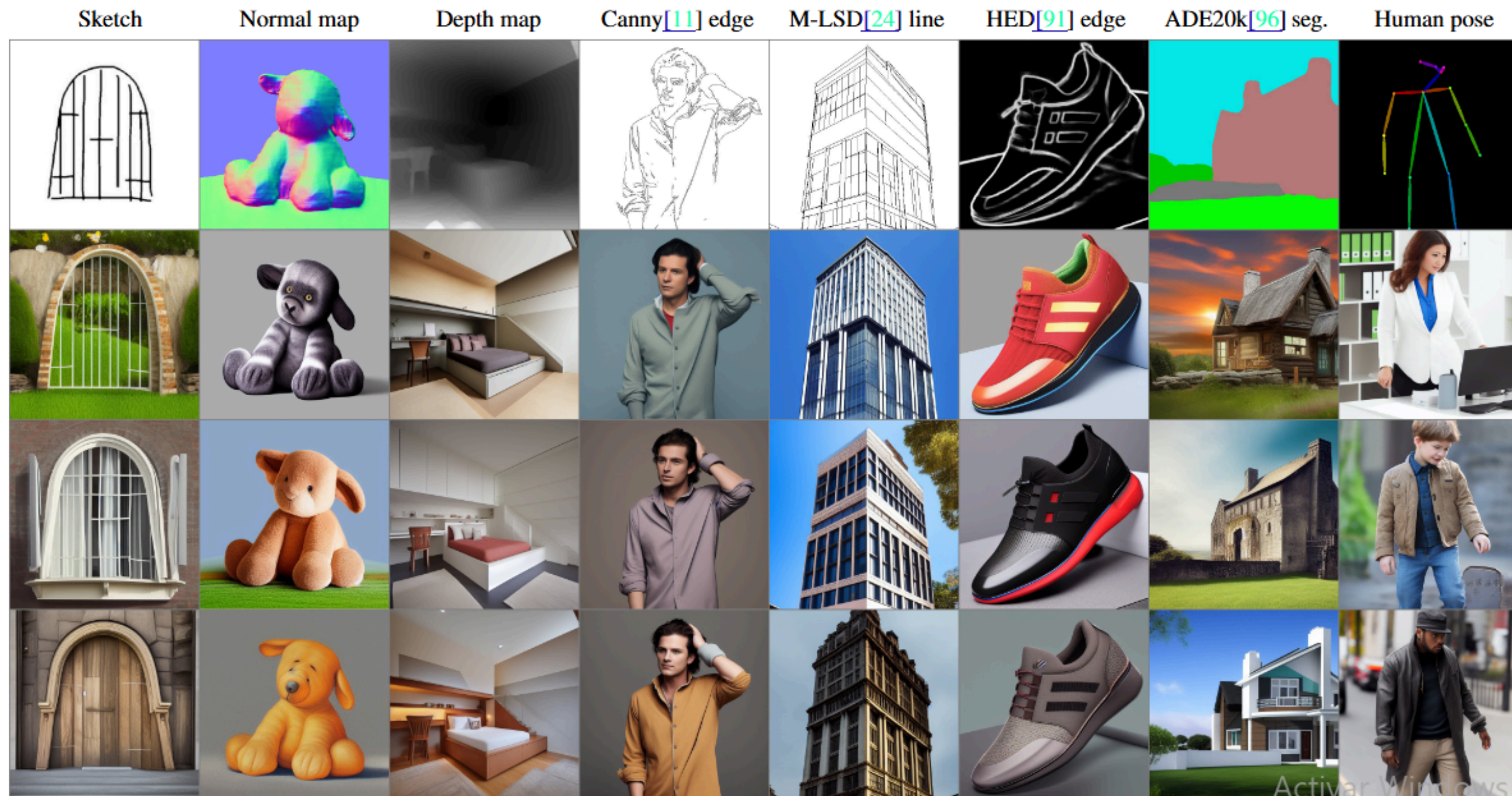


+

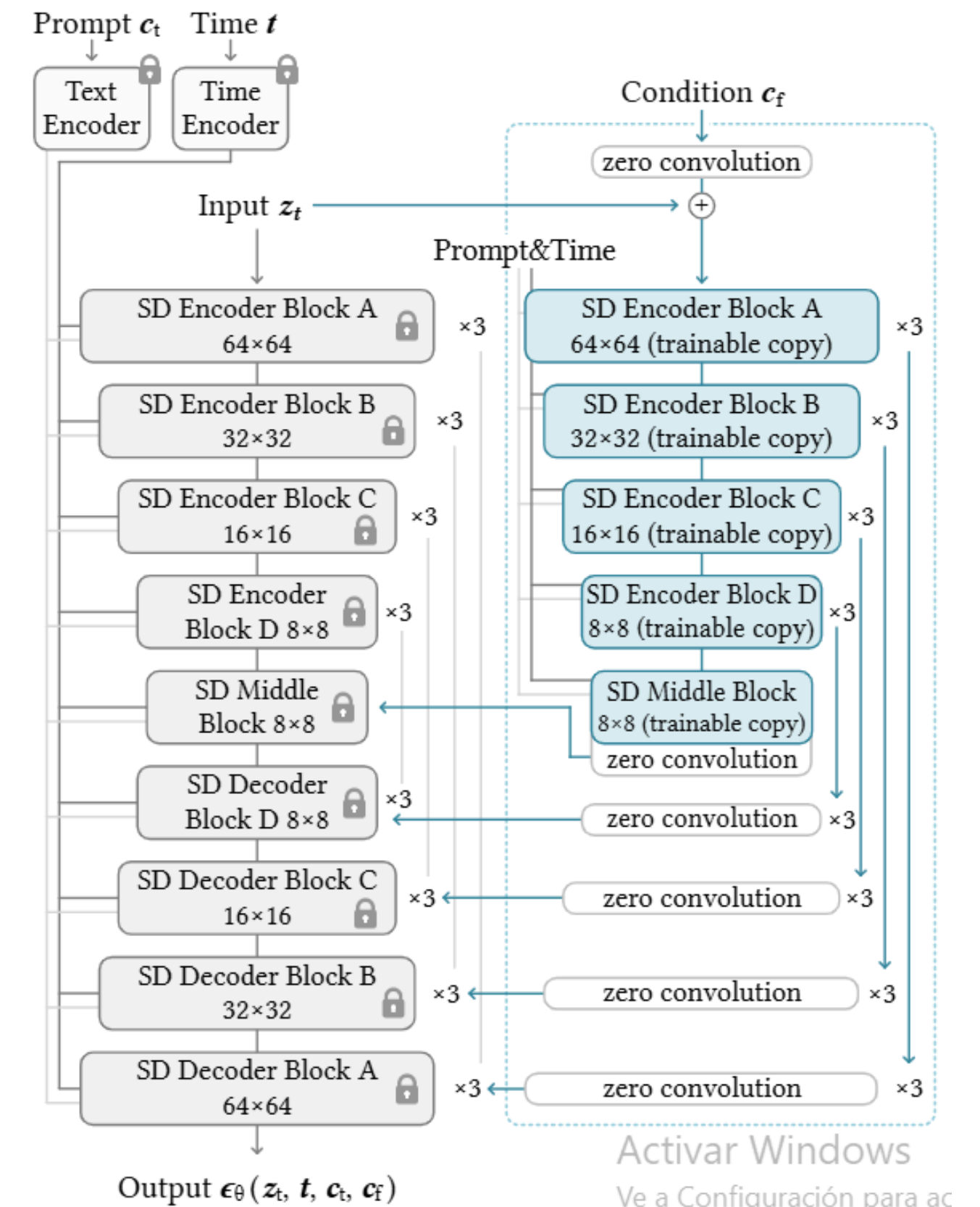
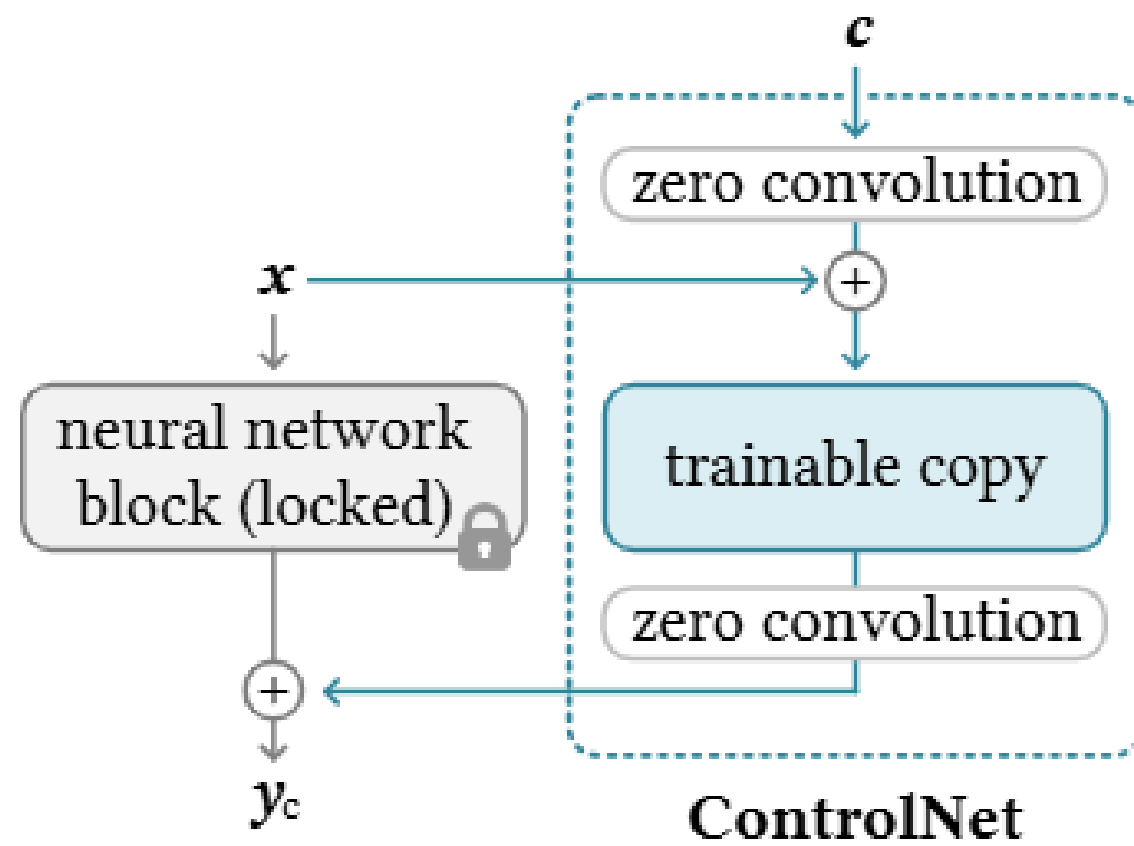
*A scene with
tilt 0.75, and
pan 0.19*



ControlNet



ControlNet



(a) Stable Diffusion

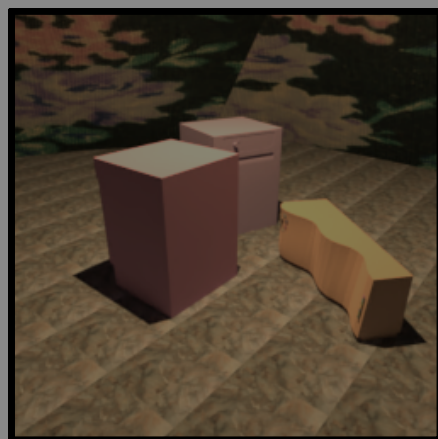
(b) ControlNet

ControlNet - *Albedos*

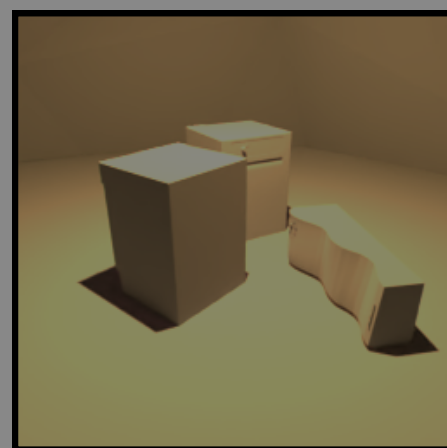
Ideally we would condition controlnet to generate a different image that maintains the albedo of the original image.

Context

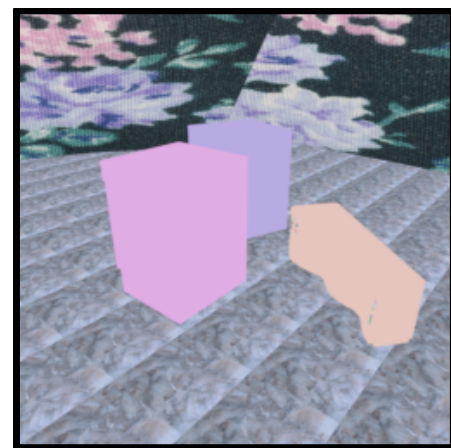
Intrinsic decomposition models



Input image



shading



albedo

$$Image = Albedo \times Shading$$

Data-driven / Generative models

Image
+
Condition



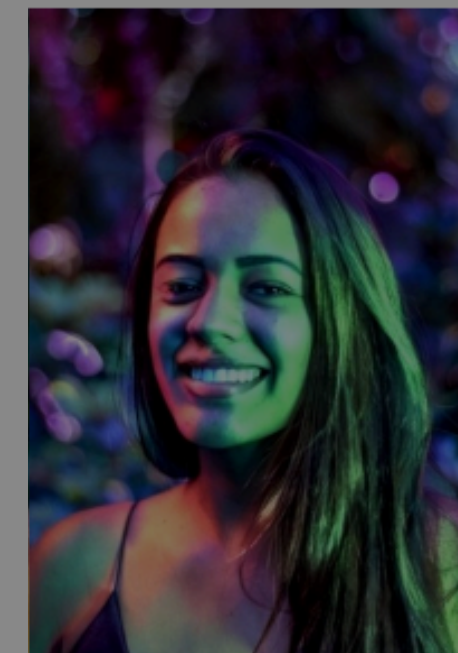
MODEL



Relit Output



MODEL



Conditioning: "Neon signs
on the street"

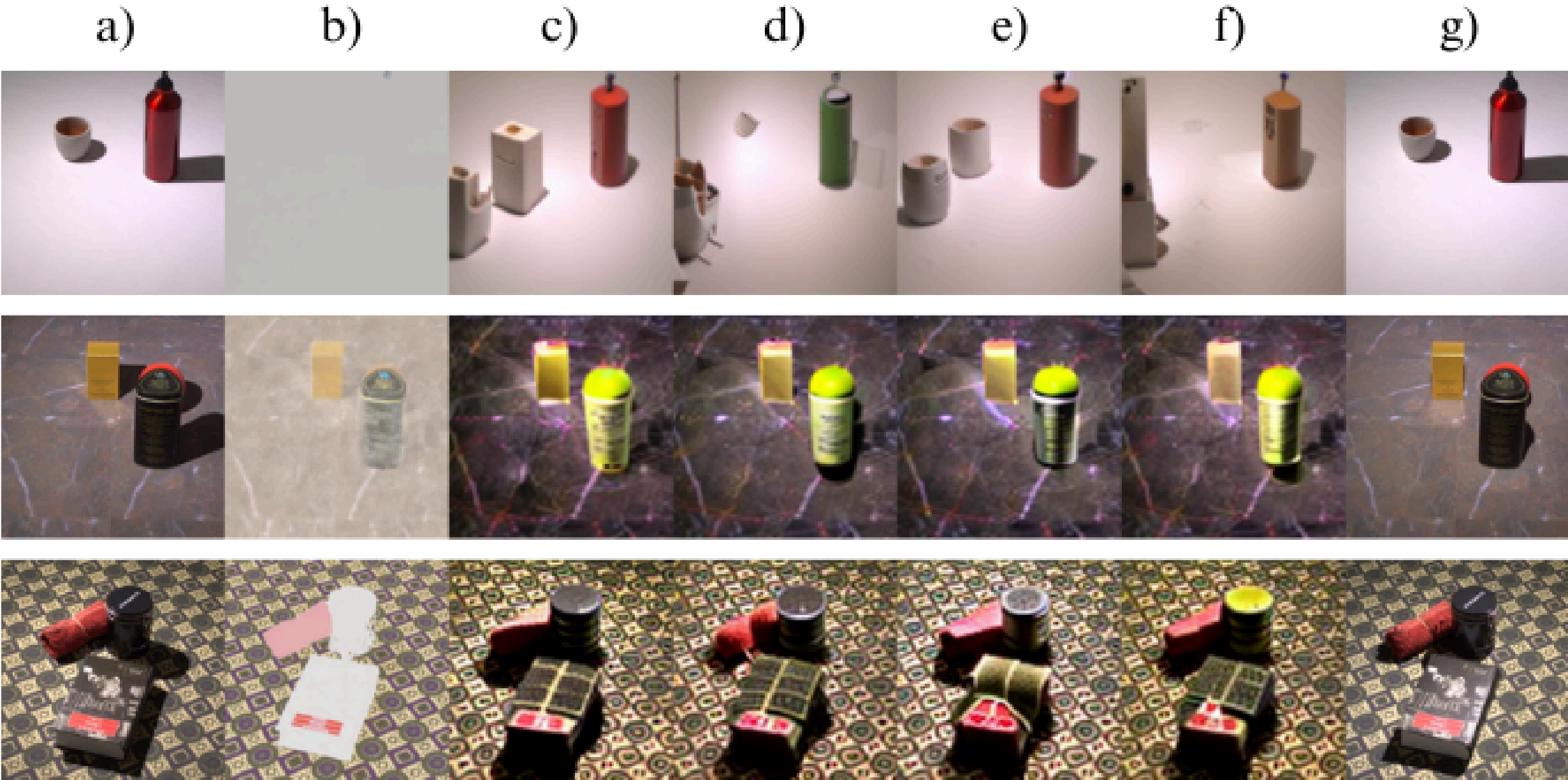
ControlNet - Albedos

To train we just need paired albedo/image data. There is no real huge dataset that has this.

We decided to use another model, RGBX, to extract the albedos from a portion of Laion-aesthetics.



██████████



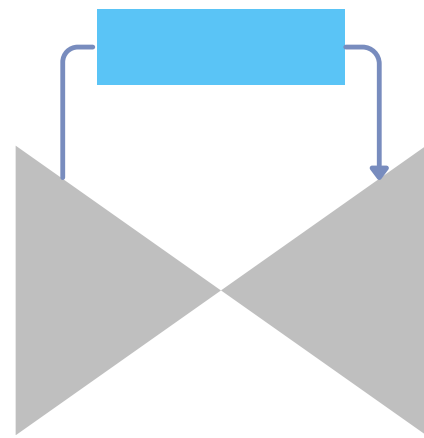
Problems:

1. A lot of texture information is still missed

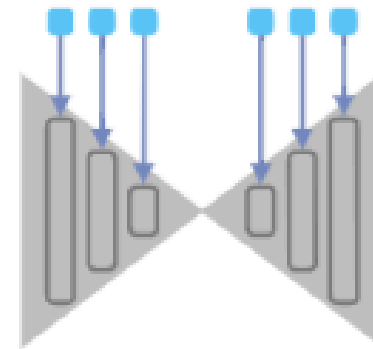
2. If the albedo is not correctly predicted it will not work.

Recap

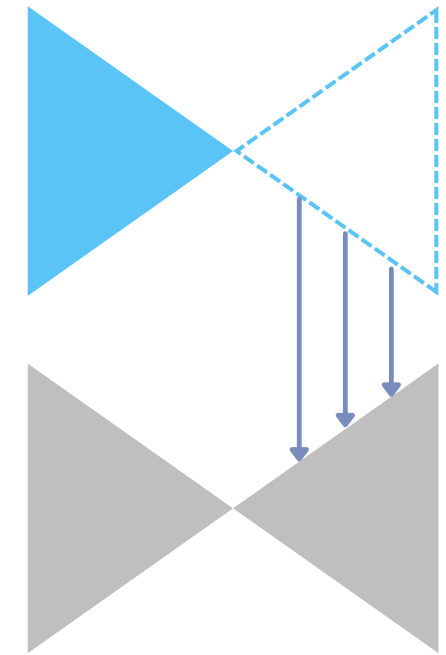
a) Concept Sliders



b) IP-Adapter

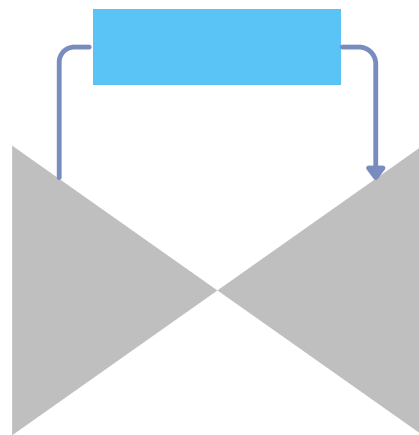


c) ControlNet



Recap. Model attempts:

a) Concept Sliders



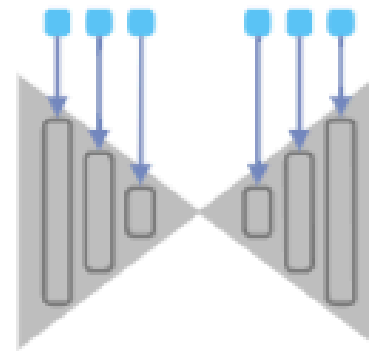
-1



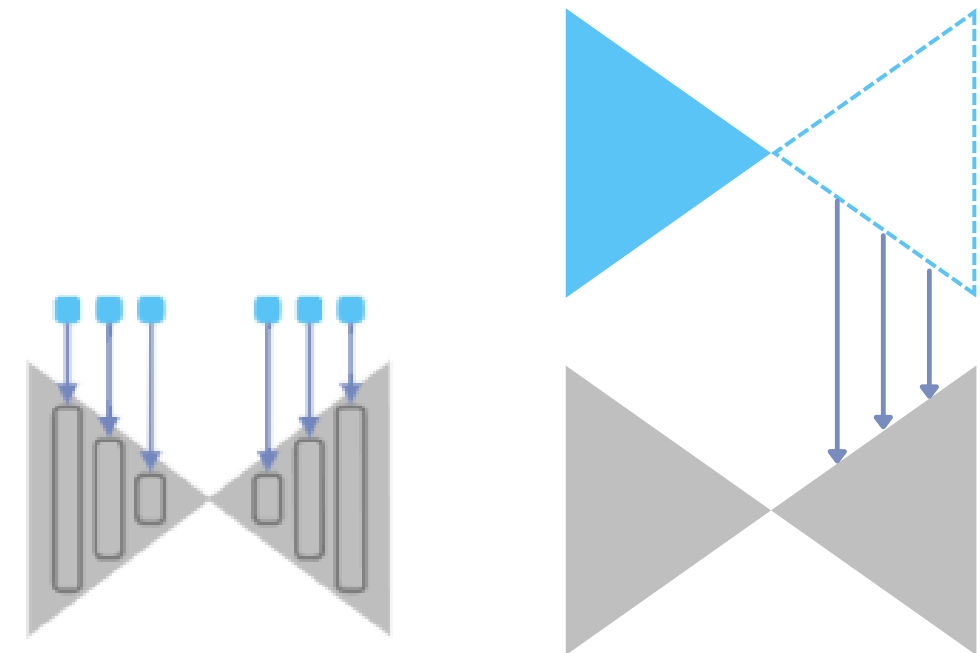
-0.75



b) IP-Adapter



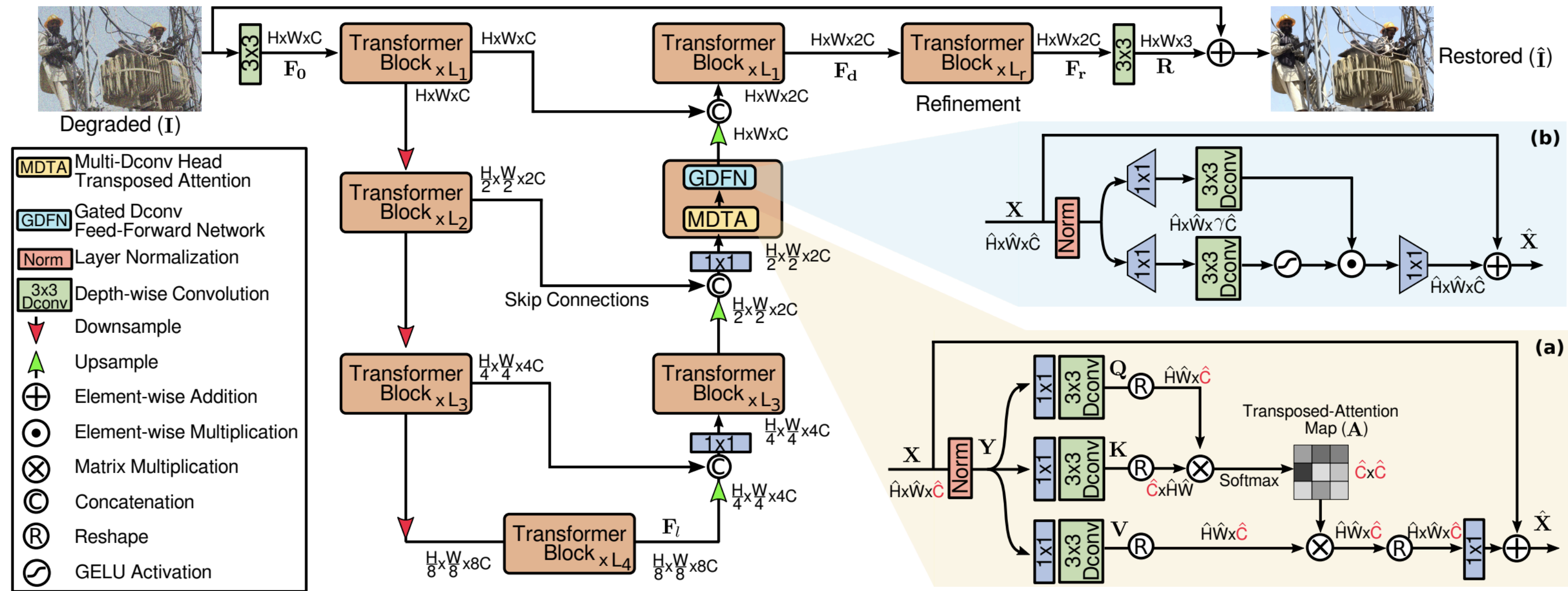
c) IP-Adapter + ControlNet



Change of approach

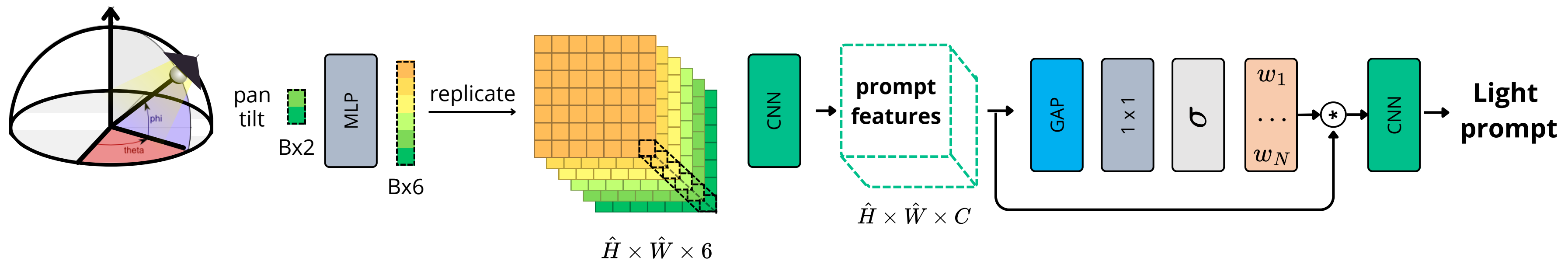
Transformer-based relighting

Restormer

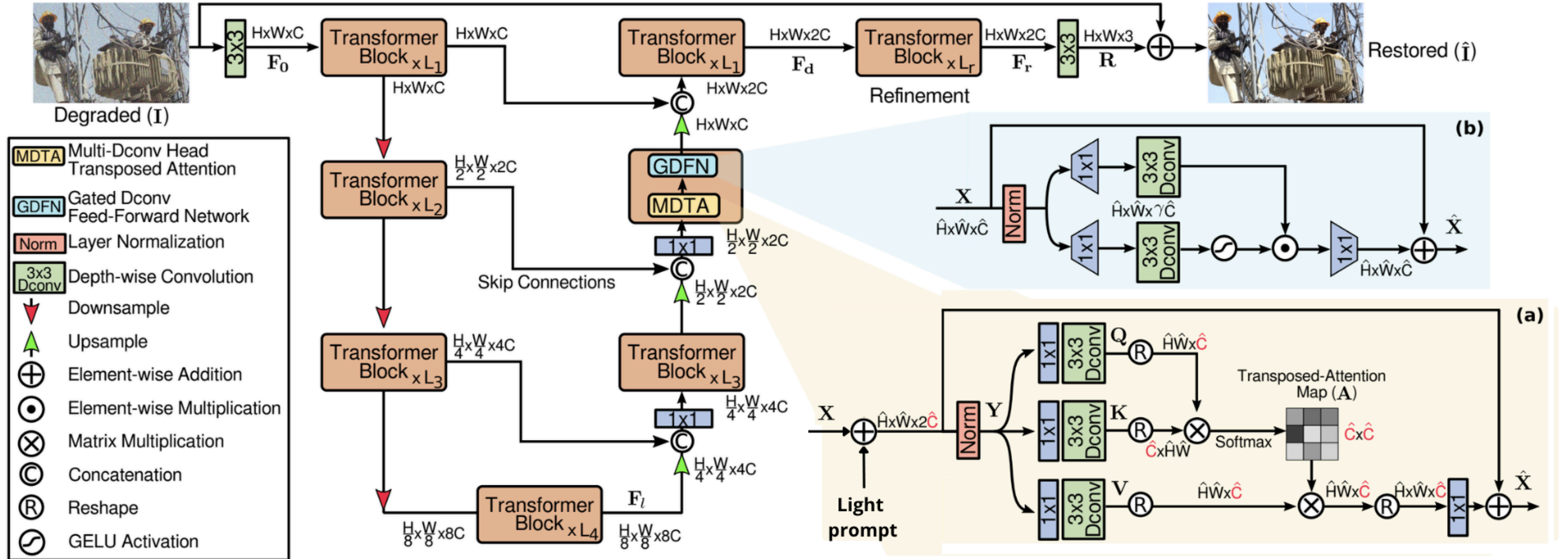


Condition strategy

How do we introduce the light information?



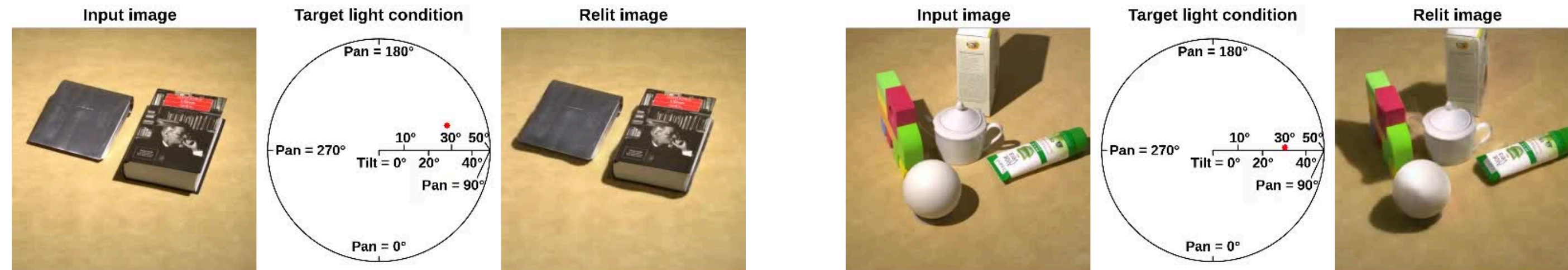
Condition strategy



Qualitative results

The model that performed the best: Restormer with condition strategy 2

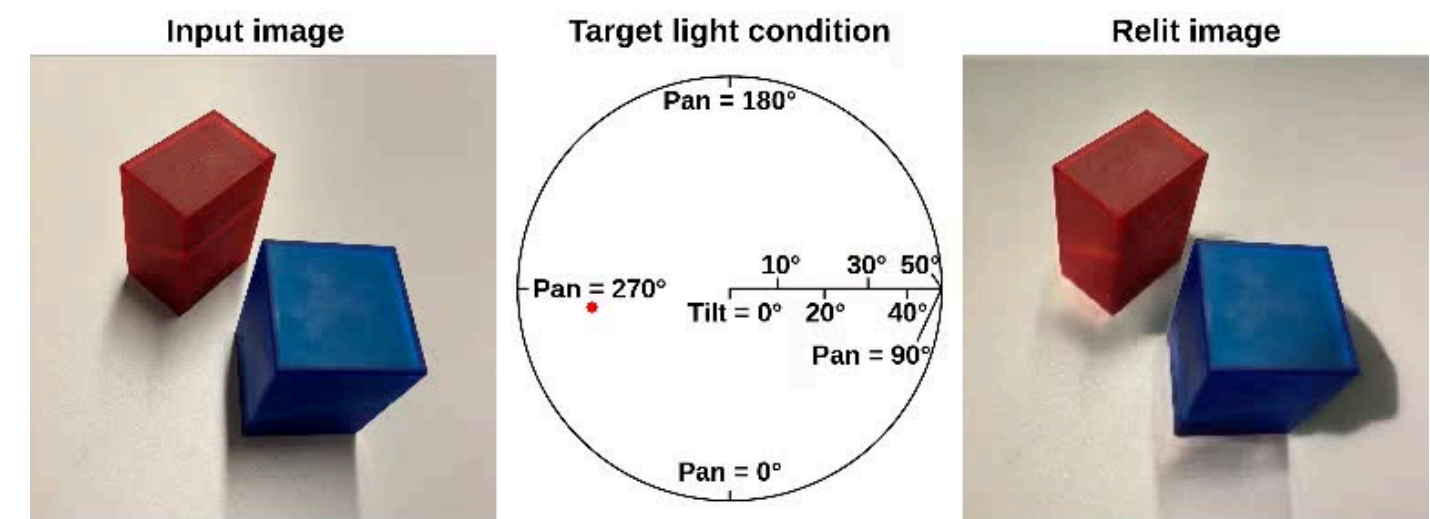
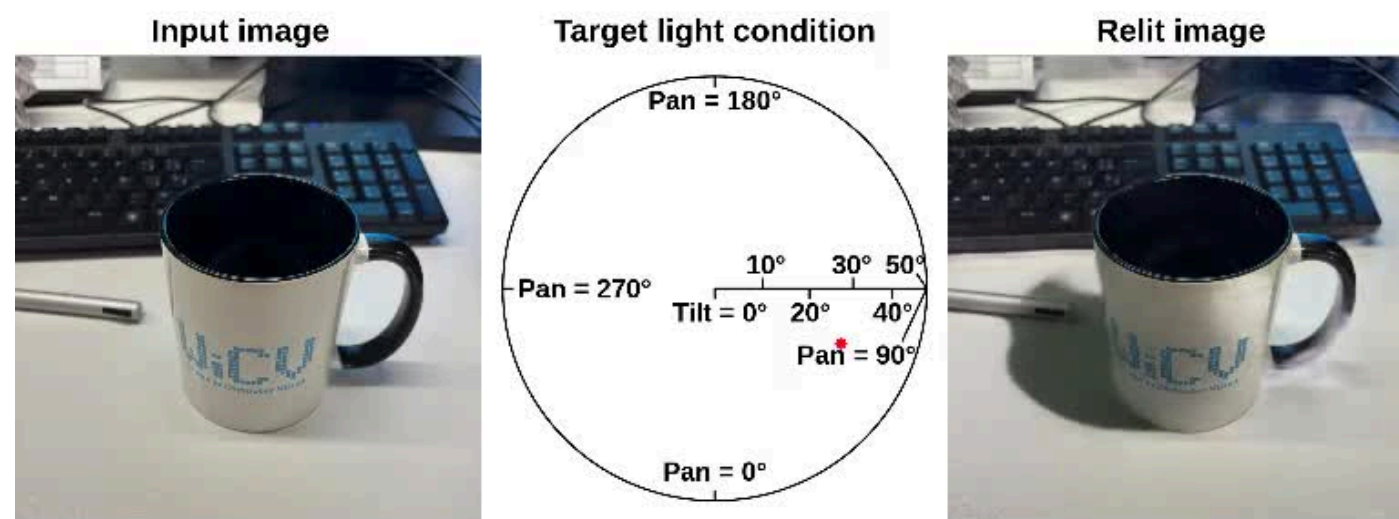
Val set



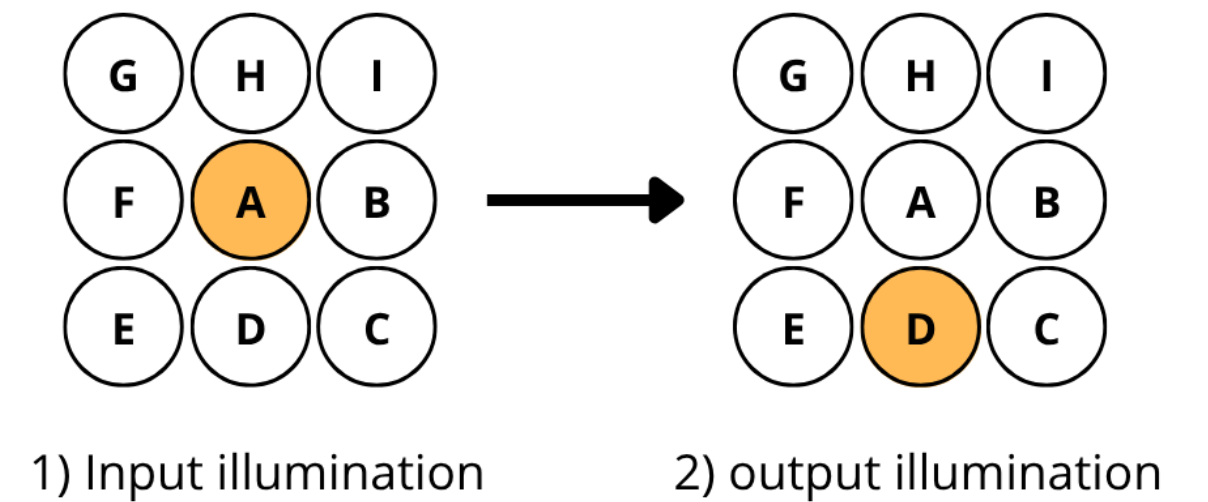
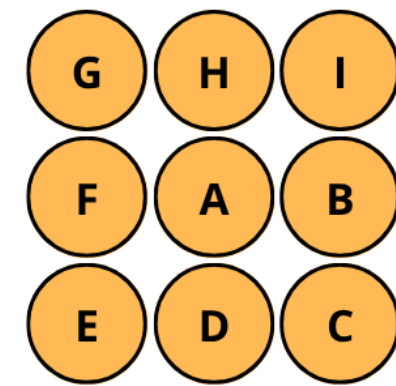
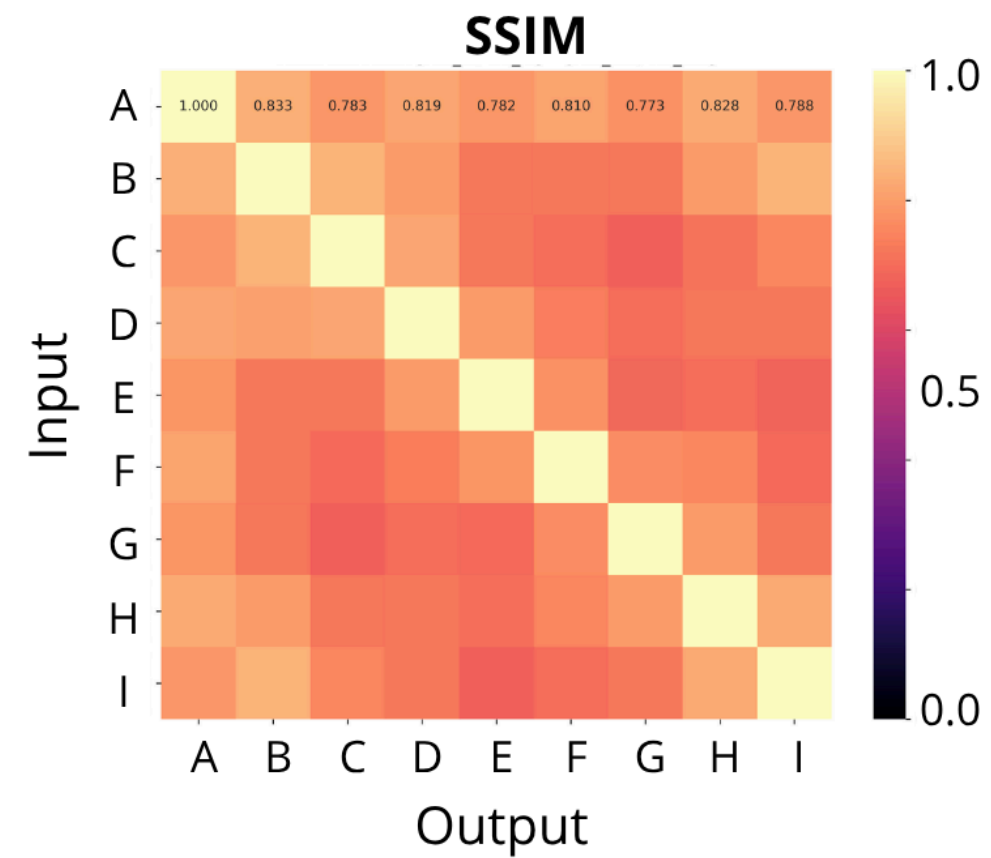
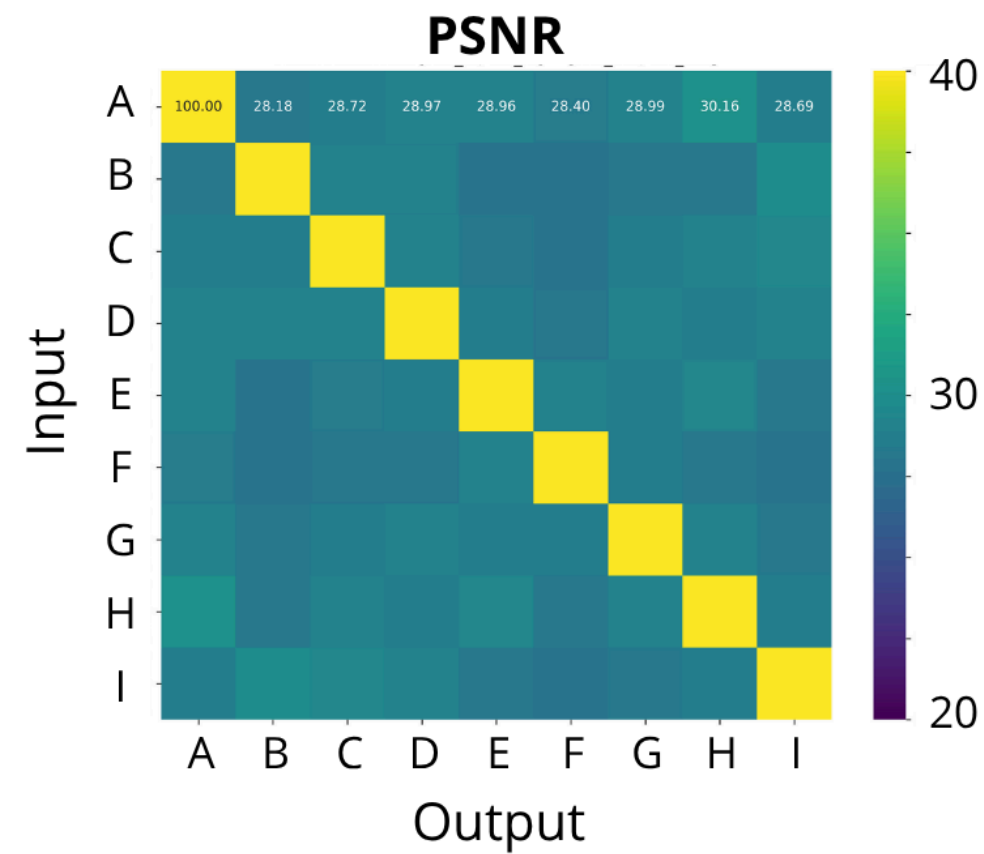
Qualitative results

The model that performed the best: Restormer with condition strategy 2

Out of domain dataset



Quantitative results



Conclusions

We accomplished our initial goals:

1. Study approaches
2. Design a pipeline with:
 - a. Model able to **relight** images
 - b. **Control** mechanisms
3. Train on a diverse dataset
4. Evaluate on unseen images

Really challenging project

Towards controllable relighting

Maria Pilligua

Supervisor: Javier Vazquez