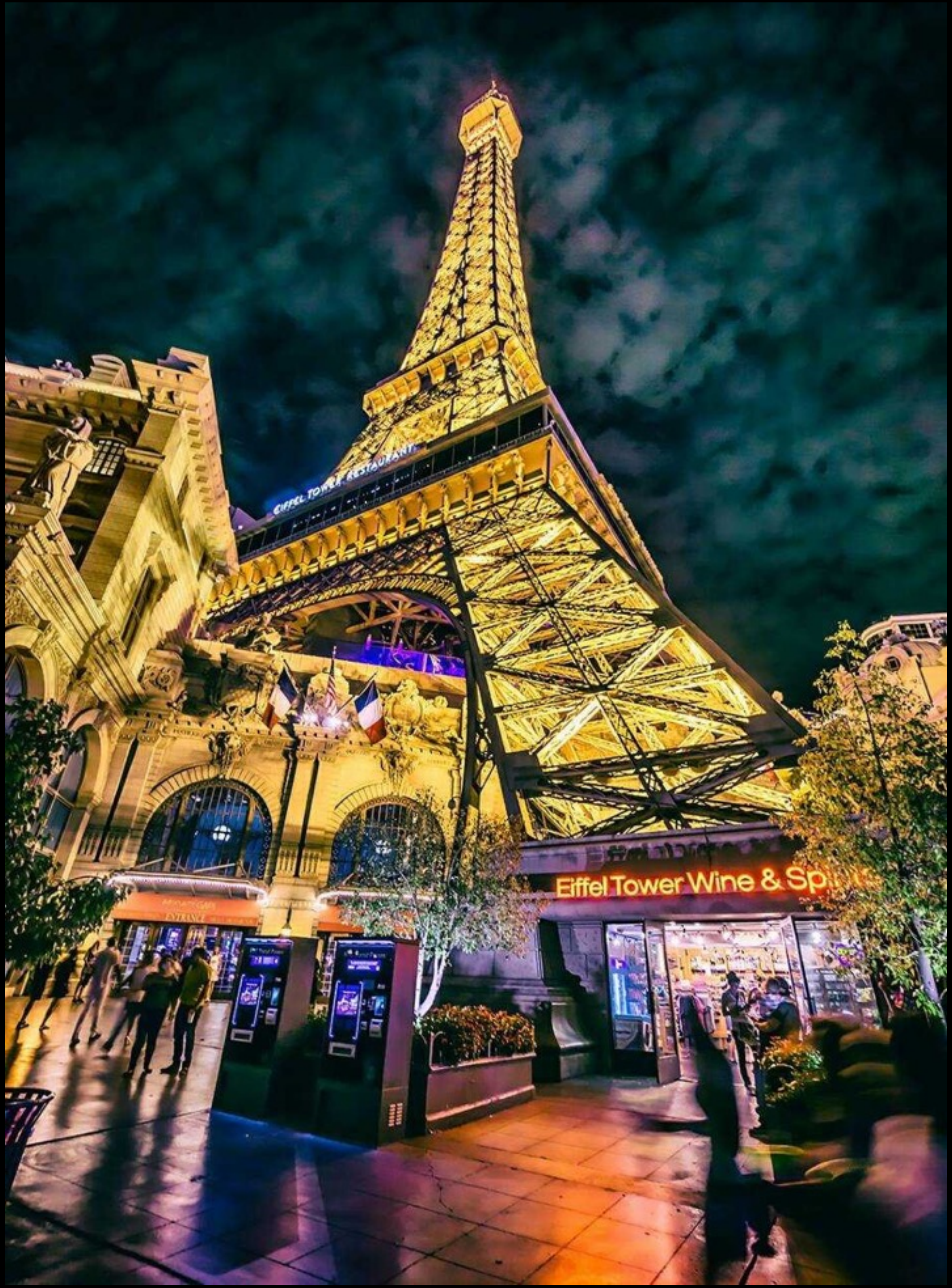
Exposure: A White-Box Photo Post-Processing Framework

Yuanming Hu[1,2] Hao He[1,2] Chenxi Xu[1,3] Baoyuan Wang[1] Stephen Lin[1]
[1]Microsoft Research [2]MIT CSAIL [3]Peking University


EIFFEL TOWER RESTAURANT
Eiffel Tower Wine &

Post-processing
a.k.a. retouching

"Magic"




EIFFEL TOWER RESTAURANT
Eiffel Tower Wine &

Exposure + 2.40


Eiffel Tower Wine &

Highlight -78


EIFFEL TOWER RESTAURANT
Eiffel Tower Wine &

White balance

Temperature 2600
Tint +23


Eiffel Tower Restaurant
Eiffel Tower Wine & Spirits


Eiffel Tower Restaurant
Eiffel Tower Wine & Spirits

Clarity + 63


Eiffel Tower Restaurant
Eiffel Tower Wine & Spirits

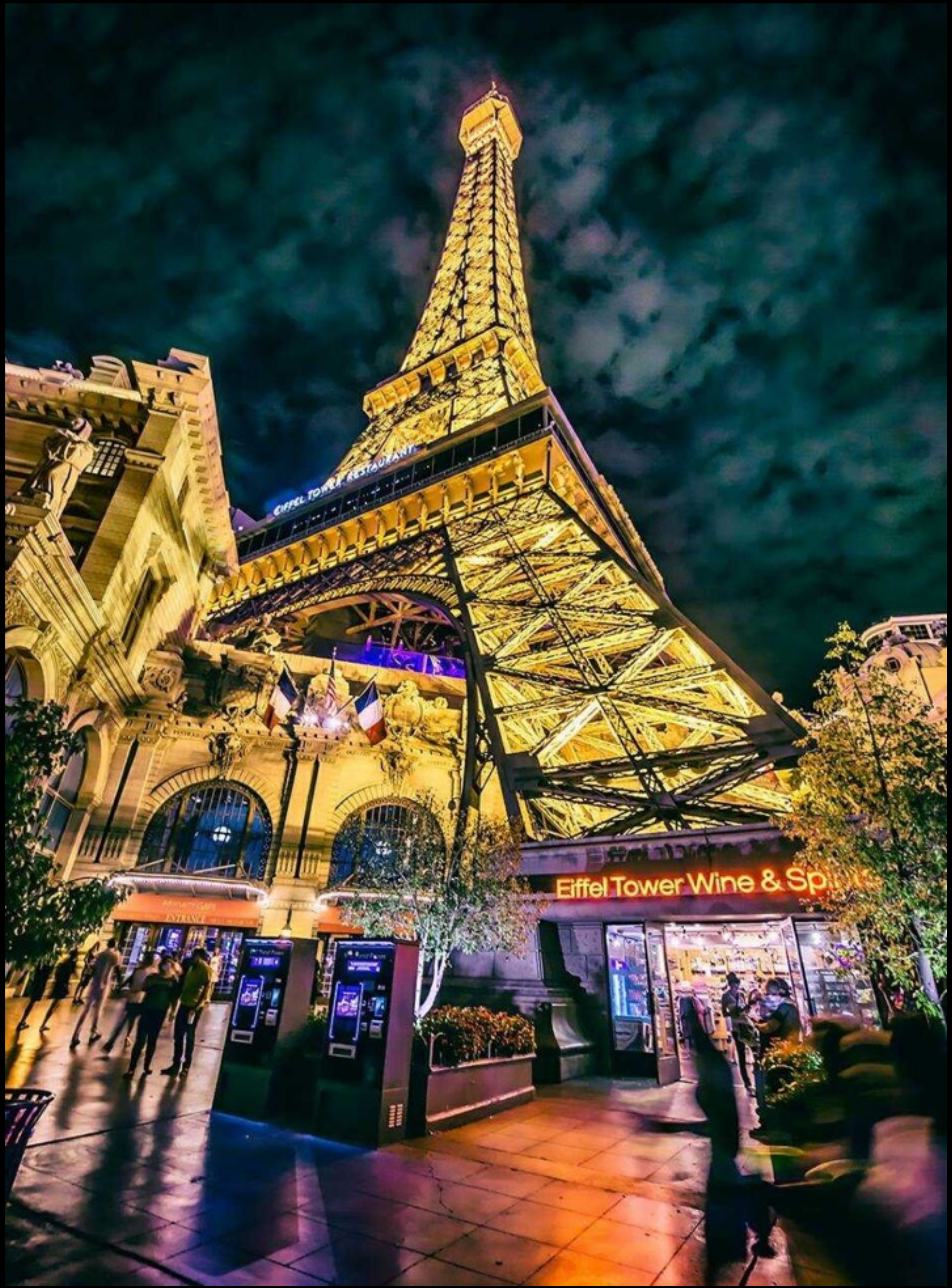

EIFFEL TOWER RESTAURANT
Eiffel Tower Wine &

Vibrance +75

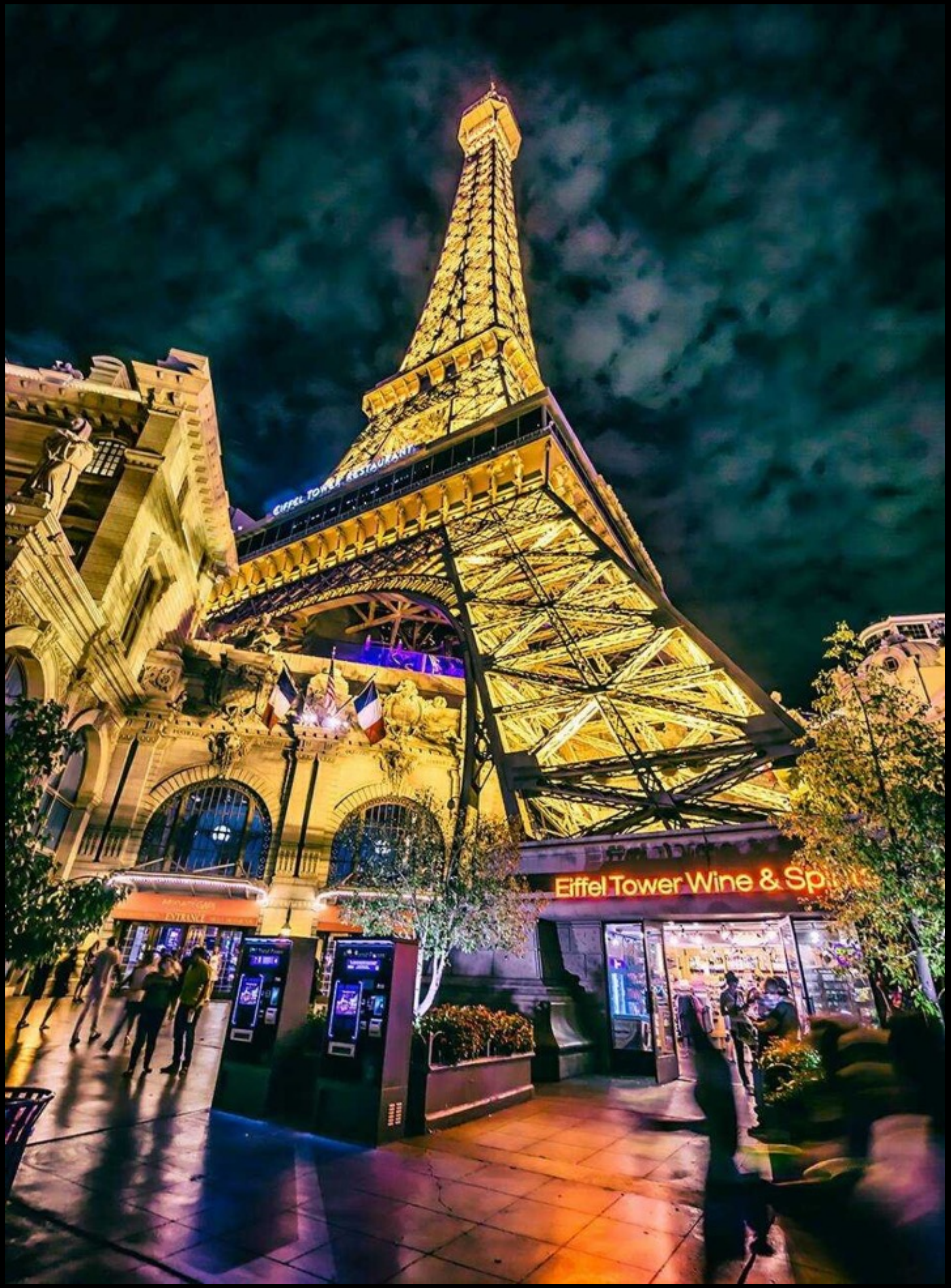

EIFFEL TOWER RESTAURANT
Eiffel Tower Wine &

Shadow + 70

A few
more steps…



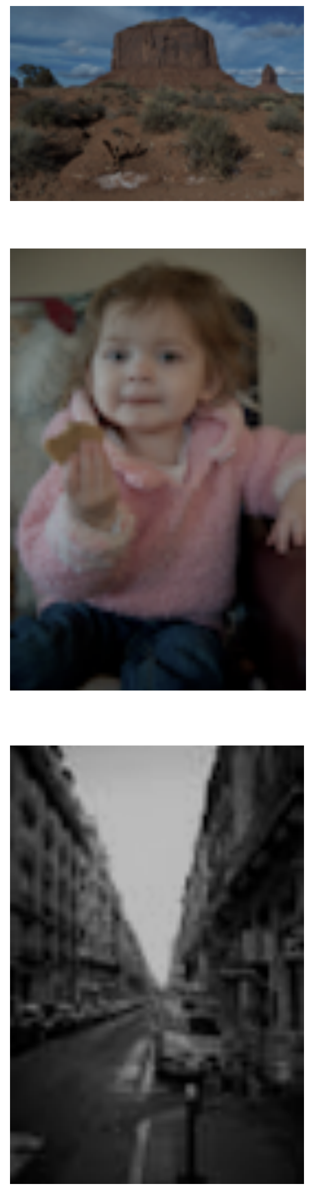
A few
more steps…



Automatic Photo Post-Processing

Training Dataset

Raw



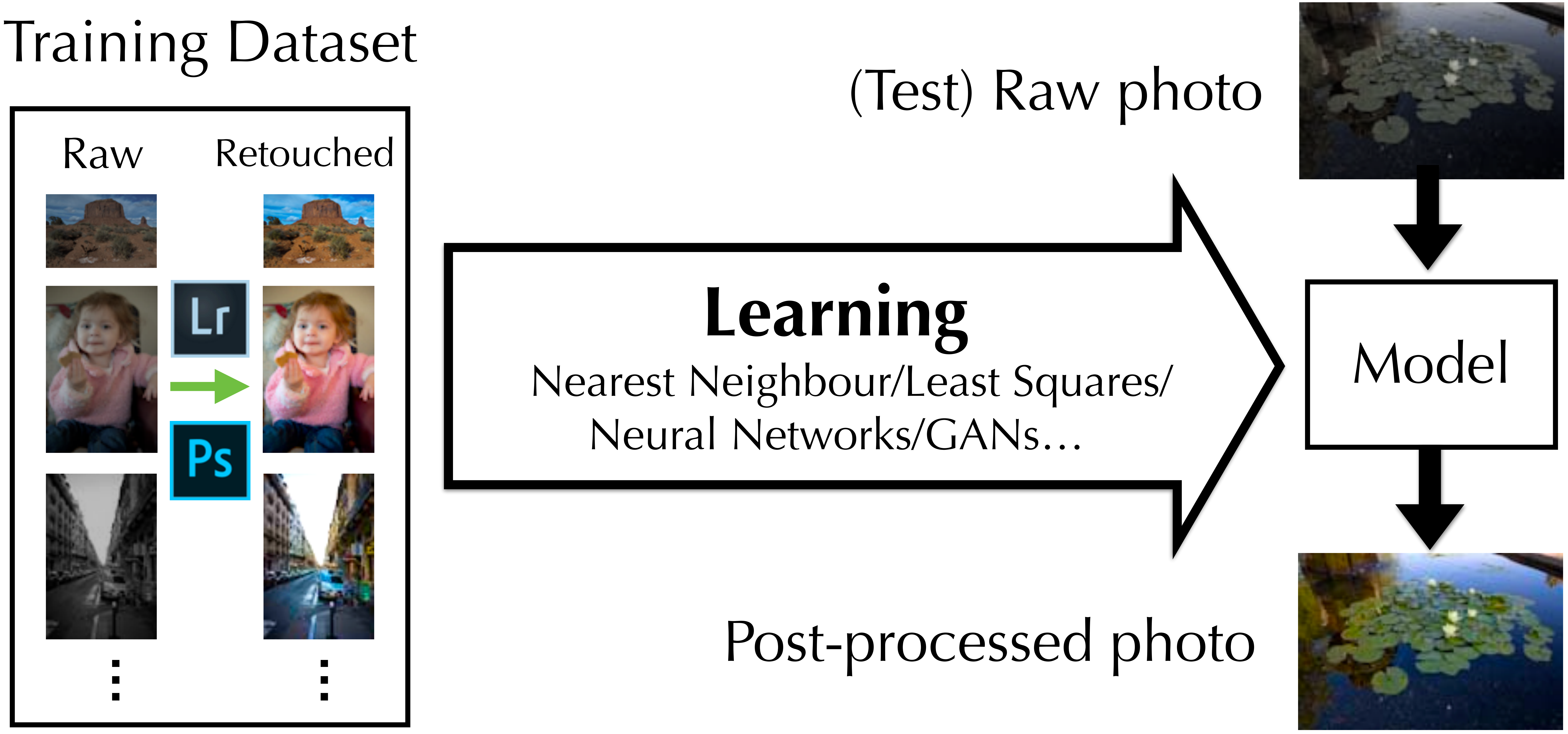
⋮

Automatic Photo Post-Processing

Training Dataset

Automatic Photo Post-Processing

Automatic Photo Post-Processing



Learning-based Photo Post-Processing

Bychkovsky et al. 2011, **Learning Photographic Global Tonal Adjustment with a Database of Input / Output Image Pairs**

MIT-Adobe FiveK Dataset

x5000

+

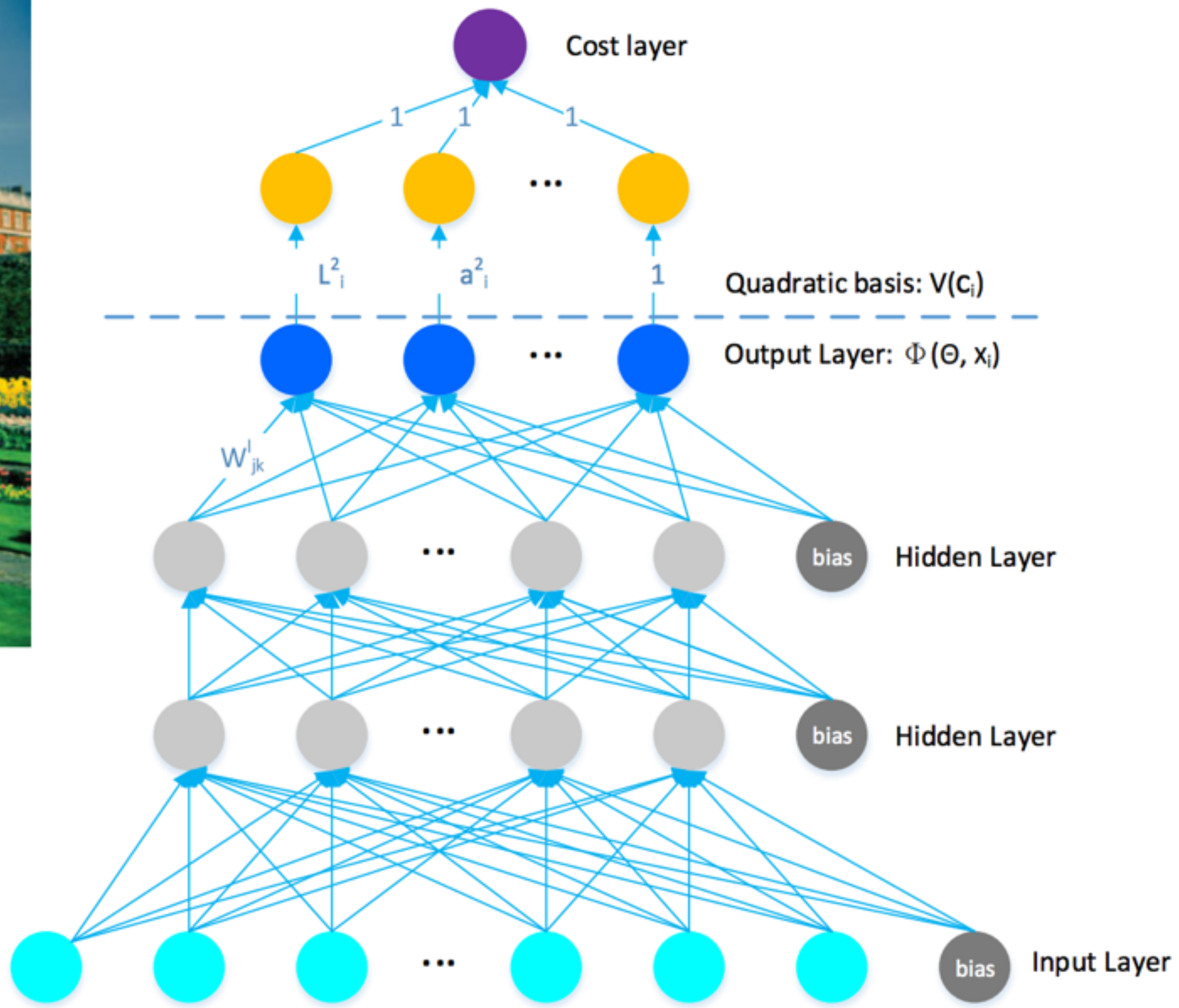
Learning-based Global Tonal Adjustment

Learning-based Photo Post-Processing

Yan et al. 2014, **Automatic Photo Adjustment Using Deep Neural Networks**

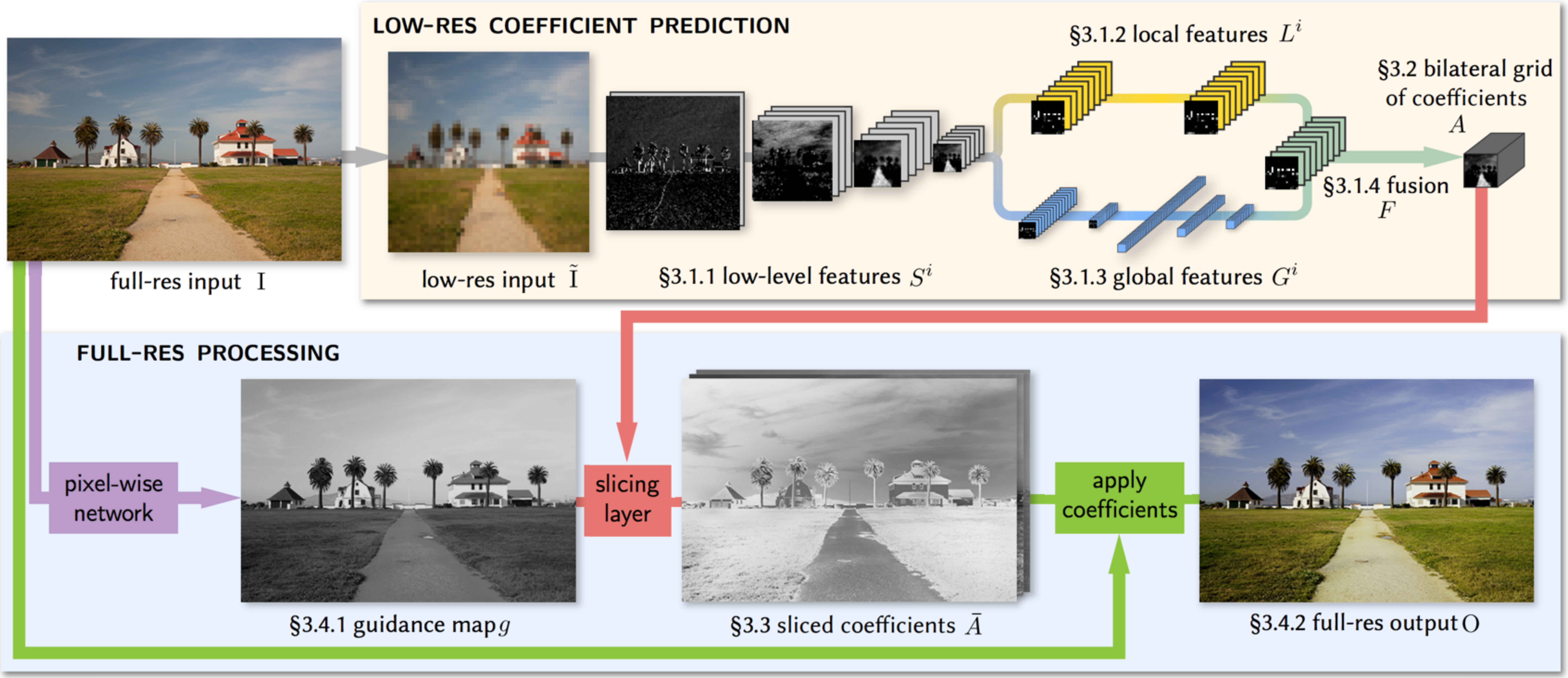
(a) (b)

Input Output

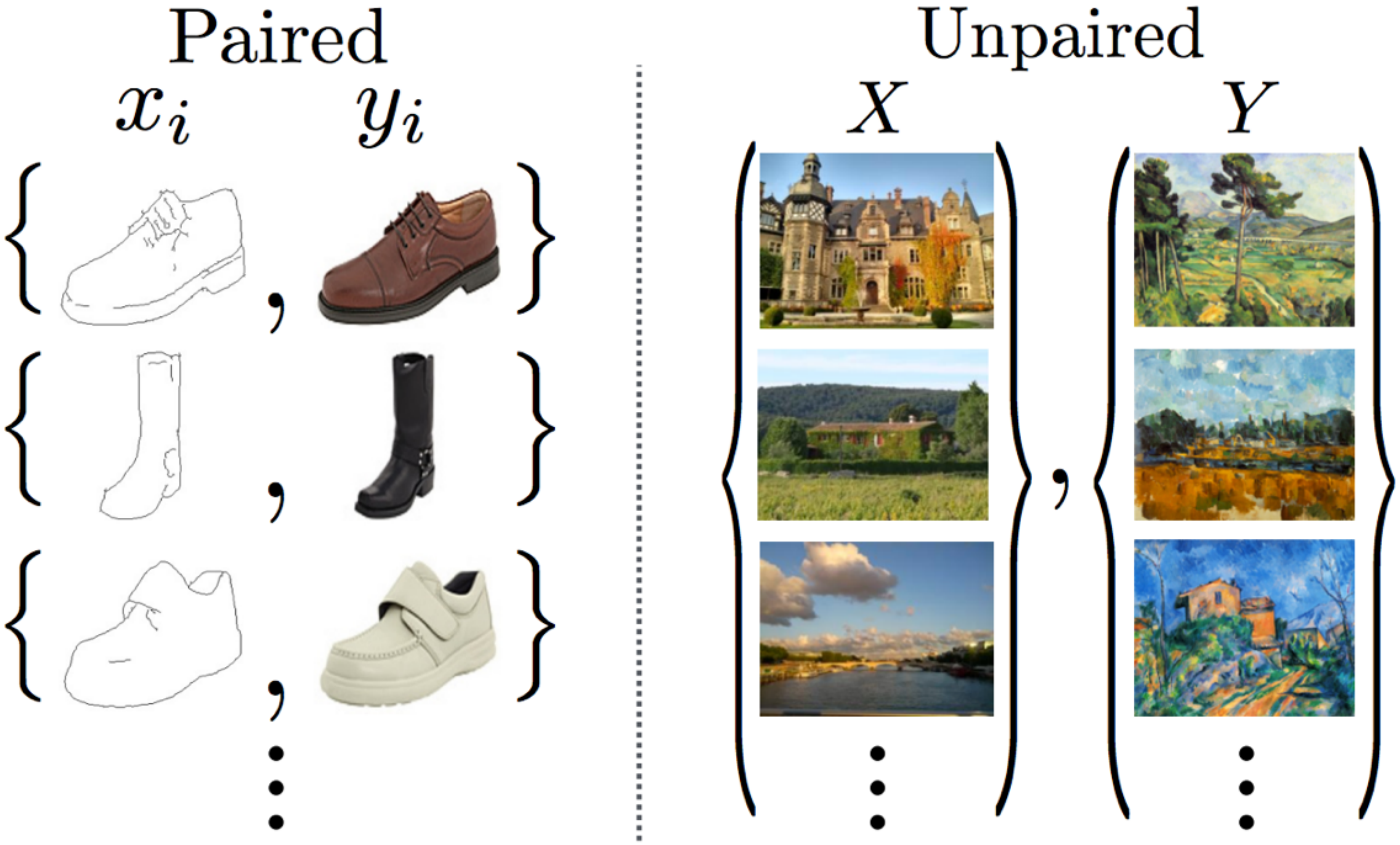


Learning-based Photo Post-Processing

Gharbi et al., **Deep Bilateral Learning for Real-Time Image Enhancement**



Paired v.s. Unpaired Image Translation

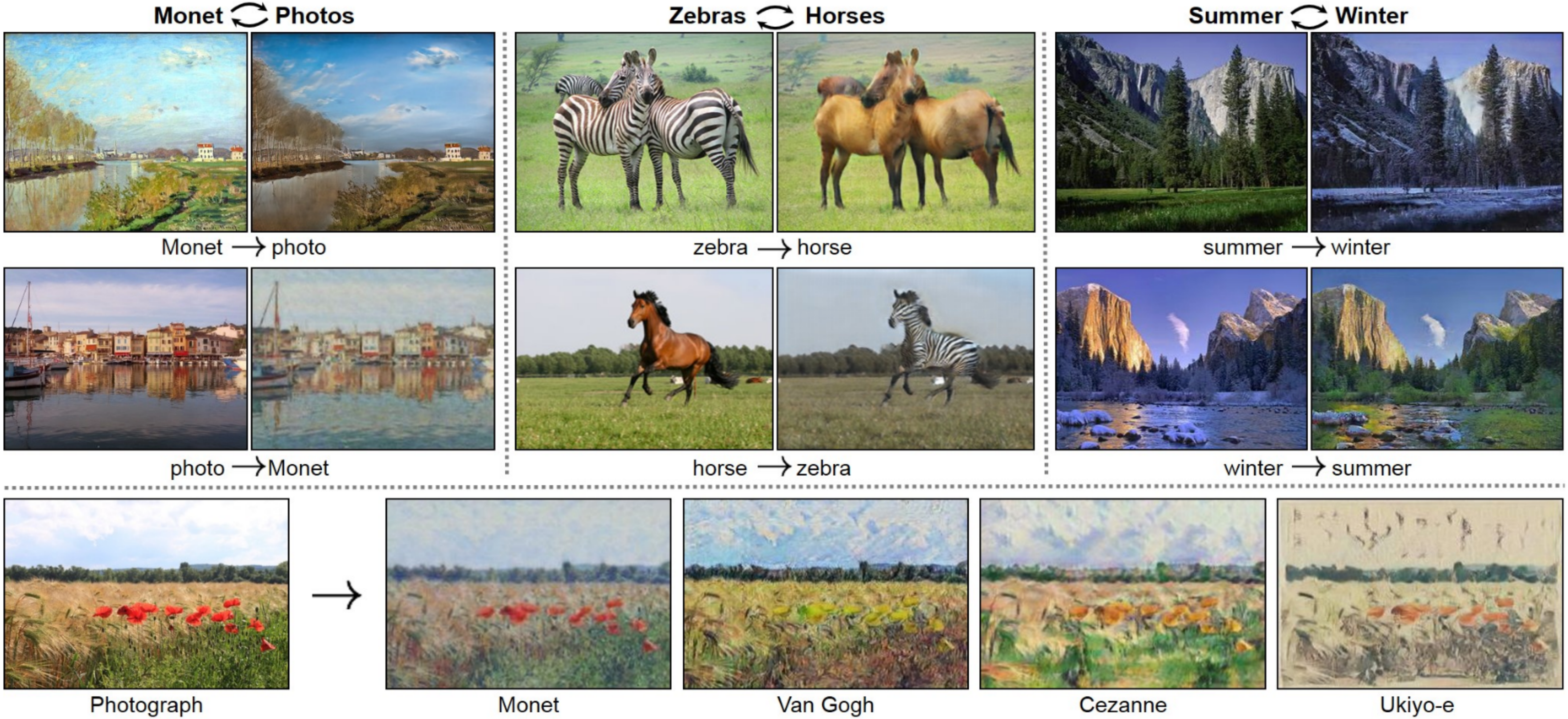


[Isola et al. 2017, **Image-to-Image Translation with Conditional Adversarial Networks**]

[Zhu et al. 2017, **Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks**]

CycleGAN

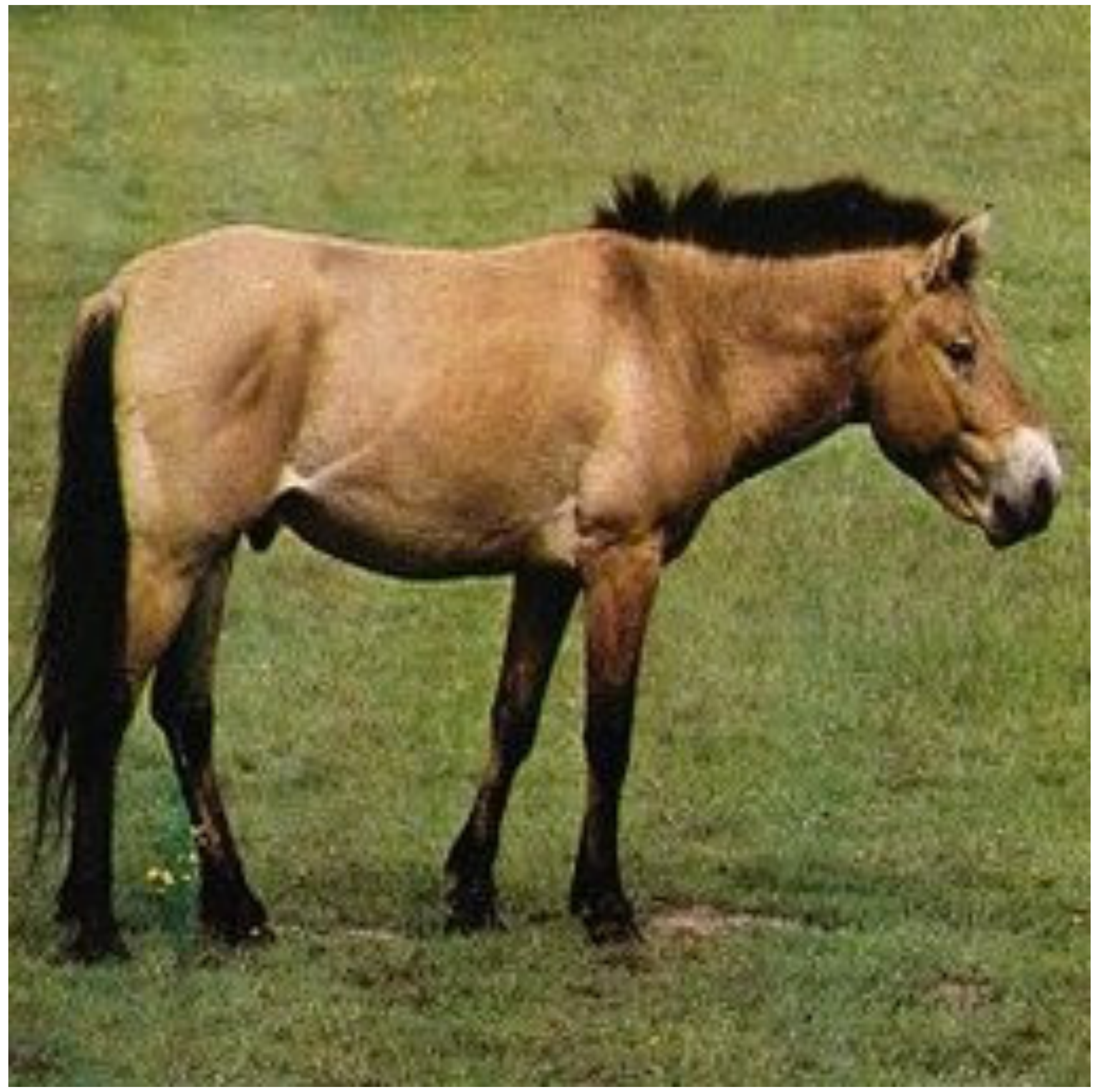
[Zhu et al. 2017, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks]

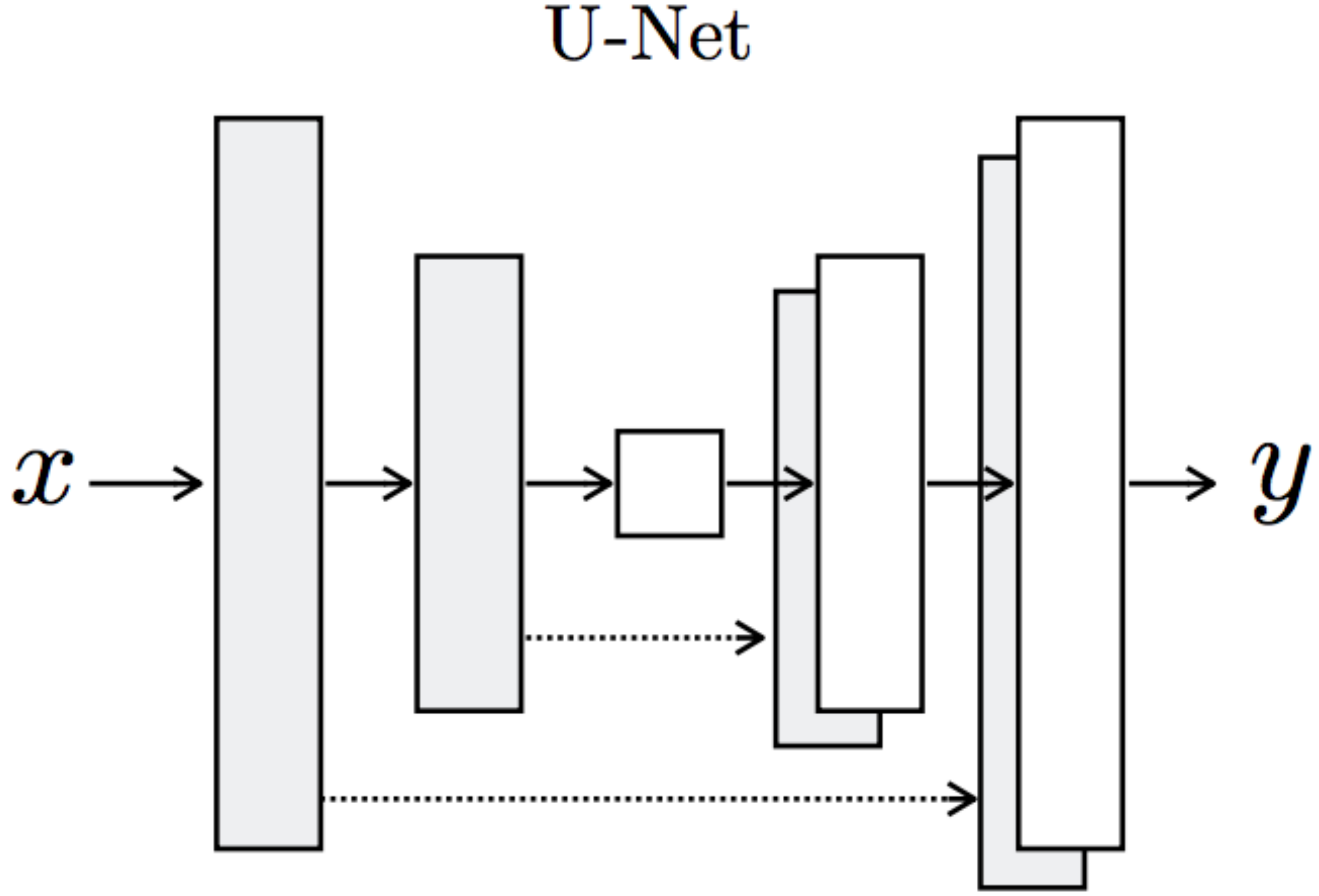


Generator

Encoder/
decoder-based
CNN

256x256 px

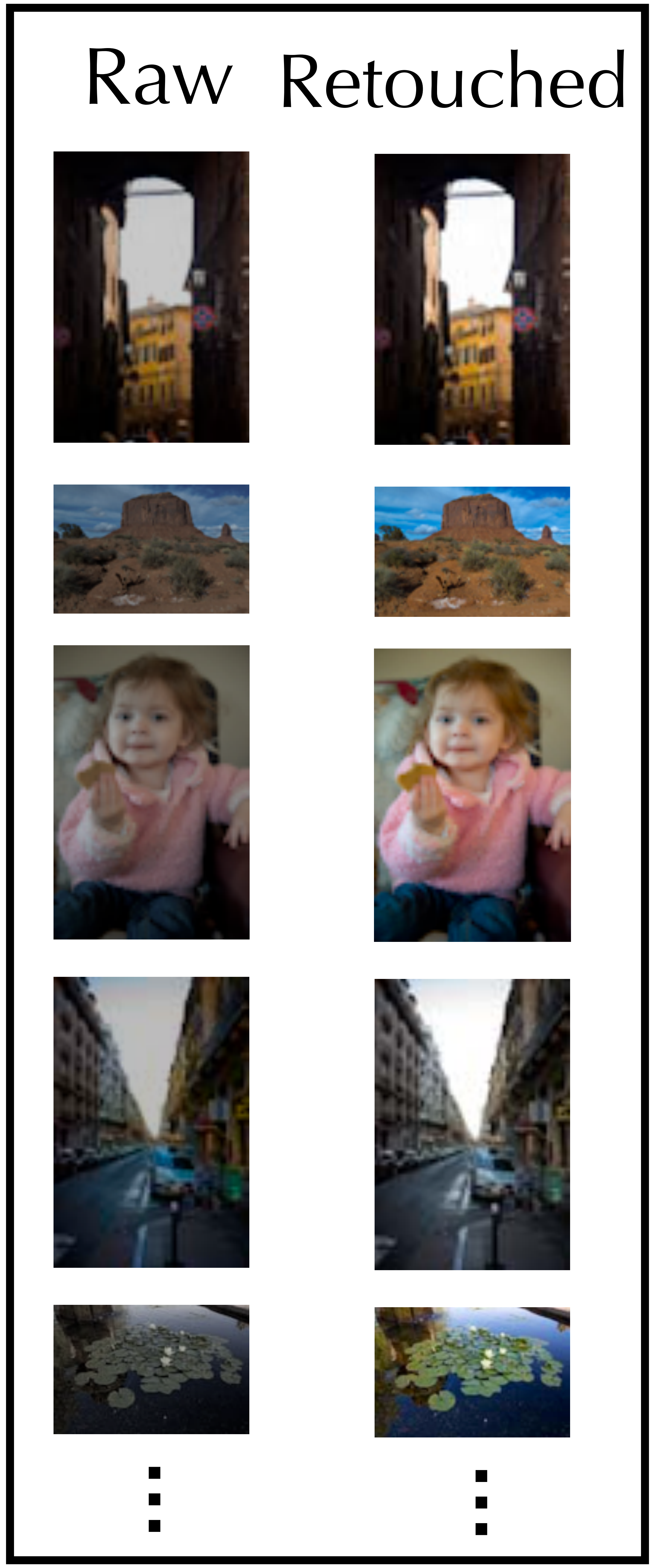




256x256 px

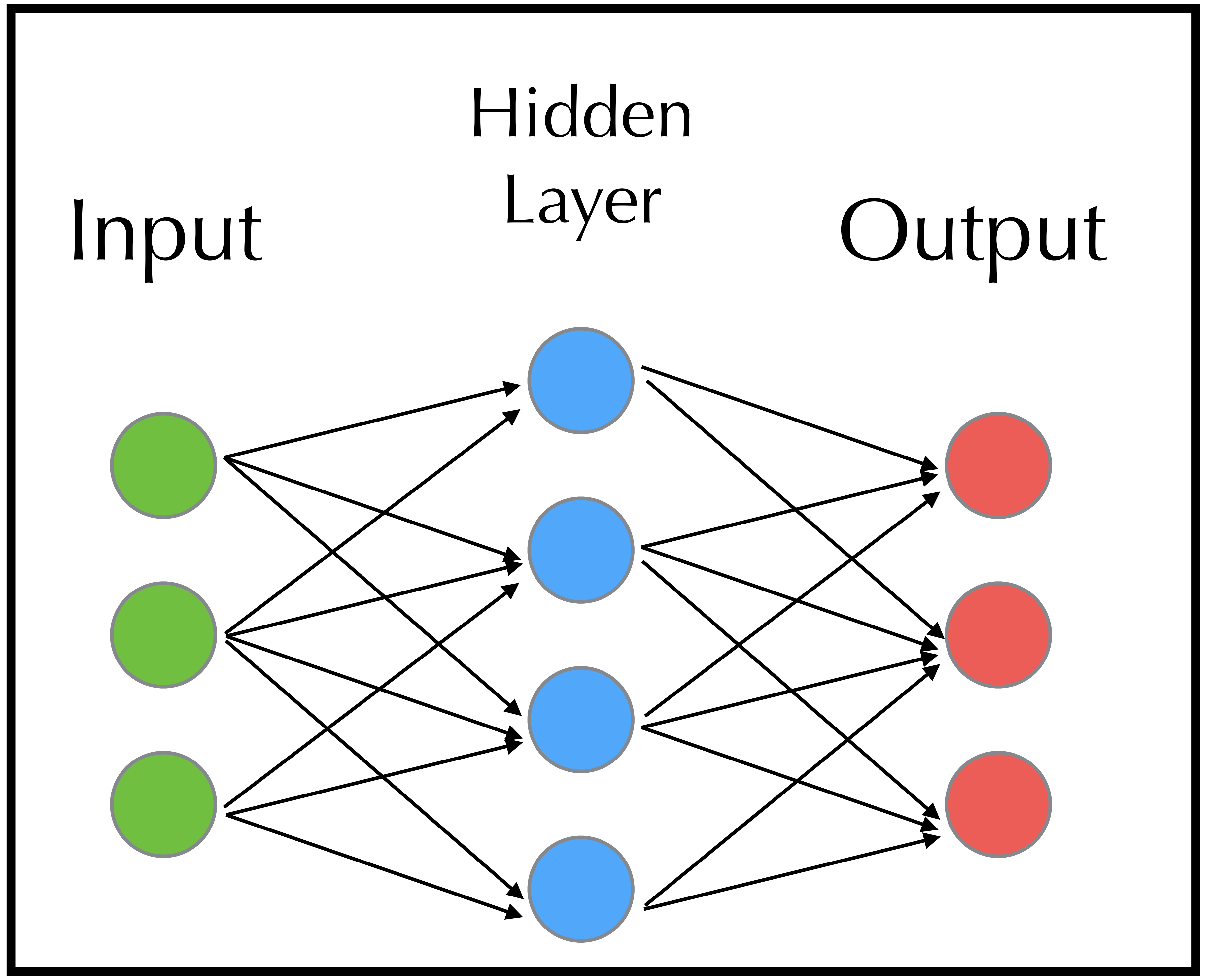
[Zhu et al. 2017, Unpaired Image-to-Image Translation
using Cycle-Consistent Adversarial Networks]

Dataset



Deep learning

Deep neural networks



Traditional deep-learning approaches generate **black boxes** (CNNs) out of existing ones (datasets).

To understand the magic of photo retouching, we need a **white box** result.

	High Resolution	Human Understandable	Unpaired Training	End-to-end Processing
Tonal Adjustment Learning Bychkovsky et al. 2011	○	○		
Color transform learning Yan et al.	○			○
Deep Bilateral Learning Gharbi et al.	○			○
CycleGAN, Zhu et al.			○	○

	High Resolution	Human Understandable	Unpaired Training	End-to-end Processing
Tonal Adjustment Learning Bychkovsky et al. 2011	○	○		
Color transform learning Yan et al.	○			○
Deep Bilateral Learning Gharbi et al.	○			○
CycleGAN, Zhu et al.			○	○
Exposure (ours)	○	○	○	○

Our Approach

Differentiable Photo Postprocessing Model

Deep Reinforcement Learning

Generative Adversarial Networks

Our Approach

Differentiable Photo Postprocessing Model

Deep Reinforcement Learning

Generative Adversarial Networks

Modelling

Our Approach

Differentiable Photo Postprocessing Model	Deep Reinforcement Learning	Generative Adversarial Networks

Modelling

Optimization

Our Approach

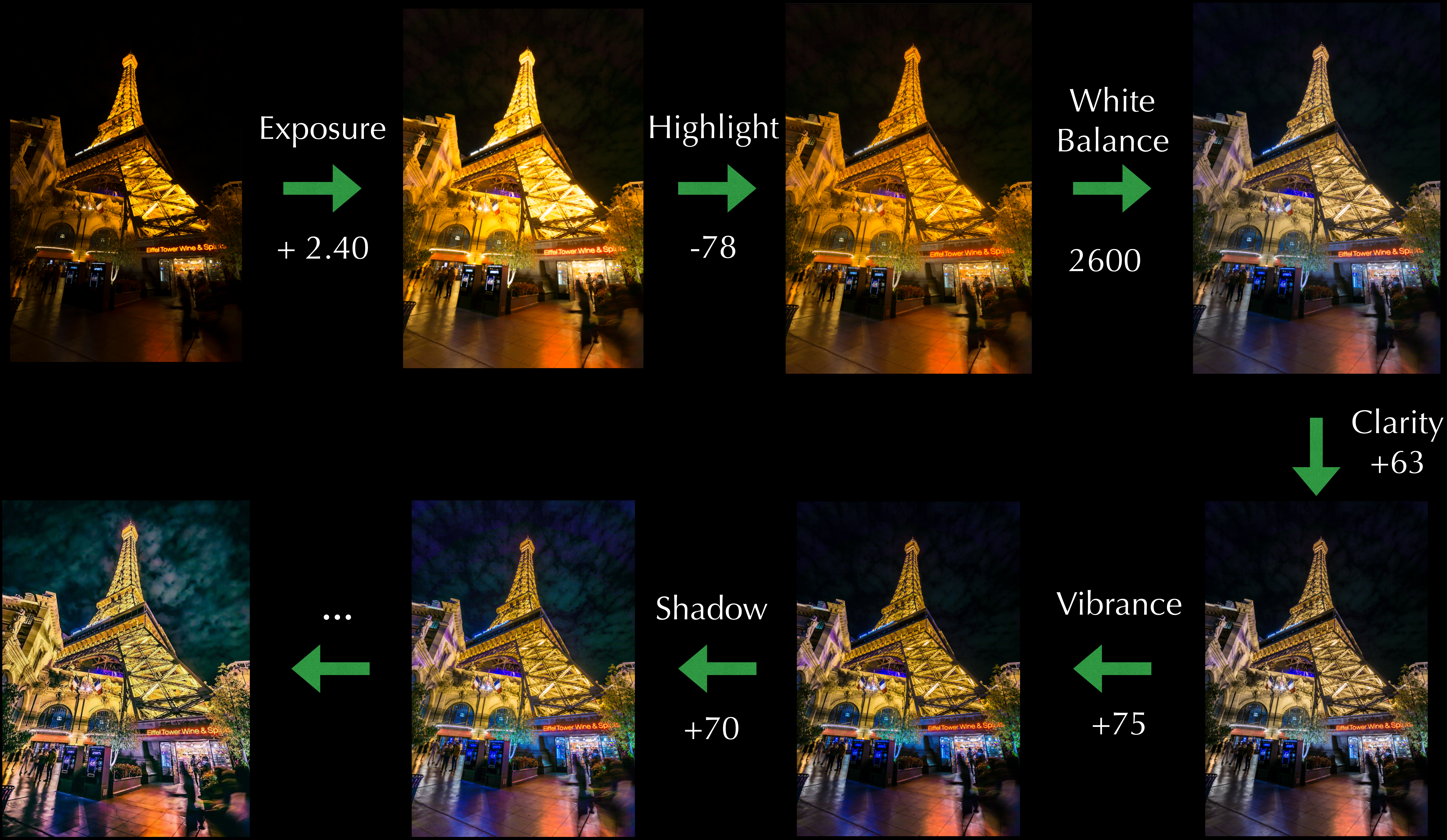
Differentiable Photo Postprocessing Model

Deep Reinforcement Learning

Generative Adversarial Networks

Modelling

Optimization


Exposure
+ 2.40
Highlight
-78
White
Balance
2600
Clarity
+63
Vibrance
+75
Shadow
+70
...


Exposure
+ 2.40
Highlight
-78
White Balance
2600
Pick a image operation, estimate the parameter. Repeat until done.
Clarity
+63
Vibrance
+75
Shadow
+70
...

Image Operations (Filters)

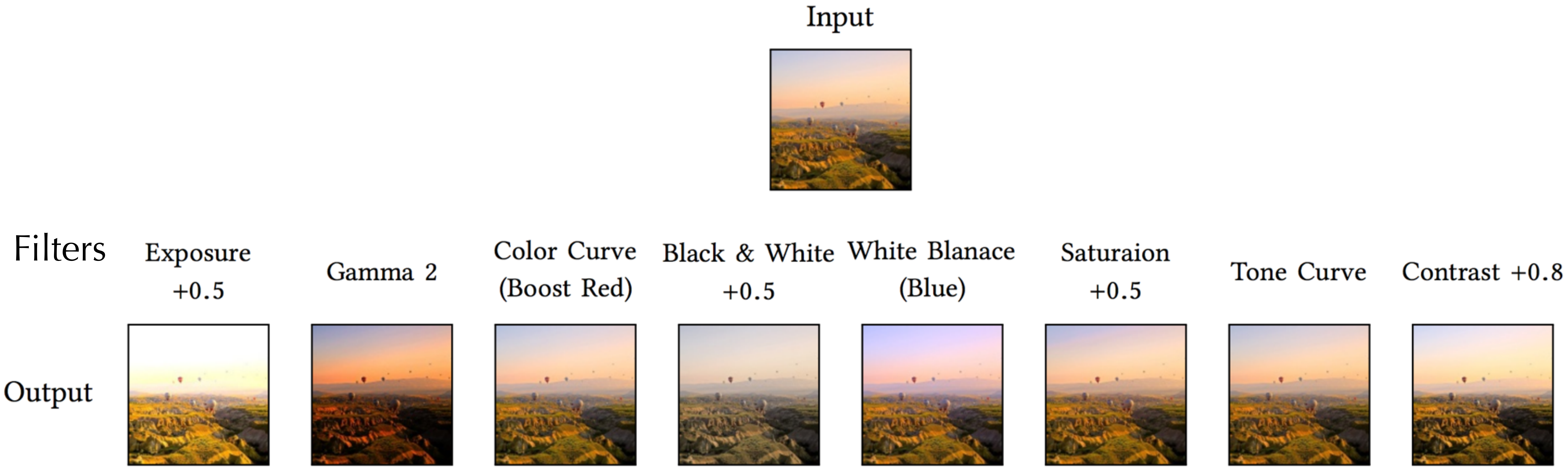


Image Operations (Filters)

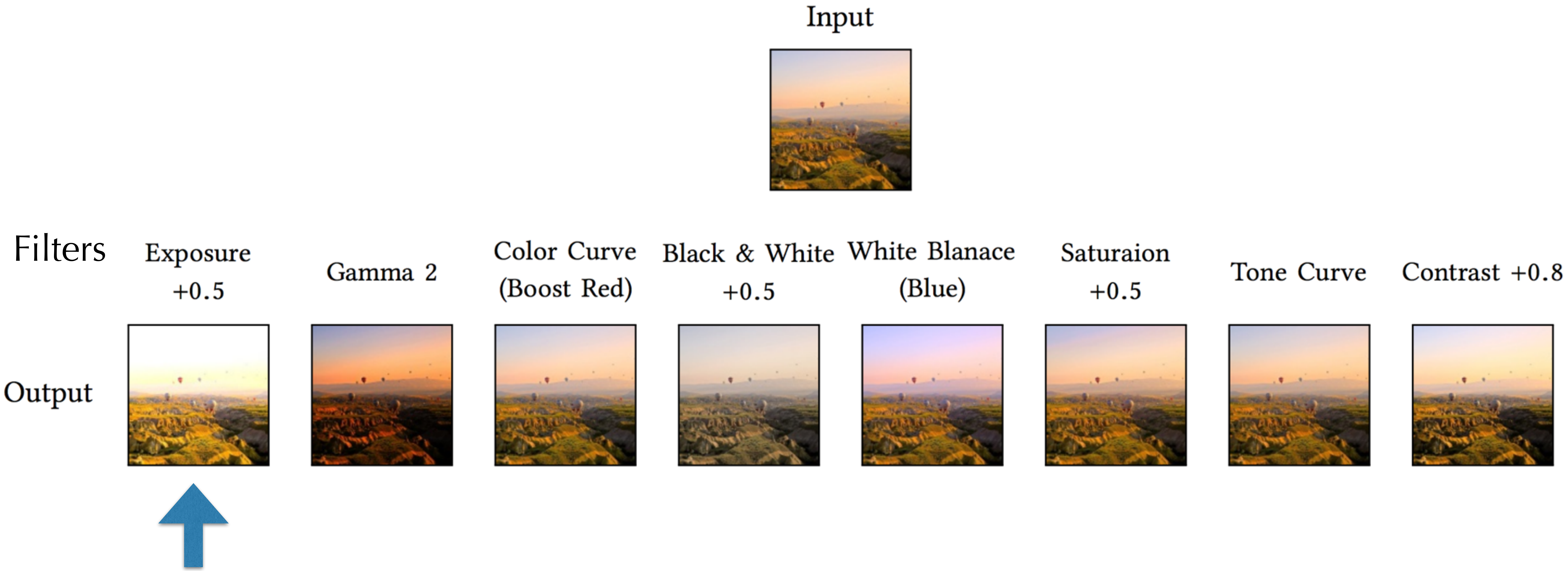


Image Operations (Filters)

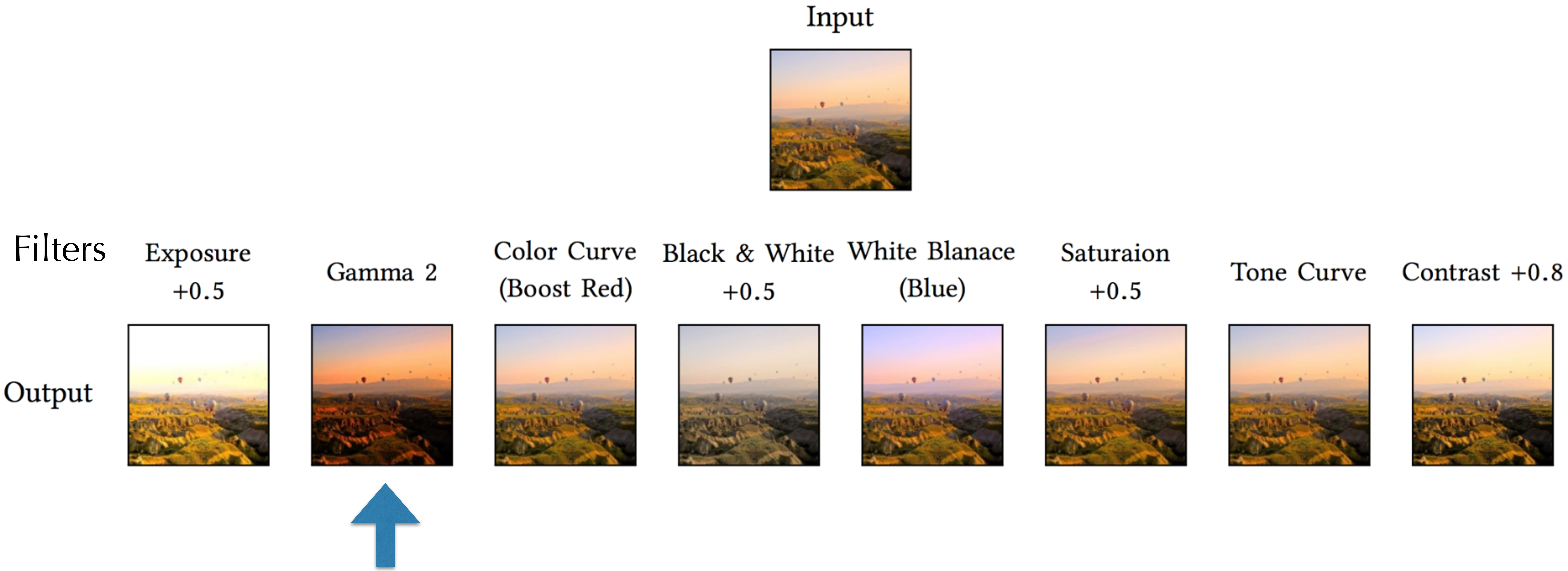


Image Operations (Filters)

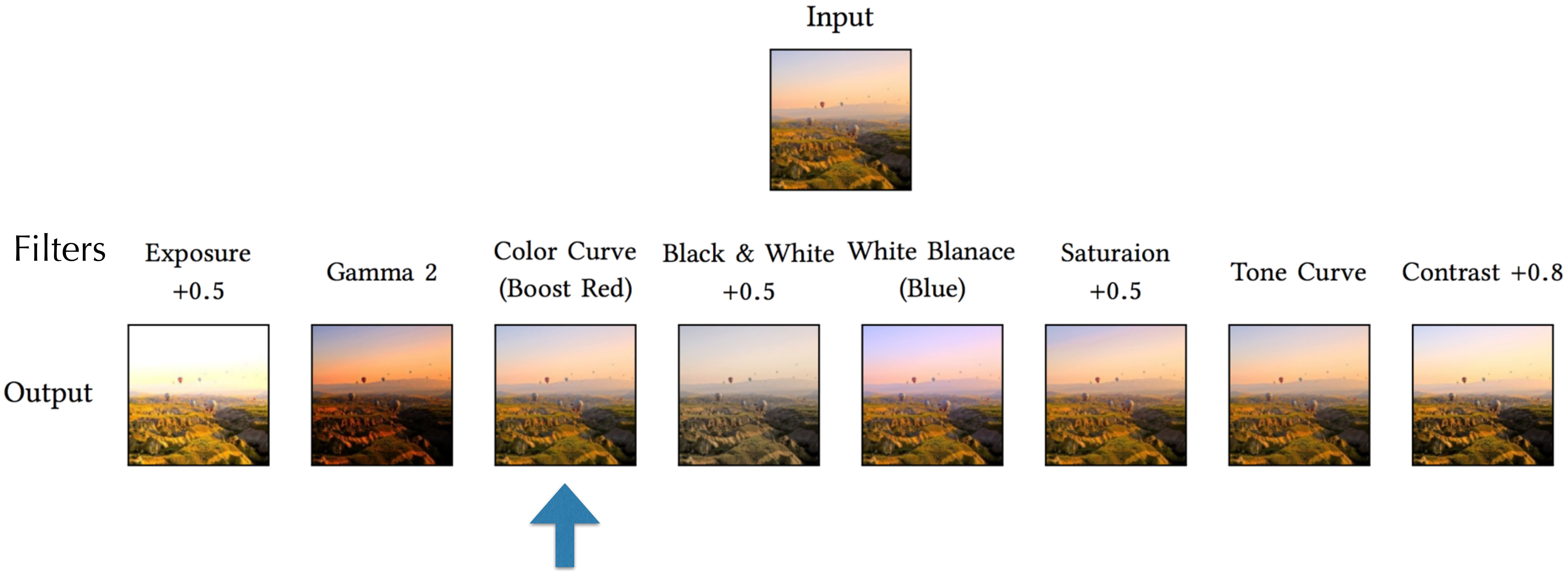
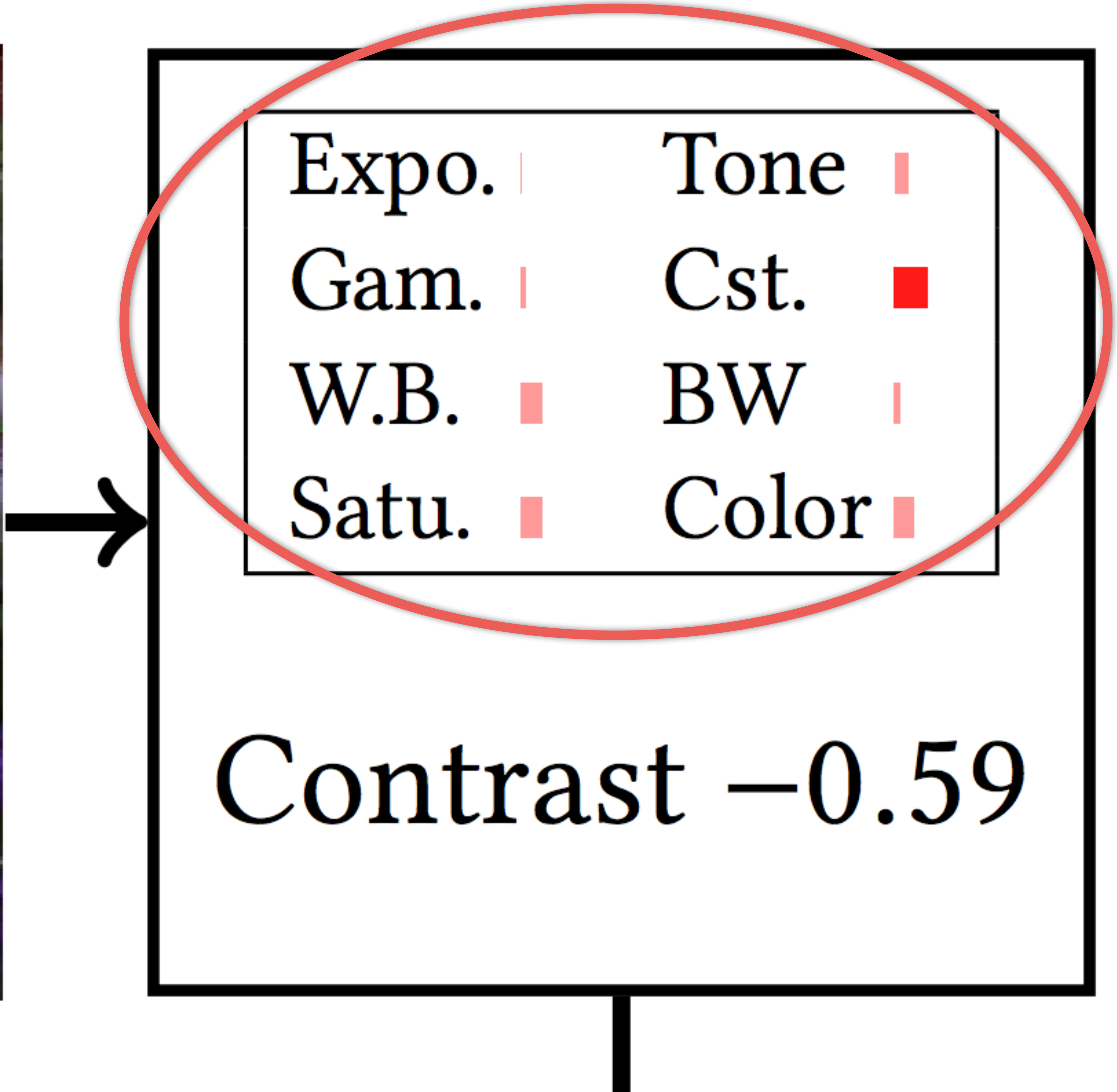


Image Operations (Filters)

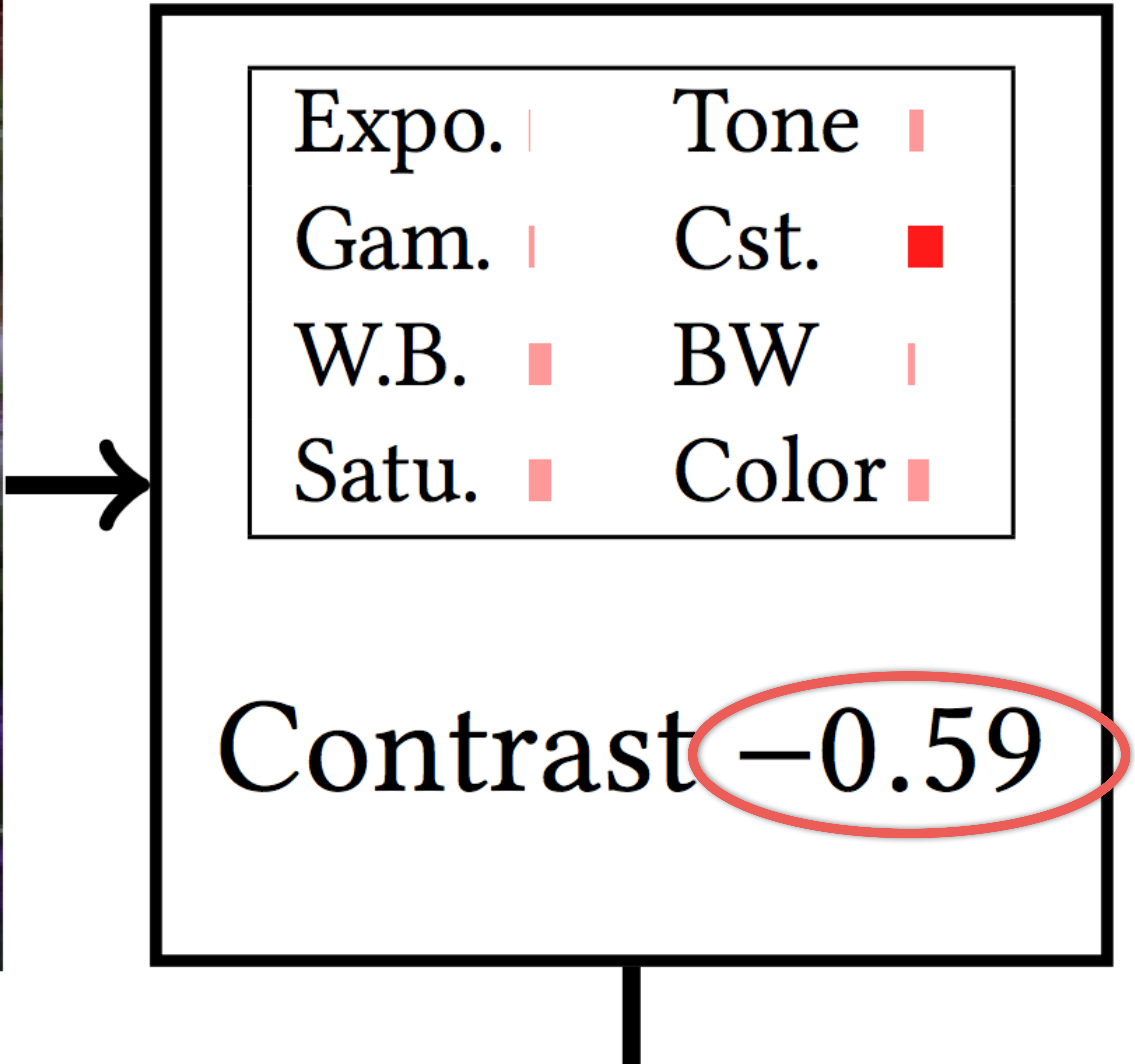
Image Operations (Filters)

Modelling Photo Post-Processing

Modelling Photo Post-Processing



Modelling Photo Post-Processing



Modelling Photo Post-Processing

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Generative Adversarial Networks

$\underbrace{\hspace{3em}}_{\text{Modelling}}$ $\underbrace{\hspace{6em}}_{\text{Optimization}}$

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Generative Adversarial Networks

Modelling

Optimization

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Generative Adversarial Networks

Modelling

Optimization

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Generative Adversarial Networks

Modelling

Optimization

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

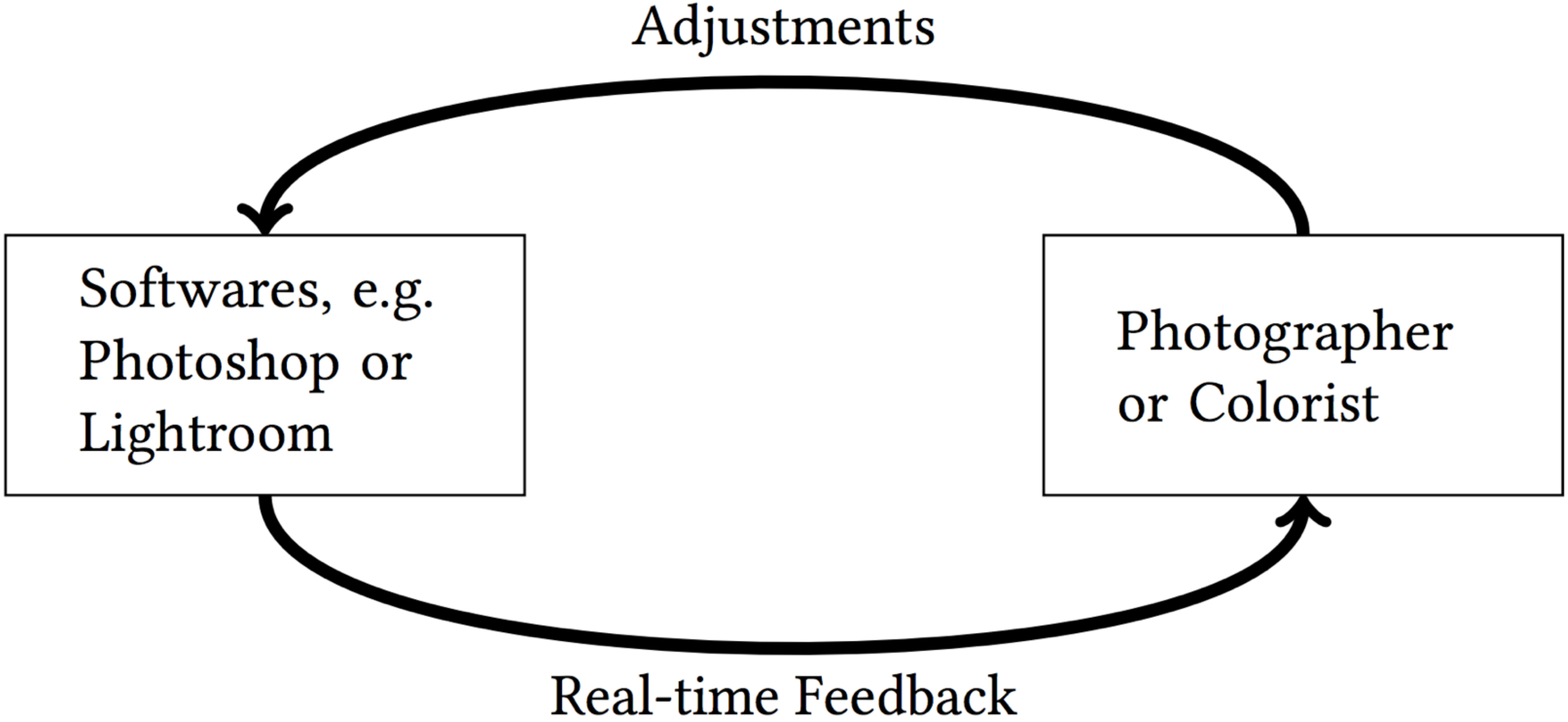
Deep Reinforcement Learning

Generative Adversarial Networks

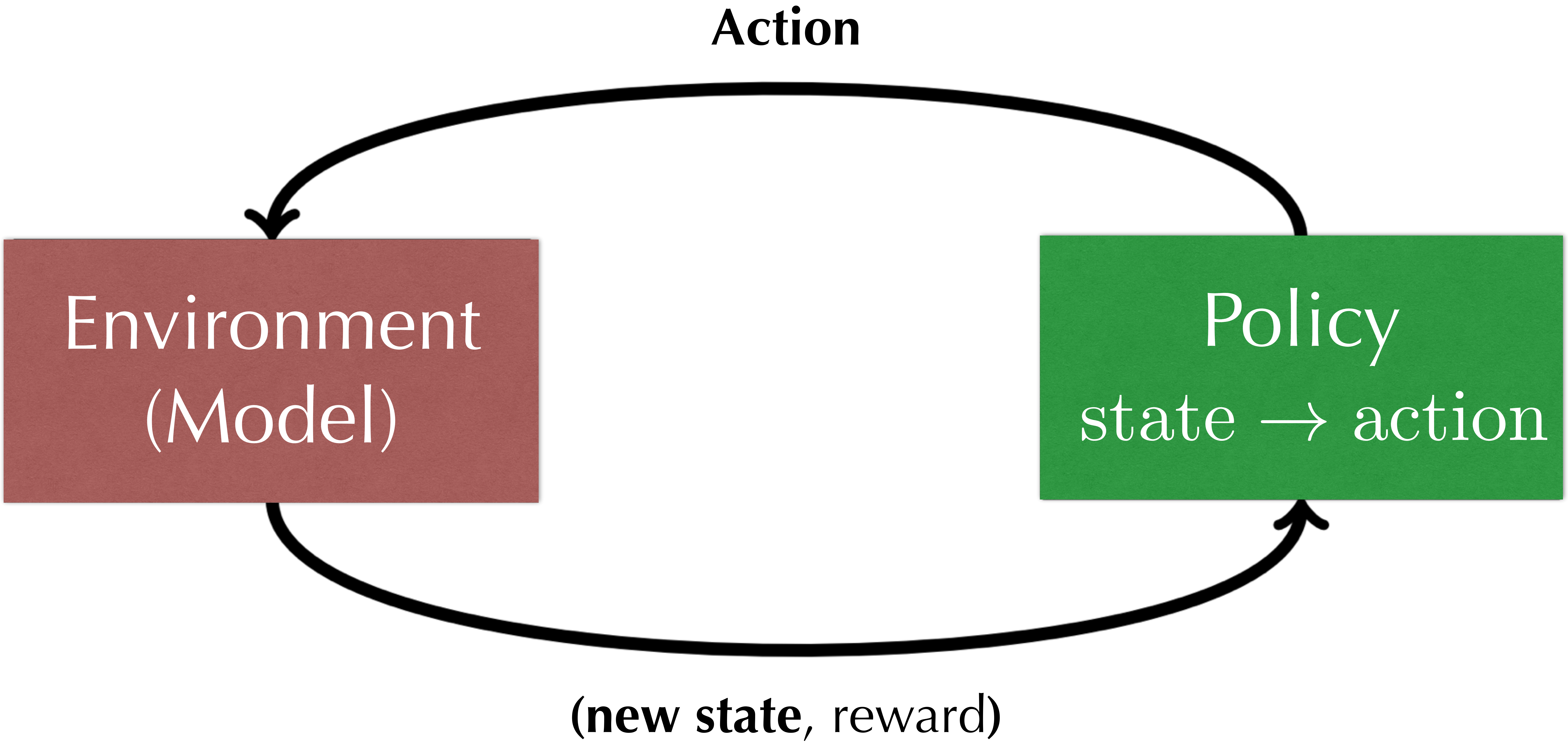
Modelling

Optimization

Learning to Make Decisions



Learning to Make Decisions



Learning to Make Decisions

Learning to Make Decisions

Gradient-based Policy Optimization

Monte-Carlo Estimation of (Stochastic) Policy Gradient

Loss Policy Value

$$\nabla_{\theta_1} J(\pi_\theta) = \mathop{\mathbb{E}}_{\substack{s\sim\rho^\pi \\ a_1\sim\pi_1(s) \\ a_2=\pi_2(s,a_1)}} \left[\nabla_{\theta_1} \log \pi_1(a_1|s) Q(s,(a_1,a_2))\right]$$

NN weights

$a_1\sim\pi_1(s)$ ← **Discrete policy (filter selection)**

$a_2=\pi_2(s,a_1)$ ← **Continuous (deterministic) policy parameter estimation**

Gradient-based Policy Optimization

Deterministic Policy Gradient Theorem

Loss Policy Value

$$\nabla_{\theta_2} J(\pi_\theta) = \mathop{\mathbb{E}}_{\substack{s\sim\rho^\pi \\ a_2=\pi_2(s,a_1)}} \left[\nabla_{a_2} Q\left(s,(a_1,a_2)\right) \nabla_{\theta_2} \pi_2(s,a_1)\right]$$

NN weights

Continuous (deterministic) policy parameter estimation

Our Approach:
Learn **image operations**, instead of **pixels**.

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**,
instead of **pixels**

Generative Adversarial Networks

$\underbrace{\hspace{3em}}_{\text{Modelling}}$ $\underbrace{\hspace{8em}}_{\text{Optimization}}$

Our Approach:
Learn **image operations**, instead of **pixels**.

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**,
instead of **pixels**

Generative Adversarial Networks

Modelling

Optimization

Our Approach:
Learn **image operations**, instead of **pixels**.

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**,
instead of **pixels**

Generative Adversarial Networks

$\underbrace{\text{Differentiable Photo Postprocessing Model}}_{\text{Modelling}}$

$\underbrace{\text{Deep Reinforcement Learning} \quad \text{Generative Adversarial Networks}}_{\text{Optimization}}$

Modelling

Optimization

Our Approach:
Learn **image operations**, instead of **pixels**.

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**,
instead of **pixels**

Generative Adversarial Networks

Modelling

Optimization

Designing the Reward Function

$r = -D($  $,$ 

$)$

Generated

Target (i.e. “ground truth”)

(Conditional) Generative Adversarial Networks (c-GANs)

Reward: Earth Mover's Distance

Reward: Earth Mover's Distance

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**, instead of **pixels**

Generative Adversarial Networks

Reward function
Training without **pairs**

Modelling

Optimization

Our Approach

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**,
instead of **pixels**

Generative Adversarial Networks

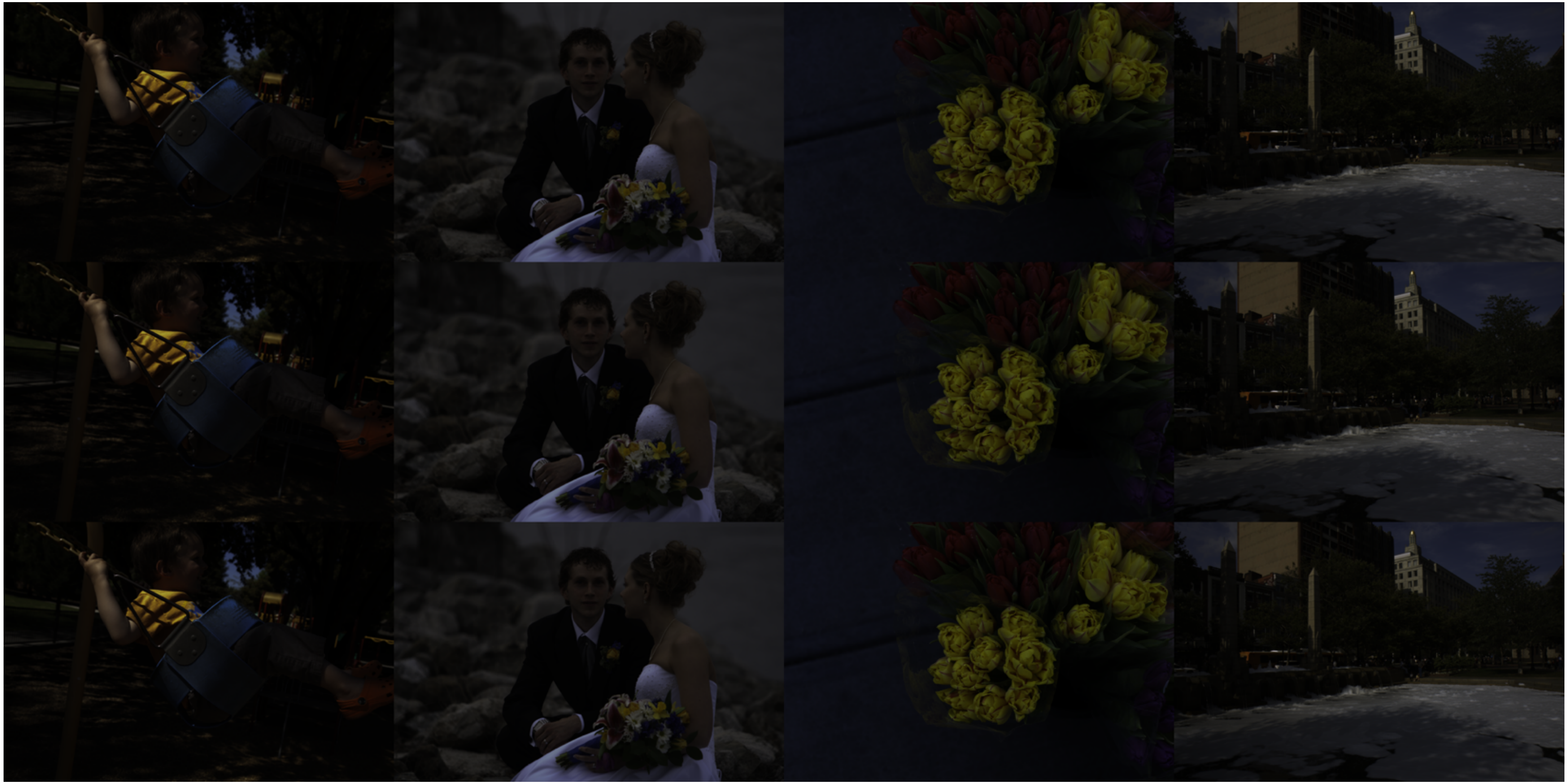
Reward function
Training without **pairs**

$\underbrace{\hspace{4em}}_{\text{Modelling}}$ $\underbrace{\hspace{8em}}_{\text{Optimization}}$

Modelling

Optimization

Results: Retouching and Stylisation



Results: Retouching and Stylisation

Results: Retouching and Stylisation

Results: Retouching and Stylisation

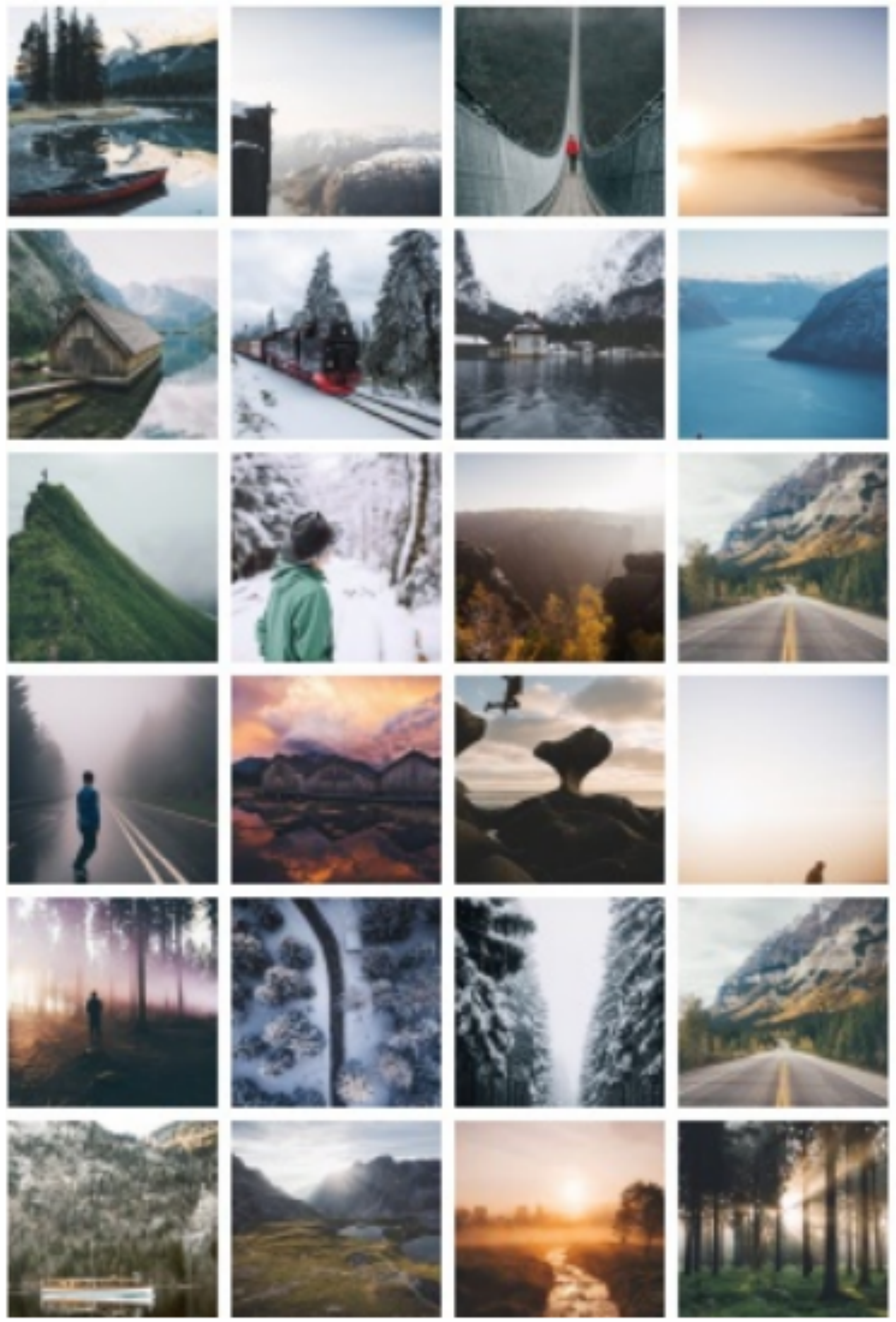
Results: Retouching and Stylisation

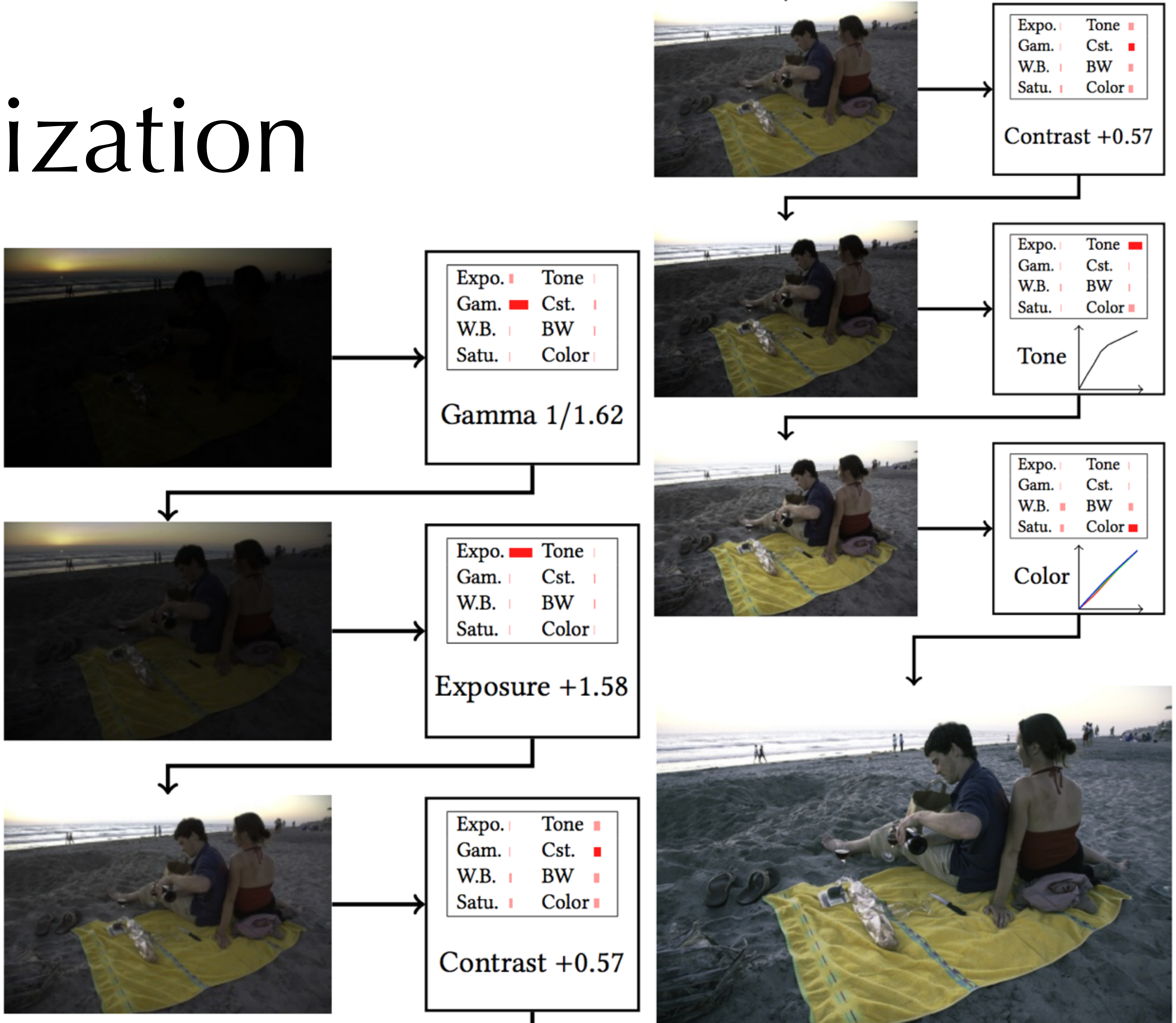
Results: Retouching and Stylisation

Results: Retouching and Stylisation

Results: Stylization

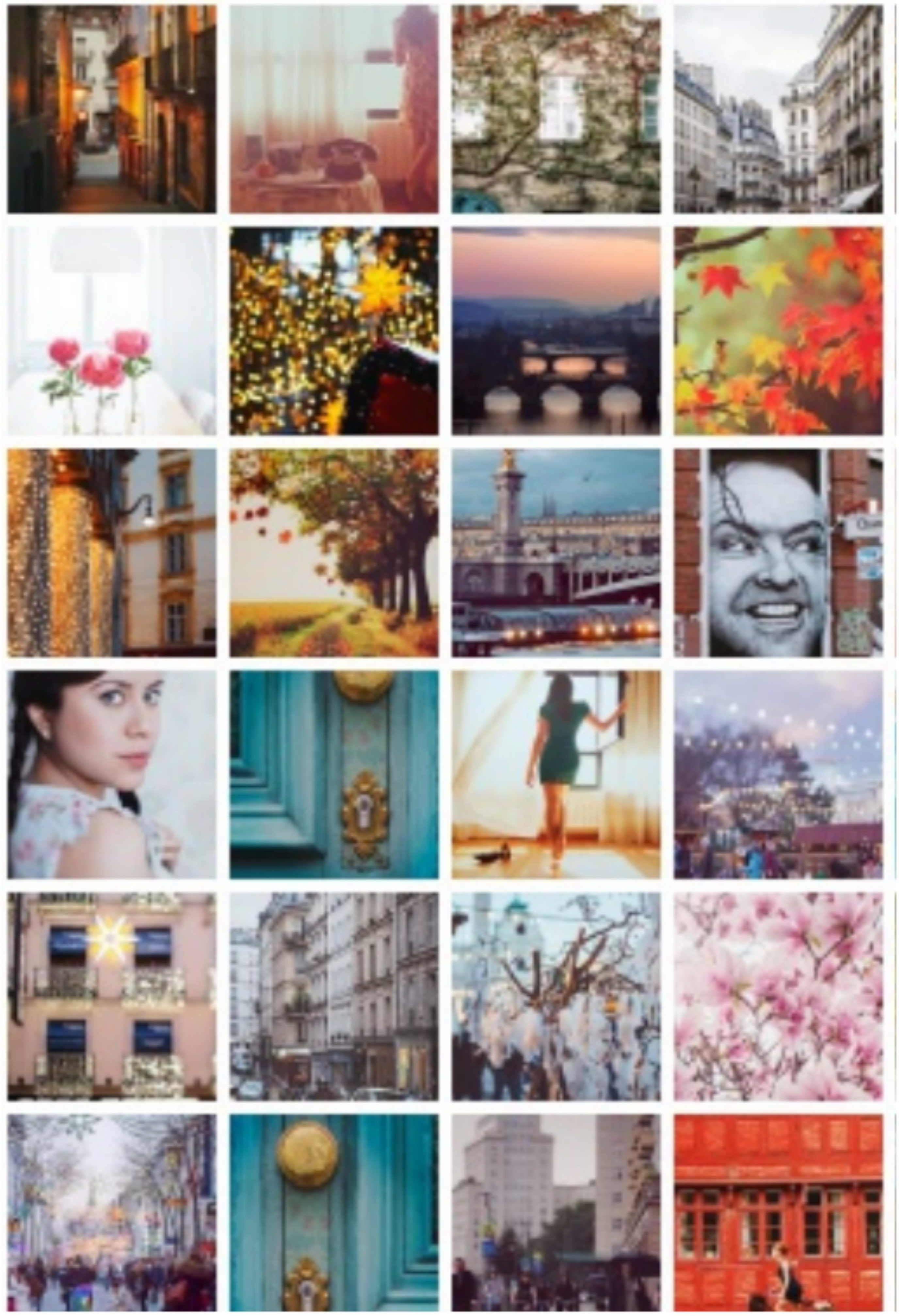
500px Artist A

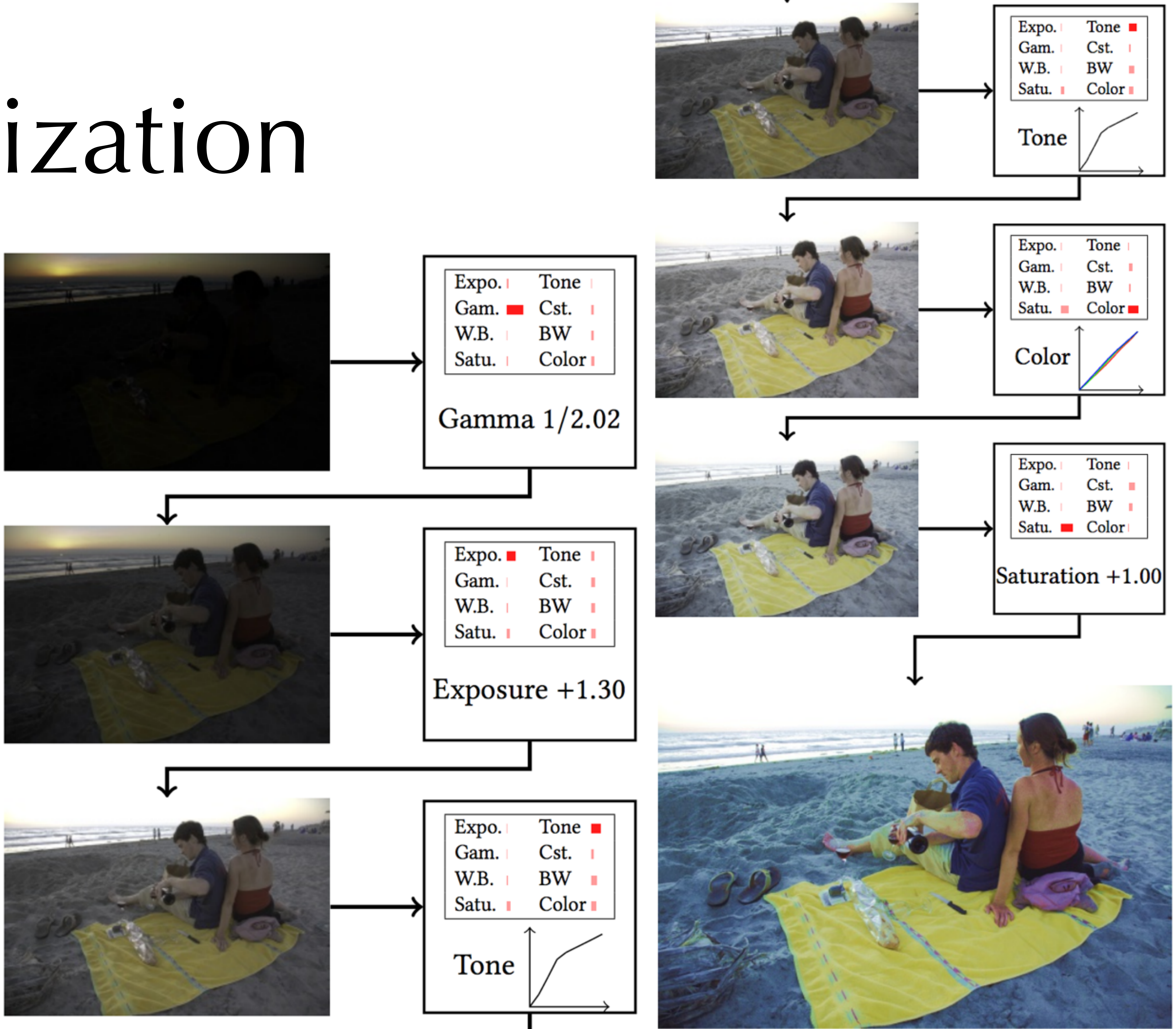




Results: Stylization

500px Artist B





An "Infinite-Resolution" GAN

An "Infinite-Resolution" GAN

An "Infinite-Resolution" GAN

Pix2pix (paired data needed)

Ours (unpaired training)

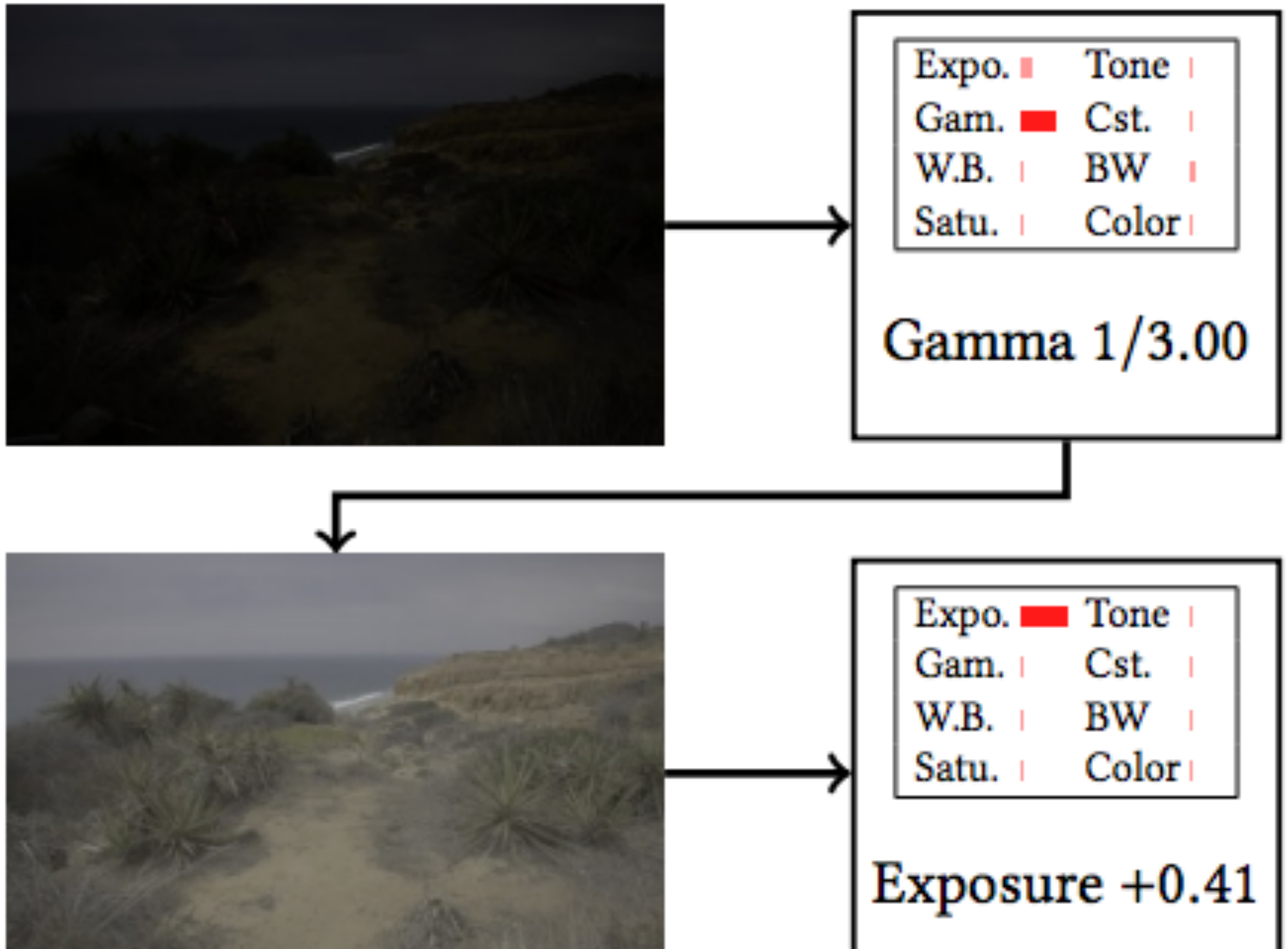
An “Infinite-Resolution” GAN

Pix2pix (paired data needed)

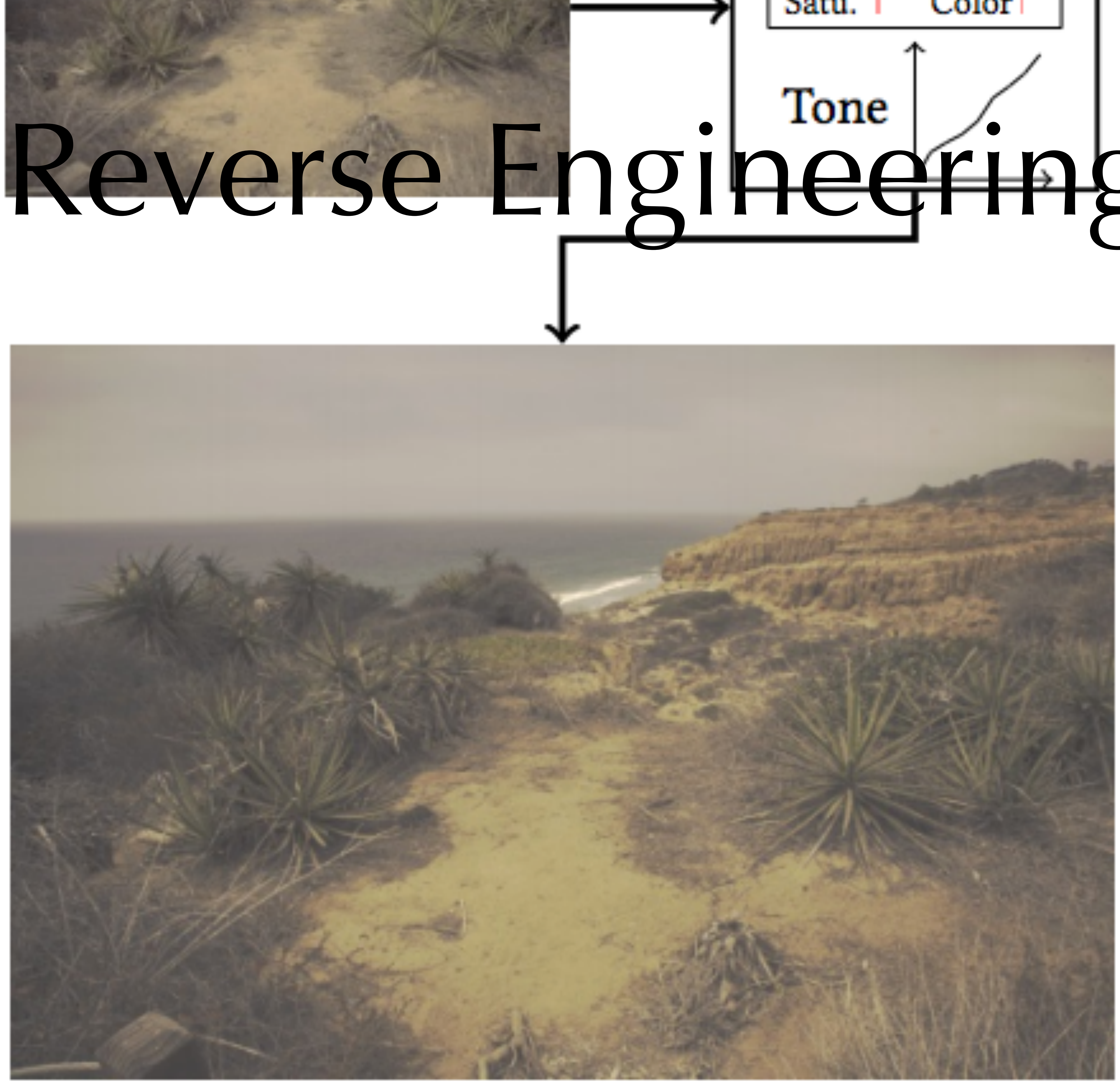
Ours (unpaired training)

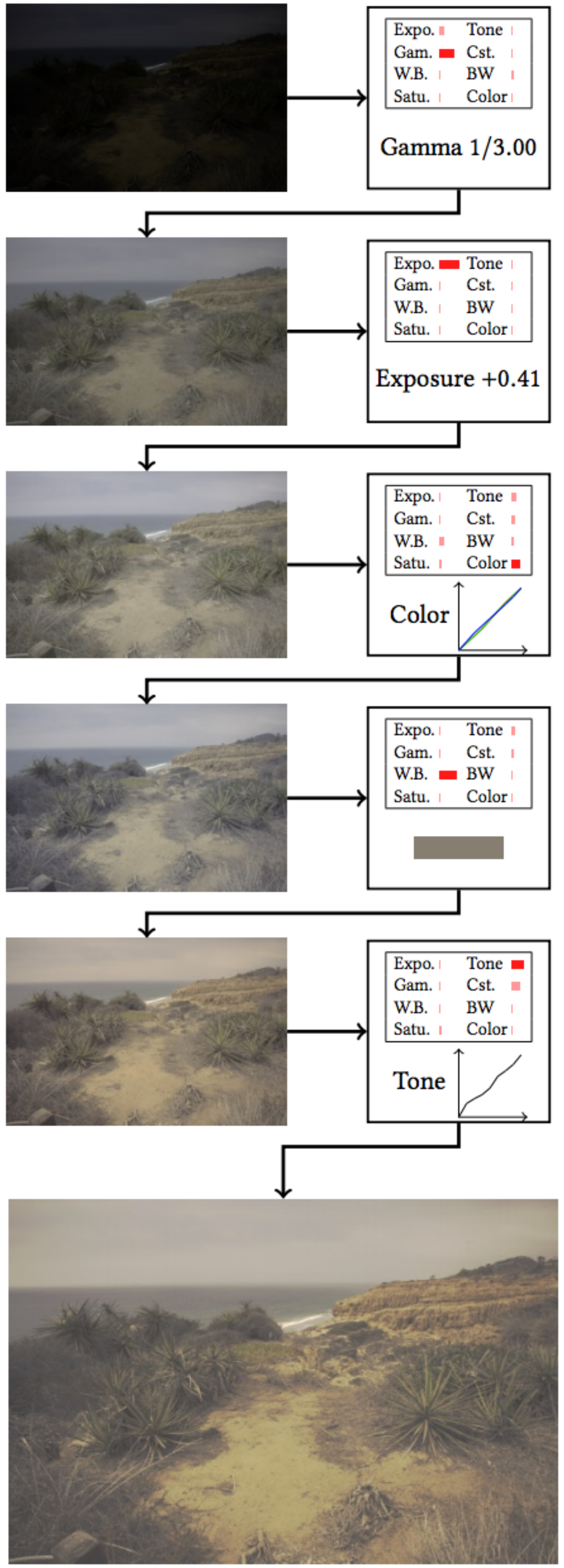
Results: Reverse Engineering

Results: Reverse Engineering

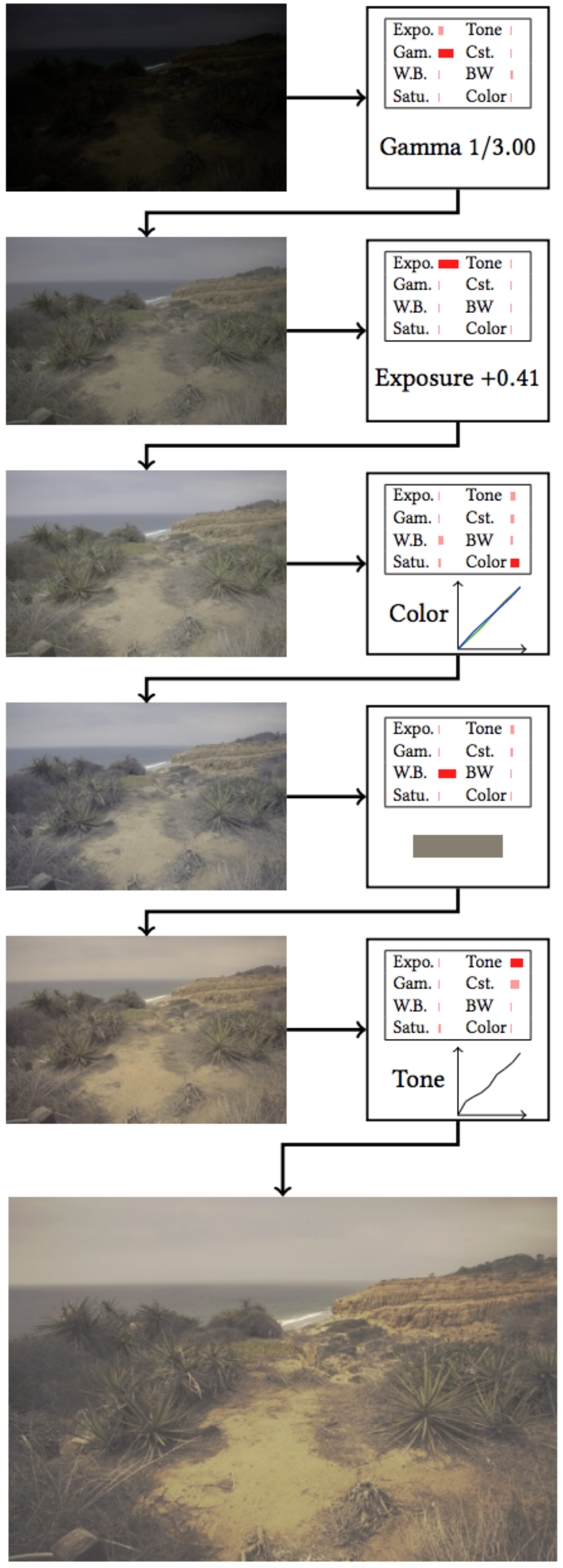


Results: Reverse Engineering



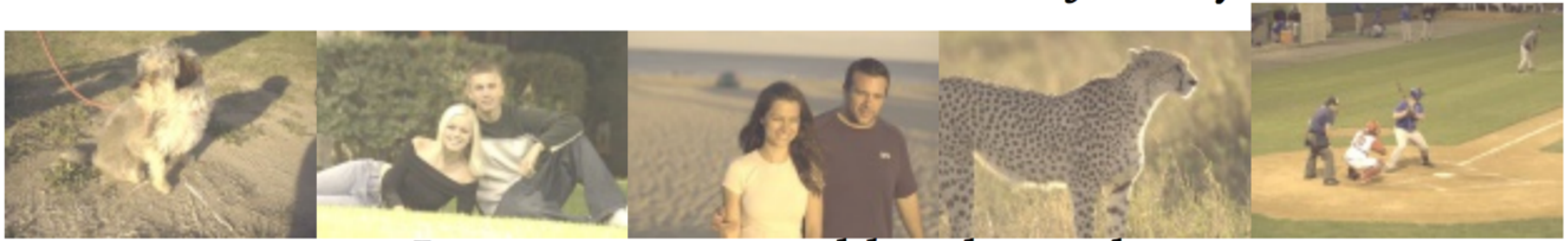


Expo. Tone
Gam. Cst.
W.B. BW
Satu. Color
Gamma 1/3.00
Exposure +0.41
Color
Tone




```python
# Step 1: Gamma
image = image ** (1 / 3.0)
# Step 2: Exposure
image = image / image.mean() * 0.6
# Step 3: Boost blue shadow
blue_shadow = image[:, :, 2] < 0.5
blue = image[:, :, 2]
blue = blue_shadow * (blue * 2) ** 0.7 / 2 + blue * (1 - blue_shadow)
image[:, :, 2] = blue
# Step 4: White balance
image = image * np.array((1.055, 0.984, 0.886)).reshape((1, 1, 3))
# Step 5: Boost shadow
shadow = image < 0.33
image = ((image * shadow * 3) ** 0.8 / 3) + image * (1 - shadow)
```

Code based on the learned trajectory



Images generated by the code

Images generated by the black-box filter

Summary

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**, instead of **pixels**

Generative Adversarial Networks

Training without **pairs**

Advantages: infinite resolution, human-understandable (reverse engineering artistic styles), unpaired training

Summary

Differentiable Photo Postprocessing Model

resolution independent
content preserving
human-understandable

Deep Reinforcement Learning

Learn image **operations**, instead of **pixels**

Generative Adversarial Networks

Training without **pairs**

Advantages: infinite resolution, human-understandable (reverse engineering artistic styles), unpaired training

Limitations: RL/GAN stability, hyper-parameters, faces

Retouch your photos with **Exposure**!

Input (tone-mapped) Retouched Input (tone-mapped) Retouched Input (tone-mapped) Retouched

Reproducible research:

https://github.com/yuanming-hu/exposure

Retouch your photos with **Exposure**!

Input (tone-mapped) Retouched Input (tone-mapped) Retouched Input (tone-mapped) Retouched

Reproducible research:
https://github.com/yuanming-hu/exposure

Thank you!

Questions are welcome!