SCGAN: Saliency Map-guided Colorization with Generative Adversarial Network

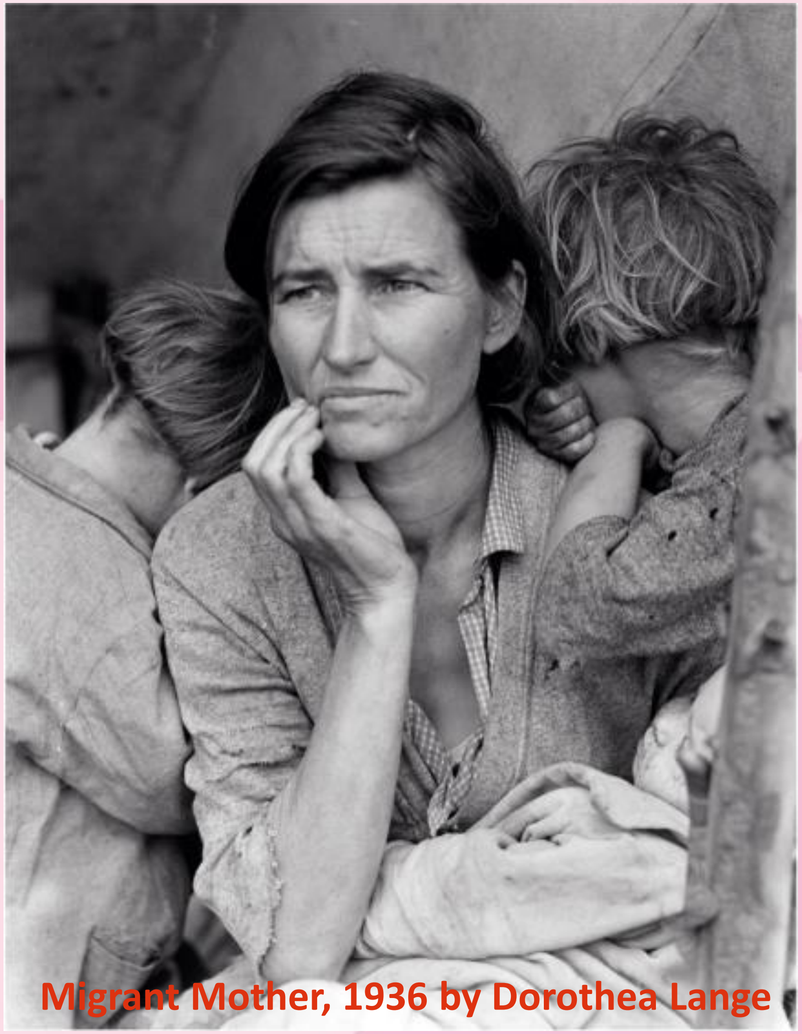
Presenter: Yuzhi Zhao
Supervisor: Dr. Lai-Man Po

5th Feb 2021

香港城市大學
City University of Hong Kong

List

1 Background: What is Colorization?



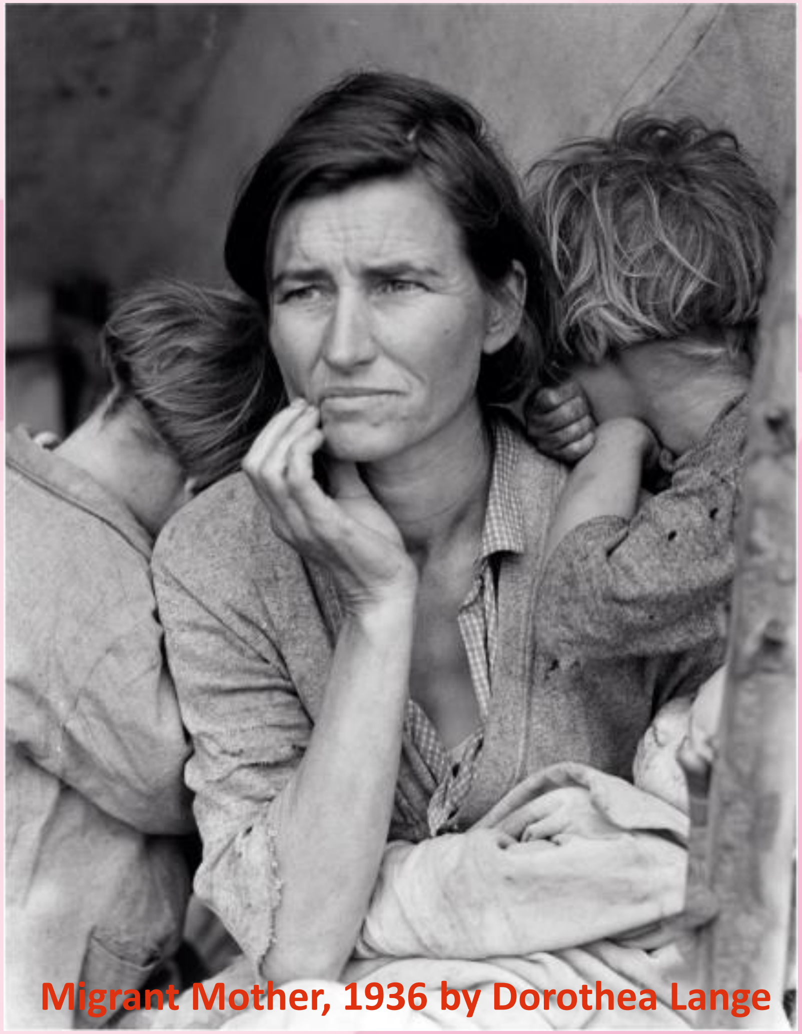
Migrant Mother, 1936 by Dorothea Lange

$\mathcal{F}$

?

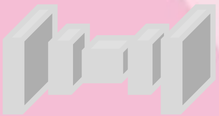
SCGAN Result

1 Background: What is Colorization?



Migrant Mother, 1936 by Dorothea Lange

$\mathcal{F}$



SCGAN
IEEE TCSVT
2020

SCGAN Result

1 Background: Related Works

Non-parametric

Transferring Color to Greyscale Images, ACM TOG, 2002

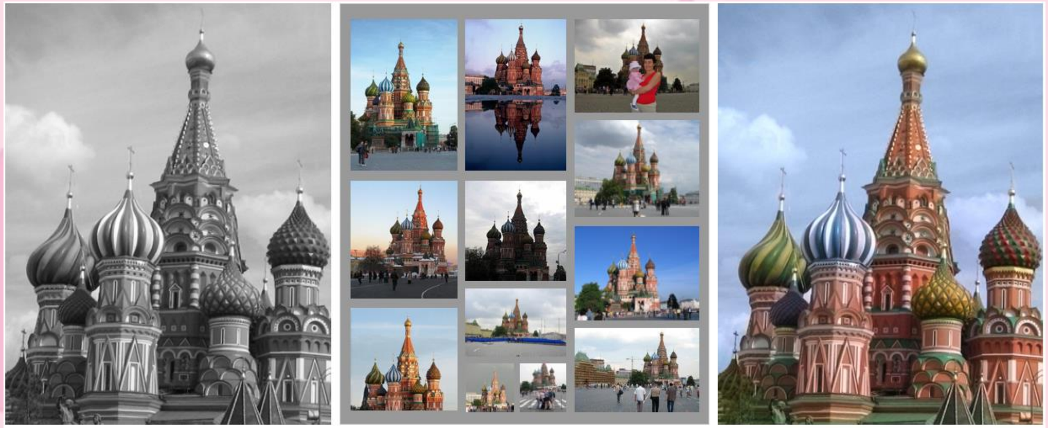
Colorization using optimization, ACM TOG, 2004

Colorization by Example, EGSR, 2005

Natural Image Colorization, EGSR, 2007

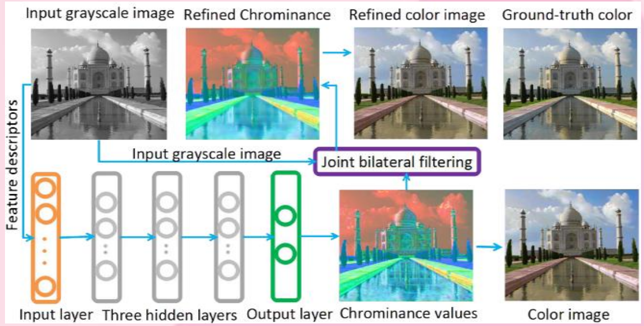
Intrinsic Colorization, ACM TOG, 2008

Image Colorization Using Similar Images, ACM MM, 2012

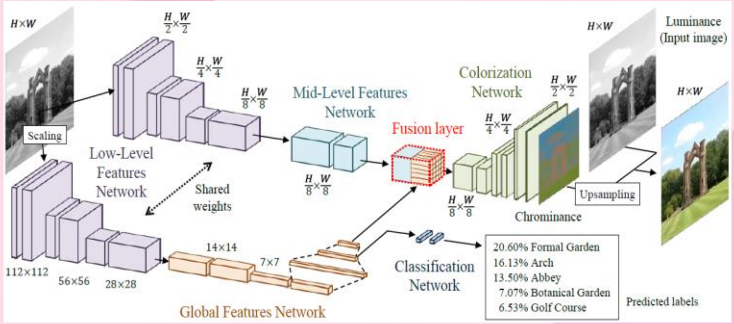


Parametric

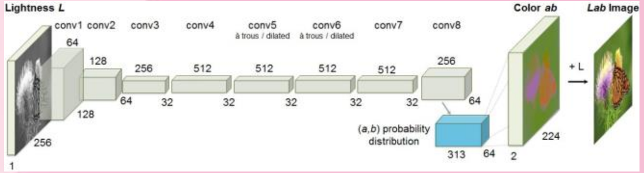
Deep Colorization,
CVPR, 2015



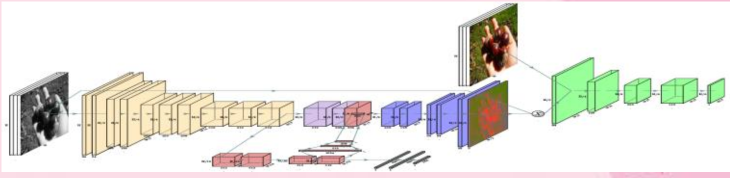
Let there be Color!,
ACM TOG, 2016



Colorful Image
Colorization,
ECCV, 2016

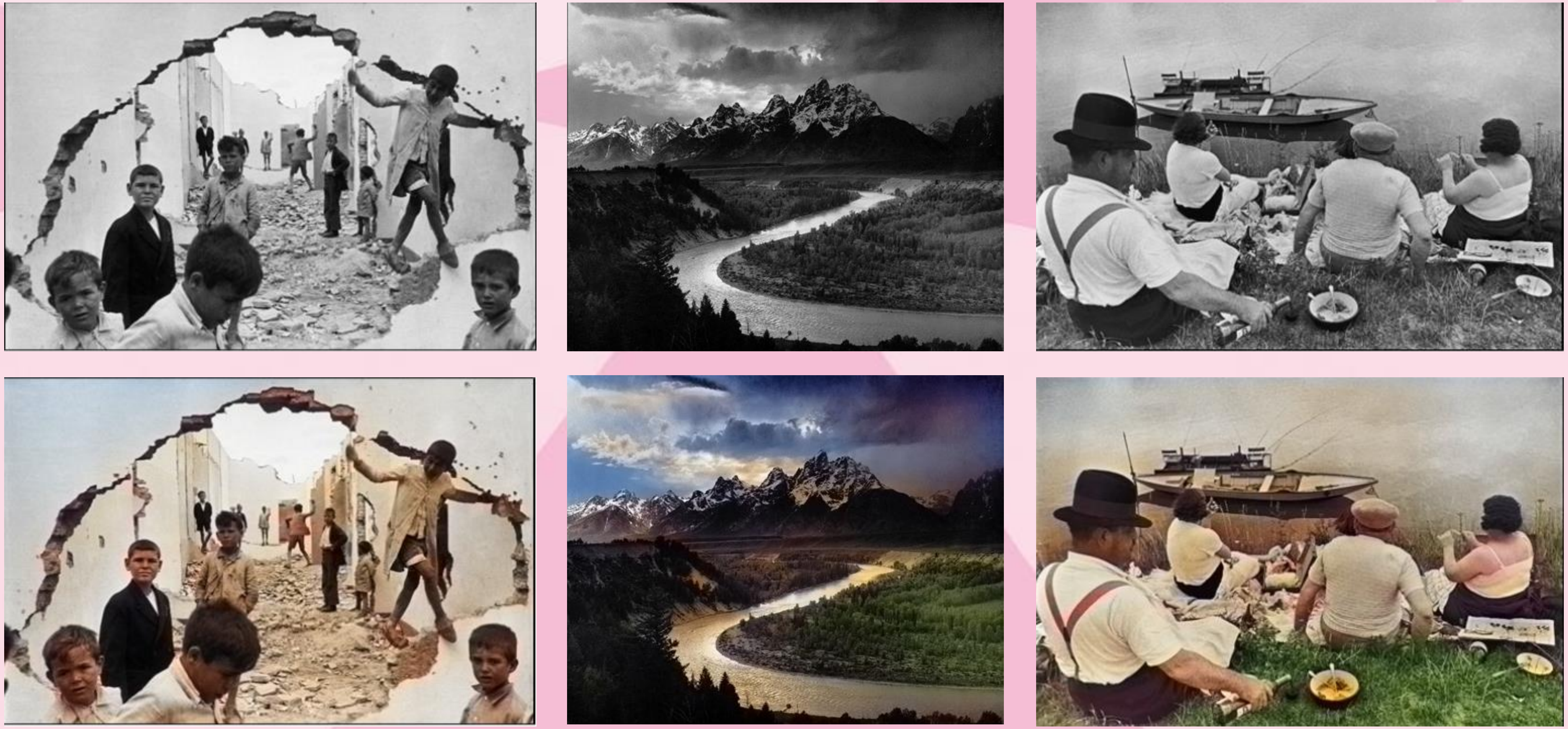


ChromaGAN,
WACV, 2020



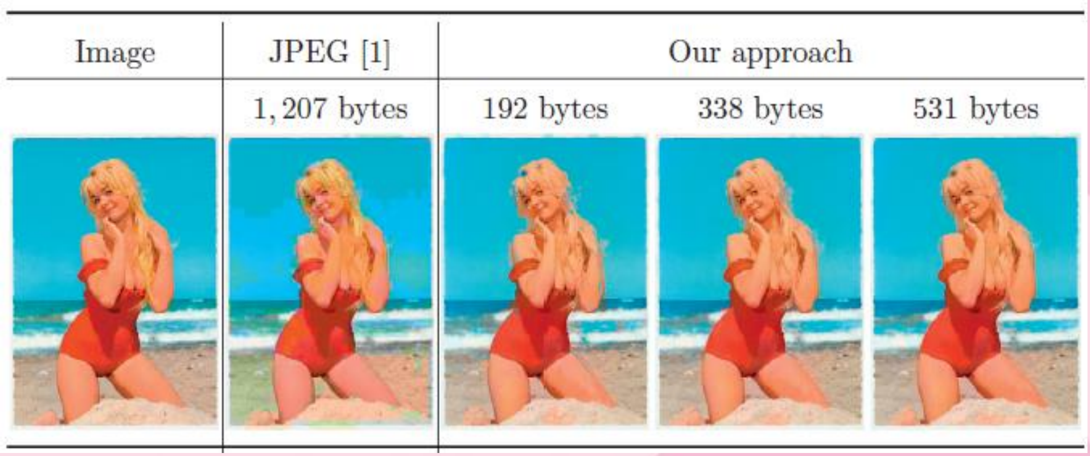
1 Background: Applications

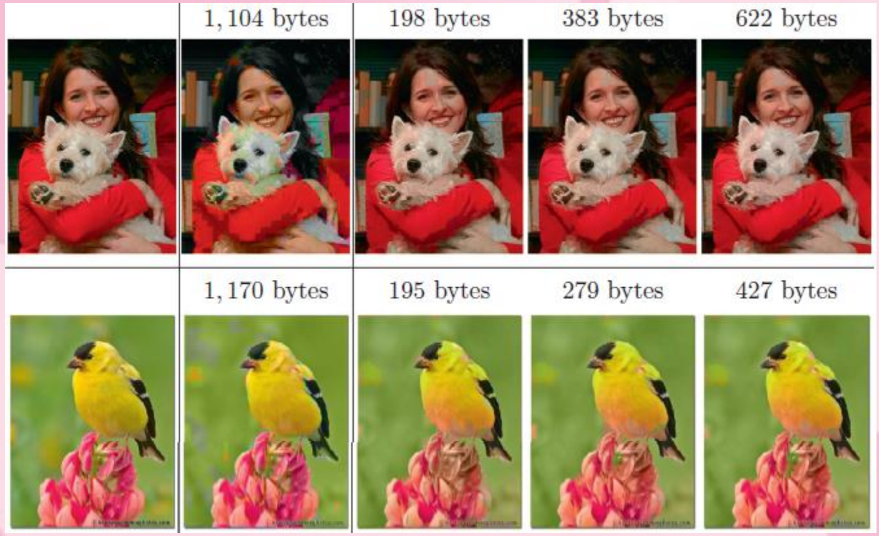
(1) Colorizing Legacy Photos



1 Background: Applications

(2) Image Compression

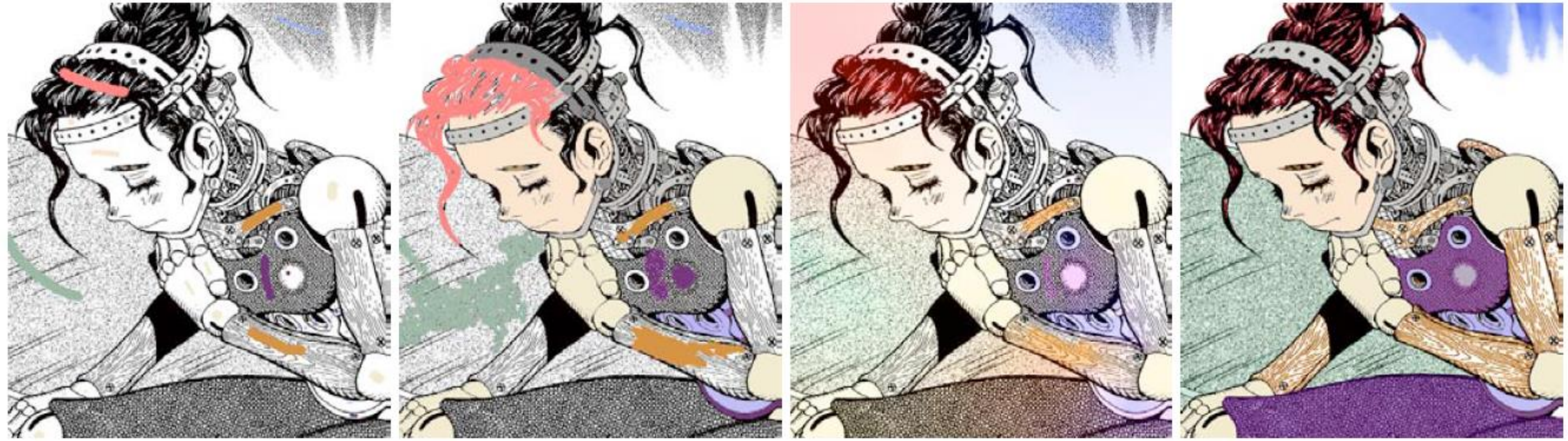




Multiple hypothesis colorization and its application to image compression, CVIU, 2017

1 Background: Applications

(3) Manga Colorization

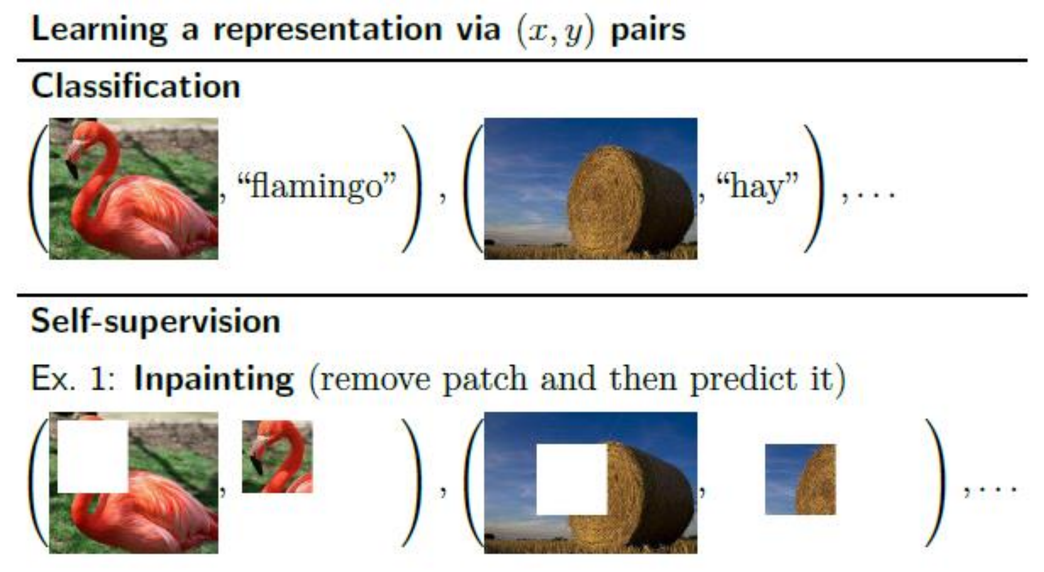


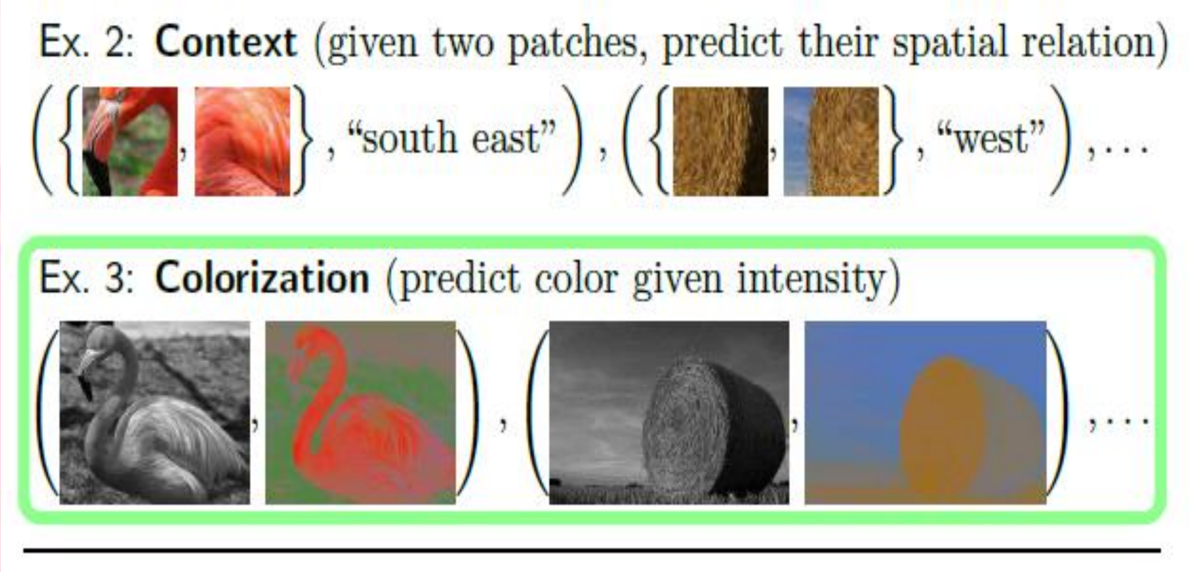
(a) Input (b) Flood-fill (c) Optimization (d) Our result

Manga Colorization, ACM TOG, 2006

1 Background: Applications

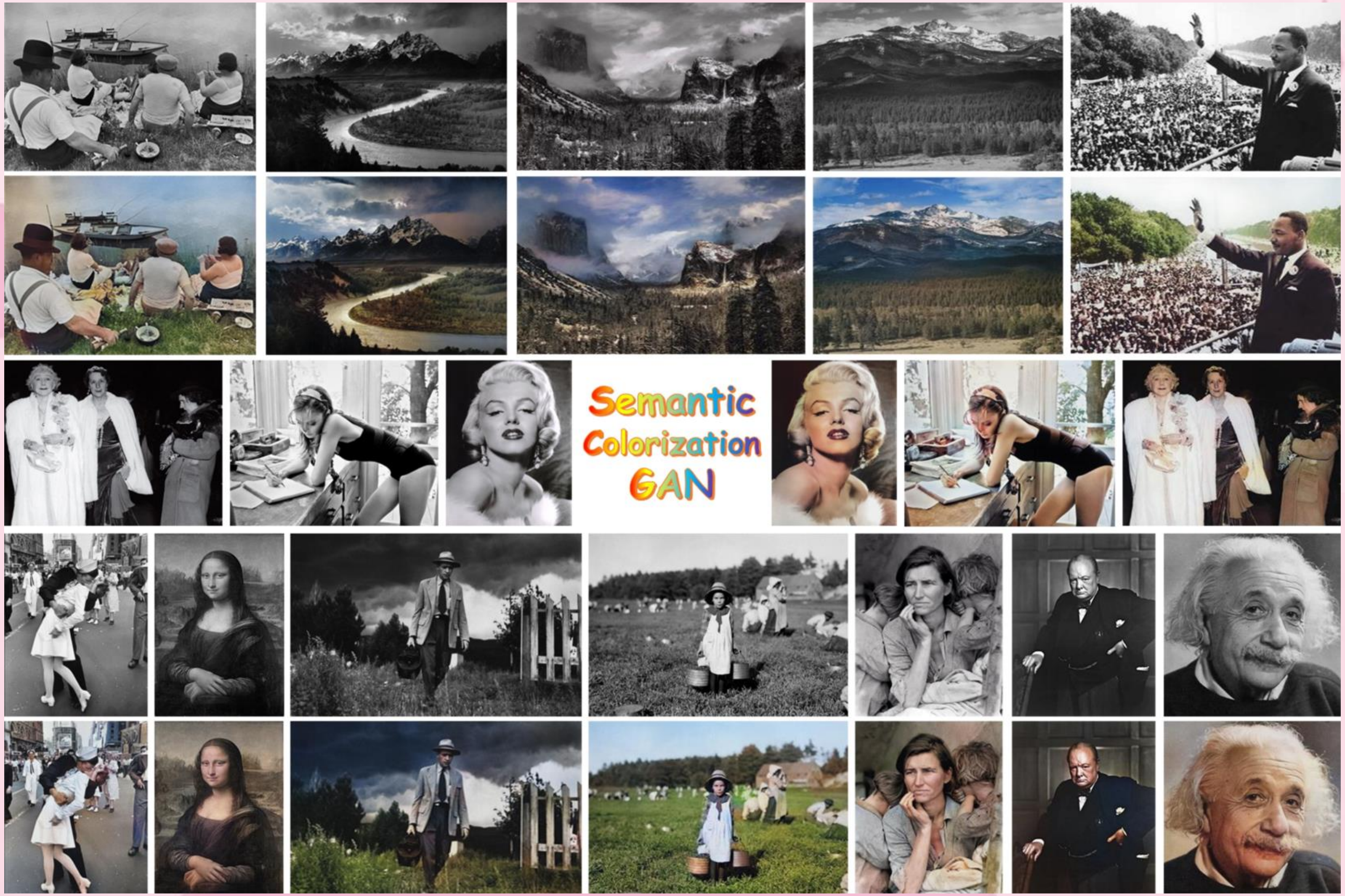
(4) As a Proxy Task for other CV tasks





Colorization as a Proxy Task for Visual Understanding, CVPR, 2017

1 Background: Representative Image



1 Background: Representative Image

code can be found at:
https://github.com/zhaoyuzhi/Semantic-Colorization-GAN

2 Motivation

(1) There exist semantic confusion problem



The road is colorized as blue | The car color bleeds | The grass color radiates





Totally unreasonable color in special scenes | The clothes color emits | Strange color embeddings



2 Motivation

(2) There exist object intervention problem

The grass / tree is green and interferes the surrounding objects

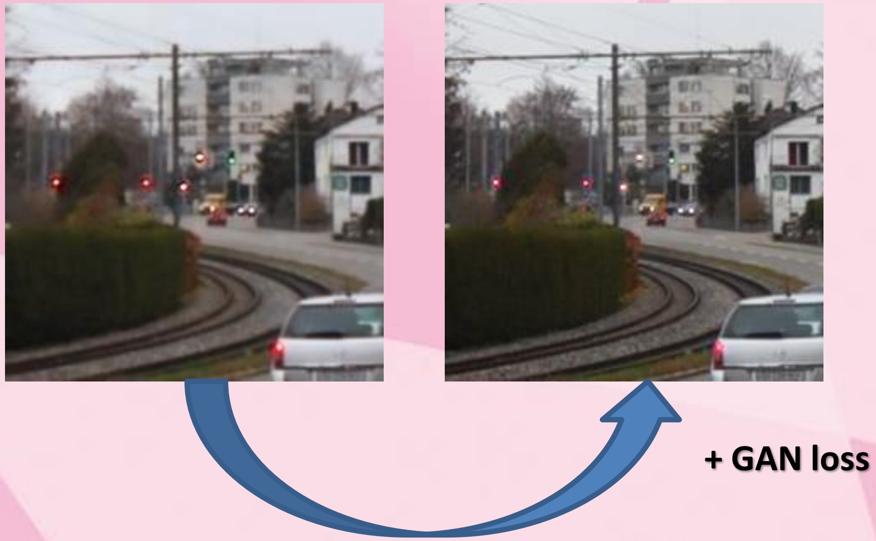
Most common object intervention problems

The sky is blue and interferes the surrounding objects

Other object intervention problems

2 Motivation

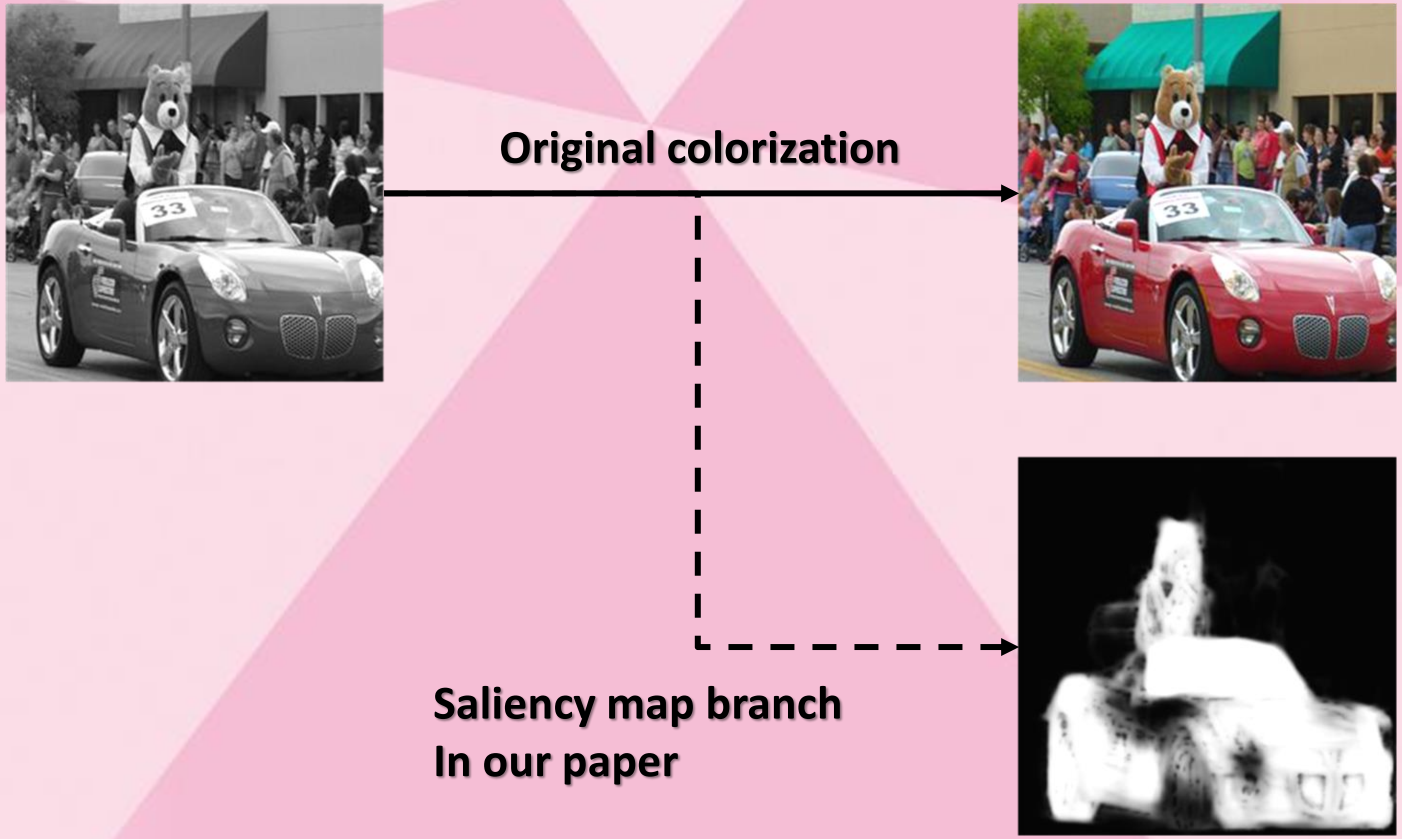
(3) Better Loss Function



(4) Some application does not have enough training data

2 Motivation

We propose to introduce a regularizer to enhance the colorization system: saliency map



2 Motivation

What is saliency map?

saliency map is the visually important region of an image, as a weak binary segmentation map

saliency map is also unsupervised when generation, thus adaptive to any datasets
(i.e., we can generate saliency maps for arbitrary images using SOTA saliency detection method)

2 Motivation

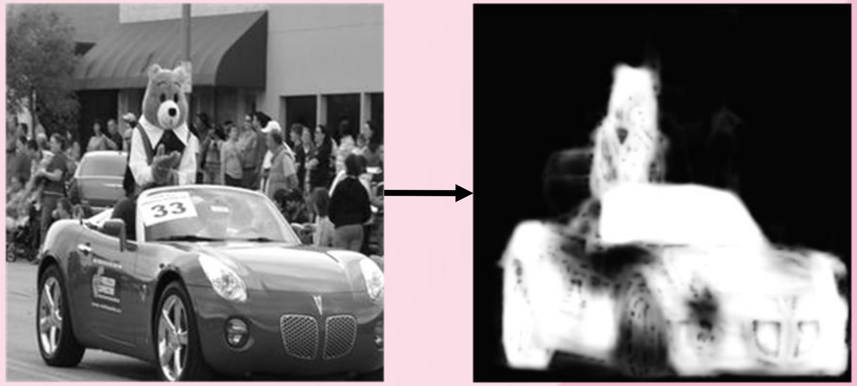
Why saliency map is useful to address those problems?

saliency map is the visually important region of an image, as a weak binary segmentation map

2 Motivation

Why saliency map is useful to address those problems?

saliency map is also unsupervised when generation, thus adaptive to any datasets



RGB images

Multispectral images

NIR images

2 Motivation

Why saliency map is useful to address those problems?

saliency map is also unsupervised when generation, thus adaptive to any datasets

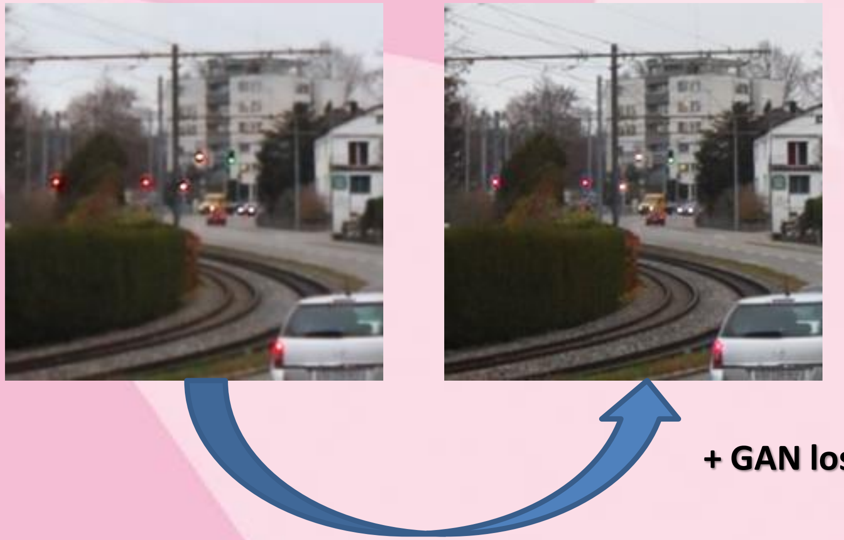
NIR images from a small dataset

2 Motivation

We also propose to use a GAN loss to sharpen image and make the output more colorful

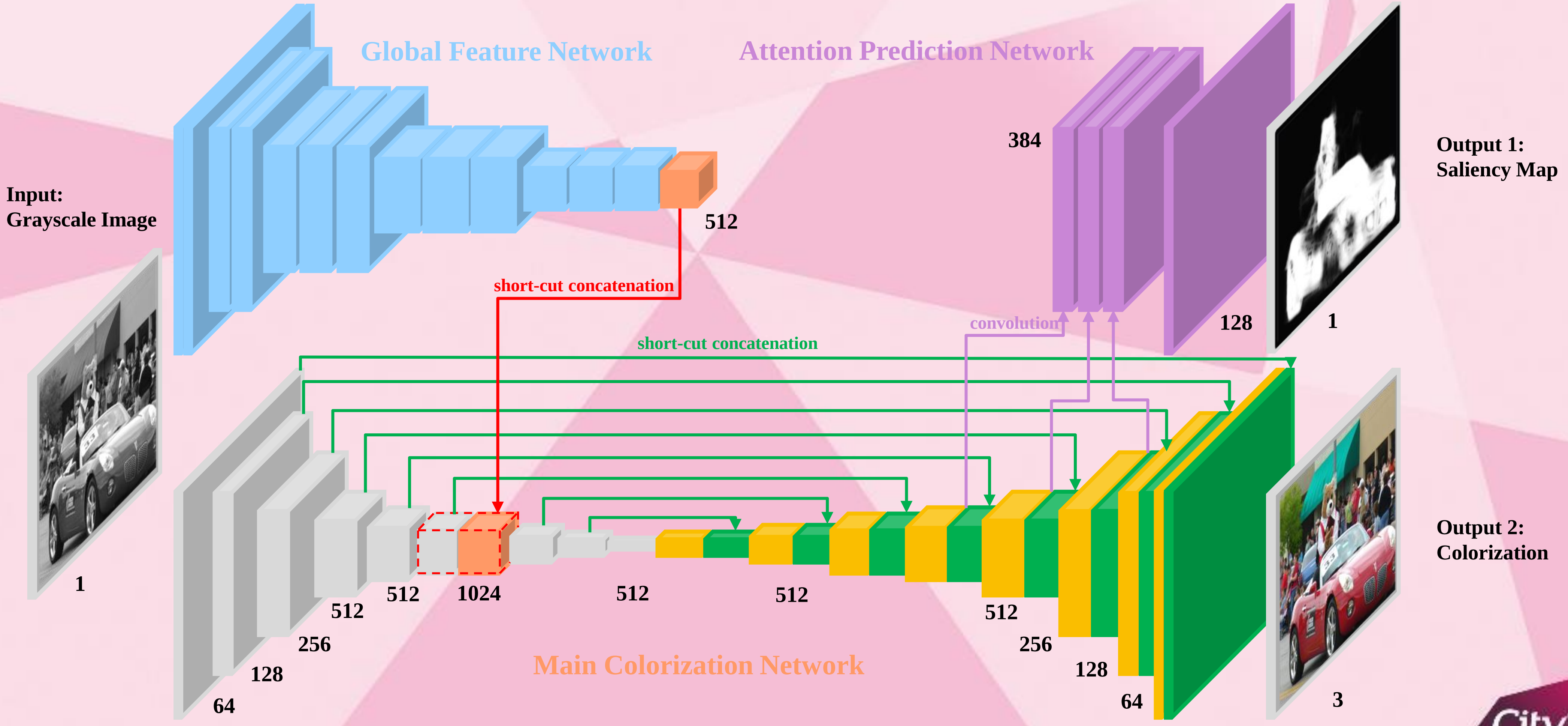
GAN is a generative model that generates an image from noise. The proposed colorization framework is a conditional GAN that generates RGB images from grayscale images

+ GAN loss



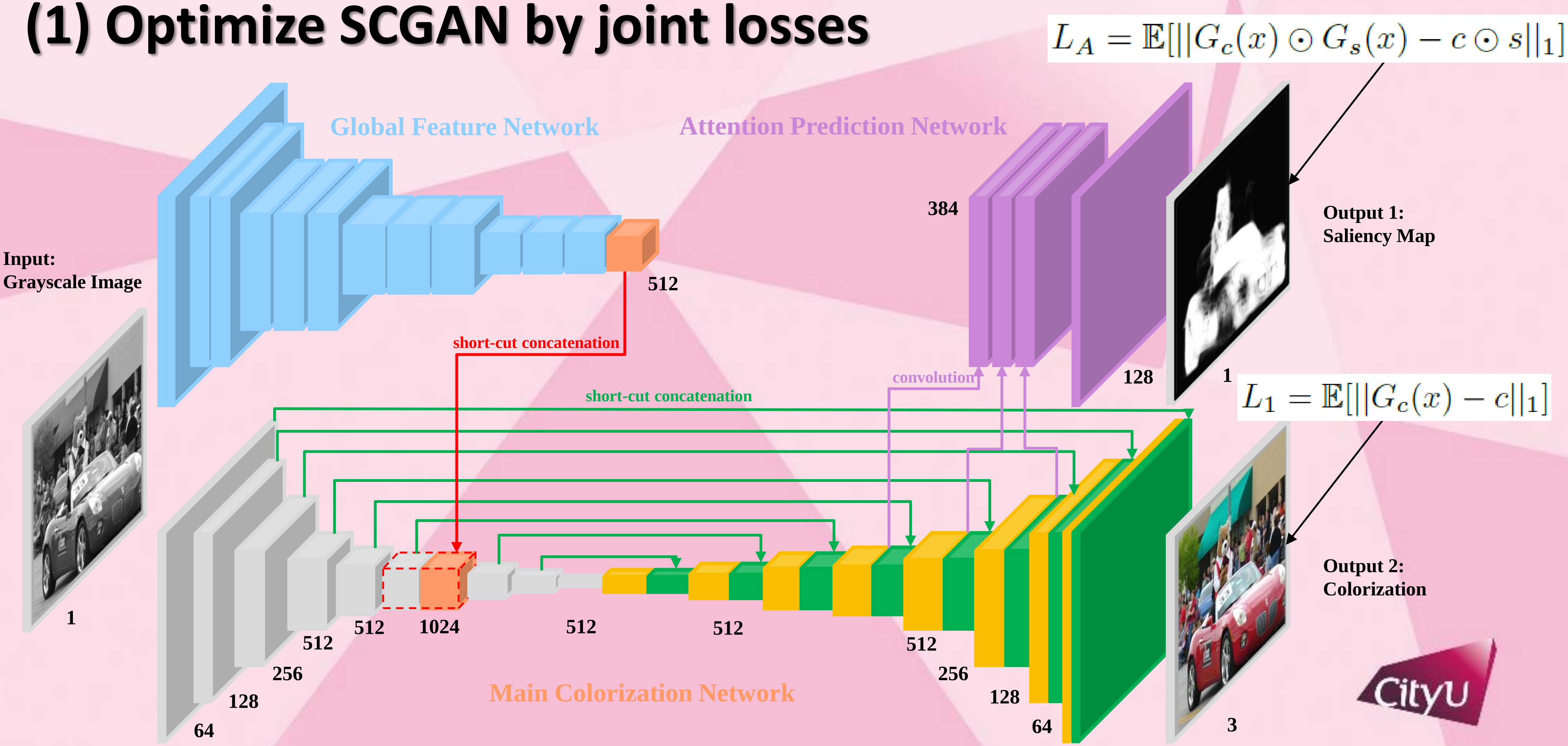
+ GAN loss

3 Methodology (SCGAN)



3 Methodology (SCGAN)

(1) Optimize SCGAN by joint losses

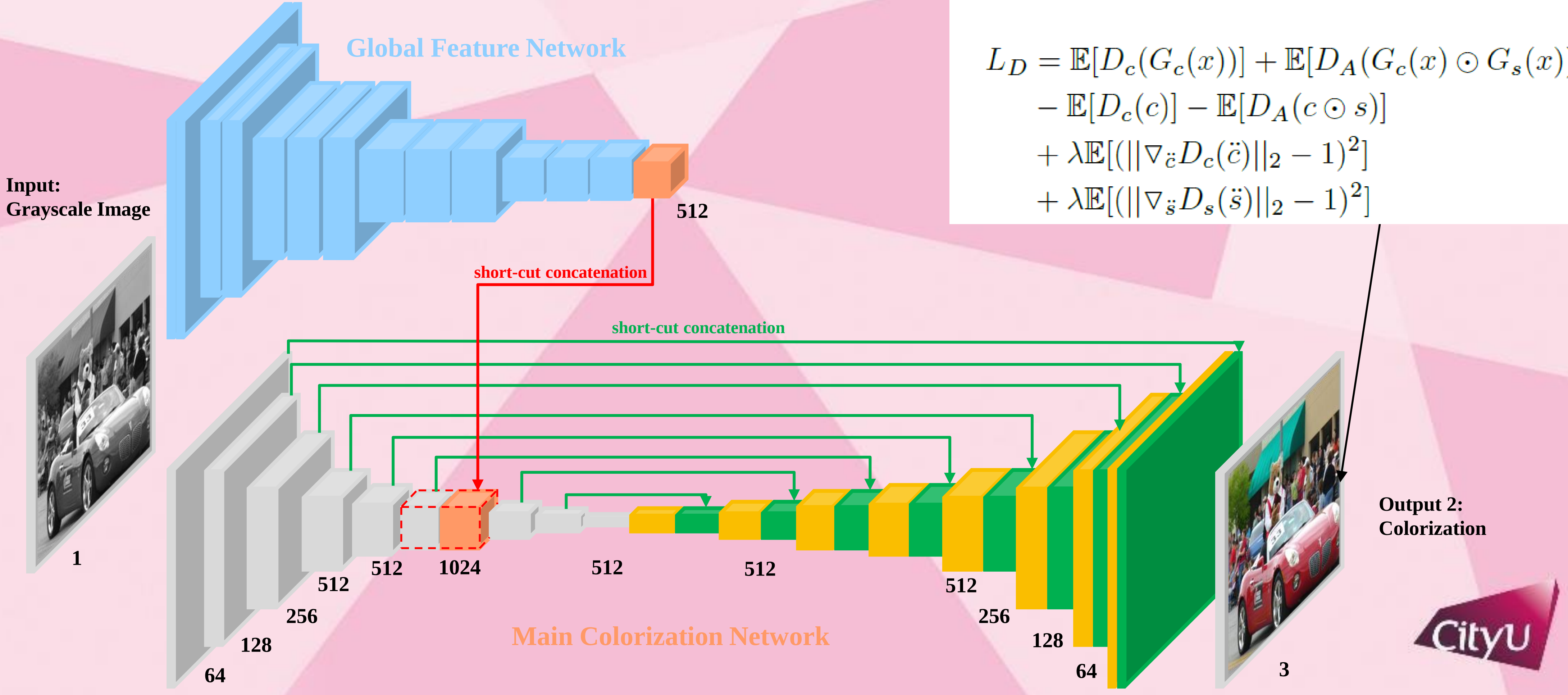


3 Methodology (SCGAN)

(2) GAN loss

$$L_G = -\mathbb{E}[D_c(G_c(x))] - \mathbb{E}[D_A(G_c(x) \odot G_s(x))],$$

$$\begin{aligned} L_D = & \ \mathbb{E}[D_c(G_c(x))] + \mathbb{E}[D_A(G_c(x) \odot G_s(x))] \\ & - \mathbb{E}[D_c(c)] - \mathbb{E}[D_A(c \odot s)] \\ & + \lambda \mathbb{E}[(||\nabla_{\ddot{c}} D_c(\ddot{c})||_2 - 1)^2] \\ & + \lambda \mathbb{E}[(||\nabla_{\ddot{s}} D_s(\ddot{s})||_2 - 1)^2] \end{aligned}$$



4 Optimization

Details:

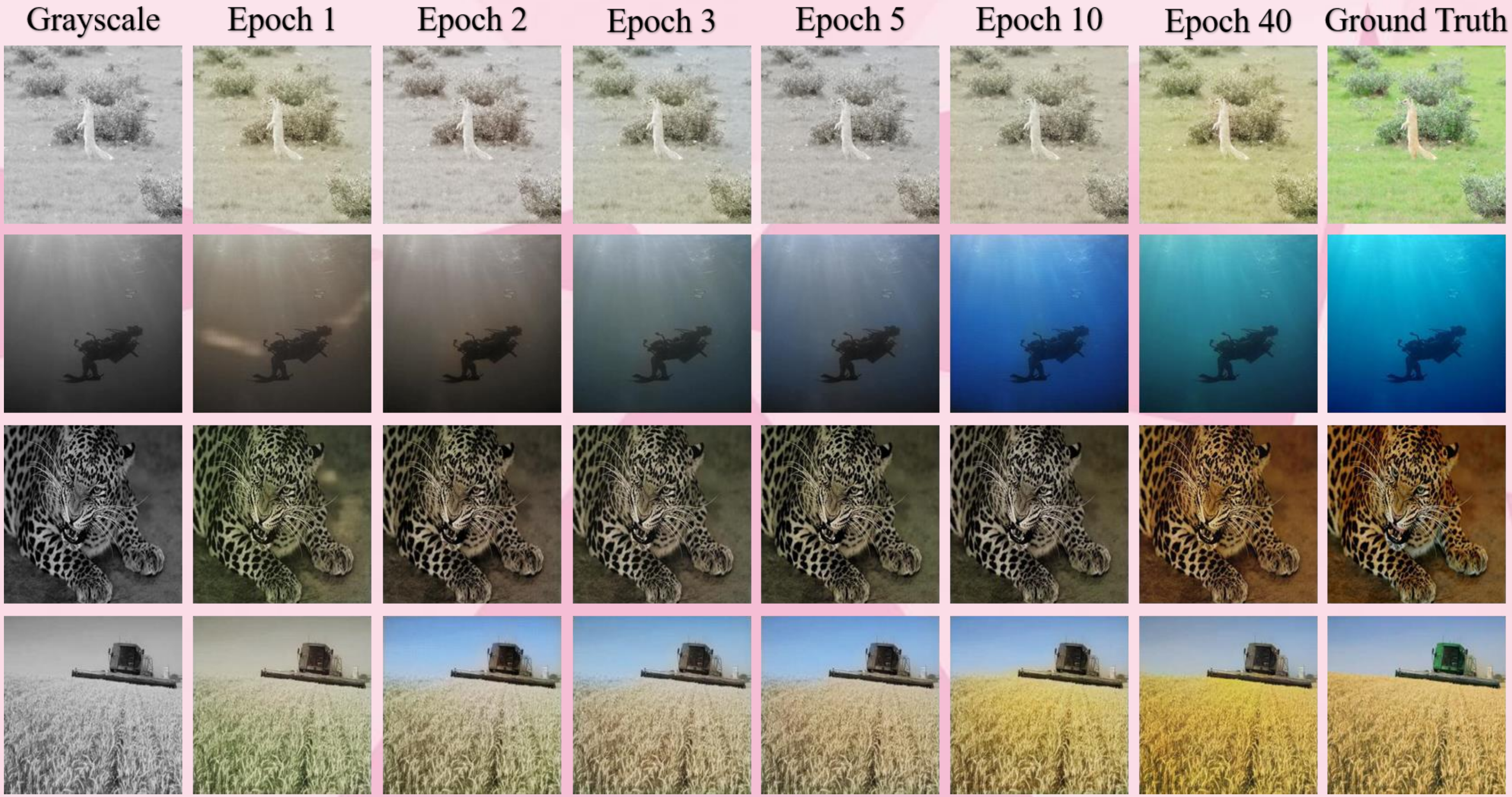
(1) Dataset: 10% of ImageNet (where other methods use the entire dataset);

(2) Adam optimizer with Xavier initialization;

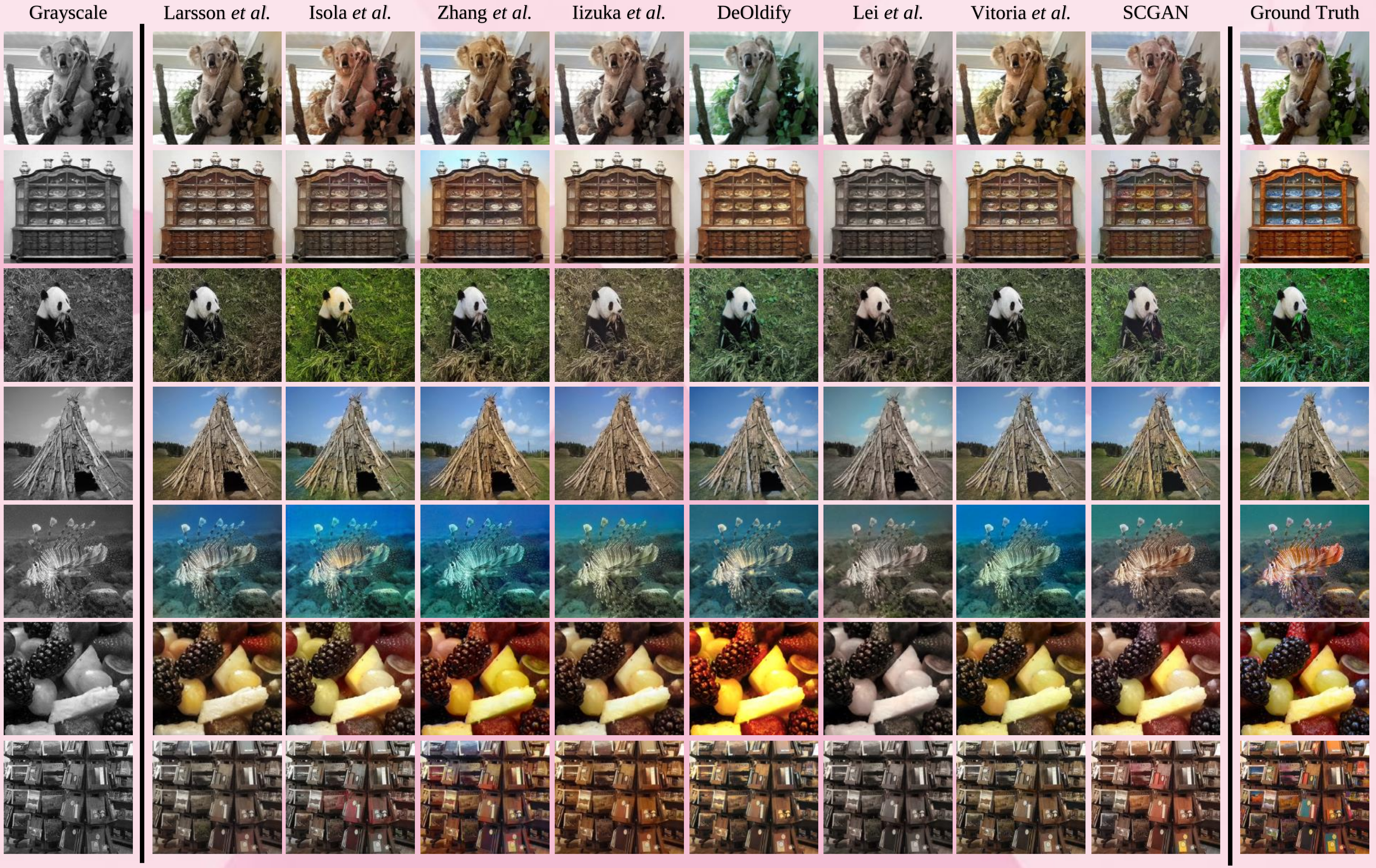
(3) Image patch 256×256;

(4) PyTorch 1.0.0 with Python 3.6.

4 Optimization



5 Comparison with State of the Art



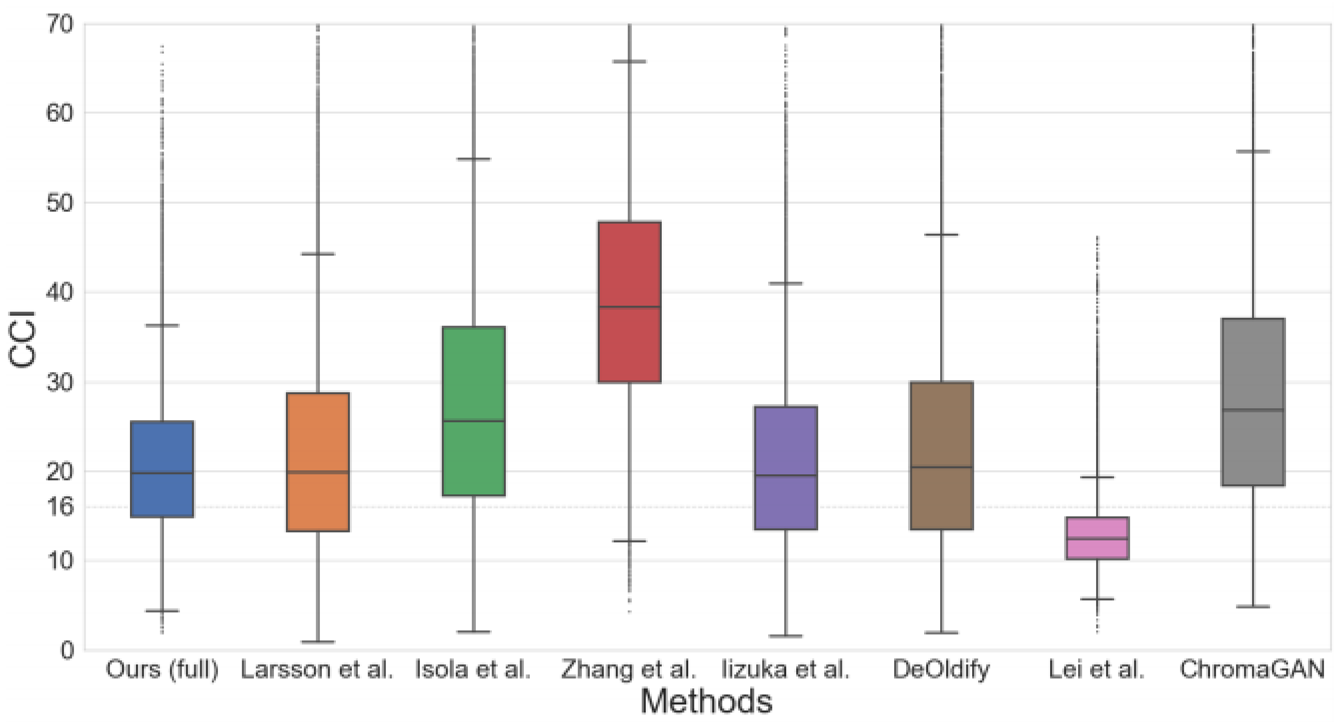
5 Comparison with State of the Art

Quantitative results

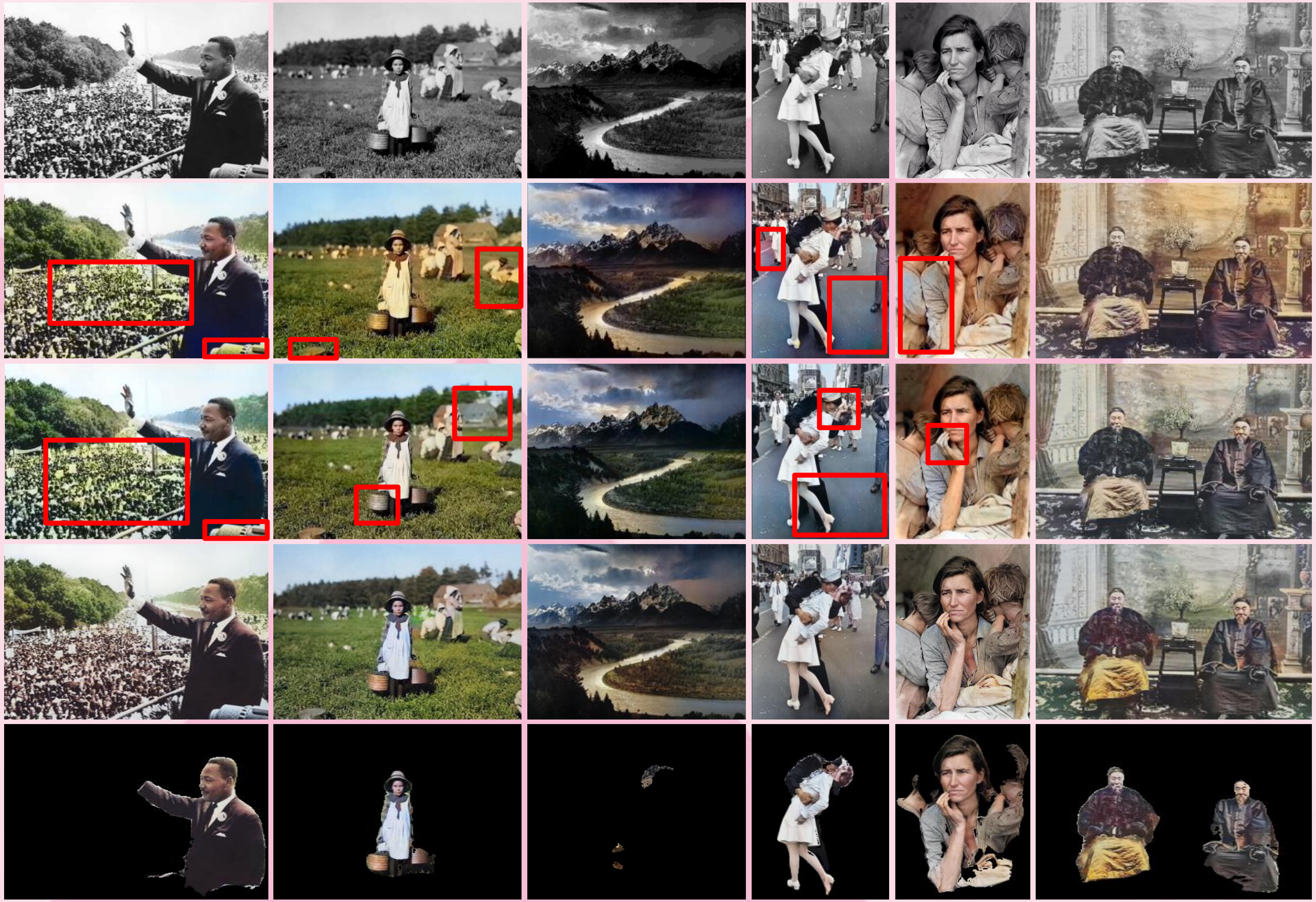
Method	PSNR	SSIM	Top-1 Accuracy	CCI Ratio	Color Naturalness	Color Bleeding Removal	Color Colorfulness
Ground Truth	/	1	63.44%	/	/	/	/
Grayscale	23.23	0.9394	49.78%	/	/	/	/
Larsson *et al.*	**24.42**	0.9229	**55.16%**	14.93%	9.14	8.58	8.22
Isola *et al.*	23.25	0.9386	52.29%	11.26%	8.56	7.93	8.96
Zhang *et al.*	22.49	0.9153	53.97%	3.300%	9.05	7.10	**9.50**
Iizuka *et al.*	24.32	0.9464	53.05%	19.60%	9.17	8.34	8.76
DeOldify	23.14	0.9194	53.45%	14.73%	9.20	8.57	9.01
Lei *et al.*	22.96	0.9146	51.46%	11.40%	8.02	7.14	7.45
Vitoria *et al.*	24.32	0.9273	53.65%	11.24%	9.03	8.24	9.17
SCGAN	23.80	**0.9473**	53.47%	**21.41%**	**9.32**	**8.68**	9.04

5 Comparison with State of the Art

Quantitative results



5 Comparison with State of the Art



Grayscale

Zhang *et al.*

Vitoria *et al.*

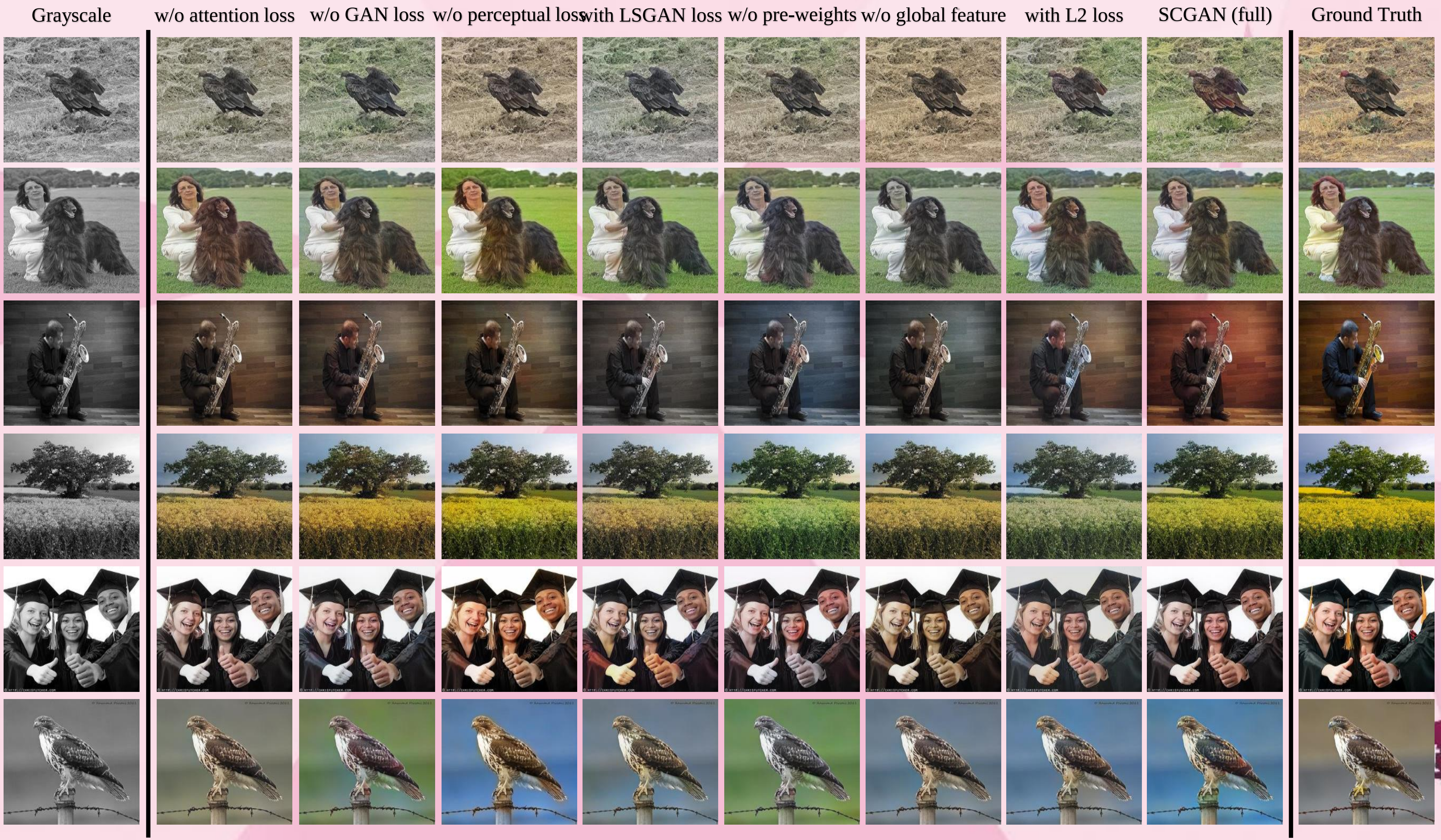
SCGAN

SCGAN

5 Comparison with State of the Art

SCGAN | Grayscale | Reference | Welsh *et al.* | Gupta *et al.* | Bugeau *et al.* | He *et al.*

6 Ablation Study



6 Ablation Study

Quantitative results

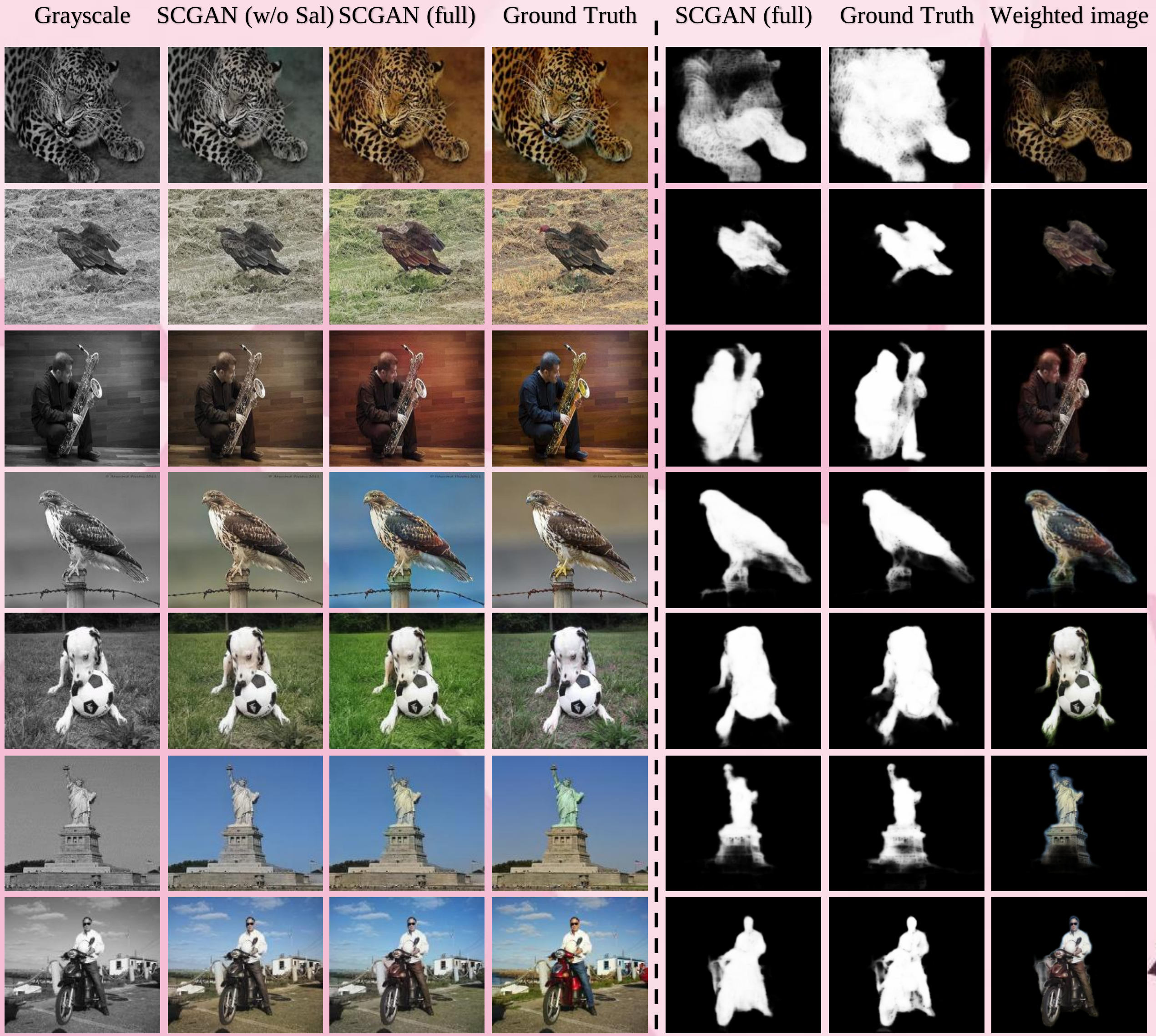
Method	PSNR	SSIM	Top-1 Acc	CCI Ratio
w/o attention loss	**23.81**	0.9368	52.34%	18.56%
w/o GAN loss	23.28	0.9305	52.89%	20.45%
w/o perceptual loss	23.80	0.9443	52.11%	21.31%
with LSGAN loss	23.46	0.9390	53.42%	20.86%
w/o pre-weights	23.15	0.9280	52.59%	18.20%
w/o global feature	23.61	0.9356	52.16%	17.55%
with L2 loss	23.67	0.9436	53.26%	19.58%
SCGAN (full)	23.80	**0.9473**	**53.47%**	**21.41%**

7 Discussion of Saliency Map-guidance Method

SCGAN trained with saliency map, i.e., SCGAN (full)

VS

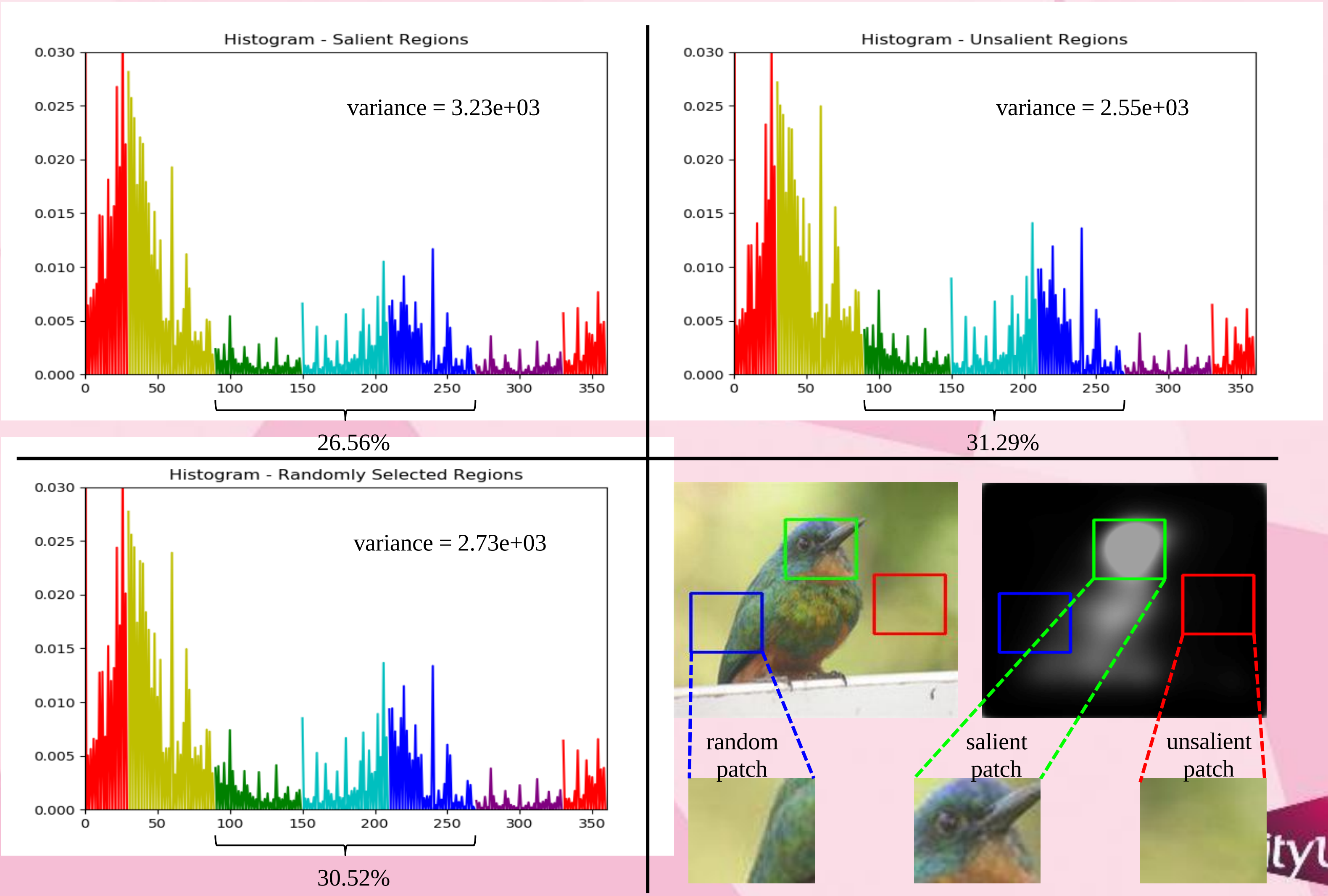
SCGAN trained without saliency map, i.e., SCGAN (w/o sal)



7 Discussion of Saliency Map-guidance Method

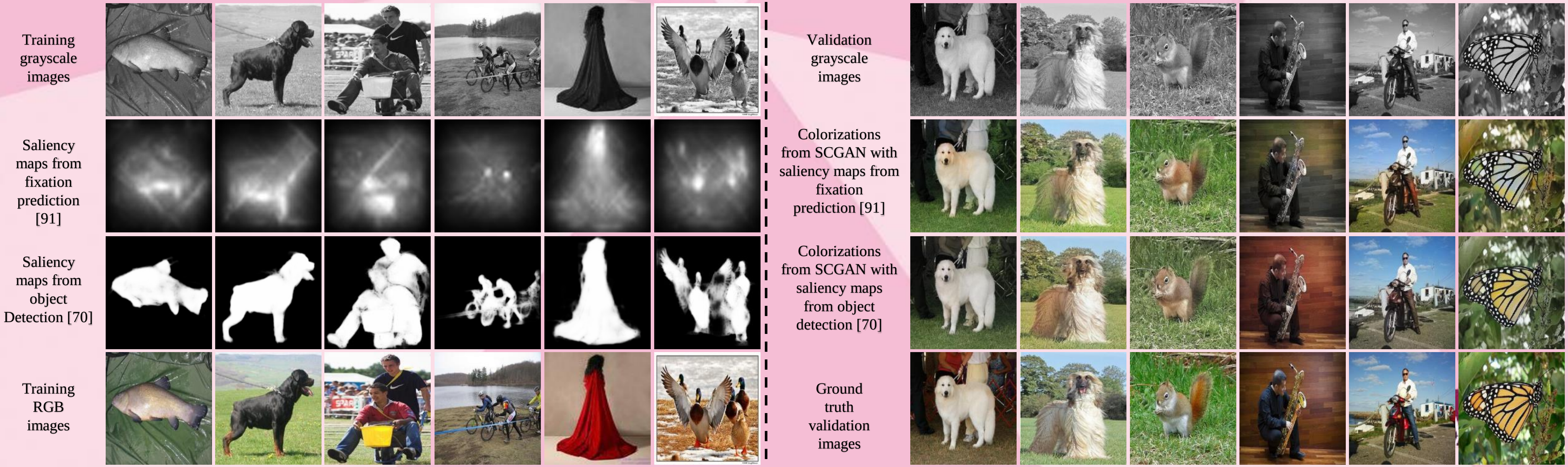
SCGAN trained with saliency map, i.e., SCGAN (full)

occupies more colorful regions, while less green and blue



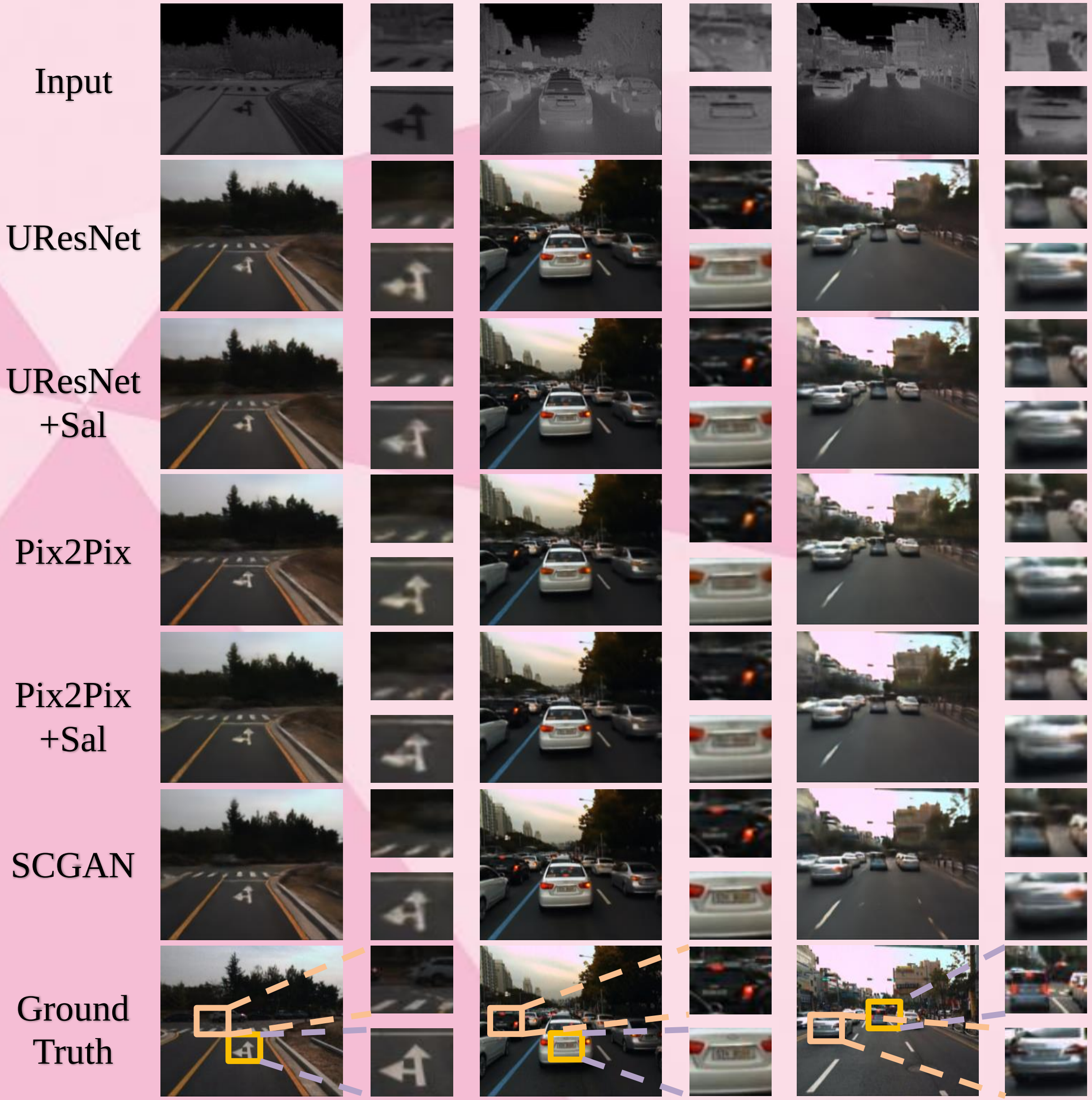
7 Discussion of Saliency Map-guidance Method

Two types of saliency maps: we use object detection saliency map



8 Other Applications: Multispectral and Legacy Images

Multispectral images with saliency map-guidance method and the proposed SCGAN



8 Other Applications: Multispectral and Legacy Images

Multispectral images with saliency map-guidance method and the proposed SCGAN

Method	PSNR	SSIM	Saliency Map Guidance
Pix2Pix	23.55	0.8165	-
Pix2Pix+Sal	23.53	0.8164	✓
UResNet	23.66	0.8219	-
UResNet+Sal	23.72	0.8244	✓
SCGAN	**24.59**	**0.8396**	✓

8 Other Applications: Multispectral and Legacy Images

Legacy images with saliency map-guidance method and the proposed SCGAN

Grayscale

SCGAN

DeOldify

ColouriseSG

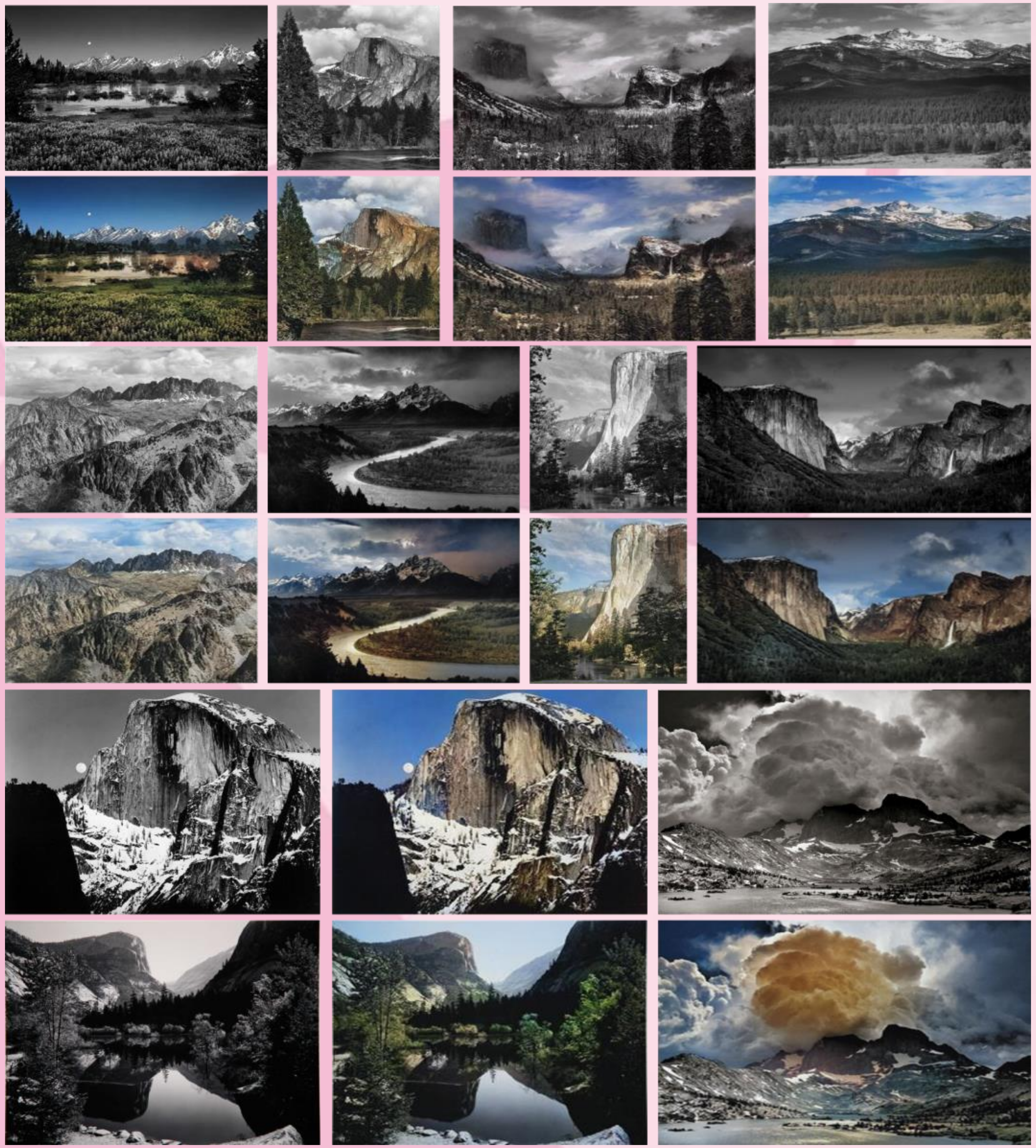
8 Other Applications: Multispectral and Legacy Images

8 Other Applications: Multispectral and Legacy Images

Legacy images with saliency map-guidance method and the proposed SCGAN

Portrait Photographs

8 Other Applications: Multispectral and Legacy Images



Legacy images with saliency map-guidance method and the proposed SCGAN

Landscape Photographs

9 Limitation and Future Work

full paper can be found at:
https://ieeexplore.ieee.org/document/9257445/keywords#keywords

專業 創新 胸懷全球
Professional·Creative
For The World