Learning Representations for Automatic Colorization

Gustav Larsson, Michael Maire, Greg Shakhnarovich

TTI Chicago / University of Chicago

ECCV 2016

Colorization

Let us first define “colorization”

Colorization

Definition 1: The inverse of desaturation.

Original

Colorization

Definition 1: The inverse of desaturation.

Original

$\xrightarrow{\text{Desaturate}}$

Grayscale

Colorization

Definition 1: The inverse of desaturation.

Grayscale

Colorization

Definition 1: The inverse of desaturation.

Original

$\xleftarrow{\text{Colorize}}$

Grayscale

Colorization

Definition 1: The inverse of desaturation. (**Underconstrained!**)

Colorize $\longleftarrow$

Original

Grayscale

Colorization

Definition 2: An inverse of desaturation, that...

Grayscale

Colorization

Definition 2: An inverse of desaturation, that...

Our Method

Colorize $\longleftarrow$

Grayscale

... is plausible and pleasing to a human observer.

Colorization

Definition 2: An inverse of desaturation, that...

Our Method

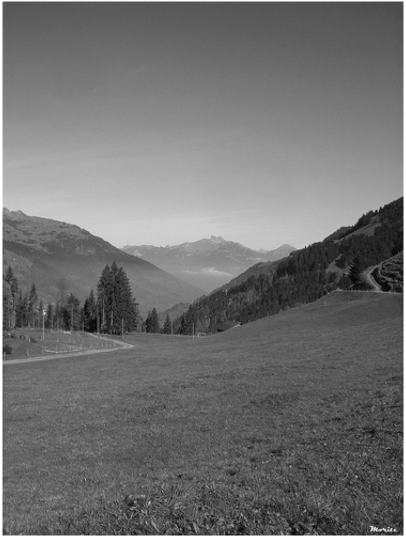
$\xleftarrow{\text{Colorize}}$

Grayscale

... is plausible and pleasing to a human observer.

- Def. 1: Training + Quantitative Evaluation
- Def. 2: Qualitative Evaluation

Manual colorization

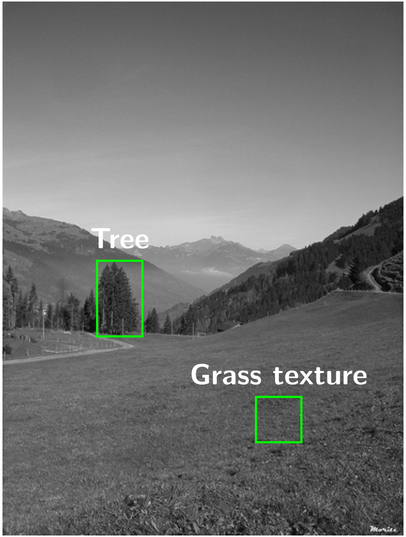


I thought I would give it a quick try...

Manual colorization

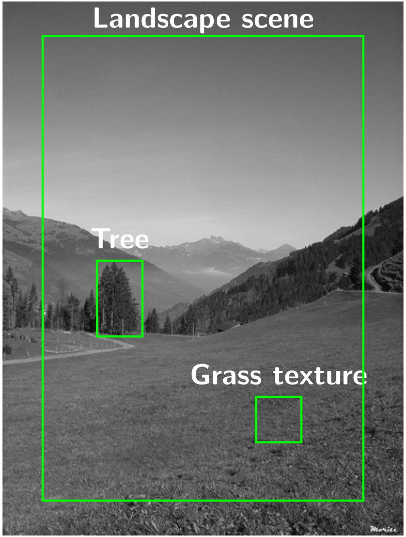
Low-level features

Manual colorization



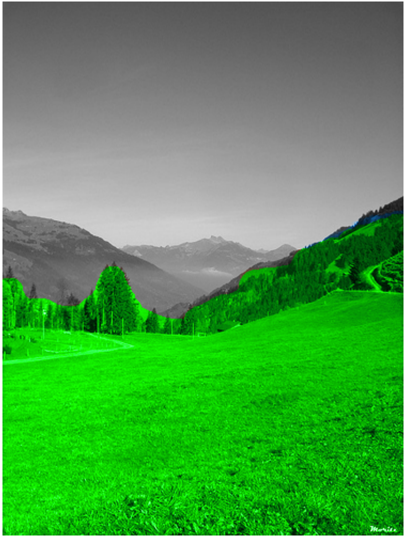
Mid-level features

Manual colorization



High-level features

Manual colorization

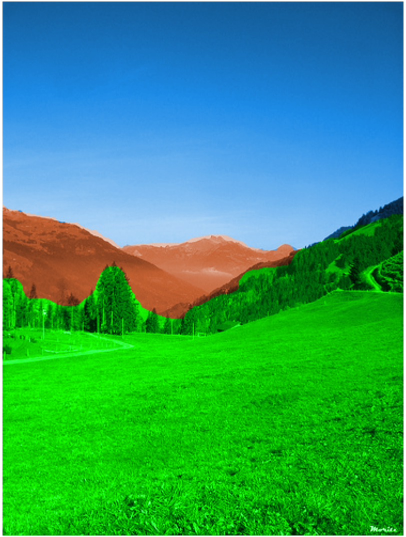


Grass is green

Manual colorization

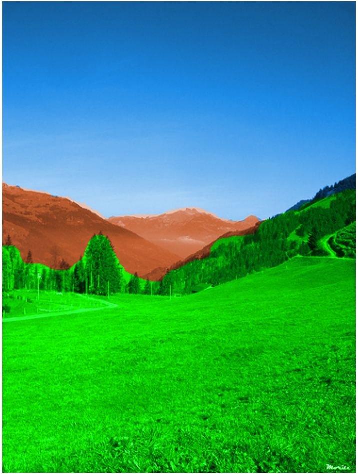
Sky is blue

Manual colorization



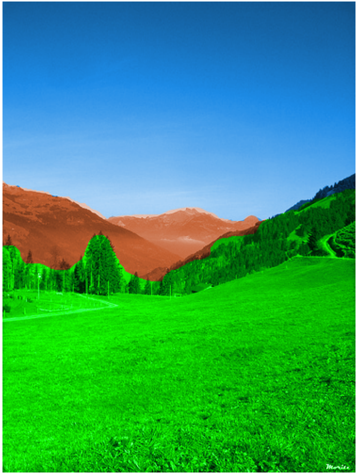
Mountains are... brown?

Manual colorization

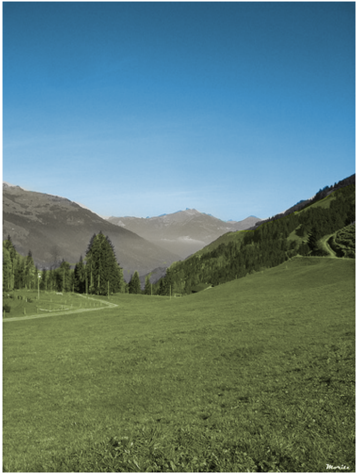


Manual ($\approx$ 15 s)

Manual colorization

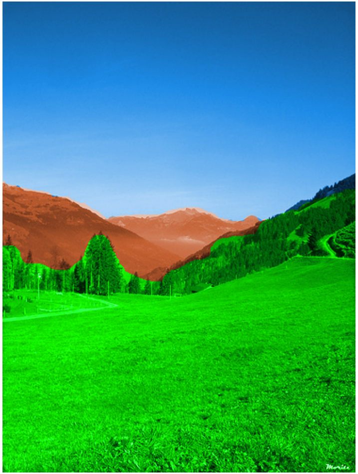


Manual ($\approx$ 15 s)

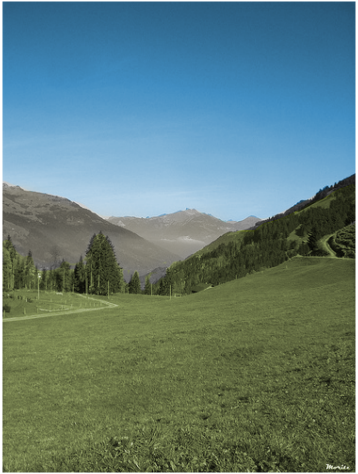


Manual ($\approx$ 3 min)

Manual colorization



Manual ($\approx$ 15 s)



Manual ($\approx$ 3 min)

Automatic ($<$ 1 s)

Our Method

Motivation

1. Colorize old B&W photographs

Motivation

1. Colorize old B&W photographs

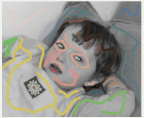
2. Proxy for visual understanding
    - Learning representations useful for other tasks

Related work

Scribble-based methods

Levin et al. (2004); Huang et al. (2005);

Qu et al. (2006); Luan et al. (2007)



Input

Output

Transfer-based methods

Welsh et al. (2002); Irony et al. (2005);

Charpiat et al. (2008); Morimoto et al. (2009);

Chia et al. (2011)

Reference

Input

Output

Prediction-based methods

Deshpande et al. (2015); Cheng et al. (2015)

Iizuka et al. (2016)

Zhang et al. (2016); Larsson et al. (2016)

Input

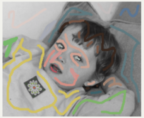
Output

Related work

Scribble-based methods

Levin et al. (2004); Huang et al. (2005);

Qu et al. (2006); Luan et al. (2007)



Input

Output

Transfer-based methods

Welsh et al. (2002); Irony et al. (2005);

Charpiat et al. (2008); Morimoto et al. (2009);

Chia et al. (2011)

Reference

Input

Output

Prediction-based methods

Deshpande et al. (2015); Cheng et al. (2015) ← **ICCV**

Iizuka et al. (2016)

Zhang et al. (2016); Larsson et al. (2016)

Input

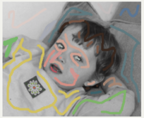
Output

Related work

Scribble-based methods

Levin et al. (2004); Huang et al. (2005);

Qu et al. (2006); Luan et al. (2007)



Input

Output

Transfer-based methods

Welsh et al. (2002); Irony et al. (2005);

Charpiat et al. (2008); Morimoto et al. (2009);

Chia et al. (2011)

Reference

Input

Output

Prediction-based methods

Deshpande et al. (2015); Cheng et al. (2015)

Iizuka et al. (2016) ← **SIGGRAPH**

Zhang et al. (2016); Larsson et al. (2016)

Input

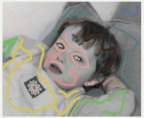
Output

Related work

Scribble-based methods

Levin et al. (2004); Huang et al. (2005);

Qu et al. (2006); Luan et al. (2007)



Input

Output

Transfer-based methods

Welsh et al. (2002); Irony et al. (2005);

Charpiat et al. (2008); Morimoto et al. (2009);

Chia et al. (2011)

Reference

Input

Output

Prediction-based methods

Deshpande et al. (2015); Cheng et al. (2015)

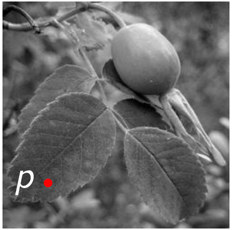
Iizuka et al. (2016)

Zhang et al. (2016); **Larsson et al. (2016)** ← **ECCV**

Input

Output

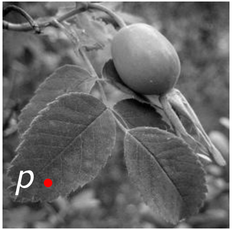
Design principles



Input: Grayscale Image

Design principles

- Semantic knowledge

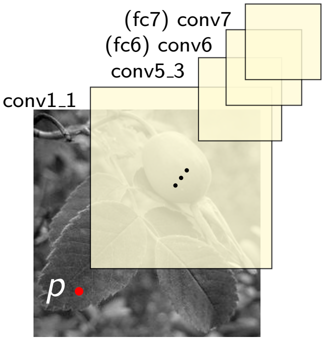


Input: Grayscale Image

Design principles

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier

VGG-16-Gray

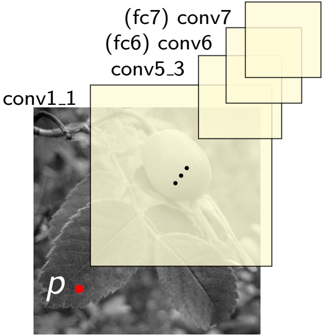


Input: Grayscale Image

Design principles

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features

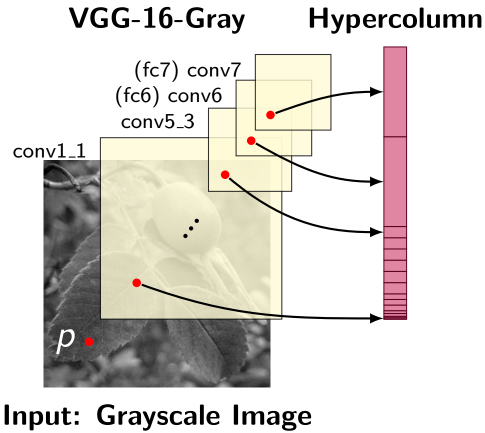
VGG-16-Gray



Input: Grayscale Image

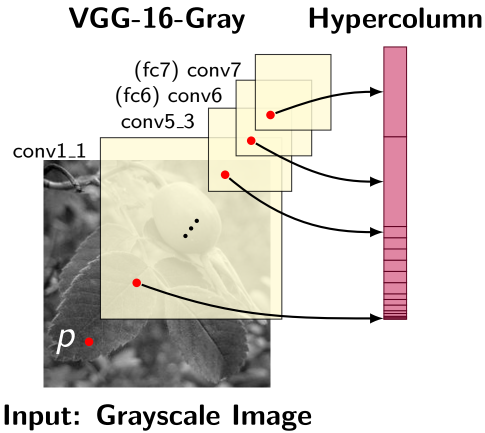
Design principles

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn



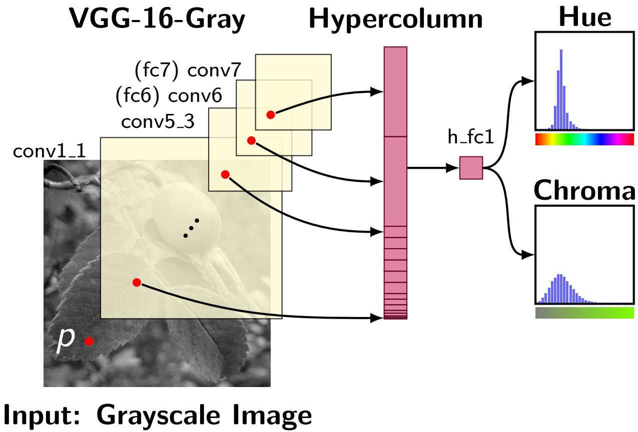
Design principles

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn
- Colorization not unique



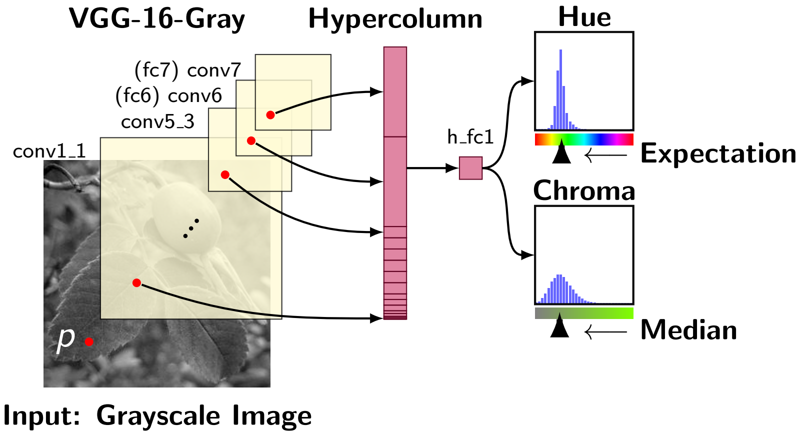
Design principles

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn
- Colorization not unique $\rightarrow$ Predict histograms



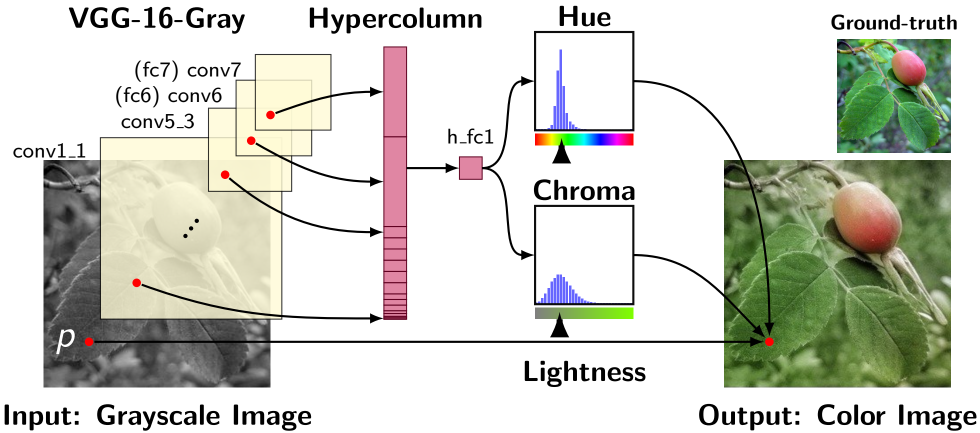
Design principles

- Semantic knowledge → Leverage ImageNet-based classifier
- Low-level/high-level features → Zoom-out/Hypercolumn
- Colorization not unique → Predict histograms



Design principles

- Semantic knowledge $\rightarrow$ Leverage ImageNet-based classifier
- Low-level/high-level features $\rightarrow$ Zoom-out/Hypercolumn
- Colorization not unique $\rightarrow$ Predict histograms



Histogram prediction

The histogram representation is rich and flexible:

Histogram prediction

The histogram representation is rich and flexible:

Histogram prediction

The histogram representation is rich and flexible:

Histogram prediction

The histogram representation is rich and flexible:

Histogram prediction

The histogram representation is rich and flexible:

Histogram prediction

The histogram representation is rich and flexible:

Training

- Start with an ImageNet pretrained network

Training

- Start with an ImageNet pretrained network

- Adapt to grayscale input

Training

- Start with an ImageNet pretrained network
- Adapt to grayscale input
- Fine-tune for colorization with log-loss on ImageNet without labels

Sparse Training

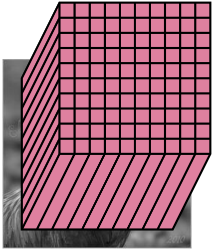
Trained as a fully convolutional network with:

Sparse Training

Trained as a fully convolutional network with:

Dense hypercolumns

- Low-level layers are upsampled
- ✗ High memory footprint

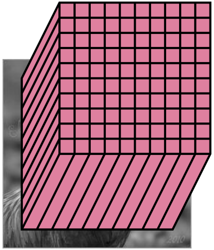


Sparse Training

Trained as a fully convolutional network with:

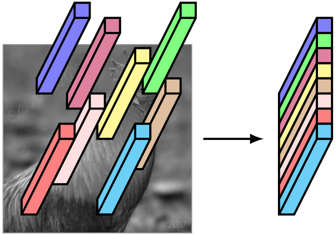
Dense hypercolumns

- Low-level layers are upsampled
- ✗ High memory footprint



Sparse hypercolumns

- Direct bilinear sampling
- ✓ Low memory footprint

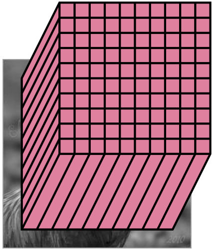


Sparse Training

Trained as a fully convolutional network with:

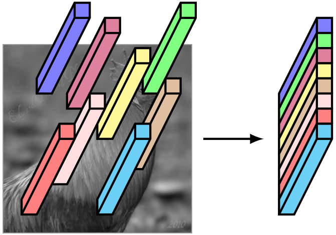
Dense hypercolumns

- Low-level layers are upsampled
- ✗ High memory footprint



Sparse hypercolumns

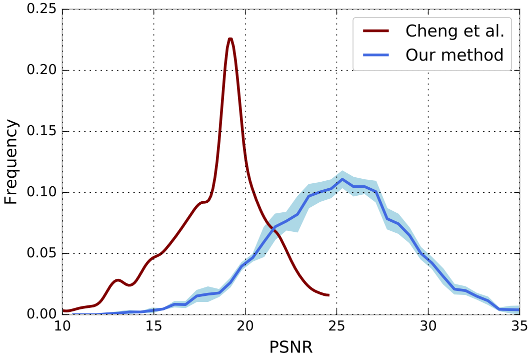
- Direct bilinear sampling
- ✓ Low memory footprint



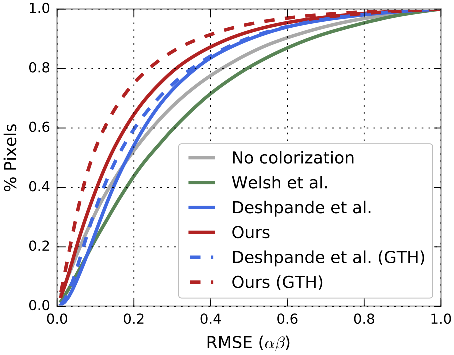
Source code available for Caffe and TensorFlow

Comparison: Previous work

Significant improvement over state-of-the-art:



vs. Cheng et al. (2015)



vs. Deshpande et al. (2015)

Comparison: Concurrent work

Model	MSE	PSNR
Zhang et al.	270.17	21.58
Baig et al.	194.12	23.72
Ours	**154.69**	**24.80**

Source: Baig and Torresani (2016) [Arxiv]

Comparison: Concurrent work

Model	MSE	PSNR
Zhang et al.	270.17	21.58
Baig et al.	194.12	23.72
Ours	**154.69**	**24.80**

Source: Baig and Torresani (2016) [Arxiv]

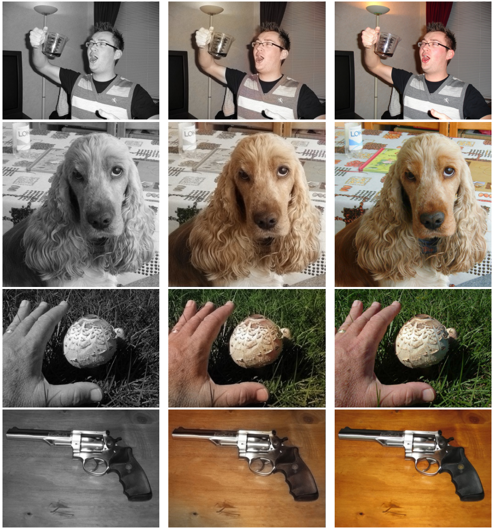
Model	AuC CMF non-rebal (%)	AuC CMF rebal (%)	VGG Top-1 Classification Accuracy (%)	Turk Labeled Real (%) mean	Turk Labeled Real (%) std
Ground Truth	100.00	100.00	68.32	50.00	–
Zhang et al.	91.57	65.12	56.56	25.16	2.26
Zhang et al. (rebal)	89.50	**67.29**	56.01	**32.25**	2.41
Ours	**91.70**	65.93	**59.36**	27.24	2.31

Source: Zhang et al. (2016) [ECCV]

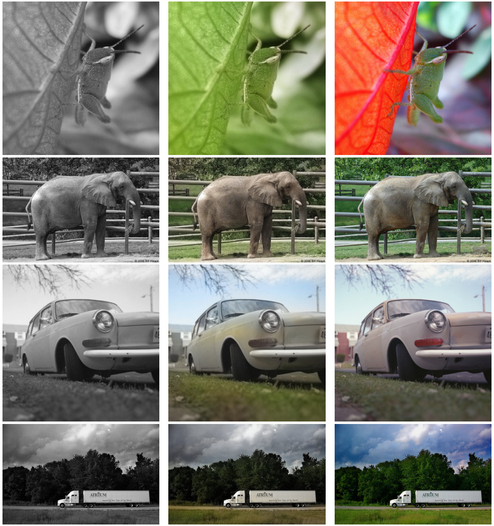

SCHOOL BUS
Copyright © 2007 James Markus - All Rights Reserved
Photomatter.com


Photomatter.com

Examples



Input Our Method Ground-truth

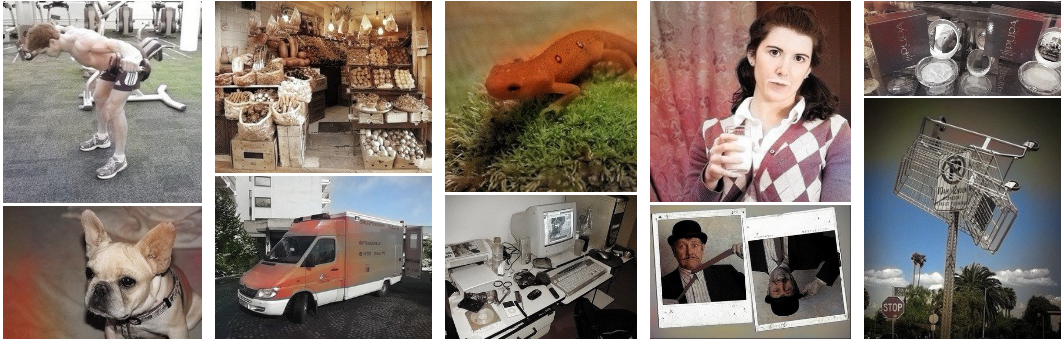


Input Our Method Ground-truth

Examples

B&W photographs

Examples



Failure modes

Self-supervision (ongoing work)

1. Train colorization from scratch

Self-supervision (ongoing work)

1. Train colorization from scratch

Initialization	RMSE	PSNR
ImageNet Classifier	0.299	24.45
Random	0.311	24.25

How much does ImageNet pretraining help colorization?

Self-supervision (ongoing work)

1. Train colorization from scratch

Initialization	RMSE	PSNR
ImageNet Classifier	0.299	24.45
Random	0.311	24.25

How much does ImageNet pretraining help colorization?

2. Use network for other tasks, such as semantic segmentation:

Self-supervision (ongoing work)

1. Train colorization from scratch

Initialization	RMSE	PSNR
ImageNet Classifier	0.299	24.45
Random	0.311	24.25

How much does ImageNet pretraining help colorization?

2. Use network for other tasks, such as semantic segmentation:

Initialization	X_{ImageNet}	Y_{ImageNet}	mIU (%)
ImageNet Classifier	✓	✓	64.0
Random			32.5

Pascal VOC 2012 segmentation val

Self-supervision (ongoing work)

1. Train colorization from scratch

Initialization	RMSE	PSNR
ImageNet Classifier	0.299	24.45
Random	0.311	24.25

How much does ImageNet pretraining help colorization?

2. Use network for other tasks, such as semantic segmentation:

Initialization	X_{ImageNet}	Y_{ImageNet}	mIU (%)
ImageNet Classifier	✓	✓	64.0
ImageNet Colorizer	✓		**50.2**
Random			32.5

Pascal VOC 2012 segmentation val

Summary

- Fully automatic colorization with state-of-the-art results
- Efficient training via sparse sampling of hypercolumns
- Promising proxy task for visual representation learning

See you at poster O-3A-04 upstairs!

Source code and demo available online:

colorize.ttic.edu
gustavla/autocolorize

```
pip install autocolorize
autocolorize grayscale.png -o color.png
```

Summary

- Fully automatic colorization with state-of-the-art results
- Efficient training via sparse sampling of hypercolumns
- Promising proxy task for visual representation learning

See you at poster O-3A-04 upstairs!

Source code and demo available online:

colorize.ttic.edu

gustavla/autocolorize

```
pip install autocolorize
autocolorize grayscale.png -o color.png
```

Summary

- Fully automatic colorization with state-of-the-art results
- Efficient training via sparse sampling of hypercolumns
- Promising proxy task for visual representation learning

See you at poster O-3A-04 upstairs!

Source code and demo available online:

demo colorize.ttic.edu

gustavla/autocolorize

```
pip install autocolorize
autocolorize grayscale.png -o color.png
```

References

Baig, M. H. and Torresani, L. (2016). Colorization for image compression. arXiv preprint arXiv:1606.06314.

Charpiat, G., Hofmann, M., and Schölkopf, B. (2008). Automatic image colorization via multimodal predictions. In ECCV.

Cheng, Z., Yang, Q., and Sheng, B. (2015). Deep colorization. In ICCV.

Chia, A. Y.-S., Zhuo, S., Gupta, R. K., Tai, Y.-W., Cho, S.-Y., Tan, P., and Lin, S. (2011). Semantic colorization with internet images. ACM Transactions on Graphics (TOG), 30(6).

Deshpande, A., Rock, J., and Forsyth, D. (2015). Learning large-scale automatic image colorization. In ICCV.

Huang, Y.-C., Tung, Y.-S., Chen, J.-C., Wang, S.-W., and Wu, J.-L. (2005). An adaptive edge detection based colorization algorithm and its applications. In ACM international conference on Multimedia.

Iizuka, S., Simo-Serra, E., and Ishikawa, H. (2016). Let there be Color!: Joint End-to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification. ACM Transactions on Graphics (Proc. of SIGGRAPH 2016), 35(4).

Irony, R., Cohen-Or, D., and Lischinski, D. (2005). Colorization by example. In Eurographics Symp. on Rendering.

Larsson, G., Maire, M., and Shakhnarovich, G. (2016). Learning representations for automatic colorization. ECCV.

Levin, A., Lischinski, D., and Weiss, Y. (2004). Colorization using optimization. ACM Transactions on Graphics (TOG), 23(3).

Luan, Q., Wen, F., Cohen-Or, D., Liang, L., Xu, Y.-Q., and Shum, H.-Y. (2007). Natural image colorization. In Eurographics conference on Rendering Techniques.

Morimoto, Y., Taguchi, Y., and Naemura, T. (2009). Automatic colorization of grayscale images using multiple images on the web. In SIGGRAPH: Posters.

Qu, Y., Wong, T.-T., and Heng, P.-A. (2006). Manga colorization. ACM Transactions on Graphics (TOG), 25(3).

Welsh, T., Ashikhmin, M., and Mueller, K. (2002). Transferring color to greyscale images. ACM Transactions on Graphics (TOG), 21(3).

Zhang, R., Isola, P., and Efros, A. A. (2016). Colorful image colorization. ECCV.