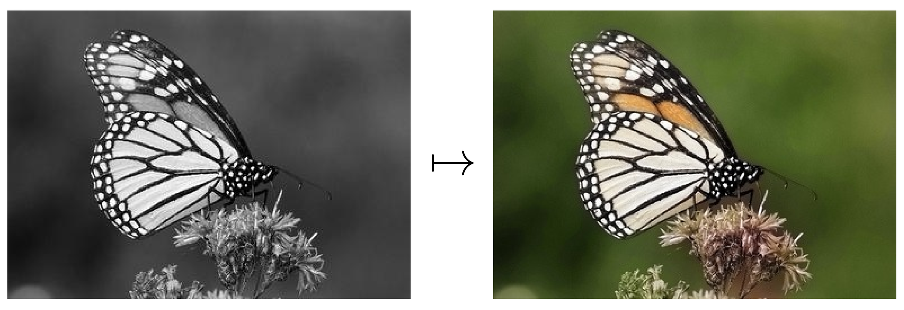
Learning Representations for Automatic Colorization

Gustav Larsson — University of Chicago
Michael Maire — TTI Chicago
Greg Shakhnarovich — TTI Chicago

Poster: O-3A-04

Overview

Problem statement:
- **Input** Grayscale image
- **Output** Plausible and pleasing color rendition



Background

- Three methods in the literature:
  - **Scribble based** User colorizes some pixels, the algorithm fills in the rest
  - **Transfer based** Color is transfered from a reference image
  - **Fully automatic** Parametric model that predicts color directly
- Recent interest in fully automatic colorization:
  - Deshpande et al. (ICCV 2015), Cheng et al. (ICCV 2015), Iizuka (SIGGRAPH 2016), Zhang et al. (ECCV 2016)

Our work

- Design principles:
  - **Semantic cues** → leverage ImageNet-based classifier
  - **Low-level/high-level cues** → Zoom-out/hypercolumn architecture
  - **Colorization not unique** → Predict color histograms
- Fully automatic. Takes less than 1 second per image.
- High-quality results. State-of-the-art on all quantitative measures.

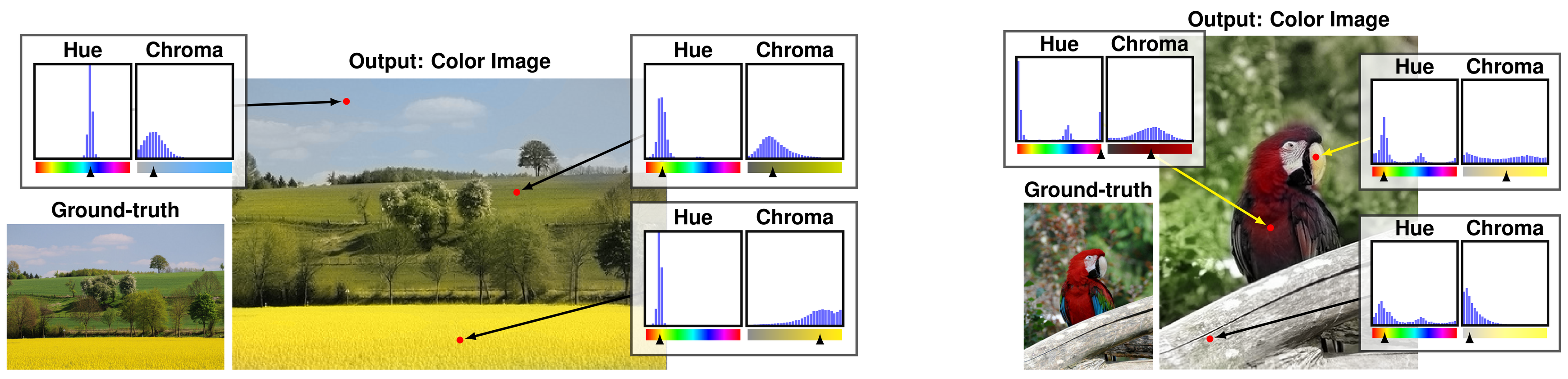
Old B&W photographs automatically colorized

Training

We employ a fully convolutional network and consider several training formulations:

- Color space
  - L*a*b
  - Hue/chroma
- Loss function
  - Regression (L_2)
  - Histogram predictions (K bins)
- Spatial sampling
  - Dense
  - Sparse

We chose **hue/chroma** with **histogram predictions** at **sparse** locations.



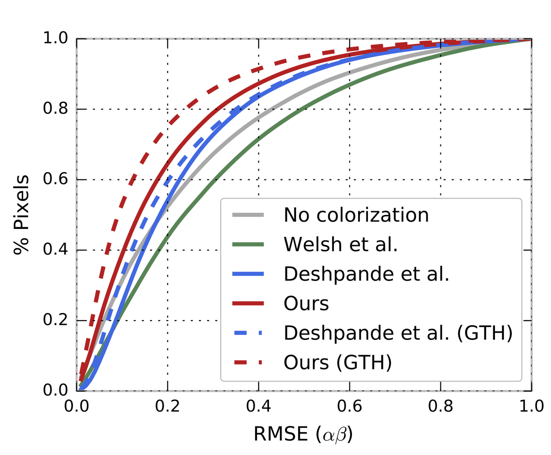
Experimental Results

Datasets

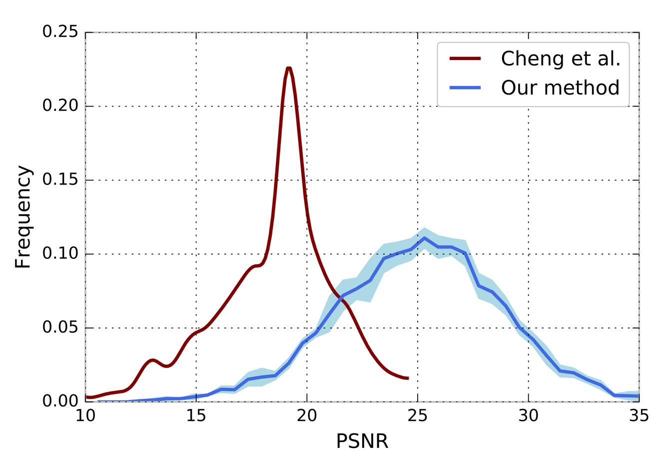
- **ImageNet** 1.3M images, 1000 categories
  - **cval1k** 1000 images for validation
- **SUN Database** Two subsets are used:
  - **SUN-A** 47 object categories
  - **SUN-6** 6 scene categories

Metrics

- **RMSE** Per-pixel root mean square error in the $\alpha\beta$ color space
- **PSNR** Peak signal-to-noise ratio in the RGB color space



SUN-6. Cumulative histogram of RMSE



SUN-A. Histogram of per-image PSNR

ImageNet/cval1k	RMSE	PSNR
No colorization	0.343	22.98
Lab, L_2	0.318	24.25
Lab, $K=32$	0.321	24.33
Lab, $K=16\times16$	0.328	24.30
Hue/chroma, $K=32$	0.342	23.77
+ chromatic fading	**0.299**	**24.45**

SUN-6	RMSE
No colorization	0.285
Welsh *et al.*	0.353
Deshpande *et al.*	0.262
+ GT Scene	0.254
Our Method	**0.211**

SUN-6 (GT Hist)	RMSE
Deshpande *et al.* (C)	0.236
Deshpande *et al.* (Q)	0.211
Our Method (Q)	0.178
Our Method (E)	**0.165**

GT Hist: Ground-truth global histogram available

Colorization as Self-supervision

- Preliminary results show that colorization can be trained from scratch (*left*)
- That model can then be used as a replacement for ImageNet pretraining (*right*)

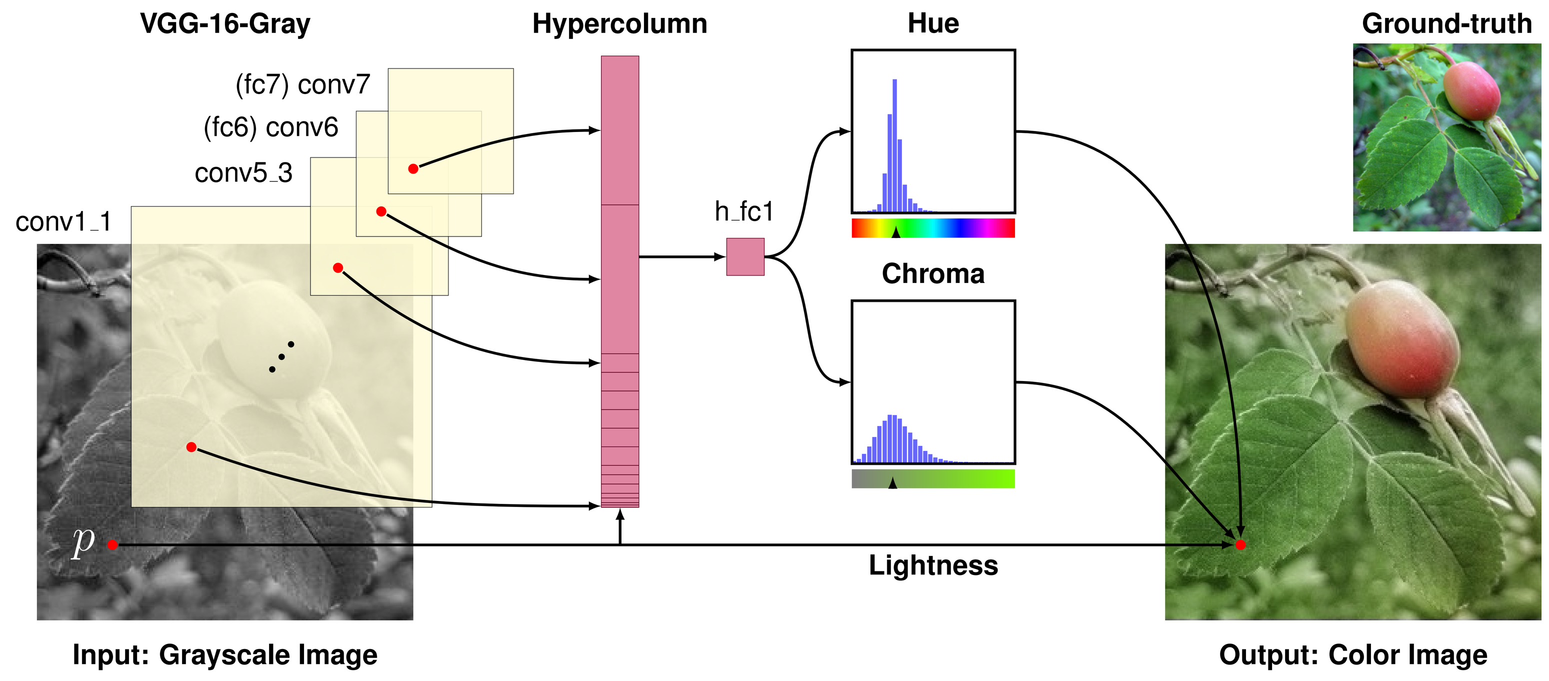
Initialization	RMSE	PSNR
Classifier	0.299	24.45
Random	0.311	24.25

ImageNet/cval1k. ImageNet pretraining helps, but is not essential.

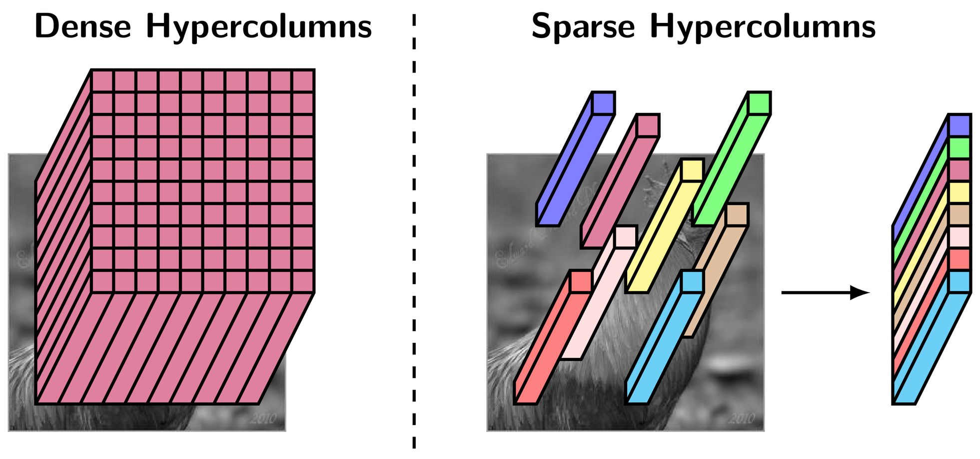
Initialization	X_{ImageNet}	Y_{ImageNet}	mIU (%)
Classifier	✓	✓	64.0
Colorizer	✓		50.2
Random			32.5

VOC 2012 segmentation validation set.

Network Architecture

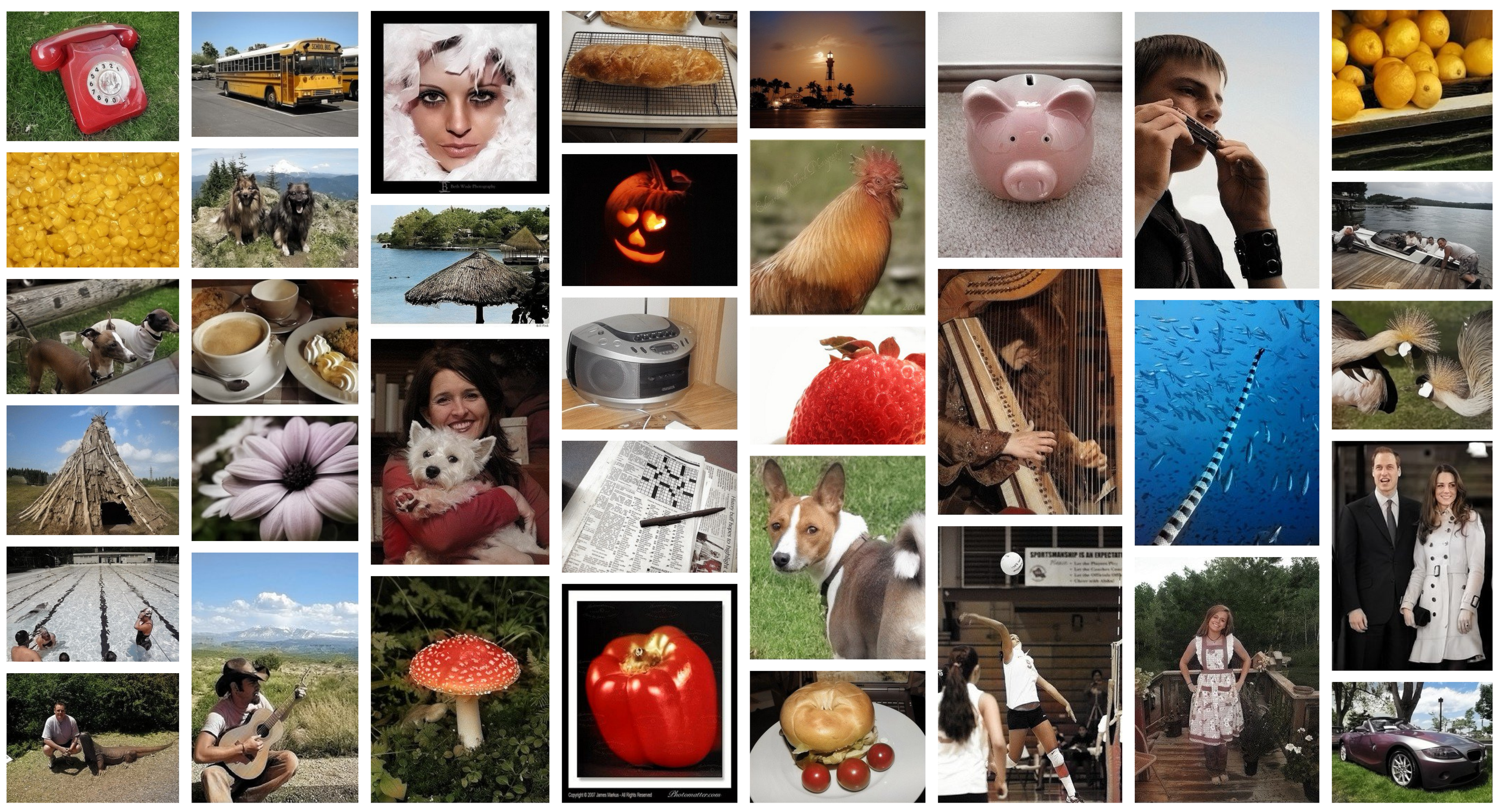


Sparse Training



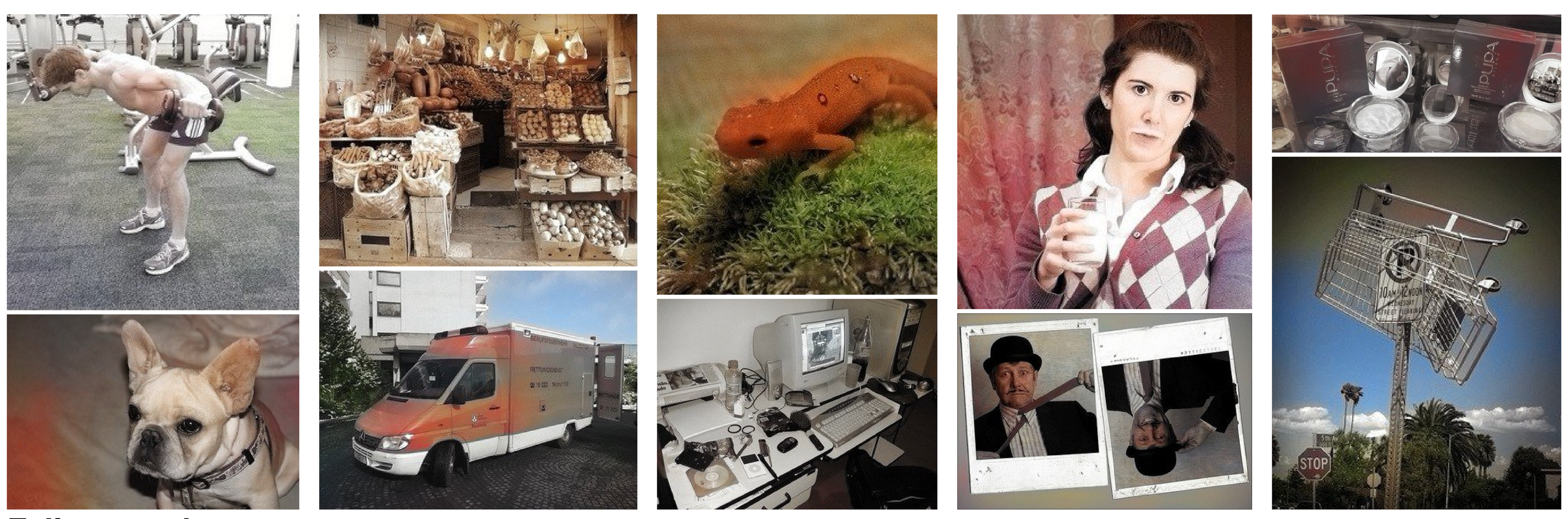
- Dense
  - Low-level layers are upsampled
  - High memory footprint
- Sparse
  - Direct bilinear sampling
  - Low memory footprint
- Allows larger images / more images per batch
- Widely applicable to image-to-image tasks
- Code available for Caffe/TensorFlow

Examples



Automatically colorized photos

Sampling multiple colorizations using the rich histogram representation's color uncertainty



Failure modes

colorize.ttic.edu
gustavla/autocolorize

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