

PixColor: Pixel Recursive Colorization

Appendix

Sergio Guadarrama
sguada@google.com

Ryan Dahl
rld@google.com

David Bieber
dbieber@google.com

Mohammad Norouzi
mnorouzi@google.com

Jonathon Shlens
shlens@google.com

Kevin Murphy
kpmurphy@google.com

Google Research
Mountain View, USA

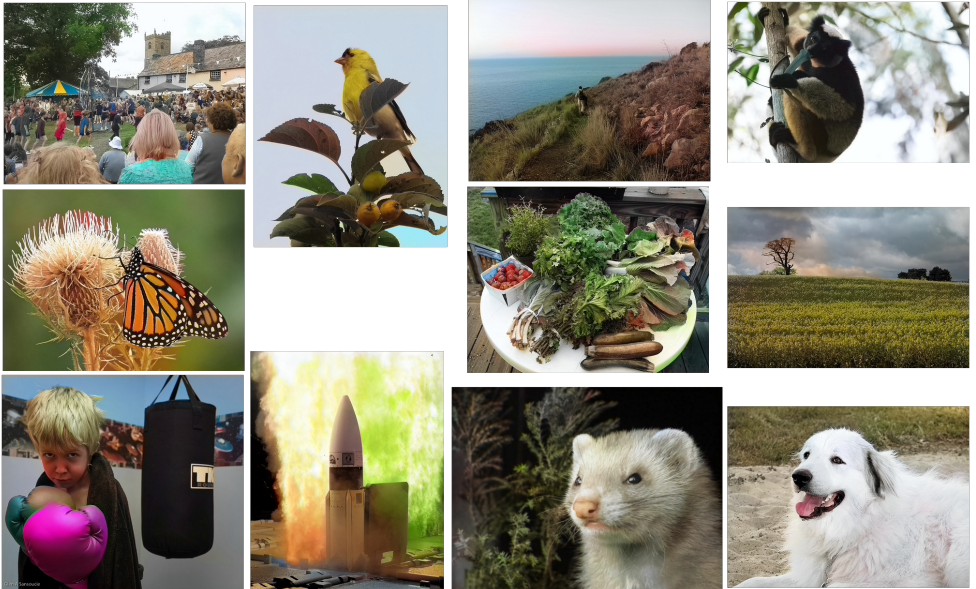


Figure 1: Selected high resolution and non-square samples.

Operation	Kernel	Strides	Feature maps
PixelCNN conditioning network	$224 \times 224 \times 3$ input		
Conv2D	7×7	2	64
$3 \times$ ResNet block	3×3	2, 1, 1	64/256 bottleneck
$4 \times$ ResNet block	3×3	2, 1, 1, 1	128/512 bottleneck
$23 \times$ ResNet block	3×3	1	256/1024 bottleneck
$3 \times$ Conv2D	3×3	1	64
Gradient Multiplier	Engaged at 100,000 steps		
PixelCNN colorization network	$28 \times 28 \times 2$ input		
Masked Conv2D	7×7	1	64
$10 \times$ Gated Conv2D Blocks	5×5	1	64
Masked Conv2D	1×1	1	1024
Masked Conv2D	1×1	1	$2 * 32$
Refinement Network	$224 \times 224 \times 3$ input		
Conv2D	3×3	2	64
$2 \times$ Conv2D	3×3	1, 2	128
$2 \times$ Conv2D	3×3	1, 2	256
$2 \times$ Conv2D	3×3	1	512
Conv2D	3×3	1	256
$4 \times$ Conv2D	3×3	2, 1, 2, 1	512
Conv2D	5×5	1	1024
Conv2D	3×3	1	512
$2 \times$ Conv2D	3×3	1	128
Bilinear Upsample			
$2 \times$ Conv2D	3×3	1	64
Bilinear Upsample			
$2 \times$ Conv2D	3×3	1	32
Conv2D	1×1	1	2
Optimizer	Adam (beta1=0.9, momentum=0.9)		
Batch size	$64 = 8 \times 8$ (8 GPUs, synchronous updates)		
Iterations	360,000		
Learning Rate	0.0003		
Colorness Threshold	0.05		
Weight, bias initialization	Truncated normal (stddev=0.1), Constant(0)		

Table 1: Hyperparameters used. The PixelCNN and Refinement networks were trained independently.

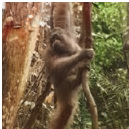
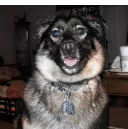
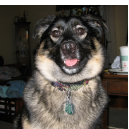
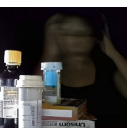
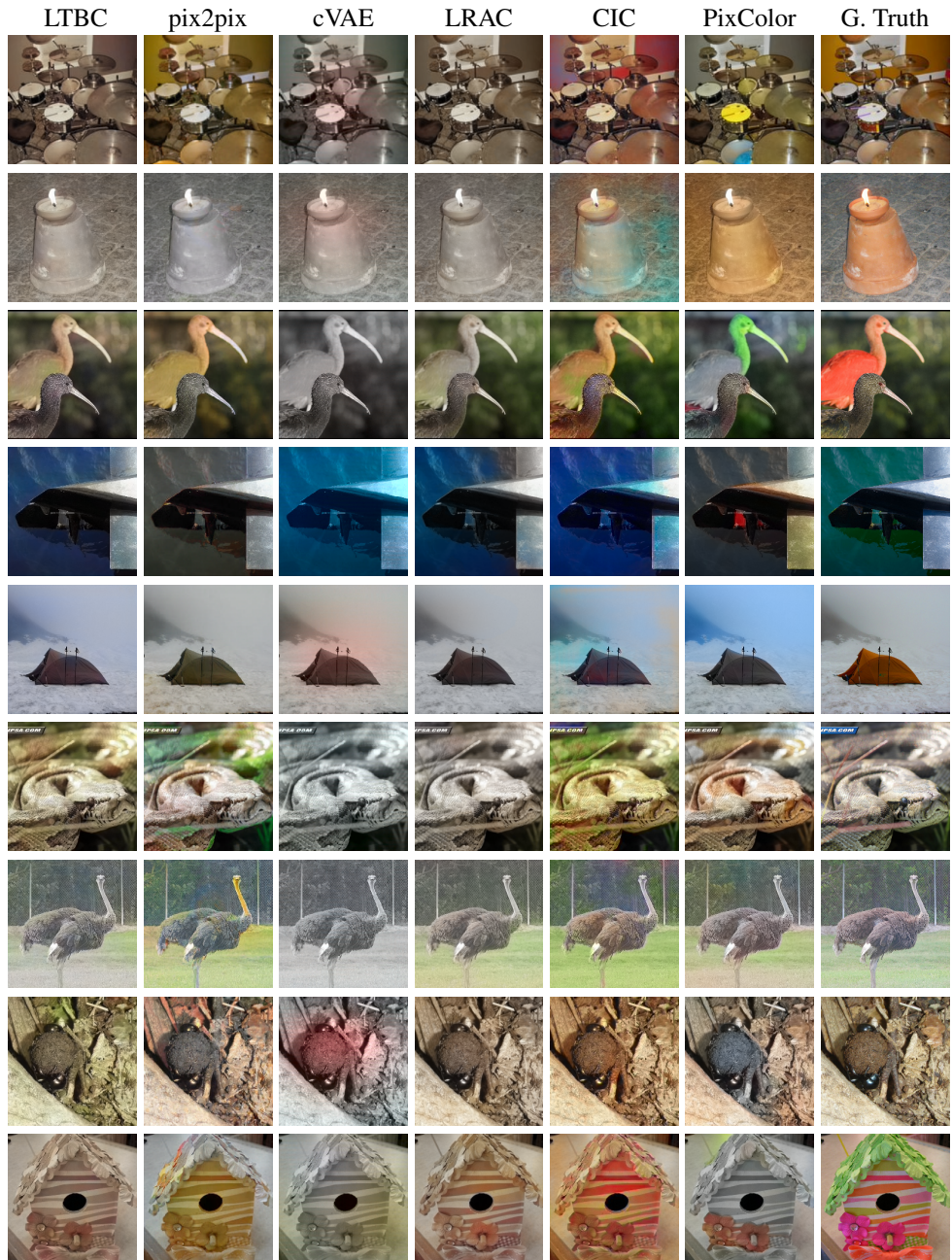
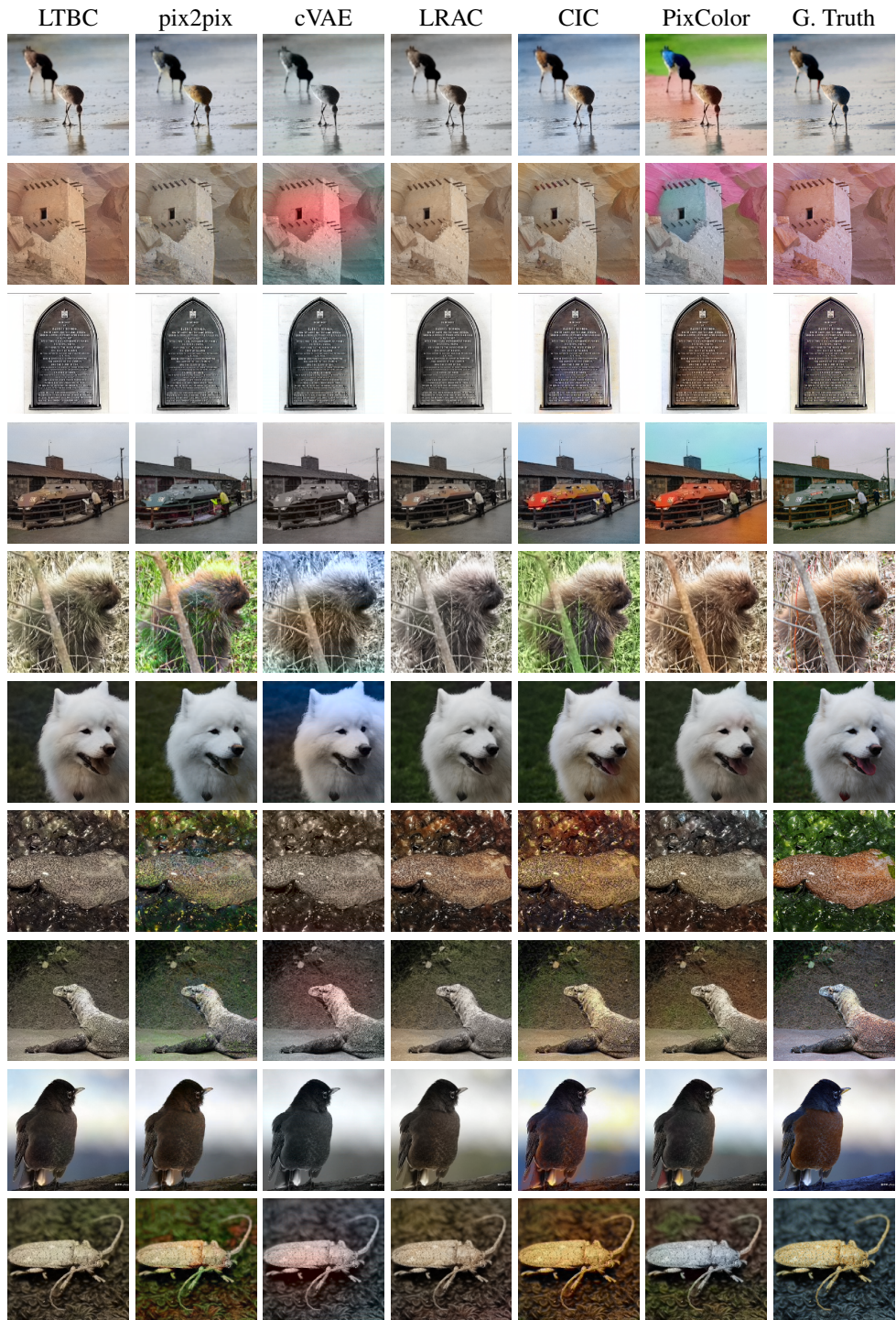
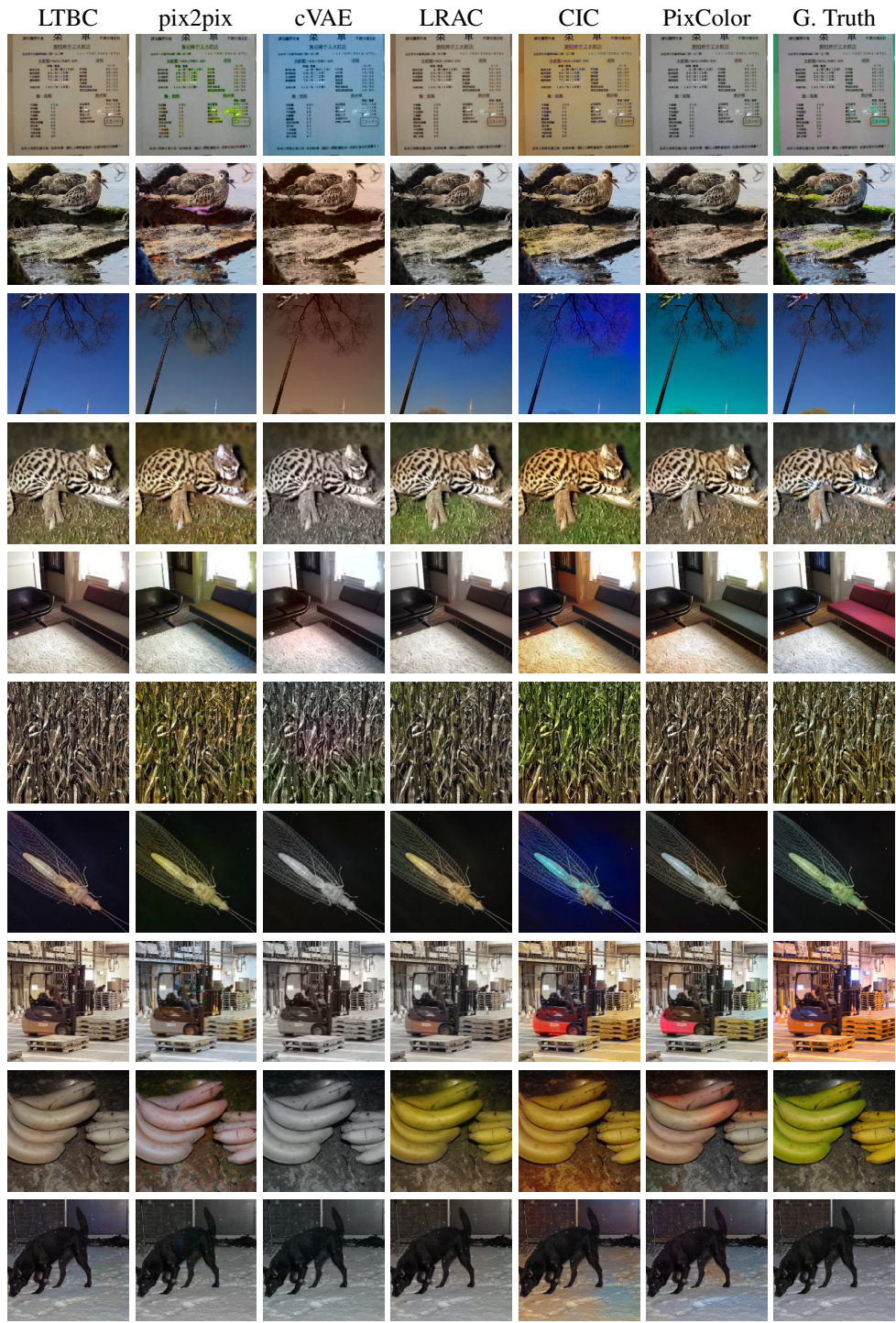
VTT Score	PixColor	G. Truth	PixColor	G. Truth	PixColor	G. Truth
0%						
20%						
40%						
60%						
80%						
100%						

Figure 2: Images illustrating each possible VTT score.

Table 2: More qualitative side by side comparison of colorizations produced by various methods (LTBC: [1], pix2pix: [1], VAE: [1], LRAC: [1], CIC: [1], PixColor: this paper, G. Truth: original color). These images are *randomly sampled* from the ImageNet test set.









References

- [1] Aditya Deshpande, Jiajun Lu, Mao-Chuang Yeh, and David A. Forsyth. Learning diverse image colorization. *arXiv:1612.01958*, 2016.
- [2] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. 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*, 35(4), 2016.
- [3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *arxiv:1611.07004*, 2016.
- [4] Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. Learning representations for automatic colorization. *ECCV*, 2016.
- [5] Richard Zhang, Phillip Isola, and Alexei A Efros. Colorful image colorization. *ECCV*, 2016.