Application of Deep Learning on Image Deblurring

Mattia Orlandi & Giacomo Pinardi

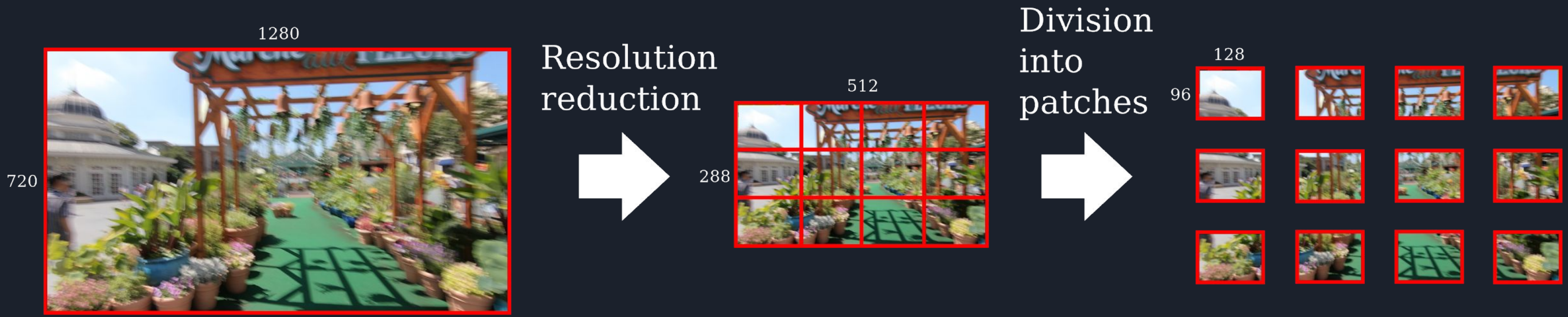
Contents

Datasets - CIFAR10

- Constructed by applying Gaussian blur with random standard deviation between 0 and 3
- Basic data augmentation: horizontal and/or vertical flip

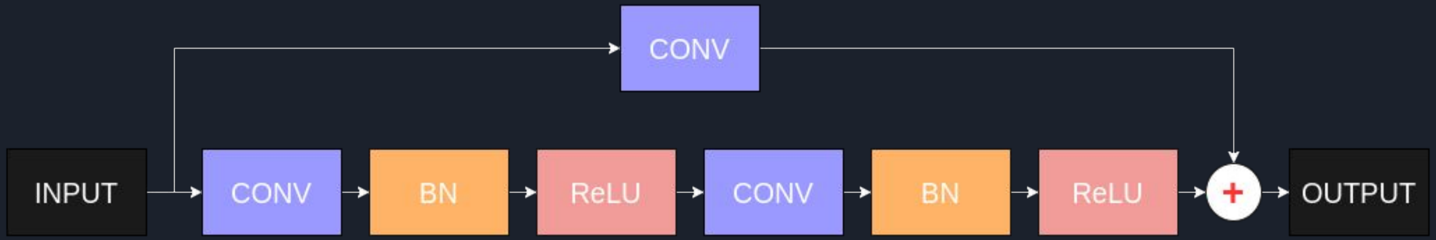
Datasets - REDS

- Resolution reduction from 1280x720 to 512x288
- Patch-based approach: each image is splitted into 12 patches analysed separately
- Basic data augmentation: horizontal and/or vertical flip



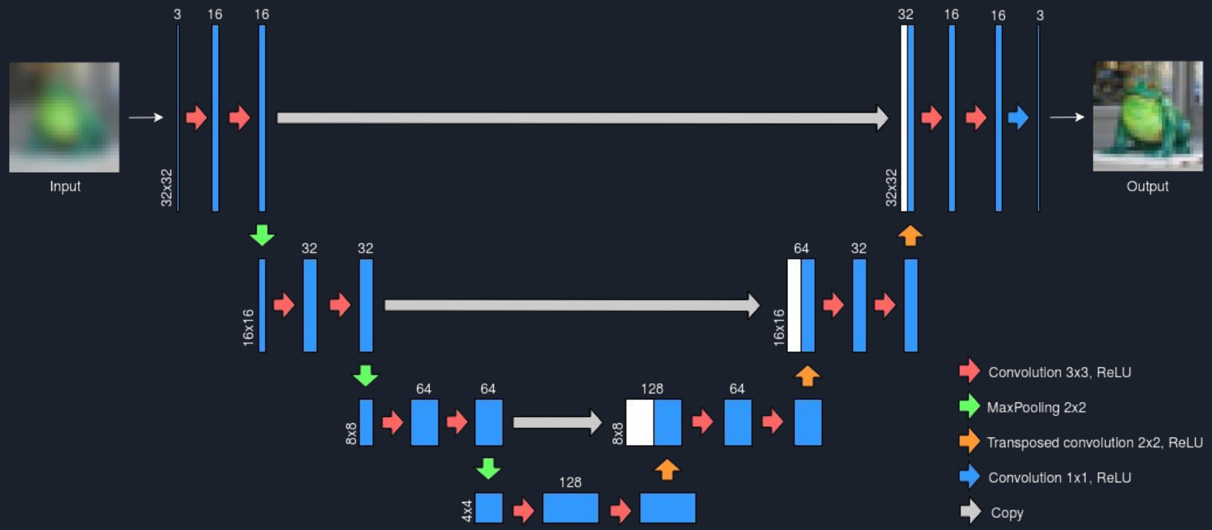
Convolutional Autoencoders - ResNet16

- It comprises an encoder and a decoder, each made of 4 residual blocks:
  - the encoder shrinks the input image using standard strided convolutions
  - the decoder enlarges the latent image using transposed strided convolutions
- Residual connections inspired by ResNet50 (for image classification)



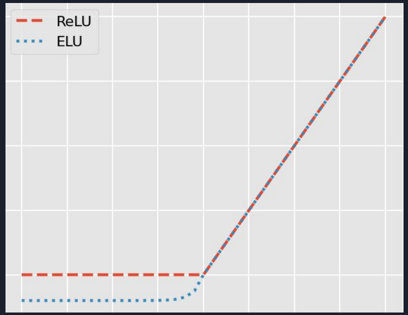
Convolutional Autoencoders - UNet16

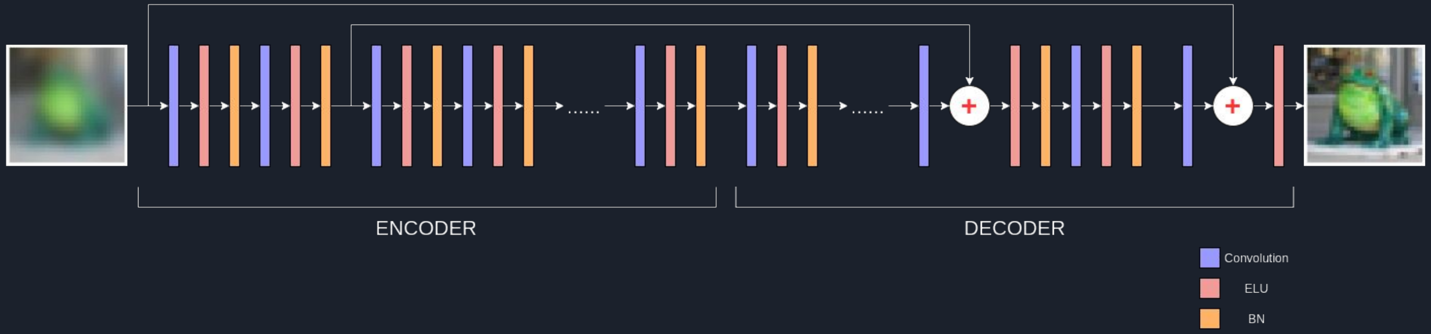
- Based on UNet model for image segmentation:
  - contracting path (convolutional layers + max pooling)
  - expansive path (common + transposed convolutions)



Convolutional Autoencoders - REDNet30

- 30 blocks made of:
  - Convolution 3x3 with stride 1 and padding
  - Exponential Linear Unit
  - Batch Normalization
- Symmetrical residual connections every 2 blocks
- No bottleneck: spatial size remains constant

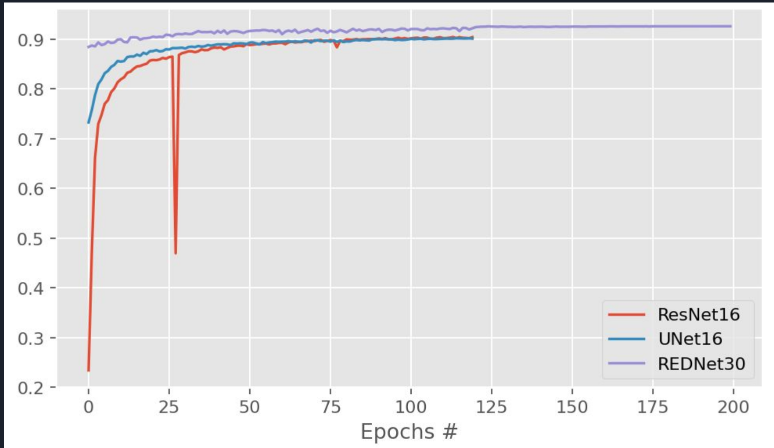




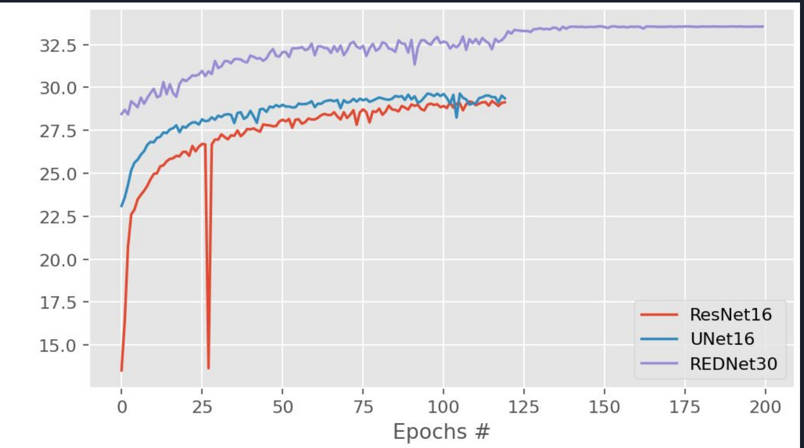
(CAE) Training phase on CIFAR-10

- Batch size: 32 (x8 TPU's replicas)
- 120 epochs (+ extra 80 epochs for REDNet30)
- Adam optimizer
- Learning rate: 0.001 (+ exp. decay in REDNet30's last 80 epochs)
- LogCosh loss function (more robust than MSE and differentiable unlike MAE)
- SSIM and PSNR metrics to assess image similarity

SSIM



PSNR



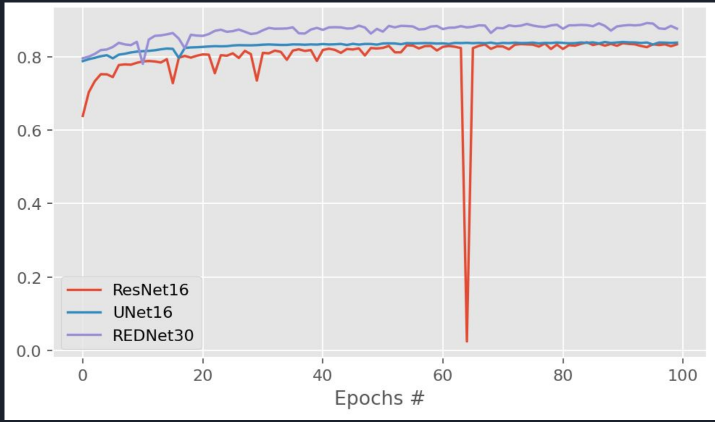
(CAE) Experimental results on CIFAR-10

Model	ResNet16	UNet16	**REDNet30**	Baseline
Loss [10^{-4}]	9.794	10.44	**6.966**	34.48
SSIM	0.9034	0.9001	**0.9258**	0.7135
PSNR	29.09	29.32	**33.52**	24.67
MSE [10^{-3}]	1.965	2.095	**1.398**	6.317
MAE [10^{-2}]	2.888	2.898	**2.165**	4.517

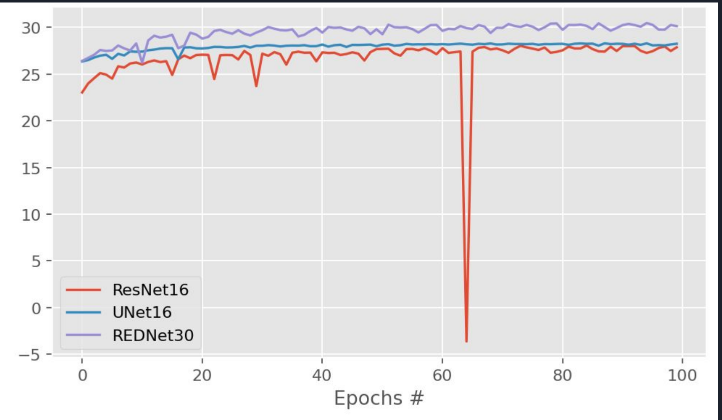
(CAE) Training phase on REDS

- 100 epochs instead of 120
- Other parameters remained unchanged

SSIM



PSNR

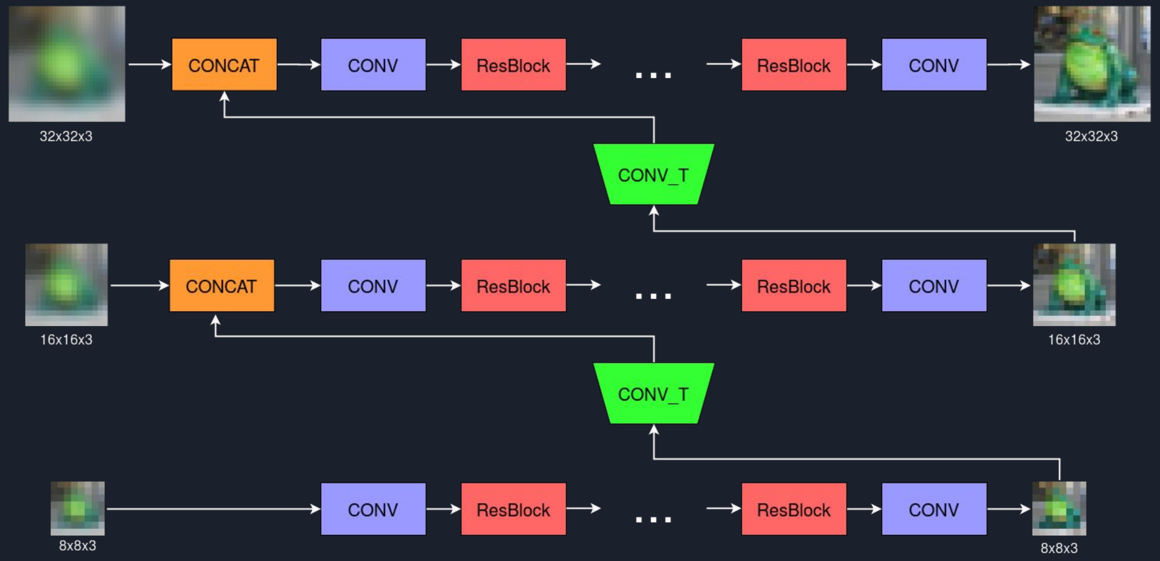


(CAE) Experimental results on REDS

Model	ResNet16	UNet16	**REDNet30**	Baseline
Loss [10^{-4}]	8.669	8.763	**5.793**	19.54
SSIM	0.8554	0.8646	**0.8976**	0.8237
PSNR	29.49	30.41	**31.98**	28.80
MSE [10^{-3}]	1.742	1.761	**1.163**	2.594
MAE [10^{-2}]	2.536	2.312	**1.884**	2.118

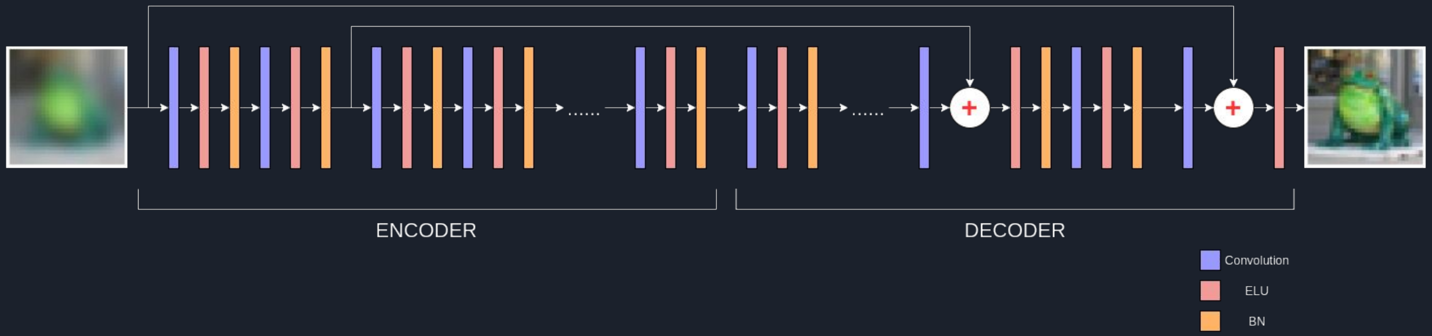
WGANs - MSDeblurWGAN generator

- It comprises three branches, each made of 19 ResBlocks:
  - branch for coarsest resolution
  - branch for middle resolution (which relies on the coarsest one)
  - branch for finer resolution (which relies on the middle one)



WGANs - REDNet30WGAN generator

- The architecture of the generator is the same as the one of REDNet30, since it performed well as a standard CAE



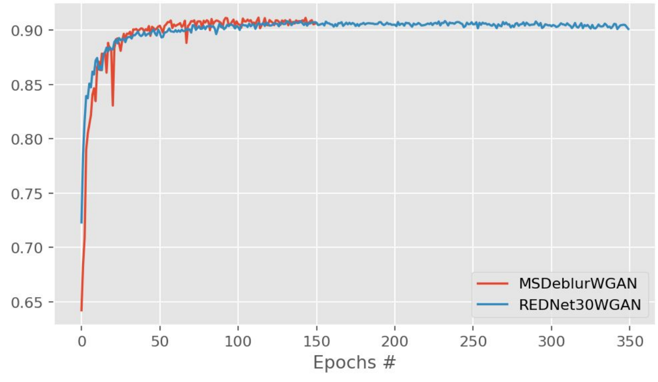
WGANs - PatchGAN critic

- The discriminator in WGANS (called *critic*) computes a 1-Lipschitz continuous function:
  - apply basic 4x4 convolution with stride 2
  - use of Layer Normalization
- To obtain the image score the last 16 values are averaged to get a single value

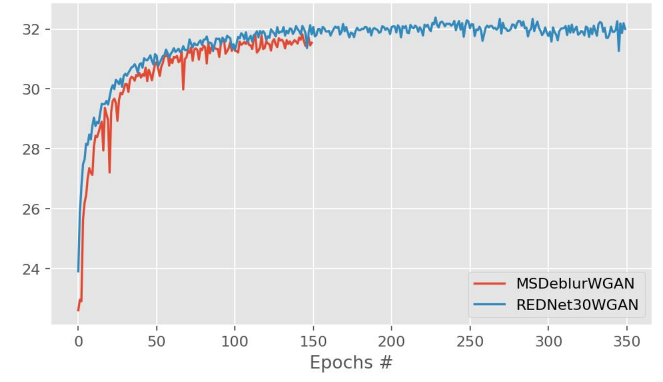
(WGAN) Training phase on CIFAR-10

- Batch size: 32 (x8 TPU's replicas)
- 150 epochs for MSDeblurWGAN, 350 epochs for REDNet30WGAN
- Adam optimizer
- Learning rate: $2 \bullet 10^{-4}$
- Content loss: LogCosh (multi-scale version for MSDeblurWGAN)
- Adversarial loss: negated critic's score
- SSIM and PSNR metrics to assess image similarity
- Gradient penalty technique to enforce 1-Lipschitz constraint
- At every generator's train step, the critic is trained 5 times

SSIM



PSNR



(WGAN) Experimental results on CIFAR-10

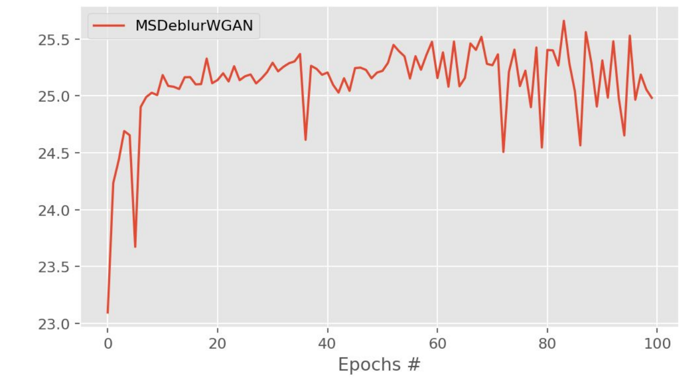
Model	MSDeblur WGAN	REDNet30 WGAN	**REDNet30**	Baseline
Loss	0.1254	-0.0057	**---**	---
SSIM	0.9054	0.9012	**0.9258**	0.7135
PSNR	31.49	31.98	**33.52**	24.67
MSE [10^{-3}]	1.734	1.900	**1.398**	6.317
MAE [10^{-2}]	2.471	2.482	**2.165**	4.517

(WGAN) Training phase on REDS

- 100 epochs (MSDeblurWGAN only)
- Number of ResBlocks reduced from 19 to 5
- Lower image resolution (256x144 instead of 512x288)

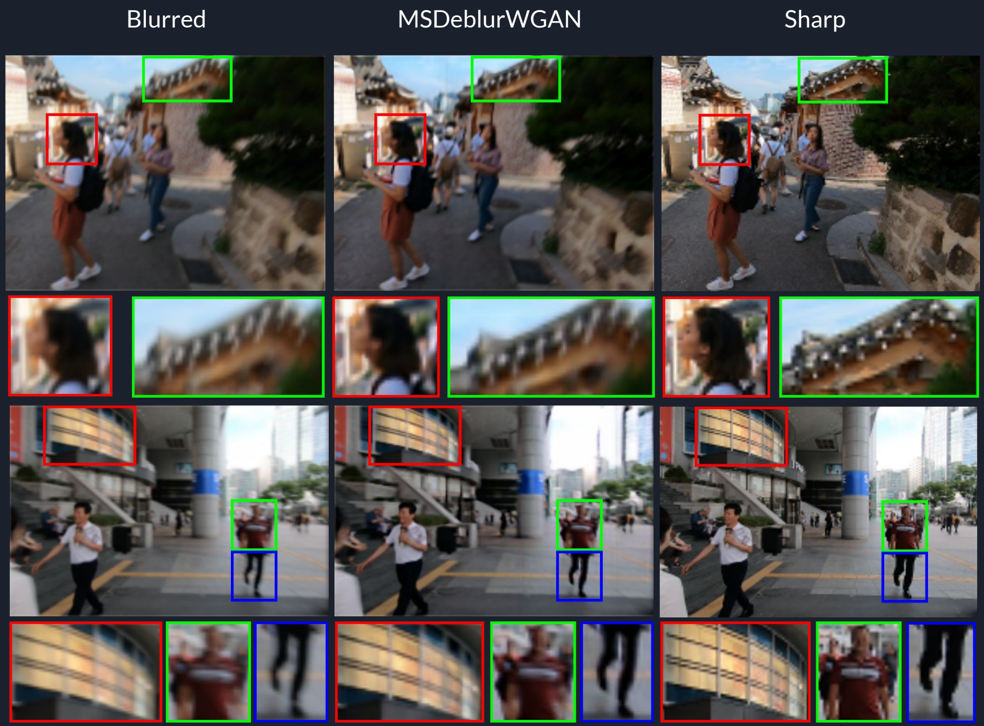
SSIM

PSNR



(WGAN) Experimental results on REDS

Model	MSDeblur WGAN	**REDNet30**	Baseline
Loss	0.2195	**---**	---
SSIM	0.8031	**0.8976**	0.8359
PSNR	27.11	**31.98**	27.68
MSE [10^{-3}]	1.729	**1.163**	3.120
MAE [10^{-2}]	2.495	**1.884**	2.390



Tools

- Google Colab’s TPU:
  - training performed in parallel across 8 replicas
  - results are reduced into a single value
- Files in tfrecord format on Google Cloud Storage
- Brain floating point:
  - half bits employed (16) w.r.t. standard 32-bit floating point
  - same dynamic range

REDNet30 on GOPRO

- REDNet30 trained on REDS dataset
- Does not reach Nah's DeepDeblur's performance
- Still able to reconstruct smaller details in GOPRO test images
- Some generalization capabilities

Blurred | Nah et al. | REDNet30