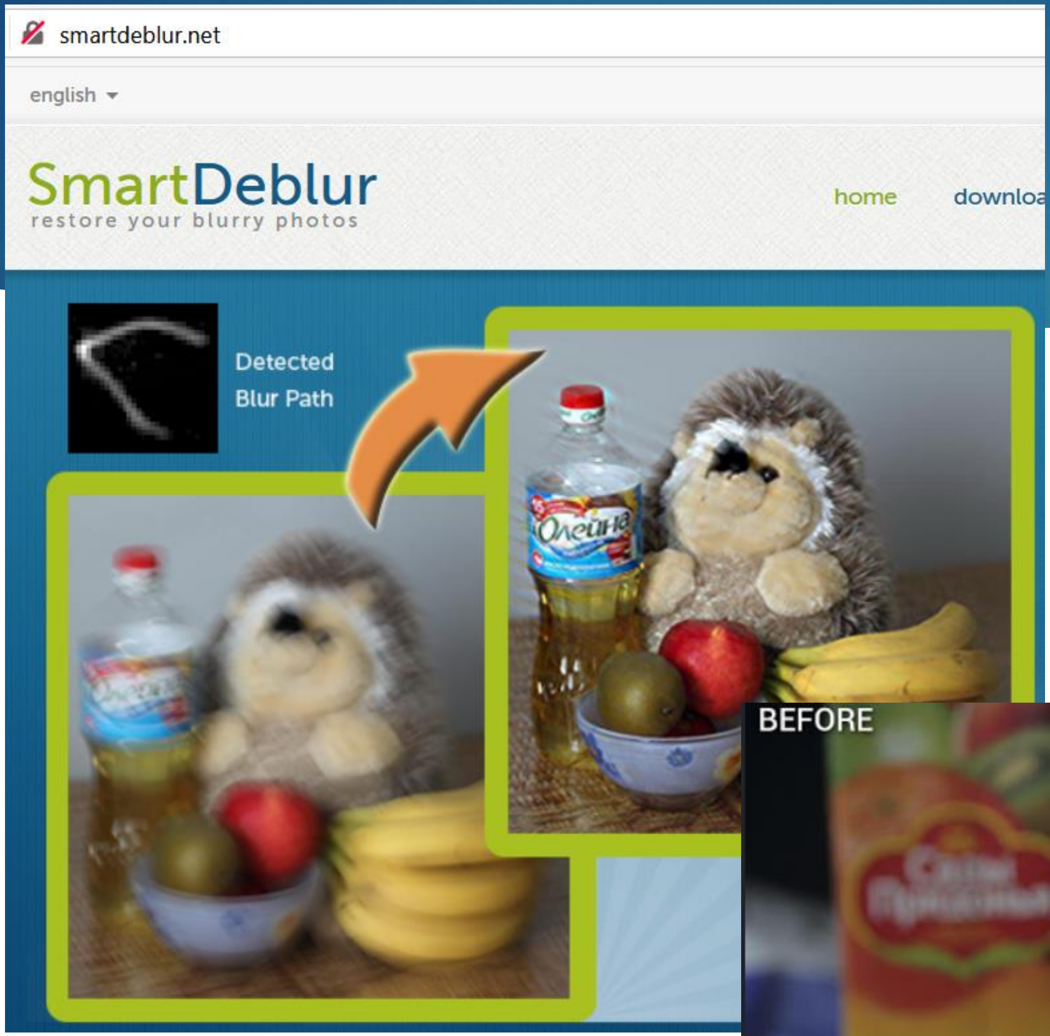
Digital Image Processing Final Project Proposal

0860801 Felipe Manosalvas
0860802 Edwin Arkel Rios
0860828 Eduin Hernandez

Group No. 19

Preface


smartdeblur.net
english
SmartDeblur
restore your blurry photos
home
Detected
Blur Path


BEFORE
AFTER
Сады
Придонья

DeblurGAN-v2 Paper

Introduction

Image Deblurring

- Progress of GANs

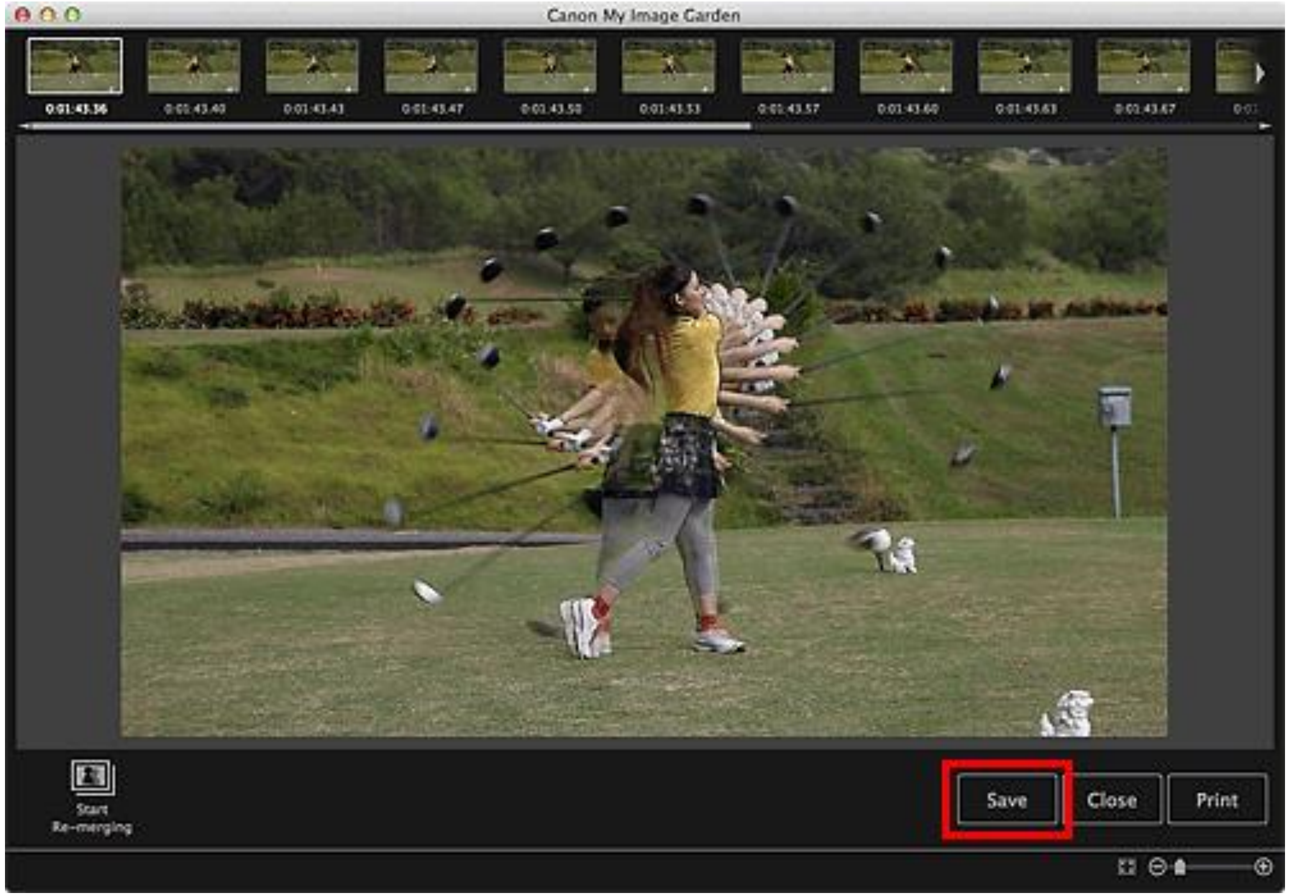
$$I_B = k(M) * I_S + N,$$

- **The project is focused in three main parts :**

1. Generate blurred images
2. A method to deblur
3. A novel method to quantify how good is a deblurred algorithm

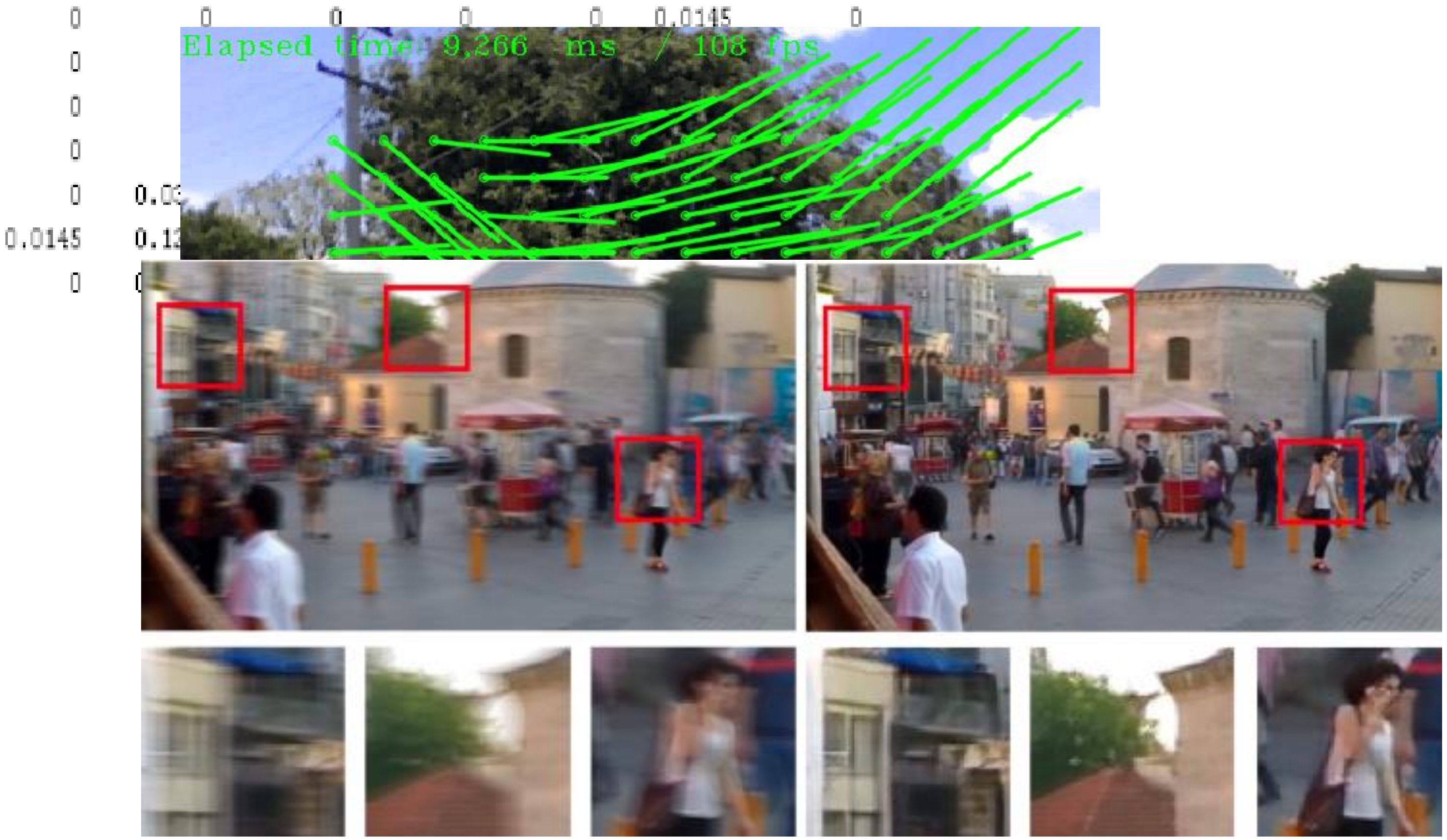
Blurred Image Generation

- Difficult to Dataset
- Take from video
- Convolving sharp images
- A novel method



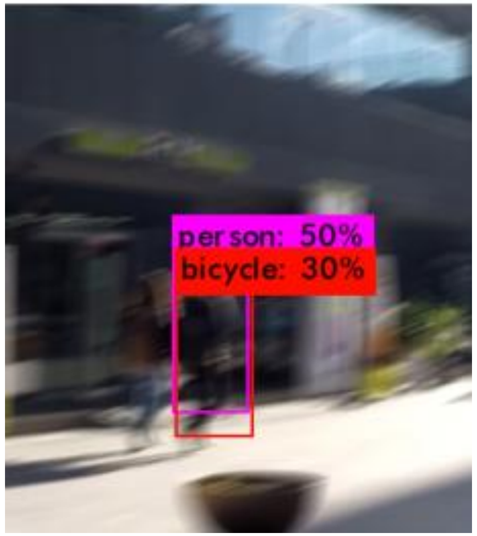
Deblurring Methods

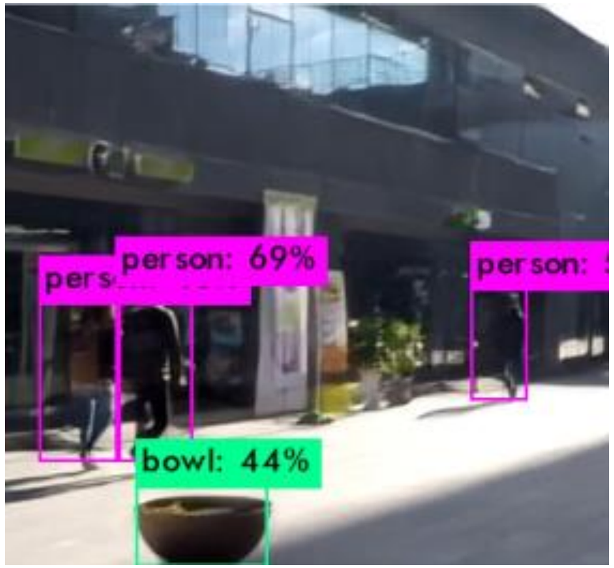
- Estimate the blur kernel
- Obtain the motion flow estimation
- Perform blind kernel-free image deblurring



Deblurring Algorithm Performance

- Run an object detection benchmark on yolo

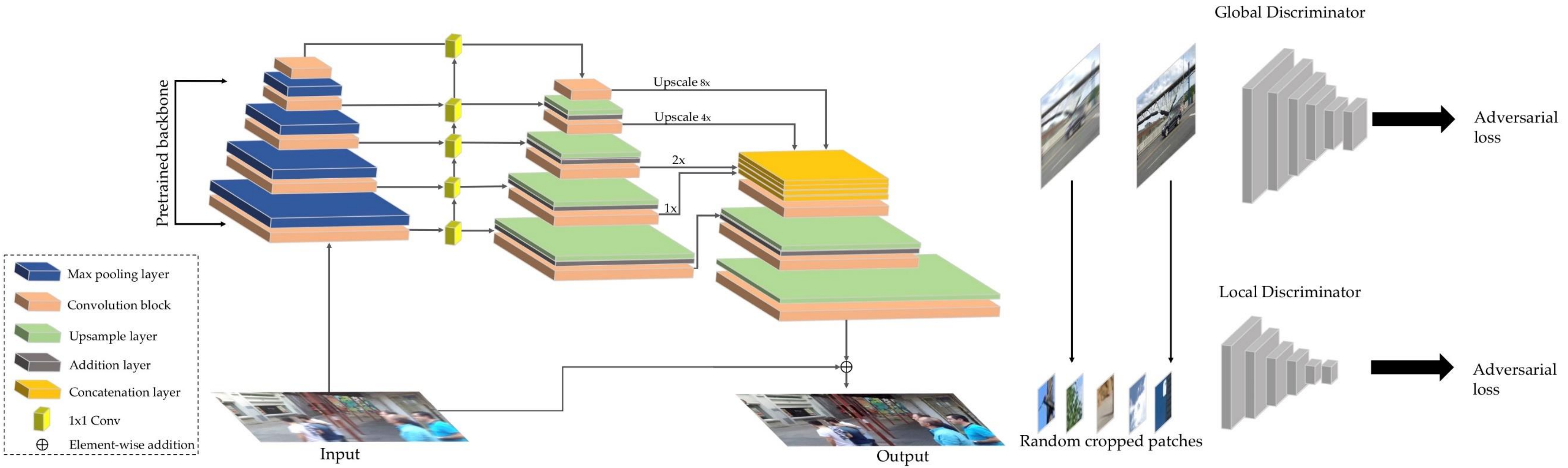




Architecture

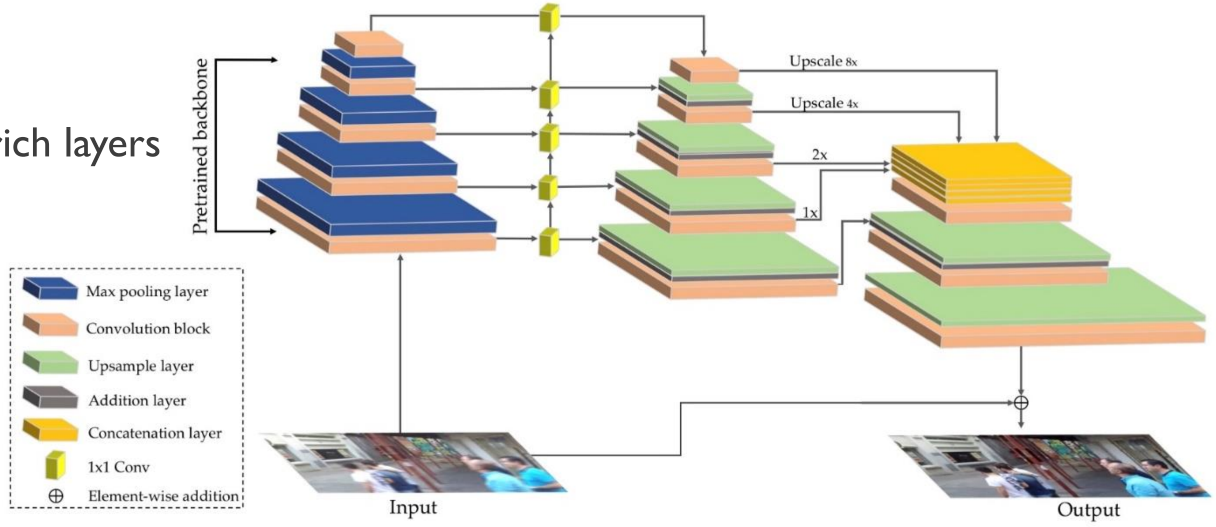
Feature Pyramid Deburring

"A lighter-weight alternative to incorporate multi-scale features"
-Kupyn et al.



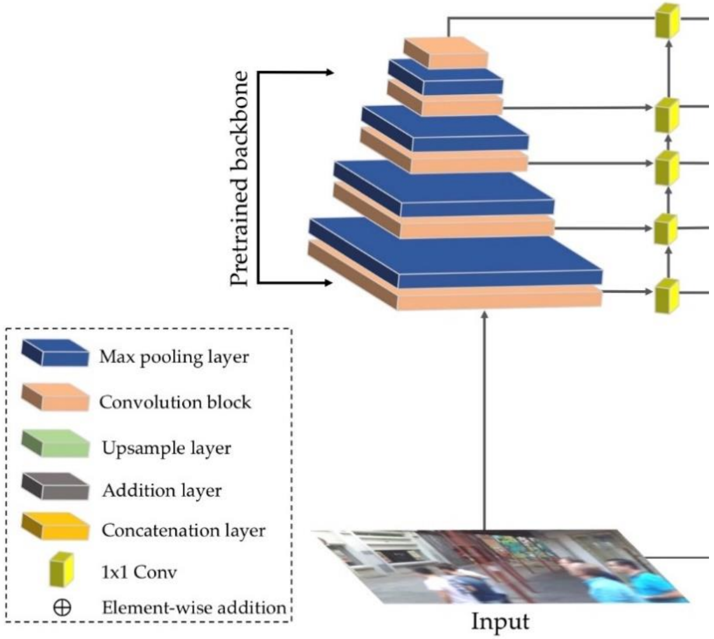
Feature Pyramid Deburring

- Bottom-up pathway
  - Convolutional network for feature extraction
  - Down sampling of special resolution
  - Semantic context information is extracted and compressed
- Top-down pathway
  - Reconstruction of higher special resolution from semantic rich layers
- Lateral Connections
  - Supplement high-resolution details and help localize objects



Backbones

- Inception-ResNet-v2
  - Pursues strong deburring performance
- MobileNet-v2
  - For mobile on device enhancement
- MobileNet-DSC
  - Replaces all normal convolution in the full network
  - Extremely Light-weight
  - Efficient Deburring



Generative adversarial networks

- Two player minimax game Discriminator and Generator
  - Goal is to capture the real data distribution

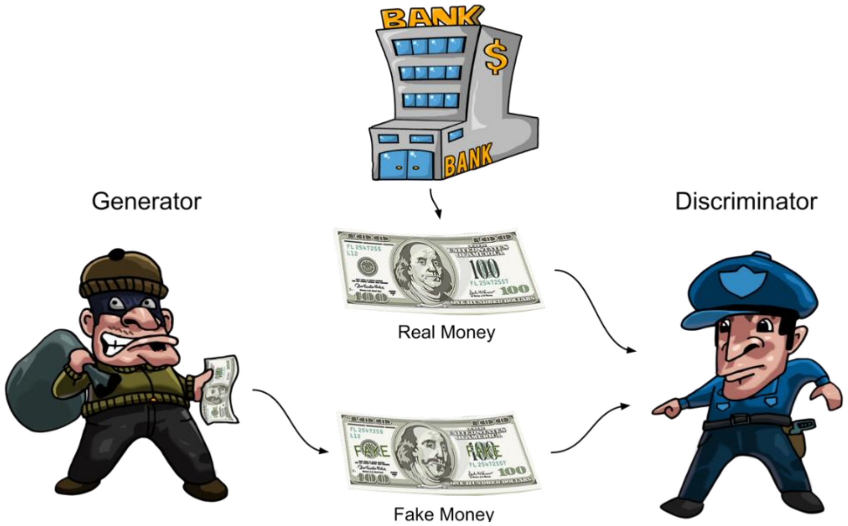
$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}\left[\log D(x)\right] + \mathbb{E}_{z\sim p_z(z)}\left[\log(1-D(G(z)))\right]$$

- Least Square GANs Discriminator
  - Loss function that provides smoother and non-saturation gradient

$$\min_D V(D) = \frac{1}{2}\mathbb{E}_{x\sim p_{data}(x)}\left[(D(x)-1)^2\right] + \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}\left[D(G(z))^2\right]$$

$$\min_G V(G) = \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}\left[(D(G(z))-1)^2\right]$$

- Relativistic GAN
  - Estimates probability that given real data is more realistic than random fake sampled data

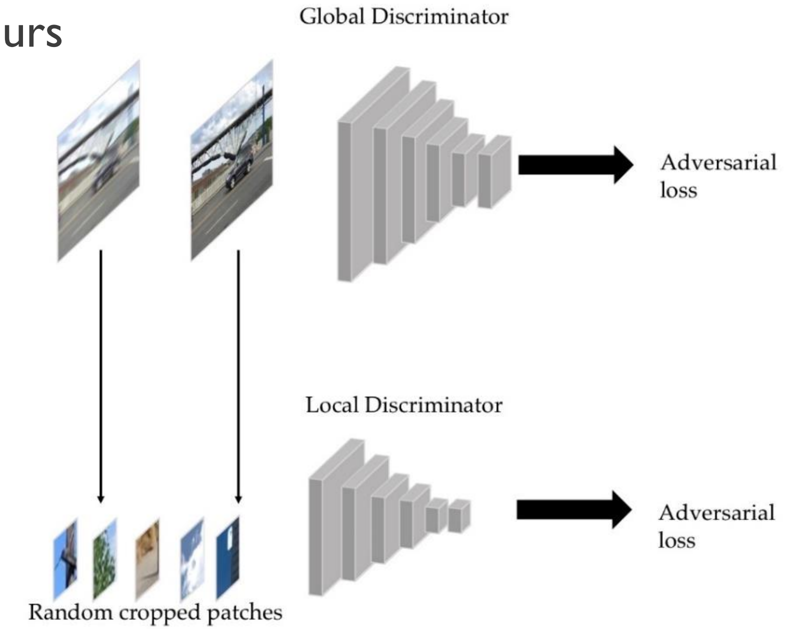


Double-Scale RaGAN-LS Discriminator

- Double Scale Discriminators to handle larger and more heterogeneous blurs
  - Local branch
    - Operates on patch levels
    - Produces sharper results
  - Global branch
    - Feed on full input image
    - Better performance for highly non-uniform blurs
- Losses

$$L_D^{RaLSGAN} = \mathbb{E}_{x \sim p_{data}(x)}\left[(D(x) - \mathbb{E}_{z \sim p_z(z)} D(G(z)) - 1)^2\right] + \mathbb{E}_{z \sim p_z(z)}\left[(D(G(z)) - \mathbb{E}_{x \sim p_{data}(x)} D(x) + 1)^2\right]$$

$$L_G = 0.5 * L_p + 0.006 * L_X + 0.01 * L_{adv}$$



Training Datasets

- GoPro:
  - GoPro Hero 4 camera
  - 240 fps video sequences
- DVD:
  - Multiple Devices
  - 71-real world videos sequences at 240 fps
- NFS:
  - iPhone 6 and iPad Pro
  - 75 videos with high frame rate
  - Youtube Videos

Experimental Evaluation

Quantitative Evaluation

- Evaluation metrics: PSNR and SSIM, inference efficiency

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

$$MSE = \frac{1}{m\,n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

$$\begin{aligned} PSNR &= 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \\ &= 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \\ &= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE) \end{aligned}$$

Quantitative Evaluation

- Second best in some tests, but fastest by far; several order of magnitudes

Table 1: Performance and efficiency comparison on the GoPro test dataset, All models were tested on the *linear* image subset.

	Sun *et al.* [43]	Xu *et al.* [51]	DeepDeblur [33]	SRN [45]	DeblurGAN [21]	Inception-ResNet-v2	MobileNet	MobileNet-DSC
PSNR	24.64	25.10	29.23	**30.10**	28.70	29.55	28.17	28.03
SSIM	0.842	0.890	0.916	0.932	0.927	**0.934**	0.925	0.922
Time	20 min	13.41s	4.33s	1.6s	0.85s	0.35s	0.06s	**0.04s**
FLOPS	N/A	N/A	1760.04G	1434.82G	678.29G	411.34G	43.75G	**14.83G**

Table 2: PSNR and SSIM comparison on the Kohler dataset.

Method	Sun [43]	DeepDeblur [33]	SRN [45]	DeblurGAN [21]	Inception-ResNet-v2	MobileNet	MobileNet-DSC
PSNR	25.22	26.48	26.75	26.10	26.72	26.36	26.35
SSIM	0.773	0.807	0.837	0.816	0.836	0.820	0.819

Table 3: Results on DVD dataset

	PSNR	SSIM	Inference Time	Resolution
WFA	28.35	N/A	N/A	N/A
DVD (single)	28.37	0.913	1.0s	960 x 540
DeblurGAN-v2 (MobileNet)	**28.54**	**0.929**	**0.06s**	1280 x 720

Visual Comparison

(a) Blurry (b) DeepDeblur [33]
(c) SRN [45] (d) DeblurGAN [21]
(e) DeblurGAN-v2 (Inception-ResNet-v2) (f) DeblurGAN-v2 (MobileNet)

Figure 7: Visual comparison on the Kohler dataset.

Ablation Study

- DeblurGAN (ResNet G, local-scale patch D, WGAN-GP + perceptual loss) modifications study:
  - Generator: adding FPN
  - Discriminator: adding global-scale
  - Loss: replacing WGANGP with RaGAN-LS and adding MSE term

Table 5: Ablation Study on the GoPro dataset, based on DeblurGAN-v2 (Inception-ResNet-v2).

	PSNR	SSIM
DeblurGAN (starting point)	28.70	0.927
+ FPN	29.26	0.931
+ FPN + Global D	29.29	0.932
+ FPN + Global D + RaGAN-LS	29.37	0.933
DeblurGAN-v2 (FPN + Global D + RaGAN-LS + MSE Loss)	**29.55**	**0.934**
Removing perceptual loss (replace 0.5 with 0 in L_G)	28.81	0.924

Final Project Proposal

DeblurGAN-v2(deo)

- Currently allows single image deblurring
- Make pipeline for video (blurring and) deblurring
- Evaluating deblurred video image quality to original

Original Video → Original Frames → Blurred Frames → Blurred Video → Blurred Frames → Deblurred Frames → Deblurred Video

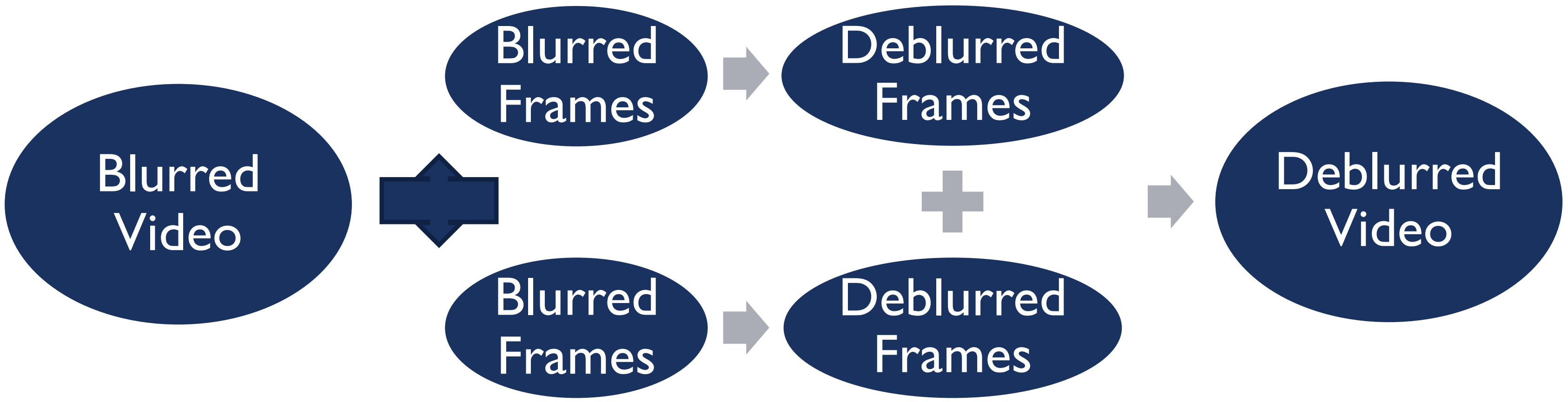
DeblurGAN-v2(deo)

- Alternative pipeline for naturally blurry video:
  - Low framerate
  - Lack of proper illumination
  - Camera movement

Original Blurred Video → Original Blurred Frames → Deblurred Frames → Deblurred Video

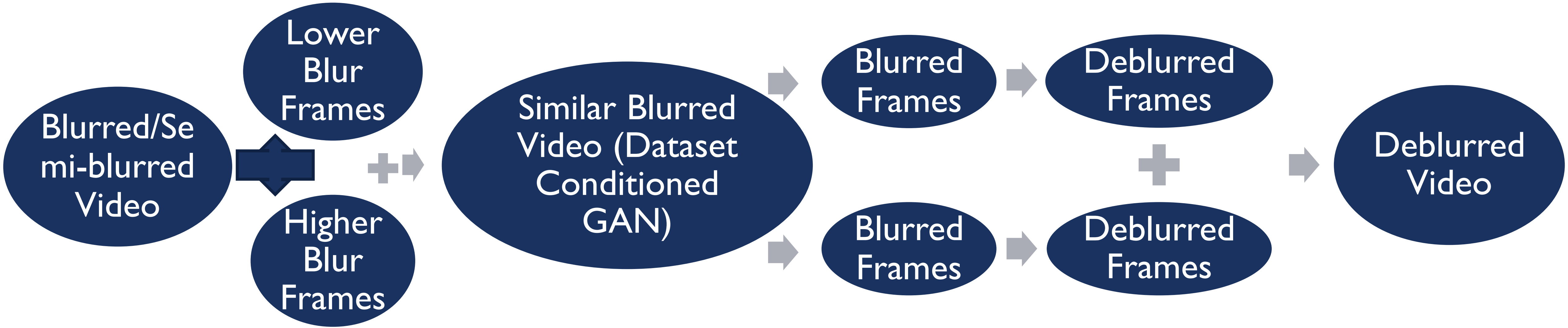
DeblurGAN-v2(deo)

- Other possible lines of improvement come from HPC (High-Performance Computing) side:
  - Parallel processing of frames for obtaining close to real-time deblurring



DeblurGAN-v2(deo)

- Even more ambitious (and advanced pipeline): fine-tuning network on the go with input video blur/deblurred image-pairs with high-learning rate/loss penalization for conditioning to new dataset



SupraFPS

- Super Resolution for FPS of videos: 30 FPS to 120 FPS (slow-motion)
- Naive/simple approach: linear/weighted interpolation of video frames

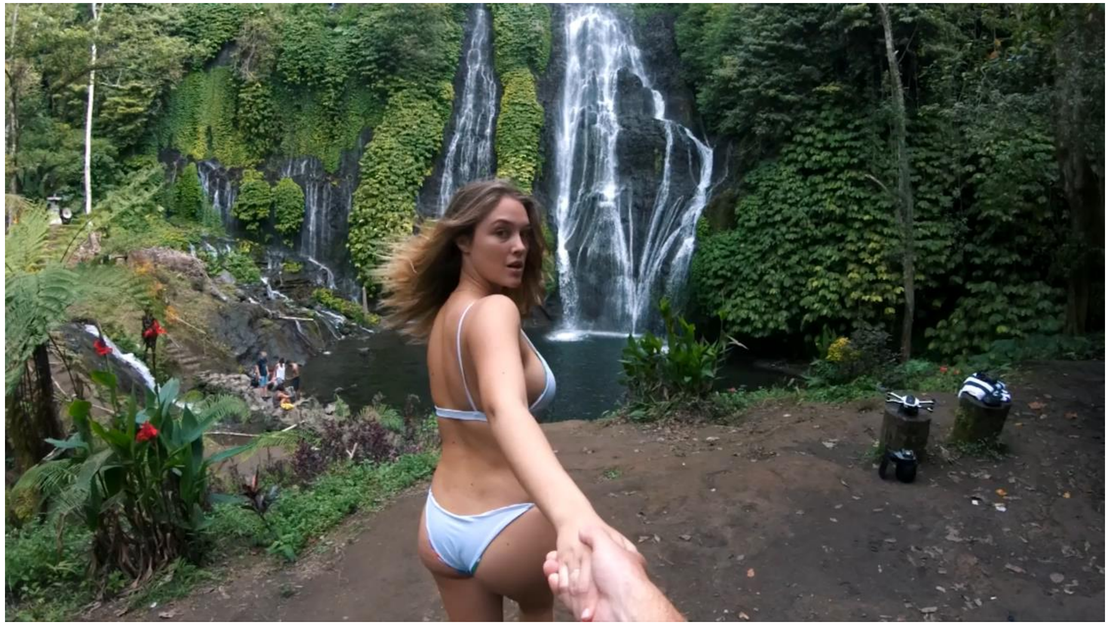
Original Video (FPS: X, time: t_1) → Original Frames → Interpolated /Blurry Frames → Deblurred Interpolated Frames → Slow-mo Video: (FPS: X, time: t_2)

*X: FPS of original video, t_1: duration in time (s) of original video, t_2 (t_2 > t_1): duration in time (s) of new video

fpsGAN

- More-advanced approaches: use kernel for motion estimation to get interpolated frames, CNNs

- GAN-boosted approach: fpsGAN
  - Real high FPS video, down-sample to low FPS pairs and use network to generate/learn difference between the real high FPS videos and the ones generated from our previous approaches
  - Use the results of fpsGAN (which are probably not that accurate) through deblurGAN-v2(deo) and compare to results through only fpsGAN



SupraFPS Advanced Pipeline

Real High-FPS and Synthetic High-FPS Videos

fpsGAN Training

fpsGAN Input: Low-FPS Video

fpsGAN Output: High-FPS Video

DeblurGAN-v2(deo) Deblurring

Deblurred-High-FPS (Final Output) Videos

References

- O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, "DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better," *arXiv:1908.03826 [cs]*, Aug. 2019.
- O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks," *arXiv:1711.07064 [cs]*, Apr. 2018.
- "PyTorch." [Online]. Available: https://www.pytorch.org. [Accessed: 07-Dec-2019].
- "NVIDIA Tesla P100: The Most Advanced Data Center Accelerator," *NVIDIA*. [Online]. Available: https://www.nvidia.com/en-us/data-center/tesla-p100/. [Accessed: 07-Dec-2019].
- "GoPro HERO6 Black: The Best Slow Motion Ever in 4K! - YouTube." [Online]. Available: https://www.youtube.com/watch?v=khPwinkQ52k. [Accessed: 07-Dec-2019].
- *TAMU-VITA/DeblurGANv2*. TAMU-VITA, 2019.