Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

V-SENSE

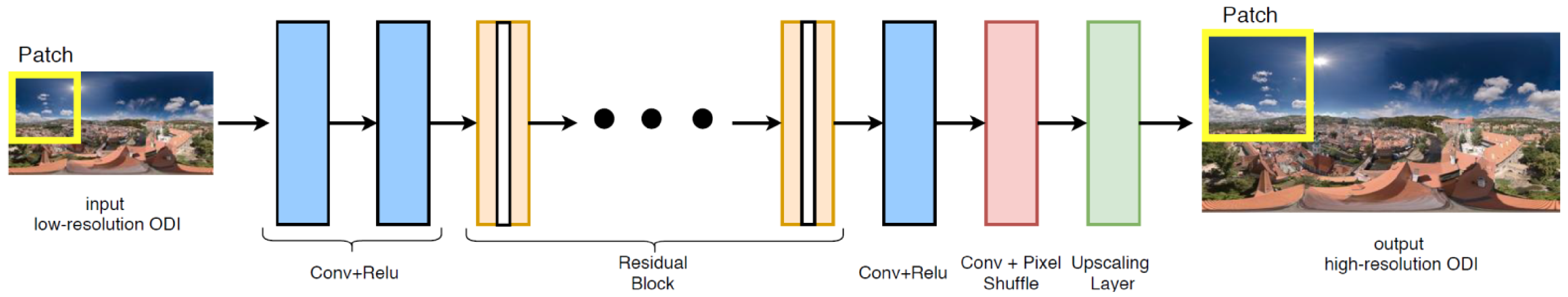
Super-resolution of Omnidirectional Images Using Adversarial Learning

Cagri Ozcinar*, Aakanksha Rana*, Aljosa Smolic

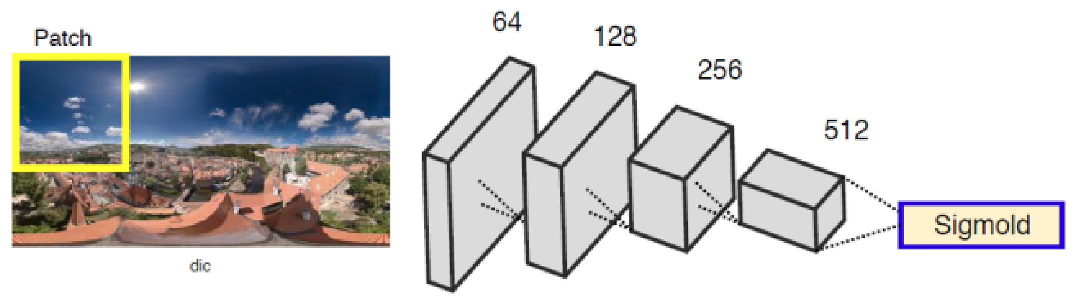
Problem and Objective

- Designing immersive virtual reality systems with ODIs is challenging as they require high resolution content.
- Proposing improved generative adversarial network (GAN) based model for super-resolution of ODIs.
- Handling artefacts obtained in the spherical observational space.

Proposed Model



Architecture of Generator Network with corresponding block labels [14].



Architecture of the used Discriminator Network, PatchGAN [17].

Loss functions

The objective function:

$$\mathbf{G}^* = \arg\min_{G}\max_{D}[\mathcal{L}_{adv}(\mathbf{G},\mathbf{D})] + \beta\mathcal{L}_{feat}(\mathbf{G}) + \gamma\mathcal{L}_{360-SS}(\mathbf{G})$$

360-SS Loss:

$$\mathcal{L}_{360-SS} = \frac{1}{K}\sum_{i=1}^{K} d^i_{360-SS},$$

$$d_{360-SS} = \frac{\sum_{x=1}^{W/r}\sum_{y=1}^{H/r}\left(SSIM(I^{x,y}_{360-sr}, \widehat{I}^{x,y}_{360-sr})q_r^{x,y}\right)}{\sum_{x=1}^{W/r}\sum_{y=1}^{H/r} q_r^{x,y}},$$

$$q_r^{x,y} = cos\frac{(y+0.5-(H/2r)\pi}{(H/r)}.$$

Feature Loss:

$$\mathcal{L}_{feat} = \frac{1}{K}\sum_{i=1}^{K}(\mathcal{F}^i(I_{360-sr}) - \mathcal{F}^i(\widehat{I}_{360-sr})),$$

Adversarial Loss:

$$\mathcal{L}_{adv} = \sum_{i=1}^{K}(-log\mathbf{D}(\mathbf{G}(I^i_{360-lr}))).$$

Training

- 3500 ODIs.
- Random crops of size 512×512.
- Data augmentation techniques such as rotation and flipping.
- 2hr training time with 12 GB NVIDIA Titan-X GPU on an Intel Xeon E7 core i7 machine.
- Inference time is 0.030 milliseconds for each ODI.

Quantitative Results

By a factor of 2 on 500 ODIs

Method	r = 2			
	SSIM	PSNR	WS-SSIM	WS-PSNR
NN	0.92 ± 0.06	29.38 ± 0.04	0.86± .03	34.34 ± .05
Bicubic	0.93 ± 0.05	30.64 ± 0.06	0.88 ± .04	35.54 ± .07
SRGAN [14]	0.94 ± 0.05	32.56 ± 0.06	0.90 ± .06	36.35 ± .06
Ours	0.95 ± 0.04	33.20 ± 0.04	0.92 ± .04	37.68 ± .05
Ours+ 360-SS loss	0.95 ± 0.03	33.56 ± 0.04	0.93 ± .06	37.96 ± .03

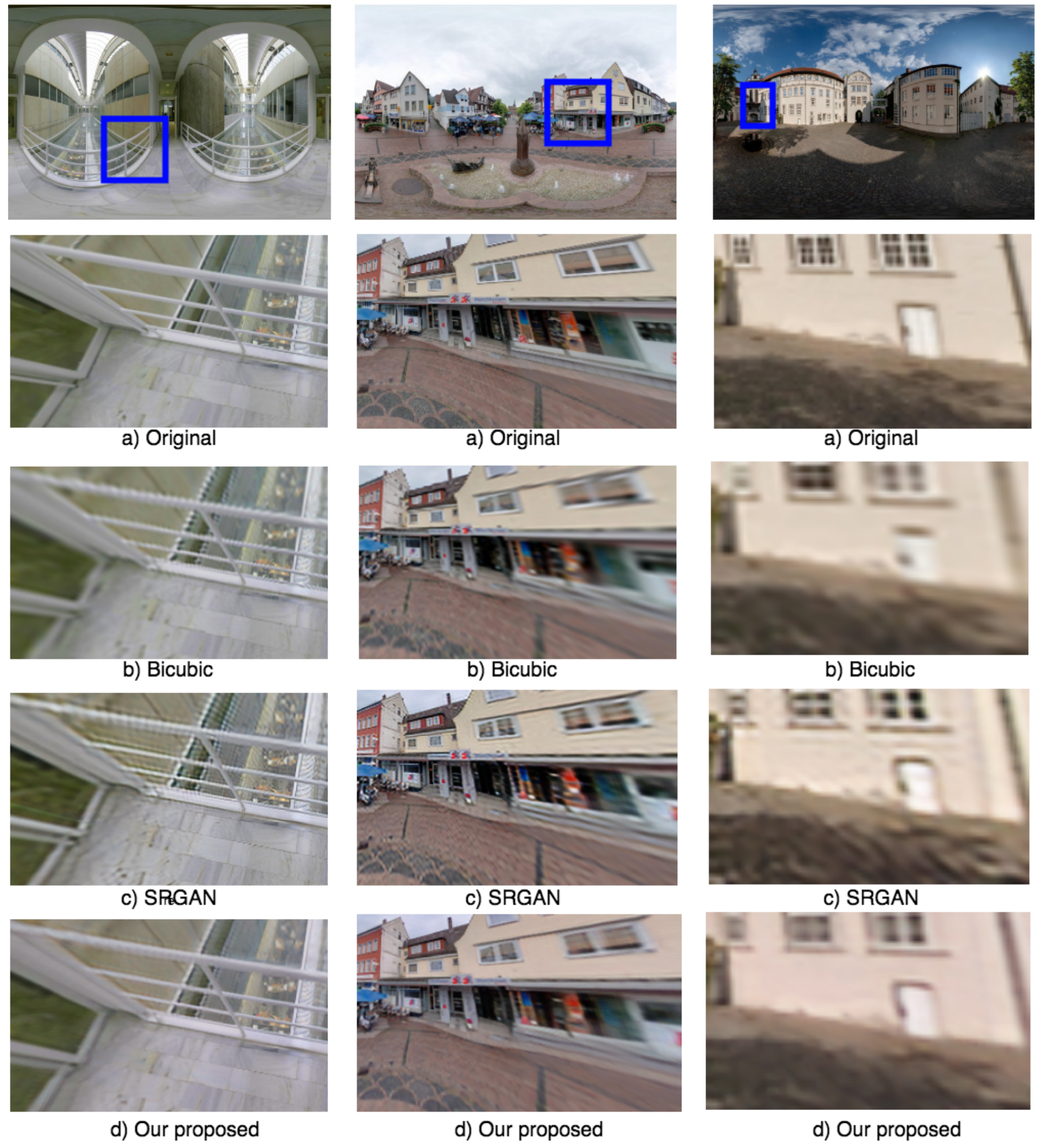
By a factor of 4 on 500

Method	r = 4			
	SSIM	PSNR	WS-SSIM	WS-PSNR
NN	0.83 ± 0.07	25.77 ± 0.05	0.71± .03	32.44 ± .05
Bicubic	0.85 ± 0.07	26.71 ± 0.03	0.74 ± .04	32.76 ± .04
SRGAN [14]	0.86 ± 0.02	27.11 ± 0.09	0.75 ± .06	34.76 ± .03
Ours	0.87 ± 0.05	27.19 ± 0.08	0.76 ± .05	35.89 ± .05
Ours+ 360-SS loss	0.87 ± 0.04	27.70 ± 0.03	0.77 ± .08	36.98 ± .06

By a factor of 8 on 500

Method	r = 8			
	SSIM	PSNR	WS-SSIM	WS-PSNR
NN	0.83 ± 0.07	23.47 ± 0.05	0.64 ± .06	31.12 ± .04
Bicubic	0.85 ± 0.07	24.26 ± 0.03	0.66 ± .07	31.83 ± .06
SRGAN [14]	0.86 ± 0.02	25.10 ± 0.09	0.70 ± .06	33.00 ± .07
Ours	0.87 ± 0.05	26.24 ± 0.08	0.73 ± .07	34.68 ± .04
Ours + 360-SS loss	0.87 ± 0.04	26.56 ± 0.03	0.75 ± .06	35.54 ± .06

Qualitative Results



References

[14] C. Ledig, *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network,” CVPR, 2017.
[17] P. Isola, *et al.*, “Image-to image translation with conditional adversarial networks,” CVPR, 2017.

contact: ozcinarc@scss.tcd.ie and ranaaa@ scss.tcd.ie