Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation

Jo, Younghyun, et al. CVPR 2021.

2021.11.25 임은우

Image super-resolution (SR)

The goal is to generate a high-resolution (HR) from its corresponding low resolution (LR) counterpart.

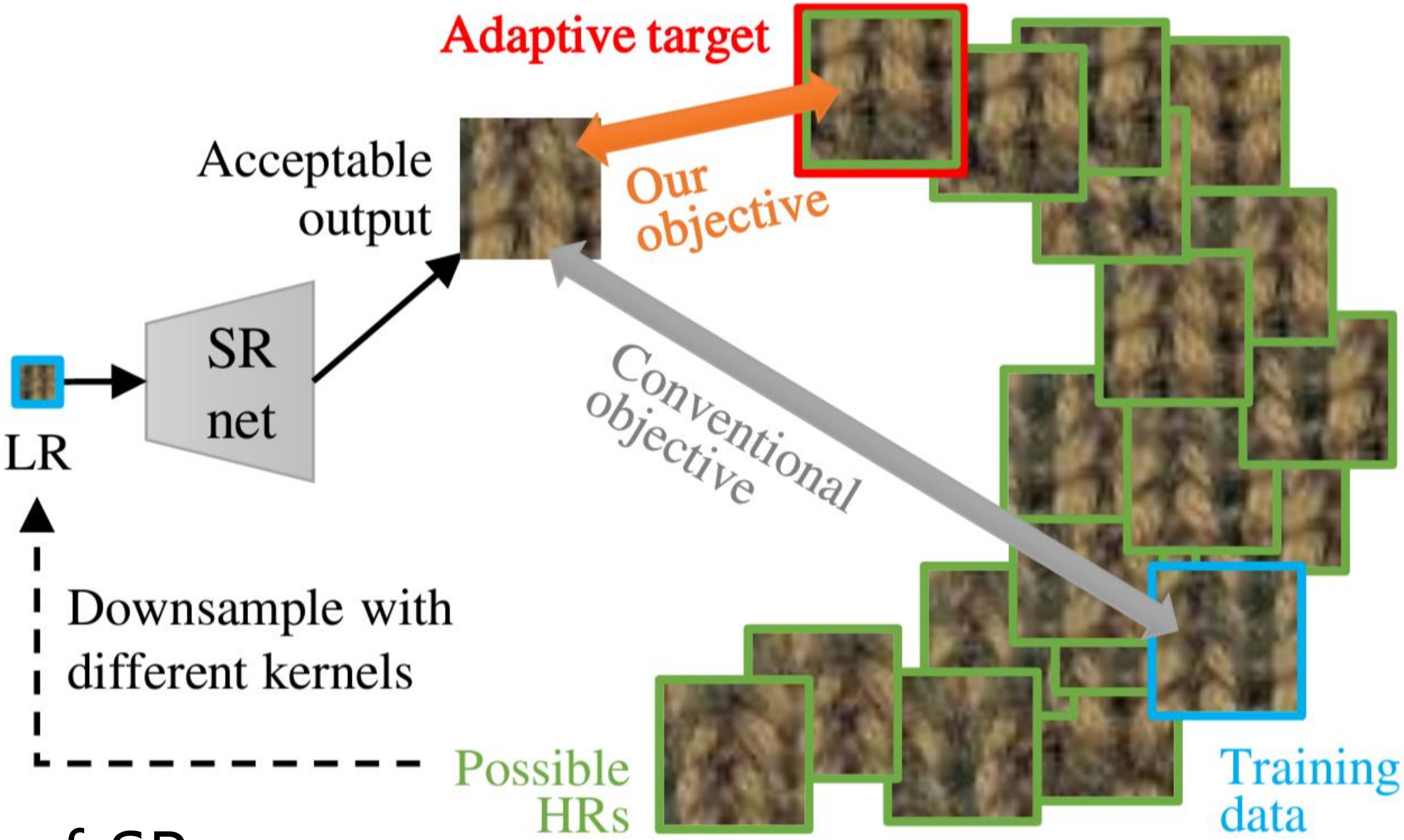
For learning based approach, a LR image I_{LR} is simulated from a ground truth (GT) HR image I_{GT}.

SR network f is trained to generate HR output $f(x_i)$ close to the given target y_i
by minimizing a loss function $\sum_i l(y_i, f(x_i))$

Notation

From dataset $D = \{(x_i, y_i)\}$, $I_{LR} = (I_{GT} * k) \downarrow_s$
where f is SR network, k is blur kernel,
$\downarrow_s$ is downsampling operator with scale factor s.

Introduction



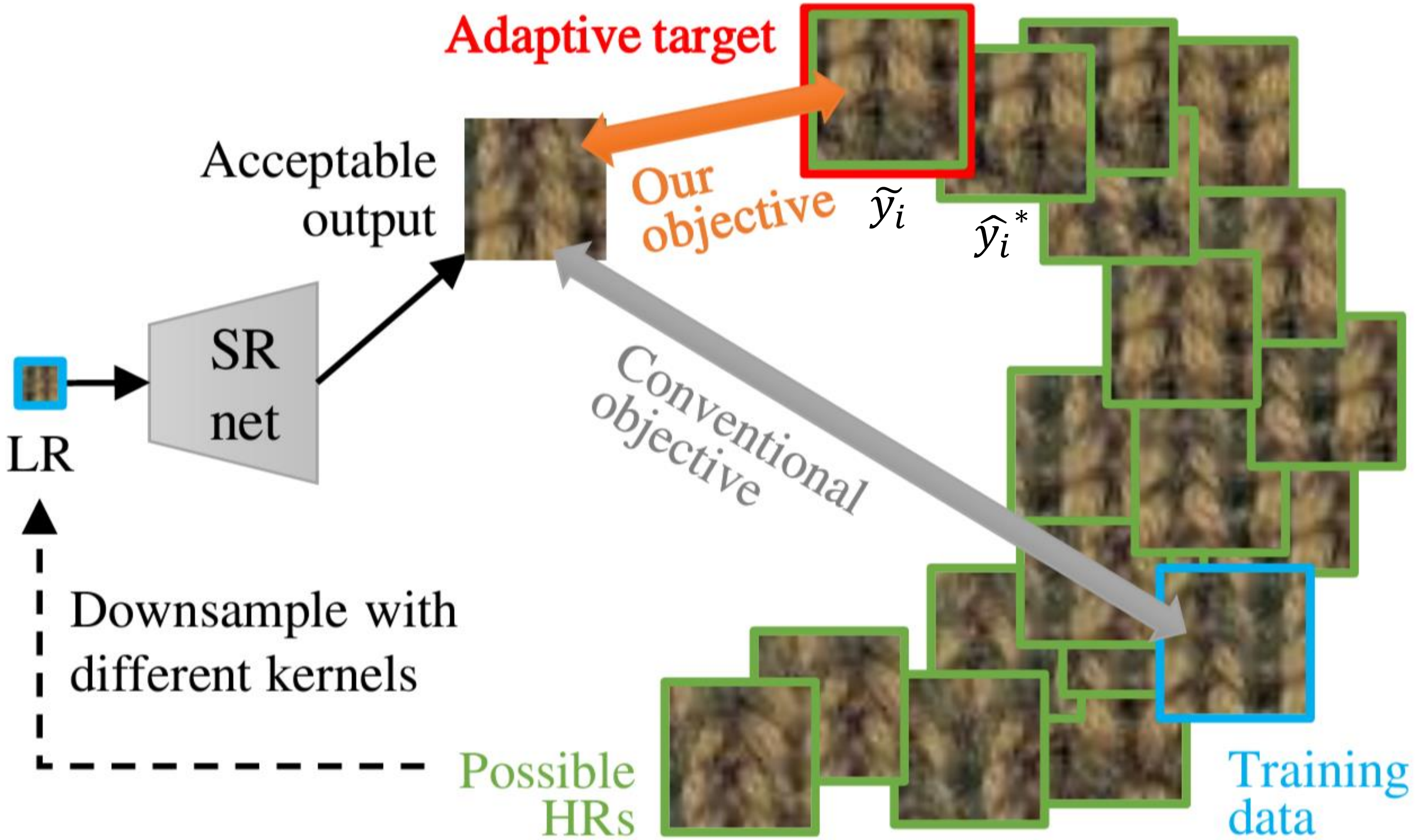
Ill-posedness of SR

There are many HR patches can result in the same LR patch.
i.e., the given target y_i is not the only solution for the input x_i (one to many mapping)

Even if SR network generated that is close to one of possible HR targets (green), the conventional loss will increase.

Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Method



$$\hat{y}_i^* = \underset{\hat{y}_i}{\operatorname{argmin}}\, l(\hat{y}_i, f(x_i)), s.t.\ x_i = (\hat{y}_i * k_i^m) \downarrow_s$$

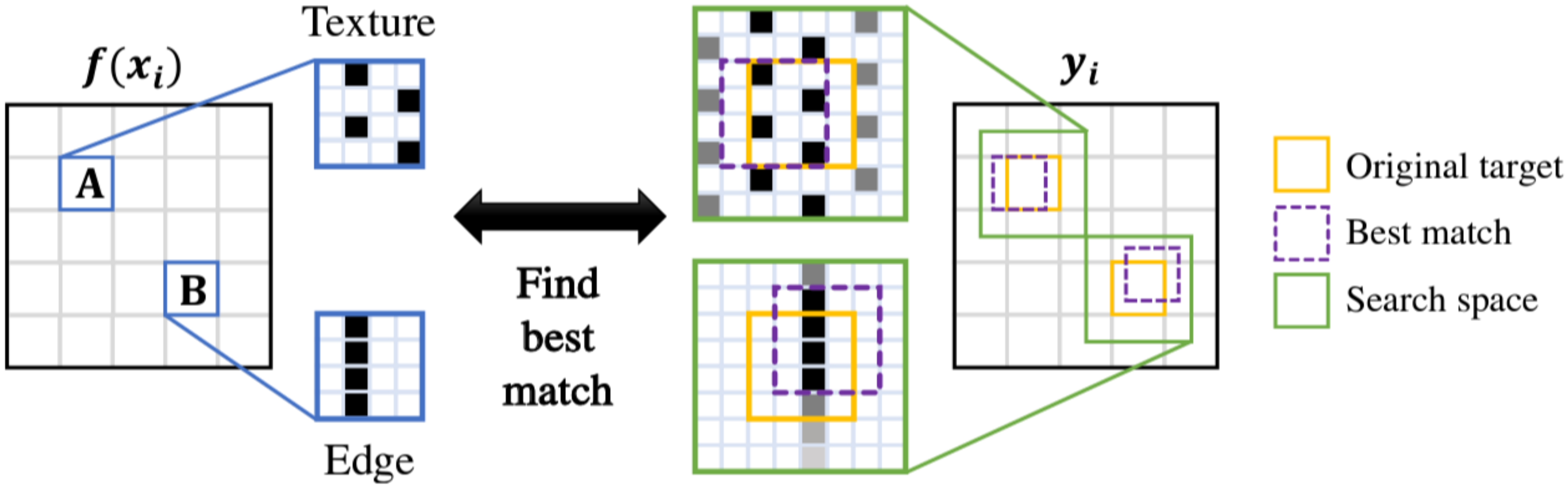
where $\hat{y}_i$ is one of the possible HR images and k_i^m is blur kernel.

Use alternative target relaxes loss by allowing various HR predictions given an LR input.

→ encourage sharp output generation

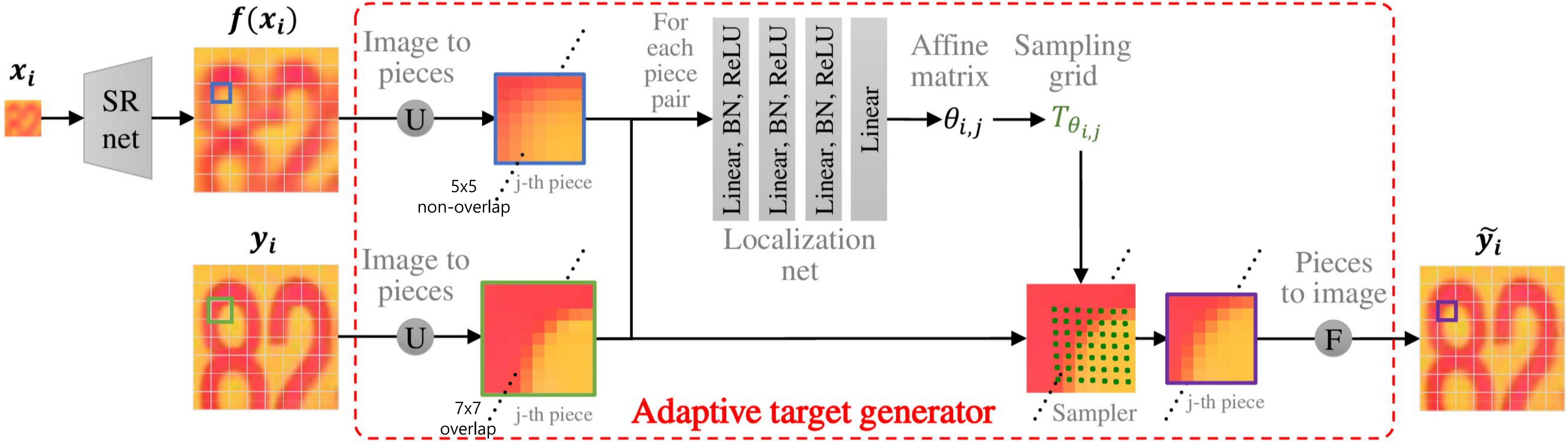
Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Adaptive Target



Assume that both targets (original and best match) are mapped to the same LR patch when downsampled.

Adaptive Target Generator (ATG)



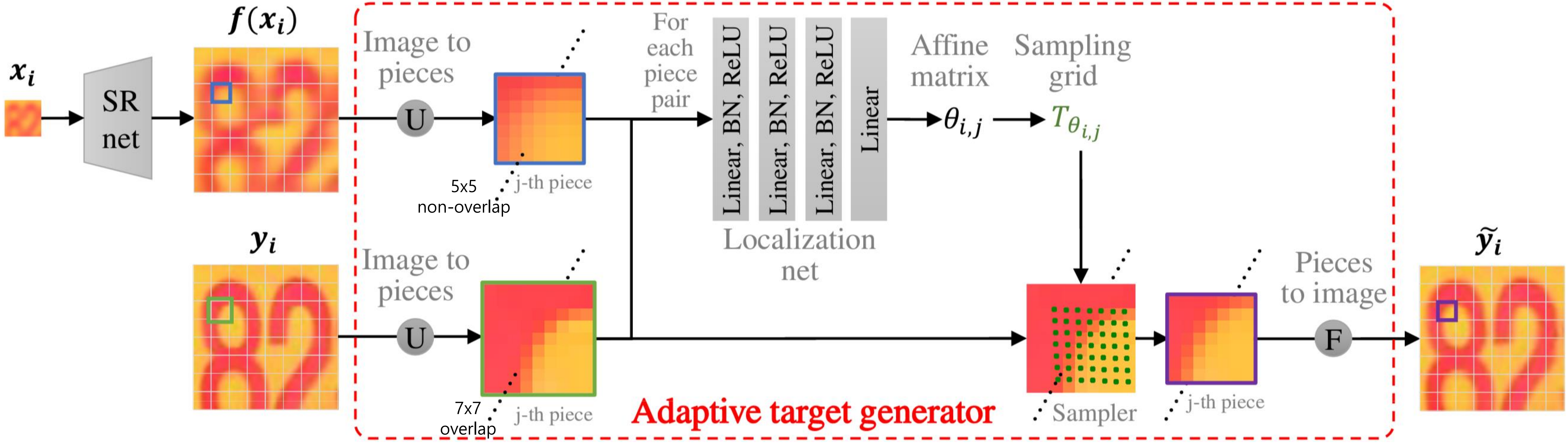
$$\tilde{y}_i = ATG(y_i, f(x_i))$$

ATG input: SR network output patch, original patch

ATG output: affine transformation matrix $\theta_{i,j}$

Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Pretraining ATG



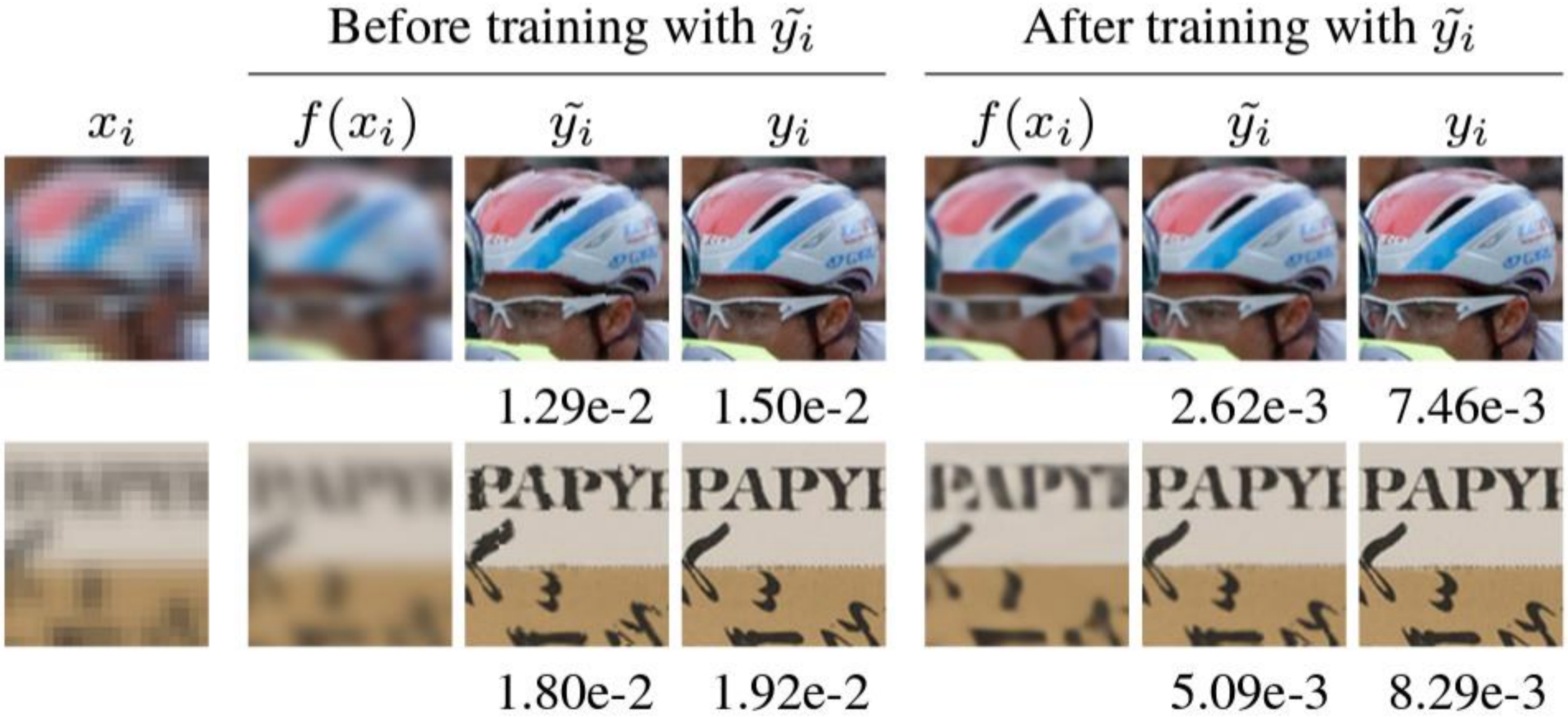
Simulate synthetic affine matrix θ^{Syn}
combining translation and rotation in random order.
(range [-1,1] pixel, [-10,10] degree)

$$\sum_i \left[\sum_j l\left(\theta_{i,j}, \theta_{i,j}^{Syn}\right) + \lambda l(\tilde{y}_i, \tilde{y}_i^{Syn}) \right]$$

Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Train SR Net with ATG

MSE



SR Net loss: $\sum_i l(y_i, \tilde{f}(x_i))$

Adaptive targets are not significantly different from original targets, and penalize the SR network less

Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Results

Method	DIV2KRK		
	PSNR	SSIM	LPIPS
Bicubic	25.37	0.6812	0.5635
KG [2] + ZSSR [34]	26.80	0.7319	0.4389
IKC [8]	27.69	0.7658	0.3856
Ours	**28.42**	**0.7854**	**0.3394**
SRMD [44]	27.63	0.7623	0.3816
MZSR [36]	24.99	0.6919	0.4096
USRNet [43]	29.80	0.8148	0.3088
Ours (ZSSR)	27.47	0.7561	0.3772
Ours (IKC)	28.33	0.7799	0.3558

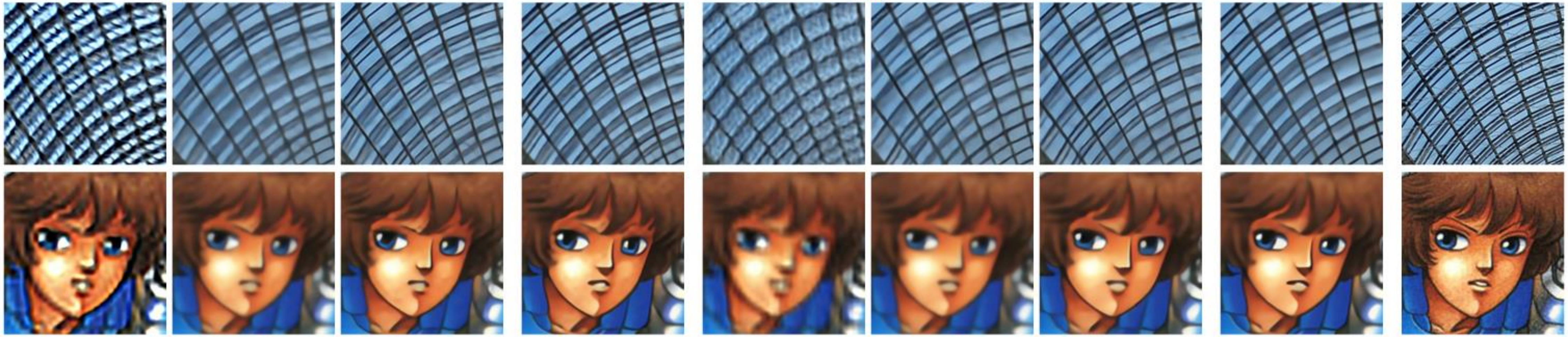
Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Results

Method	Set5			Set14			BSDS100			Urban100			Manga109		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
Bicubic	25.85	0.7289	0.4460	24.20	0.6222	0.5468	24.59	0.5940	0.6445	21.68	0.5745	0.5904	22.79	0.7097	0.4427
KG [2] + ZSSR [34]	-	-	-	-	-	-	-	-	-	22.06	0.6223	0.3788	24.75	0.7786	0.2454
IKC [8]	**31.62**	0.8808	0.2016	**28.18**	0.7608	0.3207	27.33	0.7161	0.4174	25.31	0.7487	0.2895	28.82	0.8756	0.1664
Ours	31.58	**0.8814**	**0.1932**	28.14	**0.7626**	**0.3122**	**27.43**	**0.7216**	**0.4030**	**25.72**	**0.7683**	**0.2518**	**29.97**	**0.8955**	**0.1286**
SRMD [44]	31.47	0.8799	0.1898	28.11	0.7628	0.3076	27.32	0.7200	0.4031	25.34	0.7548	0.2682	29.80	0.8925	0.1288
MZSR [36]	25.01	0.7192	0.3851	23.60	0.6140	0.4839	24.13	0.5866	0.5787	21.25	0.5674	0.5071	22.08	0.7031	0.3565
USRNet [43]	32.37	0.8932	0.1786	28.68	0.7799	0.2904	27.66	0.7349	0.3812	26.29	0.7899	0.2287	31.02	0.9119	0.1094

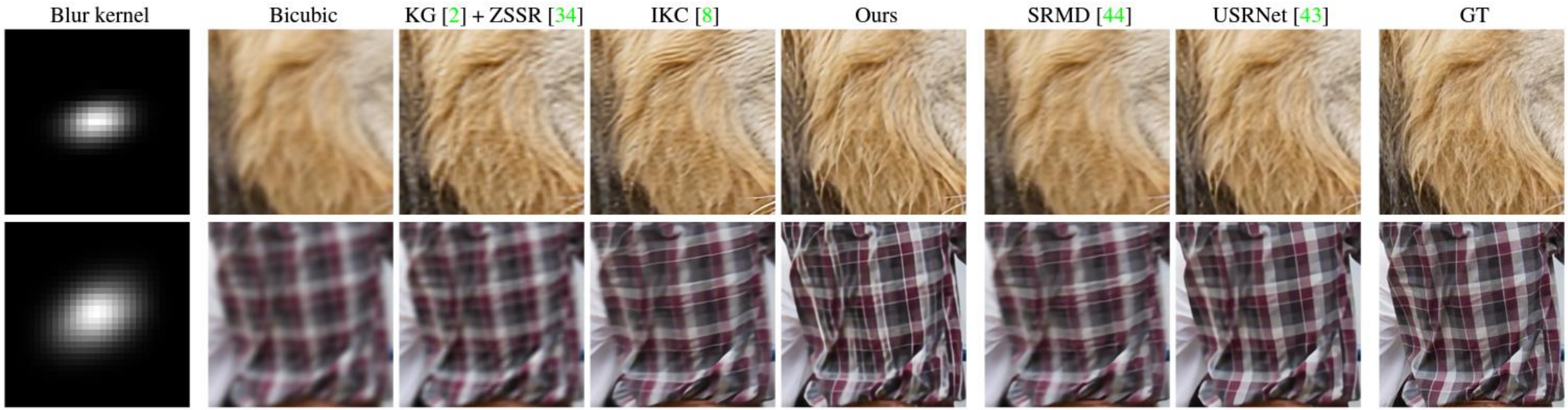
$\sigma = 1.8$ | $\sigma = 3.2$

KG [2] + ZSSR [34] | IKC [8] | Ours | USRNet [43] | KG [2] + ZSSR [34] | IKC [8] | Ours | USRNet [43] | GT



Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Results



Jo et al. "Tackling the Ill-Posedness of Super-Resolution through Adaptive Target Generation." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021

Thank you ☺