



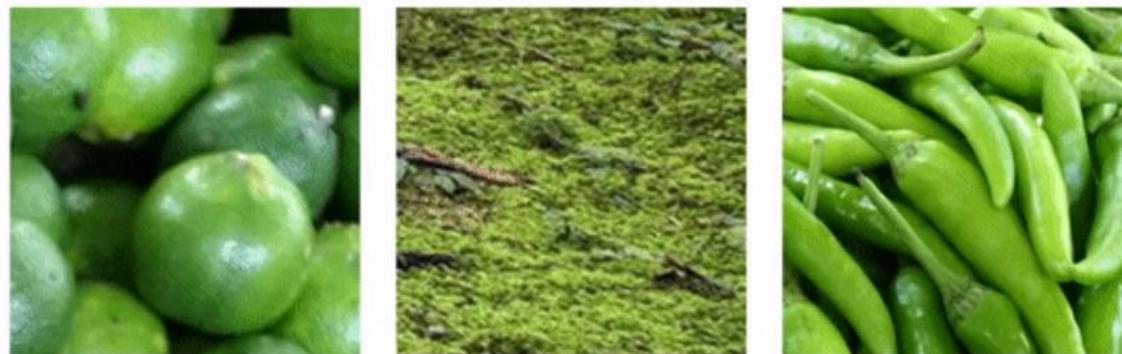
# **SRFlow-DA: Super-Resolution Using Normalizing Flow with Deep Convolutional Block**

Younghyun Jo Sejong Yang Seon Joo Kim

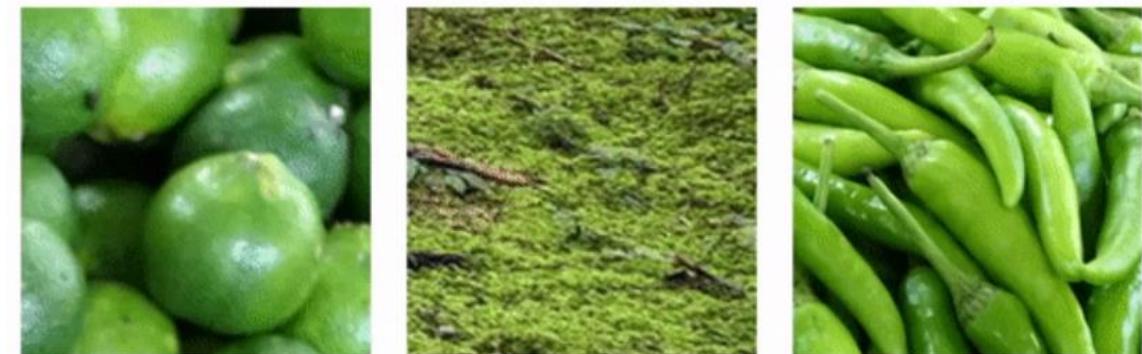
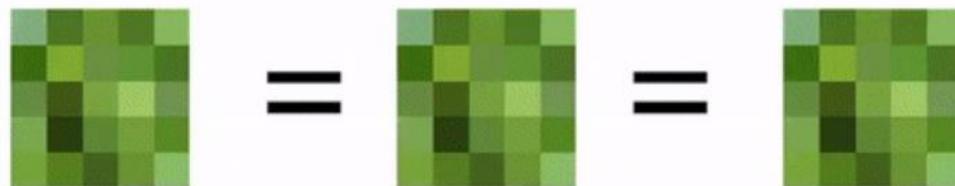
Yonsei University

NTIRE 2021  
Learning the Super-Resolution Space Challenge

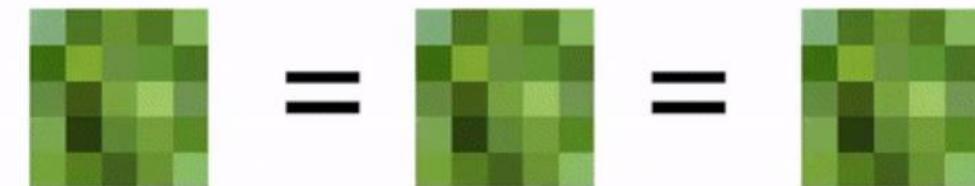
# Learning SR Space: Concept



↓ Downsampling ↓



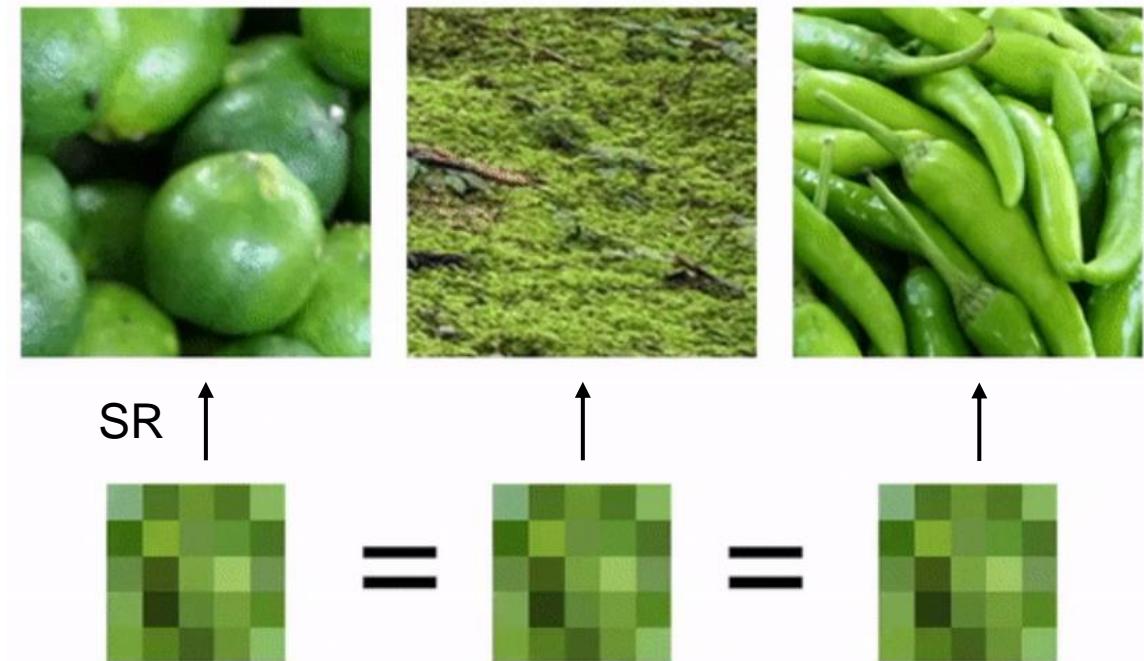
SR ↑



- Many high-resolution images can be downsampled to the same low-resolution image.
- Ill-posed.

- Most of current SR methods only map one LR to one HR.
- Deterministic mapping.

# Learning SR Space: Concept



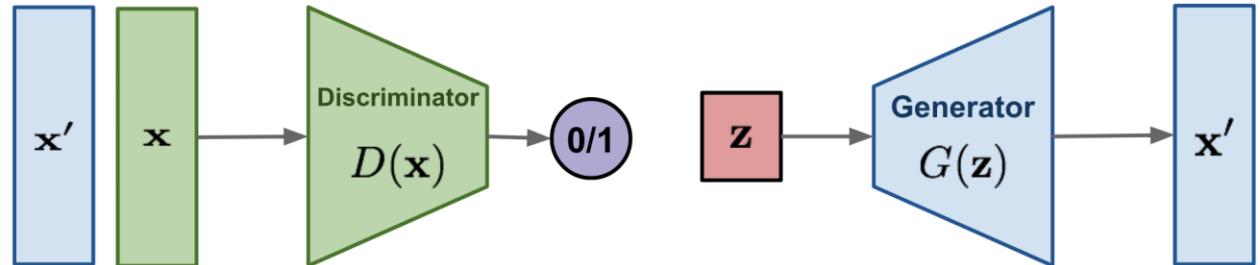
- Many high-resolution images can be downsampled to the same low-resolution image.
- Ill-posed.

- One can frame the SR problem as learning a stochastic mapping.
- **Challenge: How to make it?**

# VAE vs GAN vs Flow

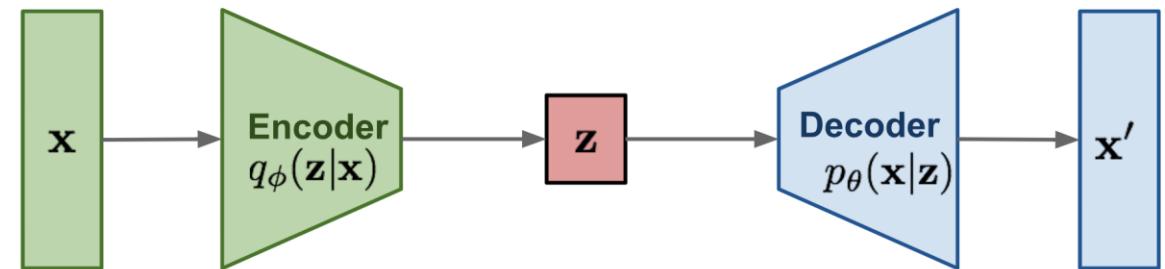
Implicit density.

**GAN:** minimax the classification error loss.



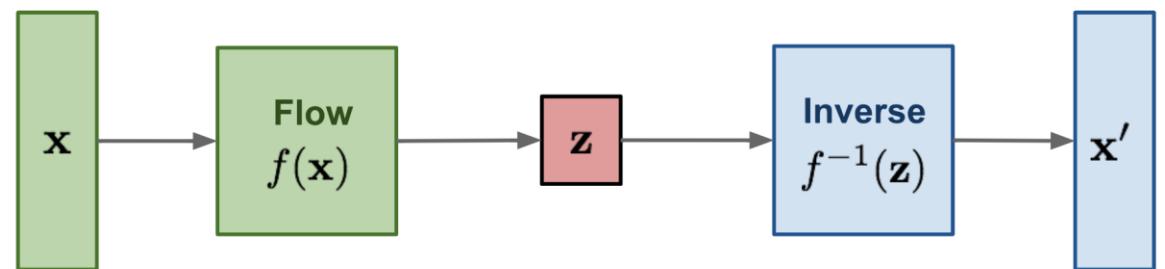
Explicit approximation.

**VAE:** maximize ELBO.



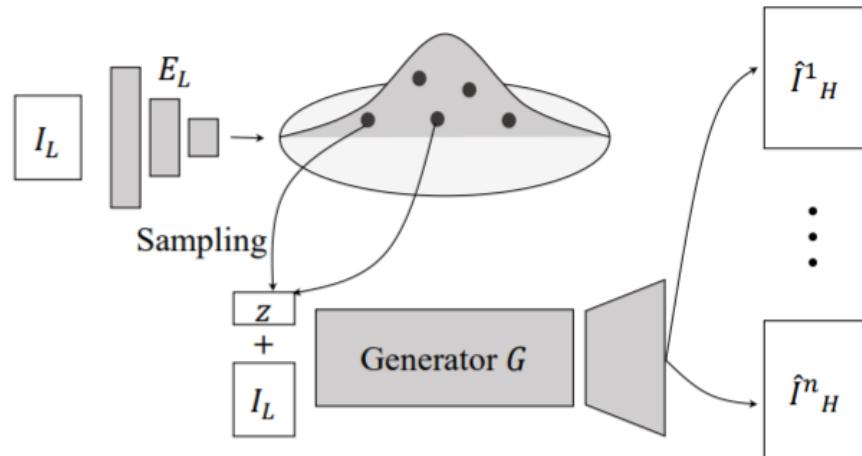
Explicit density.

**Flow-based generative models:** minimize the negative log-likelihood



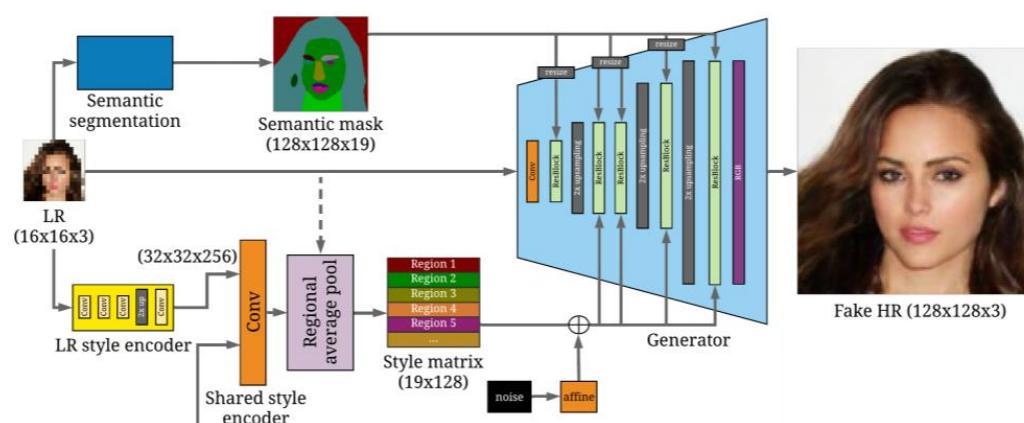
<https://lilianweng.github.io/lil-log/2018/10/13/flow-based-deep-generative-models.html>

# VAE vs GAN vs Flow



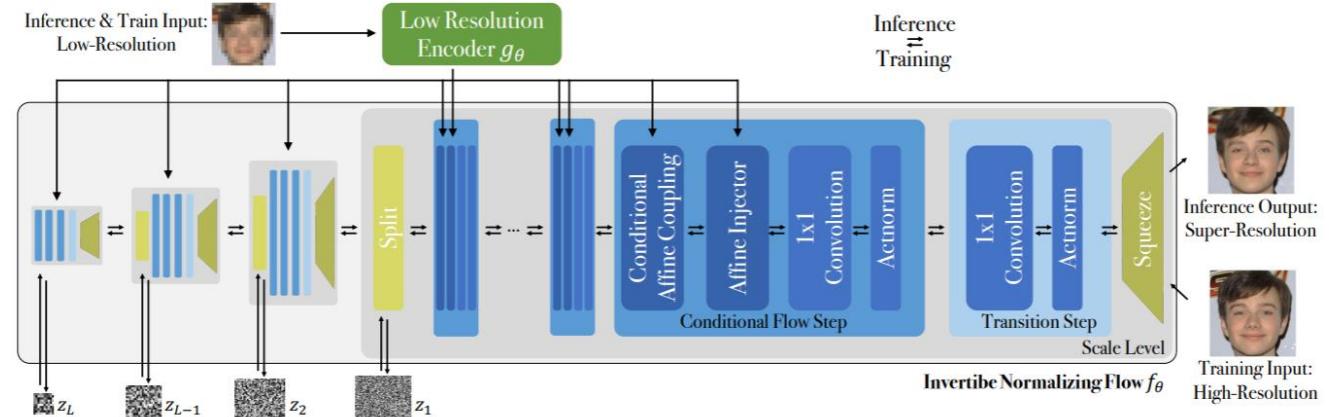
Hyun, Sangeek, and Jae-Pil Heo. "VarSR: Variational Super-Resolution Network for Very Low Resolution Images." ECCV, 2020.

- “VarSR”, VAE like, stochastic output.
- For faces, numbers.



DeepSEE: Deep Disentangled Semantic Explorative Extreme Super-Resolution. Bühler et al., ACCV 2020.

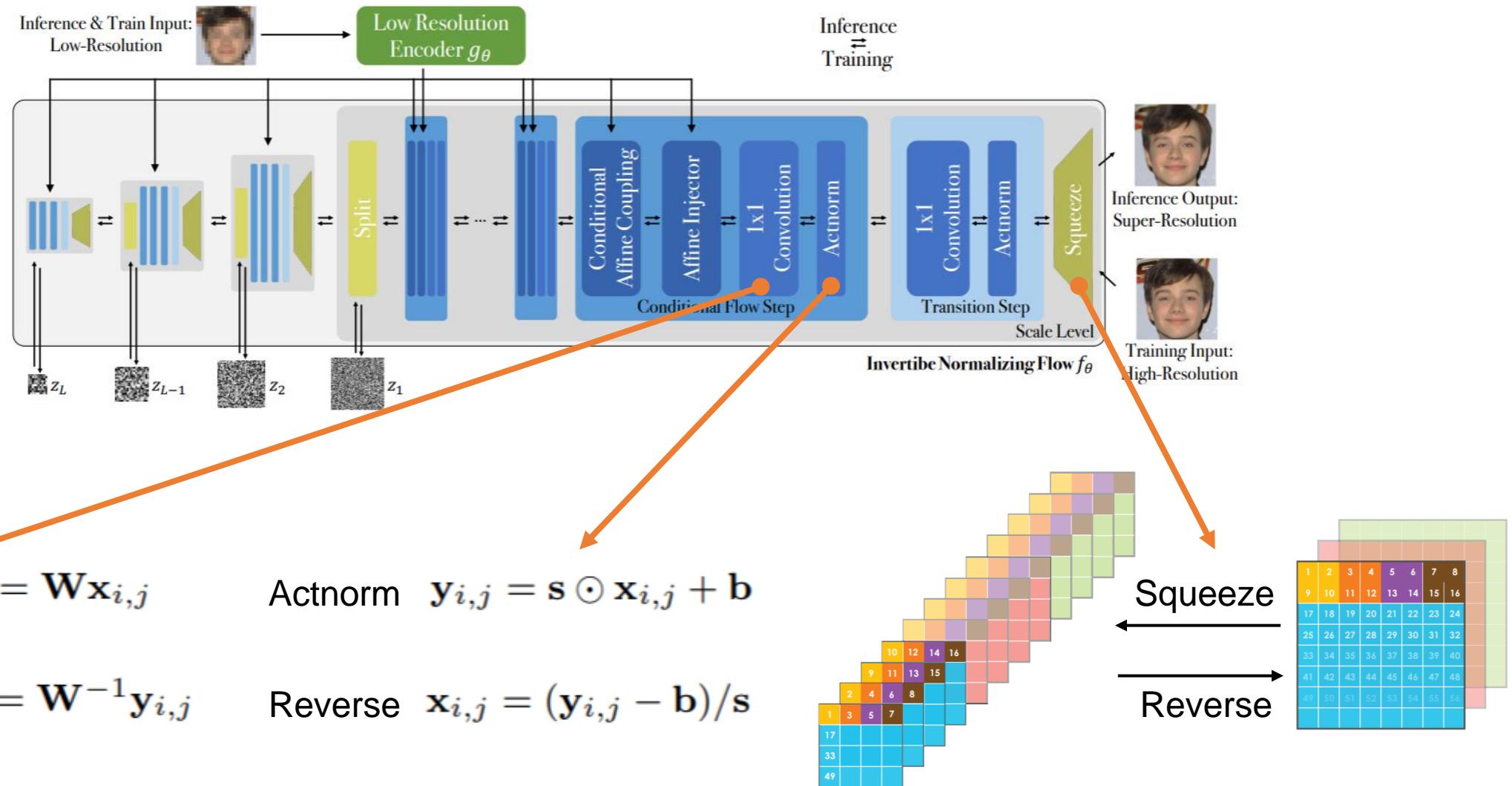
- “DeepSEE”, GAN, stochastic output.
- For faces.



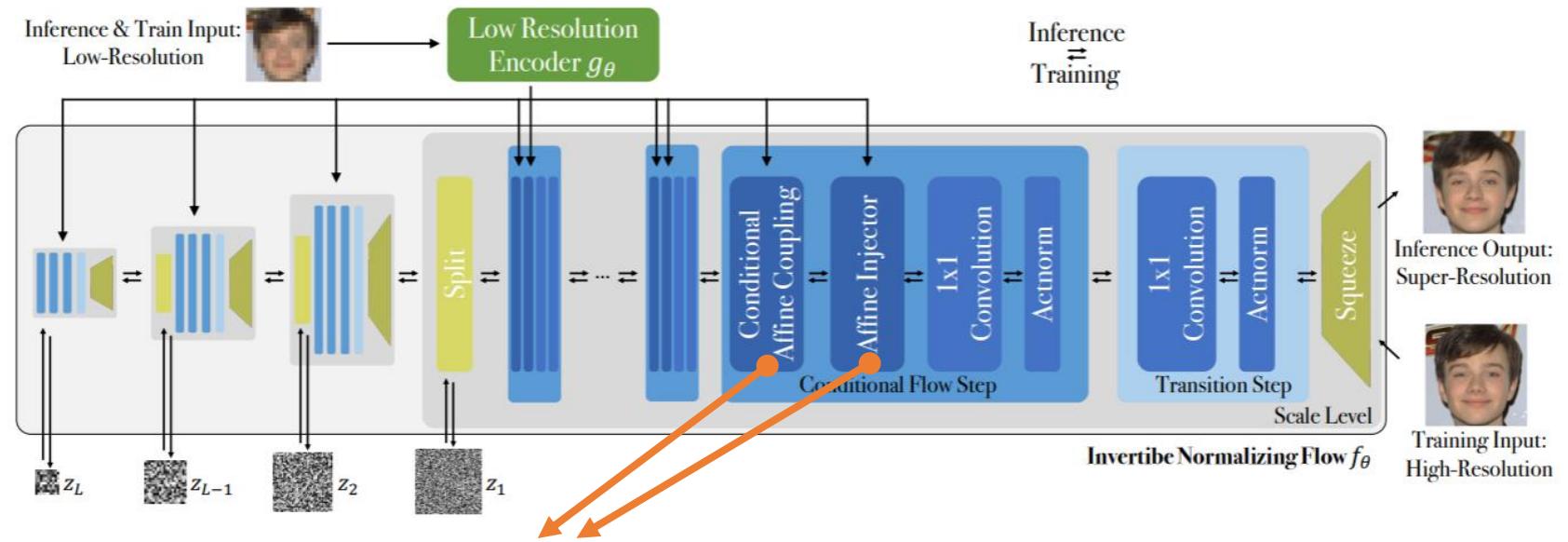
Lugmayr, Andreas, et al. "Srfow: Learning the super-resolution space with normalizing flow." European Conference on Computer Vision. Springer, Cham, 2020.

- “SRFlow”, normalizing flow, stochastic output.
- For general images.

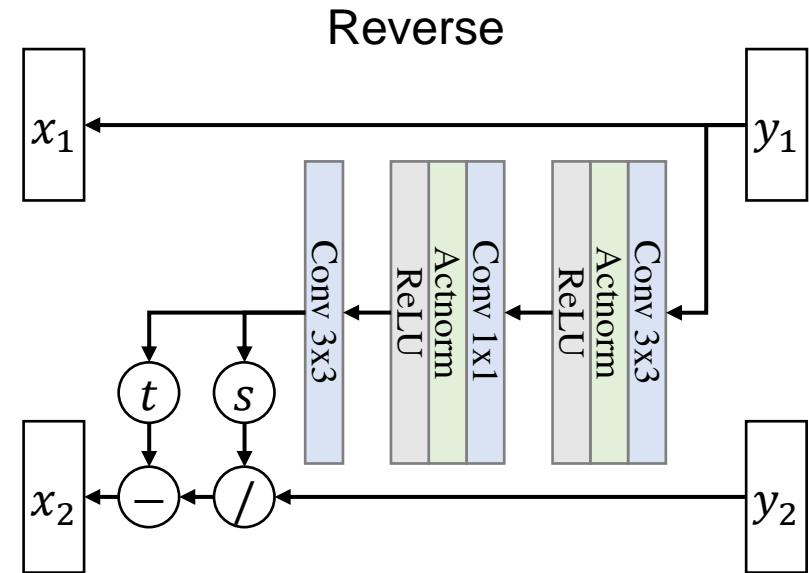
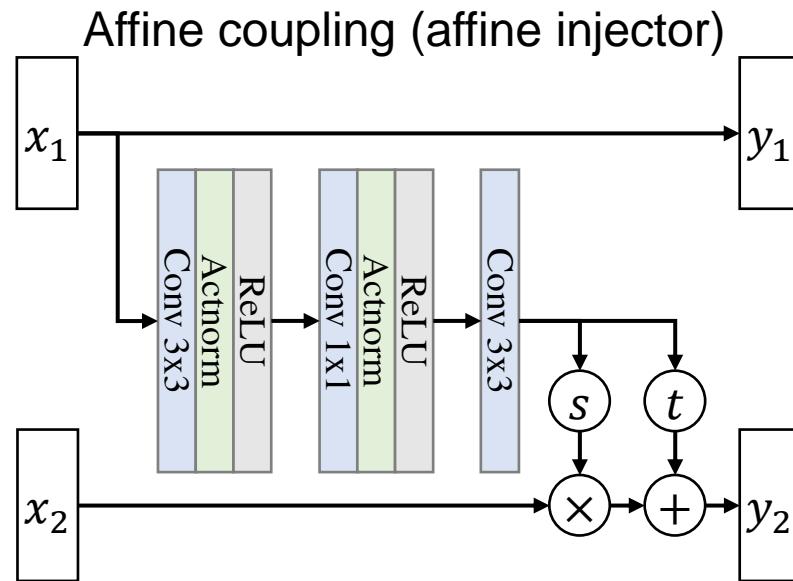
# Our Baseline: SRFlow



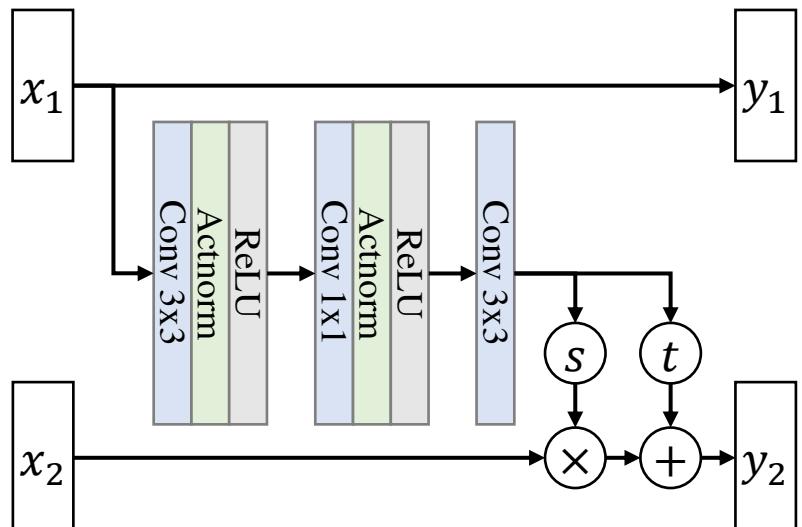
# Our Baseline: SRFlow



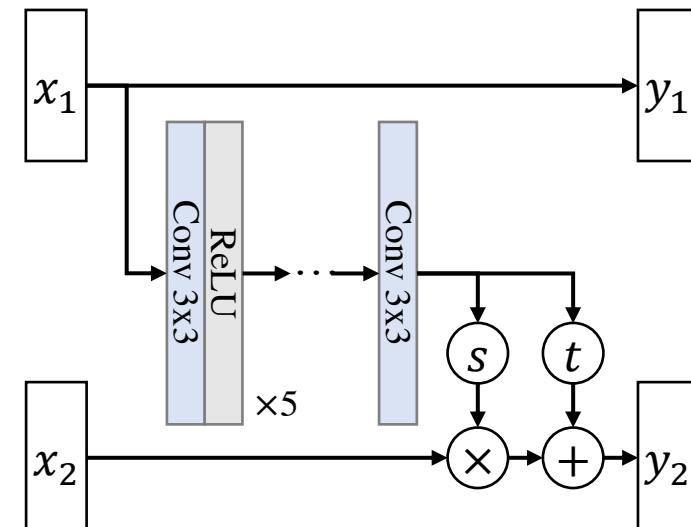
- For x8 SR.
- 16 flow steps for each level ( $K=16$ ).
- Total 4 scale level ( $L=4$ ).



## Our Approach: SRFlow-DA (Deep convolutional block in the Affine couplings)



Original affine coupling

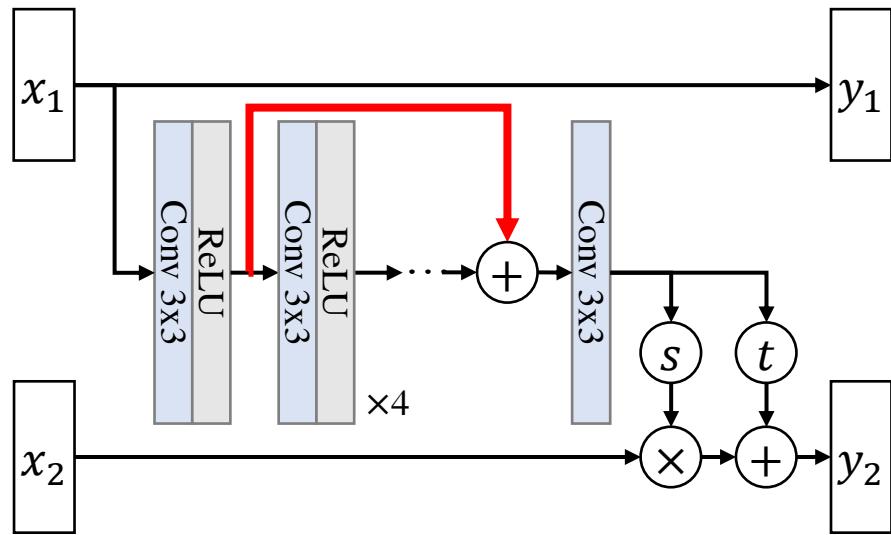


Our approach (SRFlow-DA)

- 5x5 Receptive field (RF).
- $K=16 \rightarrow$  Total 129x129 RF.
- $L=4$ .
- Params: 34.1M.
- Training time: 120h (RTX 2080 TI).
- Inference time: 1.97s (1920x1080 output).

- Stack 6 3x3 conv layers.
- Remove actnorm.
- 13x13 RF.
- $K=6 \rightarrow$  Total 145x145 RF.
- $L=3$ .
- Params: 13.3M.
- Training time: 47h (RTX 2080 TI).
- Inference time: 1.01s (1920x1080 output).

# Variants of SRFlow-DA

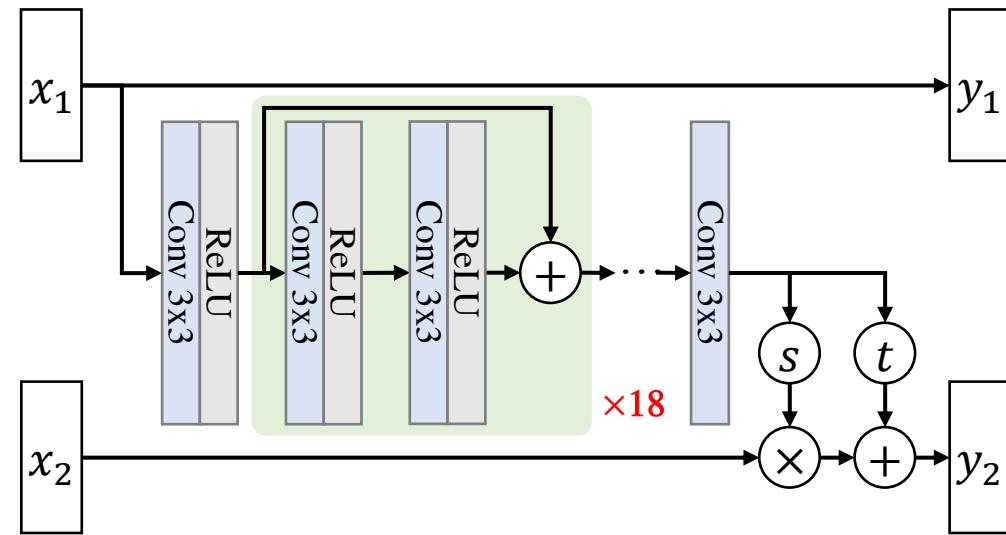


SRFlow-DA-R

- 13x13 RF.
- $K=6$ .
- $L=3$ .
- **Add a skip connection.**

SRFlow-DA-S

- 13x13 RF.
- **$K=18$ .**
- **$L=1$ .**
- For a single-scale architecture.



SRFlow-DA-D

- **77x77 RF.**
- **$K=1$ .**
- $L=3$ .
- To see if the performance can increase using only a deeper block without many flow steps.

Method	Params	Time		PSNR		LPIPS		Diversity	
		Training	Inference	LR	HR	mean	global		
Bicubic	-	-	-	38.70	26.70	0.409	0.409	0.409	0
RRDB	16.7M	-	-	49.20	<b>29.44</b>	0.253	0.253	0.253	0
ESRGAN	16.7M	-	-	39.03	26.22	0.124	<u>0.124</u>	0.124	0
SRFlow	22.8M	120h	1.98s	50.64	27.09	<b>0.120</b>	<b>0.119</b>	<b>0.089</b>	<u>25.24</u>
SRFlow-DA	8.7M	33h	1.18s	<u>50.88</u>	<u>27.57</u>	<u>0.121</u>	<b>0.119</b>	<u>0.092</u>	23.55
SRFlow-DA-R	8.7M	33h	1.19s	<b>50.92</b>	27.23	<b>0.120</b>	<b>0.119</b>	<b>0.089</b>	<b>25.50</b>
SRFlow-DA-S	8.8M	25h	0.91s	50.48	27.43	0.130	0.129	0.098	24.01
SRFlow-DA-D	6.4M	30h	1.01s	49.00	26.78	0.132	0.131	0.101	23.24

(a)  $\times 4$  SR results on DIV2K validation set.

Method	Params	Time		PSNR		LPIPS		Diversity	
		Training	Inference	LR	HR	mean	global		
Bicubic	-	-	-	37.16	23.74	0.584	0.584	0.584	0
RRDB	16.7M	-	-	45.43	<b>25.50</b>	0.419	0.419	0.419	0
ESRGAN	16.7M	-	-	31.35	22.18	0.277	0.277	0.277	0
SRFlow	34.1M	120h	1.97s	50.09	23.03	<b>0.272</b>	<b>0.276</b>	<b>0.200</b>	<b>25.28</b>
SRFlow-DA	13.3M	47h	1.01s	<u>50.91</u>	<u>23.75</u>	<u>0.261</u>	<u>0.259</u>	<u>0.198</u>	23.45
SRFlow-DA-R	13.3M	47h	1.04s	50.73	23.42	<b>0.256</b>	<b>0.253</b>	<b>0.191</b>	<u>24.57</u>
SRFlow-DA-S	15.0M	30h	0.65s	<b>52.21</b>	23.71	0.268	0.265	0.204	23.17
SRFlow-DA-D	9.5M	37h	0.80s	45.18	22.54	0.306	0.304	0.241	22.54

(b)  $\times 8$  SR results on DIV2K validation set.

Diversity: span of global and local

Worse diversity, but better global, local values

Bicubic



RRDB



SRFlow



SRFlow-DA



GT



SRFlow-DA



SRFlow-DA-R



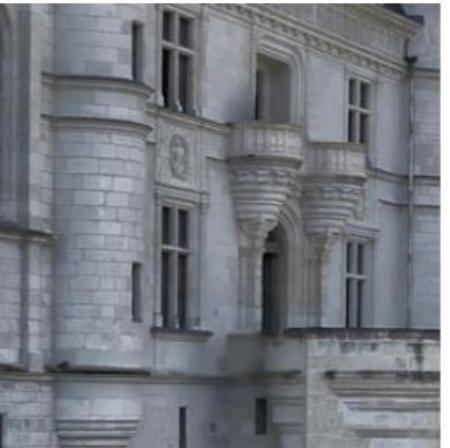
SRFlow-DA-S



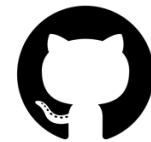
SRFlow-DA-D



GT



# Thank you



<https://github.com/yhjo09/SRFlow-DA>