FS-NCSR: Increasing Diversity of the Super-Resolution Space via Frequency Separation and Noise-Conditioned Normalizing Flow

Ki-Ung Song*, Dongseok Shim*, Kang-wook Kim*, Jae-young Lee, Younggeun Kim
Seoul National University, Seoul, Republic of Korea / Deepest, Seoul, Republic of Korea
{sk851, tlaehdtjr01, full324, jerry96, eyfydsyd97}@snu.ac.kr

Introduction

- Super-resolution models predict the missing high-frequency information of the HR images from the given LR image.
- Predicting not only high-frequency but also low-frequency information given LR images is inefficient.
- We propose FS-NCSR (Frequency Separated Noise-Conditioned Normalizing Flow for Super Resolution), which applies frequency separation to NCSR.

Methods

- Given a LR image, our goal is to learn a diverse super-resolution space corresponding to that image.
- The flow-based super-resolution model configures a mapping fθ : X → Z between the desired data distribution X and latent space distribution Z , by maximizing the negative log-likelihood:

$$-\log p_X(x) = -\log p_Z(f_\theta(x)) - \log|\det \frac{\partial f_\theta}{\partial x}(x)|. \quad (1)$$

- SRFlow [1] and NCSR [2] acheived more diverse and equally photo-realistic super-resolution results, compared to GAN-based models (e.g. ESRGAN [3]).


RGB image


High Frequency

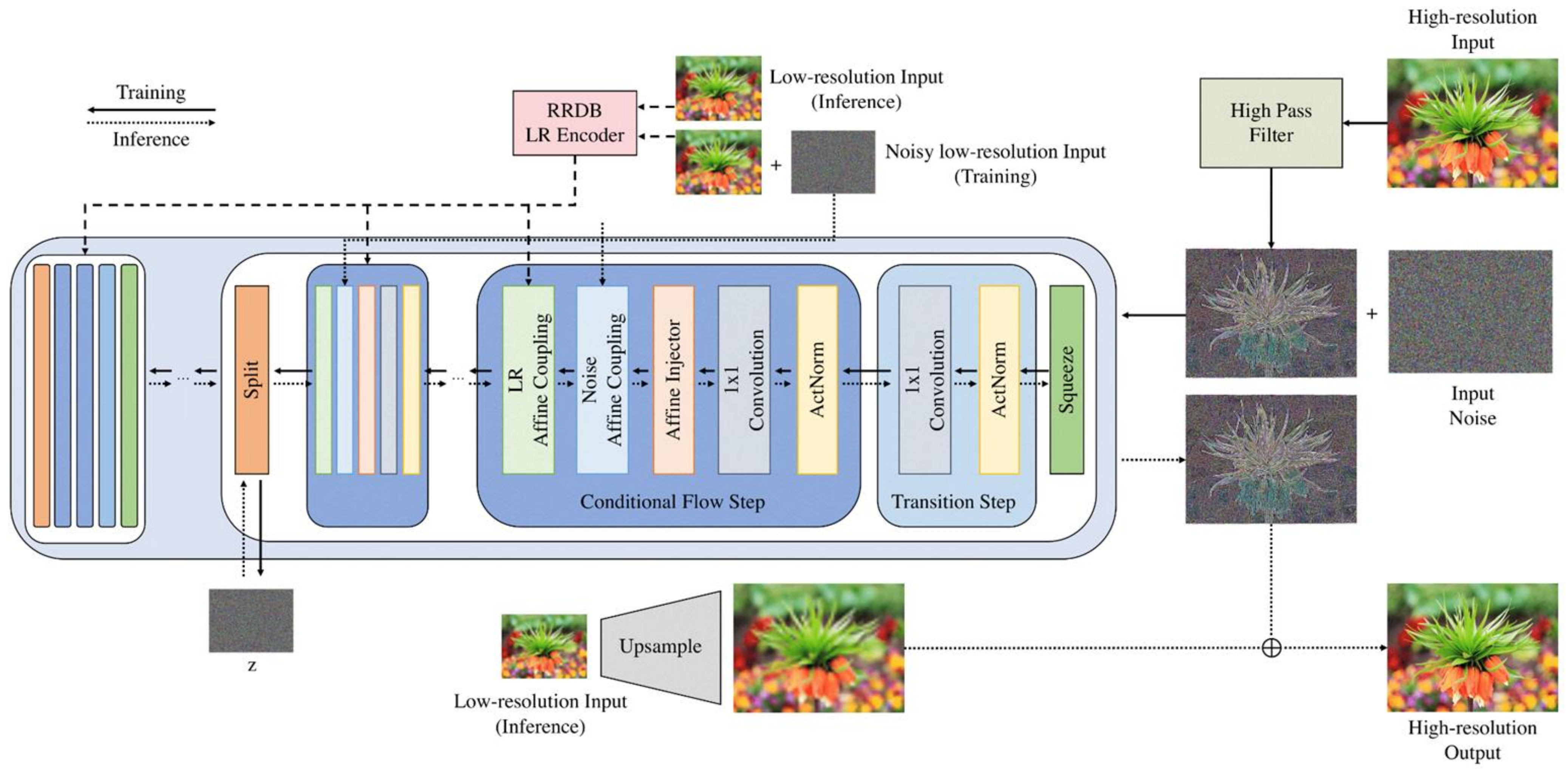
- High-frequency information is relatively sparse compared to its original RGB images.

- We adopt frequency separation to NCSR, namely FS-NCSR.
- The LR image contains sufficient information in the low-frequency domain of the desired HR image x, thus, a bicubic downsampling-upsampling process with scale factor s can be interpreted as a low-pass filer, L_s.
- With a simple low-pass filter, the LR image can be seen as a low-pass filtered image, which leads to a simple high-pass filter H_s.

$$L_s(x) = ((x)_{s\downarrow})_{s\uparrow}, \quad H_s(x) = x_{hf} = x - L_s(x). \quad (2)$$

- From the image pair of LR and HR images, the readily defined high-pass filter enables a frequency separation, an extraction of high-frequency information of the HR image.

- The below image is an overview of the FS-NCSR. FS-NCSR focuses on restoring high-frequency information with frequency separation and noise-conditioning.



Results

- FS-NCSR shows the performance in generating diverse super-resolution outputs compared to other state-of-the-art algorithms due to the frequency separation.

Model	Diversity↑	LPIPS↓	LR PSNR↑
RRDB [33]	0	0.253	49.20
ESRGAN [33]	0	0.124	39.03
ESRGAN+ [24]	22.13	0.279	35.45
SRFlow [23]	25.26	0.120	49.97
HCFlow [19]	22.73	**0.116**	49.46
NCSR [13]	26.72	0.119	**50.75**
FS-NCSR (Ours)	**29.44**	0.127	49.31

Model	Diversity↑	LPIPS↓	LR PSNR↑
RRDB [33]	0	0.419	45.43
ESRGAN [33]	0	0.277	31.35
SRFlow [23]	25.31	0.272	**50.00**
NCSR [13]	26.8	0.278	44.55
FS-NCSR (Ours)	**26.9**	**0.257**	48.90


(a) ×4 LR (b) NCSR [13] (c) FS-NCSR (Ours) (d) Ground Truth


(a) Output 1 (b) Output 2 (c) Output 3 (d) Output 4 (e) Output 5 (f) Ground Truth

Conclusion

- With a simple high pass filter using the characteristics of the LR image, FS-NCSR concentrated on filling the missing information in the high-frequency domain of the desired image by frequency separation and increased the diversity of super-resolution outputs.

[1] Lugmayr, Andreas, et al. "Srflow: Learning the super-resolution space with normalizing flow." European conference on computer vision. Springer, Cham, 2020.
[2] Kim, Younggeun, and Donghee Son. "Noise conditional flow model for learning the super-resolution space." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
[3] Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." Proceedings of the European conference on computer vision (ECCV) workshops. 2018.