NTIRE 2022 Challenge on Learning the Super-Resolution Space

ETH zürich | CVL Computer Vision Lab

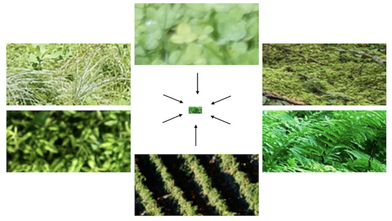
Andreas Lugmayr*, Martin Danelljan*, Radu Timofte*
ETH Zurich

Zhihao Zhao, Jinhui Tang, Zechao Li, Jinshan Pan
Nanjing University of Science and Technology

Ki-Ung Song, Dongseok Shim, Kang-wook Kim, Jae-young Lee, Younggeun Kim
Seoul National University

GitHub
git.io/SR22

Introduction



Motivation
- Super-Resolution (SR) is an ill-posed problem
- Many SR methods map a low-resolution image to *one* high-resolution image

Goals
- Stimulate research into leaning the full space of plausible SR
- Establish benchmark protocols and metrics for stochastic SR
- Probe the state-of-the-art in SR
- Compare stochastic generation technologies for Stochastic SR

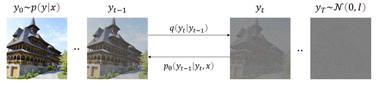
VAE-Based

Team SR_DL
- Best performing VAE-Based Method in NTIRE21
- Use adversarial loss to improve visual quality

General
- Diversity due to stochastic VAE property
- Advantage over Flow: Does not require special model architecture

Diffusion-Based

Teams IMAG_WZ and IMAG_ZW
- Submission to NTRE 22
- Condition Denoising Diffusion Probabilistic Model
- Inference:
  - Sample from pure noise
  - Denoise conditioned on low-resolution
  - Generate SR in T steps



MAG_ZW
- Uses Gamma distribution which better adapts to residual noise

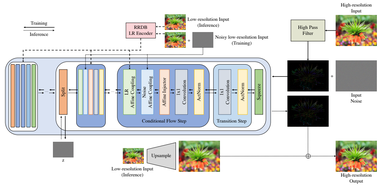
IMLE-Based

Team FutureReference
- High-Quality conditional image generation
- Diverse output
- Training objective reverses the generation process to match SR with real data
- Only Team submitting IMLE
- Submitted to NTIRE21

Flow-Based

General
- Training: Transform high-resolution image to noise conditioned on the low-resolution image
- Model is invertible and preserves probability volume
- Inference: Transform noise to high-resolution image conditioned on low-resolution image

Team Deepest
- Submission to NTRE 22
- Noise-Conditioned affine coupling
- Frequency Separation
- Model focuses on high-frequency outputs
- Add noise on the sparce high-frequency image for more diversity



GAN-Based

Team SSS
- Submitted to NTIRE21
- Best performing purely GAN-Based method
- Uses Spatial Feature Transformation layers to generate stochastic SR

General
- Struggle to generate large diversity in SR
- Adversarial loss encourages hallucinations
- No purely GAN-Based Method in NTIRE21 reached the required LR-PSNR

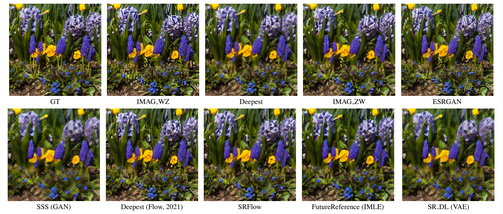
Quantitative Results

	Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
4×	IMAG_ZW	0.171	48.14	21.938$_{(3)}$	3.57$_{(2)}$	2.5
	Deepest	0.126	50.13	28.853$_{(1)}$	3.67$_{(3)}$	2.5
	IMAG_WZ	0.169	45.20	27.320$_{(2)}$	3.34$_{(1)}$	1.5
	FutureReference (IMLE)	0.165	37.51	19.636	-	-
	SR_DL (VAE)	0.234	39.80	20.508	-	-
	SSS (GAN)	0.110	44.70	13.285	-	-
	Deepest (Flow)	0.117	50.54	26.041	-	-
	SRFlow	0.122	49.86	25.008	3.62	-
	ESRGAN	0.124	38.74	0.000	3.52	-
	GT	0	∞	-	3.15	-

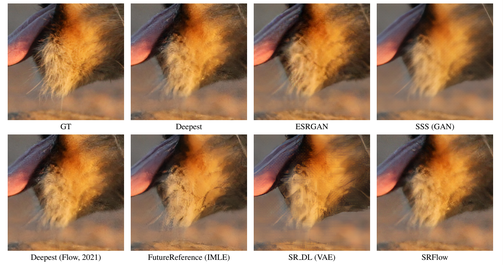
	Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
8×	Deepest	0.257	50.37	26.539	4.510	-
	FutureReference (IMLR)	0.291	36.51	17.985	4.741	-
	SSS (GAN)	0.237	37.43	13.548	4.850	-
	SR_DL (VAE-GAN)	0.311	42.28	14.817	4.797	-
	Deepest (Flow)	0.259	48.64	26.941	4.503	-
	SRFlow	0.282	47.72	25.582	4.775	-
	ESRGAN	0.284	30.65	0	4.452	-
	GT	0	∞	-	3.173	-

Visual Examples

4 × Super-Resolution



8 × Super-Resolution



Samples of IMAG_WZ

Method Comparison

4 × Super-Resolution

8 × Super-Resolution