NCSR:NOISE CONDITIONAL FLOW MODEL FOR LEARNING THE SUPER-RESOLUTION SPACE

Younggeun Kim, Donghee Son

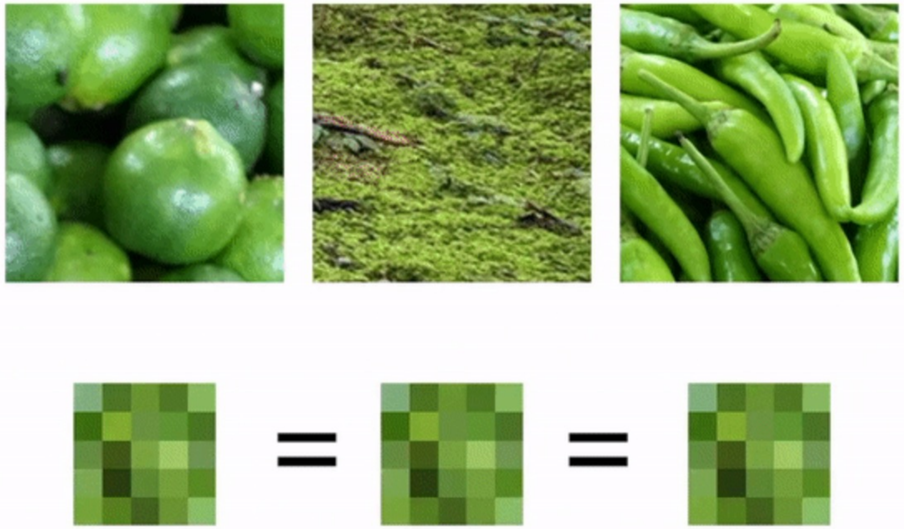
Seoul National Univerity and Deepest

Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
svnit_ntnu	0.355	27.52	$1.871_{(11)}$	-	-
SYSU-FVL	0.244	49.33	$8.735_{(10)}$	-	-
nanbeihuishi	0.161	50.46	$12.447_{(9)}$	-	-
FudanZmic21	0.273	47.20	$16.450_{(7)}$	-	-
FutureReference	0.165	37.51	$19.636_{(6)}$	-	-
SR_DL	0.234	39.80	$20.508_{(5)}$	-	-
SSS	0.110	44.70	$13.285_{(8)}$	$4.530_{(3)}$	5.5
BeWater	0.137	49.59	$23.948_{(3)}$	$4.720_{(4)}$	3.5
CIPLAB	0.121	50.70	$23.091_{(4)}$	$4.478_{(2)}$	3.0
njtech&seu	0.149	46.74	$\mathbf{26.924}_{(1)}$	$4.977_{(5)}$	3.0
Deepest	0.117	50.54	$26.041_{(2)}$	$\mathbf{4.372}_{(1)}$	**1.5**
SRFlow	0.122	49.86	25.008	4.410	-
ESRGAN	0.124	38.74	0.000	4.467	-
GT	0	∞	-	3.728	-

Table 2. Quantitative comparison of participating teams. (4×)

Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
svnit_ntnu	0.481	25.55	$4.516_{(10)}$	-	-
SYSU-FVL	0.415	47.27	$8.778_{(9)}$	-	-
FudanZmic21	0.496	46.78	$14.287_{(7)}$	-	-
FutureReference	0.291	36.51	$17.985_{(5)}$	-	-
njtech&seu	0.366	29.65	$\mathbf{28.193}_{(1)}$	-	-
SSS	0.237	37.43	$13.548_{(8)}$	$4.692_{(3)}$	5.5
SR_DL	0.311	42.28	$14.817_{(6)}$	$4.738_{(4)}$	5.0
BeWater	0.297	49.63	$23.700_{(3)}$	$5.133_{(5)}$	4.0
CIPLAB	0.266	50.86	$23.320_{(4)}$	$4.637_{(2)}$	3.0
Deepest	0.259	48.64	$26.941_{(2)}$	$\mathbf{4.630}_{(1)}$	**1.5**
SRFlow	0.282	47.72	25.582	4.635	-
ESRGAN	0.284	30.65	0	4.323	-
GT	0	∞	-	2.613	-

Table 3. Quantitative comparison of participating teams. (8×)

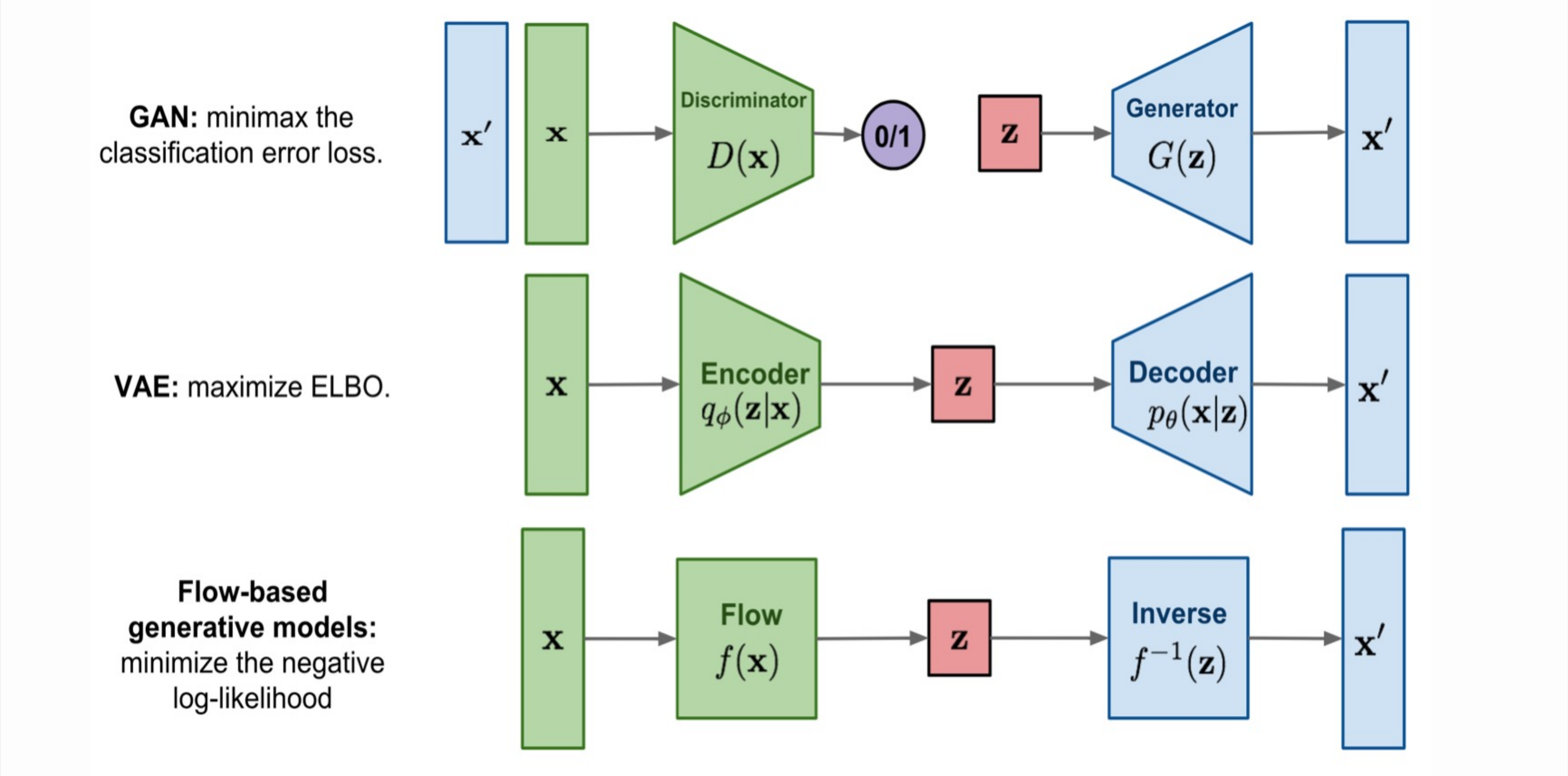


Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
svnit_ntnu	0.355	27.52	$1.871_{(11)}$	-	-
SYSU-FVL	0.244	49.33	$8.735_{(10)}$	-	-
nanbeihuishi	0.161	50.46	$12.447_{(9)}$	-	-
FudanZmic21	0.273	47.20	$16.450_{(7)}$	-	-
FutureReference	0.165	37.51	$19.636_{(6)}$	-	-
SR_DL	0.234	39.80	$20.508_{(5)}$	-	-
SSS	0.110	44.70	$13.285_{(8)}$	$4.530_{(3)}$	5.5
BeWater	0.137	49.59	$23.948_{(3)}$	$4.720_{(4)}$	3.5
CIPLAB	0.121	50.70	$23.091_{(4)}$	$4.478_{(2)}$	3.0
njtech&seu	0.149	46.74	$\mathbf{26.924}_{(1)}$	$4.977_{(5)}$	3.0
Deepest	0.117	50.54	$26.041_{(2)}$	$\mathbf{4.372}_{(1)}$	**1.5**
SRFlow	0.122	49.86	25.008	4.410	-
ESRGAN	0.124	38.74	0.000	4.467	-
GT	0	∞	-	3.728	-

Table 2. Quantitative comparison of participating teams. (4×)

Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
svnit_ntnu	0.481	25.55	$4.516_{(10)}$	-	-
SYSU-FVL	0.415	47.27	$8.778_{(9)}$	-	-
FudanZmic21	0.496	46.78	$14.287_{(7)}$	-	-
FutureReference	0.291	36.51	$17.985_{(5)}$	-	-
njtech&seu	0.366	29.65	$\mathbf{28.193}_{(1)}$	-	-
SSS	0.237	37.43	$13.548_{(8)}$	$4.692_{(3)}$	5.5
SR_DL	0.311	42.28	$14.817_{(6)}$	$4.738_{(4)}$	5.0
BeWater	0.297	49.63	$23.700_{(3)}$	$5.133_{(5)}$	4.0
CIPLAB	0.266	50.86	$23.320_{(4)}$	$4.637_{(2)}$	3.0
Deepest	0.259	48.64	$26.941_{(2)}$	$\mathbf{4.630}_{(1)}$	**1.5**
SRFlow	0.282	47.72	25.582	4.635	-
ESRGAN	0.284	30.65	0	4.323	-
GT	0	∞	-	2.613	-

Table 3. Quantitative comparison of participating teams. (8×)



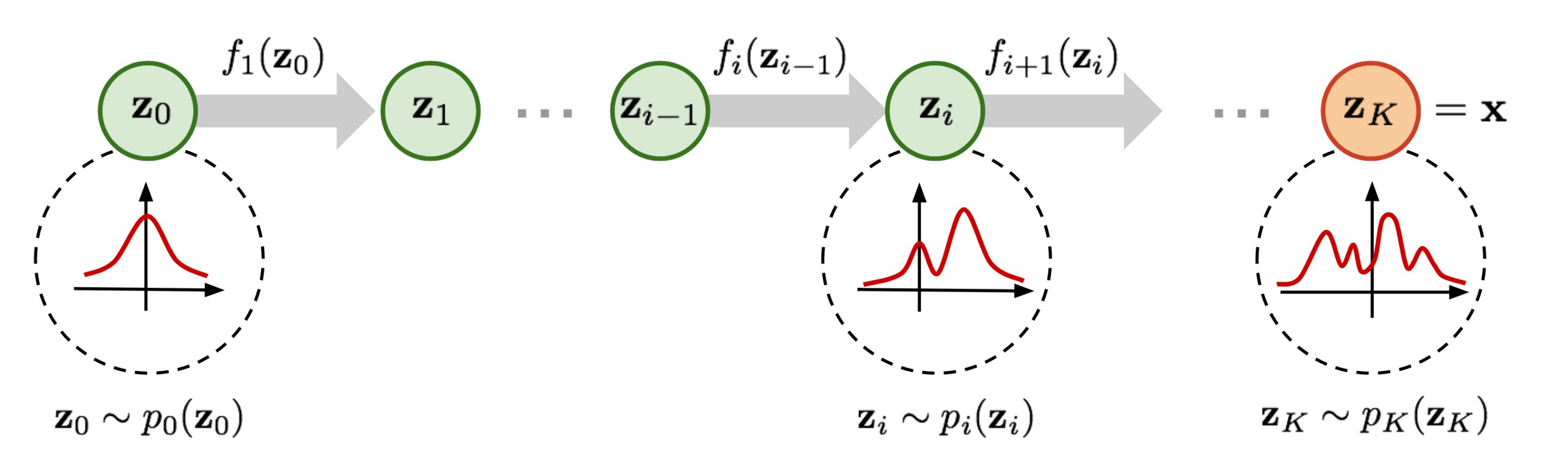
-Simple distribution to Complex distribution

-Inverse functions

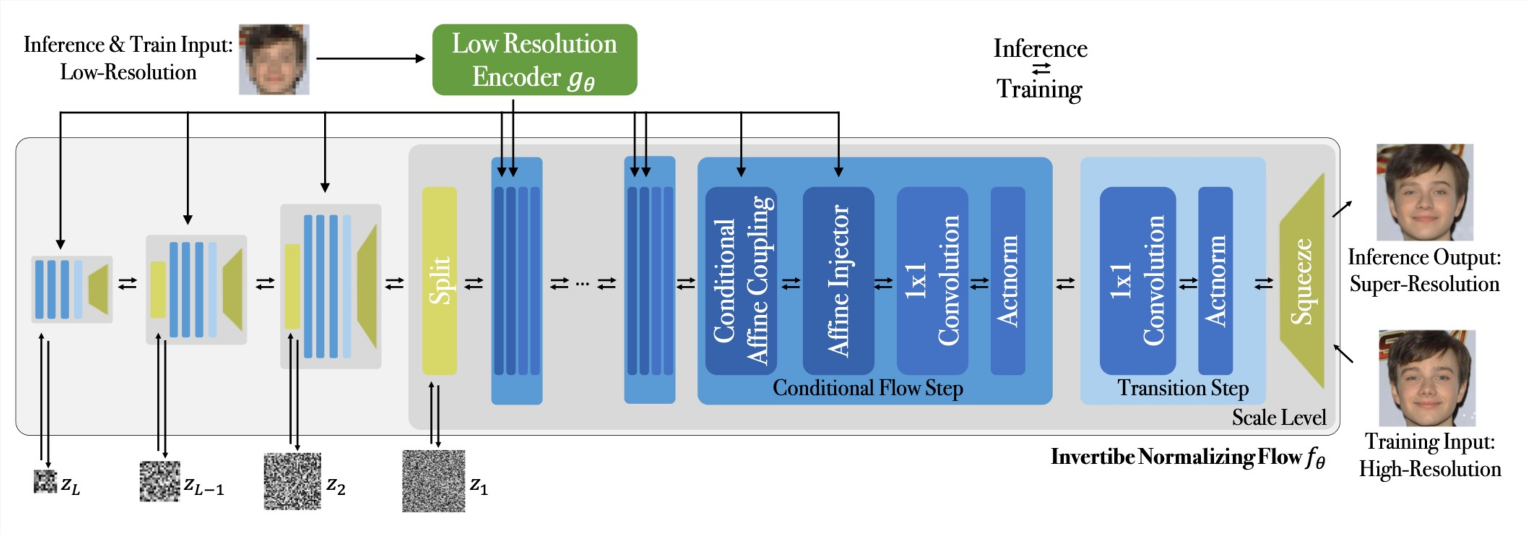
-Not deterministic model

SRFLOW

- SRFlow: Super-Resolution + Normalizing Flow
- Normalizing Flow : Generative model, $P(x|z)$
  - $X = z_k = f_k \circ \cdots f_2 \circ f_1(z_0)$
  - $\log P(X) = \log \pi_0(z_0) - \sum_{i=1}^{k} \log|\det \frac{df_i}{dz_{i-1}}|$
    - Function f_i
      - Invertible & easy to compute jacobian determinant



SRFLOW



Normalizing Flow

1. $z_1,\ z_2 = split(z)$
2. $scale, bias = fn(z_2)$
3. $z_1 = \frac{z_1}{scale} - bias$
4. $z_{new} = concat(z_1, z_2)$

Conditional Normalizing Flow

1. $z_1,\ z_2 = split(z)$
2. $scale, bias = fn(z_2,\ LR\_image)$
3. $z_1 = \frac{z_1}{scale} - bias$
4. $z_{new} = concat(z_1, z_2)$

Affine Injector

1. $scale, bias = fn(LR\_image)$
2. $z = \frac{z}{scale} - bias$

DIV2k 828
LR-PSNR: 21.53

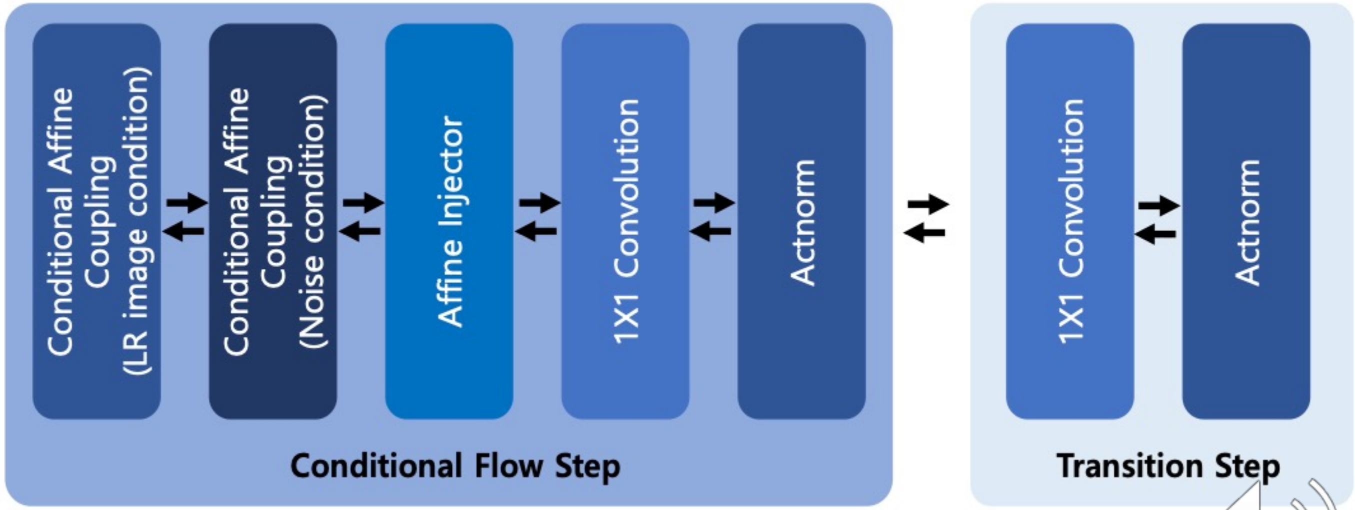
DIV2k 873
LR-PSNR: 24.11

Adding
Noise to generate
Various images

DIV2k 828
LR-PSNR: 21.53

DIV2k 873
LR-PSNR: 24.11

Noise Conditional Layer

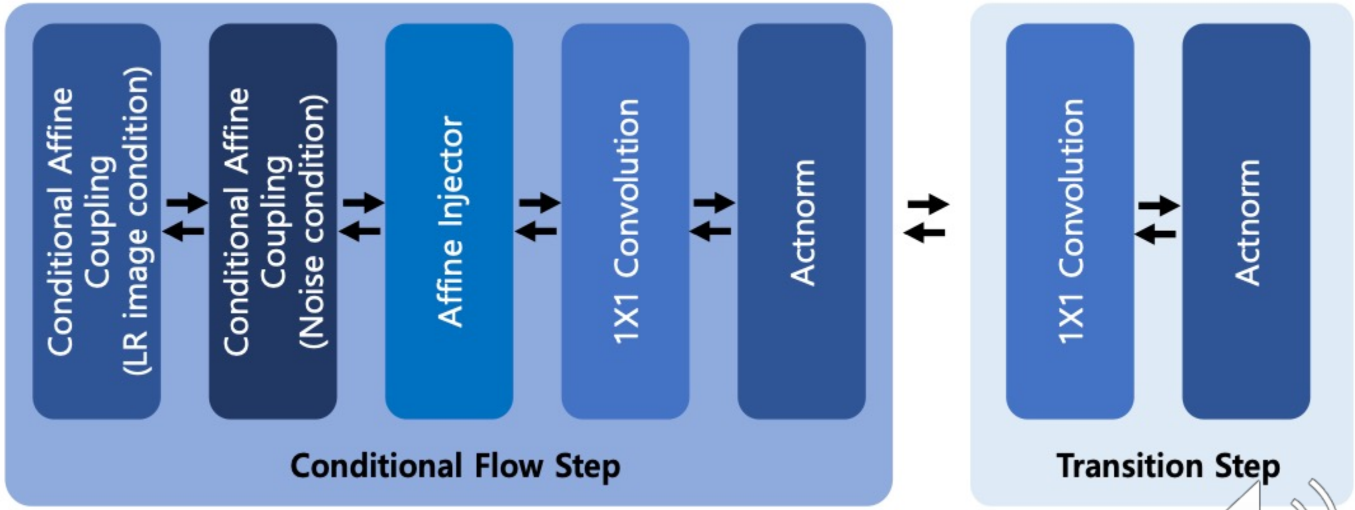


Github repo: https://github.com/younggeun-kim/NCSR

DIV2k 828
LR-PSNR: 21.53

DIV2k 873
LR-PSNR: 24.11

Model	w/o NCL	with NCL
Diversity	25.38	26.72
LPIPS	0.1228	0.1193
LR PSNR	50.08	50.75
LR PSNR-worst	47.32	49.14



NCSR

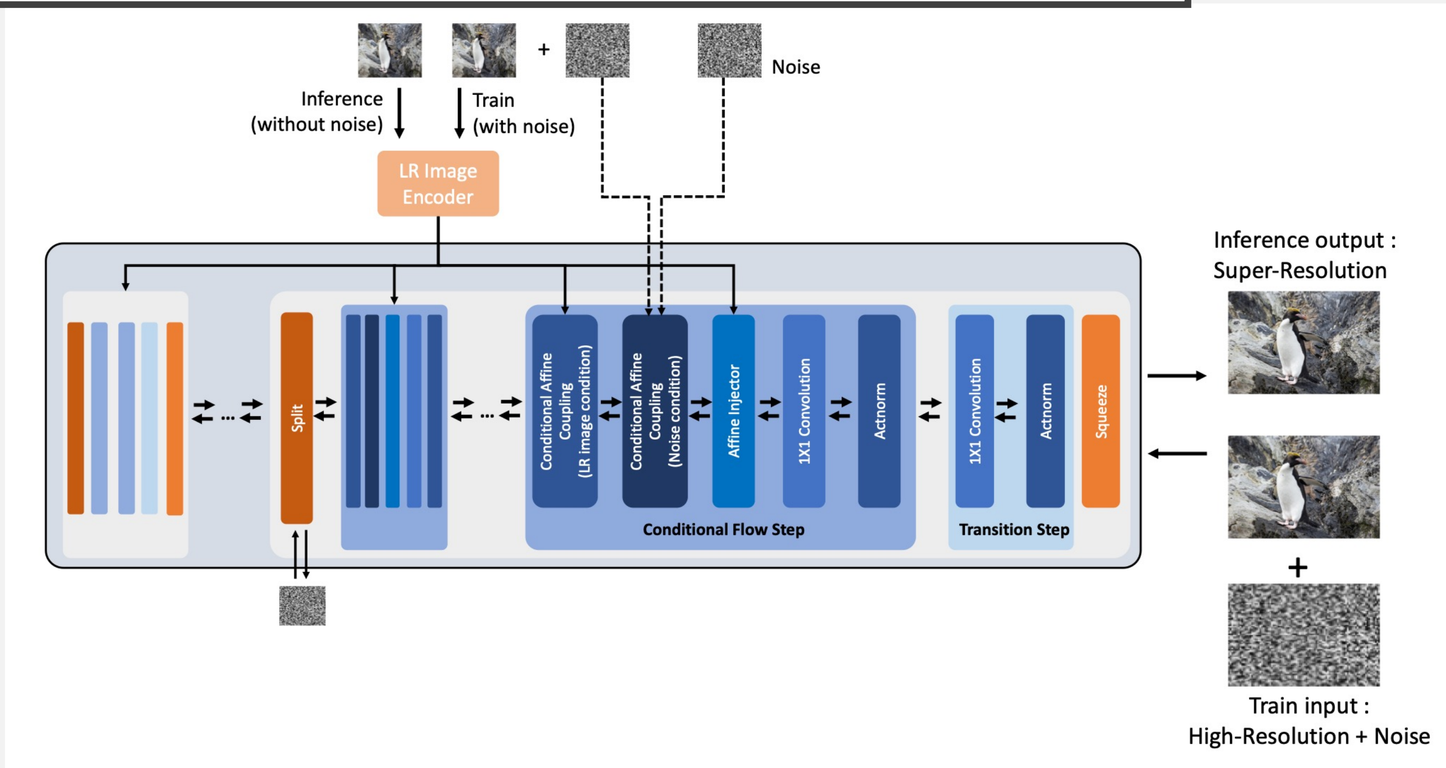
- Noise Conditional Super-Resolution Model
  - Noise perturbed data
  - Noise Conditional Flow
- 1. Noise sampling
- 2. Data Perturbation
  - $x^+ = x + v$ (GT)
  - $y^+ = y + w$ (LR), (w is dowsampled v)
- 3. Noise Conditional Flow
  - $f^{-1}(x^+|y^+, v) = z$

Noise sampling

$$c_i \sim Unif[a, b]$$
$$\Sigma_i = c_i^2 I$$
$$v_i \sim N(0, \Sigma_i)$$

NCSR

- 1. Noise sampling
- 2. Data Perturbation
  - $x^{+} = x + v$ (GT)
  - $y^{+} = y + w$ (LR)
- 3. Noise Conditional Flow
  - $f^{-1}(x^{+}|y^{+}, v) = z$



Normalizing Flow

1. $z_1,\ z_2 = split(z)$
2. $scale, bias = fn(z_2)$
3. $z_1 = \frac{z_1}{scale} - bias$
4. $z_{new} = concat(z_1, z_2)$

Conditional Normalizing Flow

1. $z_1,\ z_2 = split(z)$
2. $scale, bias = fn(z_2,\ y)$
3. $z_1 = \frac{z_1}{scale} - bias$
4. $z_{new} = concat(z_1, z_2)$

Conditional Normalizing Flow with Noise

1. $z_1,\ z_2 = split(z)$
2. $scale, bias = fn(z_2,\ y^{+}, v)$
3. $z_1 = \frac{z_1}{scale} - bias$
4. $z_{new} = concat(z_1, z_2)$

DATASET

- Train data : DF2K
  - 800 + 2650 2K resolution images
- Test / Valid : DIV2K
  - 100 2K resolution images
- Additional Dataset is allowed except DIV2K valid and test datas
  - Unsplash2K
    - Crawled from unsplash.com
    - 498 pairs
  - http://github.com/dongheehand/unsplash2K

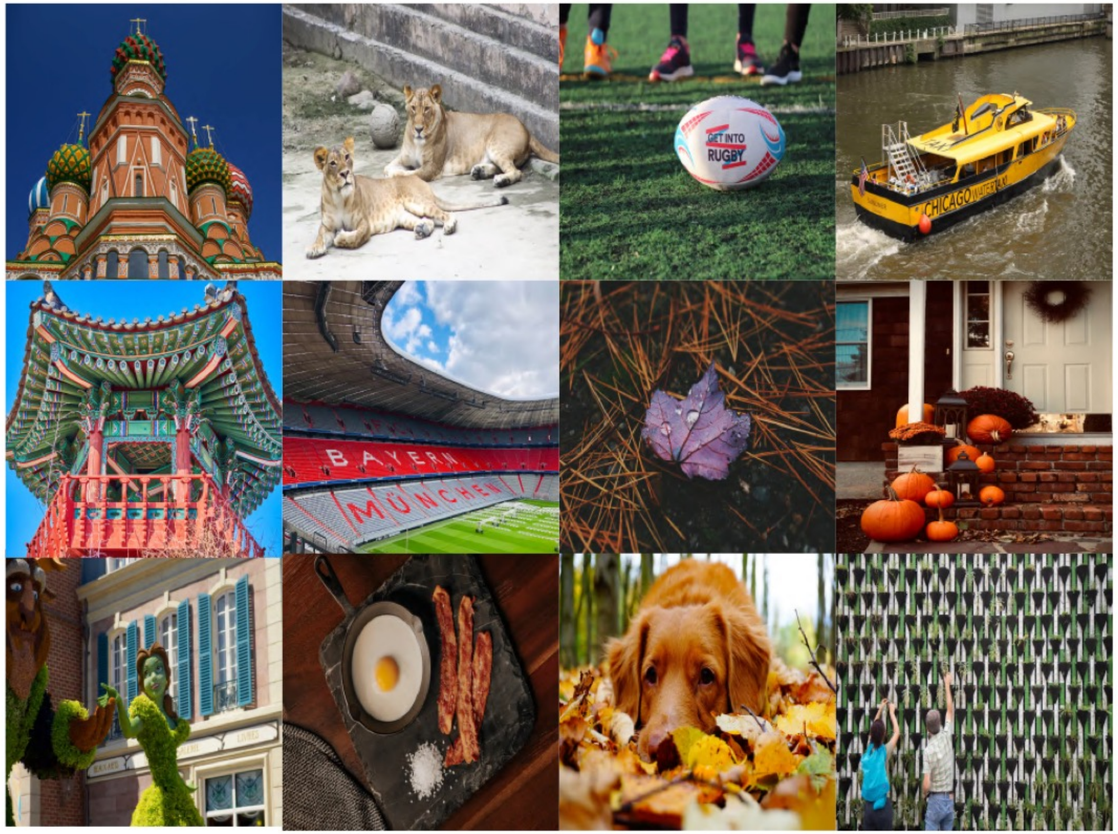


Figure 5: The sample images of Unsplash 2K

New Super Resolution Dataset

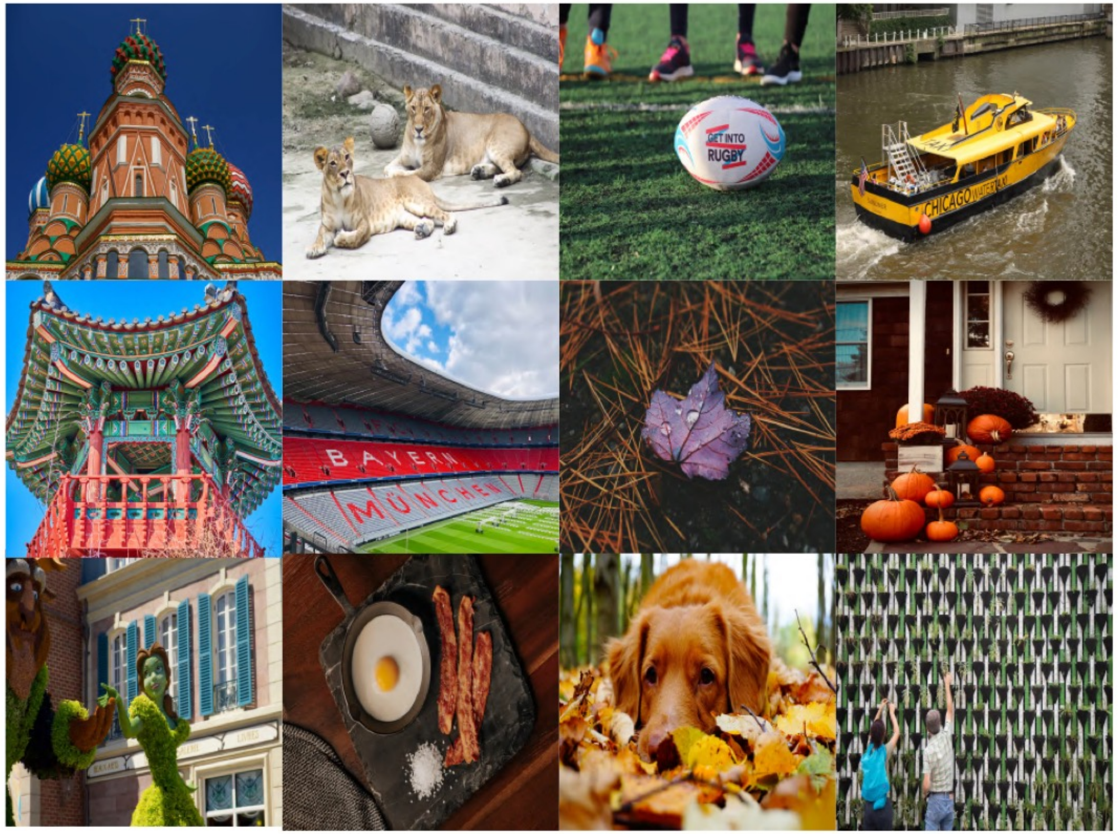


Figure 5: The sample images of Unsplash 2K

New Super Resolution Dataset

Noise Conditional Layer	✗	✗	✓	✓	✓
Std Conditional Layer	✗	✓	✗	✗	✗
Noise-free block	✗	✓	✗	✓	✓
Add extra data	✗	✗	✗	✗	✓
LR-PSNR worst	47.32	45.78	49.01	49.14	50.13

Ablation Study

HR | Bicubic | ESRGAN | SRFlow | NCSR (Ours)

HR Bicubic ESRGAN SRFlow NCSR (Ours)

Model	Diversity	LPIPS	LR PSNR
RRDB [28]	0	0.253	49.20
ESRGAN [28]	0	0.124	39.03
ESRGAN+ [26]	22.13	0.279	35.45
SRFlow [22]	25.26	0.120	49.97
NCSR (Ours)	26.72	0.119	50.75
NCSR* (Ours)	26.79	0.118	50.88

Table 1: General image SR $\times 4$ results on the 100 validation images of the DIV2K dataset

HR Bicubic ESRGAN SRFlow NCSR (Ours)

Model	Diversity	LPIPS	LR PSNR
RRDB [28]	0	0.253	49.20
ESRGAN [28]	0	0.124	39.03
ESRGAN+ [26]	22.13	0.279	35.45
SRFlow [22]	25.26	0.120	49.97
NCSR (Ours)	26.72	0.119	50.75
NCSR* (Ours)	26.79	0.118	50.88

Model	Diversity	LPIPS	LR PSNR
RRDB [28]	0	0.419	45.43
ESRGAN [28]	0	0.277	31.35
SRFlow [22]	25.31	0.272	50.00
NCSR (Ours)	26.8	0.278	44.55
NCSR* (Ours)	25.7	0.253	49.97

Table 2: General image SR $\times$8results on the 100 validation images of the DIV2K dataset

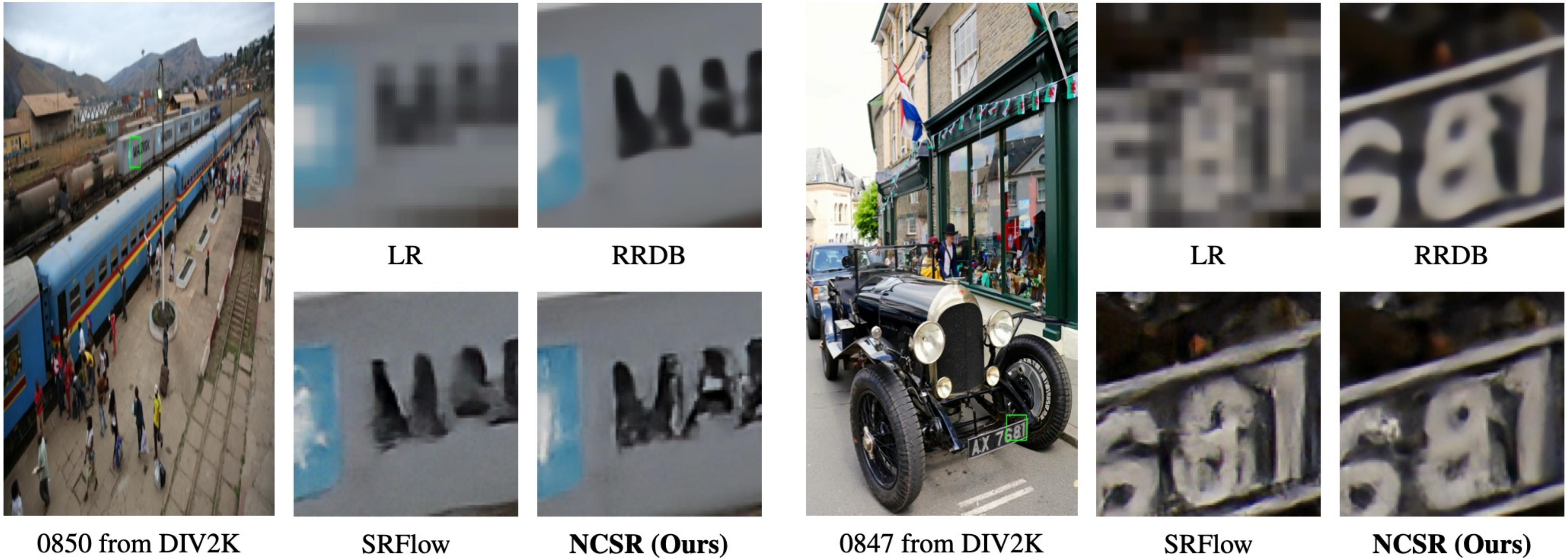


Figure 6: Qualitative comparisons with other methods for $\times 8$ SR model.

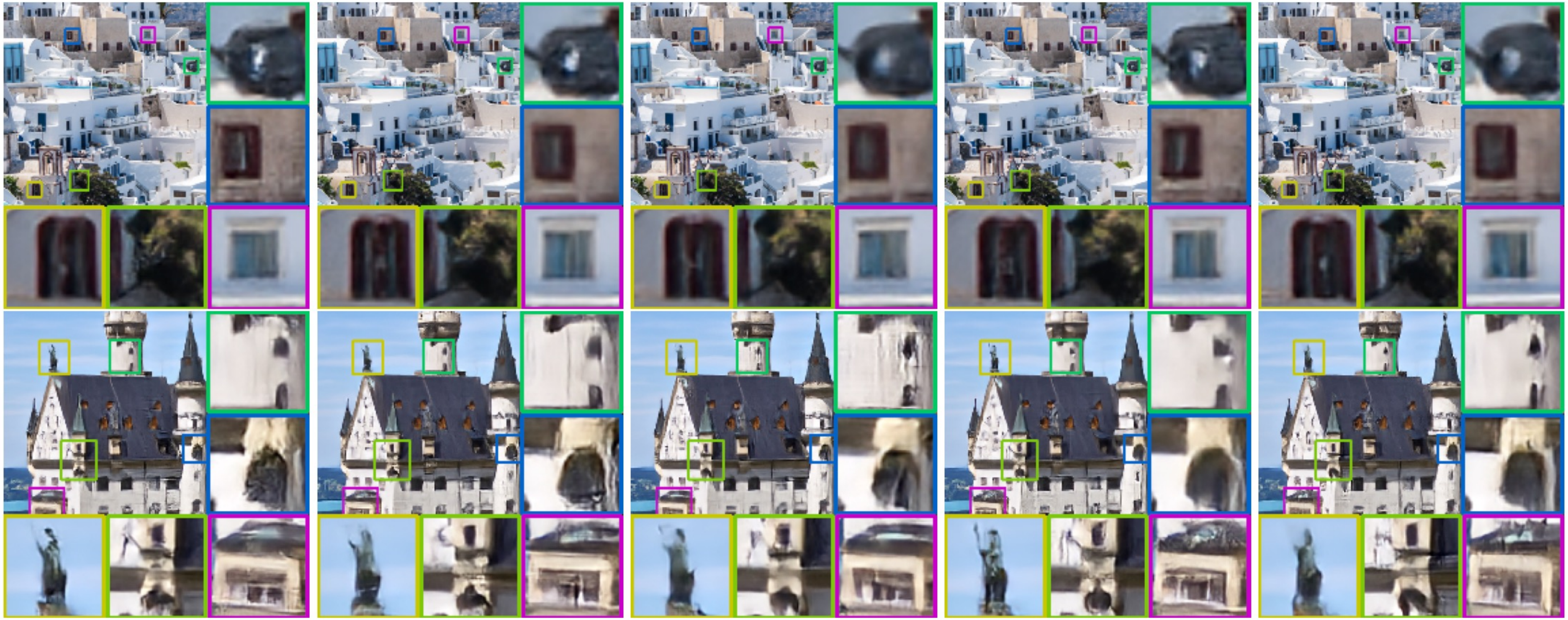


Figure 7: Random samples generated by NCSR. Upper : ×4 SR model, Lower : ×8 SR model

X4 task

Figure 2. Qualitative comparison between the participating approaches for 4× super-resolution

Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
svnit_ntnu	0.355	27.52	$1.871_{(11)}$	-	-
SYSU-FVL	0.244	49.33	$8.735_{(10)}$	-	-
nanbeihuishi	0.161	50.46	$12.447_{(9)}$	-	-
FudanZmic21	0.273	47.20	$16.450_{(7)}$	-	-
FutureReference	0.165	37.51	$19.636_{(6)}$	-	-
SR_DL	0.234	39.80	$20.508_{(5)}$	-	-
SSS	0.110	44.70	$13.285_{(8)}$	$4.530_{(3)}$	5.5
BeWater	0.137	49.59	$23.948_{(3)}$	$4.720_{(4)}$	3.5
CIPLAB	0.121	50.70	$23.091_{(4)}$	$4.478_{(2)}$	3.0
njtech&seu	0.149	46.74	$\mathbf{26.924}_{(1)}$	$4.977_{(5)}$	3.0
Deepest	0.117	50.54	$26.041_{(2)}$	$\mathbf{4.372}_{(1)}$	**1.5**
SRFlow	0.122	49.86	25.008	4.410	-
ESRGAN	0.124	38.74	0.000	4.467	-
GT	0	∞	-	3.728	-

Table 2. Quantitative comparison of participating teams. (4×)

X8 task

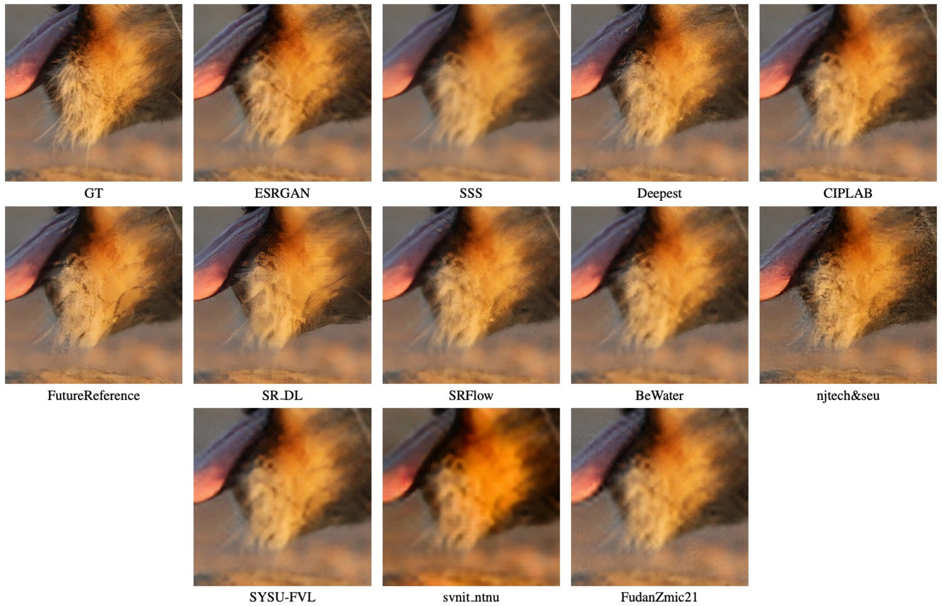


Figure 3. Qualitative comparison between the participating approaches for 8× super-resolution

Team	LPIPS	LR-PSNR	Div. Score S_{10} [%]	MOR	Final Rank
svnit_ntnu	0.481	25.55	$4.516_{(10)}$	-	-
SYSU-FVL	0.415	47.27	$8.778_{(9)}$	-	-
FudanZmic21	0.496	46.78	$14.287_{(7)}$	-	-
FutureReference	0.291	36.51	$17.985_{(5)}$	-	-
njtech&seu	0.366	29.65	$\mathbf{28.193}_{(1)}$	-	-
SSS	0.237	37.43	$13.548_{(8)}$	$4.692_{(3)}$	5.5
SR_DL	0.311	42.28	$14.817_{(6)}$	$4.738_{(4)}$	5.0
BeWater	0.297	49.63	$23.700_{(3)}$	$5.133_{(5)}$	4.0
CIPLAB	0.266	50.86	$23.320_{(4)}$	$4.637_{(2)}$	3.0
Deepest	0.259	48.64	$26.941_{(2)}$	$\mathbf{4.630}_{(1)}$	**1.5**
SRFlow	0.282	47.72	25.582	4.635	-
ESRGAN	0.284	30.65	0	4.323	-
GT	0	∞	-	2.613	-

Table 3. Quantitative comparison of participating teams. (8×)

SUMMARY

- Flow-based model(not deterministic model)
- Noise adding to generate diverse super resolution images
- Noise conditional layer to inform noise to model
- Solving dimension mismatch leads good perceptual quality
- Got 1st place in NTIRE 2021, Learning for the Super Resolution task both of X4, X8 tasks.
- Github repo: https://github.com/younggeun-kim/NCSR
- If you have question about paper or presentation, please send mail to eyfydsyd97@snu.ac.kr