Image Super-Resolution via Deep Learning

14 September 2023

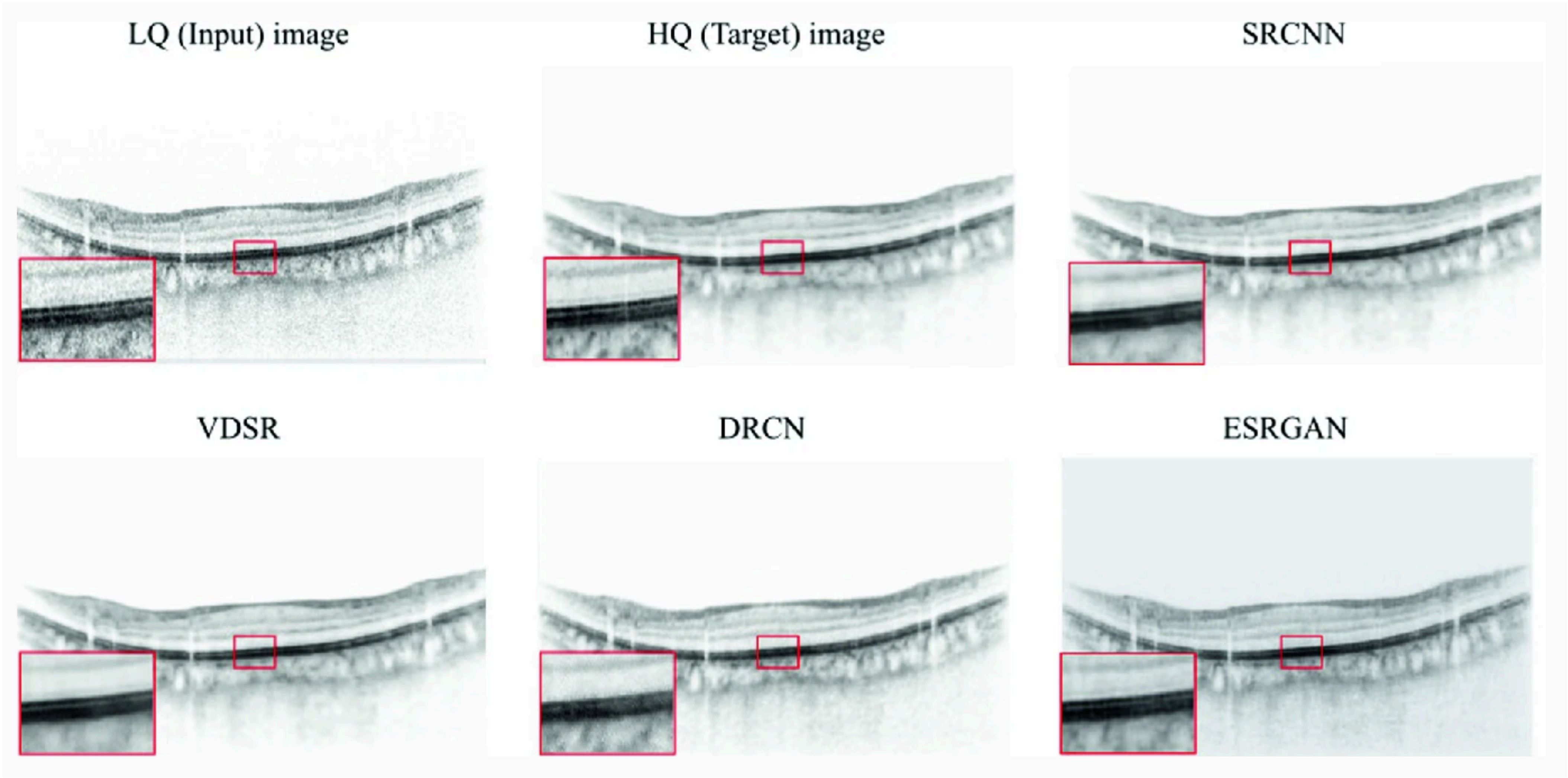
Valdivino Alexandre de Santiago Júnior

Coordenação de Pesquisa Aplicada e Desenvolvimento Tecnológico (COPDT)
Instituto Nacional de Pesquisas Espaciais (INPE)
São José dos Campos, SP, Brazil

Image Super-Resolution (SR)

- Goal: to recover high-resolution (HR) images from low-resolution (LR) ones.
- Resolution: the **dimensionality** of the image. For instance, an image has a resolution of $W \times H$ pixels.
- See: https://iterative-refinement.github.io/

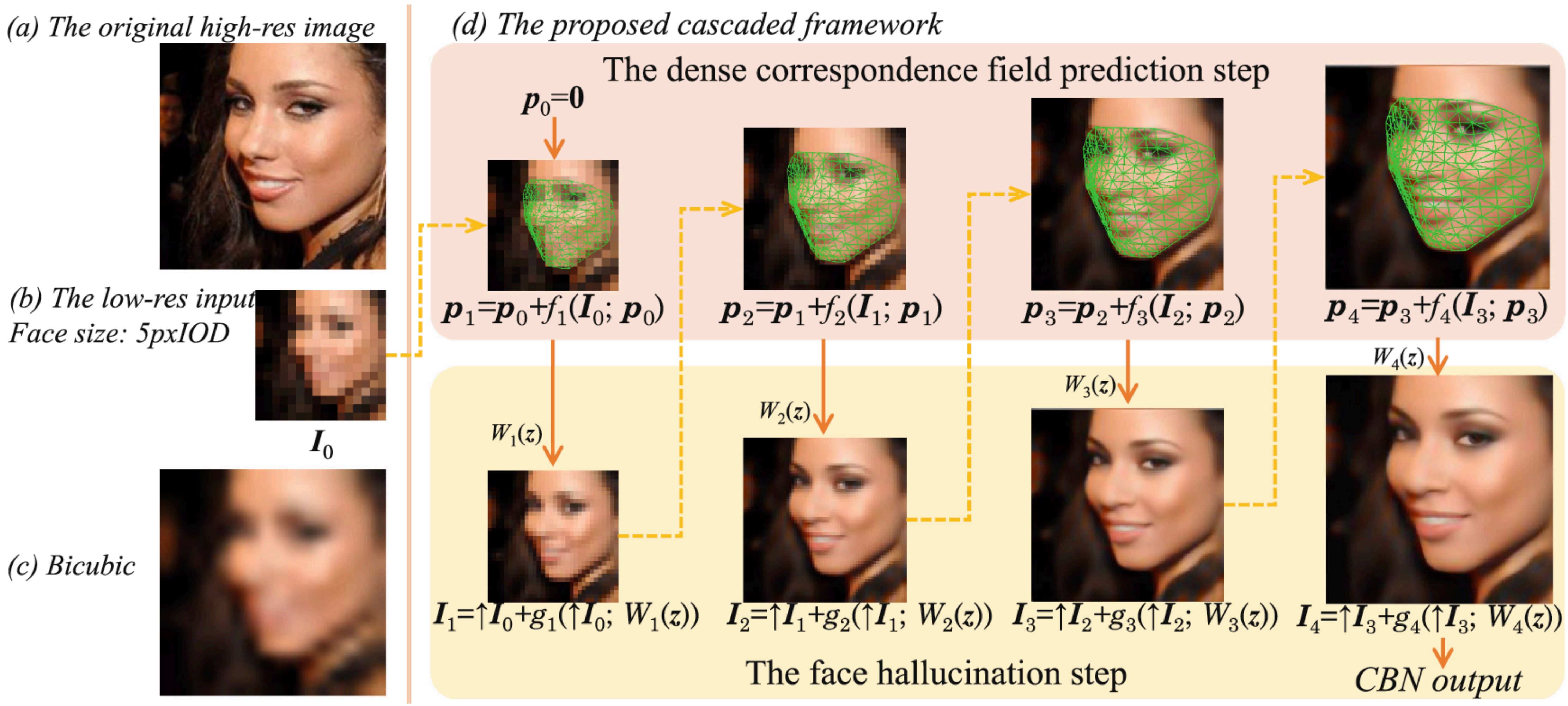
Applications



Medical Imaging
(Optical Coherence Tomography (OCT) scan)

Source: Yamashita, K.; Markov, K. Medical Image Enhancement Using Super Resolution Methods. In Proceedings of the Computational Science—ICCS, Amsterdam, The Netherlands, 3–5 June 2020; Krzhizhanovskaya, V.V., Závodszky, G., Lees, M.H., Dongarra, J.J., Sloot, P.M.A., Brissos, S., Teixeira, J., Eds.; Springer: Cham, Switzerland, 2020; pp. 496–508.

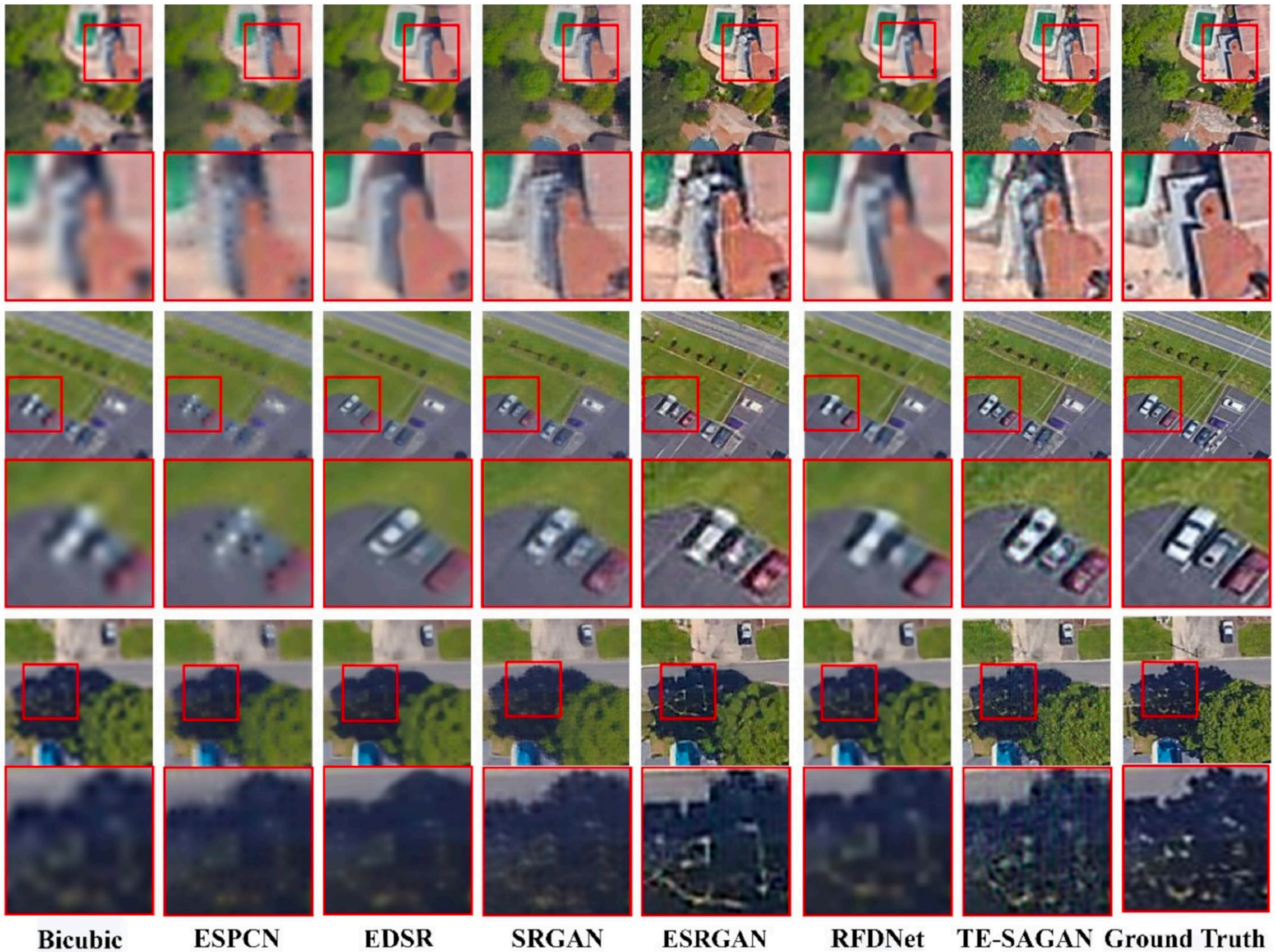
Applications



Security via Person Identification

Source: Zhu, S.; Liu, S.; Loy, C.C.; Tang, X. Deep Cascaded Bi-Network for Face Hallucination. In Proceedings of the Computer Vision—ECCV 2016, Amsterdam, The Netherlands, 11–14 October 2016; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Springer: Cham, Switzerland, 2016; pp. 614–630.

Applications



Remote Sensing*

Source: Xu, Y.; Luo, W.; Hu, A.; Xie, Z.; Xie, X.; Tao, L. TE-SAGAN: An Improved Generative Adversarial Network for Remote Sensing Super-Resolution Images. Remote Sens. 2022, 14, 2425.

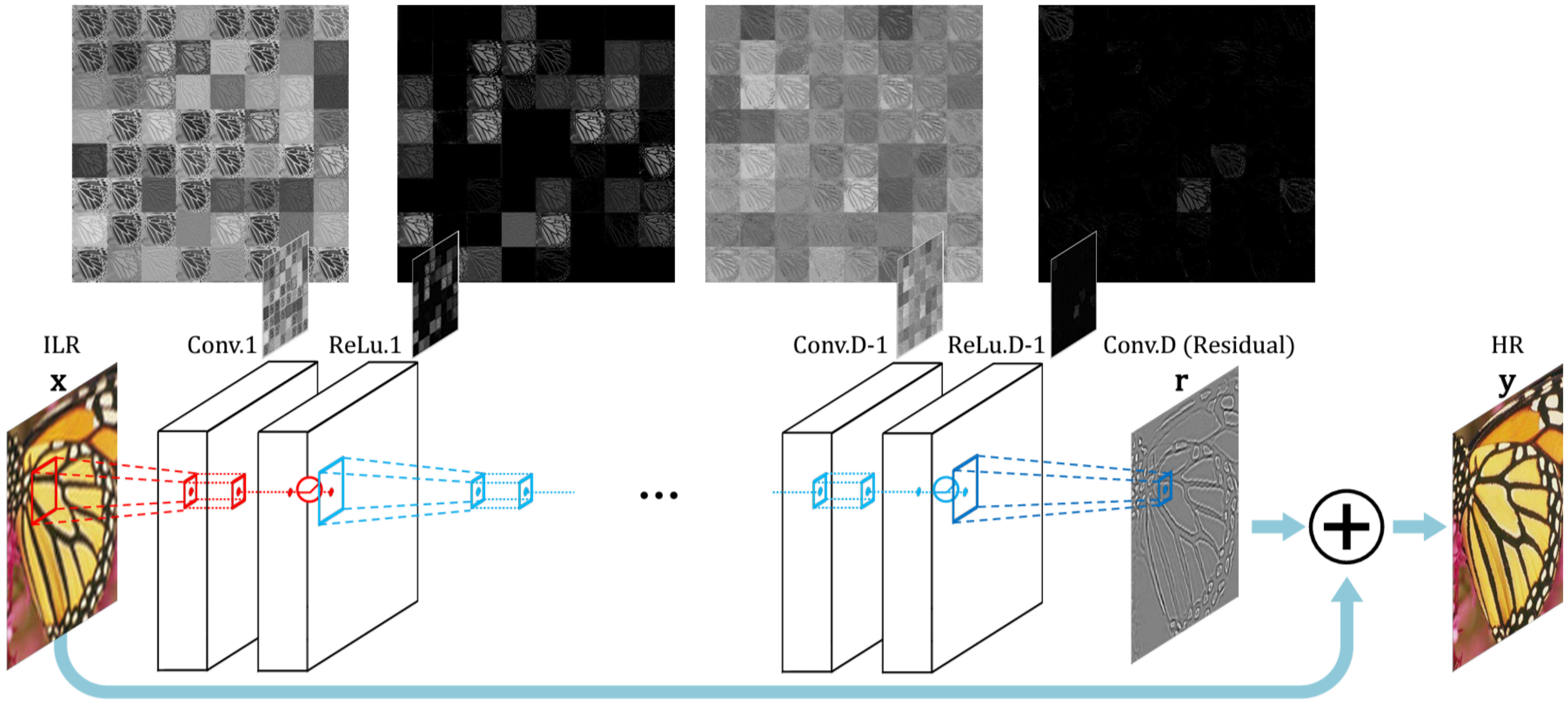
Methods for Image SR

- Classical: bicubic interpolation and Lanczos resampling , edge-based methods, …

- Deep learning (DL):
  - Convolutional neural networks (CNNs);
  - Generative adversarial networks (GANs);
  - Attention-based networks.

Supervised Image SR

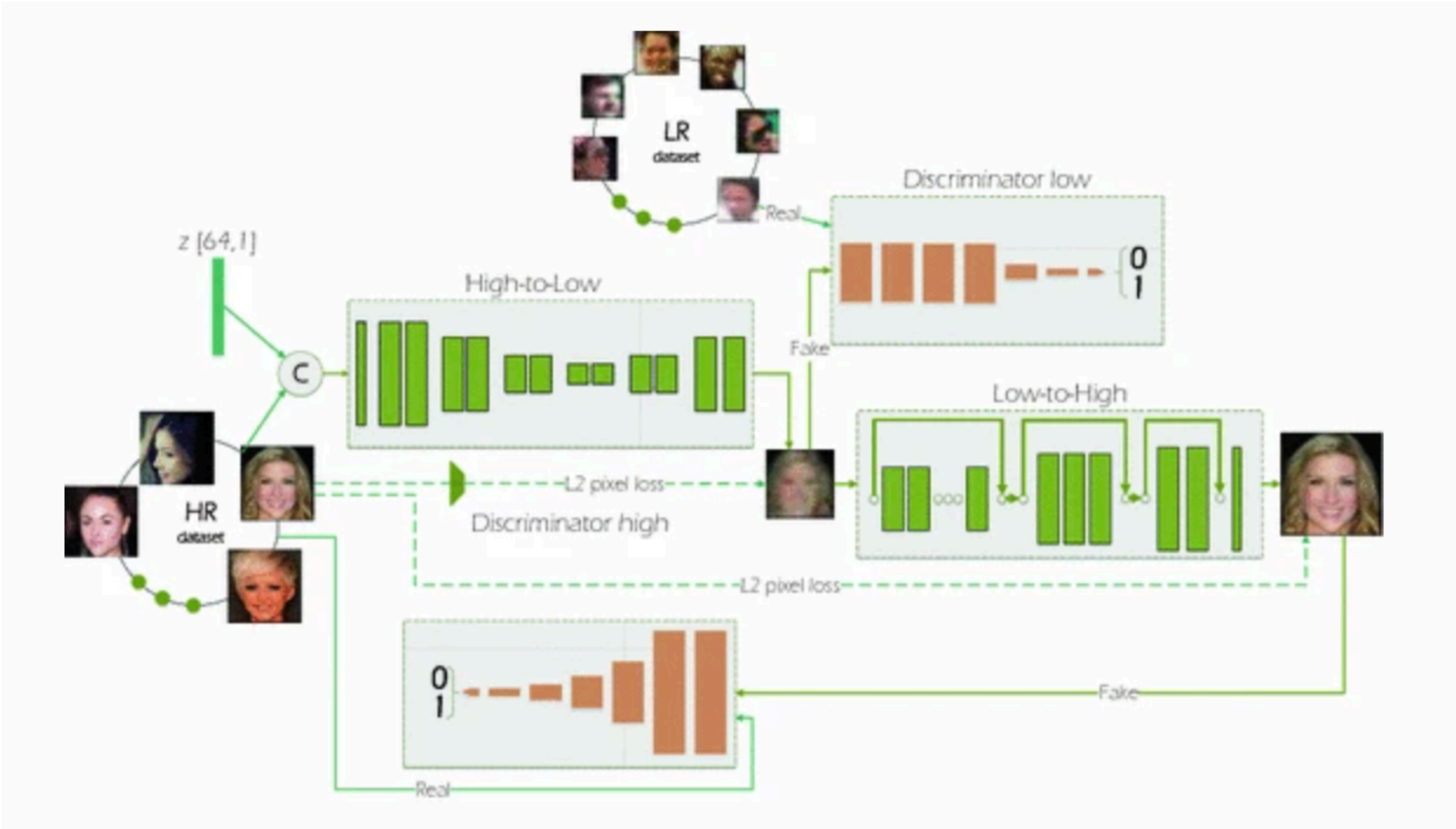
- Models: trained with both LR images and the corresponding HR ones.



Source: J. Kim, J. Kwon Lee, and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 1646–1654.

Unsupervised Image SR

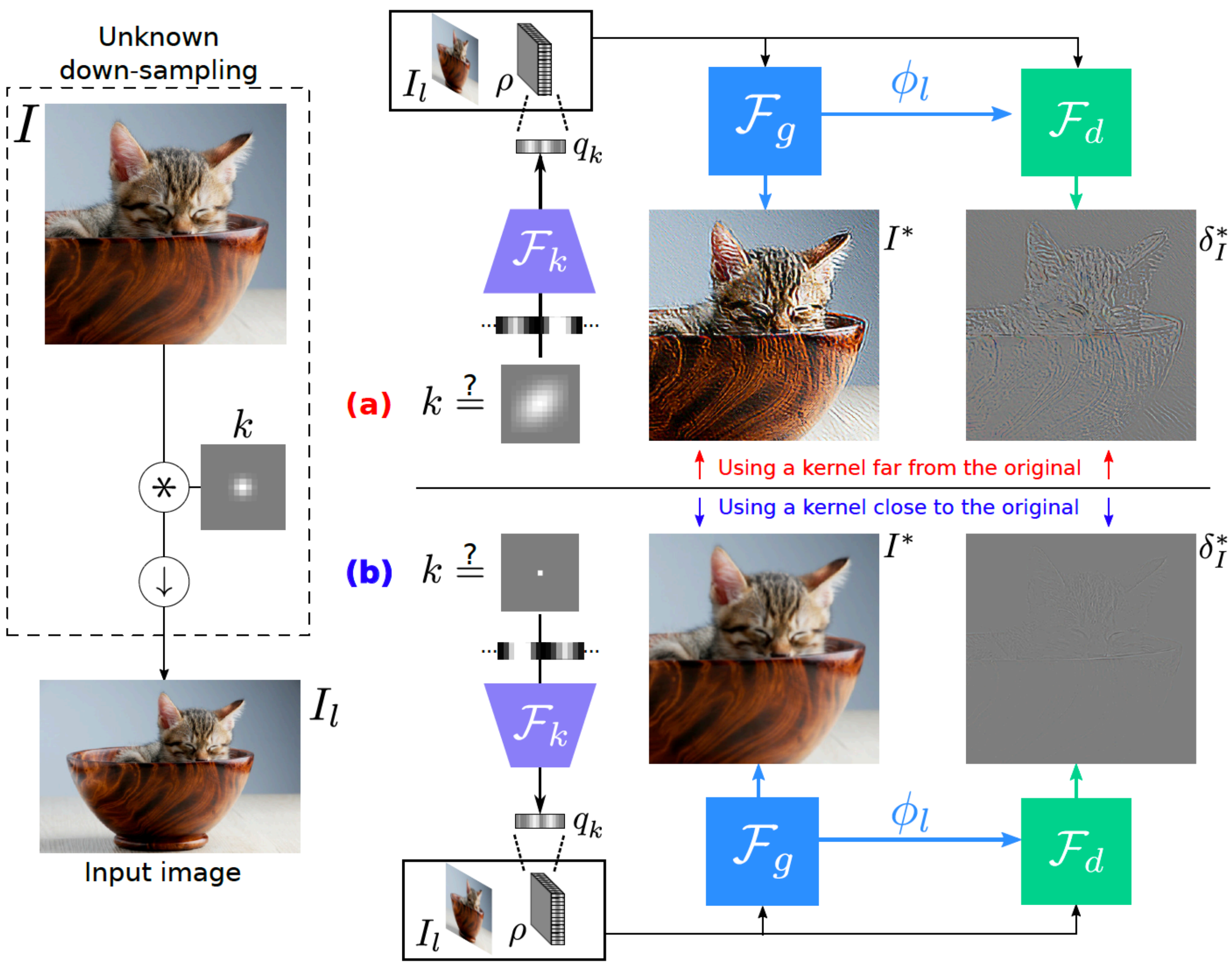
- Models: only unpaired LR-HR images are available for training.



Source: Bulat, A.; Yang, J.; Tzimiropoulos, G. To Learn Image Super-Resolution, Use a GAN to Learn How to Do Image Degradation First. In Proceedings of the Computer Vision—ECCV 2018, Munich, Germany, 8–14 September 2018; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Springer: Cham, Switzerland, 2018; pp. 187–202..

Blind Image SR

- The degradation process/kernels is/are unknown. Techniques in this context rely on LR images.



Source: Cornillère, V.; Djelouah, A.; Yifan,W.; Sorkine-Hornung, O.; Schroers, C. Blind Image Super-Resolution with Spatially Variant Degradations. ACM Trans. Graph. 2019, 38, 1–13.

Evaluating DL Tecnhiques for Image SR

- There are many techniques and experiments proposed boosted by DL techniques.

- But … studies usually do not consider high scaling factors, capping it at 2x or 4x.

Evaluating DL Tecnhiques for Image SR

- When there are exceptions: no significant diversity of images and feature spaces.
- Interesting to consider quite distinct broader domains:
  - Medical images;
  - Images obtained by satellites via sensors with different characteristics;
  - Images more "usual" like those of animal's faces.

This Study

- A **high-scale (8x)** controlled experiment which evaluates five recent **DL techniques** tailored for blind image SR:
  - Adaptive Pseudo Augmentation (APA),
  - Blind Image SR with Spatially Variant Degradations (BlindSR);
  - Deep Alternating Network (DAN);
  - FastGAN;
  - Mixture of Experts Super-Resolution (MoESR).

This Study

- LR = 128 × 128 pixels; HR = 1024 × 1024 pixels.
- **Single-image SR**: BlindSR, DAN, and MoESR;
- **Non-single but few-shot image SR**: APA and FastGAN.

This Study

- Relying basically on public sources, **14 LR image datasets** (100 samples each) from five different broader domains:
  - Aerial;
  - Fauna;
  - Flora;
  - Medical;
  - Satellite (Space).
- See: https://www.kaggle.com/datasets/valdivinosantiago/dl-blindsr-datasets

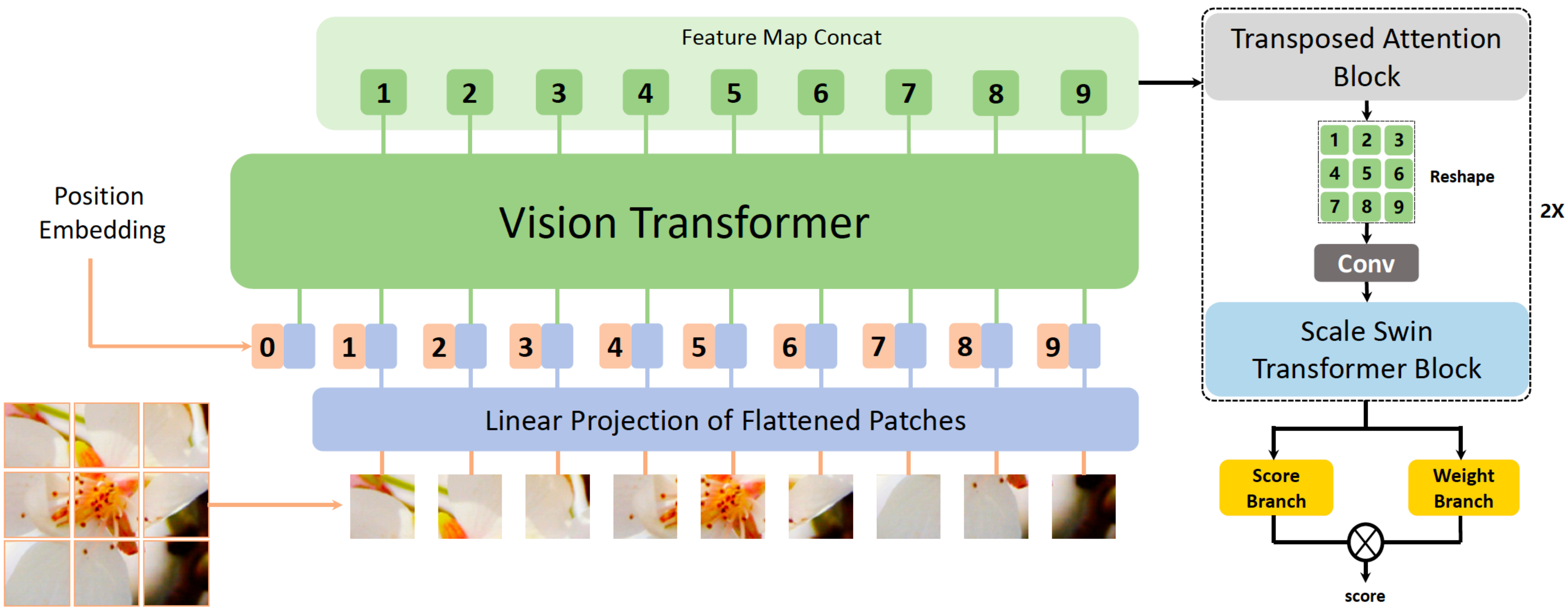
This Study

- No-reference image quality assessment (NR-IQA): image quality without a reference image (perceptual quality).

- Selected NR-IQA metrics:
  - Classical natural image quality evaluator (**NIQE**);
  - Vision transformer(ViT)-based multi-dimension attention network for no-reference image quality assessment (**MANIQA**) score.

This Study

❖ MANIQA model.



Source: Yang, S.; Wu, T.; Shi, S.; Lao, S.; Gong, Y.; Cao, M.; Wang, J.; Yang, Y. MANIQA: Multi-dimension Attention Network for No-Reference Image Quality Assessment. arXiv 2022, arXiv:2204.08958.

Project IDeepS

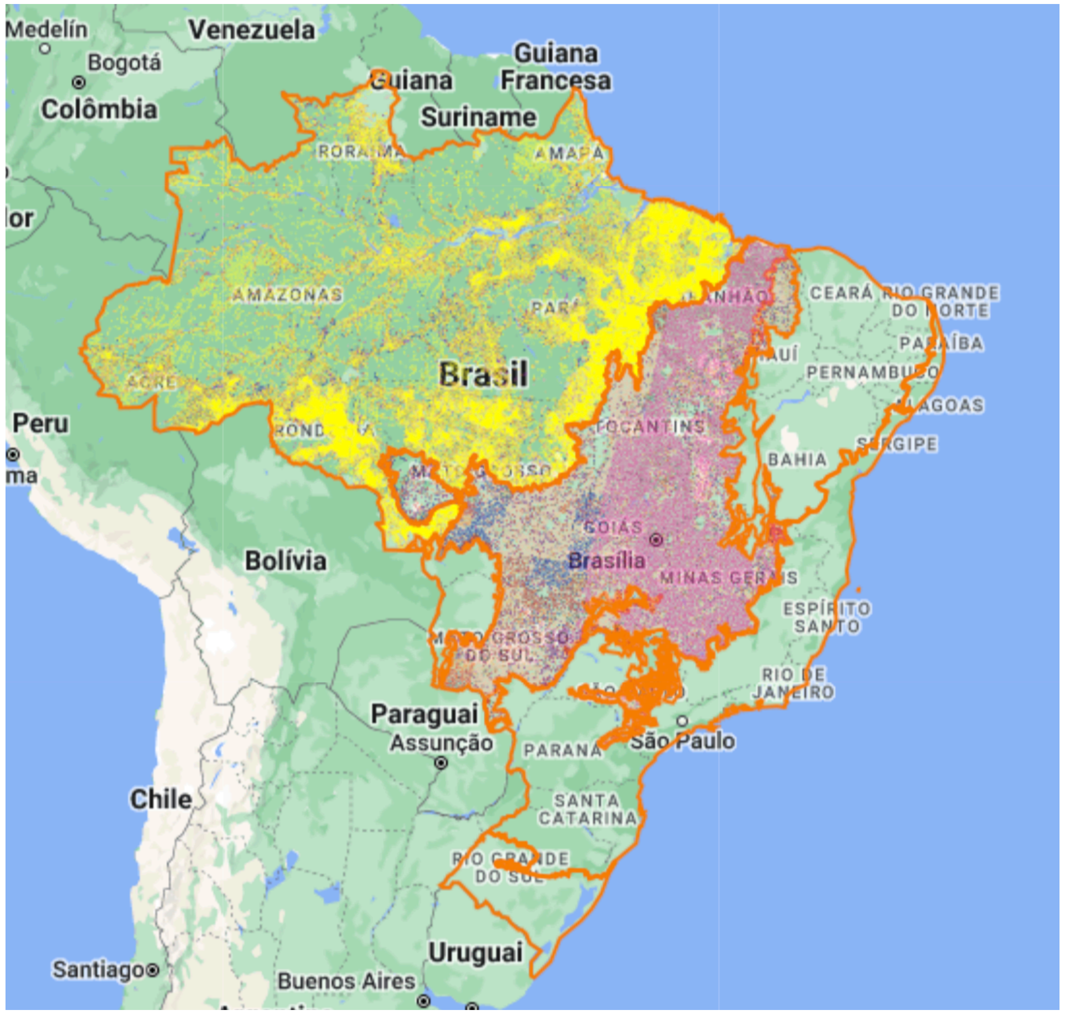
- Classificação de imagens via redes neurais profundas e grandes bases de dados para aplicações aeroespaciais.

Project IDeepS

Source: https://github.com/vsantjr/IDeepS

IDeepS: Objective 1

- Large-scale investigation, deep neural networks (DNNs), satellite image classification.



IDeepS: Objective 2

- Best DNNs, drones, autonomy.

IDeepS: Higher Objective

Recommendations/Suggestions

Remote Sensing

Drones

Research Questions (RQs)

- RQ_1—Which out of the five algorithms for blind image SR is the best regarding the metrics NIQE and MANIQA score? And which can be considered the best overall?

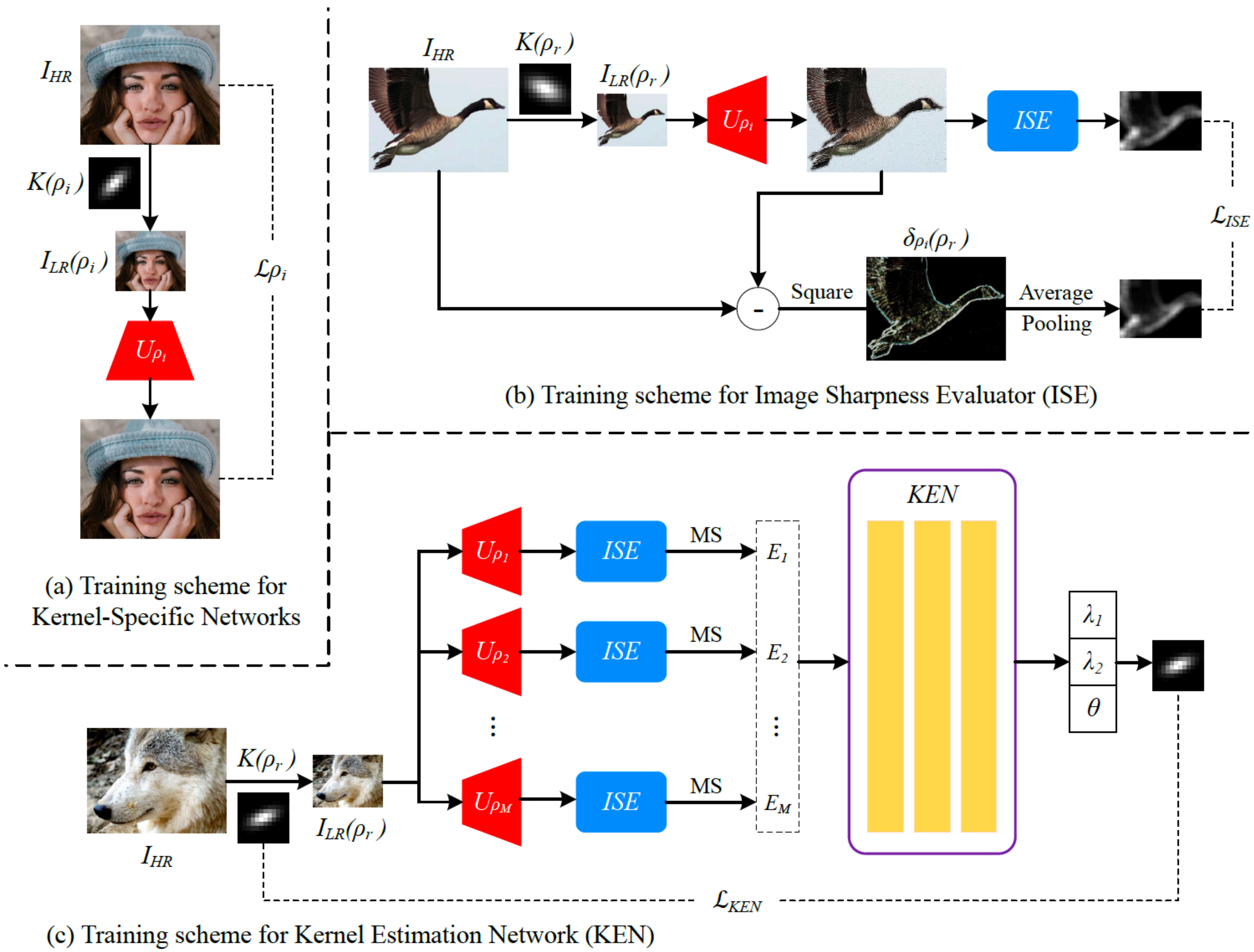
- RQ_2—Does the two top approaches present similar behaviours when deriving HR images?

Datasets: Description

Domain	Dataset	Description
Aerial	`condoaerial`	Aerial Semantic Segmentation Drone Dataset
	`massachbuildings`	Massachusetts Buildings Dataset
	`ships`	Ship Detection from Aerial Images Dataset
	`ufsm-flame`	Drone Images from UFSM and Flame Datasets
Fauna	`catsfaces`	Cats Faces Dataset
	`dogsfaces`	Dogs Faces Dataset
Flora	`flowers`	102 Category Flower Dataset
	`plantpat`	Plant Pathology 2021-FGVC8-Dataset
Medical	`melanomaisic`	SIIM-ISIC Melanoma Classification Dataset
	`structretina`	Structured Analysis of the Retina Dataset
Satellite	`amazonia1`	Cloudless Scene from Amazonia 1 Satellite Dataset
	`cbers4a`	Scene with Clouds from CBERS-4A Satellite Dataset
	`deepglobe`	Forest Aerial Images for Segmentation Dataset
	`isaid`	Instance Segmentation in Aerial Images Dataset

Datasets: Samples

Some DL Techniques: MoESR



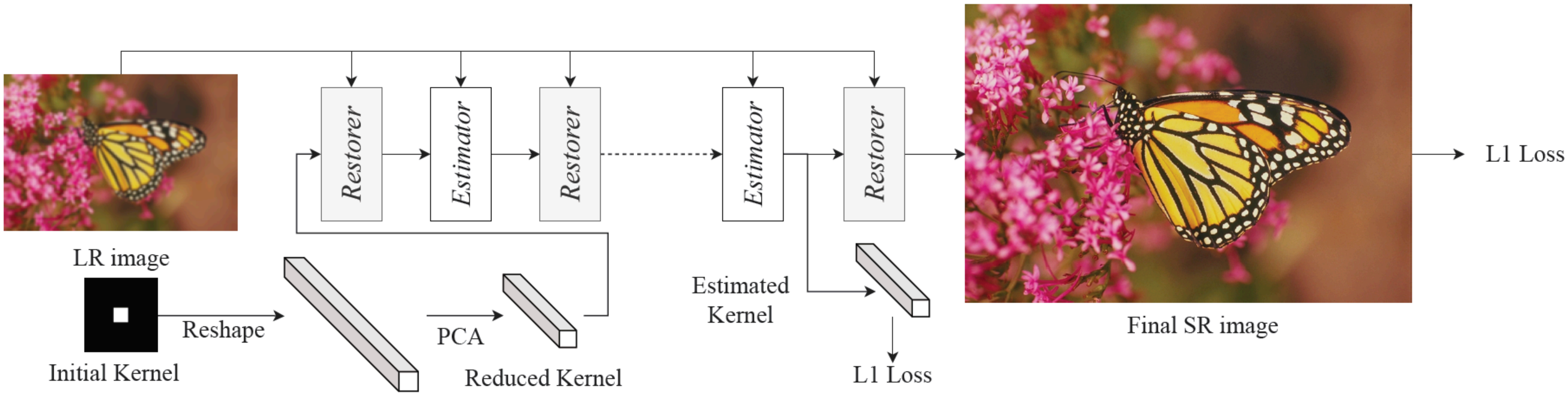
- Different experts for different degradation kernels.

- MoESR predicts the degradation kernel and super-resolve the LR image using the most adequate kernel-specific expert.

- Image Sharpness Evaluator (ISE) assesses the sharpness of the images generated by the experts.

- These evaluations are used by the Kernel Estimation Network (KEN) to estimate the kernel and select the best pretrained expert network.

Some DL Techniques: DAN



- Alternating optimisation algorithm which restores an HR image and estimates the corresponding blur kernel alternately.

- The Restorer convolutional neural module restores an HR image based on the predicted Estimator's kernel, and the Estimator convolutional neural module estimates a blur kernel with the help of the restored HR image.

Runnings

- Bull Sequana X1120 computing node of the SDumont supercomputer.
  - 4x NVIDIA Volta V100 graphics processing units (GPUs).

- Each run: 4 days, being considered the latest model when the execution exceeded this time.

Results: NIQE (↓)

- Mean NIQE values: broader domains.

Domain	Min (Best)		Max (Worst)	
	Technique	NIQE	Technique	NIQE
Aerial	MoESR	14.648013	BlindSR	22.534419
Fauna	MoESR	14.629656	BlindSR	18.961343
Flora	MoESR	14.615975	BlindSR	21.668490
Medical	APA	14.520624	BlindSR	18.717661
Satellite	MoESR	14.385091	BlindSR	24.726333

Best: MoESR, APA.
Worst: BlindSR.

Results: NIQE (↓)

- Mean NIQE values: datasets.

Domain	Dataset	Min (Best)		Max (Worst)	
Aerial	condoaerial	MoESR	14.648013	BlindSR	20.611777
	massachbuildings	MoESR	16.427236	BlindSR	22.199742
	ships	APA	17.688239	BlindSR	22.534419
	ufsm-flame	APA	15.492423	BlindSR	21.684649
Fauna	catsfaces	MoESR	14.629656	BlindSR	18.702201
	dogsfaces	MoESR	15.314877	BlindSR	18.961343
Flora	flowers	MoESR	14.615974	BlindSR	18.174120
	plantpat	APA	15.241252	BlindSR	21.668490
Medical	melanomaisic	APA	15.025945	BlindSR	18.717661
	structretina	APA	14.520624	BlindSR	16.517888
Satellite	amazonia1	APA	16.027069	BlindSR	23.698020
	cbers4a	APA	16.398990	DAN	17.438634
	deepglobe	APA	16.863639	BlindSR	24.726333
	isaid	MoESR	14.385091	BlindSR	21.504252

Best: APA, MoESR.
Worst: BlindSR.

Results: NIQE (↓)

- Improvement metric: $I\% = \dfrac{(W - B) \times 100}{B}$

- Improvement of MoESR over APA.

Dataset	MoESR	APA	$I\%$
`condoaerial`	14.648013	15.292783	4.402
`massachbuildings`	16.427236	16.760828	2.031
`catsfaces`	14.629656	15.576498	**6.472**
`dogsfaces`	15.314877	15.444758	0.848
`flowers`	14.615974	14.787114	1.171
`isaid`	14.385091	15.01201	4.358
$\overline{I\%}$			3.214

Results: NIQE (↓)

- Improvement of APA over MoESR.

Dataset	APA	MoESR	$I\%$
ships	17.688239	18.220023	3.006
ufsm-flame	15.492423	15.614556	0.788
plantpat	15.241252	15.80596	3.705
melanomaisic	15.025945	16.902496	**12.489**
structretina	14.520624	15.914241	9.598
amazonia1	16.027069	17.041395	6.329
cbers4a	16.39899	17.037916	3.896
deepglobe	16.863639	17.43862	3.410
$\overline{I\%}$			5.403

- Conclusions: APA was the best followed by MoESR. BlindSR was the worst.

Results: MANIQA (↑)

- Mean MANIQA scores: broader domains.

Domain	Max (Best)		Min (Worst)	
	Technique	**MANIQA**	**Technique**	**MANIQA**
Aerial	DAN	0.696858	FastGAN	0.409388
Fauna	MoESR	0.713253	FastGAN	0.515693
Flora	MoESR	0.698373	APA	0.430053
Medical	MoESR	0.614705	APA	0.432007
Satellite	DAN	0.736443	FastGAN	0.327089

Best: MoESR, DAN.
Worst: FastGAN, APA.

Results: MANIQA (↑)

- Mean MANIQA scores: datasets.

Domain	Dataset	Max (Best)		Min (Worst)	
Aerial	condoaerial	MoESR	0.657257	FastGAN	0.409388
	massachbuildings	DAN	0.696858	FastGAN	0.525381
	ships	DAN	0.577708	FastGAN	0.492311
	ufsm-flame	DAN	0.618042	FastGAN	0.493545
Fauna	catsfaces	MoESR	0.713253	FastGAN	0.611105
	dogsfaces	MoESR	0.638982	FastGAN	0.515693
Flora	flowers	MoESR	0.698373	FastGAN	0.533008
	plantpat	MoESR	0.606683	APA	0.430053
Medical	melanomaisic	DAN	0.542052	FastGAN	0.449874
	structretina	MoESR	0.614705	APA	0.432007
Satellite	amazonia1	MoESR	0.579970	FastGAN	0.417986
	cbers4a	MoESR	0.408773	FastGAN	0.327089
	deepglobe	MoESR	0.591723	APA	0.330667
	isaid	DAN	0.736443	FastGAN	0.445917

Best: MoESR, DAN.
Worst: FastGAN, APA.

Results: MANIQA (↑)

- Improvement metric: $I\% = \dfrac{(B - W) \times 100}{W}$

- Improvement of DAN over MoESR.

Dataset	**DAN**	**MoESR**	$\boldsymbol{I\%}$
`massachbuildings`	0.696858	0.674955	**3.245**
`ships`	0.577708	0.571233	1.134
`ufsm-flame`	0.618042	0.608577	1.555
`melanomaisic`	0.542052	0.541265	0.145
`isaid`	0.736443	0.72679	1.328
$\overline{I\%}$			1.481

Results: MANIQA (↑)

- Improvement of MoESR over DAN.

Dataset	MoESR	DAN	$I\%$
`condoaerial`	0.657257	0.653125	0.633
`catsfaces`	0.713253	0.681594	4.645
`dogsfaces`	0.638982	0.605406	5.546
`flowers`	0.698373	0.674927	3.474
`plantpat`	0.606683	0.578225	4.922
`structretina`	0.614705	0.609501	0.854
`amazonia1`	0.57997	0.546786	**6.069**
`cbers4a`	0.408773	0.403834	1.223
`deepglobe`	0.591723	0.580666	1.904
$\overline{I\%}$			3.252

- Conclusions: MoESR was the best followed by DAN. FastGAN was the worst and APA got the penultimate place.

Answering RQ_1

- RQ_1—Which out of the five algorithms for blind image SR is the best regarding the metrics NIQE and MANIQA score? And which can be considered the best overall?

- R: **Considering both metrics, NIQE and MANIQA score, we can state that MoESR was the most outstanding approach. Note that we saw contradictory performances regarding APA where it was the best strategy evaluated via NIQE and almost the worst approach, if we take into account the MANIQA score.**

Results: Behaviours

- Take the 10 HR images with the **best (highest)** MANIQA scores from MoESR and DAN.

- $|C_B|$ = cardinality, set of common best images.
- $|N_B|$ = cardinality, set of non-common best images.
- Ex: $|C_B(condoaerial)| = 7$; $|N_B(condoaerial)| = 3$

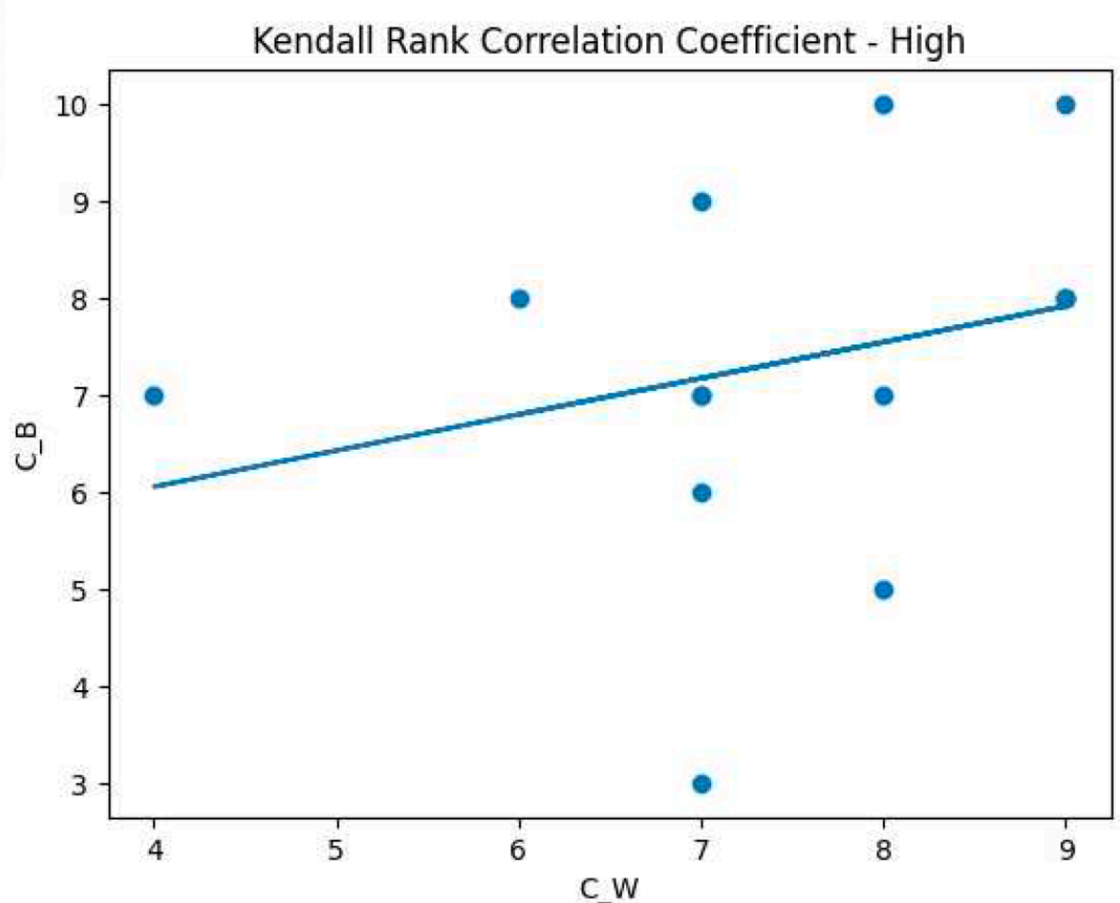
Results: Behaviours

- Take the 10 HR images with the **worst (lowest)** MANIQA scores from MoESR and DAN.

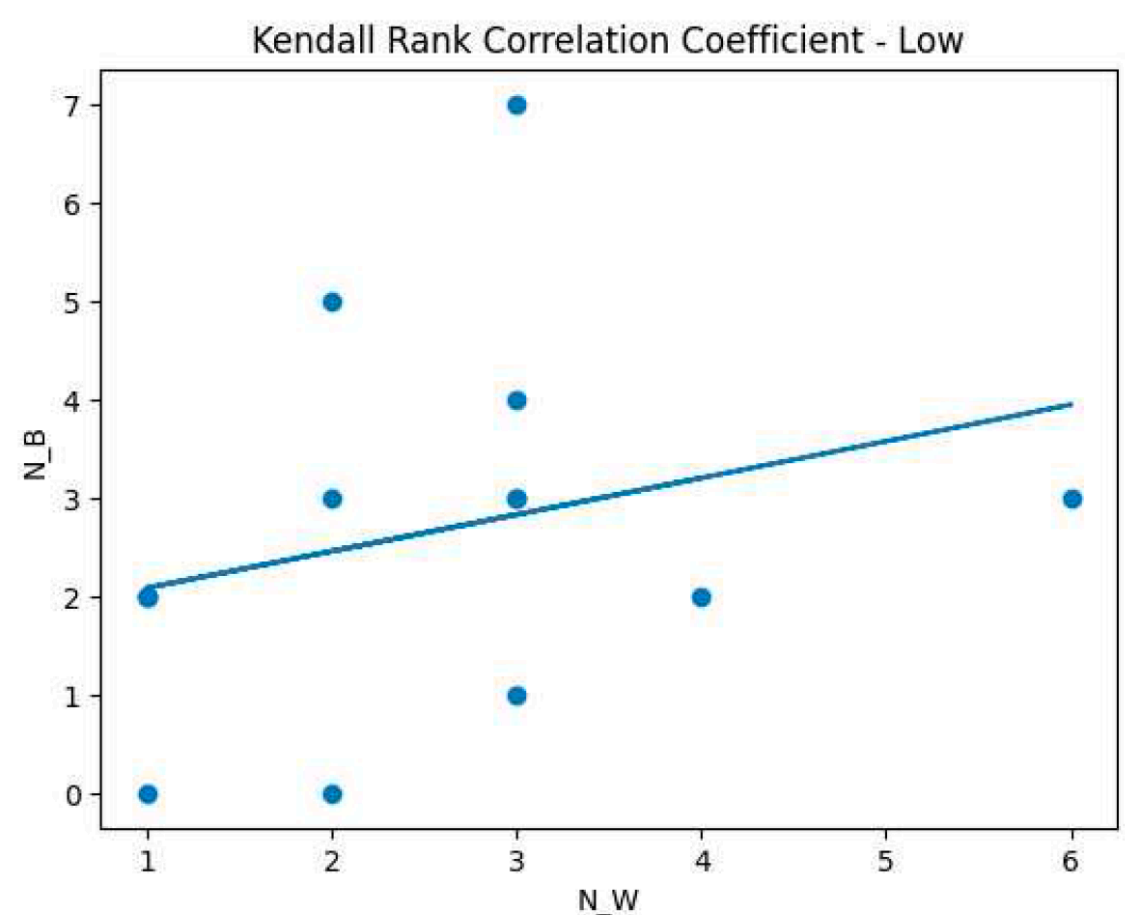
- $|C_W|$ = cardinality, set of common worst images.
- $|N_W|$ = cardinality, set of non-common worst images.
- Ex: $|C_W(amazonia1)| = 6$; $|N_W(amazonia1)| = 4$

Results: Behaviours

- Kendall's τ coefficient: $|C_W| \times |C_B|$; $|N_W| \times |N_B|$.
- Both cases: $\tau = 0.306912$ (Good correlation).



(**a**) Correlation of C sets



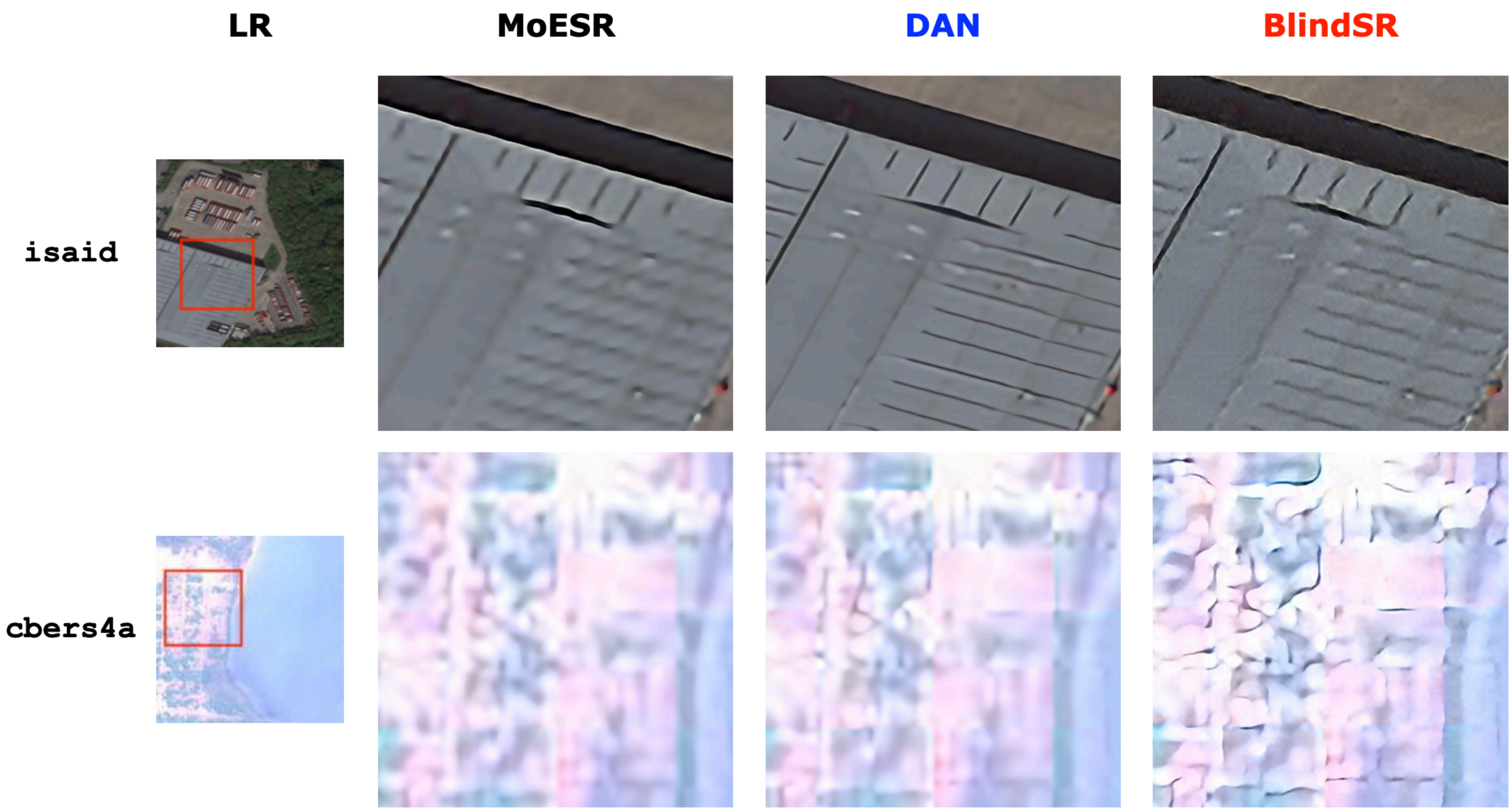
(**b**) Correlation of N sets

Answering RQ_2

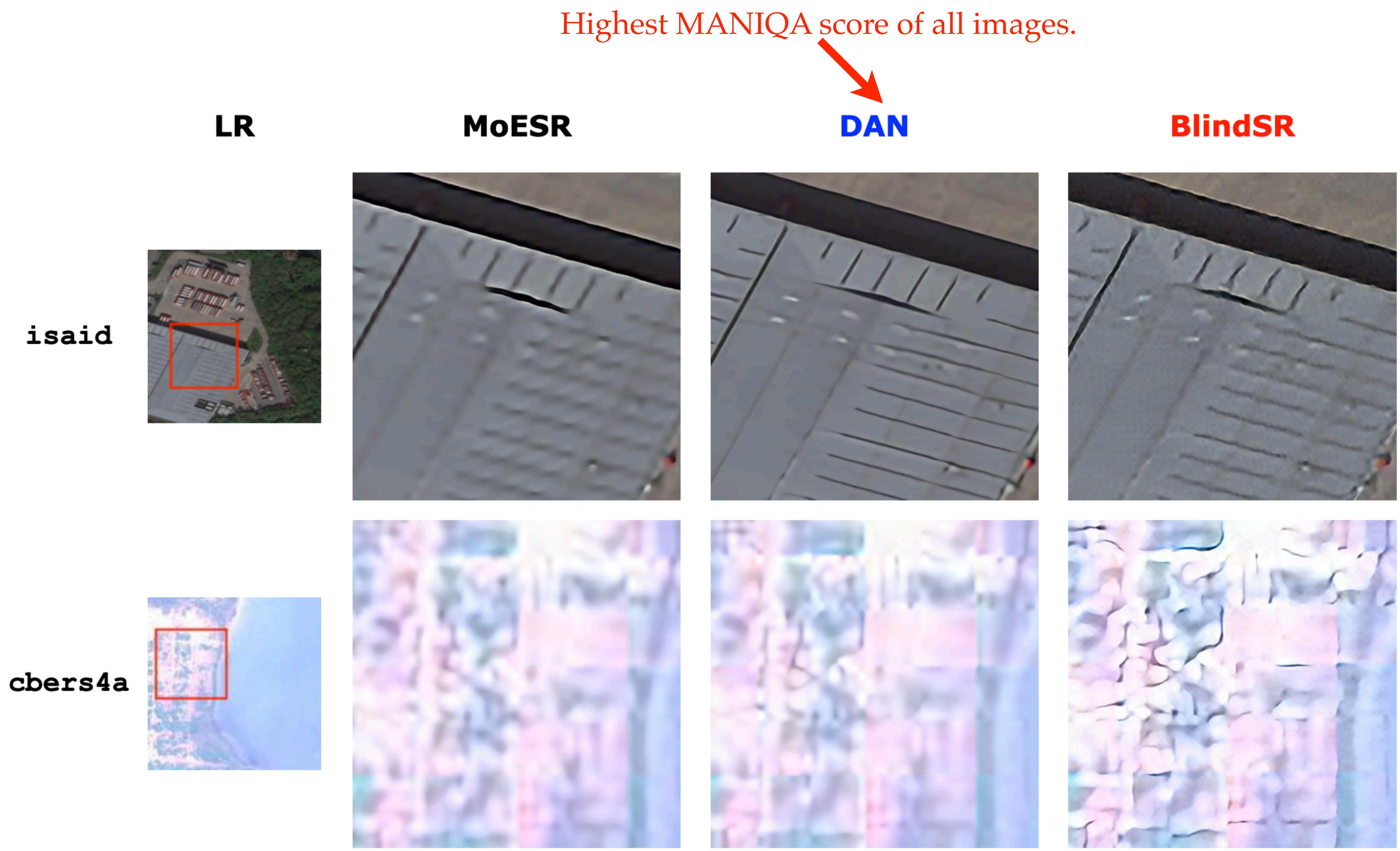
- RQ_2—Does the two top approaches present similar behaviours when deriving HR images?

- R: **The interpretation of the results is that the images detected as having the best, as well as the worst, perceptual qualities, based on the MANIQA scores, are somewhat "common" to both techniques. Hence, we can conclude that both approaches (MoESR and DAN) present similar behaviours.**

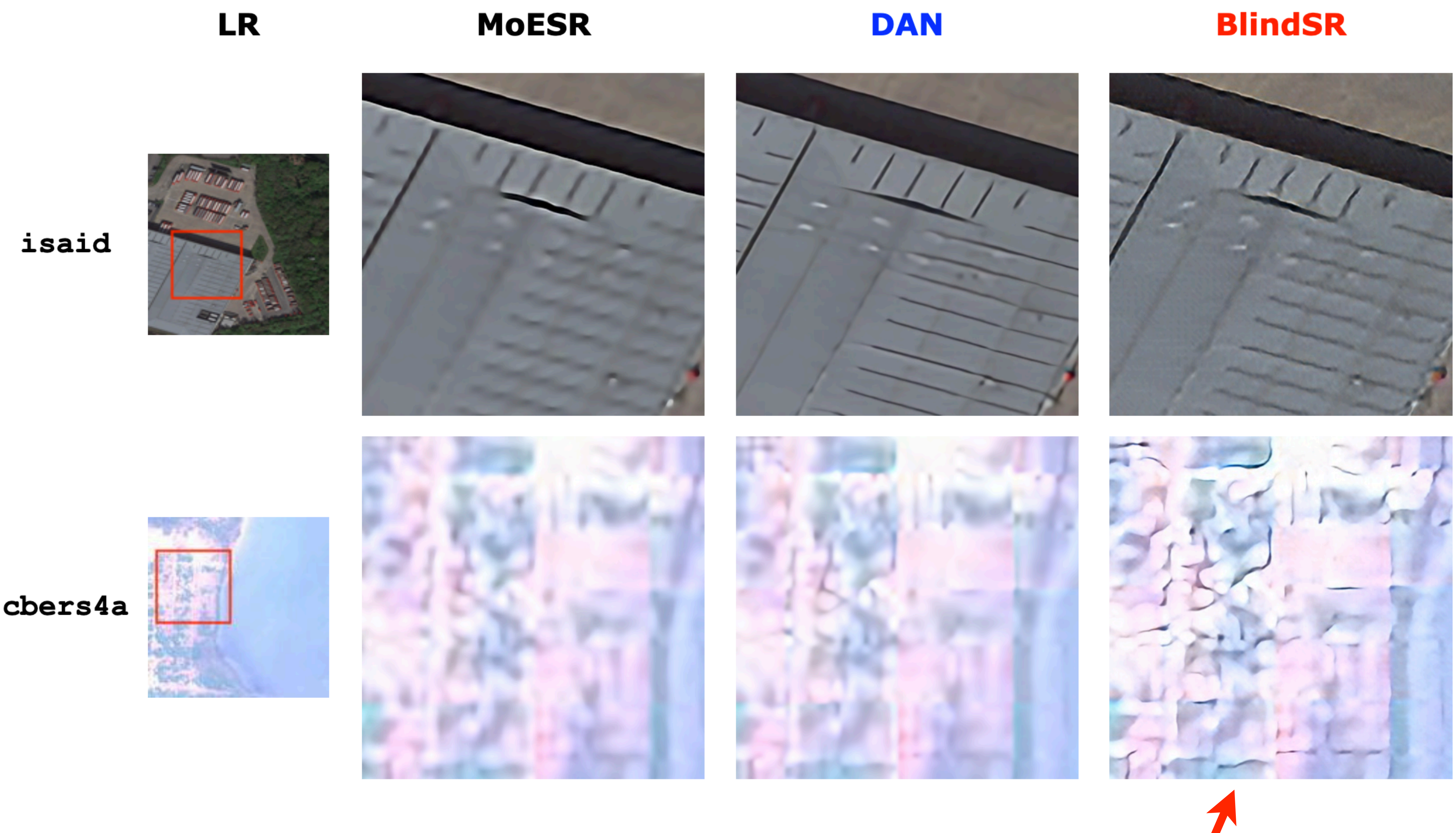
Visual Analysis



Visual Analysis

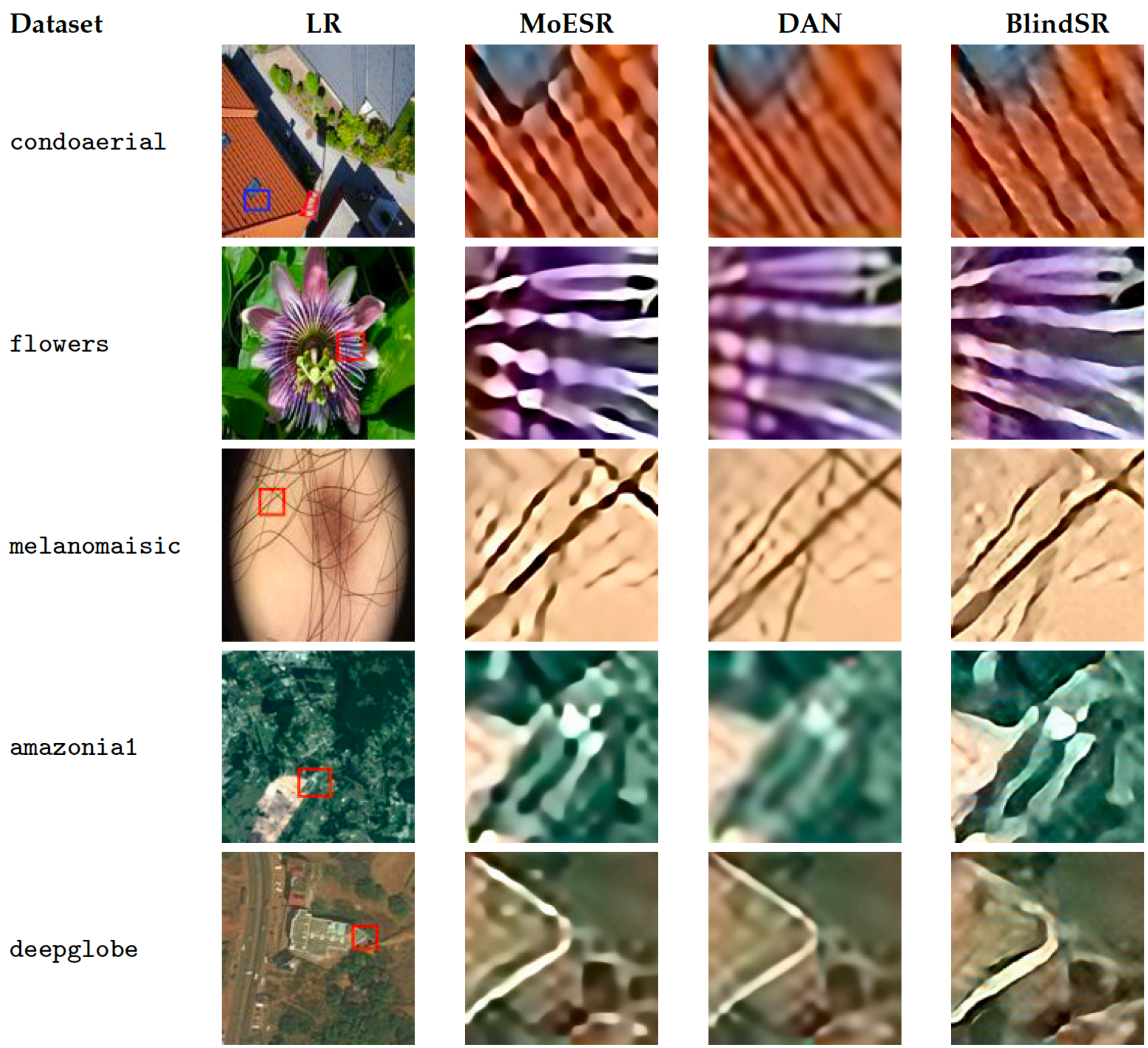


Visual Analysis



Lowest MANIQA score of all images.

Visual Analysis

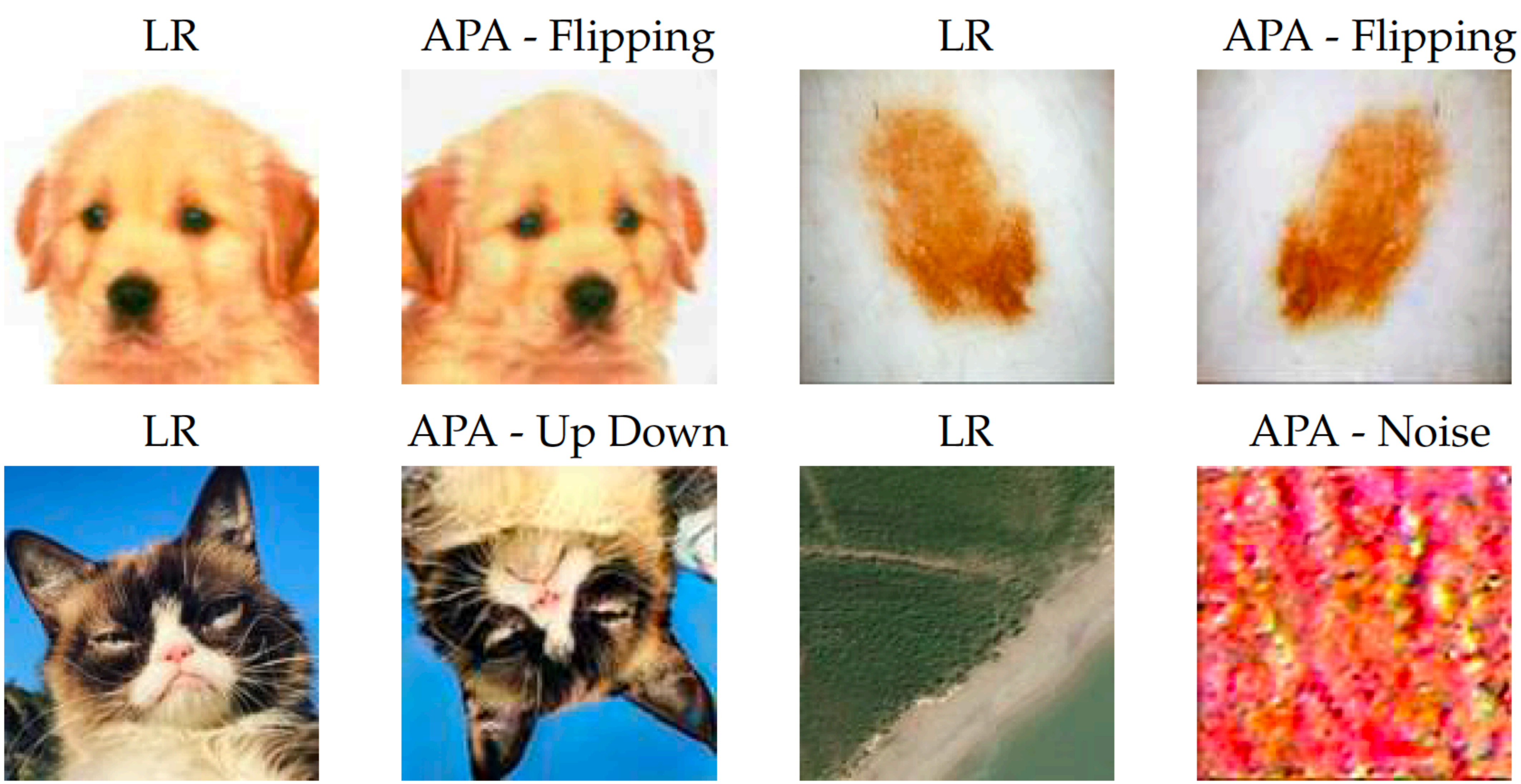


GAN-based: Issues

- APA in a custom dataset: prepare the dataset, training, and inference for generating images.
- Training was not completed, even using 4x NVIDIA Volta V100 GPUs for 4 days. Datasets are very small.
- Thus, APA is a very "heavy" model.

GAN-based: Issues

- APA: other issues.

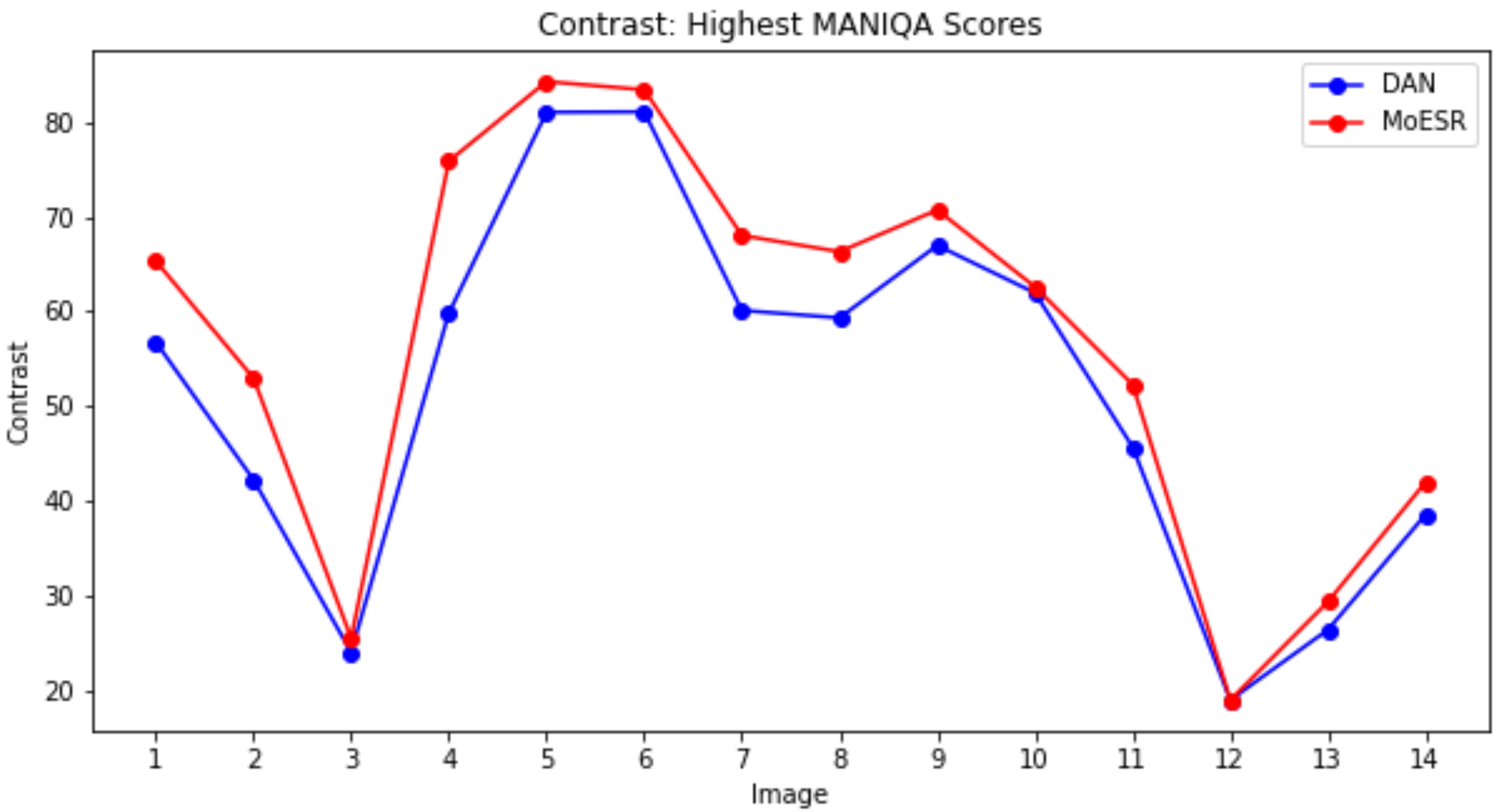


GAN-based: Issues

- FastGAN in a custom dataset: training (considerably faster than APA) and inference.
- FastGAN issue: mode collapse (also in APA).

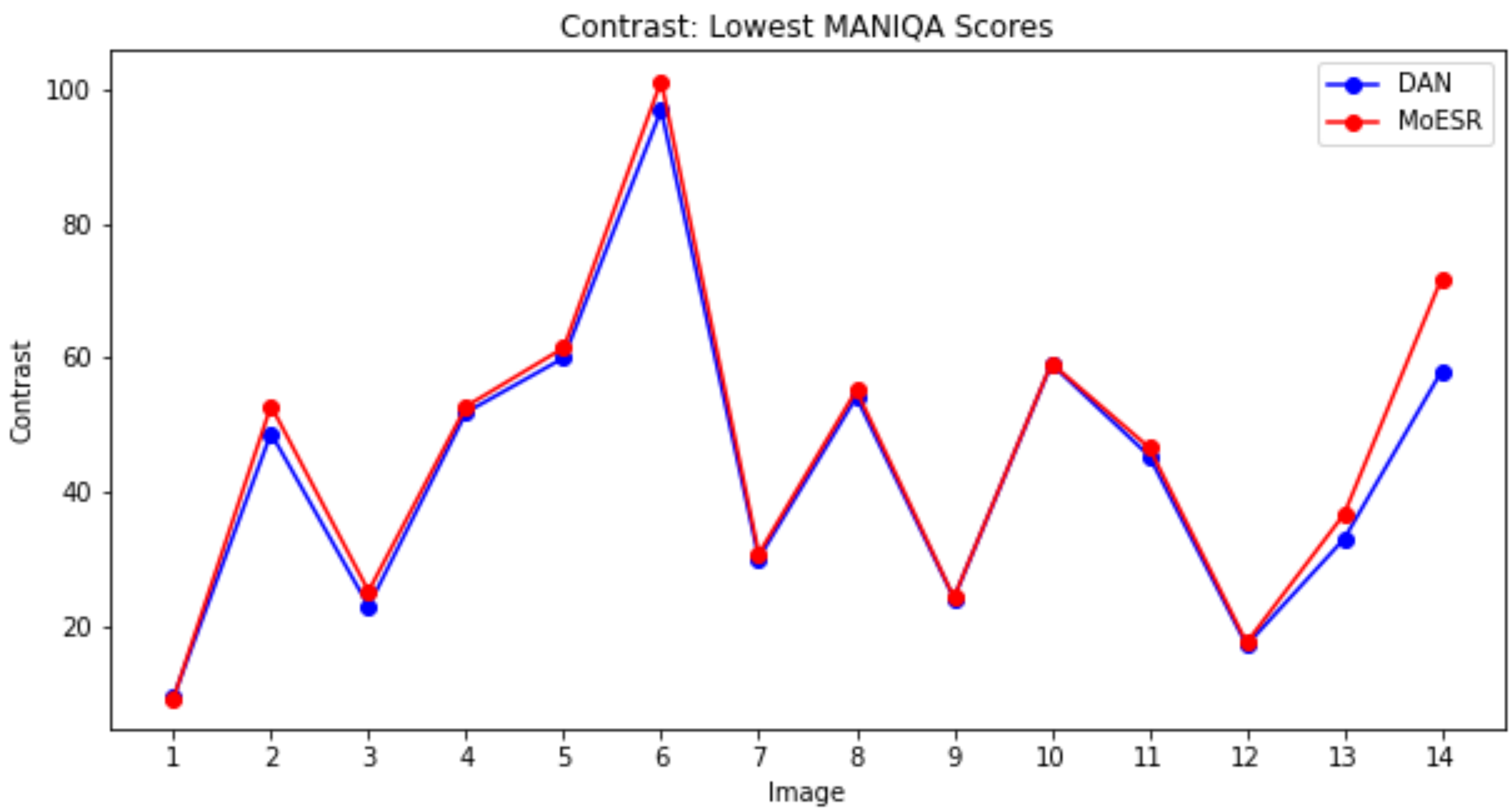
Some Explainability

- MoESR: sharper HR images than the ones of DAN.

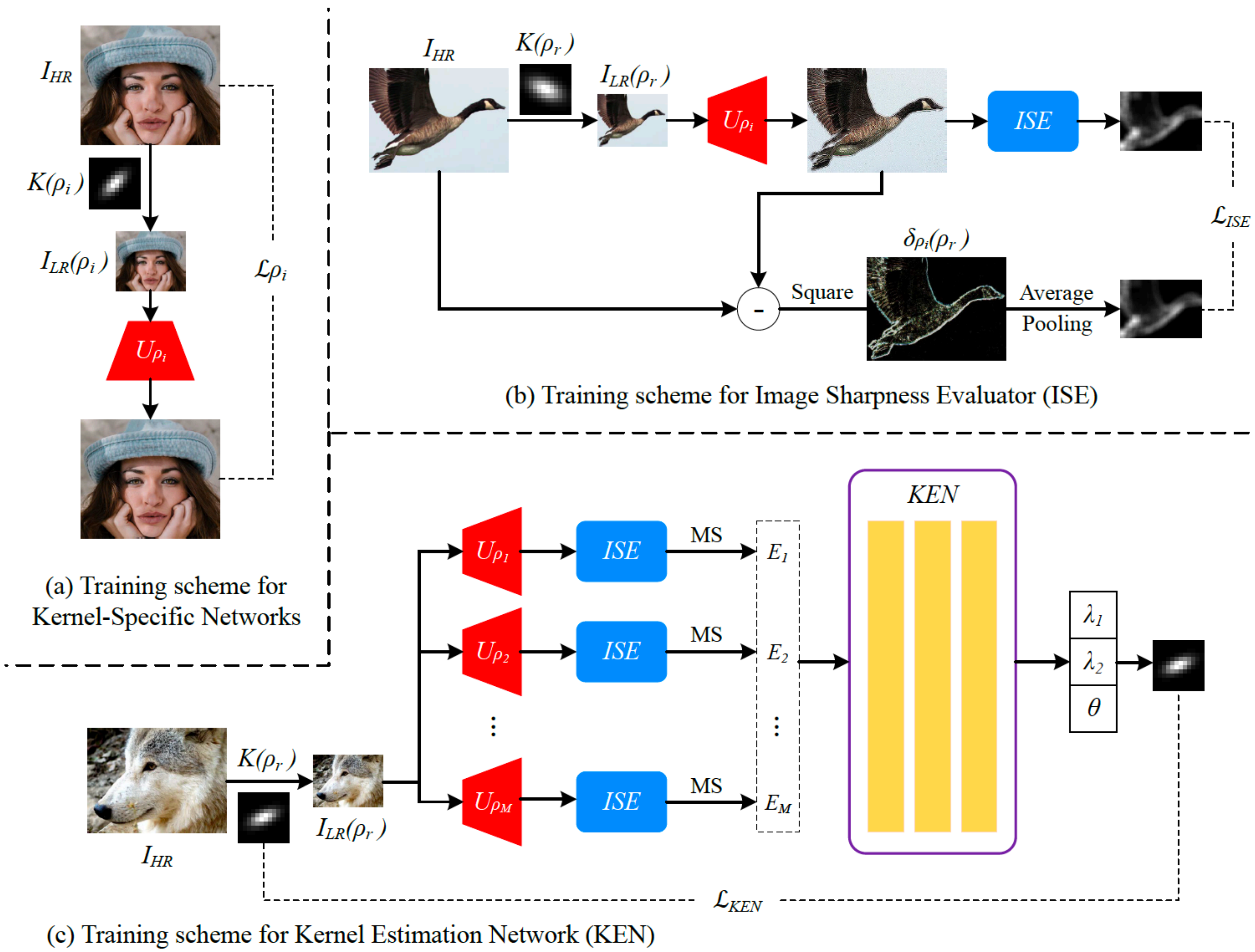


Some Explainability

- MoESR: sharper HR images than the ones of DAN.



Some Explainability: MoESR



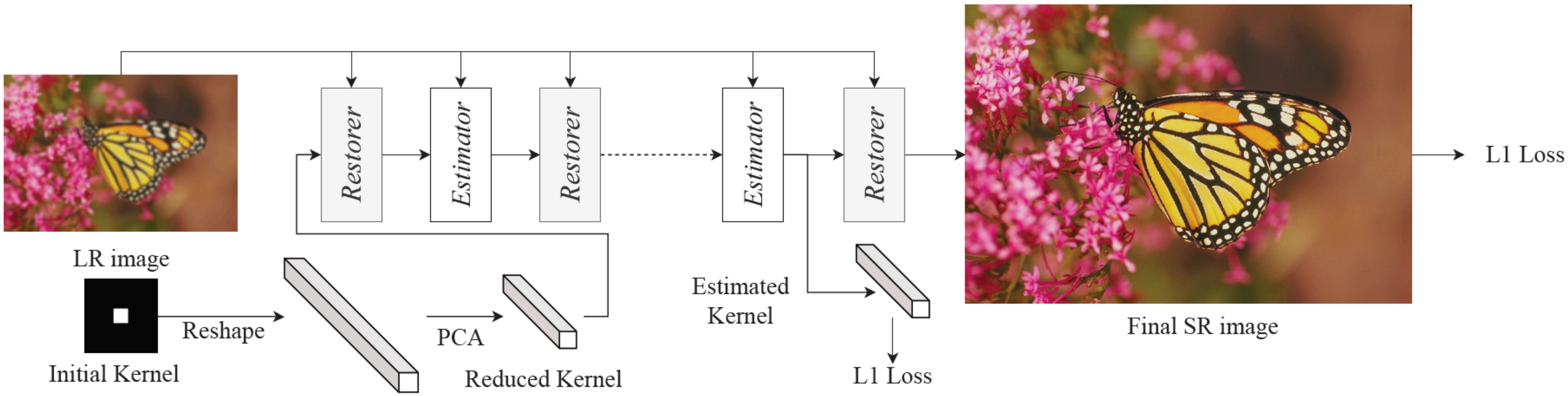
- Issue: **oversharpening**.

- ISE is trained to detect blurry or oversharpened regions and predicts errors.

- KEN uses the sharpness measures from ISE to estimate the kernel and select the best pretrained model.

- Misleading evaluations of sharpness by ISE may compromise the decision made by the KEN component.

Some Explainability: DAN



- Issue: **blurry images**.

- The kernel is initialised by Dirac function, and it is also reshaped and then reduced by principal component analysis (PCA).

- The kernel is reduced by PCA and, thus, the Estimator only needs to estimate the PCA result of the blur kernel.

- Loss of information when using PCA for dimensionality reduction, and recent evaluations show that PCA results are not as reliable and robust as it is usually assumed to be.

Final Remarks

- Independent and unbiased controlled experiments: important for professionals.

- MANIQA scores: FastGAN and APA (GAN-based approaches) were the worst techniques.

- Recommendation for blind image SR: single-image and non-GAN-based approaches are the best way to go (but it is necessary more experimentaion).

Final Remarks

- Recommendation among the DL techniques: MoESR.

- But … looking at the HR images generated by all DL techniques for all sets we can conclude that:
  - The perception quality of the images as a whole needs to improve;
  - New approaches, addressing larger scaling factors, are necessary for the future.

- Supporting code: https://github.com/vsantjr/DL_BlindSR

Article

AI

Submit to this Journal

Review for this Journal

Edit a Special Issue

Article Menu

Academic Editors

Chuan-Ming Liu

Wei-Shinn Ku

Subscribe SciFeed

Recommended Articles

Related Info Link

More by Author Links

Order Article Reprints

Open Access Article

Evaluating Deep Learning Techniques for Blind Image Super-Resolution within a High-Scale Multi-Domain Perspective

by Valdivino Alexandre de Santiago Júnior

Coordenação de Pesquisa Aplicada e Desenvolvimento Tecnológico (COPDT), Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, São Paulo 12227-010, Brazil

AI **2023**, *4*(3), 598-619; https://doi.org/10.3390/ai4030032

Received: 20 June 2023 / Revised: 14 July 2023 / Accepted: 24 July 2023 / Published: 1 August 2023

(This article belongs to the Topic Applied Computing and Machine Intelligence (ACMI))

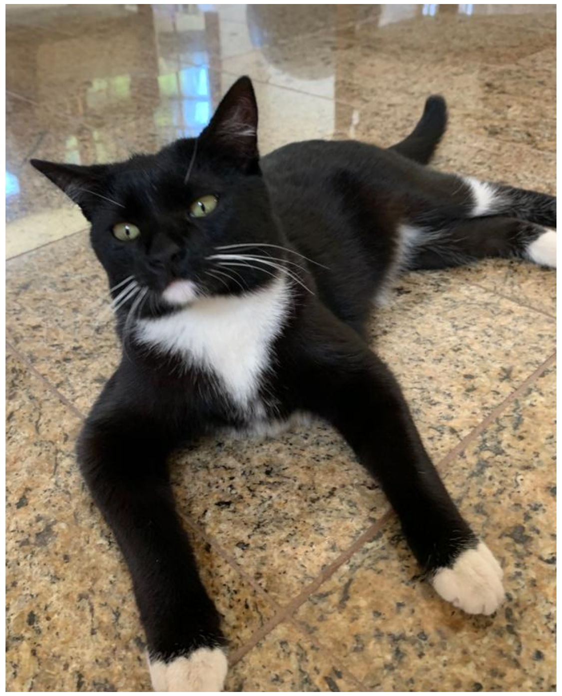
Download | Browse Figures | Versions Notes

Abstract

Despite several solutions and experiments have been conducted recently addressing image super-resolution (SR), boosted by deep learning (DL), they do not usually design evaluations with high scaling factors. Moreover, the datasets are generally benchmarks which do not truly encompass significant diversity of domains to proper evaluate the techniques. It is also interesting to remark that blind SR is attractive for real-world scenarios since it is based on the idea that the degradation process is unknown, and, hence, techniques in this context rely basically on low-resolution (LR) images. In this article, we present a high-scale (8×) experiment which evaluates five recent DL techniques tailored for blind image SR: Adaptive Pseudo Augmentation (APA), Blind Image SR with Spatially Variant

Source: https://www.mdpi.com/2673-2688/4/3/32

Thank You!



E-mail: valdivino.santiago@inpe.br

Web: http://www.lac.inpe.br/~valdivino/

GitHub: https://github.com/vsantjr