ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CVL Computer Vision Lab

Towards Practical Image Restoration and Enhancement

Dr. Shuhang GU
School of EIE, The University of Sydney

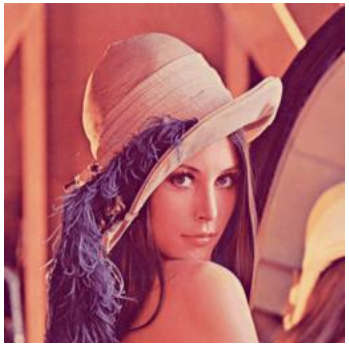
Outline

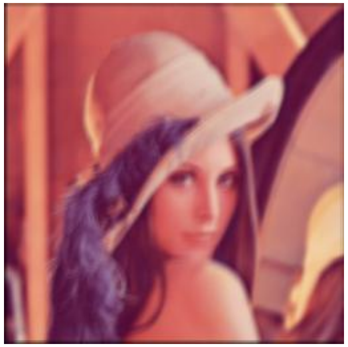
- **Background of Image Restoration and Enhancement**
- **DNN models for Image Restoration and Enhancement**
- **Towards Practical Image Restoration and Enhancement**
  - Efficient Models
  - New Application Scenarios
- **Conclusion**

Background of Image Enhancement

Classical Image Restoration Problem

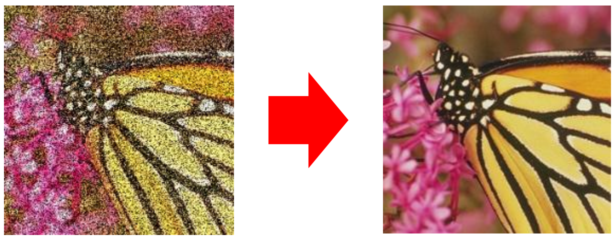
- Goal: Estimate the latent high quality image from its degraded observation.





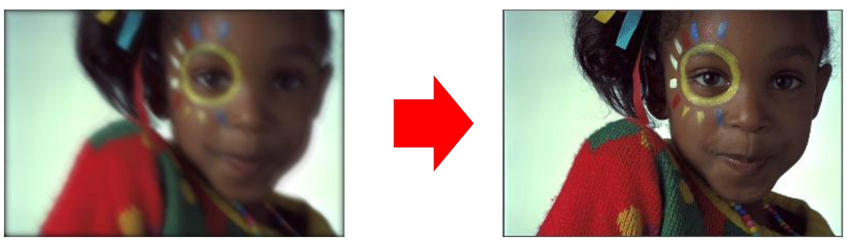
Typical Restoration Problems

- Image Denoising



- Image Super-Resolution

- Image Debluring



- Image Inpainting

From Image Restoration to Image Enhancement

Contrast Enhancement

Rain-Streak Removal

...

Explicit priors

- Maximum a posterior estimator

$$\max_{x} p(X \mid Y) = \max_{x} p(Y \mid X)\, p(X)$$

Likelihood comes from the degradation model

Independent Component Analysis

Variational Models

Markov Random Field

Sparse Representation Models

Implicit priors

Certain Operation

Implicit priors are embedded in the restoration operations.

- Low-pass filtering
- Diffusion-based methods
- Regression-based methods
- **Deep Learning methods**

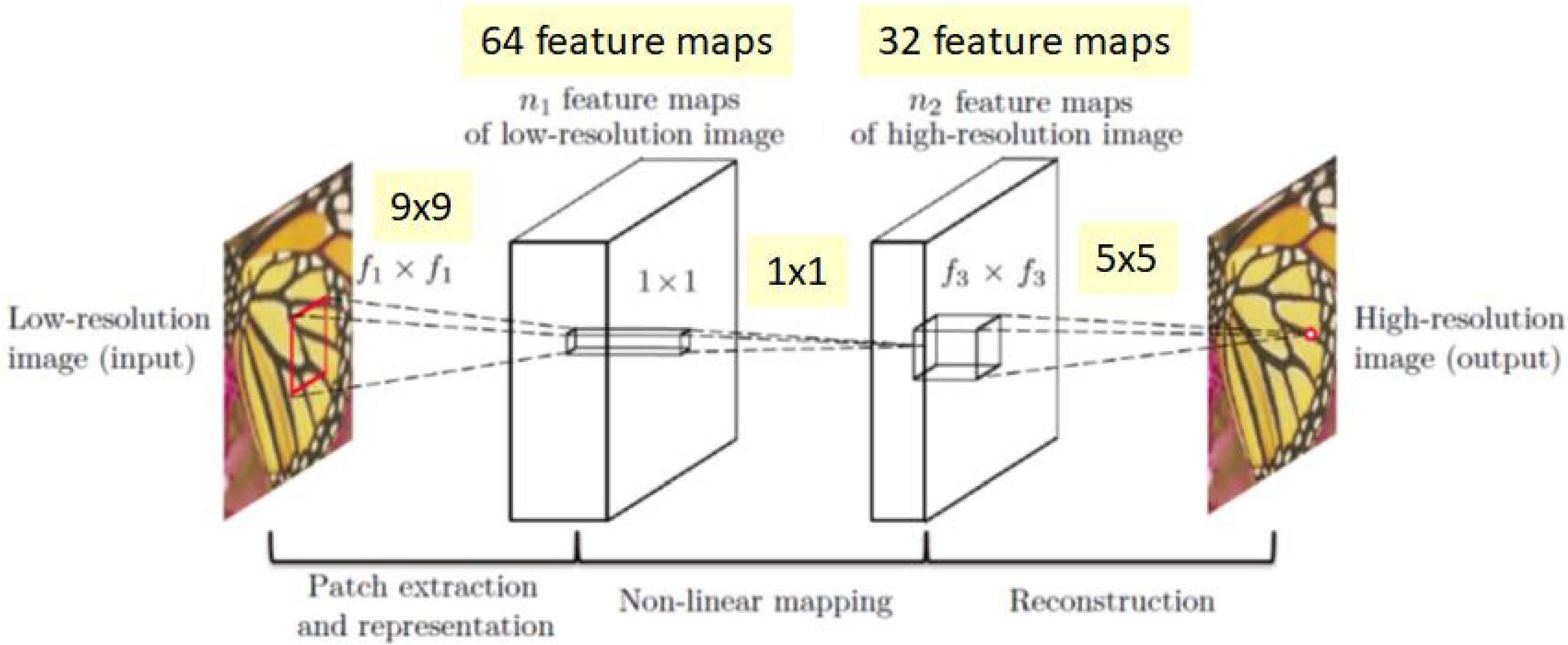
Deep Neural Networks for Image Restoration and Enhancement

DNN Restoration Models Before SRCNN

- **Jain and Seung, NIPS 2009.**
  - A small network with 4 hidden layers for image denoising.
- **Xie et al., NIPS 2012.**
  - Stack sparse denoising auto-encoder for image denoising.
- **Burger et al., CVPR 2012.**
  - Train a multi-layer perceptron for image denoising.

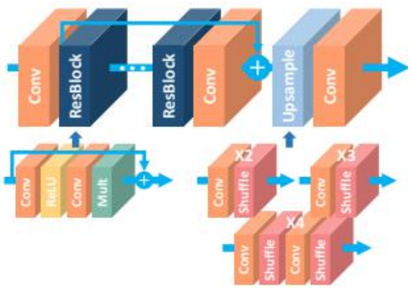
V. Jain and S. Seung, “Natural image denoising with convolutional networks”, In NIPS 2009.
J. Xie et al., “Image denoising and inpainting with deep neural networks”, In NIPS 2012.
Burger et al., "Image denoising: Can plain neural networks compete with BM3D?." CVPR 2012.

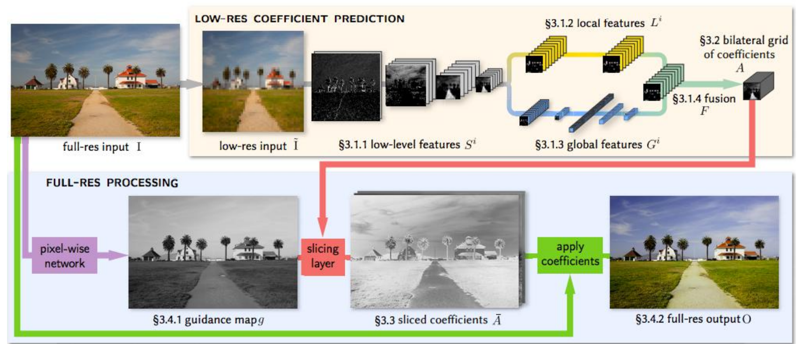
SRCNN: First S-o-t-a DNN Restoration Model

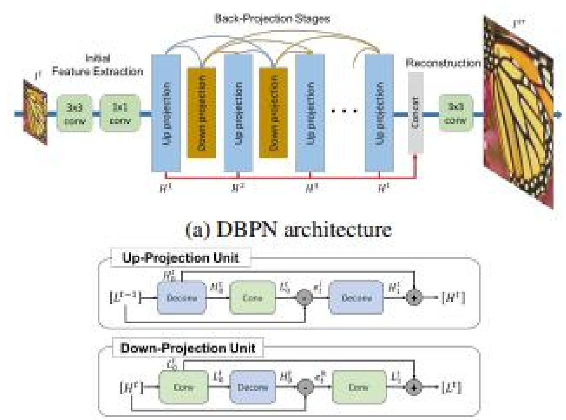


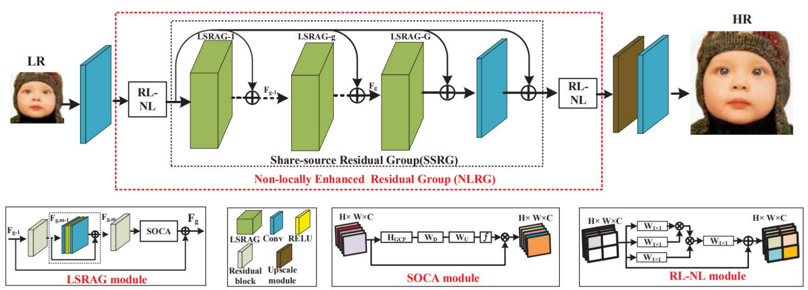
Dong, Chao, et al. "Image super-resolution using deep convolutional networks."
IEEE transactions on pattern analysis and machine intelligence 38.2 (2015): 295-307.

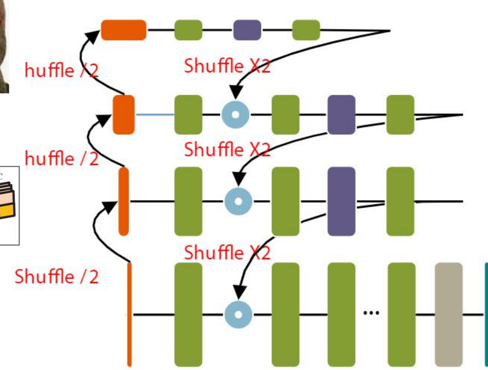
DNNs for Image Restoration and Enhancement

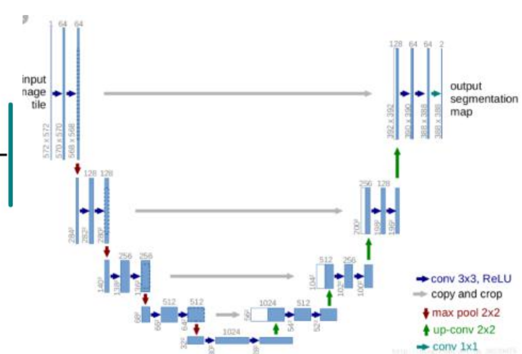


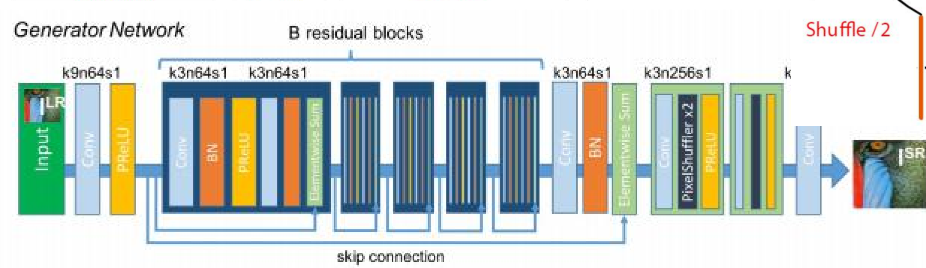












Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." CVPRW 2017.
Gharbi, Michaël, et al. "Deep bilateral learning for real-time image enhancement." TOG 2017.
Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." CVPR 2017.
Gu, Shuhang, et. al. “Self-Guided Network for fast image denoising” Submitted to ICCV 2019.
Dai, Tao, et al. "Second-order Attention Network for Single Image Super-Resolution." CVPR 2019.
Ronneberger, Olaf, et al., "U-net: Convolutional networks for biomedical image segmentation."MICCAI, 2015.
Haris, Muhammad, Gregory Shakhnarovich, and Norimichi Ukita. "Deep back-projection networks for super-resolution." CVPR 2018.

Limitations of Existing Approaches

- Efficiency

EDSR (X4)	20.05s	Winner of NTIRE 2017
DBPN (X8)	35.00s	Winner of NTIRE 2018

- Limited Scenarios
  - Mainly works only foucs on synthetic data
- Visual Quality
- Controlable Enhancement
- ...

- Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." CVPRW 2017.
- Haris, Muhammad, Gregory Shakhnarovich, and Norimichi Ukita. "Deep back-projection networks for super-resolution." CVPR 2018.
- Shuhang Gu, et al. “AIM 2019 Challenge on Extreme Image Super Resolution: Methods and Results”, ICCVW 2019.

Towards Practical

Image Restoration and Enhancement

Towards Practical
Image Restoration and Enhancement

Computation/Memory Efficient
Image Restoration and Enhancement

Towards fast enhancement networks

- Key Issues in deep enhancement networks
  - Rerceptive Field: incorporating sufficient information from input image
  - Non-linear Capacity: capture the complex relationship between input and output

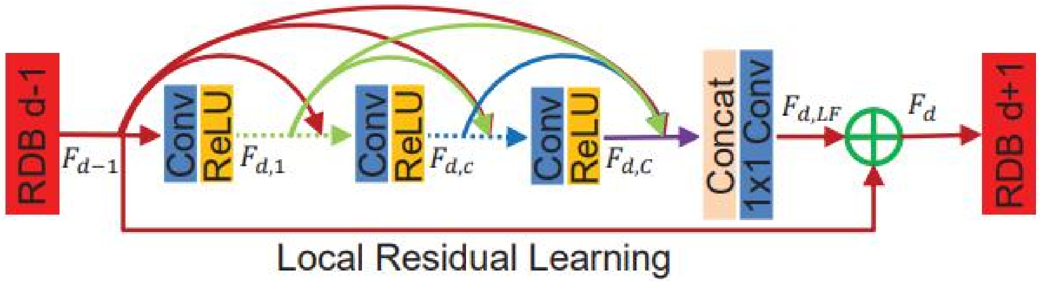
Towards fast enhancement networks

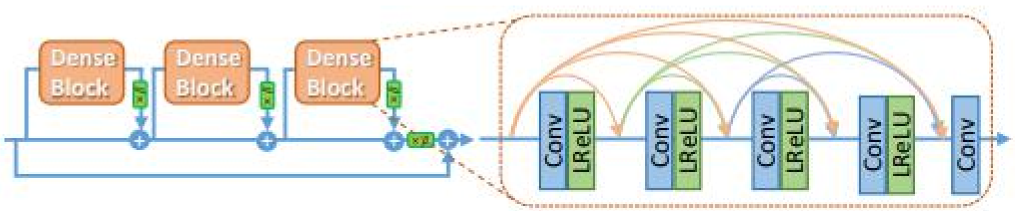
- Key Issues in deep enhancement networks
  - Rerceptive Field: incorporating sufficient information from input image
  - Non-linear Capacity: capture the complex relationship between input and output

  - Temporal Information incorporation: exploit temporal information for video manipulaiton

Non-linear Capacity Strengthen

- Adopt deeper netwrok architectures, using residual/dense blocks.



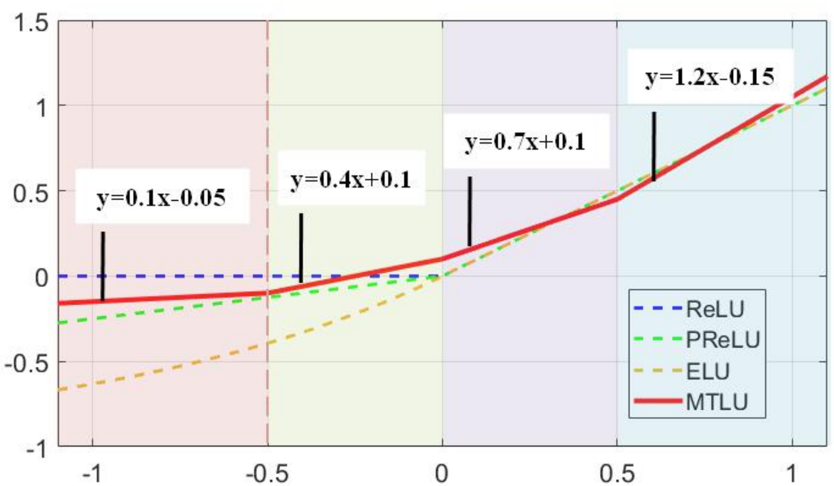


Xintao Wang et. al. Esrgan: Enhanced super-resolution generative adversarial networks
Yulun Zhang et. al. Residual dense network for image super-resolution.

Non-linear Capacity Strengthen by Parameterized Non-linear Activation Function

- Multi-bin Trainable Linear Unit

$$f(x) = \begin{cases} a_0 x + b_0, & if\ x \le c_0; \\ a_k x + b_k, & if\ c_{k-1} < x \le c_k; \\ \quad \ldots & \\ a_K x + b_K, & if\ c_{K-1} < x. \end{cases}$$

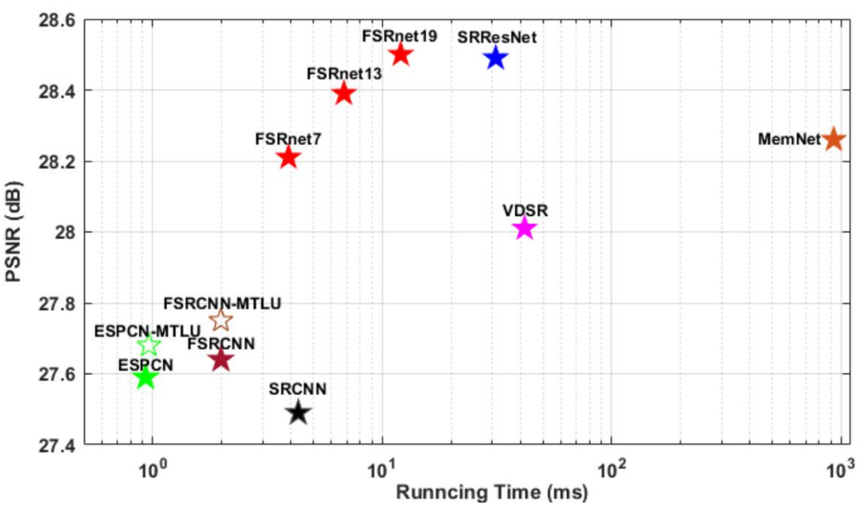


Computer Vision Lab

Gu, Shuhang, Wen Li, Radu Timofte, and Luc Van Gool. "Multi-bin Trainable Linear Unit for Fast Image Restoration Networks." *arXiv preprint arXiv:1807.11389* (2018).

Efficient Image-to-Image mapping with MTLU

- SR performance



- Denoise performance

Runtime for processing a 512x512 noisy image

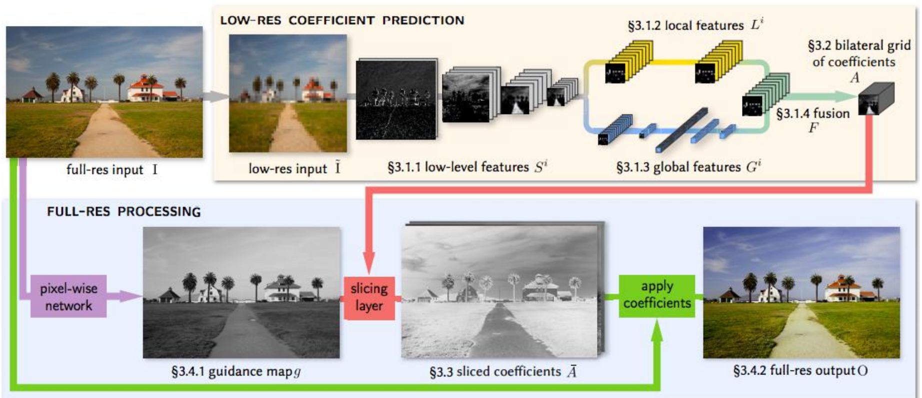
	DnCNN	FDnet
Runtime [ms]	74.8	5.0

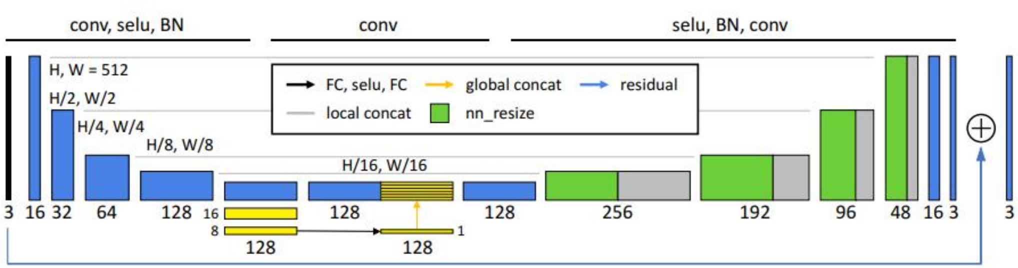
Average PSNR on Benchmark Dataset

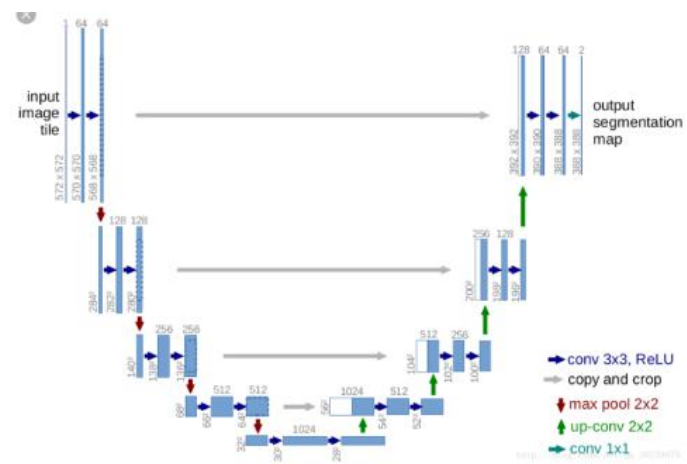
	DnCNN	FDnet
Sigma 25	29.23	29.12
Sigma 50	26.23	26.24
Sigma 70	24.64	24.76

References can be found in our paper: Gu, Shuhang, Radu Timofte, and Luc Van Gool. "Multi-bin Trainable Linear Unit for Fast Image Restoration Networks." *arXiv preprint arXiv:1807.11389* (2018).

Rerceptive Field Enlargement



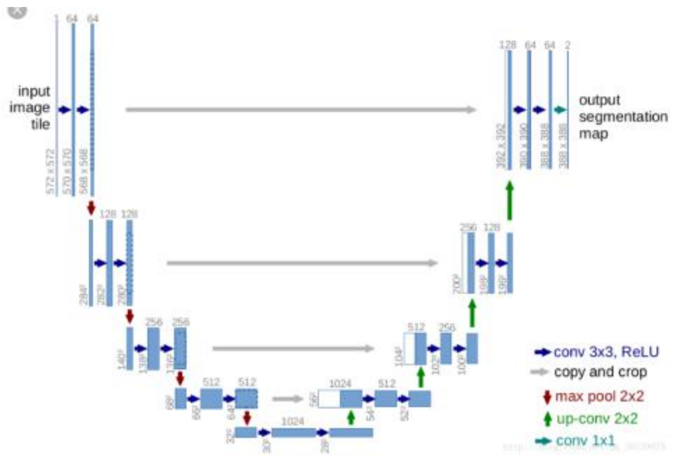




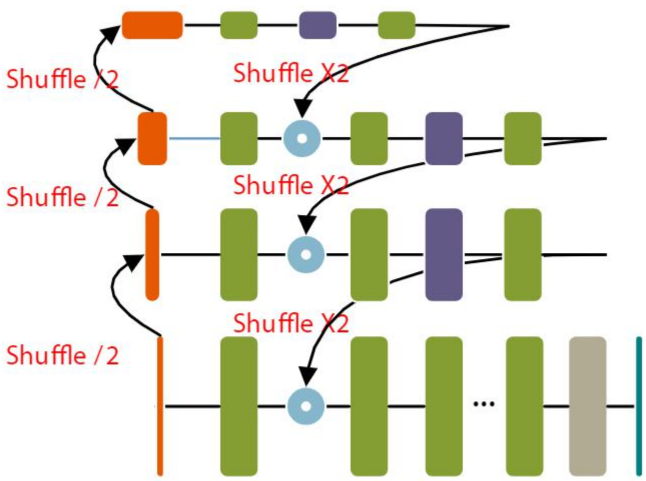
Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation."MICCAI, 2015.
Gharbi, Michaël, et al. "Deep bilateral learning for real-time image enhancement." ACM Transactions on Graphics (TOG) 36.4 (2017): 118.
Chen, Yu-Sheng, et al. "Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans." CVPR 2018.

Rerceptive Field Enlargement by Incorporating Contextual Information with SGN

- U-Net



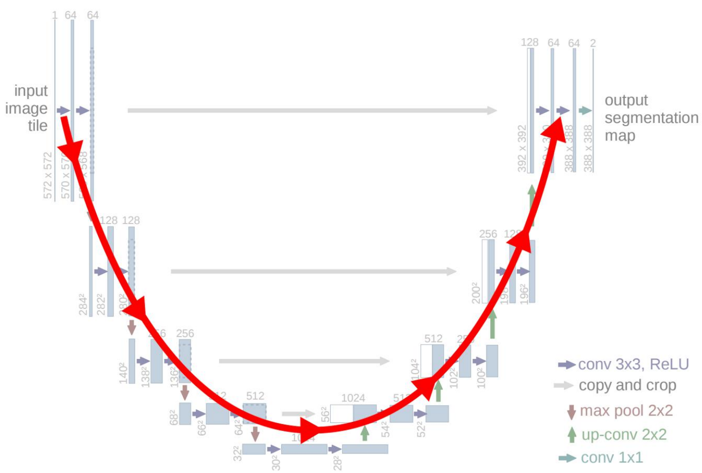
- Self-guided Network



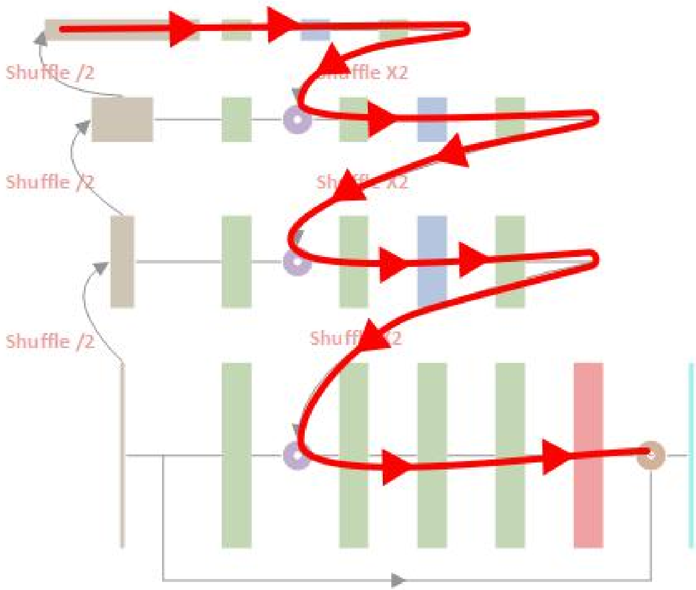
- Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation."MICCAI, 2015.
- Gu, Shuhang, et. al. “Self-Guided Network for fast image denoising” Submitted to ICCV 2019.

Rerceptive Field Enlargement by Incorporating Contextual Information with SGN

- U-Net

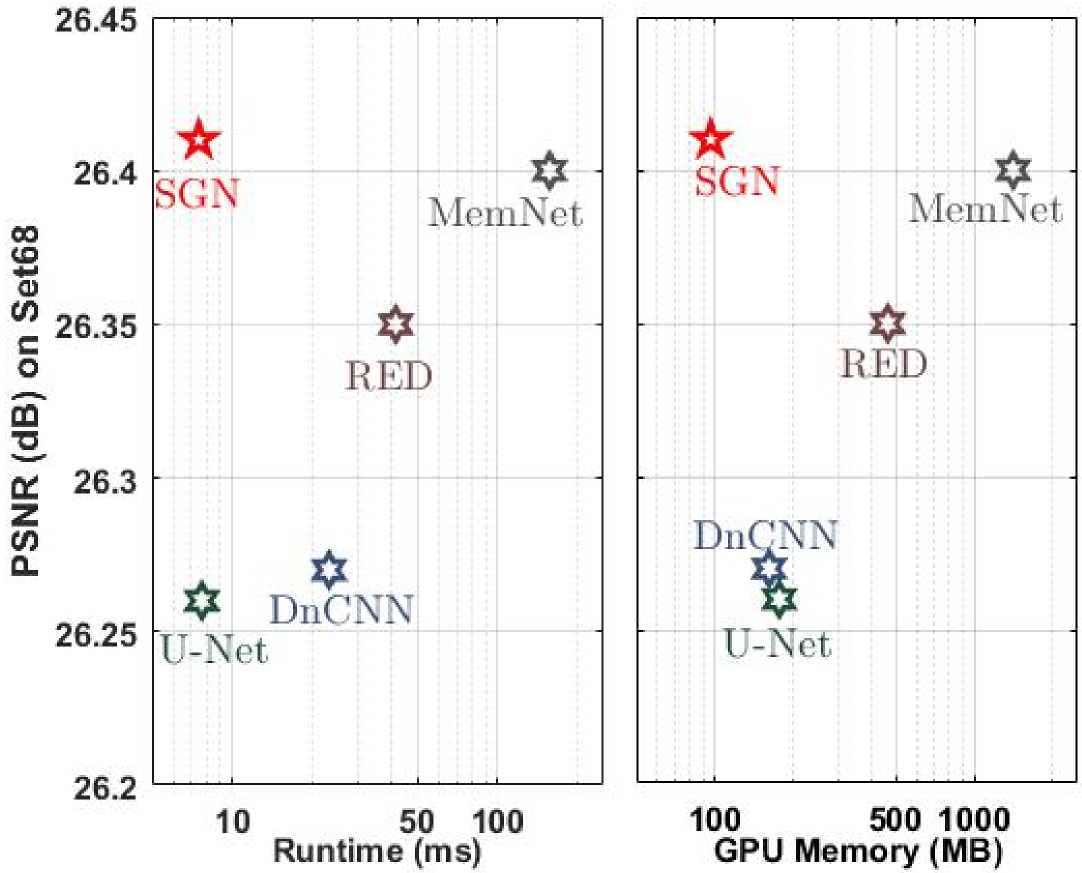


- Self-guided Network



- Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation."MICCAI, 2015.
- Gu, Shuhang, et. al. “Self-Guided Network for fast image denoising” Submitted to ICCV 2019.

Comarison between SGN and S.O.T.A Methods



References can be found in our paper: Gu, Shuhang, et. al. "Self-Guided Network for fast image denoising" Submitted to ICCV 2019.

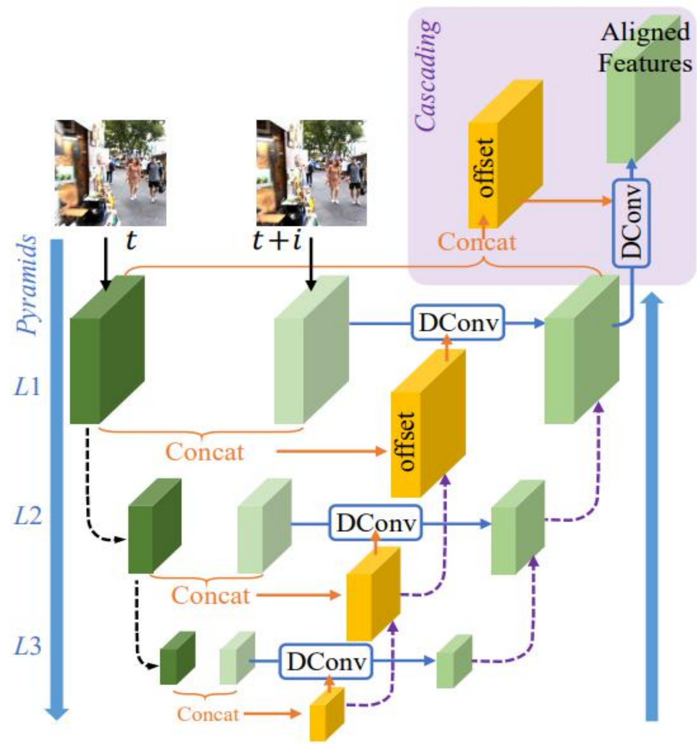
Comarison between SGN and S.O.T.A Methods

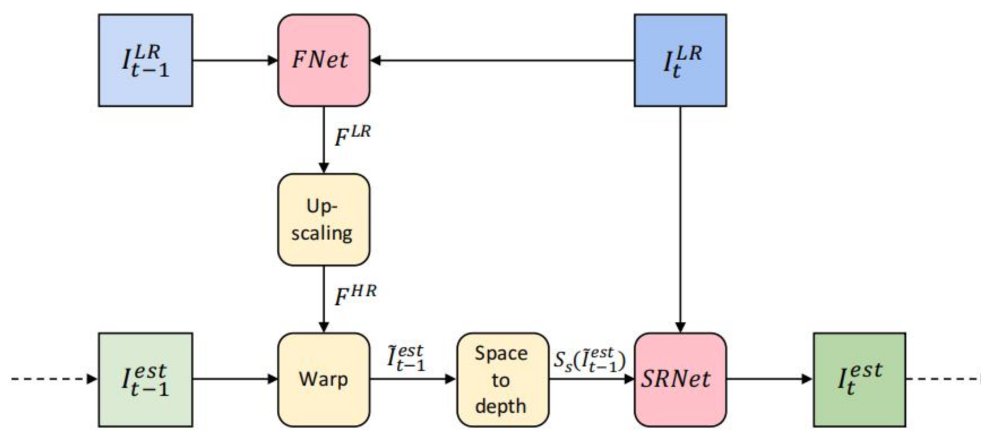
Reference Image | U-Net [33] | SGN_{L3}

CVL Computer Vision Lab

References can be found in our paper: Gu, Shuhang, et. al. “Self-Guided Network for fast image denoising” Submitted to ICCV 2019.

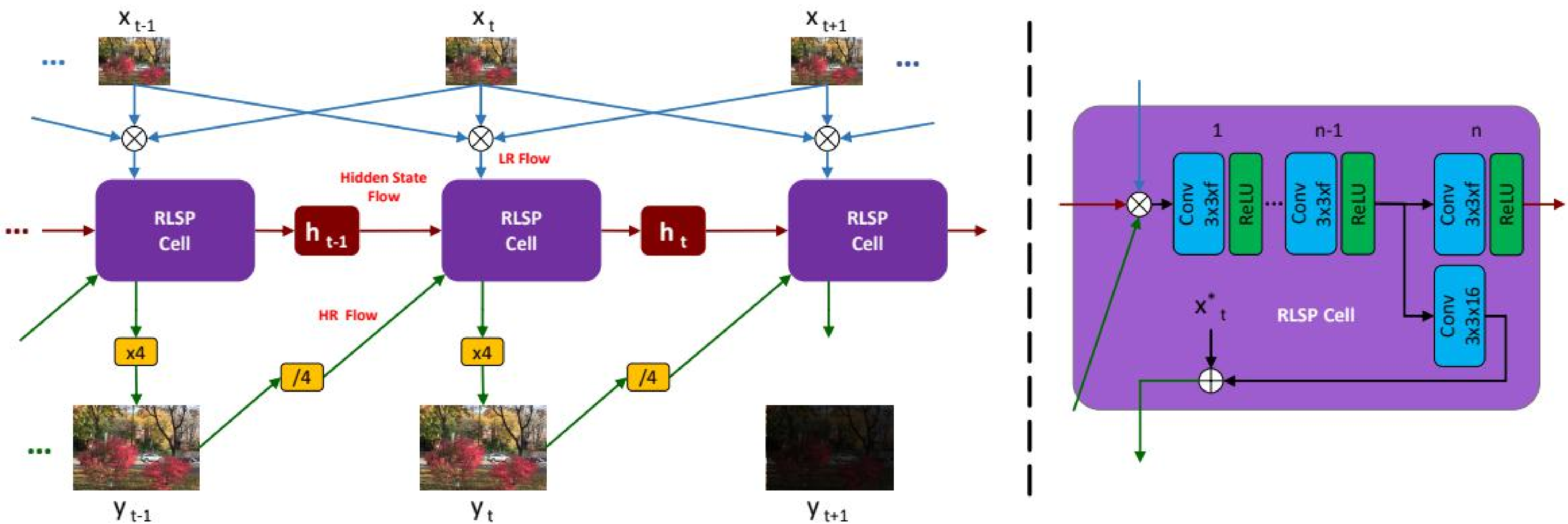
Temporal Information Incorporation





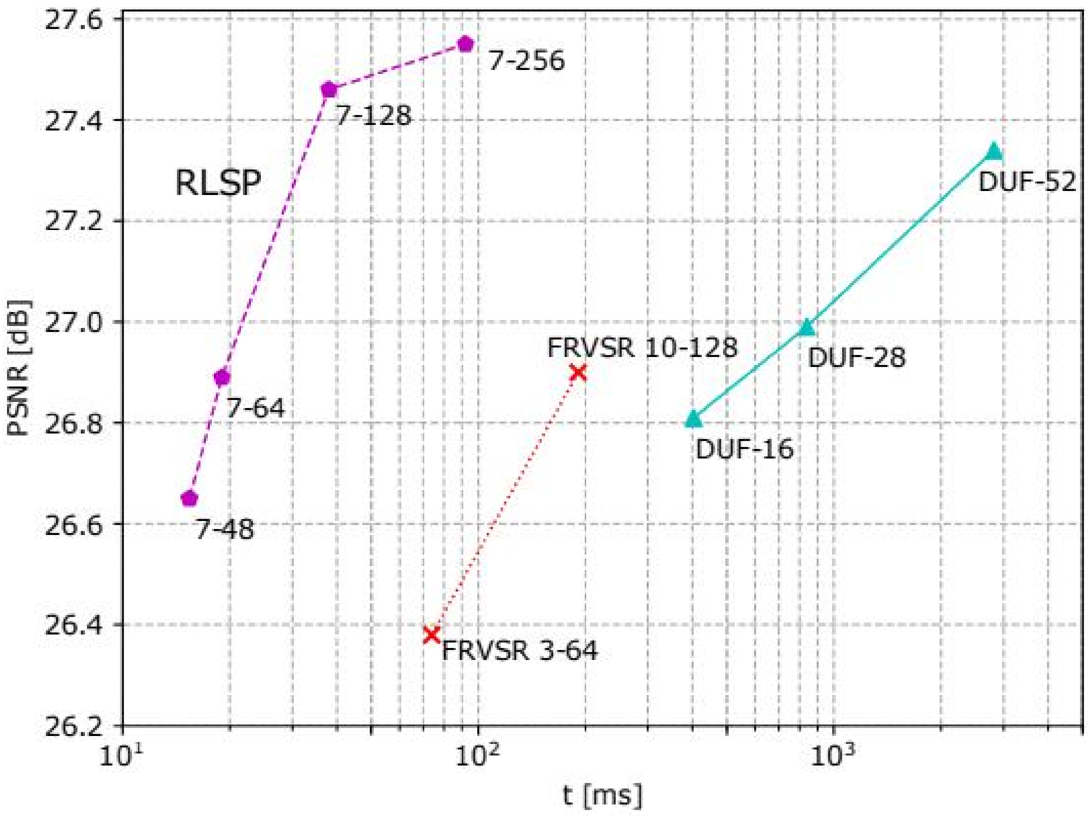
Mehdi S. M. Sajjadi et. al. Frame Recurrent Video Super-Resolution.
Xintao Wang et. al. EDVR: Video Restoration with Enhanced Deformable Convolutional Networks.

Temporal Information Incorporation by Recurrent Latent Space Propagation

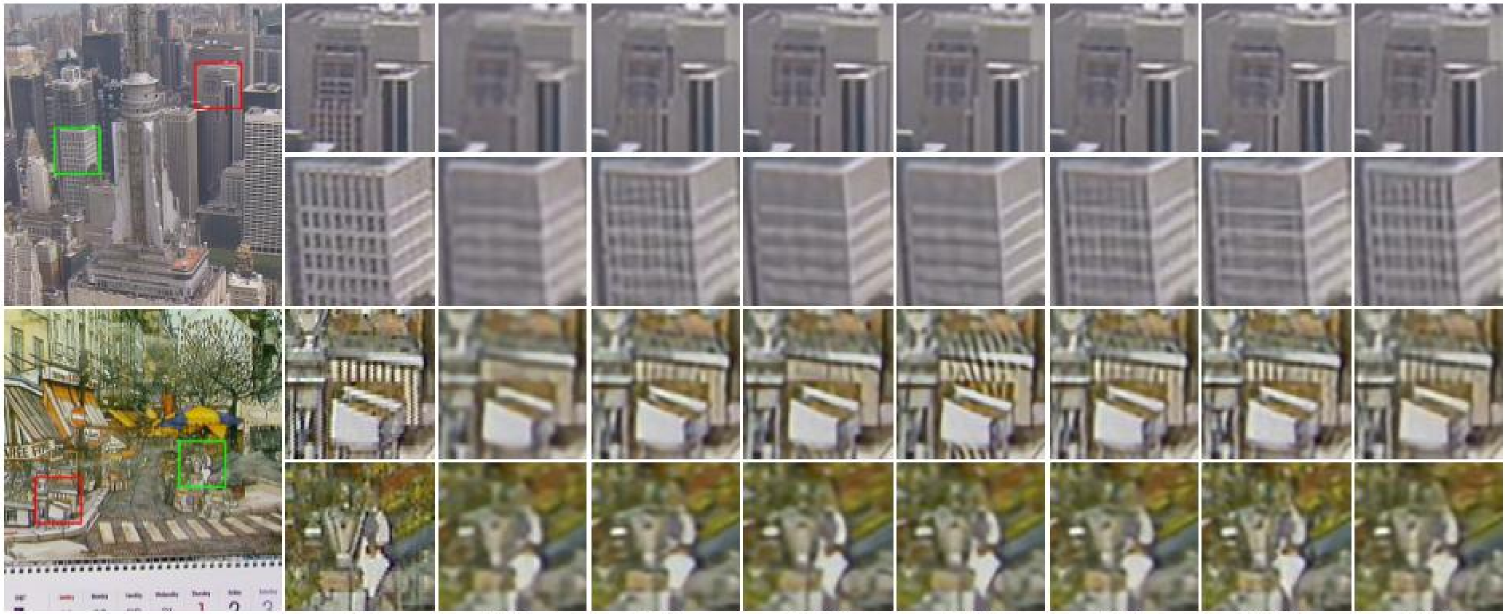


References can be found in our paper: Dario et. al. “Efficient Video Super-Resolution Through Recurrent Latent Space Propagation ” ICCVW 2019.

Efficient Video Processing with RLSP



Efficient Video Processing with RLSP



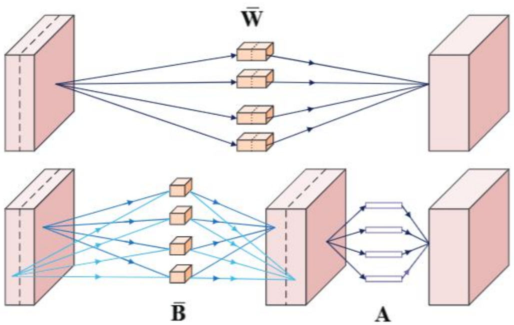
Vid4	#Frame	FLOPs	#Param.	Calendar (Y)	City (Y)	Foliage (Y)	Walk (Y)	Average (Y)	Average (RGB)
Bicubic	1	N/A	N/A	18.83/0.4936	23.84/0.5234	21.52/0.4438	23.01/0.7096	21.80/0.5426	20.37/0.5106
SPMC † [24]	3	-	-	-/-	-/-	-/-	-/-	25.52/0.76	-/-
Liu† [20]	5	-	-	21.61/-	26.29/-	24.99/-	28.06/-	25.23/-	-/-
TOFlow [26]	7	0.81T	1.41M	22.29/0.7273	26.79/0.7446	25.31/0.7118	29.02/0.8799	25.84/0.7659	24.39/0.7438
DUF-52L [11]	7	0.62T	5.82M	24.17/0.8161	28.05/0.8235	26.42/0.7758	30.91/0.9165	27.38/0.8329	25.91/0.8166
RBPN [7]	7	9.30T	12.2M	24.02/0.8088	27.83/0.8045	26.21/0.7579	30.62/0.9111	27.17/0.8205	25.65/0.7997
EDVR-L† [25]	7	0.93T	20.6M	24.05/0.8147	28.27/0.8122	26.34/0.7635	31.02/0.9152	27.35/0.8264	25.83/0.8077
PFNL† [28]	7	0.70T	3.00M	24.37/0.8246	28.09/0.8385	26.51/0.7768	30.65/0.9135	27.40/0.8384	25.67/0.8189
FRVSR 10-128 [22]	recurrent (2)	0.13T	5.05M	23.50/0.7932	27.82/0.8148	26.03/0.7621	29.79/0.9059	26.78/0.8190	25.34/0.8018
RLSP 7-256 [4]	recurrent (3)	0.09T	4.21M	24.35/0.8265	28.44/0.8387	26.76/0.7897	30.62/0.9140	27.54/0.8422	25.69/0.8153

References can be found in our paper: Dario et. al. “Efficient Video Super-Resolution Through Recurrent Latent Space Propagation ” ICCVW 2019.

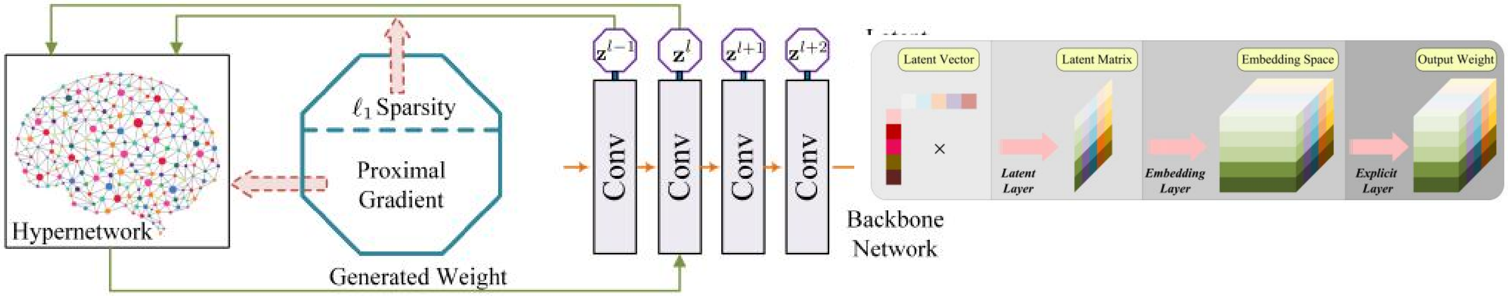
Memory Efficient Enhancement Models

- Weight Reusing
- Network compression

Filter Basis Learning (Li et al., ICCV 2019)



- Differentiable Meta Purning (Under Review)



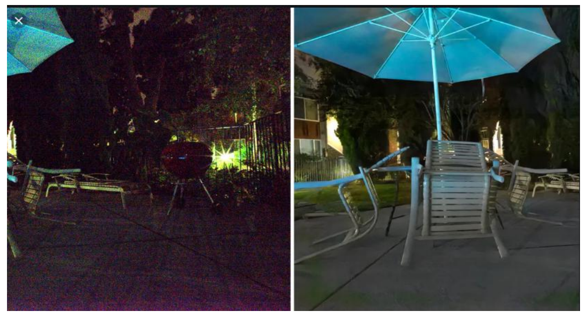
- Yawei Li, et al. “Learning Filter Basis for Convolutional Neural Network Compression”, ICCV 2019.
- Yawei Li, et al. “DHP: Differentiable Meta Purning via Hypernetworks.”, Submitted to ECCV.

Towards Practical Image Restoration and Enhancement

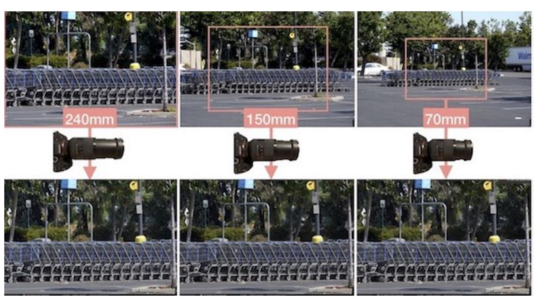
Restoration and Enhancement in More Complex Scenarios

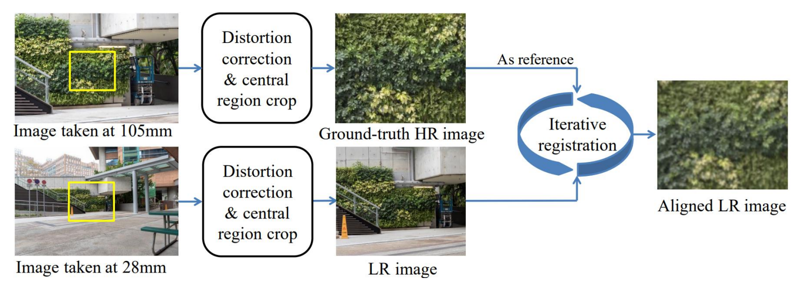
Data Collection for More Complex Enhancement Scenarios

- Better photography condtion



- Different photography conditions

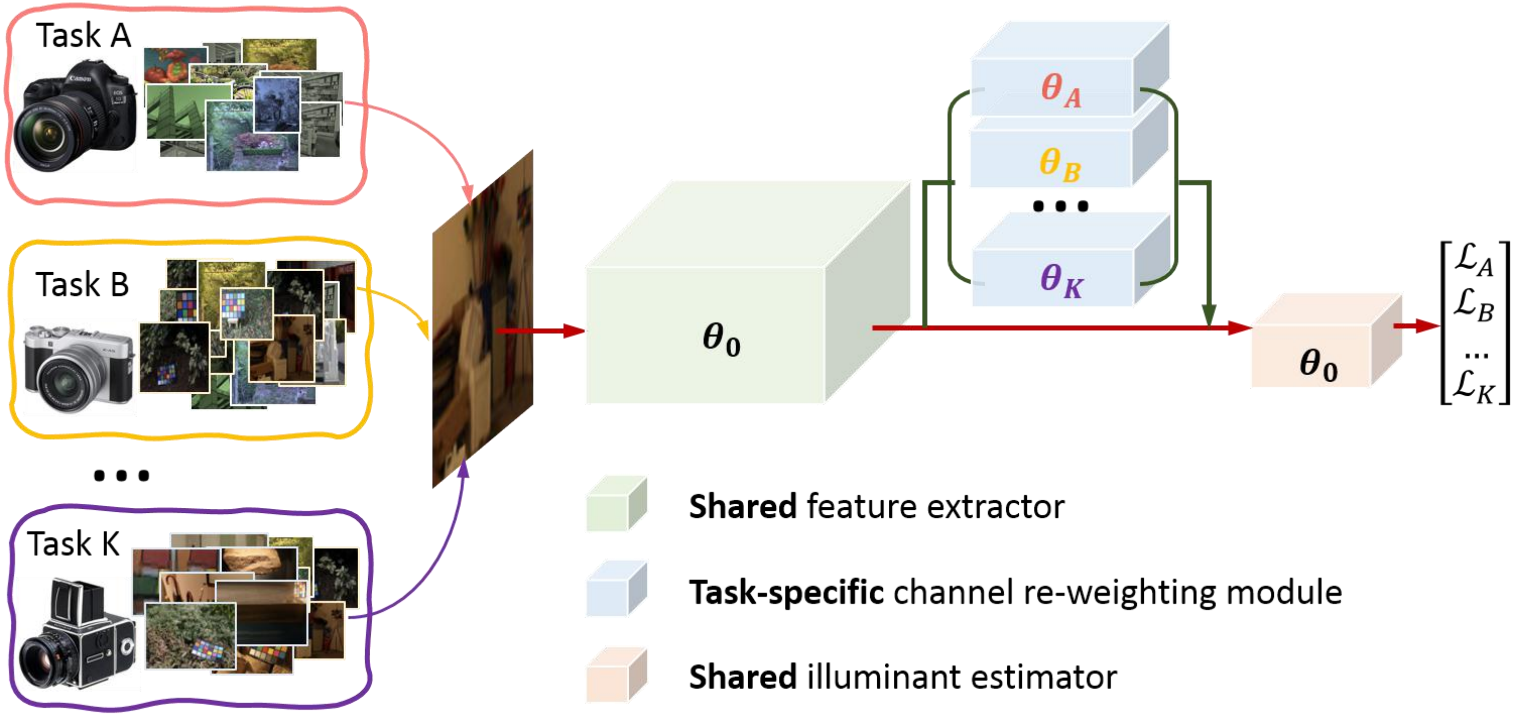




Chen, Chen, et al. "Learning to see in the dark." CVPR. 2018.
Pang, Jiahao, et al. "Zoom and learn: Generalizing deep stereo matching to novel domains." CVPR. 2018.
J. Cai*, H. Zeng*, H. Yong, Z. Cao, L. Zhang, "Toward Real-World Single Image Super-Resolution: A New Benchmark and A New Model," in ICCV 2019

Weakly Supervised Approaches

- Few-shot Color Constancy



CVPR 2020. To appear

Unsupervised Image Enhancement

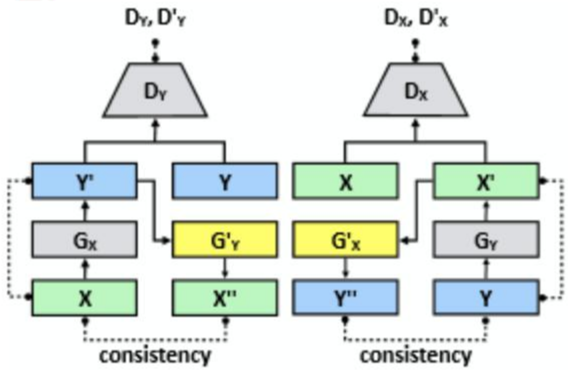
- Unsupervised learning for image quality mapping

Ignatov, Andrey, et al. "WESPE: weakly supervised photo enhancer for digital cameras." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018.

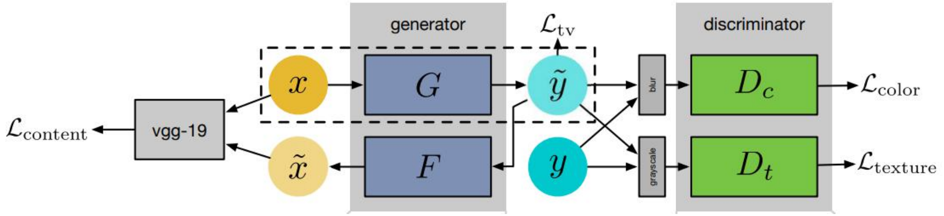
Unsupervised Image Enhancement

- Unsupervised learning for image quality mapping

DPE:



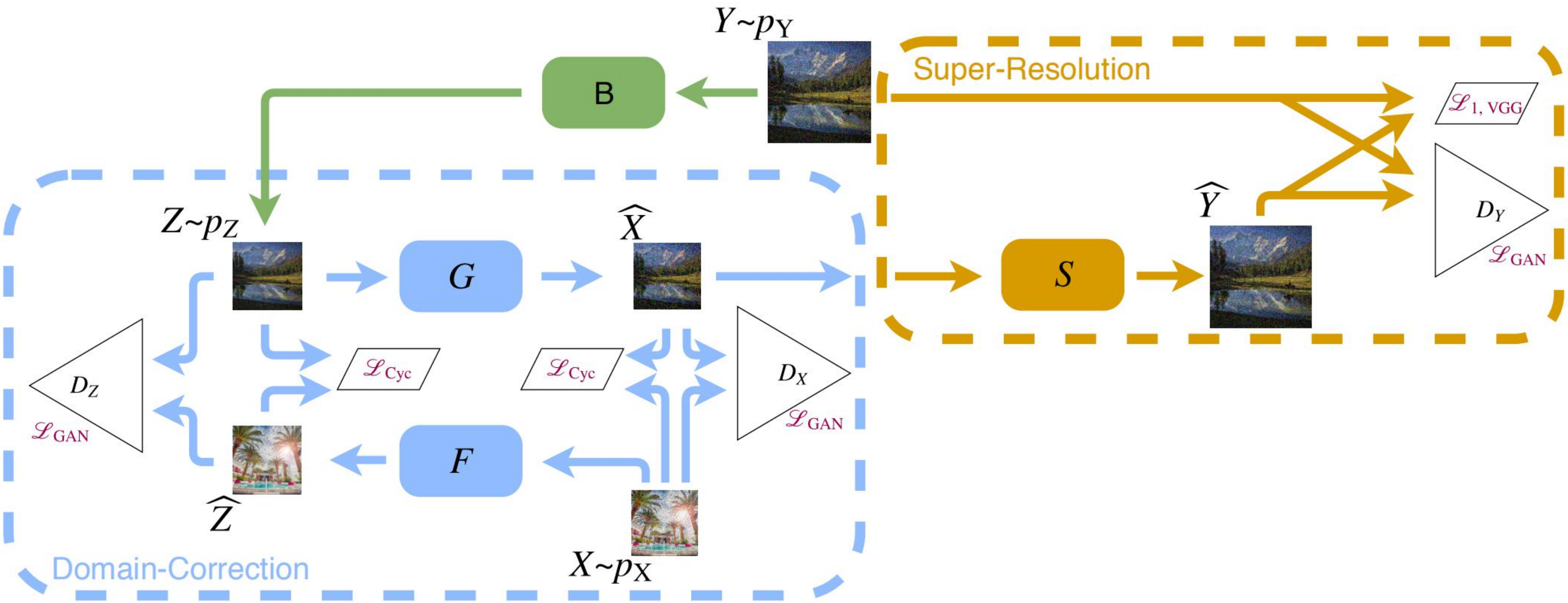
WESPE:



Deep Photo Enhancer:Unpaired Learning for Image Enhancement from Photographs with GANs. CVPR 2018.
Ignatov, Andrey, et al. "WESPE: weakly supervised photo enhancer for digital cameras." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2018.

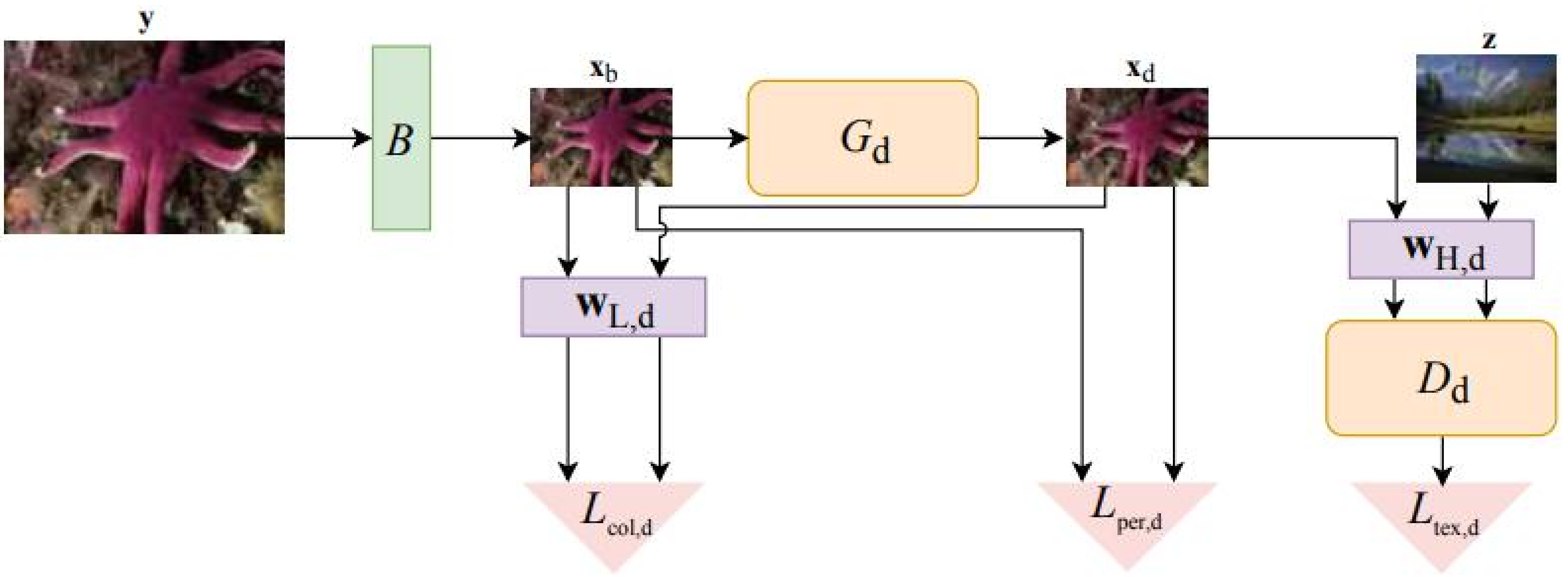
Unsupervised Super-Resolution

- Unsupervised learning for image SR



Unsupervised Image Super-Resolution

- Frequency Separation for Real-World SR



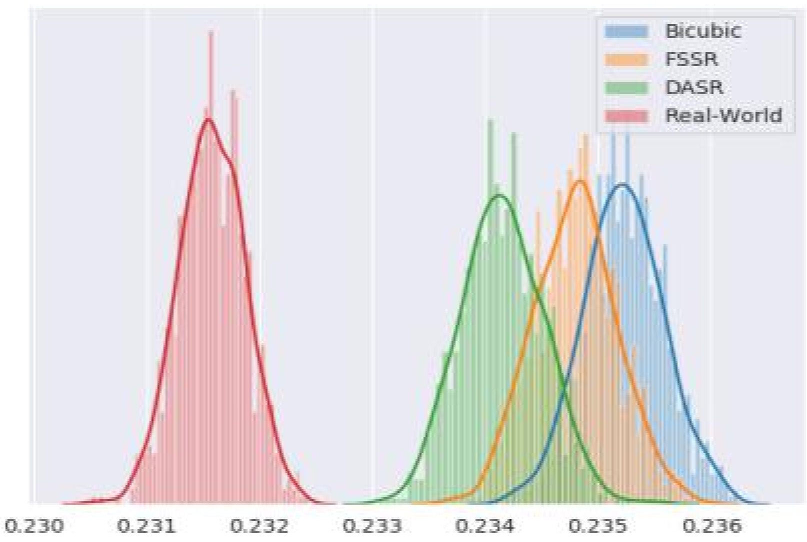
Unsupervised Image Super-Resolution

- Frequency Separation for Real-World SR

Original GT Ours Ours w/o high-pass Bicubic

Unsupervised Image Super-Resolution

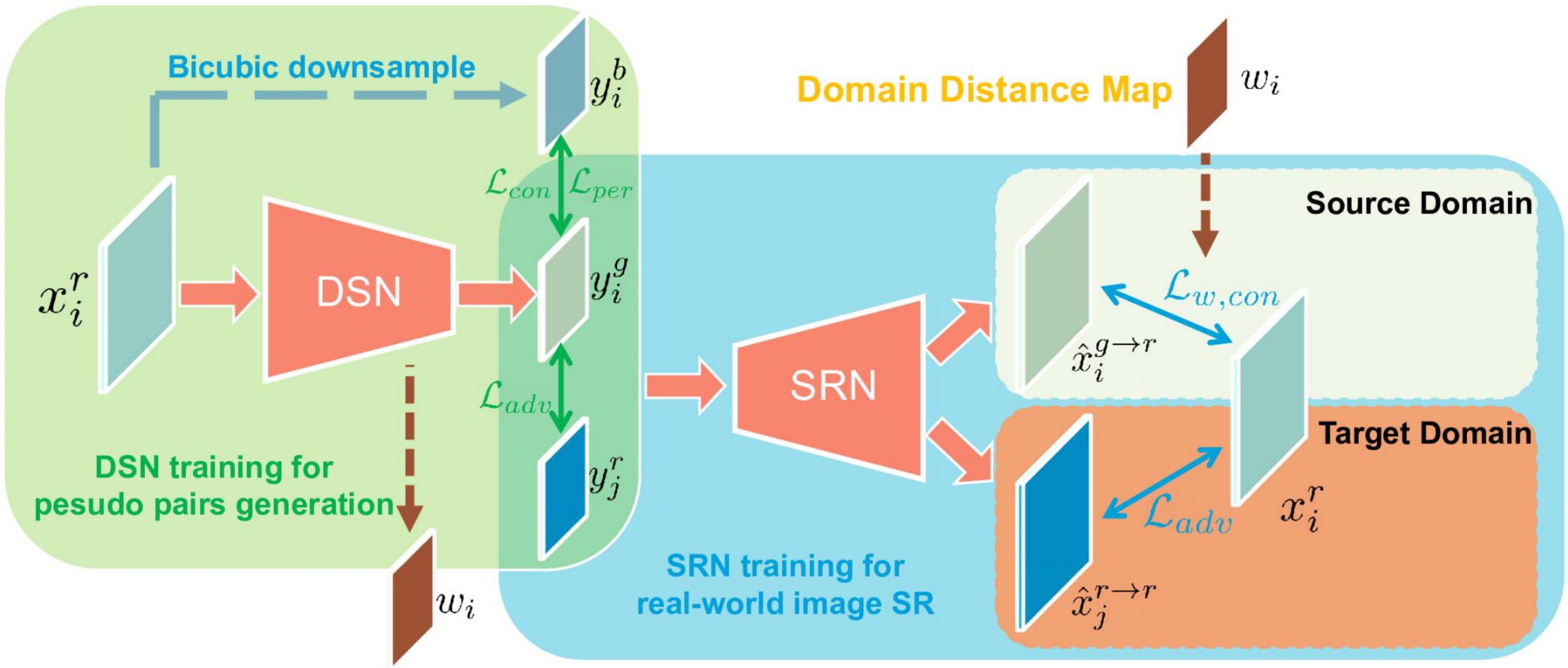
- Domain-distance Aware SR



Wei et. al. Unsupervised Real-world Image Super-Resolution via Domain-distance Aware Training. Arxiv.

Unsupervised Image Super-Resolution

- Domain-distance Aware SR

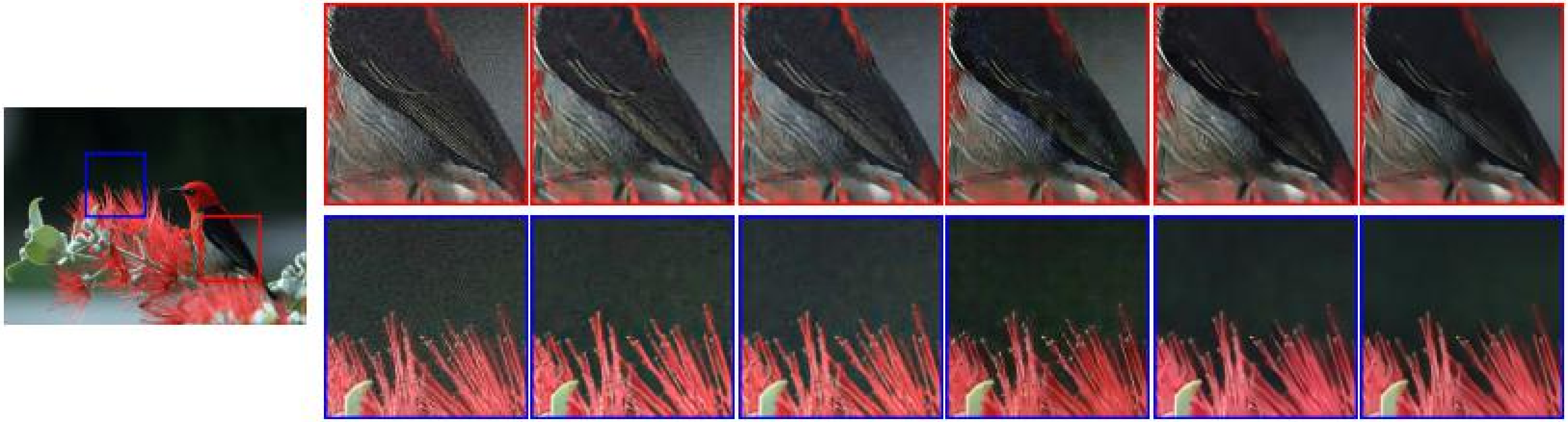


Wei et. al. Unsupervised Real-world Image Super-Resolution via Domain-distance Aware Training. Arxiv.

Unsupervised Image Super-Resolution

- Domain-distance Aware SR

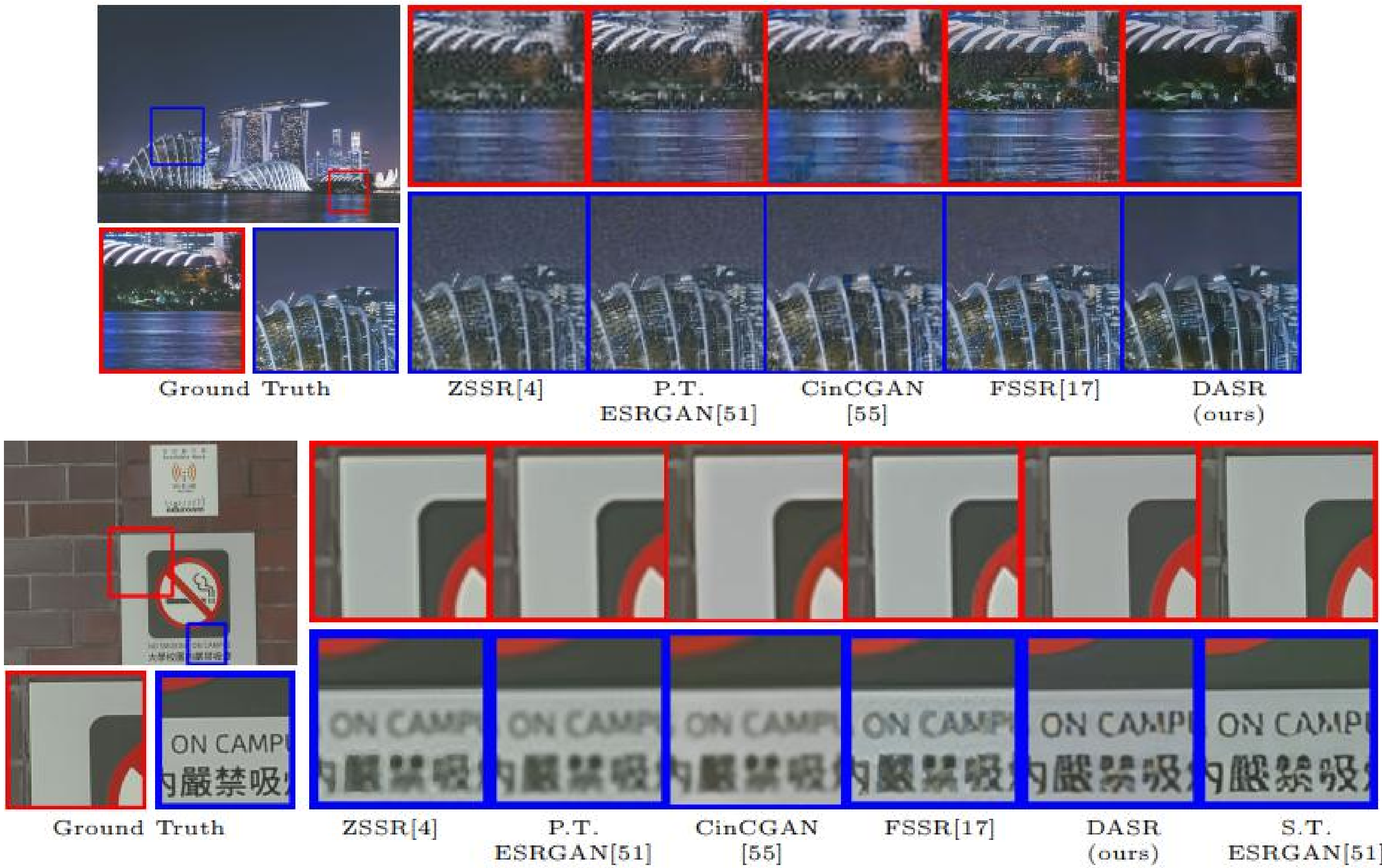
Source domain	$\{\mathcal{Y}^b, \mathcal{X}^r\}$	$\{\mathcal{Y}^b, \mathcal{X}^r\}$	$\{\mathcal{Y}^g, \mathcal{X}^r\}$	$\{\mathcal{Y}^g, \mathcal{X}^r\}$	$\{\mathcal{Y}^g, \mathcal{X}^r\}$	$\{\mathcal{Y}^g, \mathcal{X}^r\}$
Target domain	-	$\mathcal{Y}^r$	-	$\mathcal{Y}^r$	-	$\mathcal{Y}^r$
Domain gap aware	✗	✓	✗	✓	✗	✓
Weighted sup.	✗	✗	✗	✗	✓	✓
PSNR	21.382	20.820	21.910	21.805	21.452	21.600
SSIM	0.5478	0.5103	0.5555	0.5615	0.5304	0.5640
LPIPS	0.543	0.390	0.378	0.359	0.348	0.336
MOS	3.16	2.87	2.41	2.37	2.31	1.94



Wei et. al. Unsupervised Real-world Image Super-Resolution via Domain-distance Aware Training. Arxiv.

Unsupervised Image Super-Resolution

- Domain-distance Aware SR



Wei et. al. Unsupervised Real-world Image Super-Resolution via Domain-distance Aware Training. Arxiv.

Conclusion

Towards Practical Image Enhancement

- **Efficiency**
  - Computational Efficient
  - Memory Efficient
- **Application Scenarios**
  - Data Collection
  - Unsupervised Learning
- **Visual Quality**
  - Conditional Enhancement
  - ...

Thank you for your attention!