ICML International Conference On Machine Learning 40 Years

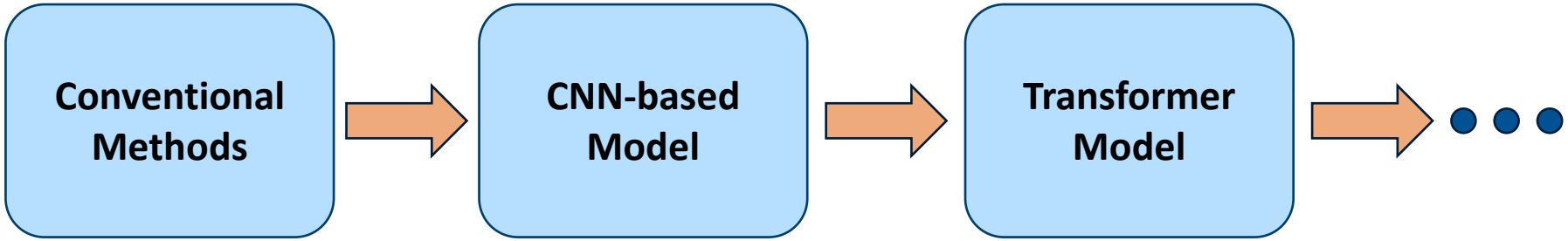
IRNeXt: Rethinking Convolutional Network Design for Image Restoration

Yuning Cui[1], Wenqi Ren[2], Sining Yang[2], Xiaochun Cao[2], Alois Knoll[1]

[1] Technical University of Munich [2] Shenzhen Campus of Sun Yat-sen University

Code: https://github.com/c-yn/IRNeXt

Motivation



Transformer Model

Global perceptive filed √ Quadratic complexity ×

Our Goal

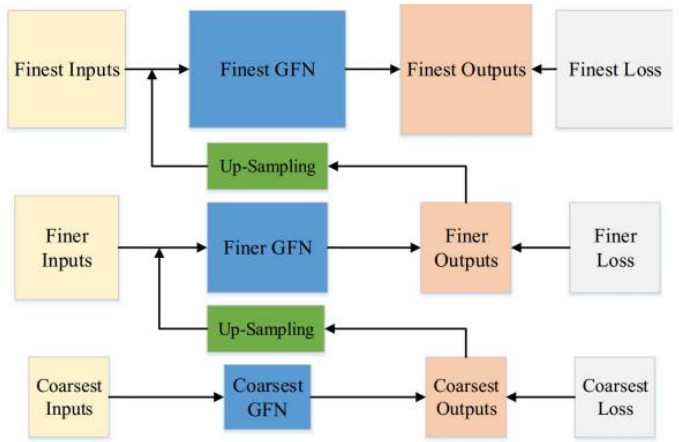
Rethinking Convolutional Network Design for Image Restoration

Comparable/Better performance √ Efficient √

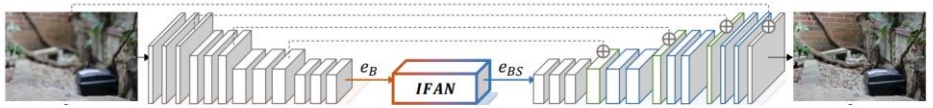
Methodology

→Revisit previous successful image restoration methods

✓Multi-scale learning



Multi-stage network



Encoder-decoder network

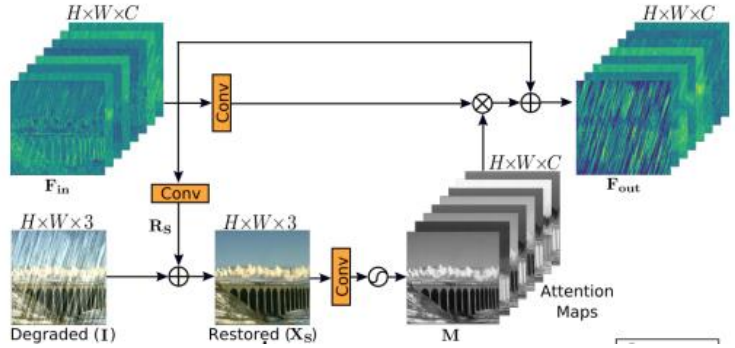
Ren, Wenqi, et al. "Gated fusion network for single image dehazing." CVPR-2018.
Lee, Junyong, et al. "Iterative filter adaptive network for single image defocus deblurring." CVPR-21.

Methodology

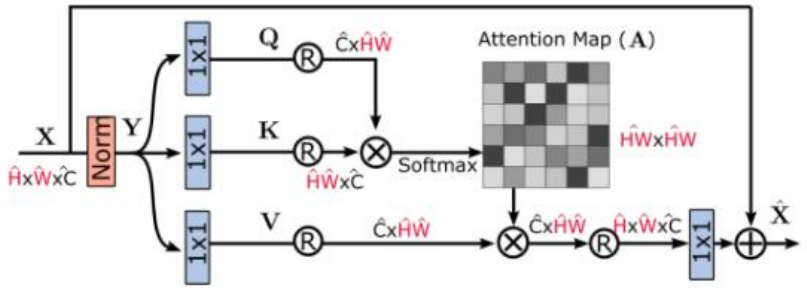
→Revisit previous successful image restoration methods

✓Multi-scale learning

✓Spatial attention



Attention w/o aggregation



Attention w/ aggregation

Zamir, Syed Waqas, et al. "Multi-stage progressive image restoration." CVPR-21.
Lee, Junyong, et al. "Iterative filter adaptive network for single image defocus deblurring." CVPR-21.

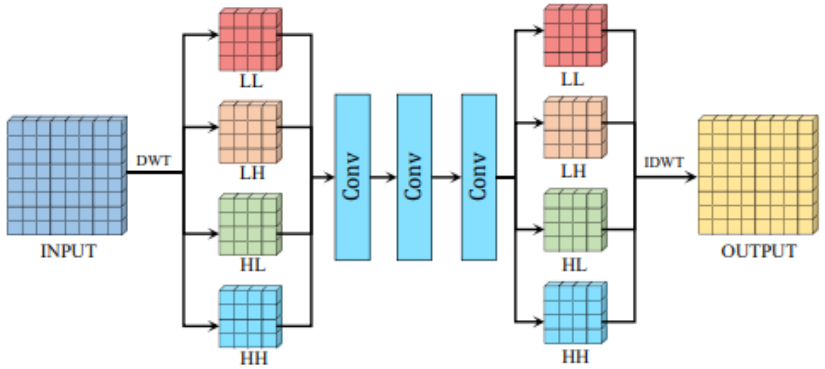
Methodology

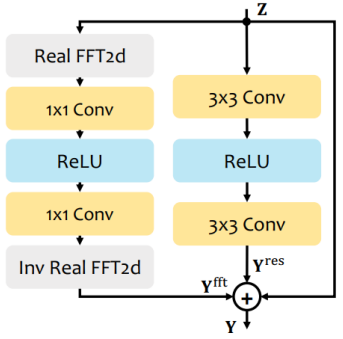
→Revisit previous successful image restoration methods

✓Multi-scale learning

✓Spatial attention

✓Frequency module

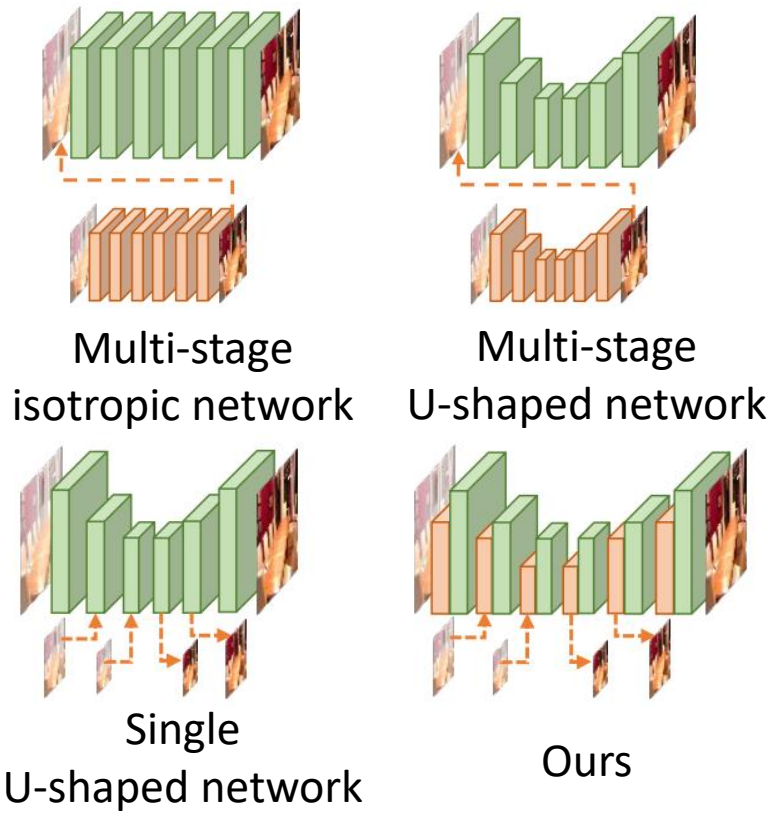




Zou, Wenbin, et al. "SDWNet: A straight dilated network with wavelet transformation for image deblurring." ICCVW-21.
Mao, Xintian, et al. "Intriguing findings of frequency selection for image deblurring." *arXiv e-prints* (2021): arXiv-2111.

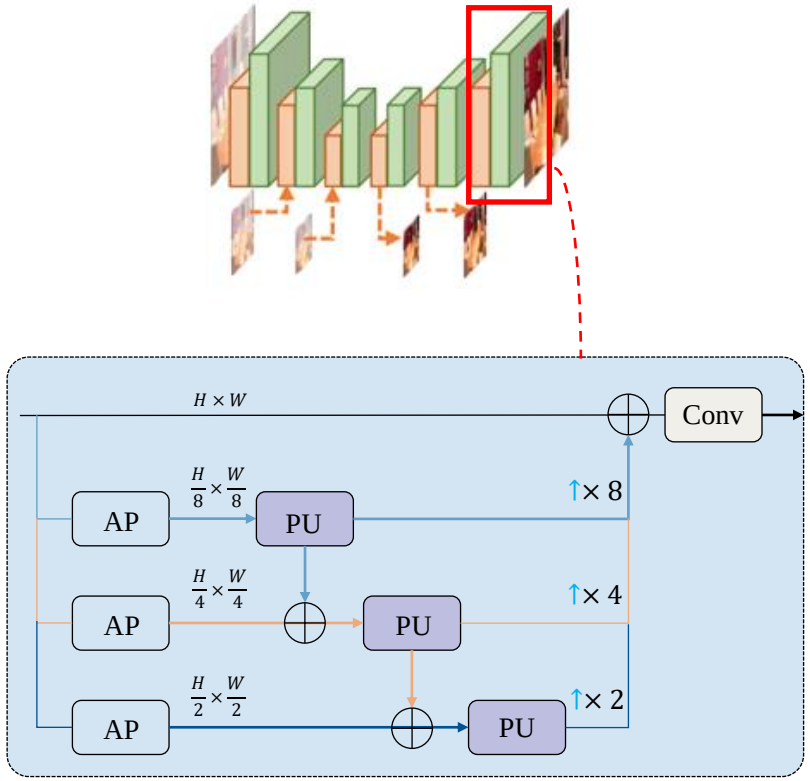
Methodology

✓Multi-scale learning



Methodology

✓Multi-scale learning



PU=Processing Unit
AP=Average Pooling

Methodology

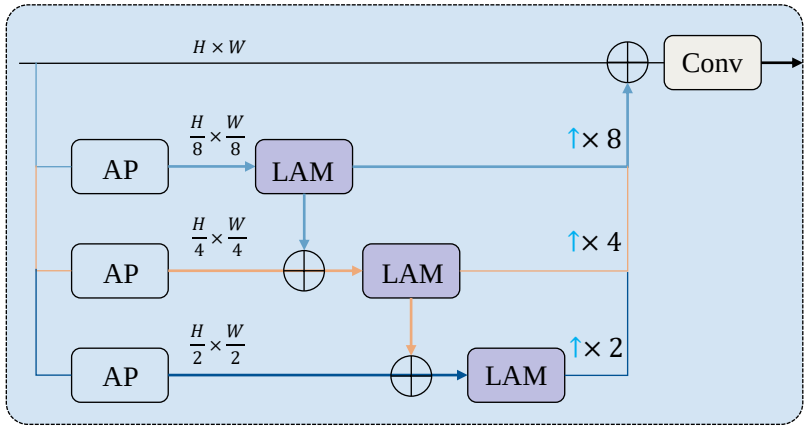
✓Multi-scale learning

✓Spatial attention

Self-attention→adaptive to the input
& quadratic complexity

Convolution→static filter
& efficient

Dynamic convolution→Softmax × Tanh √



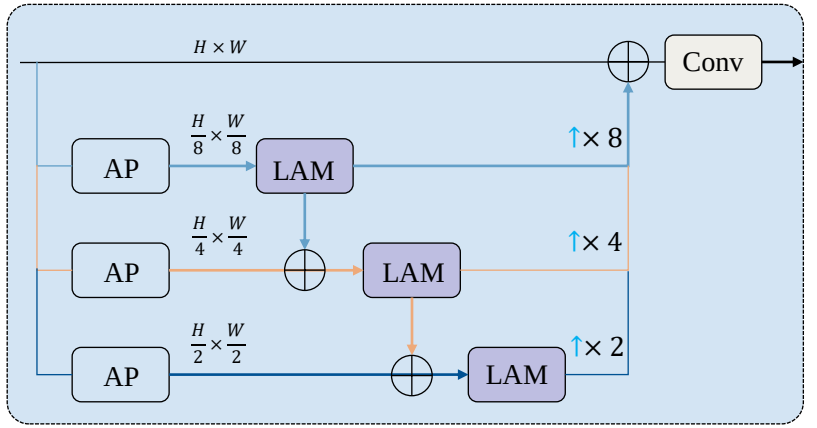
LAM=Local Attention Module
AP=Average Pooling

Methodology

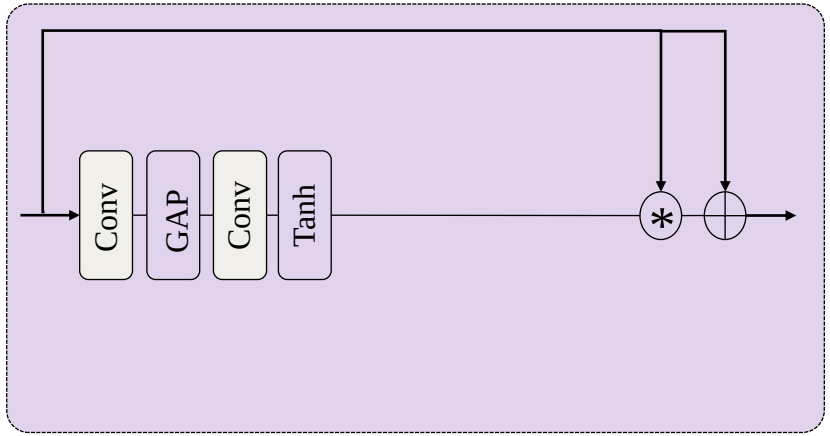
✓Multi-scale learning

✓Spatial attention

Self-attention→adaptive to the input
& quadratic complexity
Convolution→static filter
& efficient
Dynamic convolution→Softmax × Tanh √



LAM=Local Attention Module
AP=Average Pooling



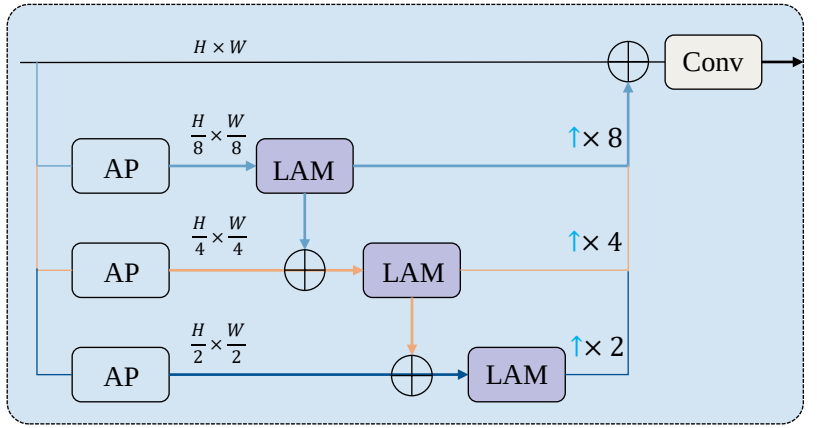
∗ Convolution

Methodology

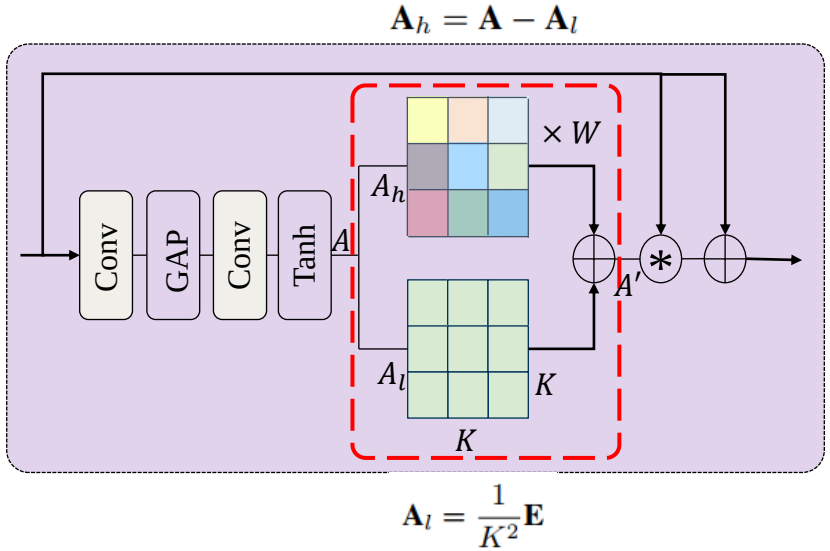
✓Multi-scale learning

✓Spatial attention

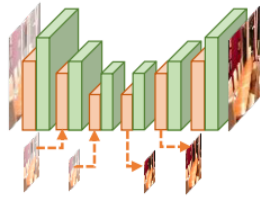
Self-attention→adaptive to the input
& quadratic complexity
Convolution→static filter
& efficient
Dynamic convolution→Softmax × Tanh √
& with frequency processing

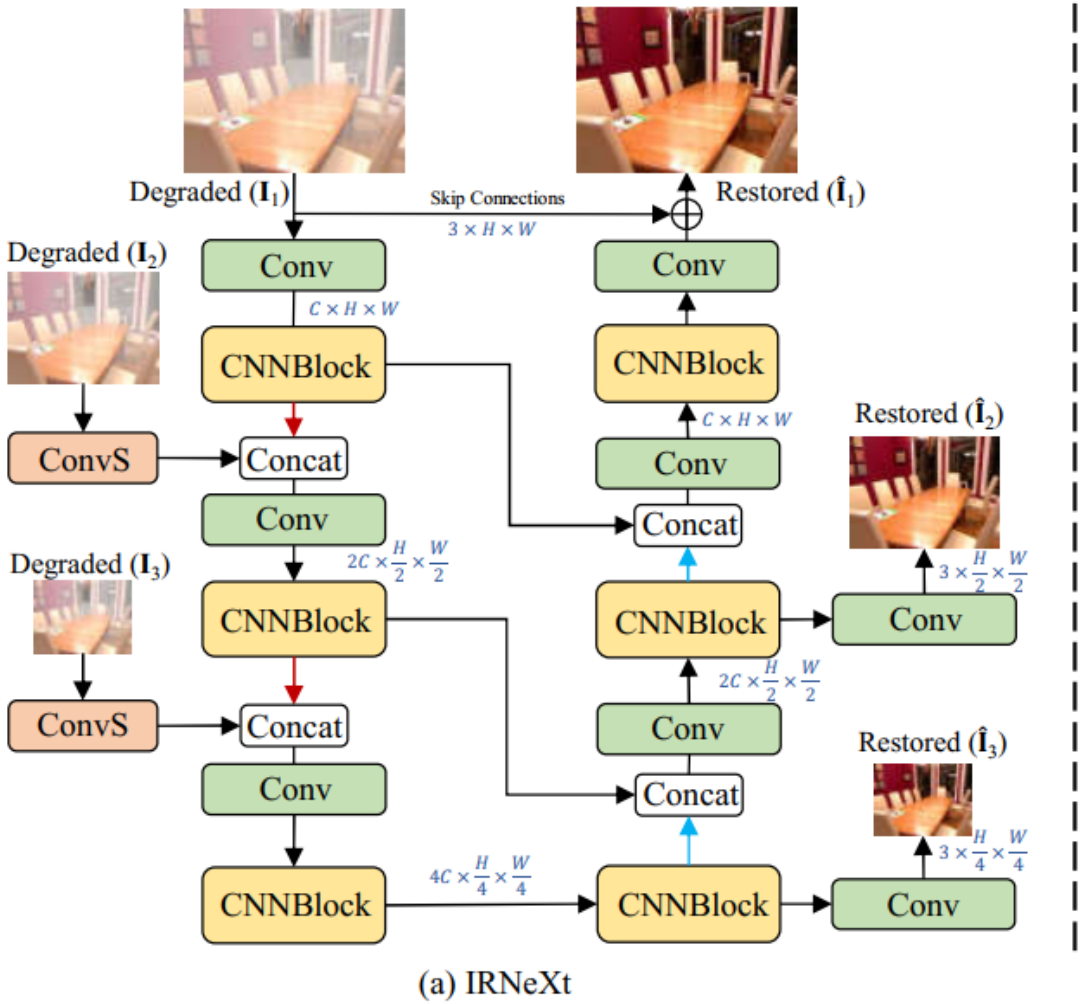


LAM=Local Attention Module
AP=Average Pooling



Architecture





(a) IRNeXt (b) ConvS (c) CNNBlock (d) Multi-Scale Module (MSM) (e) Local Attention Module (LAM)

Experiments

Image defocus deblurring

Method	Indoor Scenes				Outdoor Scenes				Combined			
	PSNR↑	SSIM↑	MAE↓	LPIPS↓	PSNR↑	SSIM↑	MAE↓	LPIPS↓	PSNR↑	SSIM↑	MAE↓	LPIPS↓
DPDNet (Abuolaim & Brown, 2020)	26.54	0.816	0.031	0.239	22.25	0.682	0.056	0.313	24.34	0.747	0.044	0.277
KPAC (Son et al., 2021)	27.97	0.852	0.026	0.182	22.62	0.701	0.053	0.269	25.22	0.774	0.040	0.227
IFAN (Lee et al., 2021)	28.11	0.861	0.026	0.179	22.76	0.720	0.052	0.254	25.37	0.789	0.039	0.217
DeepRFT (Mao et al., 2021)			-				-		25.71	0.801	0.039	0.218
DRBNet (Ruan et al., 2022)			-				-		25.73	0.791	-	0.183
Restormer (Zamir et al., 2022)	28.87	**0.882**	0.025	**0.145**	23.24	0.743	0.050	**0.209**	25.98	0.811	0.038	**0.178**
IRNeXt (Ours)	**29.22**	0.879	**0.024**	0.167	**23.53**	**0.752**	**0.049**	0.244	**26.30**	**0.814**	**0.037**	0.206

DPDD

Experiments

Image dehazing

Method	SOTS-Indoor		SOTS-Outdoor		Dense-Haze		NH-HAZE		#Param (M)
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
AOD-Net (Li et al., 2017)	20.51	0.816	24.14	0.920	13.14	0.414	15.40	0.569	0.002
GridDehazeNet (Liu et al., 2019)	32.16	0.984	30.86	0.982	13.31	0.368	13.80	0.537	0.956
MSBDN (Dong et al., 2020)	33.67	0.985	33.48	0.982	15.37	0.486	19.23	0.706	31.35
PFDN (Dong & Pan, 2020)	32.68	0.976	-		-		-		11.27
FFA-Net (Qin et al., 2020)	36.39	0.989	33.57	0.984	14.39	0.452	19.87	0.692	4.456
KDDN (Hong et al., 2020)	34.72	0.985	33.57	0.984	14.28	0.407	17.39	0.590	5.99
AECR-Net (Wu et al., 2021)	37.17	0.990	-		15.80	0.466	19.88	0.717	2.611
DeHamer (Guo et al., 2022)	36.63	0.988	35.18	0.986	16.62	0.560	**20.66**	0.684	132.45
DehazeFormer-L (Song et al., 2022)	40.05	**0.996**	-		-		-		25.44
MAXIM (Tu et al., 2022)	38.11	0.991	34.19	0.985	-		-		14.1
FSDGN (Yu et al., 2022)	38.63	0.990	-		16.91	0.581	19.99	0.731	2.73
PMNet (Ye et al., 2022)	38.41	0.990	34.74	0.985	16.79	0.510	20.42	0.730	18.90
IRNeXt (Ours)	**41.21**	**0.996**	**39.18**	**0.996**	**17.60**	**0.659**	20.55	**0.813**	5.46

Experiments

Image desnowing

Method	CSD		SRRS		Snow100K	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DesnowNet (Liu et al., 2018a)	20.13	0.81	20.38	0.84	30.50	0.94
CycleGAN (Engin et al., 2018)	20.98	0.80	20.21	0.74	26.81	0.89
All in One (Li et al., 2020)	26.31	0.87	24.98	0.88	26.07	0.88
JSTASR (Chen et al., 2020)	27.96	0.88	25.82	0.89	23.12	0.86
HDCW-Net (Chen et al., 2021c)	29.06	0.91	27.78	0.92	31.54	0.95
SMGARN (Cheng et al., 2022)	31.93	0.95	29.14	0.94	31.92	0.93
MSP-Former (Chen et al., 2022)	33.75	0.96	30.76	0.95	33.43	**0.96**
TransWeather (Valanarasu et al., 2022)	31.76	0.93	28.29	0.92	31.82	0.93
IRNeXt (Ours)	**37.29**	**0.99**	**31.91**	**0.98**	**33.61**	0.95

11.45 dB Input | PSNR GT | 11.90 dB DesnowNet | 18.30 dB JSTASR | 31.24 dB HDCW-Net | **40.74 dB IRNeXt**

Experiments

Image deraining

Method	Rain100L		Rain100H		Test100	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DerainNet (Fu et al., 2017a)	27.03	0.884	14.92	0.592	22.77	0.810
SEMI (Wei et al., 2019)	25.03	0.842	16.56	0.486	22.35	0.788
DIDMDN (Wei et al., 2019)	25.23	0.741	17.35	0.524	22.56	0.818
UMRL (Yasarla & Patel, 2019)	29.18	0.923	26.01	0.832	24.41	0.829
RESCAN (Li et al., 2018b)	29.80	0.881	26.36	0.786	25.00	0.835
PreNet (Ren et al., 2019)	32.44	0.950	26.77	0.858	24.81	0.851
MSPFN (Jiang et al., 2020)	32.40	0.933	28.66	0.860	27.50	0.876
MPRNet (Zamir et al., 2021)	36.40	0.965	30.41	0.890	30.27	0.897
HINet (Chen et al., 2021b)	37.20	0.969	30.63	0.893	30.26	0.905
DRT (Liang et al., 2022)	37.61	0.948	29.47	0.846	27.02	0.847
MAXIM (Tu et al., 2022)	38.06	**0.977**	30.81	**0.903**	31.17	**0.922**
IRNeXt (Ours)	**38.14**	0.972	**31.64**	0.902	**31.53**	0.919

13.84 dB Input | PSNR GT | 27.57 dB UMRL | 37.41 dB MPRNet | 32.22 dB MAXIM | **37.76 dB IRNeXt**

Experiments

Image motion deblurring

Method	PSNR	SSIM	FLOPs/G	Params/M	Time/s
DBGAN (Zhang et al., 2020)	31.10	0.942	759.85	**11.6**	1.447
DMPHN (Zhang et al., 2019a)	31.20	0.940	-	21.7	0.405
MIMO-UNet++ (Cho et al., 2021)	32.68	0.959	617.64	16.1	1.277
MPRNet (Zamir et al., 2021)	32.66	0.959	777.01	20.1	1.148
Restormer (Zamir et al., 2022)	32.92	0.961	140.99	26.1	1.218
Stripformer (Tsai et al., 2022)	33.08	**0.962**	170.46	20.0	1.054
IRNeXt (Ours)	**33.16**	**0.962**	**114.79**	13.21	**0.255**

GoPro

Method	PSNR	SSIM
SRN-DeblurNet (Tao et al., 2018)	32.53	0.840
MIMO-UNet (Cho et al., 2021)	32.73	0.846
MIMO-UNet+ (Cho et al., 2021)	33.37	0.856
MPRNet (Zamir et al., 2021)	33.61	0.861
Restormer (Zamir et al., 2022)	33.69	0.863
Uformer (Wang et al., 2022)	33.98	0.866
IRNeXt (Ours)	**34.08**	**0.869**

RSBlur

Thank you!