西安电子科技大学
XIDIAN UNIVERSITY

IAPR ICPR 2018
INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
In Beijing, China — August 20th-24th 2018

TWO-STAGE CONVOLUTIONAL NETWORK FOR IMAGE SUPER-RESOLUTION

{ZHENG HUI, XIUMEI WANG AND XINBO GAO}
SCHOOL OF ELECTRONIC ENGINEERING XIDIAN UNIVERSITY

ABSTRACT


Low-resolution image

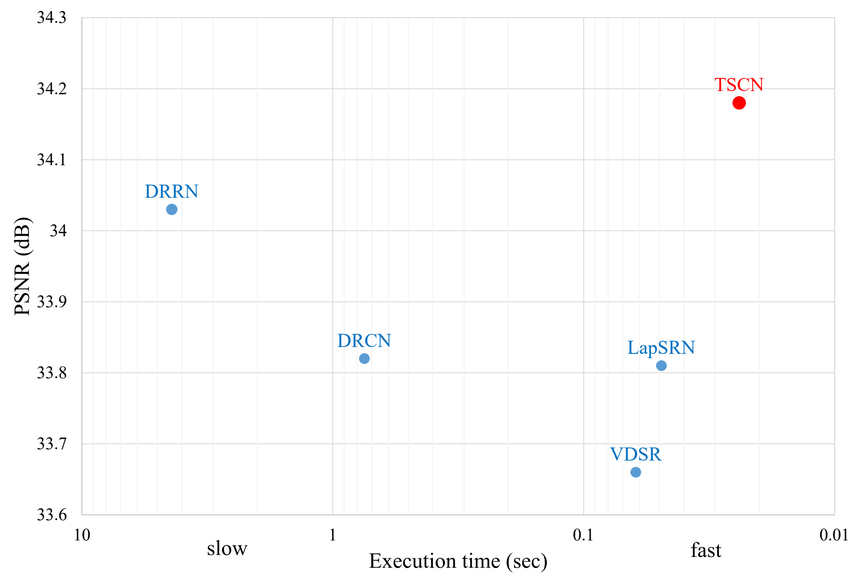
Super-Resolution


High-resolution image

- Deep convolutional neural networks have recently advanced the state-of-the-art on the issue of single image super-resolution.
- An accurate and lightweight deep model is devised for image super-resolution.
  - The proposed method constructs the multi-path information fusion (MIF) module to collect abundant information from the feature maps of the input, output and intermediary in a module.
  - To improve the SR performance, a refinement network with local residual topology architecture is introduced.
  - To speed up the inference, the proposed approach adopts less number of filters.
- Experimental results show that the proposed method achieves fast inference time and state-of-the-art SR results on four benchmark datasets simultaneously.

EXECUTION TIME

The average PSNR and inference time for upscaling 3 × on Set5



AVERAGE PSNR/SSIM/IFC RESULTS

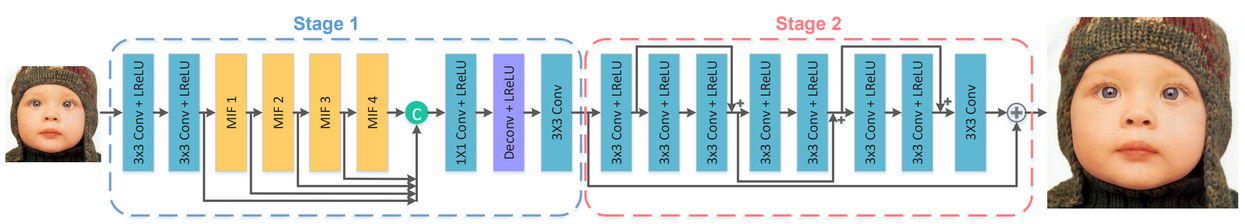
Method	Scale	Set5 PSNR / SSIM / IFC	Set14 PSNR/SSIM/IFC	BSD100 PSNR/SSIM/IFC	Urban100 PSNR/SSIM/IFC
Bicubic	× 2	33.66 / 0.9299 / 6.083	30.24 / 0.8688 / 6.105	29.56 / 0.8431 / 5.619	26.88 / 0.8403 / 6.245
FSRCNN		37.00 / 0.9558 / 8.047	32.63 / 0.9088 / 7.731	31.50 / 0.8906 / 7.082	29.85 / 0.9009 / 8.026
VDSR		37.53 / 0.9587 / 8.580	33.03 / 0.9124 / 8.159	31.90 / 0.8960 / 7.494	30.76 / 0.9140 / 8.629
DRCN		37.63 / 0.9588 / 8.783	33.04 / 0.9118 / 8.370	31.85 / 0.8942 / 7.577	30.75 / 0.9133 / 8.959
LapSRN		37.52 / 0.9591 / 9.010	32.99 / 0.9124 / 8.501	31.80 / 0.8952 / 7.715	30.41 / 0.9103 / 8.907
DRRN		37.74 / 0.9591 / 8.670	33.23 / 0.9136 / 8.280	32.05 / 0.8973 / 7.513	31.23 / 0.9188 / 8.889
TSCN		**37.88 / 0.9602 / 9.175**	**33.28 / 0.9147 / 8.729**	**32.09 / 0.8985 / 7.871**	**31.29 / 0.9198 / 9.442**
Bicubic	× 3	30.39 / 0.8682 / 3.580	27.55 / 0.7742 / 3.473	27.21 / 0.7385 / 3.138	24.46 / 0.7349 / 3.620
FSRCNN		33.16 / 0.9140 / 4.964	29.43 / 0.8242 / 4.549	28.52 / 0.7893 / 4.030	26.42 / 0.8064 / 4.842
VDSR		33.66 / 0.9213 / 5.203	29.77 / 0.8314 / 4.691	28.82 / 0.7976 / 4.151	27.14 / 0.8279 / 5.159
DRCN		33.82 / 0.9226 / 5.336	29.76 / 0.8311 / 4.782	28.80 / 0.7963 / 4.184	27.15 / 0.8276 / 5.314
LapSRN		33.81 / 0.9220 / 5.194	29.79 / 0.8325 / 4.662	28.82 / 0.7980 / 4.057	27.07 / 0.8275 / 5.156
DRRN		34.03 / 0.9244 / 5.394	29.96 / 0.8349 / 4.870	**28.95** / 0.8004 / 4.235	**27.53 / 0.8378** / 5.440
TSCN		**34.18 / 0.9256 / 5.544**	**29.99 / 0.8351 / 4.970**	**28.95 / 0.8012 / 4.350**	27.46 / 0.8362 / **5.559**
Bicubic	× 4	28.42 / 0.8104 / 2.329	26.00 / 0.7027 / 2.237	25.96 / 0.6675 / 1.978	23.14 / 0.6577 / 2.361
FSRCNN		30.71 / 0.8657 / 2.986	27.59 / 0.7535 / 2.707	26.96 / 0.7128 / 2.359	24.60 / 0.7258 / 2.895
VDSR		31.35 / 0.8838 / 3.542	28.01 / 0.7674 / 3.106	27.29 / 0.7251 / 2.679	25.18 / 0.7524 / 3.462
DRCN		31.53 / 0.8854 / 3.543	28.02 / 0.7670 / 3.098	27.23 / 0.7233 / 2.633	25.14 / 0.7510 / 3.465
LapSRN		31.54 / 0.8852 / 3.559	28.09 / 0.7700 / 3.145	27.32 / 0.7275 / 2.677	25.21 / 0.7562 / 3.530
DRRN		31.68 / 0.8888 / 3.700	28.21 / 0.7721 / 3.249	27.38 / 0.7284 / 2.746	**25.44** / 0.7638 / 3.669
TSCN		**31.82 / 0.8907 / 3.766**	**28.28 / 0.7734 / 3.286**	**27.42 / 0.7301 / 2.792**	**25.44 / 0.7644 / 3.715**

TIME PERFORMANCE

Methods	Scale	Set5 PSNR / TIME	Set14 PSNR / TIME	BSD100 PSNR / TIME	Urban100 PSNR / TIME
TSCN_I	×2	37.87 / 0.017	33.26 / 0.028	32.08 / 0.017	31.23 / 0.071
TSCN	×2	37.88 / 0.028	33.28 / 0.046	32.09 / 0.028	31.29 / 0.130
TSCN_I	×3	34.14 / 0.013	29.96 / 0.018	28.93 / 0.012	27.38 / 0.043
TSCN	×3	34.18 / 0.025	29.99 / 0.037	28.95 / 0.023	27.46 / 0.104
TSCN_I	×4	31.78 / 0.011	28.25 / 0.015	27.41 / 0.010	25.40 / 0.034
TSCN	×4	31.82 / 0.023	28.28 / 0.034	27.42 / 0.021	25.44 / 0.094

CODE

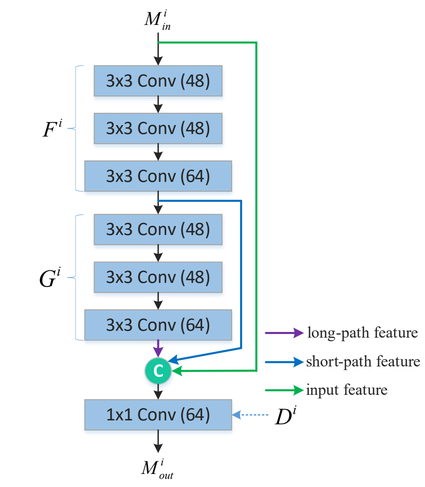
FRAMEWORK



Loss function $\mathcal{L}_{MAE} = \frac{1}{N}\sum_{i=1}^{N}\|I_i - \hat{I}_i\|_1$ — *Training*

$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N}\|I_i - \hat{I}_i\|_2^2$ — *Fine-tuning*

DETAILS OF MIF MODULE



The distillation unit of *i*-th module

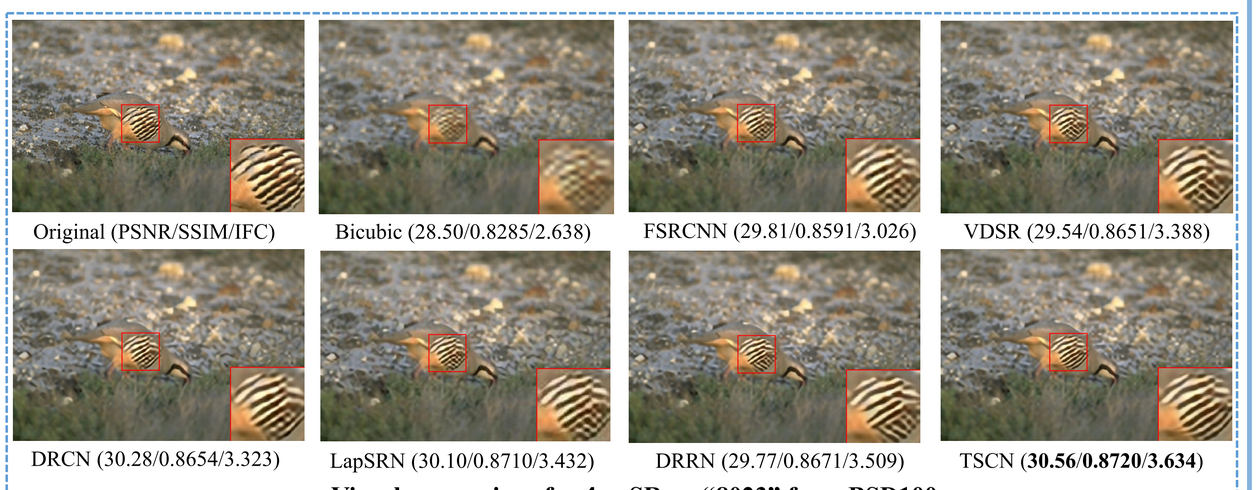
$$M_{out}^i = D^i\left(\left[M_{in}^i, F^i(M_{in}^i), G^i\left(F^i(M_{in}^i)\right)\right]\right)$$

the input

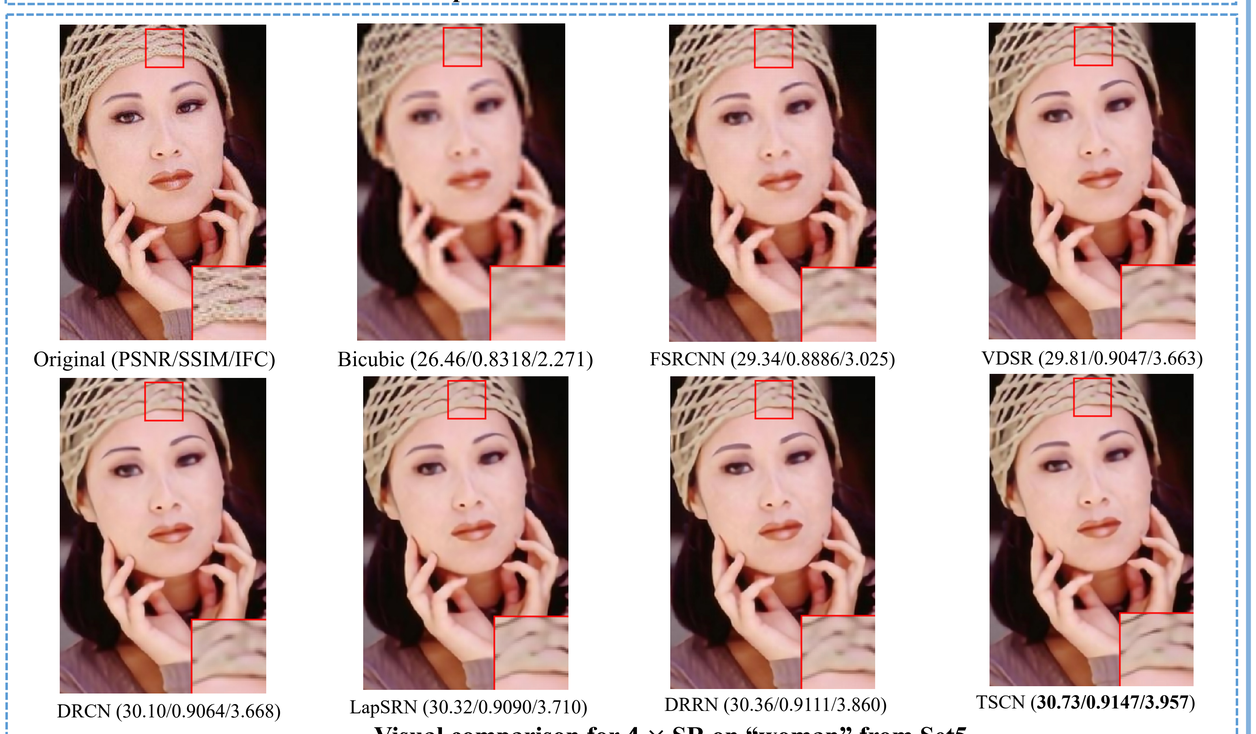
the first three layers in a module

the second three layers in a module

VISUAL RESULTS


Original (PSNR/SSIM/IFC)
Bicubic (28.50/0.8285/2.638)
FSRCNN (29.81/0.8591/3.026)
VDSR (29.54/0.8651/3.388)
DRCN (30.28/0.8654/3.323)
LapSRN (30.10/0.8710/3.432)
DRRN (29.77/0.8671/3.509)
TSCN (**30.56/0.8720/3.634**)

Visual comparison for 4 × SR on "8023" from BSD100.


Original (PSNR/SSIM/IFC)
Bicubic (26.46/0.8318/2.271)
FSRCNN (29.34/0.8886/3.025)
VDSR (29.81/0.9047/3.663)
DRCN (30.10/0.9064/3.668)
LapSRN (30.32/0.9090/3.710)
DRRN (30.36/0.9111/3.860)
TSCN (**30.73/0.9147/3.957**)

Visual comparison for 4 × SR on "woman" from Set5.

TRAINING DETAILS

Item	Detail
Input channels	1 (Y)
Training images	291
Mini-batch size	64
Learning rate	10^{-4}, halved at every 5×10^5 iterations
Data augmentation	Rotate, Flip and Downscale (×40)
Optimizer	ADAM ($\beta_1 = 0.9$)
Dataset	91 images + BSD200
Implementation	Caffe

Patch size

Scale	Patch size
2	$35^2/70^2$
3	$25^2/75^2$
4	$19^2/76^2$

Kernel size of transposed convolution

Scale	Kernel size
2	4×4
3	5×5
4	8×8