Image Super-Resolution Using Deep Feedforward Neural Networks in Spectral Domain

26.03.2018

MS Thesis Presentation

Presenter: Onur Aydın

Advisor: Assoc. Prof. Selim Aksoy & Asst. Prof. R. Gökberk Cinbiş

Outline

- 1. Introduction
- 2. Related Work
  - Deep Learning and Machine Learning Background
  - Related Work on Super Resolution
- 3. Method
  - Super-Resolution in Spectral Domain
  - Artifact Reduction in Spatial Domain
- 4. Experiments and Results
- 5. Concluding Remarks and Future Works

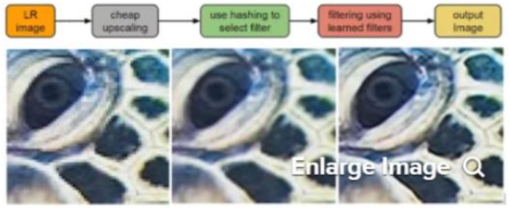
PHOTOGRAPHY

Google AI expands your photos to shrink your mobile data usage

The internet giant develops a machine learning method to improve image details and cut Google+ network burden by a third.

BY STEPHEN SHANKLAND / JANUARY 12, 2017 7:00 PM PST

It's fun to mock image-processing impossibilities where "Blade Runner" or CSI investigators zoom into photos to see far more detail than a photo could possibly have recorded.



Google's RAISR uses a multistep process to intelligently guess how to expand images without just producing a blurry larger version.

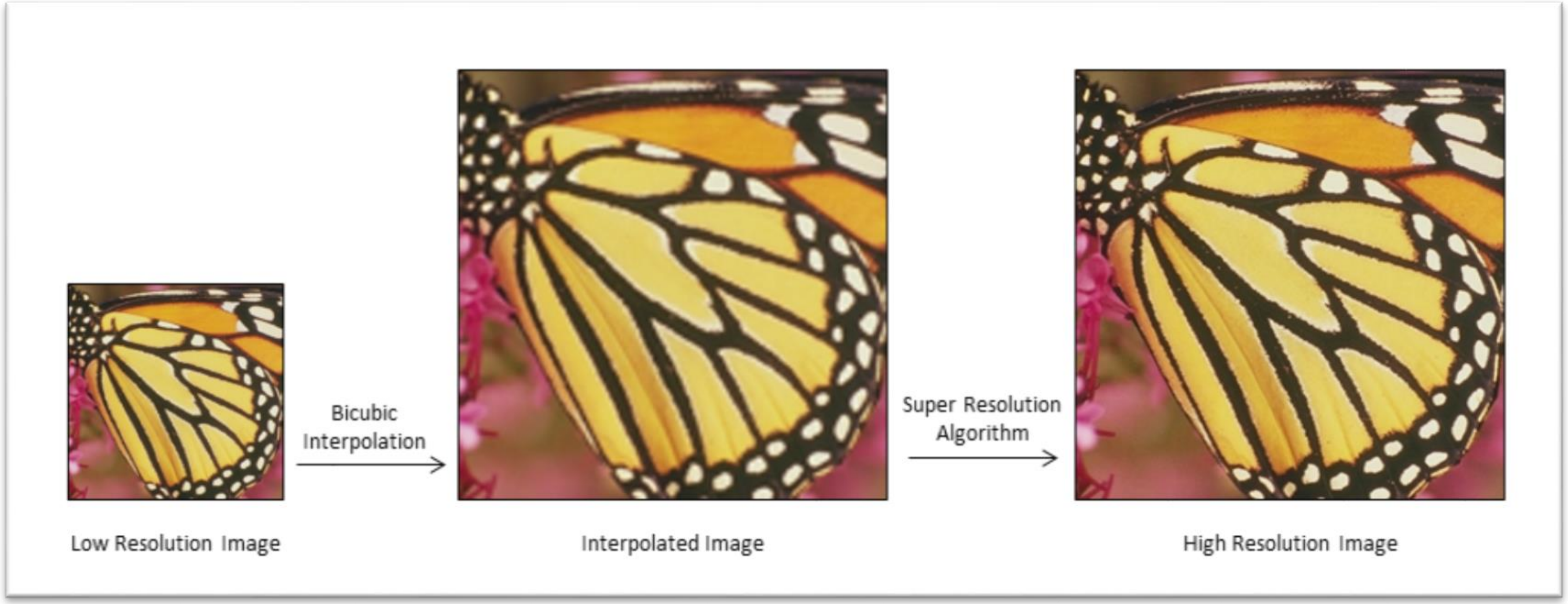
Google

But guess what? Not only does Google have technology that can do something like that kind of photo enhancement, it's also using it to keep you from gobbling through your mobile phone data plan.

1. Introduction

Single-Image Super-Resolution Problem:

- Learning a mapping from LR images to HR images with minimum perceptual loss
- Super-resolution is an ill-posed inverse problem
- Multiple high-resolution image candidates can be produced from single low-resolution image



1. Introduction

Single-Image Super-Resolution Problem:

- Problems about Super-Resolution Problem:
  - Image Super-Resolution or Image Completion?
  - Infinitely many Solutions
  - Metric Problems
  - Consecutive Networks, Pre-trained Networks
  - Ensemble Learning, Post-processing, Pre-processing

1. Introduction

Single-Image Super-Resolution Problem:

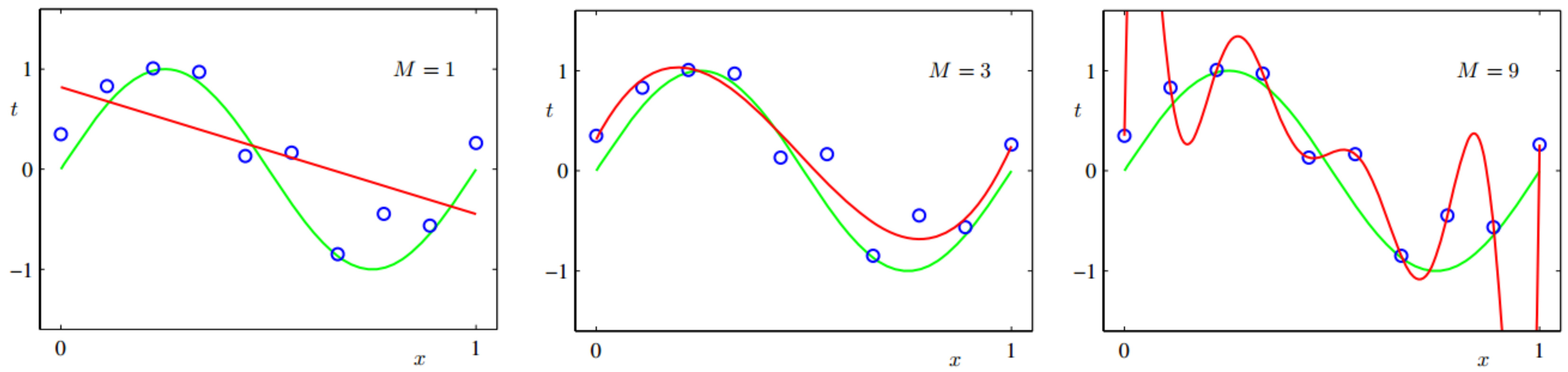
- Three important factors for Super-Resolution Problem:

- Precision
- Efficiency
- Flexibility of Scale Factor

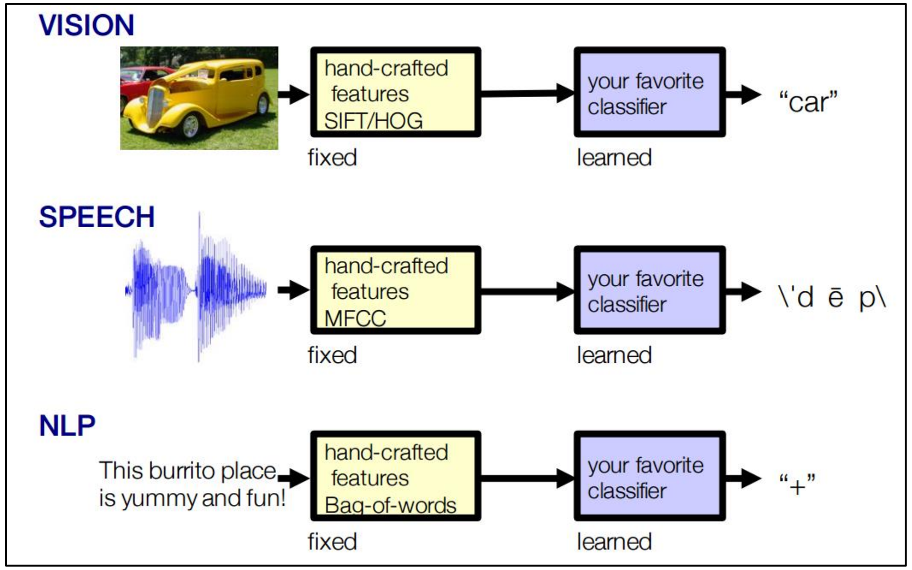
2. Related Work – Machine Learning

'Machine Learning gives computers the ability to learn without being explicitly programmed.'

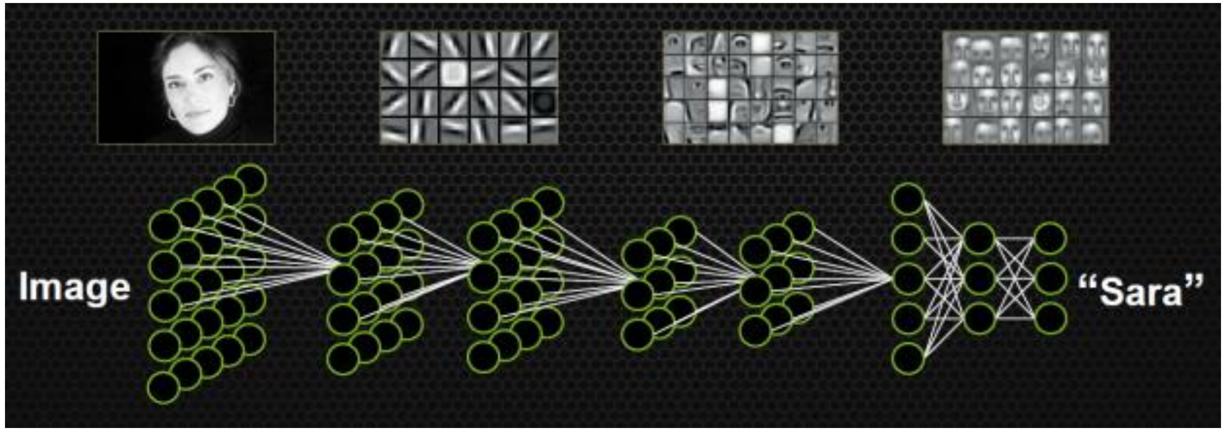
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- Semi-supervised Learning, Active Learning, One-shot Learning…



2. Related Work – Deep Learning



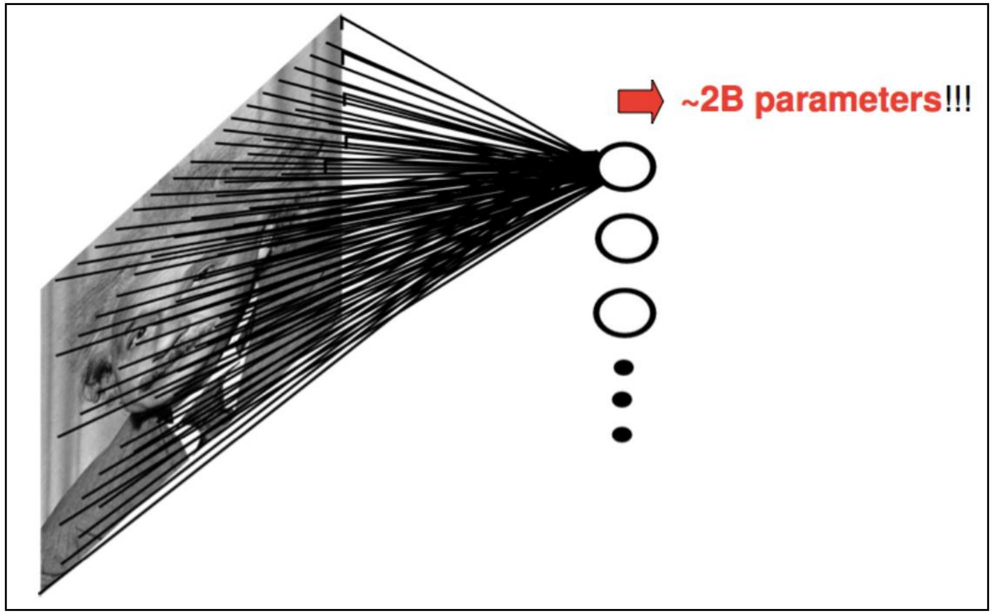
Traditional Machine Learning



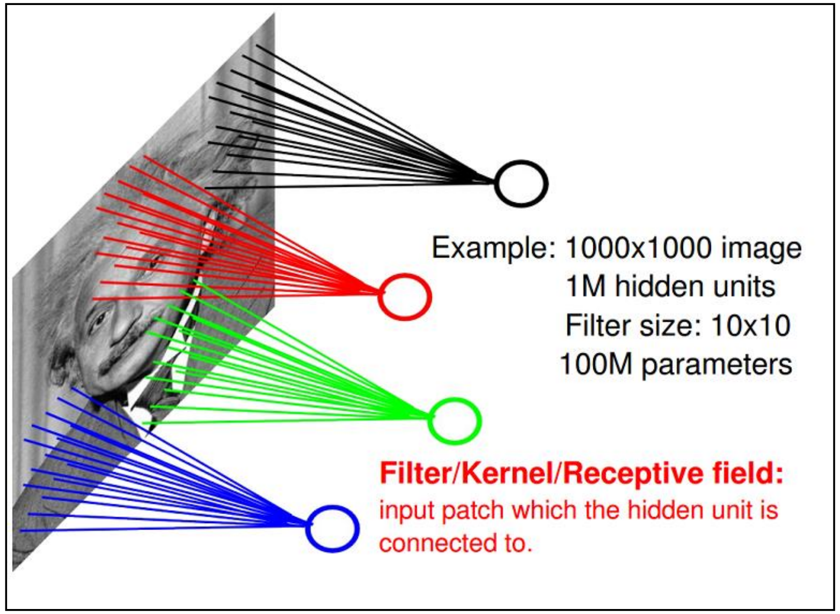
Deep Learning (End-to-end Learning)

2. Related Work – Deep Learning

Feedforward Network



Convolutional Network



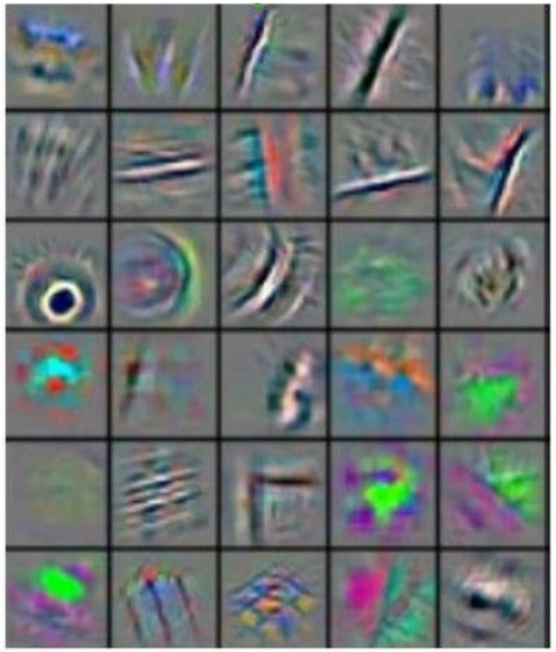
Matrix Multiplication

Convolution Operation

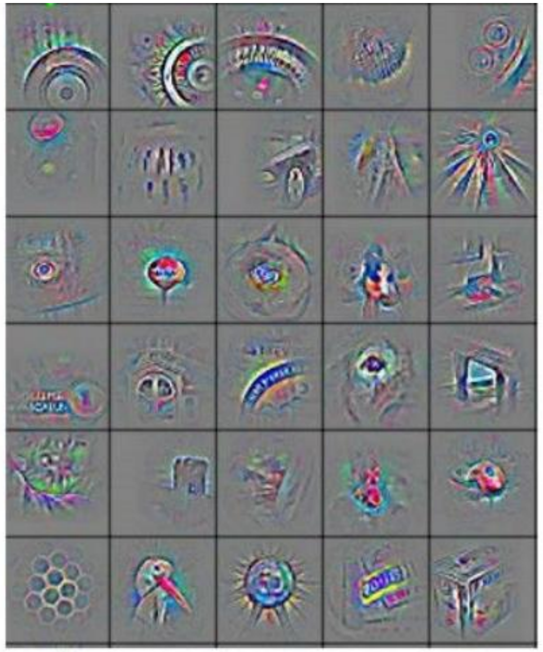
2. Related Work – Deep Learning

The Filters learned by a CNN for a classification problem

Low level Structures



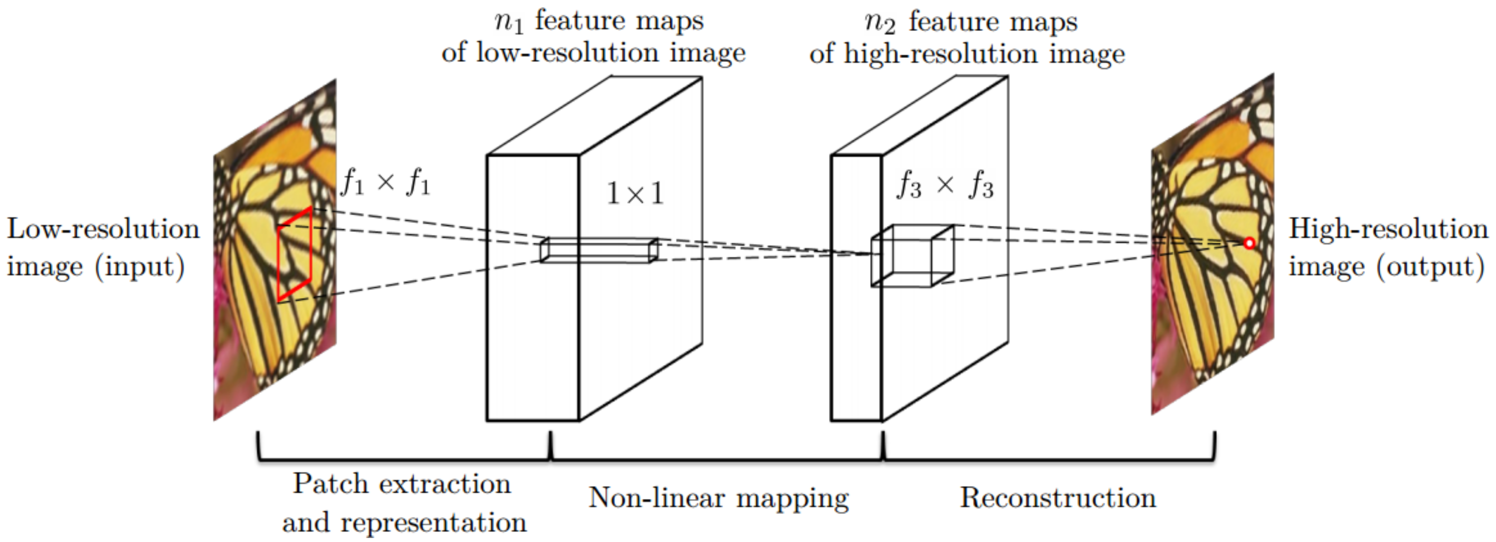
More Complex Structures



Highly Complex, Abstract Concepts

2. Related Work – Super Resolution

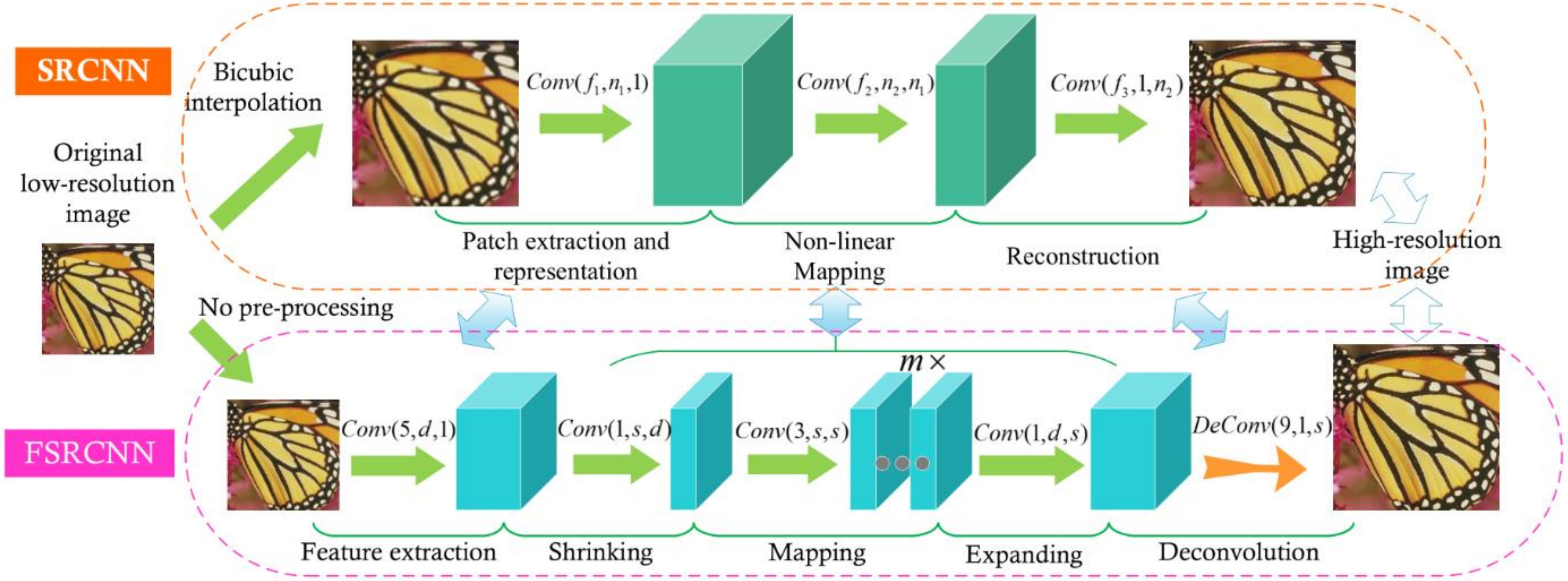
- SRCNN
  - First Deep Learning Architecture in the Literature
  - 3 Convolutional Layers are used
  - Bicubic interpolation is applied before reconstruction
  - L2 Loss function is used
  - For each scale factor, a separate network is trained



C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 2, pp. 295{307, 2016.

2. Related Work – Super Resolution

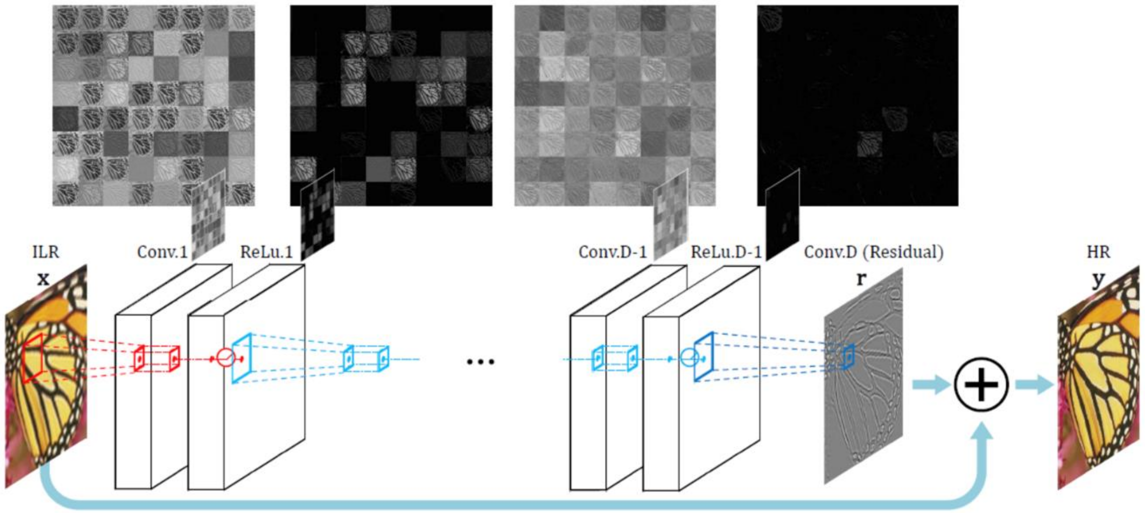
- FSRCNN
  - Faster and deeper version of SRCNN
  - Bicubic Interpolation is not used in preprocessing step
  - Instead, a Deconvolutional Layer is learned in final layer for upsampling
  - 8 Convolutional Layers are used



C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in European Conference on Computer Vision, pp. 391{407, Springer, 2016.

2. Related Work – Super Resolution

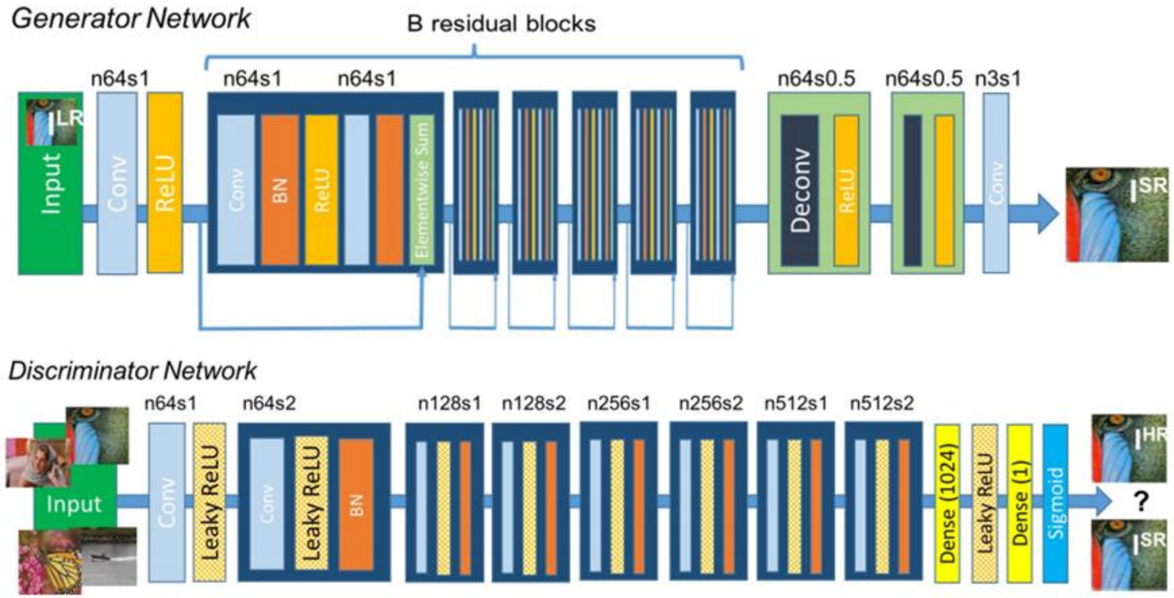
- VDSR
  - Exactly deeper version of SRCNN
  - Uses 20-layer convolutional neural network
  - Uses one residual connection from input layer to output layer
  - Trained with higher learning rate than SRCNN
  - Only one network is trained for different scale factors using scale augmentation
  - Bicubic interpolation is done in preprocessing step



J. Kim, J. Kwon Lee, and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1646{1654, 2016.

2. Related Work – Super Resolution

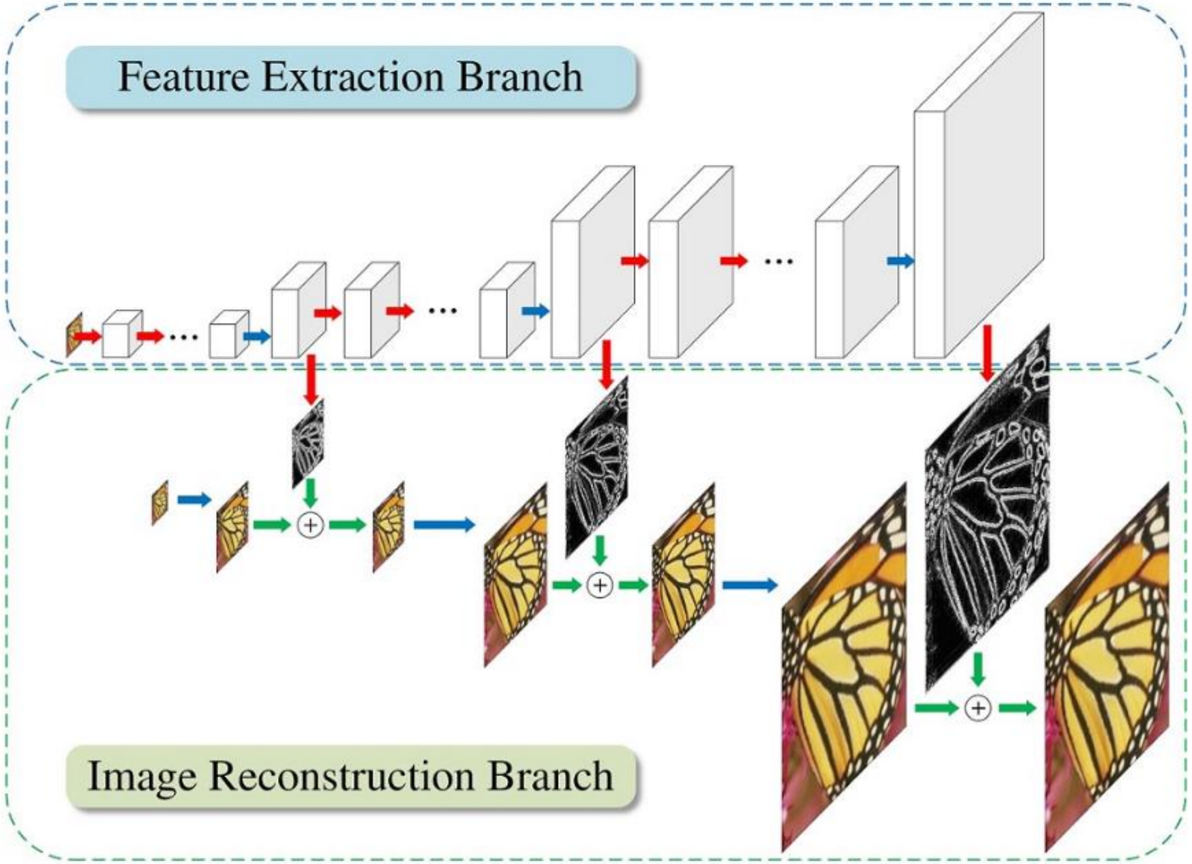
- SRGAN
  - Super resolution network is trained in adversarial setting
  - More complex network architectures are used for generator and discriminator networks than SRCNN
  - Generator Network learns to generate high-resolution image from low-resolution one
  - Discriminator Network tries to discriminate generated images from real images



C. Ledig, L. Theis, F. Husz´ar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., "Photo-realistic single image super-resolution using a generative adversarial network," 2016.

2. Related Work – Super Resolution

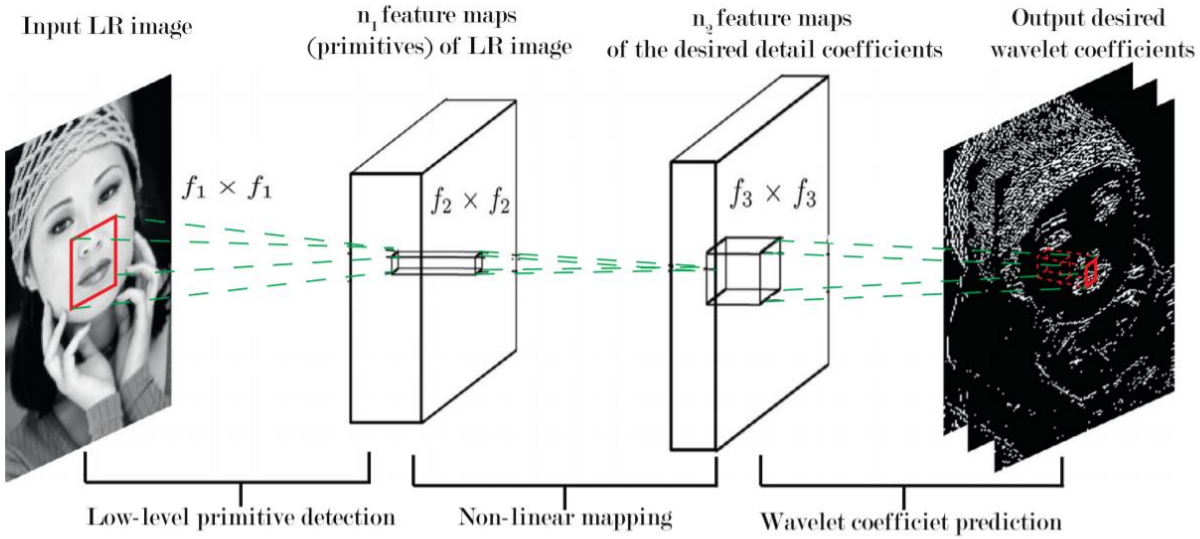
- LapSRN
  - Progressive reconstruction is done
  - No interpolation is applied in preprocessing step
  - Uses 27-layer convolutional neural network
  - Uses residual connections
  - Charbonnier loss function is used



W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep laplacian pyramid networks for fast and accurate super-resolution," arXiv preprint arXiv:1704.03915, 2017.

2. Related Work – Super Resolution

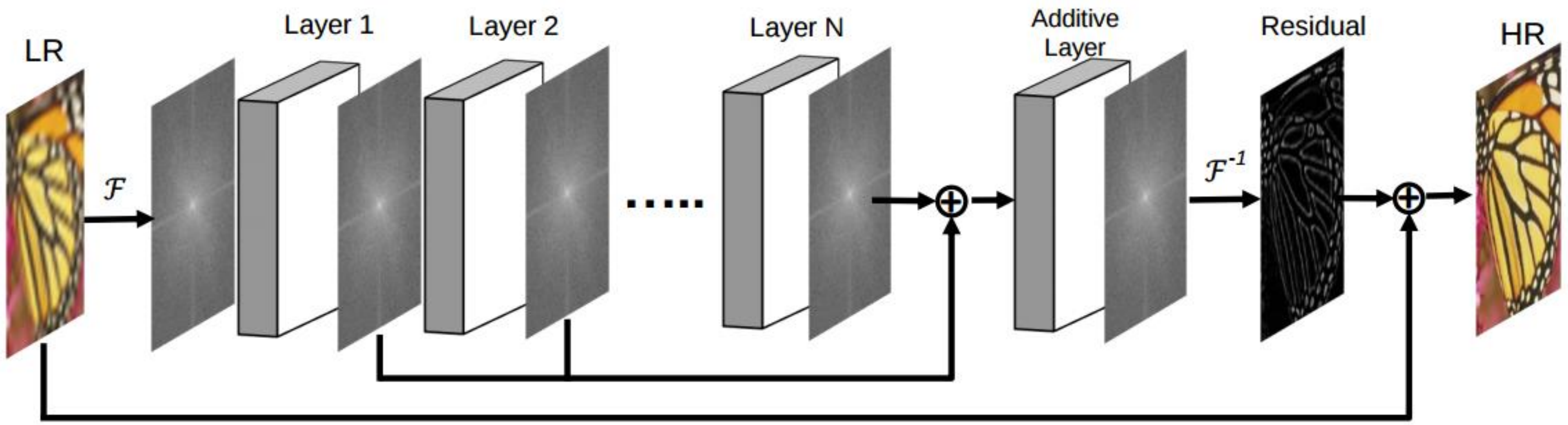
- CNNWSR
  - Super resolution network is trained for predicting discrete wavelet transform coefficients
  - Uses 3-layer convolutional neural network like SRCNN
  - Scale factor limitations due to discrete wavelet transform



N. Kumar, R. Verma, and A. Sethi, "Convolutional neural networks for wavelet domain super resolution," Pattern Recognition Letters, vol. 90, pp. 65{71, 2017.

2. Related Work – Super Resolution

- FourierSR
  - Very recent and currently **unpublished** work
  - A Super-resolution mapping is learned in Fourier Domain
  - Convolution operations are replaced with product operations
  - Additionally, a novel non-linear activation function is developed for frequency domain methods



J. Li, S. You, and A. Robles-Kelly, "A frequency domain neural network for fast image super-resolution," arXiv preprint arXiv:1712.03037, 2017

3. Method – Artifact Reduction

- Super-Resolution in Spectral Domain
- Artifact Reduction in Spatial Domain

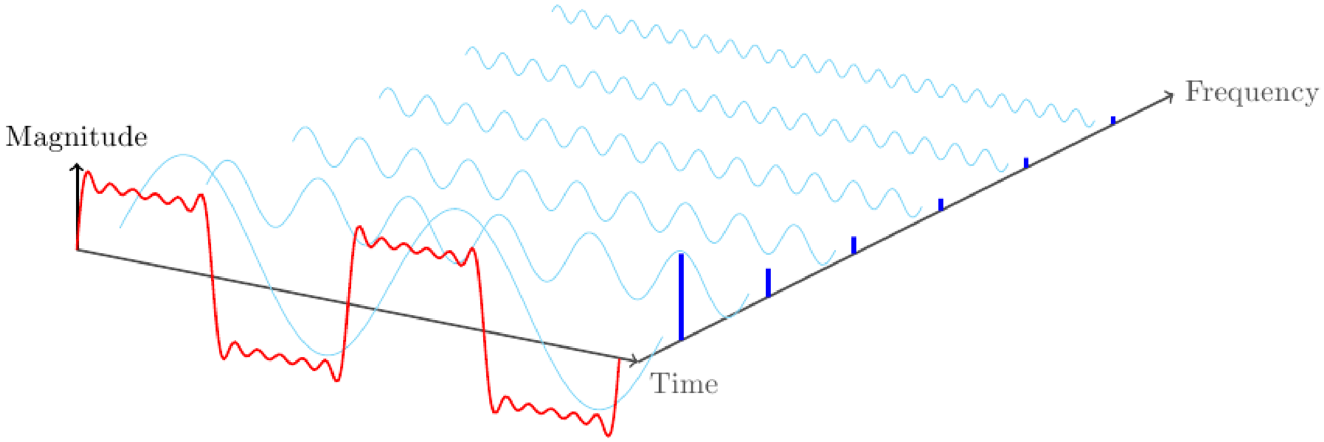
Bicubic Interpolation

Initial Super Resolution

Artifact Reduction

3. Method – Super Resolution

- Fourier Transform
  - Decomposes any differentiable real/complex signal into different complex exponential functions which oscillate at different frequencies.
  - Represents a time/spatial domain signal in frequency domain



$$F(w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-jwx}dx$$

$$f(x) = \int_{-\infty}^{\infty} F(w)e^{jwx}d\tau$$

3. Method – Super Resolution

- Discrete Fourier Transform
  - Discrete version of Fourier Transform

$$F[u] = \frac{1}{M} \sum_{y=0}^{M-1} f[x] e^{-2\pi j (\frac{ux}{M})}$$

$$f[x] = \sum_{v=0}^{M-1} F[u] e^{2\pi j (+\frac{ux}{M})}$$

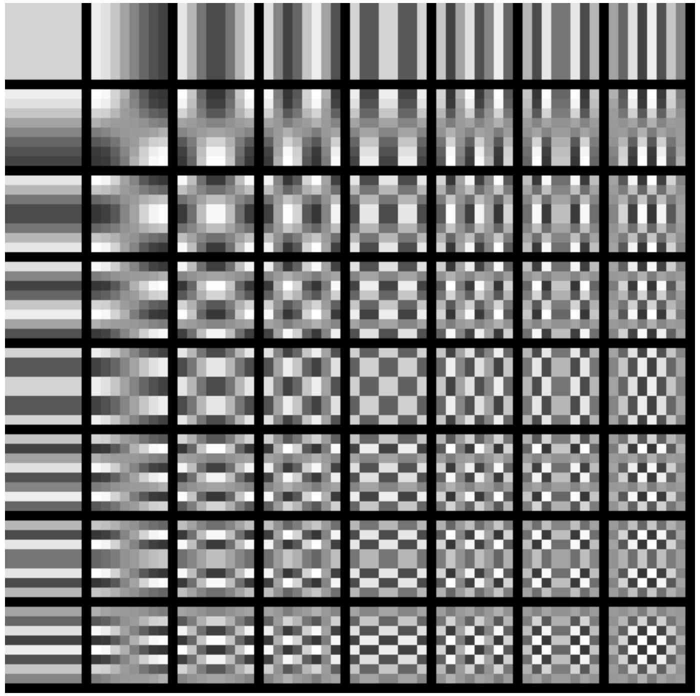
- 2D Discrete Fourier Transform
  - Discrete and 2D version of Fourier Transform

$$f[x, y] = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F[u, v] e^{2\pi j (\frac{xu}{N} + \frac{yv}{M})}$$

$$F[u, v] = \frac{1}{NM} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f[x, y] e^{-2\pi j (\frac{xu}{N} + \frac{yv}{M})}$$

3. Method – Super Resolution

- Discrete Cosine Transform
  - Very similar to Discrete Fourier Transform
  - A real-valued transform unlike complex-valued DFT
  - Basis functions are constant, real-valued cosine functions
  - Has energy compaction property
  - An approximation to Principal Component Analysis (PCA)



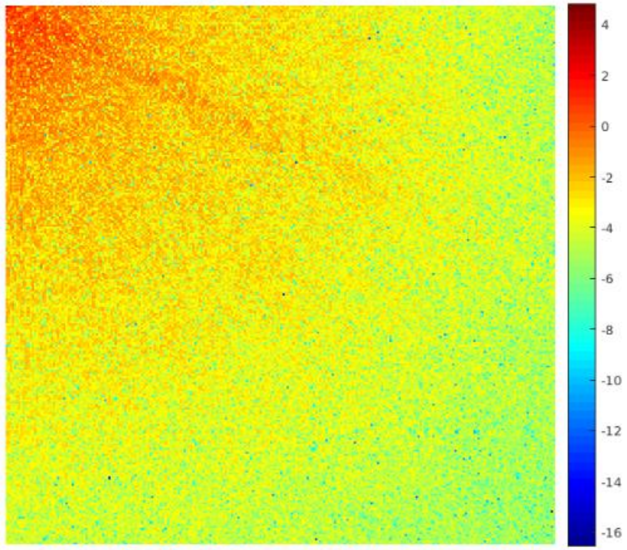
$$F[u] = a(u) \sum_{x=0}^{N-1} f[x] cos\left(\frac{\pi(2x+1)u}{2N}\right)$$

$$f[x] = \sum_{x=0}^{N-1} a(u) F[u] cos\left(\frac{\pi(2x+1)u}{2N}\right)$$

3. Method – Super Resolution

- 2D Discrete Cosine Transform

  - 2D version of DCT
  - Used in JPEG Compression

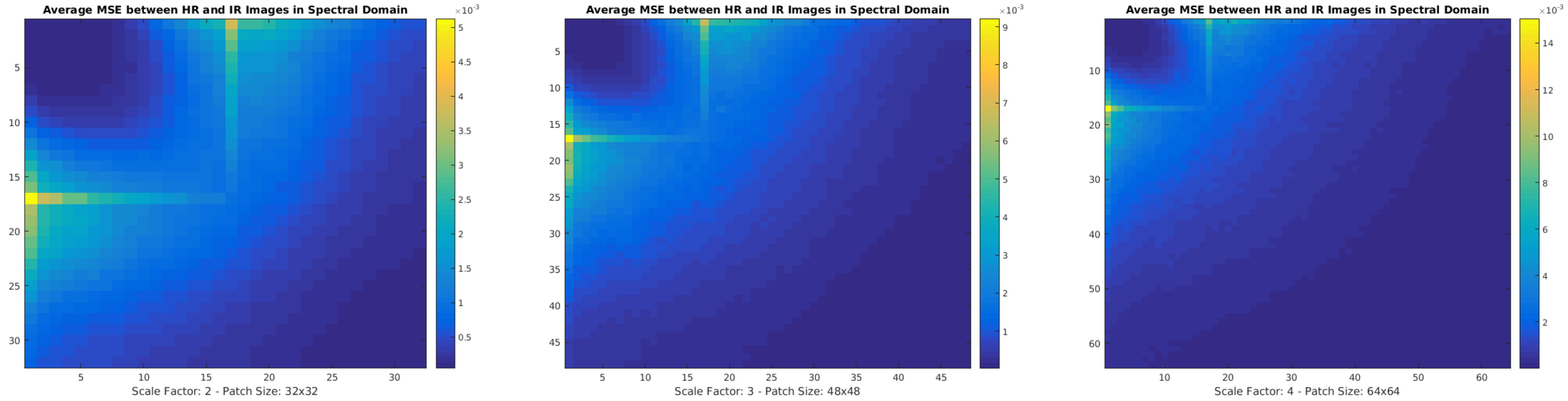


$$f[x,y] = \sum_{x=0}^{N-1}\sum_{y=0}^{M-1} a(u)a(v)F[u,v]cos\left(\frac{\pi(2x+1)u}{2N}\right)cos\left(\frac{\pi(2y+1)v}{2M}\right)$$

$$F[u,v] = a(u)a(v)\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f[x,y]cos\left(\frac{\pi(2x+1)u}{2N}\right)cos\left(\frac{\pi(2y+1)v}{2M}\right)$$

3. Method – Super Resolution

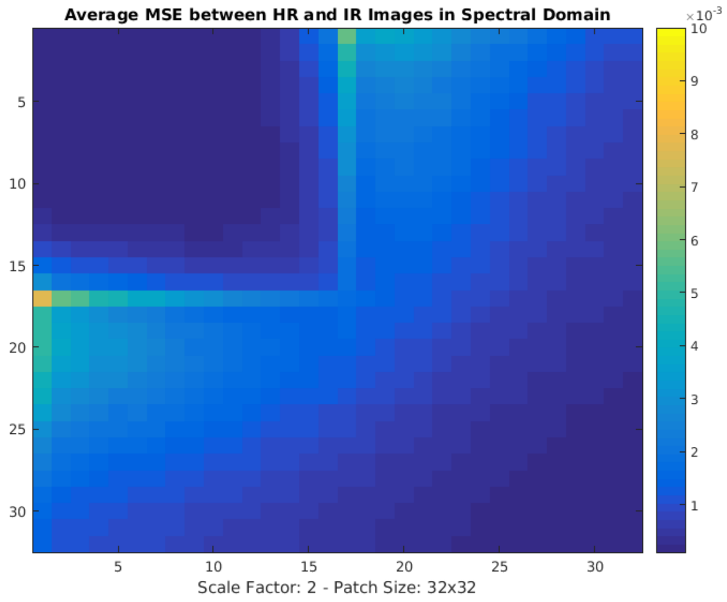
- Super-Resolution in Spectral Domain
  - Bicubic Interpolation + Divide Images into Patches + 2D DCT
  - Mean-Square Error Analysis in Spectral Domain



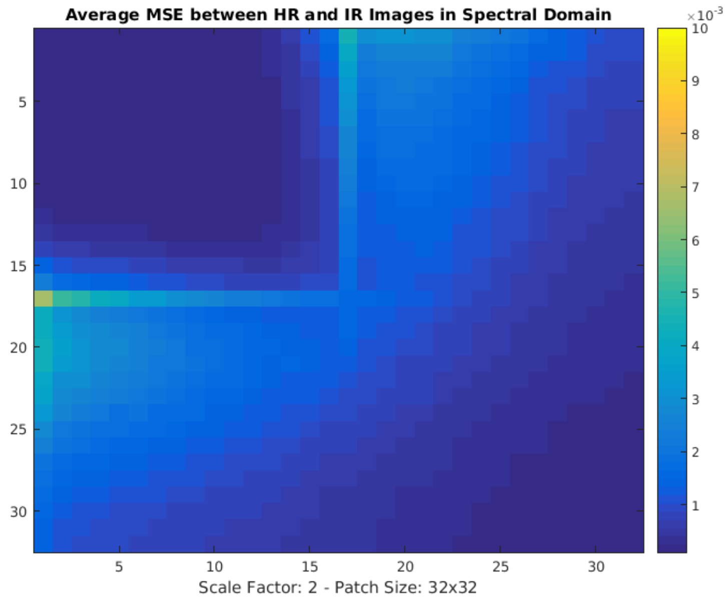
Bicubic Interpolation Bicubic Interpolation Bicubic Interpolation

3. Method – Super Resolution

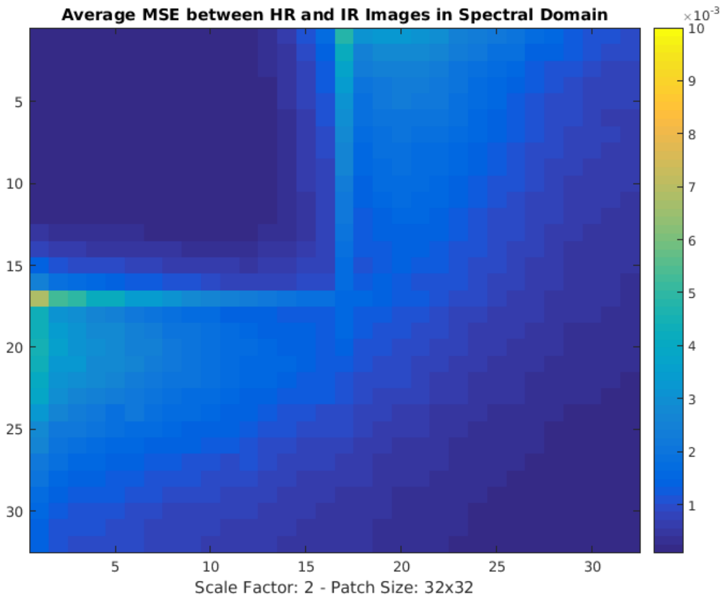
- Super-Resolution in Spectral Domain
  - Bicubic Interpolation + Divide Images into Patches + 2D DCT
  - Mean-Square Error Analysis in Spectral Domain



SRCNN



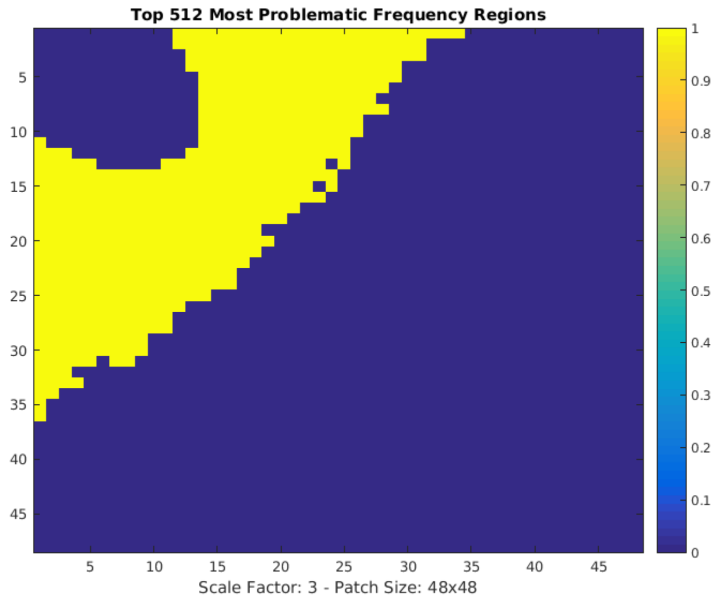
VDSR

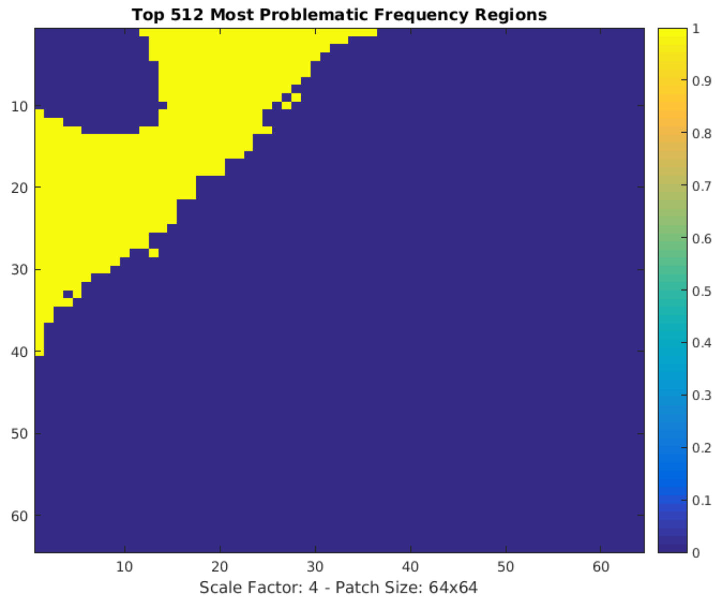


LapSRN

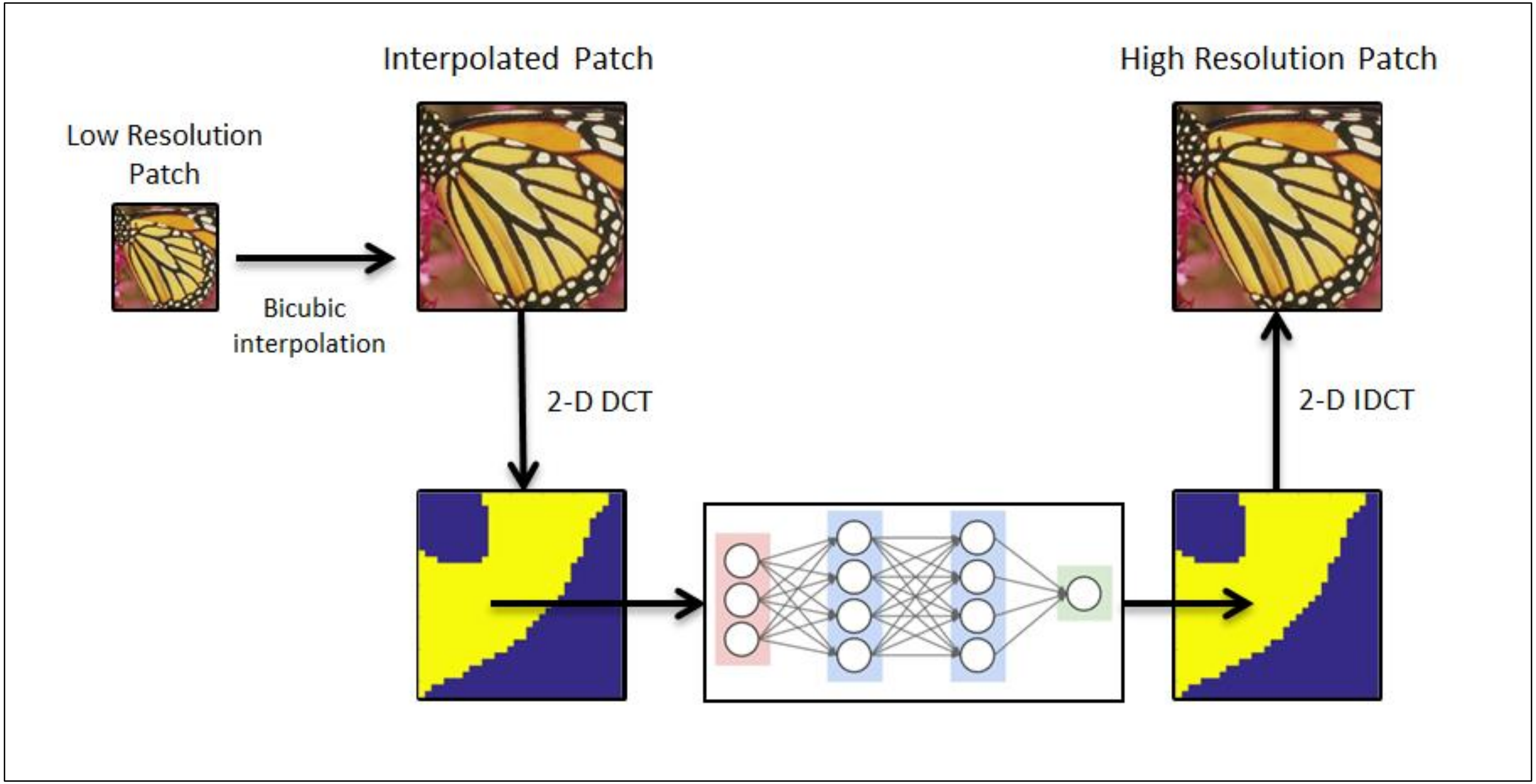
3. Method – Super Resolution

- Super-Resolution in Spectral Domain
  - Find most problematic region in spectral domain





3. Method – Super Resolution



3. Method – Super Resolution

- Super-Resolution in Spectral Domain
  - A feedforward neural network is trained to fix the problematic frequency regions.
  - Fully Connected Network is preferred over Convolutional Neural Network:
    - There is no homogeneous pixel distribution and local correlations in the spectral domain
    - The problematic region is not a rectangular region

3. Method – Super Resolution

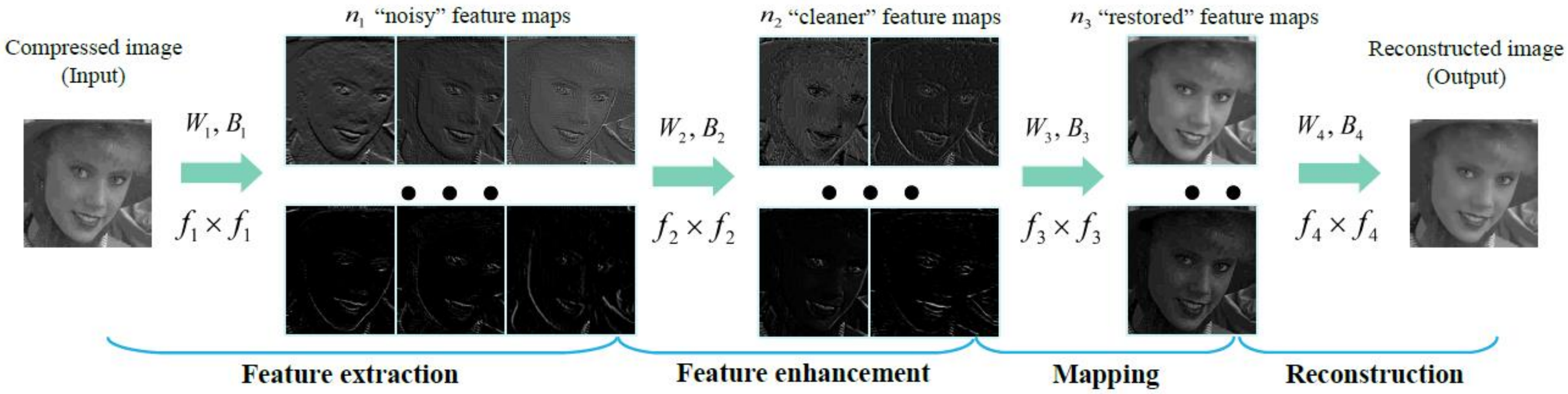
- Super-Resolution in Spectral Domain

Network Properties:

- Loss Function: Mean-square error
- Optimizer: Adam Optimizer
- Network Architecture: 4 Layered Fully Connected Layer with 512 hidden units
- Initialization: Xavier Initializer
- Regularization: Dropout
- Activation Function: Hyperbolic Tangent (due to range of DCT)

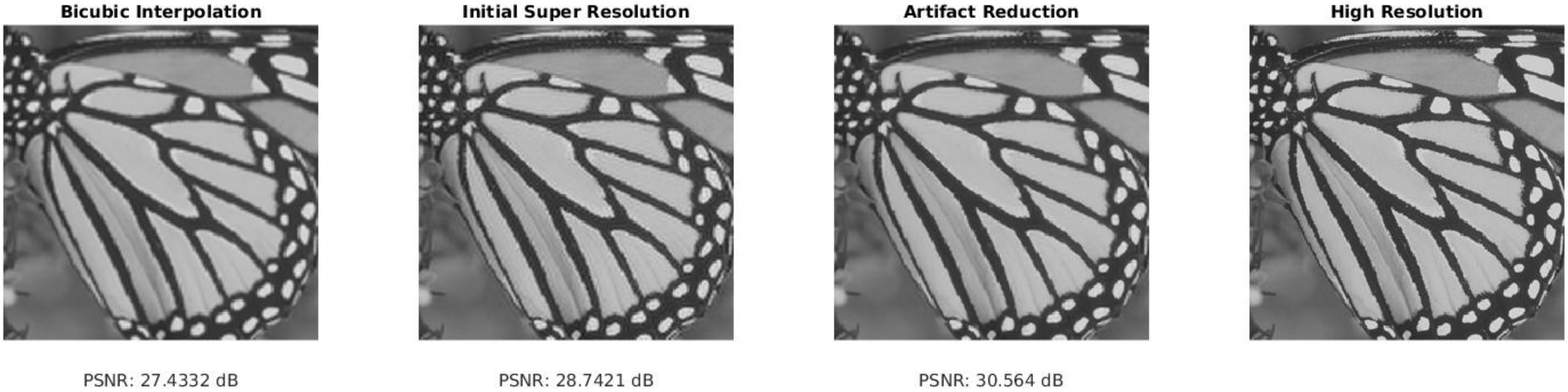
3. Method – Artifact Reduction

- Artifact Reduction in Spatial Domain
  - Processes in frequency domain introduce ringing and blocking artifacts in spatial domain
  - A pre-trained artifact reduction convolutional network is used
  - While artifacts are removed, the resolution of the images increased



3. Method – Artifact Reduction

- Artifact Reduction in Spatial Domain



4. Experiments and Results

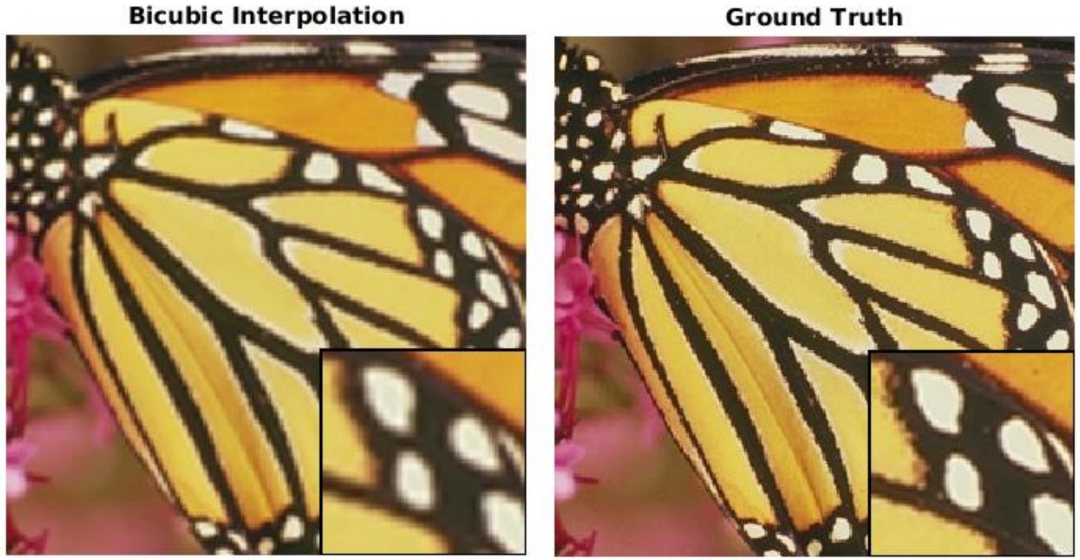
- Evaluation Metrics
  - PSNR

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right)$$

$$MSE = \frac{\sum_{M,N}[I_1(m,n) - I_2(m,n))]^2}{MN}$$

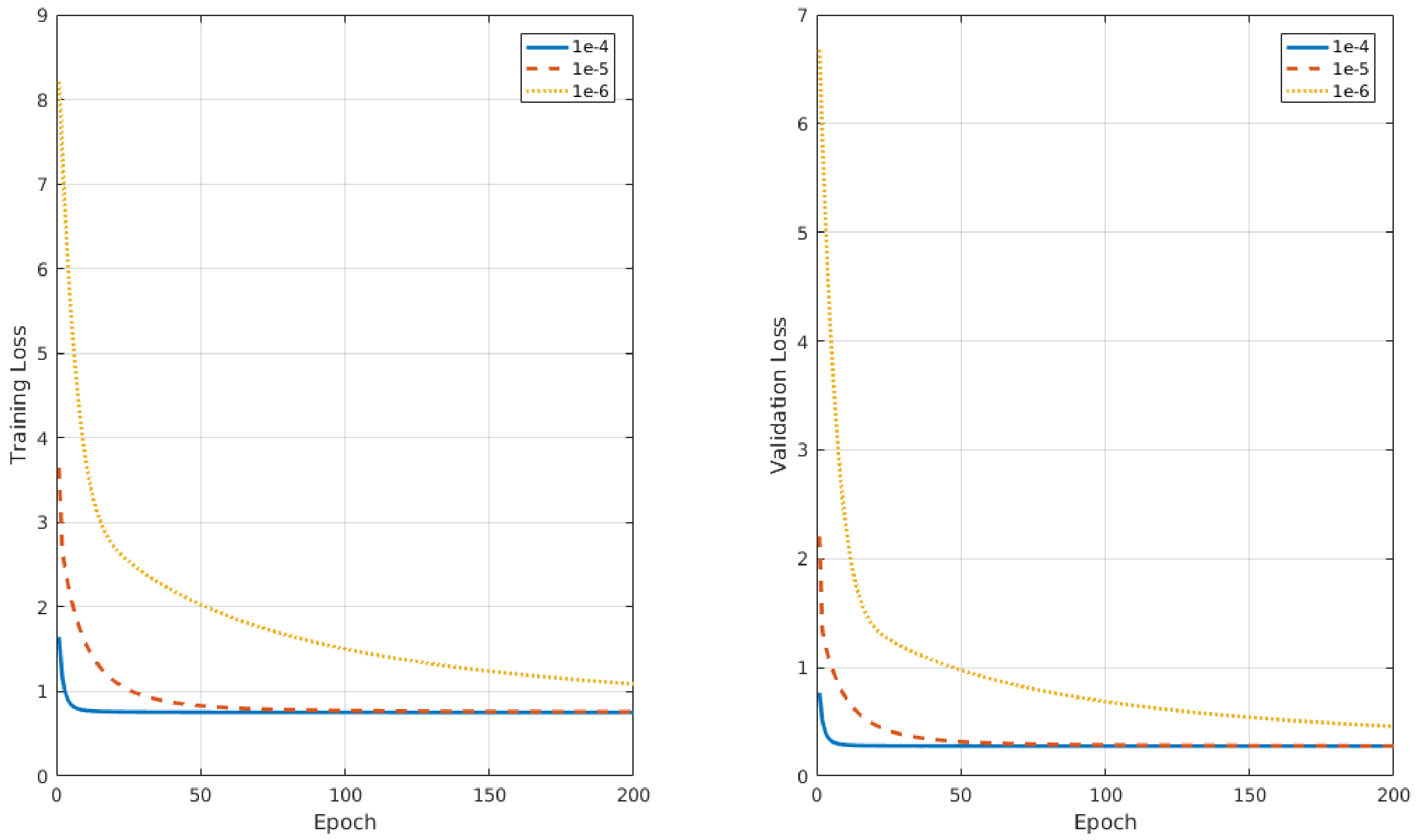
  - SSIM

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$



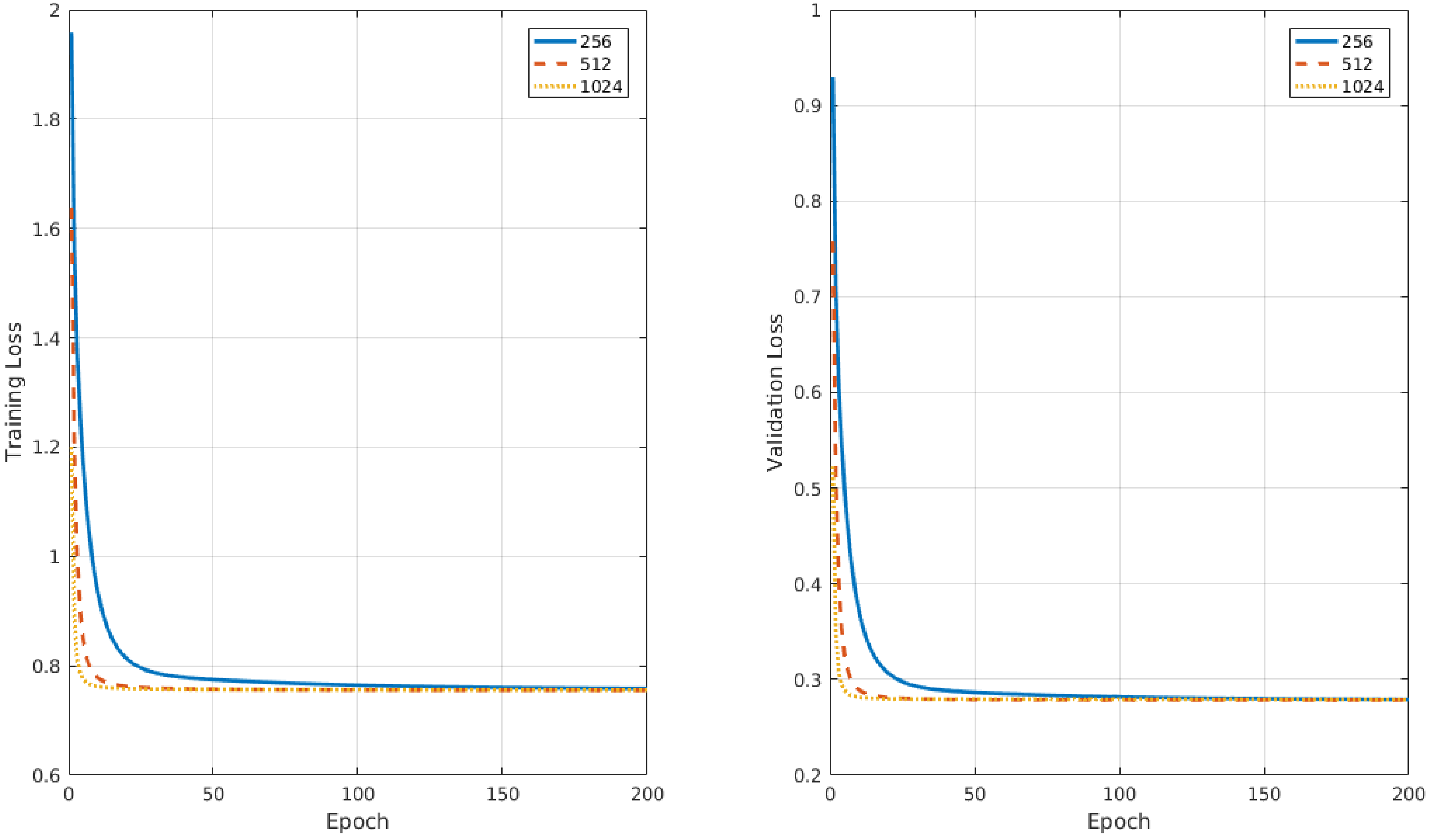
4. Experiments and Results

Learning Rate

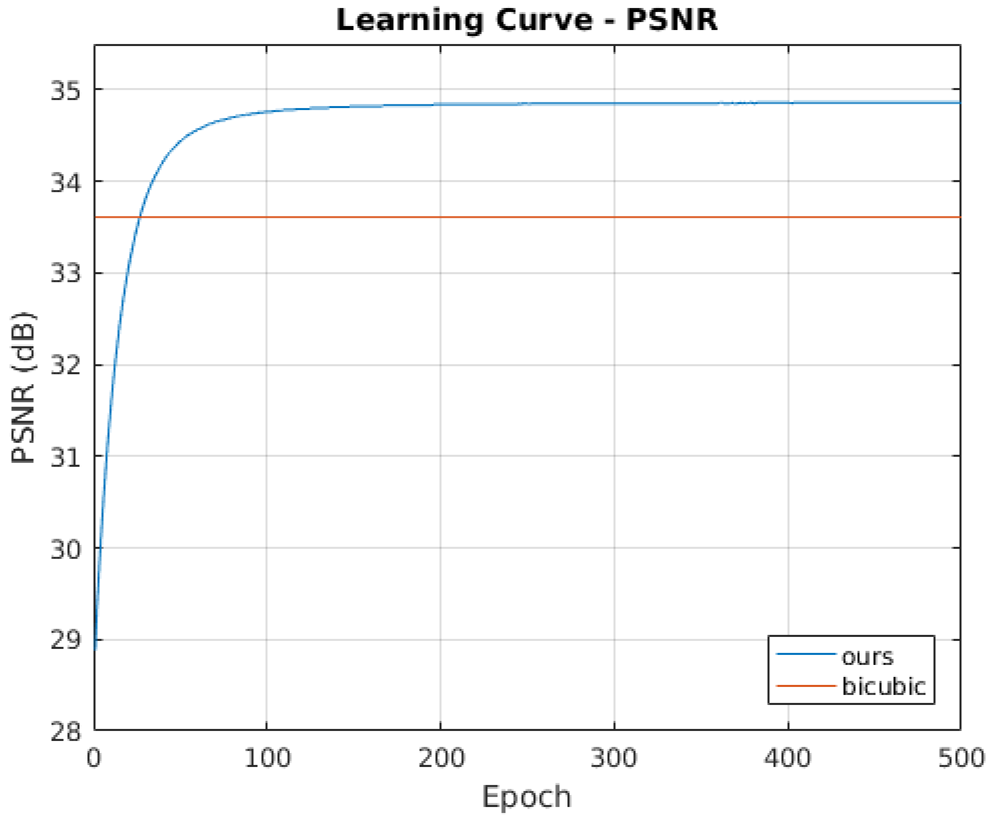


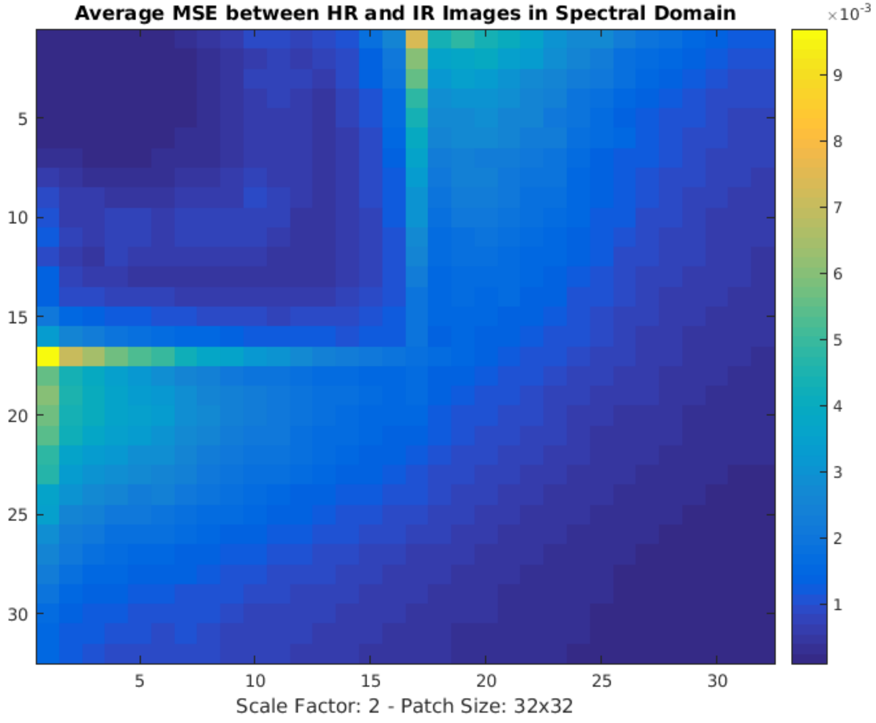
4. Experiments and Results

Number of Hidden Units in each Layer



4. Experiments and Results





4. Experiments and Results

PSNR - SSIM

Algorithm	Scale	Set5	Set14	BSDS100	Urban100
Bicubic	2	33.69 - 0.931	30.25 - 0.870	29.57 - 0.844	26.89 - 0.841
FourierSR	2	35.20 - 0.943	31.40 - 0.895	30.58 - 0.877	-
Ours	2	35.53 - 0.953	31.64 - 0.904	30.64 - 0.884	28.15 - 0.882
A+	2	36.60 - 0.955	32.32 - 0.906	31.24 - 0.887	29.25 - 0.895
RFL	2	36.59 - 0.954	32.29 - 0.905	31.18 - 0.885	29.14 - 0.891
SelfExSR	2	36.60 - 0.955	32.24 - 0.904	31.20 - 0.887	29.55 - 0.898
SRCNN	2	36.72 - 0.955	32.51 - 0.908	31.38 - 0.889	29.53 - 0.896
FSRCNN	2	37.05 - 0.956	32.66 - 0.909	31.53 - 0.892	29.88 - 0.902
VDSR	2	37.53 - 0.959	33.05 - 0.913	31.90 - 0.896	30.77 - 0.914
LapSRN	2	37.52 - 0.959	33.08 - 0.913	31.80 - 0.895	30.41 - 0.910

4. Experiments and Results

PSNR - SSIM

Algorithm	Scale	Set5	Set14	BSDS100	Urban100
Bicubic	3	30.41 - 0.869	27.55 - 0.775	27.22 - 0.741	24.47 - 0.737
FourierSR	3	31.42 - 0.883	28.32 - 0.802	27.79 - 0.772	-
Ours	3	31.44 - 0.906	28.41 - 0.828	27.78 - 0.788	24.78 - 0.781
A+	3	32.62 - 0.909	29.15 - 0.820	28.31 - 0.785	26.05 - 0.799
RFL	3	32.47 - 0.906	29.07 - 0.818	28.23 - 0.782	25.88 - 0.792
SelfExSR	3	32.66 - 0.910	29.18 - 0.821	28.30 - 0.786	26.45 - 0.810
SRCNN	3	32.78 - 0.909	29.32 - 0.823	28.42 - 0.788	26.25 - 0.801
FSRCNN	3	33.18 - 0.914	29.37 - 0.824	28.53 - 0.791	26.43 - 0.808
VDSR	3	33.67 - 0.921	29.78 - 0.832	28.83 - 0.799	27.14 - 0.829
LapSRN	3	33.82 - 0.922	29.87 - 0.832	28.82 - 0.798	27.07 - 0.828

4. Experiments and Results

PSNR - SSIM

Algorithm	Scale	Set5	Set14	BSDS100	Urban100
Bicubic	4	28.43 - 0.811	26.01 - 0.704	25.97 - 0.670	23.15 - 0.660
FourierSR	4	29.35 - 0.827	26.62 - 0.727	26.42 - 0.696	-
Ours	4	29.21 - 0.852	26.55 - 0.755	26.33 - 0.721	23.42 - 0.701
A+	4	30.32 - 0.860	27.34 - 0.751	26.83 - 0.711	24.34 - 0.721
RFL	4	30.17 - 0.855	27.24 - 0.747	26.76 - 0.708	24.20 - 0.712
SelfExSR	4	30.34 - 0.862	27.41 - 0.753	26.84 - 0.713	24.83 - 0.740
SRCNN	4	30.50 - 0.863	27.52 - 0.753	26.91 - 0.712	24.53 - 0.725
FSRCNN	4	30.72 - 0.866	27.61 - 0.755	26.98 - 0.715	24.62 - 0.728
VDSR	4	31.35 - 0.883	28.02 - 0.768	27.29 - 0.726	25.18 - 0.754
LapSRN	4	31.54 - 0.885	28.19 - 0.772	27.32 - 0.727	25.21 - 0.756

Scale Factor = x2

4. Experiments and Results

Bicubic Interpolation

PSNR: 27.4389 dB

A+

PSNR: 32.0289 dB

Ours

PSNR: 30.5692 dB

SRCNN

PSNR: 32.7765 dB

LapSRN

PSNR: 34.2258 dB

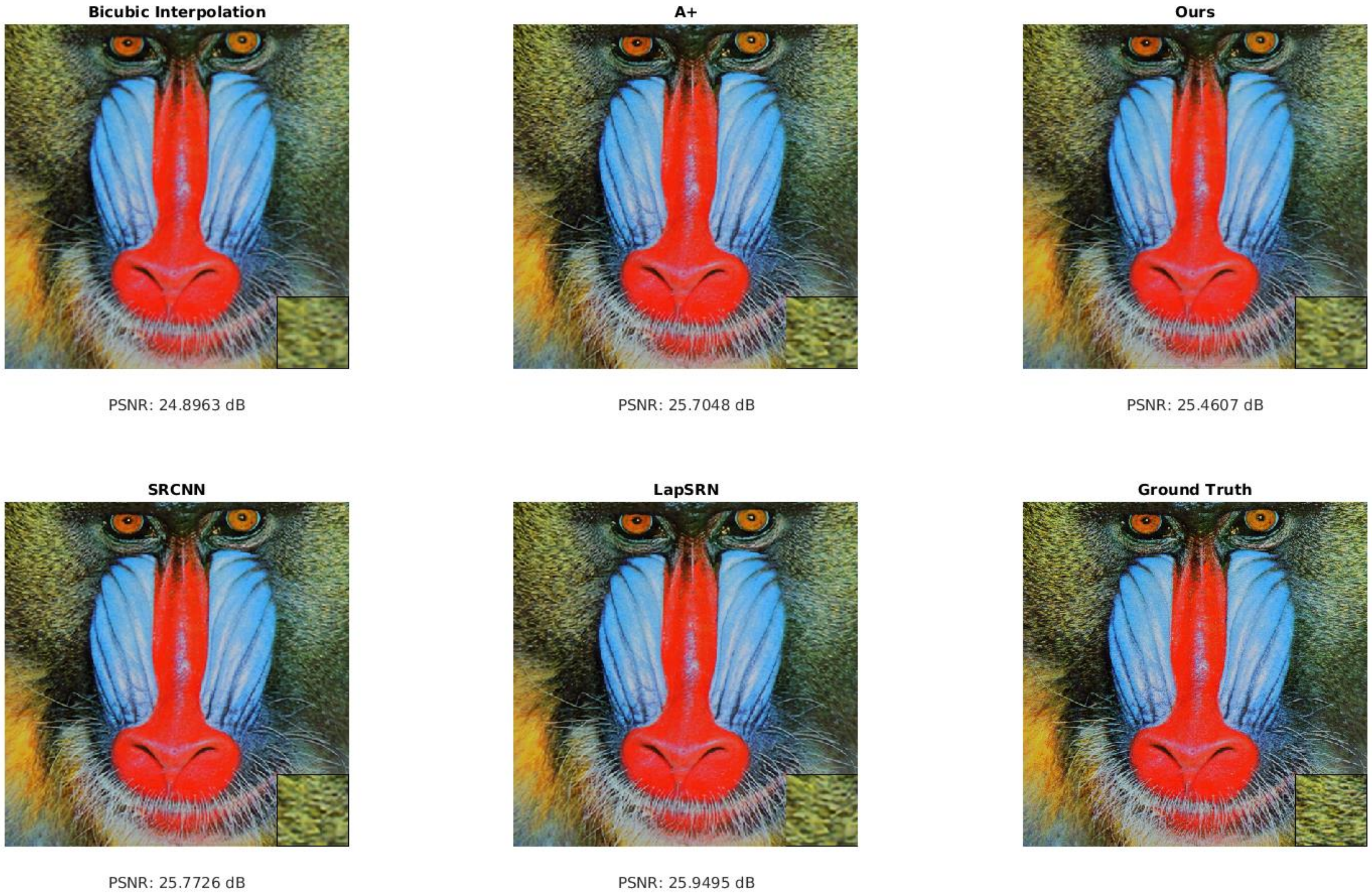
Ground Truth

Scale Factor = x2

4. Experiments and Results

Scale Factor = x2

4. Experiments and Results



Scale Factor = x2

4. Experiments and Results

Bicubic Interpolation

PSNR: 27.9799 dB

A+

PSNR: 28.6183 dB

Ours

PSNR: 28.2206 dB

SRCNN

PSNR: 28.53 dB

LapSRN

PSNR: 28.2795 dB

Ground Truth

Scale Factor = x2

4. Experiments and Results

Bicubic Interpolation

PSNR: 34.7173 dB

A+

PSNR: 36.6207 dB

Ours

PSNR: 35.8357 dB

SRCNN

PSNR: 36.414 dB

LapSRN

PSNR: 37.0189 dB

Ground Truth

4. Experiments and Results

4. Experiments and Results

4. Experiments and Results

Bicubic Interpolation

PSNR: 26.2526 dB

A+

PSNR: 26.4738 dB

Ours

PSNR: 26.6357 dB

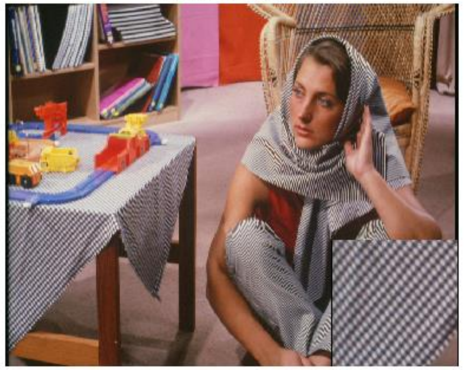
SRCNN

PSNR: 26.5429 dB

LapSRN

PSNR: 25.8492 dB

Ground Truth

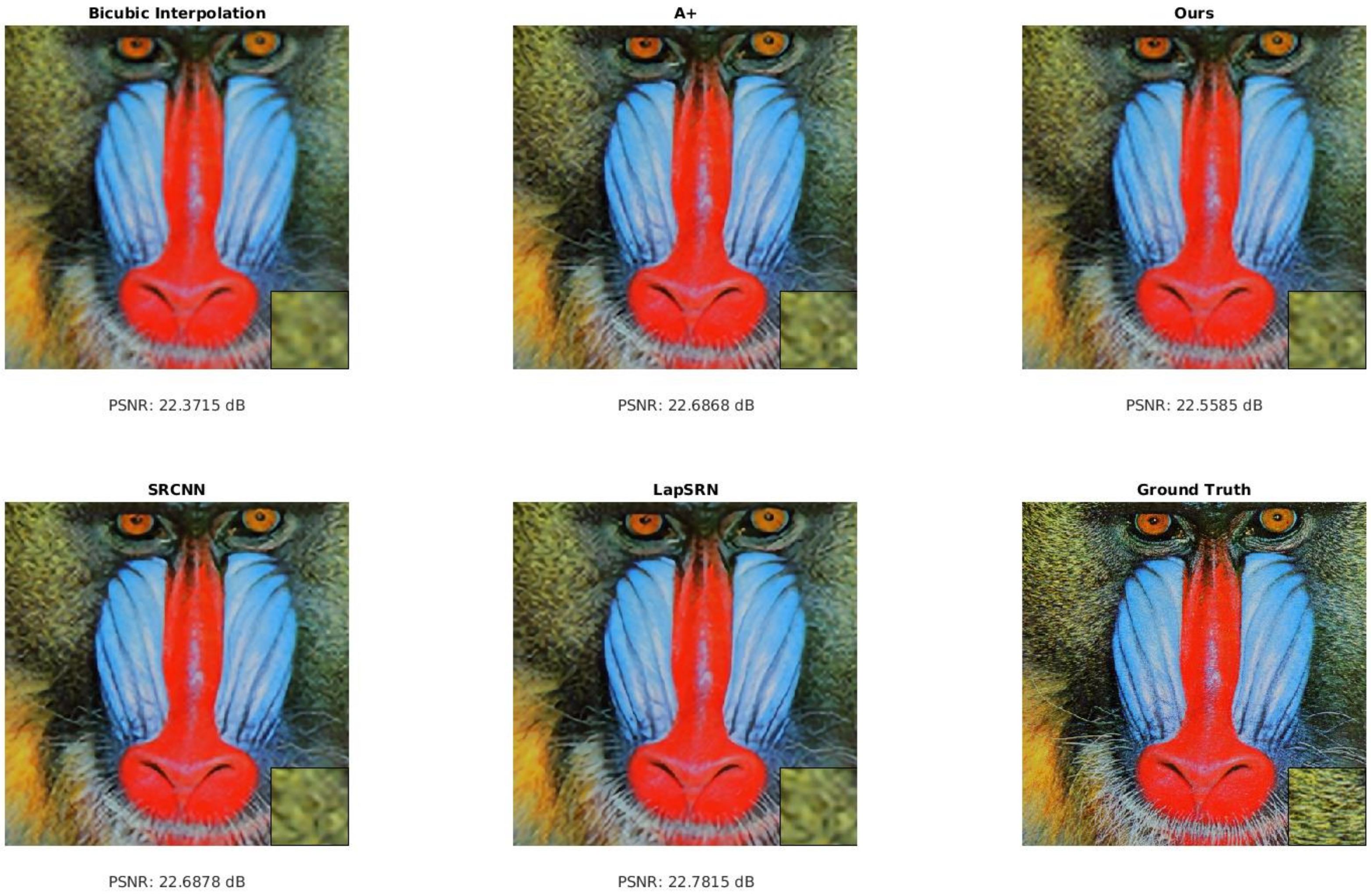


Scale Factor = x4

4. Experiments and Results

Scale Factor = x4

4. Experiments and Results



Scale Factor = x4

4. Experiments and Results

Bicubic Interpolation

PSNR: 25.1283 dB

A+

PSNR: 25.6807 dB

Ours

PSNR: 25.3998 dB

SRCNN

PSNR: 25.6869 dB

LapSRN

PSNR: 25.6849 dB

Ground Truth

5. Conclusion

- Neural Networks and Fourier Transform
  - Super Resolution in Spectral Domain
  - Artifact Reduction in Spatial Domain

- Future Works:
  - Combine spatial domain and spectral domain in SR problem
  - Better prediction model in the spectral domain might be developed
  - The spectral domain would allow us to convert super-resolution problem to image completion problem.
  - Develop a completely different model without bicubic interpolation.
  - New approaches might be developed in spectral domain for other computer vision problems.

Thank you!