Image Super-Resolution Using Deep Convolutional Networks

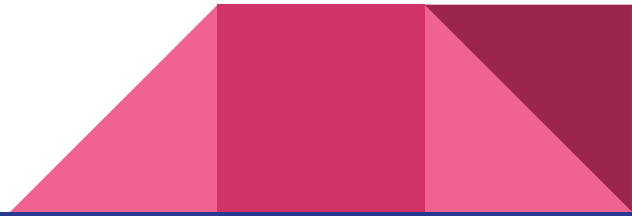
Chao Dong, The Chinese University of Hong Kong.
Chen Change Loy, The Chinese University of Hong Kong.
Kaiming He, Microsoft Research,
Xiaoou Tang, The Chinese University of Hong Kong.

Presented by
J.Nikhil (2018csm1011)

Paper appeared in **ECCV 2014 – European Conference on Computer Vision** (Citations Count - 1529)

Outline

1. Problem statement.
2. Related Work.
3. Prerequisite Concepts.
    a) Convolution neural networks (CNN).
    b) Bicubic Interpolation.
4. Proposed Approach.
5. Model and Performance Trade-offs.
6. Conclusion.



1. Problem statement

- Aim is to Convert Single Low Resolution Image to High Resolution Image.
- Classic problem in Computer Vision.
- Not a Well defined problem because solution is not Unique.
- One Solution : Constraint the Solution space by using prior Information.

1.Problem Statement

Low Resolution Image

Original / PSNR

Y

Mapping F(Y)

Goal is to Create end to end Mapping Between Y and X

High Resolution Image

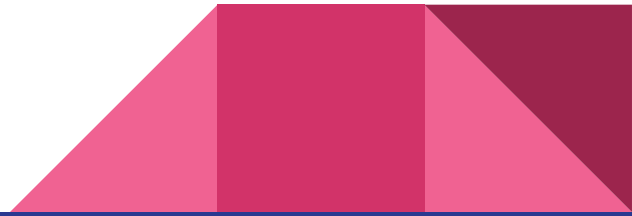
SRCNN / 27.95 dB

X

Source : Images from paper

2. Related Work

- External example based methods - Learn mapping between Low and High resolution Images using external datasets.
- Internal example based methods use Self similarity property.
- Using Self similarity approach generate patches from the Input image.
- Different mapping functions used in above methods are kernel Regression, Random forest etc..

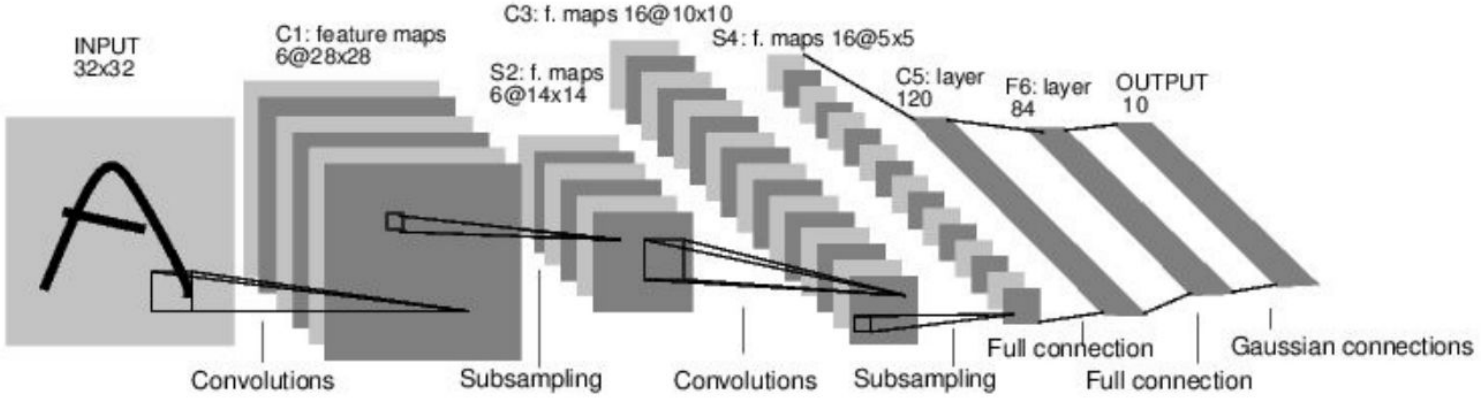


3.Concepts required

a) Convolution Neural Networks (CNNs)

- CNNs, like neural networks, are made up of neurons with learnable weights and biases.
- Each neuron receives several inputs, takes a weighted sum over them, pass it through an activation function and responds with an output.
- CNNs have wide applications in image and video recognition, recommender systems and natural language processing.

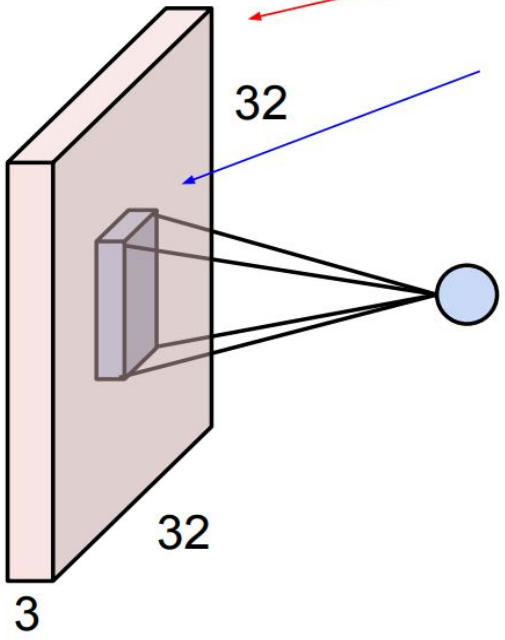
Convolutional Neural Networks



[LeNet-5, LeCun 1980]

*Source: from Andrej Karpathy course.

Convolution Layer



32x32x3 image

5x5x3 filter w

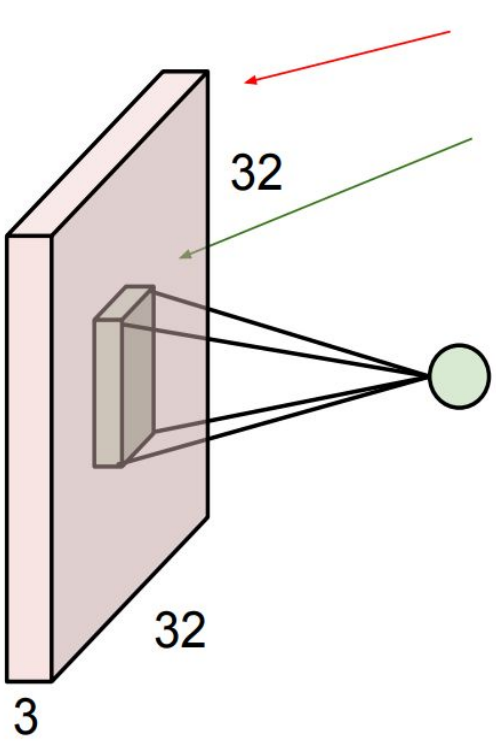
1 number:
the result of taking a dot product between the filter and a small 5x5x3 chunk of the image (i.e. 5*5*3 = 75-dimensional dot product + bias)

$$w^T x + b$$

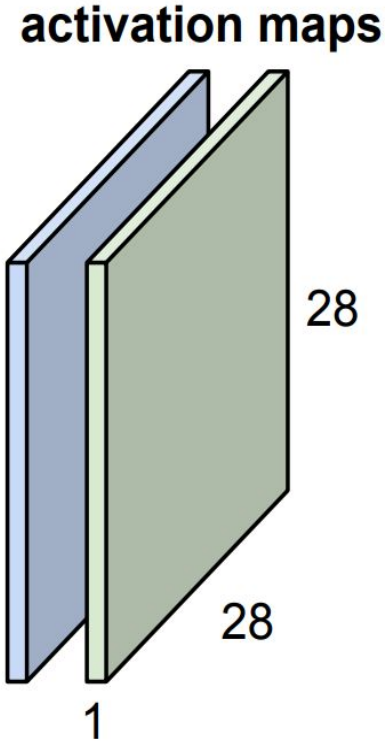
*Source: from Andrej Karpathy course.

Convolution Layer

consider a second, green filter

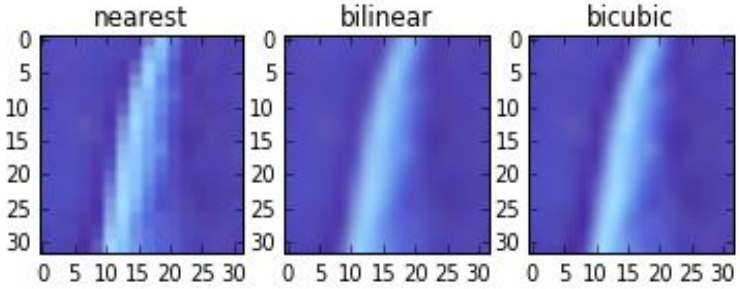


convolve (slide) over all spatial locations



*Source: from Andrej Karpathy course.

3. Bicubic Interpolation

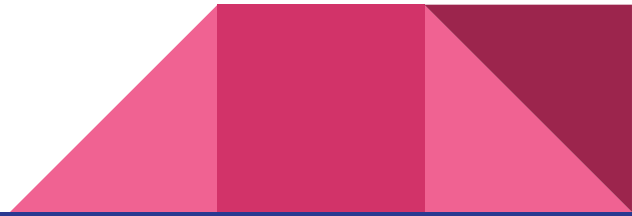


- Interpolation is a method of Constructing New data points within range of Discrete Data Points.
- Several Interpolation techniques exist like Bilinear Interpolation, Bicubic Interpolation etc..
- In proposed approach they have used Bicubic Interpolation which takes the pixel $y = ax^3+bx^2+cx+d$.
- It is observed in above diagram that edges are represented in Cubic polynomial.

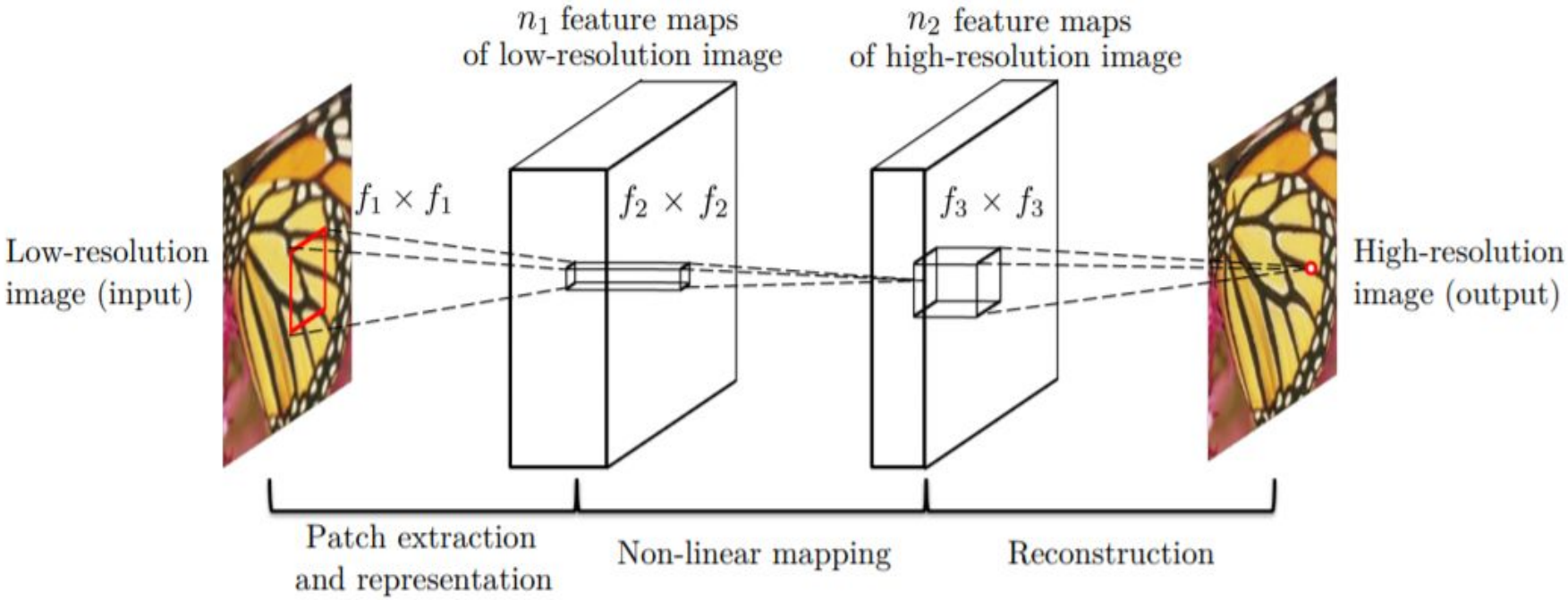
Source : Image from Google

4. Proposed Approach

- Involves Three Phases :
  - Patch Extraction and Representation.
  - Non-Linear Mapping.
  - Reconstruction.
- Initially we upscale image using Bicubic Interpolation.
- As specified before, we wish to learn mapping F between low resolution Image and High Resolution Image through above three phases.



4. Proposed Approach



Source : Image from paper.

4. Proposed Approach - (a) Patch Extraction and representation

- Convolve the image with filters and each output is referred as Basis.
- We can Understand it as Feature map which represent the Patch in low resolution image.
- How the Feature maps are generated for Low Resolution Image?

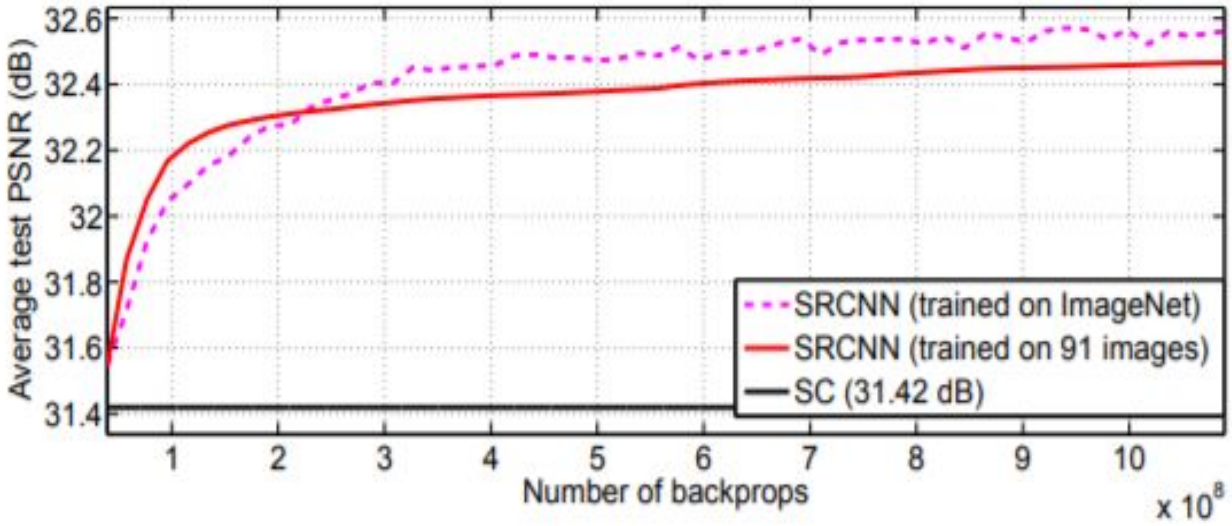
4. Proposed Approach - (b) Non Linear Mapping

- Each Feature map generated, wil be mapped to a new Feature map.
- This will be converted in later stage as a patch in High Resolution Image.
- Apply ReLU activation for each value in values generated in the Matrix.
- Values in the Weights in each Layer are updated to Minimise MSE or Maximise PSNR.

4. Proposed Approach - (c) Reconstruction & Back Propagation

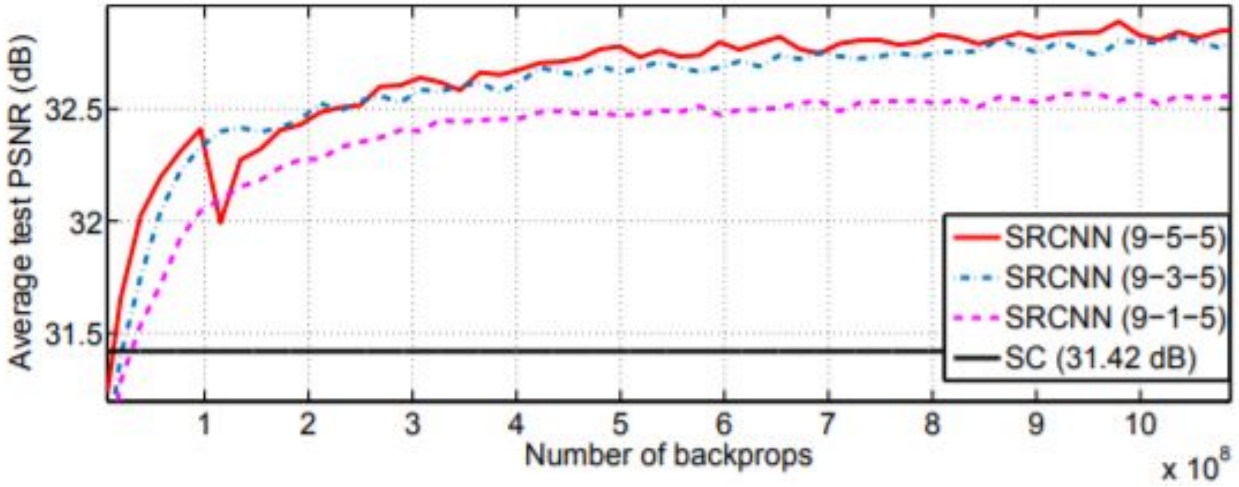
- In Reconstruction Phase, the transformed values are updated back into C Channel Image chosen Initially.
- The weight values updation Based on the standard Backpropagation algorithm used in Neural Networks.
- One Backprop means the weights updated after each Feed Forward Operation.
- Number of Backprops affect the convergence to Local Minimum or Global Minimum as Neural Networks is not a Convex Optimisation.

5. Model and Performance Analysis - Varying Number of Samples.



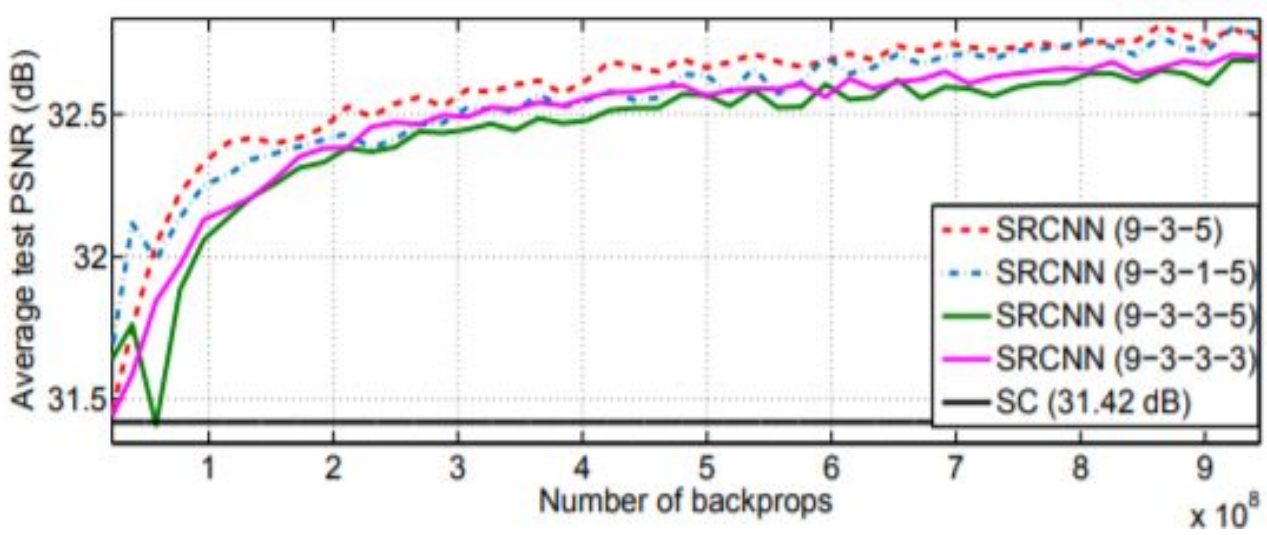
It is Observed that Training data even with 91 Images is as Comparable to 5 Million Images in ImageNet data.

5. Model and Performance Analysis - Varying Filter Size in F2



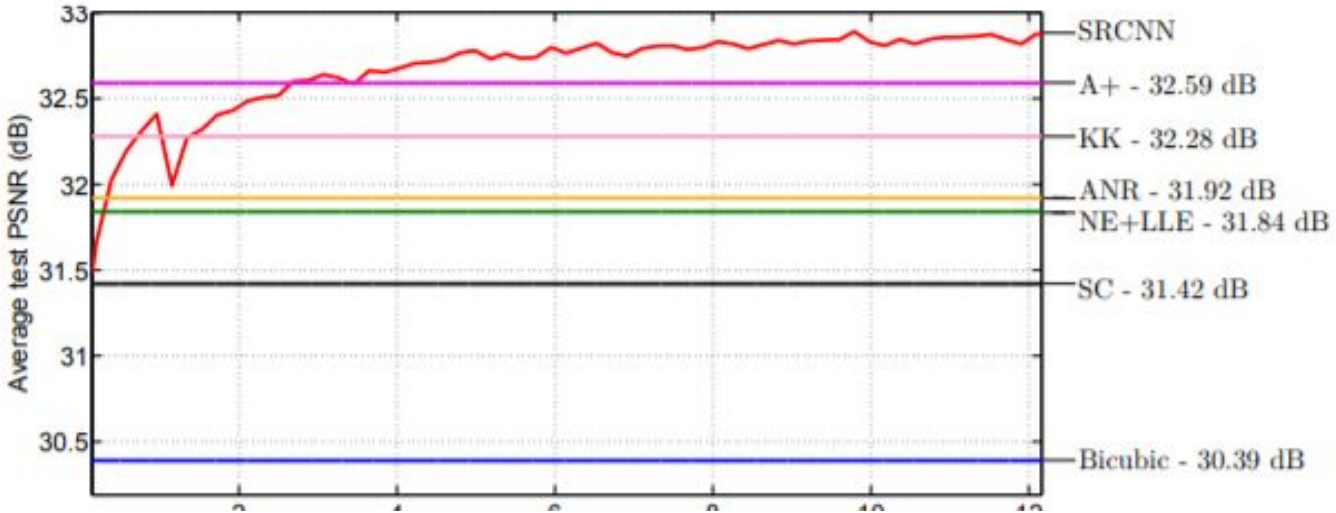
As there is Increase in Filter Size, there is Improvement in Average Test PSNR.

5. Model and Performance Analysis - Varying Number of Layers



As we Increase Number of Layers, it is not Guaranteed to Increase Performance.

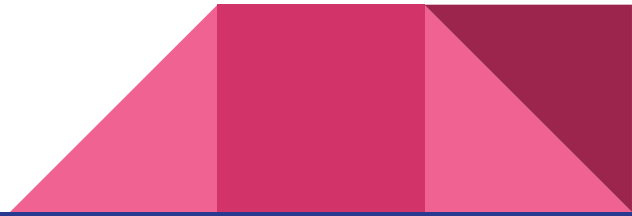
5. Model and Performance Analysis - With State of Art Algorithms



Proposed method- SRCNN achieved a better performance than previous State of Art Algorithms.

6. Conclusion

- Official Implementation - link here (Matlab code)
- Proposed a Convolution Neural network which learns end to end mapping between Low resolution Image and High resolution Image.
- Pre-processing - Bicubic Interpolation.
- Activation Function - ReLU
- Proposed approach showed better average PSNR ratio than state of art methods.
- Authors Improved SRCNN in Later years and named it is FSRCNN.



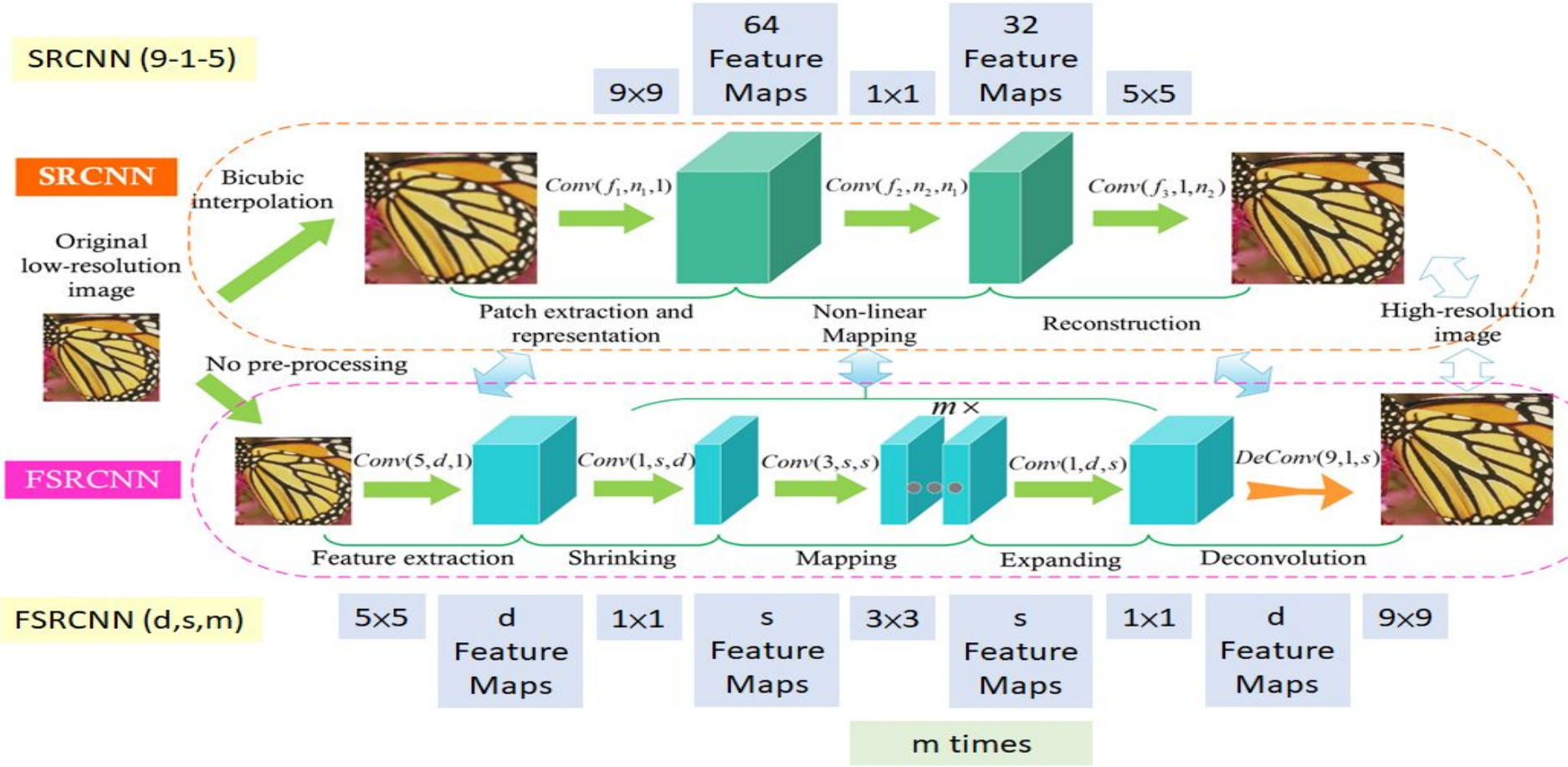


Image : FSRCNN Paper.

Thank You