PEC
EXPLORE INNOVATE EXCEL

PROJECT REPORT

(August-December,2019)

Image Super-Resolution Using Deep Convolutional Neural Networks (SRCNN)

Submitted by:

Abhimanyu Parmar -17103076

Jagmanjot Singh – 17103117

Harnoor Singh – 17103119

Shresth - 17103120

Under the guidance of:

Mrs. Amandeep Kaur

Assistant Professor

Punjab Engineering College

DECLARATION

We hereby declare that the project work entitled "**Image Super-Resolution using deep convolutional networks (SRCNN)**" is an authentic record of our work carried out during 5th semester as a part of our minor project, under the guidance of **Mrs. Amandeep Kaur** (during August to December, 2019)

Date: _____

Certified that the above statement made by the student is correct to the best of our knowledge and belief.

Mrs. Amandeep Kaur

Assistant Professor

Punjab Engineering College

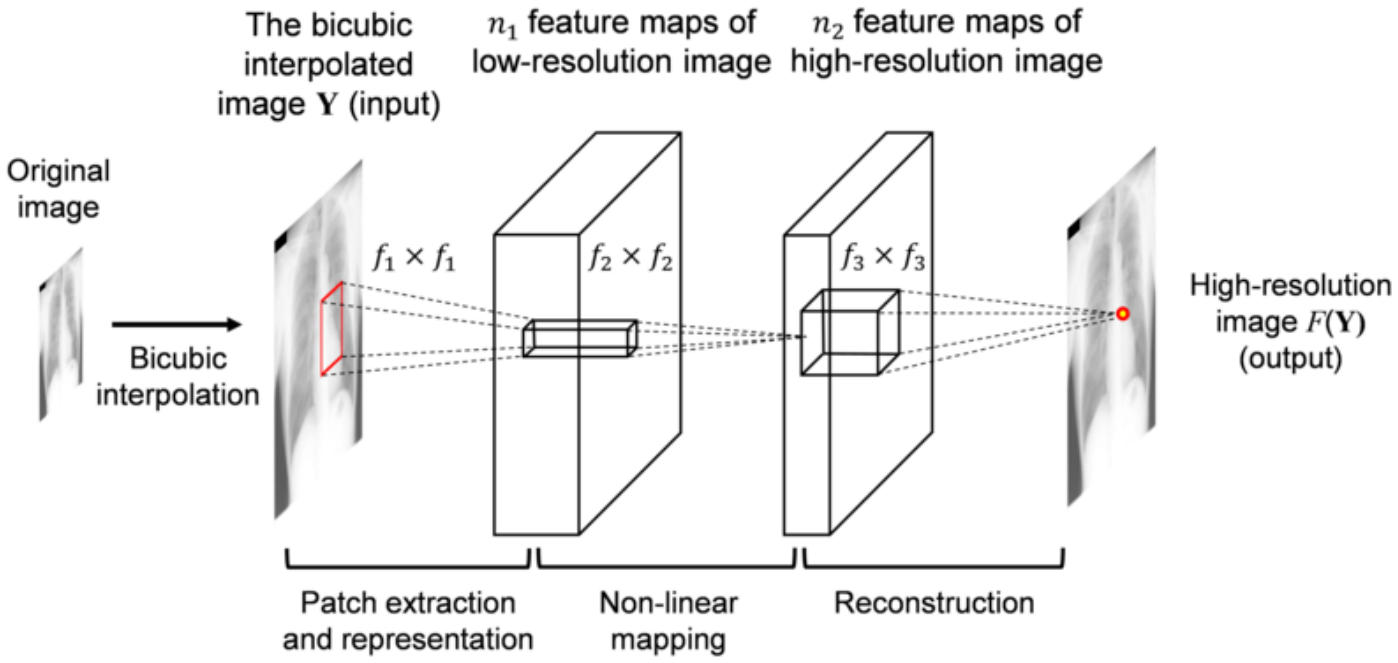
(Mentor)

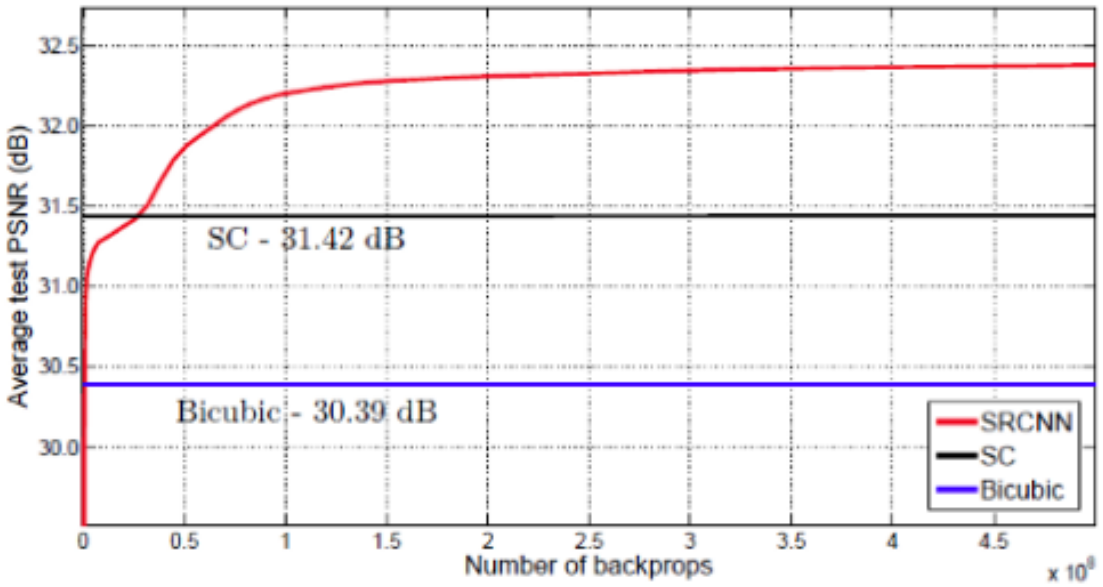
Contents

Abstract

We propose a deep learning method for single image super-resolution (SR). Our method directly learns an end-to-end mapping between the low/high-resolution images. The mapping is represented as a deep convolutional neural network (CNN) that takes the low-resolution image as the input and outputs the high-resolution one. We further show that traditional sparse-coding-based SR methods can also be viewed as a deep convolutional network. But unlike traditional methods that handle each component separately, our method jointly optimizes all layers. Our deep CNN has a lightweight structure, yet demonstrates state-of-the-art restoration quality, and achieves fast speed for practical on-line usage. We explore different network structures and parameter settings to achieve trade-offs between performance and speed.

List of Figures





List of Abbreviations

SRCNN	Super resolution convolutional neural network.
SR	Super resolution
CNN	Convolutional neural network
PSNR	Peak signal to noise ratio
FSRCNN	Fast-super resolution convolutional neural network

Introduction

Single image super-resolution (SR), which aims at recovering a high-resolution image from a single low-resolution image, is a classical problem in computer vision. We consider a convolutional neural network that directly learns an end-to-end mapping between low- and high-resolution images.

We name the proposed model **Super-Resolution Convolutional Neural Network (SRCNN)**. The proposed SRCNN has several appealing properties. Its structure is intentionally designed with simplicity in mind, and yet provides superior accuracy. Furthermore, the patch extraction and aggregation are also formulated as convolutional layers, so are involved in the optimization. We demonstrate that deep learning is useful in the classical computer vision problem of super resolution, and can achieve good quality and speed.

We present a fully convolutional neural network for image super-resolution. The network directly learns an end-to-end mapping between low and high-resolution images, with little pre/postprocessing beyond the optimization. The present work adds to the initial version in significant ways. We improve the SRCNN by introducing larger filter size in the non-linear mapping layer, and explore deeper structures by adding nonlinear mapping layers.

Our approach primarily consists of three operations:

1. Patch extraction and representation
2. Non-linear mapping
3. Reconstruction

Problem Statement

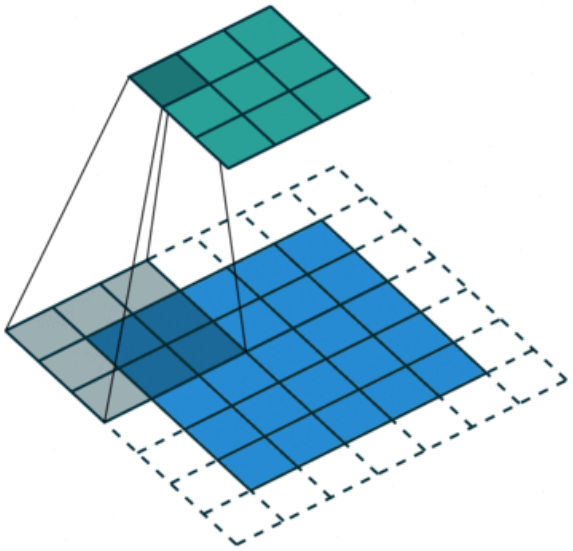
To check the efficiency and accuracy of the produced results in various models, we considered a quality measurement method called PSNR (**Peak Signal to Noise Ratio**). This metric is a de-facto standard used in Super Resolution research.

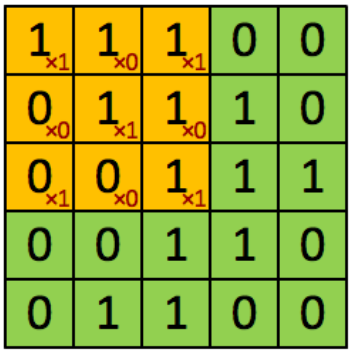
1) PSNR value obtained in traditional methods were very low, therefore results were very poor and were distorted.
2) Using SRCNN, we were able to achieve high PSNR values which resulted in better results than the traditional methods.
3) Even a simple three-layer SRCNN was able to beat most non-deep-learning methods when measured on standard benchmark datasets using PSNR
4) SRCNN uses end to end mapping between low/high resolution images.

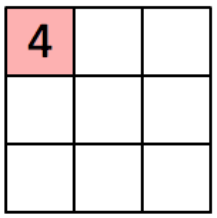
Proposed Work: Algorithm

Aiming at the problem that the existing high-resolution algorithm has too long training time, poor reconstruction performance and slow running speed, a new image super-resolution reconstruction algorithm based on convolutional neural network is proposed.

The algorithm uses low-resolution images as the network input, the higher-order representation of the image is learned using the convolution operation, the high-resolution image is up-sampling by the deconvolution operation, and the residual structure is added to the network, so that the entire network can converge better.







Model/Approach

1. Patch extraction and representation: Extracts patches from the low-resolution image Y and represents each patch as a high-dimensional vector. These vectors comprise a set of feature maps, of which the number equals to the dimensionality of the vectors.

$$F_1(\mathbf{Y}) = \max(0, W_1 * \mathbf{Y} + B_1)$$

The first Layer

Size of W1: c×f1×f1×n1

Size of B1: n1

2. Non-linear mapping: Nonlinearly maps each high-dimensional vector onto another high-dimensional vector. Each mapped vector is conceptually the representation of a high-resolution patch. These vectors comprise another set of feature map.

$$F_2(\mathbf{Y}) = \max(0, W_2 * F_1(\mathbf{Y}) + B_2)$$

The second layer

Size of W2: n1×1×1×n2

Size of B2: n2

3. Reconstruction: Aggregates the above high-resolution patch-wise representations to generate the final high-resolution image. This image is expected to be similar to the ground truth X.

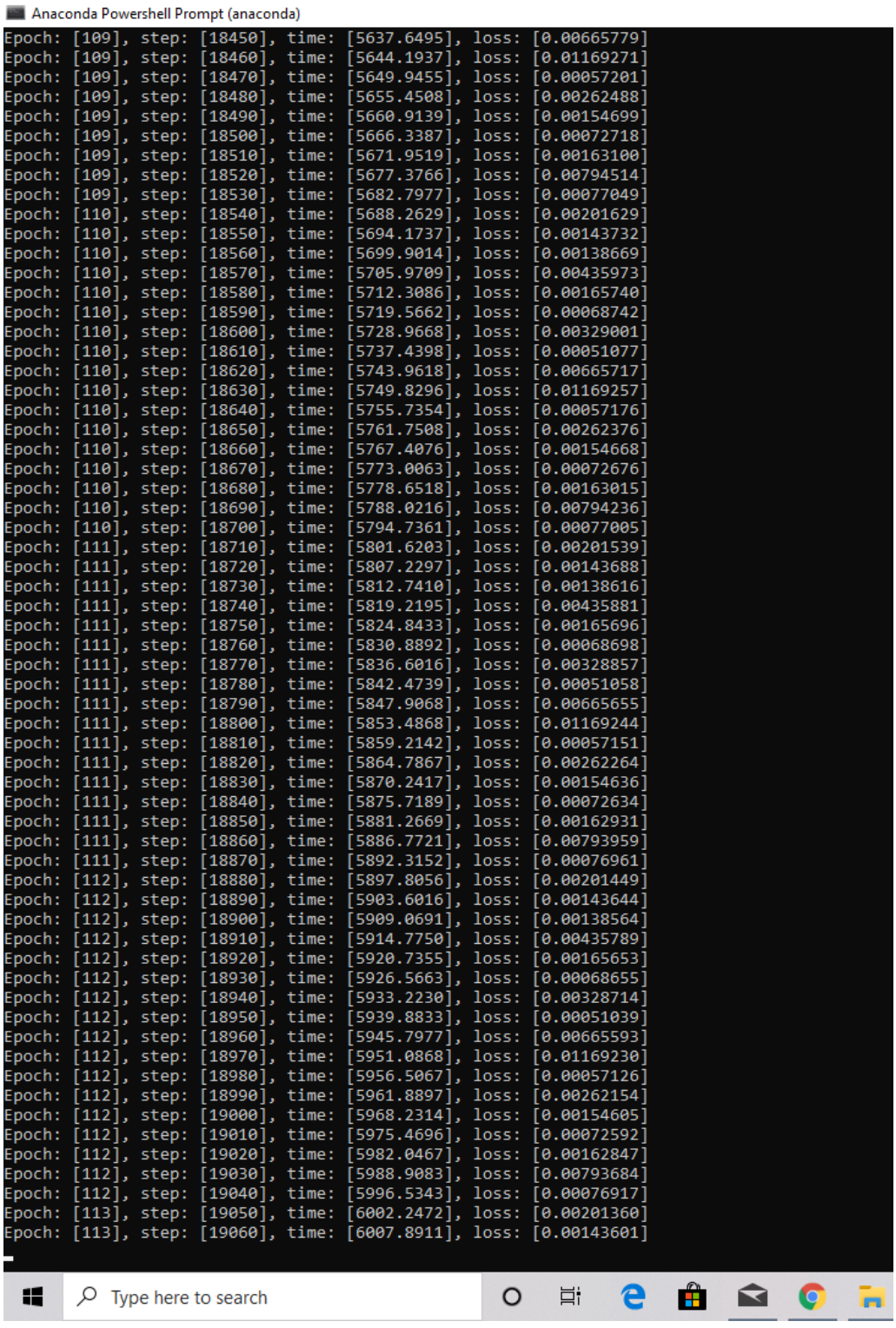
$$F(\mathbf{Y}) = W_3 * F_2(\mathbf{Y}) + B_3$$

The third layer

Size of W3: n2×f3 ×f3×c

Size of B3: c

Implementation



```
def build_model(self):
    model = Sequential()
    model.add(Conv2D(64,9,padding='same',input_shape=(self.image_size,self.image_size,self.c_dim)))
    model.add(Activation('relu'))
    model.add(Conv2D(32,1,padding='same'))
    model.add(Activation('relu'))
    model.add(Conv2D(self.c_dim,5,padding='same'))
    optimizer = Adam(lr=self.learning_rate)
    model.compile(optimizer=optimizer, loss='mean_squared_error', metrics=['accuracy'])
    return model
```

Results & discussions

Bicubic:

Predicted:

As shown in Table below, the proposed SRCNN yields the highest scores in most evaluation matrices in all experiments. When we take a look at other evaluation metrics, we observe that SC, to our surprise, gets even lower scores than the bicubic interpolation on IFC and NQM. It is clear that the results of SC are more visually pleasing than that of bicubic interpolation. Thus, SRCNN achieves the best performance among all methods and scaling factors.

The average results of PSNR (dB), SSIM, IFC, NQM, WPSNR (dB) and MSSIM on the Set14 dataset.

Eval. Mat	Scale	Bicubic	SC [50]	NE+LLE [4]	KK [25]	ANR [41]	A+ [41]	SRCNN
PSNR	2	30.23	-	31.76	32.11	31.80	32.28	**32.45**
	3	27.54	28.31	28.60	28.94	28.65	29.13	**29.30**
	4	26.00	-	26.81	27.14	26.85	27.32	**27.50**
SSIM	2	0.8687	-	0.8993	0.9026	0.9004	0.9056	**0.9067**
	3	0.7736	0.7954	0.8076	0.8132	0.8093	0.8188	**0.8215**
	4	0.7019	-	0.7331	0.7419	0.7352	0.7491	**0.7513**
IFC	2	6.09	-	7.59	6.83	7.81	**8.11**	7.76
	3	3.41	2.98	4.14	3.83	4.23	**4.45**	4.26
	4	2.23	-	2.71	2.57	2.78	**2.94**	2.74
NQM	2	40.98	-	41.34	38.86	41.79	**42.61**	38.95
	3	33.15	29.06	37.12	35.23	37.22	**38.24**	35.25
	4	26.15	-	31.17	29.18	31.27	**32.31**	30.46
WPSNR	2	47.64	-	54.47	53.85	54.57	**55.62**	55.39
	3	39.72	41.66	43.22	43.56	43.36	44.25	**44.32**
	4	35.71	-	37.75	38.26	37.85	38.72	**38.87**
MSSSIM	2	0.9813	-	0.9886	0.9890	0.9888	0.9896	**0.9897**
	3	0.9512	0.9595	0.9643	0.9653	0.9647	0.9669	**0.9675**
	4	0.9134	-	0.9317	0.9338	0.9326	0.9371	**0.9376**

Future work

1. We are already trying to extend our domain from grayscale to RGB pictures, which will further enhance the performance.
2. Our perceptual loss experiments show that PSNR may not be a good metric to use for evaluating super resolution networks. In our opinion, more research needs to be done on different types of perceptual loss.
3. Further we will also try to implement our model into video super resolution, which will convert low resolution video into high resolution video in real time.
4. We will use FSRCNN (fast-super resolution neural network) which is a more effective process as it contains five-part convolutional layer and is more effective than SRCNN.

Conclusion

We have presented a novel deep learning approach for single image super-resolution (SR). We show that conventional sparse-coding-based SR methods can be reformulated into a deep convolutional neural network. The proposed approach, SRCNN, learns an end-to-end mapping between low- and high-resolution images, with little extra pre/post-processing beyond the optimization. With a lightweight structure, the SRCNN has achieved superior performance than the state-of-the-art methods. We conjecture that additional performance can be further gained by exploring more filters and different training strategies. Besides, the proposed structure, with its advantages of simplicity and robustness, could be applied to other low-level vision problems, such as image deblurring or simultaneous SR+denoising. One could also investigate a network to cope with different upscaling factors.

References

[1] Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing 54(11), 4311–4322 (2006)

[2] Bevilacqua, M., Roumy, A., Guillemot, C., Morel, M.L.A.: Lowcomplexity single-image super-resolution based on nonnegative neighbor embedding. In: British Machine Vision Conference (2012)

[3] Burger, H.C., Schuler, C.J., Harmeling, S.: Image denoising: Can plain neural networks compete with BM3D? In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2392–2399 (2012)