ECS795P Deep Learning and Computer Vision, 2018

Name: Tanmaiyii Rao
College ID: 140361229

Course Work 1: Image Super-resolution Using Deep Learning

1. Suppose the settings of a SRCNN as: f1=9, f2=3, f3=5, how many pixels of the low-resolution image are utilized to reconstruct a pixel of the high-resolution image with the SRCNN? (10% of CW1)

Answer – In our particular example, the input (low resolution image) is 255 pixels. The formula we use to calculate the number of pixels utilized to reconstruct the high-resolution image is -

- Output size = (Input size - f1- f2 - f3 + 3)^2 – c

In this project, we have used a greyscale image so it consists of only 1 channel. Thus c =1. When we expand the formula and rearrange it we get -

Output size = input size – (f1-1) – (f2 -1) – (f3-1)

Output size = 255 – (9 -1) – (3 -1) – (5 -1)

Output size = 255 -14

Output size = 241

Thus, 241 pixels of the low-resolution image have been utilized to reconstruct the high-resolution image.

2. Why the deep convolutional model is superior to perform image super--resolution? Give one reason to explain it. (10%of CW1)

Answer- SRCNN learns an end-to-end mapping between low- and high-resolution images. This method differs fundamentally from existing external example-based approaches, in that SRCNN does not explicitly learn the dictionaries for modelling the patch space. These are implicitly achieved via hidden layers. Furthermore, the patch extraction and aggregation are also formulated as convolutional layers so are involved in the optimization. The entire SR pipeline is fully obtained through learning, with little pre/post processing. Moreover, the SRCNN is faster at speed, it is not only a quantitatively superior method, but also a practically useful one. Another important aspect is that SRCNN can process three color channels (either in YCbCr or RGB color space) simultaneously. With a lightweight structure, the SRCNN has achieved superior performance than the state-of-the-art methods. A typical and basic setting for SRCNN is f1 = 9 , f2 = 1 , f3 = 5 , n1 = 64 , and n2 = 32 . On the whole, the estimation of a high resolution pixel utilizes the information of $(9 + 5 - 1)^2$ = 169 pixels. Clearly, the information exploited for reconstruction is comparatively larger than that used in existing external example-based approaches, e.g., using $(5 + 5 - 1)^2$ = 81 pixels. This is one of the reasons why the SRCNN gives superior performance. Most importantly when compared to other methods, SRCNN gives the highest PSNR value and even gives a superior performance for other metrics as shown in the figures below.

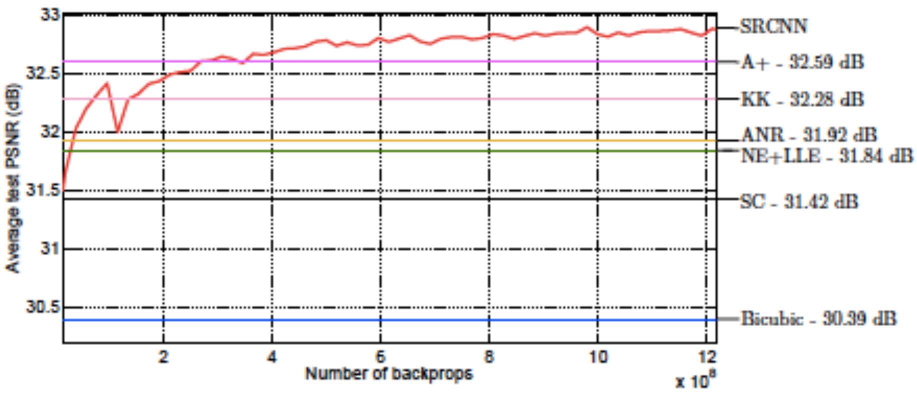

Fig. 10. The test convergence curve of SRCNN and results of other methods on the Set5 dataset.

TABLE 2

The average results of PSNR (dB), SSIM, IFC, NQM, WPSNR (dB) and MSSIM on the Set5 dataset.

Eval. Mat	Scale	Bicubic	SC [50]	NE+LLE [4]	KK [25]	ANR [41]	A+ [41]	SRCNN
PSNR	2	33.66	-	35.77	36.20	35.83	36.54	**36.66**
	3	30.39	31.42	31.84	32.28	31.92	32.59	**32.75**
	4	28.42	-	29.61	30.03	29.69	30.28	**30.49**
SSIM	2	0.9299	-	0.9490	0.9511	0.9499	**0.9544**	0.9542
	3	0.8682	0.8821	0.8956	0.9033	0.8968	0.9088	**0.9090**
	4	0.8104	-	0.8402	0.8541	0.8419	0.8603	**0.8628**
IFC	2	6.10	-	7.84	6.87	8.09	**8.48**	8.05
	3	3.52	3.16	4.40	4.14	4.52	**4.84**	4.58
	4	2.35	-	2.94	2.81	3.02	**3.26**	3.01
NQM	2	36.73	-	42.90	39.49	43.28	**44.58**	41.13
	3	27.54	27.29	32.77	32.10	33.10	**34.48**	33.21
	4	21.42	-	25.56	24.99	25.72	**26.97**	25.96
WPSNR	2	50.06	-	58.45	57.15	58.61	**60.06**	59.49
	3	41.65	43.64	45.81	46.22	46.02	**47.17**	47.10
	4	37.21	-	39.85	40.40	40.01	41.03	**41.13**
MSSSIM	2	0.9915	-	0.9953	0.9953	0.9954	**0.9960**	0.9959
	3	0.9754	0.9797	0.9841	0.9853	0.9844	**0.9867**	0.9866
	4	0.9516	-	0.9666	0.9695	0.9672	0.9720	**0.9725**

3. Please explain the physical meaning of peak signal-to-noise ratio (PSNR) in the context of image super-resolution. PS: place here the ground truth (GT) image, and the high-resolution images by SCRNN (HR-SRCNN) and bicubic interpolation (HR-BI) for reference. Also put the PSNR value below the high-resolution images. (10% of CW1)

Answer - The peak signal to noise ratio (PSNR) is a widely-used metric for quantitatively evaluating image restoration quality, and is at least partially related to the perceptual quality. The PSNR block computes the peak signal-to-noise ratio, in decibels, between two images. This ratio is often used as a quality measurement between the original and a compressed image. The higher the PSNR, the better the quality of the compressed, or reconstructed image. The PSNR values for the reconstructed images using bicubic interpolation (BI) and SRCNN are given along with the respective images below. As can be seen, it is apparent that SRCNN gives a better PSNR value and a better quality image. The images have also been provided in the source files.

GT

HR-BI
(PSNR=20.453967418499577)

HR-SRCNN
(PSNR=21.77124830643141)

References

- Dong, C., Loy, C., He, K. and Tang, X. (2016). Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), pp.295-307.

- Uk.mathworks.com. (2018). *Compute peak signal-to-noise ratio (PSNR) between images - Simulink - MathWorks United Kingdom*. [online] Available at: https://uk.mathworks.com/help/vision/ref/psnr.html [Accessed 10 Feb. 2018].