Lossy compression pipeline using Convolutional Neural Network as retrieval process

Luc Senecal

Introduction

Super-resolution techniques aim to construct better quality images from low resolution images.
Most advance algorithm use neural network trained from pair of images, one with a high resolution (HR) and one with a low resolution (LR), the reconstructed image being labelled as super-resolution (SR) image. The idea of this process, namely from a high resolution image obtaining first a low resolution and reconstruct it to match the original is not without recalling the compression operation. In this poster, I investigate the idea to use Convolution Neuronal Network to establish a lossy compression mechanism.

Description of the Model

The proposed compression pipeline would be as follow :

- Sub-sample an input image reducing its spatial size
- Reconstruct the original image by using a CNN for super-resolution

To trained the CNN, I used a modified version of the implementation of Shi et al. (2016). The neural network consists of 4 layers :

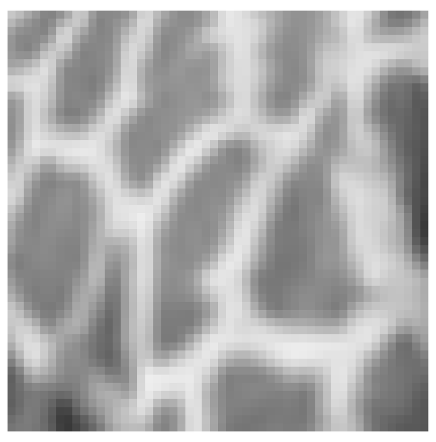
- A first layer with 64 feature maps of block 5x5
- A second layer with 64 feature maps of block 3x3
- A third layer with 32 feature maps of block 3x3
- And finally a sub-pixel convolution layer

The network was trained for 100 epochs using the Keras library on python with a learning rate of 0,001 using the Relu as activation function.
I used the The Berkeley Segmentation Data Set and Benchmark 500 (BSDSB500). The network was trained on 300 images from the set with the remaining 200 utilized for testing purposes. The low resolution images were scaled by a factor of 2, corresponding to a compression factor of 4.

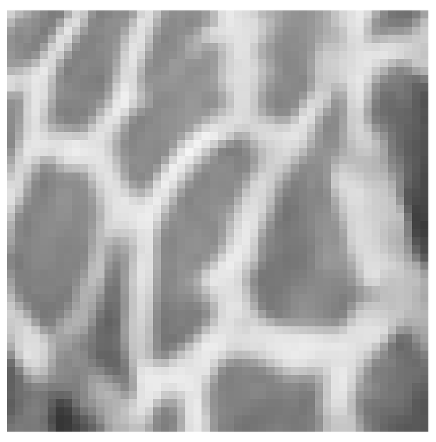
Visual results

Here are presented two images that have followed the pipeline process described in the previous section.
For convenience, only zoomed parts of these images are shown. I also include the results from an other CNN algorithm from the Gigapixel software for comparison.
As can be seen, both images reconstruct with convolutional neural network present better visual results than the bicubic interpolation.

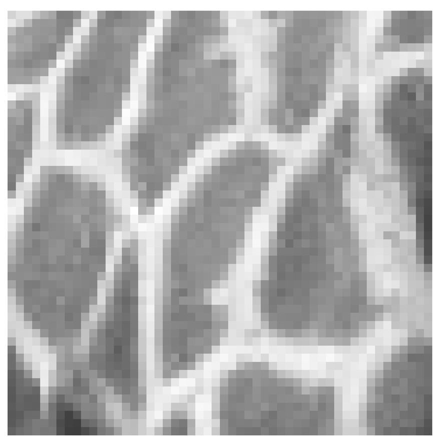
Original



Bicubic

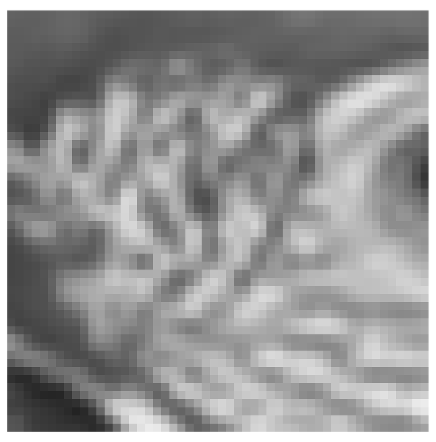


Trained model



Gigapixel

Original



Bicubic

Trained model

Gigapixel

A Metric problem ?

To show more accurate results, all images were submit to PSNR and SSIM metrics. Results are shown in the next table where P stand for the perroquet picture, G for giraffes picture, LR+ for the bicubic interpolation of the low resolution image and Giga for the Gigapixel software :

	P LR+	P SR	P Giga	G LR+	G SR	G Giga
PSNR	34.5231	36.5475	33.2074	26.6013	27.6860	24.5461
SSIM	0.9445	0.9550	0.9145	0.8254	0.8657	0.7449

Surprisingly, and for both images, PSNR values are really close for the CNN computation and the simple bicubic algorithm.
While, the visual aspect has definitely improved, PSNR metric do not corroborate these observations. SSIM values shows a more notifiable improvement but only for the trained CNN and improvement are relatively low.
As it is, the compression pipeline used here enable a poor compression ratio of 4 for a PSNR of 31.3678 (mean PSNR of 10 images from the test set).

However, one might suggest that PSNR and SSIM metrics can be irrelevant for a visual grading of images. To that extend, I used other metrics shown in the next section.

Using no-reference metrics

In this section, two new different types of metric where used :

- No-reference metric using perception based feature
- No-reference metric using trained model

For the non reference metric we used the PIQE metric (Venkatanath N et al. (2015)). For the model trained metric, we use BRISQUE metric (Mittal et al. (2012)) and NIQE metric (Mittal et al. (2013)). I used the Matlab implementation. Results are presented in the next table with the same convention as previous section. Both trained model metrics are positive values and show good result for low score. PIQE metric range from 0 to 100 and have its own grid for determining good results (lower is better).

	P HR	P LR+	P SR	P Giga
BRISQUE	1.88	33.87	29.46	28.73
NIQE	3.62	4.02	3.84	3.93
PIQE	14.86	59.04	50.81	22.15
	G HR	G LR+	G SR	G Giga
BRISQUE	13.04	39.39	40.67	22.75
NIQE	2.58	4.54	5.14	3.76
PIQE	31.61	46.72	30.70	25.43

All three metrics rate the SR image and the Giga image of the perroquet as having a better quality than the bicubic interpolation. For the perroquet, the PIQE metric even label the Giga image as "Good" with a score of 22.15, close to the limit of 20 for "Excellent" image. However, for the giraffes images, the two trained metrics find the bicubic interpolation to be a better quality image, while the PIQE metric find the giraffes Giga image better than the high resolution image (original image).

Conclusion

The idea for a compression pipeline using sub-sampling for compression then using Convolutional Neural Network in order to retrieve the original image is explored.
From the nature of Super-Resolution algorithm, the set up of such pipeline is easy and achievable. On one hand, I have found that the process has poor performance when measured with PSNR and SSIM metrics. On the other hand, metrics more focus on observer perception have suggest better performance for CNN reconstruction.
One key element to keep in mind, in the simplicity of the network used here (trained for less than an hour). More complicated networks that would need longer training session might provide better results.
However, the low compression rates for low PSNR values found in this work strongly suggest that compression pipeline with CNN will unlikely be used in the future.

References

A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 2012.

A. Mittal, R. Soundararajan, and A. C. Bovik. Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters*, 2013.

Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. 2016.

Venkatanath N, Praneeth D, Maruthi Chandrasekhar Bh, S. S. Channappayya, and S. S. Medasani. Blind image quality evaluation using perception based features. In *2015 Twenty First National Conference on Communications (NCC)*, 2015.