International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | March - 2023 Impact Factor: 7.185 ISSN: 2582-3930

Image Super-Resolution Via a Convolutional Neural Network

B.Sai Bhavani[1], B.RennyPaul[2], B.GandhiPrakash[3], B.Satishkumar[4]
Dept. of E.C.E,

R.V.R. and J.C College of Engineering

1 Abstract

We proposed 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 tradeoffs between performance and speed. Moreover, we extend our network to cope with three color channels simultaneously, and show better overall reconstruction quality.

keywords:Super-resolution, Deep Convolutional Neural Network, Sparse codin.

2 Introduction

Image super-resolution is the task of recovering a high-resolution image from a lower-resolution image. This problem is notable for its applications in security as well as in medical imaging, especially since image reconstruction offers a methodology for correcting imaging system imperfections.
For our project, we implement SRCNN and refine the model in order to improve the quality of the output images, as measured by peak signal-to-noise ratio (PSNR). The input to our algorithm is a low-resolution image, which we feed through a convolutional neural network (CNN) in order to produce a high-resolution image. Traditional methods for image upsampling rely on low-information, smooth interpolation between known pixels. Such methods can be treated as a convolution with a kernel encoding no information about the original photograph. Although they in- crease the resolution of an image, they fail to pro- duce the clarity desired in the super-resolution task. Convolutional Neural Networks (CNNs) are a gen- eralization of such algorithms, using learned kernels with nonlinear activations to encode general charac- teristics about photographs that can add structure lost in the low-resolution input. CNN architectures such as SRCNN [2] have been successfully applied to the super-resolution Overall, the contributions of this study are mainly in three aspects:

1) We present a fully convolutional neural network for image super-resolution. The network directly learns an end-to-end mapping between lowand high-resolution images, with little pre/postprocessing beyond the optimization.
2) We establish a relationship between our deeplearning-based SR method and the traditional sparse-coding-based SR methods. This relationship provides a guidance for the design of the network structure.
3) We demonstrate that deep learning is useful in the classical computer vision problem of superresolution, and can achieve good quality and speed.

3 Related work

Convolutional Neural Networks (CNN) are everywhere. It is arguably the most popular deep learning architecture. The recent surge of interest in deep learning is due to the immense popularity and effectiveness of convnets.CNN is now the go-to model on every image related problem. In terms of accuracy they blow competition out of the water. It is also successfully applied to recommender systems, natural language processing and more. The main advantage of CNN compared to its predecessors is that it automatically detects the important features with- out any human supervision.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | March - 2023 Impact Factor: 7.185 ISSN: 2582-3930

Common SISR methods[7] can be divided into three categories: interpolation-based methods, reconstruction-based methods, and learning-based methods, the last of which has received the major- ity of recent attention for its ability to yield accu- rate results given prior knowledge without the slow speed of reconstruction-based super-resolution meth- ods. Very recently, deep learning-based SISR meth- ods have shown the most promise.
Bicubic interpolation [3] is an exact, iterative, interpolation-based method commonly used because of its simplicity and speed. This is a good baseline with which to compare deep learning-based methods, and it also happens to be part of the required preprocessing for the SRCNN method. The key drawback of bicubic interpolation as super-resolution is that it does not make use of outside knowledge and thus does not lend any additional information to the upscaled image. It also has a smoothing effect, which ultimately limits the performance of SRCNN
To circumvent the issues posed by the initial bicubic upscaling, we [6] developed ESPCN (a more successful version of its predecessor, FSRCNN), which extracts feature maps in low-resolution space and introduces an efficient sub-pixel deconvolution layer whose learned filters upscale the previous layer (rather than the usual downscaling). ESPCN is also more efficient because it works on smaller-dimensional layers, and it can be used on real-time HD video.
A subsequent class of models expanded on the ESPCN approach, and took advantage of its superior performance, to increase the depth of super-resolution models and increase the solution space, which was found to greatly improve accuracy. DRCN and VDSR models [4] achieve the increased depth by cascading small filters many times (both networks are roughly 20 layers), which has the result of preserving contextual information over large image regions.
The current state of the art is achieved by models like EDSR and MDSR [5], which take advantage of the insight that whereas many CNN architectures are designed to recognize highly abstract inner representations, the task of super-resolution is localized and largely scale-free. As such, the developers of these methods utilize a number of “tricks” to accelerate training and improve performance, including sharing weights between models designed for different scales. Although more advanced methods than SRCNN exist, our aim by using this simpler model is to reimplement a performant super-resolution model, and by attempting to improve it, gain practical insight into how to better apply deep learning to the task of super-resolution.

4 Flow chart

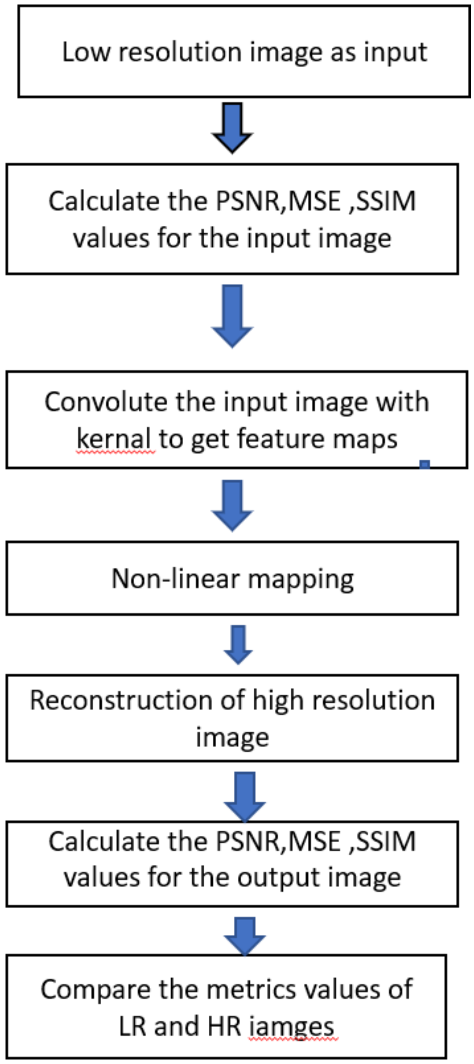

Figure 1: flowchart

5 Dataset and Features

We draw our dataset from DIV2K [1], a dataset of 1000 high-resolution images with diverse subjects. Due to the constraints on GPU resources, we used a subsample of 100 images, with a 60/20/20 train/val/test split. Because the images in the DIV2K dataset have varying sizes, we first center crop them to 800x800 and then downsize to 224x224. In order to generate our low-resolution images, we resize them to half that size and then resize them back up to 224x224, producing an image of lower resolution.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | March - 2023 Impact Factor: 7.185 ISSN: 2582-3930

Figure 2: Example of input image (left) and target image (right).

6 Architecture

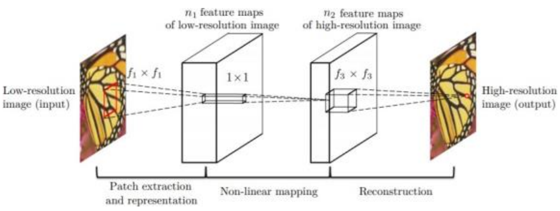

Figure 3: SRCNN architecture

SRCNN comprises three convolutional blocks corresponding to patch extraction, non-linear mapping, and reconstruction.

1) Patch extraction and representation: this operation extracts (overlapping) patches from the lowresolution 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.

2) Non-linear mapping: this operation 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 maps.

3) Reconstruction: this operation 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.

The first layer convolves the input image with a 9x9 kernel with padding and expands the three-channel image into 64 feature maps. The second layer applies a 1x1 kernel to condense to 32 feature maps. The third layer applies a 5x5 kernel to generate the output image. Both the first and second convolutional layers are succeeded by a ReLU activation function. Convolutional blocks can be treated as a subset of linear models with a small number of parameters. The 9x9 kernels of the first SRCNN layer apply thesame 81 parameters (values for each pixel of the ker-nel) to a full image, such that each pixel of the result-ing image corresponds to the level of similarity be-tween the corresponding 9x9 pixel patch of the inputimage and that kernel. The nonlinear mapping stepconnects these input feature maps to a reduced set ofoutput feature maps. The 5x5 kernels can be viewedas higher-resolution/granularity image features represented by the pixels of the output feature maps.

6.1 Metrics

For training and validation, we used mean squared error (MSE) for our loss function.

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n}(h(x^i;\theta) - y(i))^2$$

In order to evaluate the performance of the model, we also used peak signal-to-noise ratio (PSNR), a common metric for evaluating the quality of the image restoration,

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

where MAX2 I is the maximum possible pixel value of the image. While the numerator captures the sig- nal strength, the denominator represents the noise strength. The higher the peak signal-to-noise ratio, the better the reconstruction. Consequently, minimizing MSE leads to a maximization of PSNR.

7 Results

By using the provided mathematical expressions Peak Signal to Noise Ratio(PSNR), Mean Square Error(MSE),Structural Similarity Index(SSIM) are calculated.The metrics values of degraded image and output image are compared. PSNR value is increased and MSE value is decreased and SSIM value is increased. Higher the value of PSNR higher will be the resolution.

8 conclusion

We have presented a novel deep learning approach for single image super-resolution (SR). We show that conventional sparse-coding-based SR meth- ods can be reformulated into a deep convolu- tional 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.

References

1. Eirikur Agustsson and Radu Timofte. “NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study”. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. July 2017.

2. Chao Dong et al. “Image Super-Resolution Using Deep Convolutional Networks”. In: CoRR abs/1501.00092 (2015). arXiv: 1501.00092. URL: http://arxiv.org/abs/1501.00092

3. R. Keys. “Cubic convolution interpolation for digital image processing”. In: IEEE Transactions on Acoustics, Speech, and Signal Processing 29.6 (1981), pp. 1153–1160

4. Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. “Accurate image super-resolution using very deep convolutional networks”. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 1646–1654

5. Bee Lim et al. “Enhanced deep residual networks for single image super-resolution”. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017, pp. 136–144.

6. Wenzhe Shi et al. “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network”. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 1874–1883

7. Wenming Yang et al. “Deep learning for single image super-resolution: A brief review”. In: IEEE Transactions on Multimedia 21.12 (2019), pp. 3106–3121

8. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: ACM Multimedia. pp. 675–678 (2014)

9. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105 (2012)