As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 8(03):258-263, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0803048
Available online at: http://www.ijmtst.com/vol8issue03.html

Analytical Modeling for Single Image Super Resolution

Ashutosh Dubey[1*] | Chirag Keshri[1] | Divya Arora[1] | Poornima Suryavanshi[1] | Monica Bhutani[2]

[1]Department of Electronic and Communication Engineering, Bharati Vidyapeeth's College of Engineering, India.
[2]Assistant Professor, Department of Electronic and Communication Engineering, Bharati Vidyapeeth's College of Engineering, India.
[*]Corresponding Author Email ID: ashud12311@gmail.com

To Cite this Article
Ashutosh Dubey, Chirag Keshri, Divya Arora, Poornima Suryavanshi and Monica Bhutani. Analytical Modeling for Single Image Super Resolution. International Journal for Modern Trends in Science and Technology 2022, 8(03), pp. 258-263. https://doi.org/10.46501/IJMTST0803048

Article Info
Received: 16 February 2022; Accepted: 19 March 2022; Published: 24 March 2022.

ABSTRACT

For the super-resolution of a single image, the project used a Deep Learning Model (Convolution Neural Network). The method aspires for a direct end-to-end mapping between low-resolution and high-resolution images. The model is essentially a CNN-based deep learning model, with the input being a low-resolution image and the output being the equivalent high-resolution image. The model has a relatively light construction, yet it has state-of-the-art restoration quality and is fast enough for practical online use. We experiment with alternative network architectures and parameter choices to accomplish trade-offs between display and speed. Furthermore, we improve the overall reconstruction quality by extending our network to handle three color channels simultaneously.

KEYWORDS: Super Resolution, Deep Learning, Computer Neural Network, Computer Vision.

1. INTRODUCTION

The difficulty of creating a high-resolution image from a low-resolution image is a well-known one in Computer Vision. Because several solutions exist for any given low-resolution image, this problem is fundamentally ill-posed. To put it differently, it is an inverse problem with no single solution. The solution space is often strained with prior solid information to solve this problem. To begin, most models adopt the example-based methodology. They change the interior characteristics of the same images or get functions from other low resolution/high-resolution pairs.

The external techniques can be devised for general picture super-resolution or designed to accommodate field-specific tasks, i.e., face hallucinating, according to the training examples provided. The sparse-coding-based technique is one of the exemplary external example-based SR methods. The image created is decrypted by a low-resolution function. The sparse coefficients are processed into a high-resolution dictionary for rebuilding high-resolution patches. The overlapping re-constructed patches are combined (e.g., by weighted averaging). The project proves that the above pipeline is similar to a deep convolutional neural network. Inspired by this fact, we considered a CNN, which can directly learn the end-to-end mapping between the low- and high-resolution image. Our method is essentially different from the previous methods based on external samples in that it does not explicitly learn the dictionary used to model the patch space.

These are implemented through the hidden layer in a roundabout way. Patch extraction and aggregation are also stated as a convolutional layer, implying that optimization is required. The complete SR pipeline is obtained completely through learning, with just about little pre-processing/post-processing in this method. The proposed model has several appealing properties. First, its structure is designed with simplicity in mind and provides superior precision compared with state-of-the-art example-based methods. Figure 1 itself shows a comparison of an example. Secondly, this method achieves fast speed for practical online usage with moderate filters and layers, Since our method is fully fed forward and does not involve solving a consumption optimization problem, it is faster than numerous existing example-based strategies.

Third, investigations suggest that when more prominent and more diversified datasets are provided, as well as (ii) a more comprehensive and in-depth model is applied, the network's restoration quality can be further improved. Larger datasets/models, on the other hand, can cause problems for existing example-based techniques. Overall, the study's contributions are mainly in three aspects: We are presenting a fully convolutional neural network for image super-resolution. With very little processing beyond the optimization, this network learns an end-to-end tracking between low-resolution and high-resolution images. We establish an association between our deep learning-based SR method and the traditional sparse-coding-based SR methods. This relationship guides the design of the network structure. We show that deep learning effectively achieves high quality and speed in the classic computer vision problem of super-resolution. The present work enhances the initial version in significant ways. Firstly, we improve the SRCNN by introducing superior filter size in the nonlinear mapping layer and explore deeper structures by adding nonlinear mapping layers. Secondly, considerable new analyses and instinctive explanations are added to the initial results.

OBJECTIVES

Complicated deep neural network models have given better super-resolution performance than conventional super-resolution techniques. However, they are difficult to implement in reality, particularly on devices with lower complexity, power, and memory.

The method aspires for a direct end-to-end mapping between low-resolution and high-resolution images. This paper aims to convert a low-resolution image into a high-resolution image using Deep Learning (CNN). We propose an alternative deep learning approach for single image super-resolution (SR). We show that the traditional SR method based on sparse coding can be reformulated as a deep convolutional neural network., with almost no additional pre-/post-processing other than optimization.

2. RELATED WORK

There are numerous works that have been done related to image processing machine learning algorithms.

The super-resolution algorithms are of four types - prediction models, edge-based methods, statistical image methods, and patch-based methods. They are already tested and investigated in Yang et al.'s work [1]. The internal example-based methods decreased the images' similar characteristics and produced exemplar patches from the input image.

It is mentioned in Glasner's work [2], and many improved versions are put forward to speed up the work. The external example-based models train functionality between low- and high-resolution patches from external datasets. These ideas disagree on how to train a dictionary to operate with low and high-resolution images and on how representation methods should be done in such work.

In the work of Freeman et al. [3], the dictionaries are presented as low/high-resolution patch pairs. Therefore, according to the algorithm, the nearest neighbors of the input patch are found within the low-resolution space, with its corresponding high-resolution spot used for reconstruction. Other mapping functions such as kernel regression, simple function, random forest, and anchored neighborhood regression are proposed to improve mapping accuracy and speed [4].

Most of the sparse coding methods and their improvements [5-7] are among the current state-of-the-art SR methods. The main optimization focus is the patches; the patch mining and combination steps are considered pre/post-processing and handled separately. The majority of SR algorithms focus only on one channel among the three channels. For color images, the methods mentioned above first transform the problem to a different space like YCbCr, and SR is applied only on the Y channel, i.e., the luminous channel. Super-resolve all channels.

A Convolutional Neural Network may be a Deep Learning algorithm that may grip an input image, assign importance to diverse aspects/objects within the image, and differentiate one from the other. The pre-processing required during a ConvNet is lesser than other classification procedures. While in nascent methods, filters are hand-engineered, with enough training, ConvNet has the major ability to learn from these filters [10]. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons within the Human Brain and was stimulated by the organization of the graphic area.

Individual neurons answer stimuli only during a restricted region of the field of vision referred to as the Receptive Field. A collection of such fields overlaps to hide the visual cortex [10]. 2.3 Deep Learning for Image Restoration Various studies have been performed for deep learning techniques for changing low-resolution images into high-resolution images. The multi-layer perceptron (MLP), which also has fully connected layers (as opposed to convolutional), is being used to create realistic images. [11]. More intimately connected to our study, the convolutional neural network is beneficial for natural picture denoising [11] and suppressing noisy patterns. These restoration challenges are more or less denoising-driven [4]. Cui et al. [12] recommend using a core example-based approach to embed auto-encoder networks in their super-resolution pipeline[2]. Although each layer of the cascade involves independent optimization of the self-similarity search process and, as a result, the auto-encoder, the deep model isn't expressly envisioned as an end-to-end solution. On the contrary, the proposed model optimizes end-to-end mapping. Further, our model is faster at speed. It is not only a quantitatively larger method but also an essentially useful one [4].

There are four types of layers in a convolutional neural network - Convolutional layer, ReLU layer, Fully connected layer, Pooling layer.First, we have our convolutional layer, which is the central aspect of processing images in the oral evolution network; that's why we have it. Then that's going to be feeding in, and then we have the ReLU layer is the method behind how layers get activated, what makes a neuron fire. Pooling is a neural network term that is very commonly used, which means pooling out information. Then finally, we have our fully connected layer that's where the output will come out.

The convolutional layer has several filters and performs convolution operations. Each pixel is represented as a matrix of pixels. Take into account the following 5x5 image whose pixel values are merely 0 and 1 currently; basically, while we are dealing with colors, that's all the kind of image that pops in color processing.

3. CONVOLUTION NEURAL NETWORK

Owing to the increase in image and graphic data, their application in many fields ranging from GIS mapping to criminal detection, the research, and algorithms of getting a high-resolution image from a low-resolution image in the field of super-resolution is growing day by day. Deep learning technologies have shown good performance in image processing and resolution. Recent studies lean more towards a convolutional neural network (CNN) based super-resolution techniques instead of conventional pixel-based interpolation techniques.

As CNN-based deep learning technologies have shown good performance in image resolution, ant CNN-based super-resolution techniques have been developed. Recently, learning-based image SR using convolutional neural network (CNN) has achieved significant improvement and competes favorably than the other methods and the great success of CNN in computer vision researchers.

Available image super-resolution methods are based on: Interpolation, Reconstruction, Learning.

Defining Convolution Neural Network

Convolutional neural networks (CNNs) are deep learning neural networks that learn the relationships between input and output data. When given an input image, a machine learning algorithm assigns priority to different image components based on learnable weights and biases, allowing them to be distinguished from others.

CNN collects information from the images by extracting features. The following are the components of any CNN - Input layer(which is a grayscale image), Output layer involves multi-class labels, Hidden layers include ReLU, i.e., rectified linear unit, convolution layers, fully connected Neural Network, and pooling layers.

ReLU

ReLU implies a rectified linear unit. RELU is a phenomenon of imposing an activation function to enhance the nonlinearity of a network independent of the receptive field's layers of Convolution.

It allows faster training of the data, and Leaky ReLU is utilized to tackle the issue of vanishing gradient. Other activation functions include Randomized Leaky ReLU, Exponential Linear Units (ELU), Parameterized ReLU, etc.

4. ANALYTICAL MODELING

Complicated deep neural network models have given better super-resolution performance than conventional super-resolution techniques. However, they are difficult to implement in reality, particularly on devices with lower complexity, power, and memory due to the enormous network parameters and convolution operations of deeper and much denser networks.

Taking a single low-resolution image, firstly, we pre-process it to the desired size by applying bicubic interpolation. Suppose we represent the interpolated image as 'y'. Then the aim is to recover' f(y) 'from 'y', which is very close to the high-resolution version of the original image. So, we can say y is a low-resolution image used to recover x, which is already a high-resolution image.

We have to train the model and make a function 'f,' which conceptually consists of three operations :

- Patch extraction and representation: This process helps to extract patches from low-resolution images, then represent each patching in the form of a high-dimensional vector. These vectors have a set of feature maps, which equals the dimensions of the vector.
- Nonlinear mapping: This operation is harnessed to map each high-dimensional vector upon another high-dimensional vector. Each mapped vector represents a patch that is of high resolution. These vectors themselves consist of other sets of feature maps.
- Reconstruction: For generating the final high-resolution image, which is intended to be equivalent to the ground truth x, this operation aggregates high-resolution patch-wise representations. In our experiment, we will finally demonstrate these operations from a convolutional neural network.

5. METHODOLOGY

We have to train the model and make a function 'f,' which conceptually consists of three operations as mentioned above in analytical modeling.

Process Description

The following diagram makes it easier to understand how we proceed.

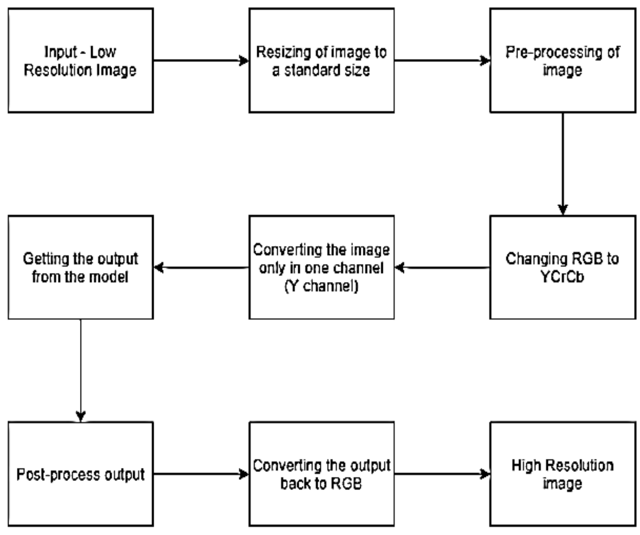

Figure 1 : Process Flow Diagram

- Patch extraction and representation : Usually, to restore an image, a common strategy is to extract patches, and then a collection of trained bases to match them. For example, Principal component analysis (PCA), Discrete Cosine Transform (DCT), Haar, etc. is almost similar to Convolution of the image by a set of filters, all of which form a basis. In our formulation, we include the optimization of bases in optimizing the network.Our first layer is expressed as a function of f1,f1: f1[y)] = max [0, w1 ∗ y + b1].

  Where w1 and b1 stand for filters and biases, respectively,'∗' stands for convolution. w1 denotes n1 support c×f1 ×f1 filters, where c denotes the number of channels in the input image and f1 denotes the spatial size of the filter. Hence w1 applies n1 convolutions on the image, and each Convolution's kernel size is written as c × f1 × f1. B1 is an n1 dimensional vector having biases; each element is associated with a filter.So, the output consists of n1 feature maps. We implement the Rectified Linear Unit (ReLU, max (0, x)) on the filtered image.

- Non-linear mapping : n1 dimensional feature is extracted in the first layer of each patch. The second operation, mapping of each of the n1-dimensional vectors to an n2 dimensional vector, is performed in the second operation. The second operation is similar to the application of n2 filters with spatial support of 1×1 dimensions. The above interpretation is only applicable for filters of 1×1dimension. But it is convenient to generalize filters of larger dimensions, say 3×3 and 5×5. In such scenarios, the nonlinear mapping is not performed on the input image's patches; rather, nonlinear mapping is performed on a 3×3 or 5×5 "patch" of the feature maps. The second layer's functioning is as follows, $f2(y) = \max [0, w2 * f1[y] + B2]$.

Where, w2 is made up of n2 filters with sizes of n1 ×f2 ×f2 and B2 is n2-dimensional. Each of the n2-dimensional vectors in the output represents a high-resolution patch that will be used to reconstruct the image.To increase nonlinearity, more convolutional layers can be added. But it will increase the model complexity (n2 × f2 × f2 × n2 for one layer), so it will take more time to train them.

- Reconstruction : The average of the predicted overlapping high-resolution patches is often taken to obtain the final full resolution image in typical reconstruction methods. Taking an average of predicted overlapping high-resolution patches can be taken as a predefined filter on a set of feature maps correspondingly to each position is the flattened vector form of a high-resolution patch.
- Influenced by the above fact, here we define a convolutional layer to obtain our final high-resolution image:

$$f[y] = [\, w3 * f2[y]] + b3.$$

Where, w3 - c filters of dimensions n2 × f3 × f3; b3 implies a c-dimensional vector. We expect the filters to behave as a 16-averaging filter if the representations of the high-resolution patches are in the image domain. If the representations of the high-resolution patches are in different domains, we expect w3 to perform similarly to projecting the coefficients onto the picture domain and then averaging.Whatever the representations, w3 is always a set of linear filters.

- Notwithstanding that different intuition inspires the above three operations, they all take us to the same form as a convolutional layer. By integrating all these actions together, we construct a convolutional neural network.

6. FUTURE SCOPE AND CONCLUSION

This paper used Super-Resolution Convolutional Neural Network to enhance the degraded image. In the proposed method, SRCNN learns an end-to-end mapping between low-resolution and high-resolution images, with no additional pre/post-processing than optimization. We have achieved superior performance with its lightweight structure than the state-of-the-art methods. Test output by using the SRCNN method is shown below:

Figure 2 : Image 1

Figure 3: Image 2

Table 1 : Classification Report for Test Image 1

Parameters	Degraded	SRCNN
PSNR (Peak Signal to Noise Ratio) measure)	32.82935	32.86873
MSE (Mean Square Error)	101.68723	100.76937
SSIM (Structural Similarity Index	0.916427	0.916010

Table 2: Classification Report For Test Image 2

Parameters	Degraded	SRCNN
PSNR	35.55319	38.69629
MSE	54.55319	26.33732
SSIM	0.96978	0.98143

Hence, this paper aims to convert a low-resolution image into a high-resolution image using Deep Learning (CNN). We propose an alternative deep learning approach for single image super-resolution (SR). We show that the traditional SR method based on sparse coding can be reformulated as a deep convolutional neural network., with almost no additional pre-/post-processing other than optimization.

With its lightweight structure, SRCNN has achieved superior performance than state-of-the-art methods. We speculate that additional performance can be obtained by exploring more filters and different training strategies. In addition, the suggested structure has the advantages of simplicity and resilience and may apply to various low-level vision challenges, such as image deblurring or synchronous SR+ denoising. People can also research a network to deal with different upgrade factors.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Yang, C.Y., Ma, C., Yang, M.H.: Single-image super-resolution: A benchmark. In: European Conference on Computer Vision, pp. 372–386 (2014)

[2] Glasner, D., Bagon, S., Irani, M.: Super-resolution from a single image. In: IEEE International Conference on Computer Vision. pp. 349–356 (2009)

[3] Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-based superresolution. Computer Graphics and Applications 22(2), 56–65 (2002)

[4] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence.38(2):295-307(2015)

[5] Timofte, R., De Smet, V., Van Gool, L.: Anchored neighborhood regression for fast example-based super-resolution. In: IEEE International Conference on Computer Vision.pp. 1920–1927 (2013)

[6] Timofte, R., De Smet, V., Van Gool, L.: A+: Adjusted anchored neighborhood regression for fast super-resolution. In: IEEE Asian Conference on Computer Vision (2014)

[7] Yang, J., Wang, Z., Lin, Z., Cohen, S., Huang, T.: Coupled dictionary training for image super-resolution. IEEE Transactions on Image Processing 21(8), 3467–3478 (2012)

[8] Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse regression and natural image prior. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(6), 1127–1133 (2010)

[9] Dai, S., Han, M., Xu, W., Wu, Y., Gong, Y., Katsaggelos, A.K.: Softcuts: a soft edge smoothness prior for color image superresolution. IEEE Transactions on Image Processing 18(5), 969–981 (2009)

[10] 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)

[11] Cui, Z., Chang, H., Shan, S., Zhong, B., Chen, X.: Deep network cascade for image superresolution.In: European Conference on Computer Vision, pp. 49–64 (2014).

[12] N. Akhtar and A. Mian, "Hyperspectral recovery from RGB images using Gaussian processes", IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 1, pp. 100-113, Jan (2020).

[13] S. Mei, R. Jiang, X. Li and Q. Du, "Spatial and Spectral Joint Super-Resolution Using Convolutional Neural Network," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 7, pp. 4590-4603, July (2020).

[14] J. Liu, Y. Xue, S. Zhao, S. Li and X. Zhang, "A Convolutional Neural Network for Image Super-Resolution Using Internal Dataset," in IEEE Access, vol. 8, pp. 201055-201070, (2020)

[15] Abbass, M.Y., Kwon, KC., Alam, M.S. et al. Image super resolution based on residual dense CNN and guided filters. Multimed Tools Appl 80, 5403–5421 (2021).