

# A Review on Medical Image Super Resolution with Application of Deep Learning

Kajol Singh

M Tech Scholar

Bansal Institute of Science and Technology,

Bhopal (M.P), India

kajolsingh01415@gmail.com

Manish Saxena

Professor

Bansal Institute of Science and Technology,

Bhopal (M.P), India

manish.saxena2008@gmail.com

**Abstract:** Super resolution problems are often discussed in medical imaging. The spatial resolution of medical images is insufficient due to limitations such as image acquisition time, low radiation dose or hardware limitations. Various super-resolution methods have been proposed to solve these problems, such as optimization or learning-based approaches. Recently, deep learning methodologies have become a thriving technology and are evolving at an exponential rate. We believe we need to write a review to illustrate the current state of deep learning in super-resolution medical imaging. In this article, we provide an overview of image resolution and the deep learning introduced in super resolution. This document describes super resolution for single images versus super resolution for multiple images, evaluation metrics and loss functions.

**Keywords:** LR, HR, CNN, DNN.

## I. INTRODUCTION

Despite the rapid development of imaging technology, imaging devices still have limited achievable resolution due to numerous theoretical and practical limitations. Super-Resolution (SR) technology offers a promising approach to computer-assisted imaging to generate high-resolution (HR) images over an existing low-resolution (LR) image or image sequences widely used in video surveillance, medical diagnosis of imaging are common. As well as radar imaging systems. SR has two main categories, namely Single Image SR (SISR) and Multiple Image SR (MISR) [1] [2]. SISR is useful in many ways because it contains an unlimited number of LR input images [3]. For this reason, this study focuses mainly on the SISR problem.

SISR is an underdetermined and therefore relatively difficult inverse problem because a given LR input image can have multiple solutions based on different texture details in its corresponding HR image. Two problems need to be addressed to produce high quality SR images. The first problem concerns

the conservation of unsatisfactory edges and the restoration of texture caused by insufficiently regulated stresses, since a single image provides very limited characteristic information. The second problem is related to the difficulty of quantitatively evaluating the estimated parameters due to the inability to measure their fundamental truth [4–5].

In recent years, deep neural networks (DNNs) have been used extensively in SR and have shown superior performance. Various fundamental methods have been developed for SR, such as non-uniform interpolation, frequency domain, and machine learning-based reconstruction approaches. While these methods can provide optimal or near-optimal images that increase resolution, they cannot guarantee improvements in detail, such as: Loss of high-frequency information and edge blurring.

To solve these problems, deep learning-based SR methods are being developed, as the mapping relationships of image characteristics from LR to HR can be fully examined and the reconstruction results show robustness and reliability, remarkable stability in multi-space -ladder. [6][7].

### A. Overview of Image Super Resolution

In the modern world, nothing has changed in the field of cameras and image sensors in recent years. Rapid advances in technology have created a sense of hope that in the near future cameras will be able to behave similarly to the human visual system in detecting natural scenes. Any camera can easily capture light intensity between 28 and 214. The term stop is used to measure intensity, which is actually the base 2 logarithm in the dynamic range. A dynamic range greater than 24 can be easily perceived by the human visual system, while 8-12 is the dynamic range that can be captured by a digital SLR camera.

With some overexposures and in exposure conditions, it may be necessary to make various settings in the camera using a high dynamic range.

The dynamic range of a natural scene is quite high compared to that of the camera. The lens range is quite small compared to our expectations for capturing a scene. Compared to HDR (High Dynamic Range Visuals), regardless of the bit depth of the camera, it is recommended to use LDR (Low Dynamic Range Visuals). HDR images are created by shooting images with different exposure times multiple times and stitching them together. This is done to extract an HDR image from an LDR image.

## II. LITERATURE REVIEW

Song et al. [8] is utilized for relative investigation reason. In this work PET pictures of cerebrum are taken as information for execution assessment of proposed model. The creator utilized picture patches to extricate the spatial area data and gave it into CNN model to distinguish the haze parts in PET pictures. The model was planned with 20 layer of CNN with different organization boundary mix.

KlemanGrm et al. [9] in this work named it as Cascaded super goal. This strategy is more vigorous than the bi cubic, closest neighbor introduction technique. For SR up scaling procedure CSRIP model is utilized and the lingering assessment is utilized with Identity earlier. The Cascaded Up gives 8x amplified picture and it likewise delivers an awesome outcome in contrast with the focused on picture which is scaled through SR module.

Choi et al. [10] proposed an engineering named as EUSR which comprises of four sections show in Discriminator and Score Predictor Multi scaling model and Multipass Manner. Different scaled pictures from a solitary model is created that is the reason Multi scale displaying is utilized.

Xunxiang et al. [11] proposed a CNN model named as SRCNN. With the pipeline of three stages that is covering patches, SRCNN produces high goal yield picture, thick removed and taking away mean. For preparing reason the Standard information (Set5) is utilized. Inadequate coding is utilized for the profound learning and the bi cubic scaling utilized for the up scaling and for quicker reaction time feed forward network is being utilized. One of the significant benefits of this model is its quicker computational quality.

George Seif et al. [12] for the acknowledgment reason for picture the creator proposed SICNN that pre-owned 12 x 14 size picture to accomplish an amplification of 8x with great perceptual quality. Along these lines, for the personality

measurements hyper circle space is utilized, for the preparation informational collection CNN utilized and similarly for the component extraction Euclidian distance is utilized.

## III. DEEP LEARNING

Traditionally, machine learning models have been trained to perform useful tasks based on manually designed functions extracted from raw data or functions learned from other simple machine learning models. Through deep learning, computers automatically learn useful representations and functions directly from raw data, bypassing this manual and this difficult step. By far the most common models in deep learning are different types of artificial neural networks, but there are others as well. The main common feature of deep learning methods is their focus on learning functionality: machine learning of data representations. This is the main difference between deep learning approaches and more "classic" machine learning. Recognition of functions and the execution of a task are combined in one problem and, therefore, both are improved in the same training process. See [13], [14] for a general overview of the region.

In medical imaging, the interest in deep learning is primarily driven by convolutional neural networks (CNNs) [15], a powerful method for learning useful representations of images and other structured data. Typically, before CNNs could be used effectively, these functions had to be developed manually or created from less powerful machine learning models.

Once it became possible to use the features learned directly from the data, many features of the handcrafted images tended to be left out as they proved nearly useless compared to the feature detectors found by CNN. There are strong preferences built into CNNs based on their construction, which helps us understand why they are so powerful. So let's take a look at the building blocks of CNN.

## IV. DEEP LEARNING IN SUPER RESOLUTION

### A. *Single image super resolution versus multiple image super resolution*

There are two types of super-resolution methods: single-image and multi-image methods. The goal is to create a high-resolution image from one or more low resolution images. Multi-image super-resolution is based on the fusion of information between low resolution images displaced by subpixels and generally allows for higher reconstruction accuracy.

Multi-image methods generally use global / local geometric or photometric relationships between multiple low-resolution images. Existing techniques include interpolation-based methods, frequency domain methods [16] or regularization-based methods [17].

As described below, deep learning approaches have greatly improved the performance of single-frame super-resolution methods. There are very few deep learning methods applied to multi-image super resolution.

In [18] the authors use a deep residual network to improve the results of an evolution model for the super resolution of different satellite images. Below we focus on the super resolution of a single image.

*B. Evaluation metrics and loss functions*

Loss functions are used as rebuild score metrics and for model optimization. Peak Signal-to-Noise Ratio (PSNR) is the most widely used measure of reconstruction quality for super resolution. Let L be the maximum pixel value, N the number of pixels, I the basic truth image and  $\hat{I}$  the reconstruction, the PSNR is defined as:

$$PSNR = 10 \cdot \text{Log}_{10} \left( \frac{L^2}{\frac{1}{N} \sum_{i=1}^N (I(i) - \hat{I}(i))^2} \right)$$

The PSNR refers to the mean square error (MSE) or the loss function L2 and provides information on the differences at the pixel level. L1 loss is more robust against outliers [19].

A trained neural network with this loss can converge faster and perform better [20]. These losses often lead to poor performance to very accurately represent the quality of the reconstruction in real images.

Therefore, other functions are used to get better quality results. The structural similarity index is also a widely used image quality index suitable for the human visual system. Measure structural similarity between images based on luminance, contrast, and texture.

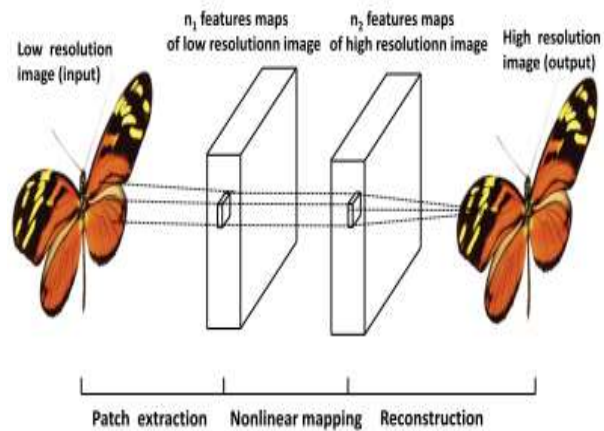


Fig. 1. Schema of SRCNN: a network consists of 3 layers. The first layer is used for patch extraction and representation, the layer in the middle corresponds to nonlinear mapping, the third layer reconstructs final SR images.

The poor perception quality of the high resolution images obtained by optimizing the root mean square error led to objective functions based on MSE in a transformed space. The loss of perception is based on the characteristics of a deep architecture.

The high resolution network is optimized by minimizing the MSE in the function space generated by a pre-trained VGG-16 network. The loss of functionality encourages the output image to be perceptually similar to the real image rather than forcing the pixels to match exactly. The loss of texture, which corresponds to the correlations between the different feature channels, has been proposed to produce more realistic textures.

For super resolution, the opposite losses are used to train the goose. Discriminator and generator learning is performed alternately. The following losses based on cross entropy are used for the generator,  $L_{generator}$ , and the discriminator,  $L_{discriminator}$  respectively:

$$L_{generator} = -\log D(\hat{I})$$

$$L_{discriminator} = -\log D(I) - \log D(\hat{I})$$

Where  $\hat{I}$  is the generated image and  $I$  is the basic truth image. Use the opponent's losses based on least squares for a more stable training process and better results. The different losses presented above are often combined, but the choice of weighting coefficients remains a problem.

V. TRENDS AND CHALLENGES

In addition to the promising performance achieved by DL algorithms in SISR, some key challenges and inherent trends remain as follows.

#### A. *Lighter Deep Architectures for Efficient SISR:*

Although the high accuracy of advanced depth models has been achieved for SISR, it is still difficult to use these models in real-world scenarios, mainly due to massive parameters and calculations. To solve this problem, we need to design shallow depth models or reduce the size of existing deep models for RHIS with fewer parameters and calculations at the cost of little or no performance degradation. Therefore, it is expected that researchers in the future will focus more on reducing the size of NNs in order to speed up the SISR process.

#### B. *More Effective DL Algorithms for Large-scale SISR and SISR with Unknown Corruption:*

In general, the DL algorithms proposed in recent years have significantly improved the performance of traditional SISR activities. However, the large scale of SISR and SISR with unknown corruption, the two main challenges in the SR community, still lack very effective corrective measures. DL algorithms are believed to be capable of solving many unattended inferences or problems, which is essential for addressing both of these challenges. Therefore, by harnessing the great power of DL, more efficient solutions to these two challenging problems are expected.

#### C. *Theoretical Understanding of Deep Models for SISR:*

The success of deep learning is said to be due to learning powerful representations. To date, however, we are unable to fully understand these representations and the deep architectures are treated as black boxes. In DL-based SISR, deep architectures are often considered a universal approximation and learned representations are often omitted for simplicity. This behavior is not useful for further research. Therefore, we need to focus not only on how a deep model works, but also on why and how it works. In other words, more theoretical studies are needed.

#### D. *More Rational Assessment Criteria for SISR in Different Applications:*

In many applications we need to design the desired objective function for a particular application. In most cases, however, we are unable to provide an explicit and precise definition for the assessment of application requirements, which leads to obfuscation of optimization goals. Many jobs, albeit for

different purposes, simply use MSE as an endpoint, which in many cases turned out to be a bad endpoint. We believe that in the future it will be very important to have clear definitions of notations in different applications. Based on these criteria, we can design more targeted optimization goals and compare algorithms more rationally in the same context.

## VI. CONCLUSION

This paper provides an overview of image resolution and the deep learning introduced in super resolution. This document describes super resolution for single images versus super resolution for multiple images, evaluation metrics and loss functions. Finally, the most important trends and challenges were discussed, lighter depth architectures for efficient SISR, more efficient DL algorithms for SISR and large-scale SISR with unknown corruption, and more rational evaluation criteria for SISR in different applications. Deep learning methods show great potential for solving SR in the medical imaging field. Despite many challenges, the performance of SR techniques is becoming more and more promising.

## REFERENCES

- [1] Zhengqiang Xiong, Manhui Lin "Single image super-resolution via Image Quality Assessment-Guided Deep Learning Network" PLoS One. 15(10): e0241313, 2020.
- [2] Ledig, C.; Theis, L.; Huszar, F.; Caballero, J.; Cunningham, A.; Acosta, A.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [3] Zhang K.; Tao D.; Gao X.; Li X.; Li J. Coarse-to-fine learning for single-image super-resolution. IEEE transactions on neural networks and learning systems 2016, 28, 1109–1122.
- [4] Yang, C.Y.; Ma, C.; Yang, M.H. Single-image super-resolution: A benchmark. European Conference on Computer Vision. Springer, 2014, pp. 372–386.
- [5] Zhang K.; Tao D.; Gao X.; Li X.; Xiong Z. Learning multiple linear mappings for efficient single image super-resolution. IEEE Transactions on Image Processing 2015, 24, 846–861. 10.1109/TIP.2015.2389629.
- [6] Ma C.; Yang C.Y.; Yang X.; Yang M.H. Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding* 2017, 158, 1–16.
- [7] Yang W.; Zhang X.; Tian Y.; Wang W.; Xue J.H.; Liao Q. Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia* 2019, 21, 3106–3121.
- [8] Tzu-An Song, Samadrita Roy Chowdhury, Fan Yang, and Joyita Dutta. 2020. Super-Resolution PET Imaging Using Convolutional Neural Networks. *IEEE Transactions on Computational Imaging*, 6, pp. 518-528
- [9] KlemanGrm and Walter J. Scheirer. 2020. Face Hallucination Using Cascaded Super-Resolution and Identity Priors. *IEEE Transactions On Image Processing*, 29.
- [10] Jun-Ho Choi et al. 2019. Deep Learning-based Image Super-Resolution Considering Quantitative and Perceptual Quality. *Neuro Computing Elsevier*.

- [11] Xunxiang Yao, QiangWu, Peng Zhang. 2019. Adaptive rational fractal interpolation function for image super-resolution via local fractal analysis. *Image and Vision Computing*82, pp. 39-49.
- [12] George Seif et al., 2018. Edge-based loss function for single image super-resolution. *IEEE. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 1468-1472.
- [13] Y. LeCun, Y. Bengio, G. Hinton “Deep learning” *Nature*, 521 (2015), p. 436.
- [14] I. Goodfellow, Y. Bengio, A. Courville “Deep learning” MIT Press (2016)
- [15] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner “Gradient-based learning applied to document recognition” *Proc IEEE*, 86 (1998), pp. 2278-2324
- [16] Nasrollahi K, Moeslund T. Super-resolution: a comprehensive survey. *Mach Vis Appl* 2014;25:1423–68.
- [17] Belekos S, Galatsanos N, Katsaggelos A. Maximum a posteriori video superresolution using a new multichannel image prior. *IEEE Trans Image Process* 2010;19:1451–64.
- [18] Kawulok M, Benecki P, Piechaczek S, Hryczenko K, Kostrzewa D, Naleoa J. Deep learning for multiple-image super-resolution. Preprint arXiv:1903.00440, 2019.
- [19] Lim B, Son S, Kim H, Nah S, Lee KM. Enhanced deep residual networks for single image super-resolution. In: *The IEEE conference on computer vision and pattern recognition (CVPR) workshops*, vol. 1; 2017. p. 3.
- [20] Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for neural networks for image processing. *IEEE Trans Comput Imaging* 2017;3.