JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

REVIEW OF SUPER RESOLUTION METHODS

[1]Gautami Kumbhar, [2]Priyadarshini Gopal, [3]Sphoorti Kadapa, [4]Tanya Raina,
Prabu I, Bindu Sadanandan, Dr. N. F. Shaikh
Modern Education Society's College of Engineering, Pune, Maharashtra, India

Abstract: In recent years, Super Resolution using traditional and deep learning methods has evolved greatly for images. The paper is an effort to review and throw light on several of the traditional approaches of Super Resolution and eventually delve into recent advances in the approaches using deep learning. In this survey, super resolution of images is categorised into the following methods: traditional approaches and Deep Learning methods (Convolutional Neural Networks (CNN), and Generative Adversarial Networks (GAN) and their variations/improvements).

Keywords: Image Super Resolution, Classical Methods, Deep Learning, CNN, GAN.

INTRODUCTION

High quality images have become a necessity today. They are needed for clear and accurate analysis of images, which would lead to precise conclusions. The effort required to analyse images becomes far less whenever the image being used happens to be of high resolution or sharp/clear, which would ultimately lead to efficient results. But, such clear or highly resolved images are not always easily available. This triggers the need to extract the essential data from the low resolved images present. Super Resolution (SR) of an image is defined as the process of obtaining a High Resolution (HR) image from an easily accessible Low Resolution (LR) image.

The approaches that are currently available for super-resolving images can be categorised as Single-Image Super-Resolution (SISR) and Multiple-Image Super-Resolution (MISR). SISR (Single Image Super Resolution) entails super resolving a single low resolution image, while using multiple low-resolution images is termed as MISR (Multiple Image Super Resolution). Out of the multiple applications of this approach, one is to aid in the analysis of satellite images. There are a number of applications where these can be used. For example, military surveillance, monitoring vegetation and animals, monitoring areas prone to natural calamities, etc.

Since the advancement of satellite image processing, remote sensing is becoming more relevant in modern life. But the inadequacy of current imaging sensors and the complexities of atmospheric conditions results in significant hurdles in remote sensing applications. As a result, super-resolution techniques for improving the quality of low-resolution remote sensing photos have received a lot of consideration.

Despite this, the use of SR in satellite images remains a largely unexplored domain. This is because of the many challenges that arise while resolving these images. Smaller objects like cars, buildings, ships, etc. become very difficult to identify accurately as these may /or may not have a very small spatial extent in low resolution satellite images. Some images may be rotation invariant, i.e. they are not affected when arbitrarily rotated, and can pose difficulty while identifying their orientation. Cloud removal also proves to be a major challenge while resolving satellite images. One of the major challenges for SR in satellite images is acquiring enough data to train the machine learning models.

This paper discusses the super-resolution techniques used right from the beginning including traditional techniques interpolation-based, reconstruction-based, learning-based methods followed by a discussion of the current technologies, such as Deep Learning using CNN and GAN. More precisely, the many different super resolution techniques are examined and explored. Furthermore, the applications of super-resolution are examined, and the possibilities of research in the future are proposed.

HISTORICAL OVERVIEW

A number of remote sensing applications require the satellite image to be of high resolution so that it can easily be processed and analysed. The major applications of high-resolution images are – better perception of it for human interpretation, and improving the representation of the image data for computers to process. In other words, Super Resolution can be generalised as the process of obtaining a High-resolution image from Low-resolution image/s.

Computational efficiency (due to a large number of unknowns, SR reconstruction necessitates costly matrix manipulations), image registration (a fundamental image processing problem that becomes more difficult in the SR setting), structural rigidity aspects (image degradation model parameters cannot be accurately estimated, and selectivity to anomalies may result in graphically disturbing artefacts), and performance limits (analysis of the performance limit of SR techniques may be difficult due to its complexity) are some of the known challenges in SR.

There are several approaches that can be used to super resolve the images and overcome these obstacles. These include the Image observation model, super-resolution in the frequency domain, interpolation restoration approaches, example-based approach, set theoretic restoration, etc. There are also various statistical approaches [1] like maximum likelihood, joint MAP restoration, Bayesian treatments, etc.

The main idea behind SR is to incorporate non-redundant information from numerous low-resolution frames to produce a high-resolution image. High-resolution images can be obtained in a variety of ways, which are categorised into:

1. Interpolation based:

Working in the wavelet domain makes it easier to use sub-bands for resolution improvements, and the authors applied various transforms such as Stationary Wavelet Transform (SWT), Discrete Wavelet Transform (DWT), and Curvelet Transform (CT) as needed. There are different interpolation methods, such as Nearest Neighbour, Bilinear, Bicubic, and so on [2]. However, among these numerous interpolations, Bicubic interpolation provides a high resolution image but introduces unwanted artefacts such as blurring [11], therefore the wavelet domain is useful to reduce the quantity of blurring artefacts [3].

In [3,] bicubic interpolation is employed to magnify the low-resolution input image. The image's edges are sharpened using the stationary wavelet transform (SWT). The sub-bands created by SWT would be altered by varying the boost value. The sub-band is then merged with Inverse Stationary Wavelet Transform (ISWT) to provide a high-resolution image.

SWT sub-bands are used for picture registration because regardless of how its input is adjusted, standard DWT does not provide the same outcome[4]. Following registration, a 45-degree rotation and up-sampling are suggested. Upsampling is performed by a factor of two. Missing coefficients are interpolated once upsampling creates space in the Curvelet domain. A single high-resolution image is formed by utilising the curvelet transform interpolation. On LR images, both SWT and DWT can be applied [5]. SWT develops each sub-band of an image's ratios, whereas DWT creates each sub-band half of the image's ratios. Before DWT, bicubic was used, however, it has no influence on the image size. Both transformations' output sub-bands will be sent through a Gaussian filter to reduce the blurring effect. The HR image will be generated after Inverse Discrete Wavelet Transform (IDWT) upsampling.

2. Reconstruction based:

Linear equations are created and pixel values can be related to HR and LR images utilising these equations.[3] Buffering amplitude self-interpolation was used and a new method has been developed for an edge directed SISR algorithm [8]. The proposed HR will preserve the features of the image while concealing aberrations. A Hybrid approach that integrates the advantages of the standard Maximum Likelihood estimator with the ability of the Projection onto convex sets to solve this problem has been proposed.[9] It has been discovered that an image with increased resolution may be restored using numerous unmoving blurred, decimated, and noisy photos. Traditional restoration tools could be directly extrapolated under this concept.

3. Learning-based techniques:

These techniques learn the correlations among LR and HR image overlays from a library of LR and HR image pairs, then apply them to a new LR image to recover its HR version [8], [6], [7], [10]. Example-based Super Resolution reconstruction approaches that use direct picture examples as a prerequisite for accurate regularisation have just been introduced. The lexicon is learned from intrinsic pictures and then applied to HR reconstruction [6]. The method necessitates a wider dictionary. A subspace technique for compact dictionary learning was proposed as a solution to this problem.[7] The amount of time spent is reduced, and the quality is improved. This approach, on the other hand, fails to eliminate undesirable artefacts. The learning method based on the subspace methodology has been extended to learn numerous bases or sub dictionaries with de-blurring, which suppresses undesired results[10]. It provides a single-image image-based SR technique depending on the knowledge of a cluster of LR-HR mapping relationships. The learnt mapping functions turn the low resolution input picture into a high resolution picture effectively and efficiently.

CHALLENGES

In order to develop high resolution imaging systems, one must overcome the challenge of diminishing returns. In particular, the chips that are required for imaging and optical parts important to catch high-resolution images become restrictively costly, sometimes costing up to millions of dollars. [16] Super-resolution is the term commonly applied to the issue of rising above or solving for the limits of optical imaging frameworks using algorithms for the purpose of image processing, which apparently is moderately modest to accomplish. Such algorithms will undoubtedly become more widely used in any situation where optical imaging machines of higher quality are either unavailable or prohibitively costly.

Super imposition of a sequence of LR or low quality images containing noise in order to achieve HR images is the central idea around which Super-Resolution revolves. Early work on Super-Resolution demonstrated that if there is the existence of a relative sub-pixel motion that remains between the under-sampled input images, aliasing effects in high-resolution fused images can be decreased (or even totally eradicated). [16] Notwithstanding, as opposed to the faint domain description of this early work, it tends to be seen that, as a rule, super-resolution is a computationally complicated and mathematically not well-presented problem. The advent of Deep learning provides a more profound view into these challenges.

DEEP LEARNING IN SR

Deep learning is a learning system based on the way humans learn. Deep learning is extremely beneficial when gathering, analysing, and interpreting large amounts of data; it makes the process efficient and faster.

Deep learning primarily may be thought of as a way to automate the identification of correlation between sensors. Deep learning algorithms are developed in a layer of increasing complexity and abstraction, unlike typical machine learning techniques, which are linear.

The main purpose of SR is to collect the information that is lacking in LR photos by super-resolving it, which gives rise to a number of potential uses. Today, there are plenty of algorithms or structures that are used for SR. Of these, the most popular ones can be grouped into two categories – Convolution Neural Networks (CNN) and Generative Adversarial Networks (GAN). These methods have a high ability for self-learning approaches to representing image data. [17].

1. Convolution Neural Networks (CNN)

Artificial Neural Networks (ANNs) as stated in [21] are computational processing systems that are based on the biological nervous systems that human brains function. Convolutional Neural Network (CNN) is one among the most important types of ANN architecture. These are generally employed for tough pattern recognition tasks in images and, due to their highly detailed yet basic architecture, provide an easier way to get started with ANNs.

CNNs consist of neurons that optimise themselves through learning. Each neuron will perform an operation based on the continuous input that it receives. This operation serves as the foundation for countless ANNs.

A CNN can be divided into these -

1. The input layer, as in other kinds of ANN, will store the image's pixel values.

2. The output of the neurons of the input is found by the convolutional layer. It calculates the scalar product between the weights and the region related to the input volume. An activation function is applied to the output of the previous layer's activation such as the rectified linear unit (ReLu) or sigmoid.

3. The pooling layer will down-sample along the spatial dimensions of the provided input, lowering the number of parameters inside that activation even further.

4. The fully-connected layers will then perform similar functions as typical ANNs, attempting to generate class scores from the activations for classification. It is also proposed that ReLu be utilised between these layers to boost performance.

2. Super Resolution Convolutional Neural Networks (SRCNN)

With the advancement of machine learning technologies, researchers are now focusing on a new learning-based method known as deep learning. Such as Super Resolution Convolutional Neural Networks (SRCNN) proposed by Dong et al. [22], has been a pioneer to researchers for developing more CNN-based algorithms. SRCNN solves the gradient-vanishing problem while consuming less memory and shortening the model's run time [23].

In deep learning or convolutional neural networks (CNN), we usually use CNN for image classification. Single image super resolution (SISR) is another application of SRCNN. In essence, with this better SR method, we can acquire a higher quality of a larger image even if we just have a little image, to begin with.

For super-resolving images, SRCNN proposes a three-layer CNN. The SRCNN architecture comprises three components: Feature extractor, non-linear mapping, and reconstruction. The training of the model is done to minimise the Mean-Squared Error gap between the restored and the referenced images on a pixel by pixel basis. A variety of model architectures and hyperparameters are tested and traded for performance and speed in the paper. [22]

3. Generative Adversarial Networks (GAN)

Ian J. Goodfellow [19], in 2014 introduced The Generative Adversarial Networks (GAN) which has then served as the backbone for deep learning algorithms.

A generator and a discriminator are the two models that make up GAN. Neural networks are commonly used to implement both of these. However, any type of discrete system that transfers data from one space to another can fulfil these goals.

For fresh data example generation, the goal of the generator is to capture the actual case distribution. The discriminator is usually a binary classifier that accurately distinguishes created instances from genuine examples. At a stationary position that is simultaneously a minimum for the generator and a maximum for the discriminator, the optimization grinds to a standstill.

The purpose of optimising GAN is to achieve Nash equilibrium [20] which is a concept in game theory that states that even if a player knows their opponent's strategy, they will stick to it since they have no reason to alter it.

4. Super Resolution GAN (SRGAN)

Despite significant progress in SISR (Single Image Super Resolution) using enhanced and efficient computational models, one of the primary difficulties to be tackled was the preservation of finer texture features when an image was super resolved at a higher upscaling factor.

The behaviour of the optimization-based super-resolution algorithms depends on which objective function is picked. The mean squared reconstruction error has been heavily researched nowadays. The estimates that emerged have excellent PSNR values. But these are perceptually not efficient enough as they fall short in high-frequency features and the quality desired at higher resolution.

An image super-resolution framework that uses a Generative Adversarial Network, SRGAN, is capable of producing images that are photo-realistic as well as natural for upscaling factors of 4x. SRGAN uses a function consisting of content as well as adversarial loss to achieve this. The function is called the perceptual loss function. The solution of the natural image manifold is forwarded by the adversarial loss using a discriminator. The network of the discriminator is pre-trained to be able to distinguish between an image that is photorealistic and original, and an image that has been super-resolved.

Furthermore, rather than pixel space similarity, a perceptual similarity that is motivated using the content loss is applied. On public benchmarks, the deep residual network is capable of procuring lifelike features from highly down-sampled photos. SRGAN exhibits tremendously substantial improvements in perceptual quality in a comprehensive mean-opinion-score (MOS) test.

The MOS scores achieved using the SRGAN are similar to those obtained with any state-of-the-art approach than those procured by the original high-resolution images.[22]

5. Enhanced Super Resolution GAN (ESRGAN)

These days, ESRGAN [17], which uses Residual-in-Residual Dense Block (RRDB), after the removal of batch normalisation and allows the prediction of the relative realness by the discriminator, instead of the absolute value[12], is commonly used for Single Frame Super Resolution. By utilising features prior to activation, it lowers perceptual loss, meaning that it could provide better control over brightness consistency and texture recovery. Although one of the main disadvantages of the method is that it is difficult to implement in real-time, as it requires high computational resources for fast processing. Applying deep learning methods to large datasets, i.e., big data also proves to be difficult. The methods also need to deal with the real-life complexities that arise while processing LR images in real-time. The structures that are currently used for SR need to be improved and enhanced to make them more robust, effective, efficient, as well as capable of handling large datasets without using too many resources [14]. The image quality of the recovered image of SRGAN is improved using the following modifications:

1. The first part of the modifications includes removal of the Batch Normalisation (BN). This has proved to improve the performance as well as decrease the complexity of computation in the various tasks that are PSNR-oriented (this includes SR). This is because the batch normalisation is known to bring artefacts.

2. The second part of the modification includes replacement of the residual block by using the Dense Block (that originates in the DenseNet), for the enhancement of the network.

Because of these improvements, the visual quality is better than SRGAN, with more genuine and lifelike textures. Furthermore, we improved perceived loss by utilising features prior to activation that provides greater supervision and hence restores more precise brightness and accurate textures.

CONCLUSION

This paper intends to explore various techniques used for Super Resolution of Images that have been implemented through the decades. This paper aims to shed light on the various techniques used for Super Resolution, from approaches that use statistical concepts, mostly favoured in the past, to currently used approaches that utilise machine learning and deep learning concepts. The paper also aims to throw light on the approaches that are used for Satellite Image Super Resolution. These can be used for research purposes in SR or any other field that can benefit from them.

REFERENCES

[1] Yang J.; Huang T. . “Image super-resolution: Historical overview and future challenges in Super-Resolution Imaging” (pp. 1-33). CRC Press. (2017)

[2] K. Ramavat; M. Joshi; P. Swadas. “A survey of Super resolution Techniques.” International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 12 Dec -2016.

[3] F. Arif; T. Sarwar. "Super-Resolution Using Edge Modification through Stationary Wavelet Transform." 2014 18th International Conference on Information Visualisation. IEEE, 2014. D. Davis & R. Roy. “Hybrid Super Resolution using SWT and CT”. International Journal of Computer Applications (0975 – 8887), Volume 59– No.7-2012.

[4] D. Davis; R. Roy. "Hybrid Super Resolution using SWT and CT." International Journal of Computer Applications, 59.7 (2012).

[5] S. Naik; N. Patel. "Single image super resolution in spatial and wavelet domain." arXiv preprint arXiv:1309.2057 (2013).

[6] G. Bhosale; A. Deshmukh; S. Medasani. "Image SuperResolution with Direct Mapping and De-Noising." Emerging Applications of Information Technology (EAIT), 2014 Fourth International Conference of. IEEE, 2014.

[7] H. Chen; B. Jiang; W. Chen. "Image super-resolution based on patch structure." Image and Signal Processing (CISP), 4th International Congress on. Vol. 2. IEEE, 2011.

[8] L. Wang; et al. "Edge-directed single-image super-resolution via adaptive gradient magnitude self interpolation." IEEE Transactions on Circuits and Systems for Video Technology 23.8 (2013): 1289-1299.

[9] M. Elad; and A. Feuer. "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images." IEEE transactions on image processing 6.12 (1997): 1646-1658.

[10] K. Zhang; et al. "Learning multiple linear mappings for efficient single image super resolution." IEEE Transactions on Image Processing 24.3 (2015): 846-861.

[11] S. Jeong, I. Seokhwa, I. Yoon, and J. Paik. "Multiframe example-based super-resolution using locally directional self similarity." IEEE Transactions on Consumer Electronics 61.3 (2015): 353-358.

[12] Jolicoeur-Martineau, A. “The relativistic discriminator: a key element missing from standard gan.” arXiv preprint arXiv:1807.00734 (2018).

[13] S. Farsiu; D. Robinson; M. Elad; P. Milanfar. “Advances and Challenges in Super-Resolution.” Wiley Publications, Inc. 15 March 2004

[14] Li, Xiaofang & Wu; Yirui & Zhang; Wen & Wang; Ruichao & Hou; Feng. (2020). “Deep learning methods in real time image super-resolution: a survey. Journal of Real-Time Image Processing.” 17.10.1007/s11554-019-00925-3.

[15] Dong, C.; Loy, C.C.; He, K.; Tang, X. “Learning a deep convolutional network for image super-resolution.” In European Conference on Computer Vision; Springer: Cham, Switzerland, 2014; pp. 184–199.

[16] Ooi, Y.K.; Ibrahim, H. “Deep Learning Algorithms for Single Image Super-Resolution: A Systematic Review.” Electronics 2021, 10, 867.

[17] XintaoWang; Ke Yu; Shixiang Wu; J. Gu; Y. Liu; C. Dong, C. Change Loy, Y. Qiao; X. Tang, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks”, in Lecture Notes in Computer Science Book (LNCS, volume 11133), January 2019.

[18] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. ECCV Workshops. September 1st 2018.

[19] I. Goodfellow; J. Pouget-Abadie; M. Mirza; B. Xu, D. WardeFarley; S. Ozair, A. Courville; Y. Bengio, “Generative adversarial nets,” in Neural Information Processing Systems, pp. 2672– 2680, 2014.

[20] L. Ratliff; S. Burden; S. Sastry, “Characterization and computation of local nash equilibria in continuous games,” in Annual Allerton Conference on Communication, Control, and Computing, pp. 917–924, 2013.

[21] K. O'Shea; R. Nash “An Introduction to Convolutional Neural Networks,” arXiv preprint arXiv:1511.08458 (2015).

[22] C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 105-114, doi:10.1109/CVPR.2017.19.