Survey on CNN based Super Resolution Methods

Rafaa Amen Kazem[1], Jamila H. Suad[2], Huda Abdulaali Abdulbaqi[3]

[1]Mustansiriyah University, College of Science, Computer Science Department, Baghdad, Iraq

*Corresponding Author: Rafaa Amen Kazem
Email: rafah.amen.kazem@gmail.com

Article Info	Abstract
	Super Resolution is a field of image analysis that focuses on boosting the resolution of photographs and movies without compromising detail or visual appeal, instead enhancing both. Multiple (many input images and one output image) or single (one input and one output) stages are used to convert low-resolution photos to high-resolution photos. The study examines super-resolution methods based on a convolutional neural network (CNN) for super-resolution mapping at the sub-pixel level, as well as its primary characteristics and limitations for noisy or medical images.

Introduction

Over the last few decades, viewing equipment such as plasma display panels (PDPs), liquid crystal displays (LCDs), and light emitting diodes (LEDs) (light emitting screens) has advanced tremendously, displaying images with high spatial and temporal parameters that are crystalline evident and visually good. However, despite the high request for HD content, it is not always available due to a variety of factors such as bandwidth constraints, different noise sources, and different compression methodologies, to name a few. Another important factor is the general availability of high-quality imaging equipment, and healthcare data is no longer modest (Al-falluji et al., 2017).

The data has grown enormously as a result of remarkable developments in picture collection technologies, making image analysis both tough and interesting. This quick advancement in medical pictures and processes necessitates lengthy and tedious work on the part of a particular medical expert, is vulnerable to human error, and may differ significantly between specialists. The use of machine learning techniques to automate the diagnostic process is an alternate solution; nevertheless, to cope with a complex problem, standard machine learning algorithms are insufficient. The successful union of computing at a high level and machine learning promises the potential to process enormous amounts of medical picture data in order to provide accurate and timely diagnoses.

Deep learning will not only assist in the identification and extraction of features, but also in the creation of new ones; additionally, it will not only detecting malady, but will also measure the predictive target and provide actionable prediction models to assist the clinician efficiently (Razzak et al., 2018). The term "ultra HD" refers to a set of algorithms aimed at increasing the resolution of photographs and movies. Upscaling and zooming are the same thing,

27

however Ultra Resolution Methods attempt to accomplish it without sacrificing photo or video detail or visual look. The main goal of ultra-high resolution is to find the value of a missing pixel in a high-resolution image. The two categories of super-resolution methods are multiple image super-resolution methods (Borman & Stevenson, 1998; Park et al., 2003; Mancas-Thillou & Mirmehdi, 2007; Tian & Ma, 2011) and single-image super-resolution methods.

The presumption behind multiple image hyper-resolution methods is that the multiple images used to create a high-resolution image are geometric transformations and non-aligned versions of each other, and that properly combining them will result in a more detailed image that can be included in an image class (Al-falluji et al., 2017). That is, ultra-high resolution techniques are applied to transform several low-resolution (LR) images to a high-resolution (HR) image. High-resolution photographs may provide more detailed information to everyone, making them more useful in a variety of applications such as satellite imagery, medical images, and so on. Several approaches in the field of sophisticated color digital image processing have emerged as a result of the increased technological interest in picture reconstruction (Shukla et al., 2020).

In terms of current work and future directions, this study presents a review of convolutional neural network-based methods with improved resolution in image analysis challenges. It covers the fundamentals of deep learning as well as the most up-to-date approaches in the field of image processing and analysis.

CNN based super resolution methods

Researchers employed the most up-to-date technology to build a high-resolution image using CNN channels. CNN optimizes all steps of the super-resolution process and tests the non-linear mapping function in the picture space between low-resolution and high-resolution. Very Deep super-resolution (VSDR), deep recursive convolutional network (DRSN), super-resolution convolutional neural network (SRCNN), and fast super resolution convolutional neural network are the most recent methods proposed using CNN (FSRCNN) (Al-falluji et al., 2017).

Table 1. Demonstrates super resolution methods based on CNN

Authors	dataset	Method	Gap	Results
Jiwon Kim, 2016 [9]	T91 Set5, Set14, B100 and Urban100.	very deep convolutional network (VDSR) + Residual Learning .	Due to the sluggish convergence rate, training an extremely deep network is difficult. The bicubic-upscaled LR photos must be processed.	PSNR=37.53 dB , SSIM= 0.9587
Yulun Zhang 2018 [10]	DIV2 Set5, Set14, B100, Urban100 and Manga109	very deep residual channel attention networks (RCAN)	Due to the sluggish convergence rate, training a deep network is difficult. LR bicubic-upscaled photos must be processed.	PSNR= 34.83 dB, SSIM= 0.9296
Jiwon Kim, 2016 [11]	T91 Set5, Set14, B100 and Urban100.	deeply-recursive convolutional network (DRCN)	Due to exploding/vanishing gradients and increased ejection between layers and parameters, using a traditional gradient descent method is extremely problematic, resulting in overfitting. You'll also need to handle the LR photos that have been bicubic-upscaled.	PSNR= 37.63 dB, SSIM= 0.9588

Yang and Lu [12] 2020	Set5, Set14, BSD100, Urban100 and Manga109	a low- and high-frequency fusing network that is profoundly recursive (DRFFN)	A larger network necessitates more processing and memory resources, making network training more difficult and extending prediction time.	PSNR=38.16 dB SSIM=0.9649
Chao Dong, 2015 [14]	T91 Set5 Set14 BSD200	Super-Resolution Convolutional Neural Network (SRCNN)	the high computational cost is still preventing them from practical use, which requires the performance of real-time (24 fps).quality	PSNR = 36.66 dB, SSIM = 0.9542.
Chi-Hieu Pham, 2019 [20]	Kirby 21 NAMIC image of Brain web	(3DSRCNN)	The magnitude of the training correction and the number of subjects, as well as the context of the MRI of the brain, the acquisition of big data sets and their combination with uniform acquisition settings, are all factors to consider. The disadvantage of the 3D SRCNN is that it takes longer to forecast the SR scan.	PSNR= 39.01 dB
Ben Garber 2020 [15]	DIV2K	Super-Resolution Convolutional Neural Network (SRCNN)	it fails to produce clarity desired in the super-resolution task. The lack of availability of GPUs also limited us to a relatively shallow neural network	PSNR = 26.442 dB
Chao Dong, 2016 [17]	T91 and General-100 Set5 Set14 BSD200	Fast Super-Resolution Convolutional Neural Networks (FSRCNN).	original LR image require bicubic interpolation . computational complexity	PSNR=37.00 dB, SSIM=0.9558.
VanitaMane 1,2020 [18]	T1-weighted structural MRI images (brain MRI)	(3D FSRCNN)	The magnitude of the training correction and the number of subjects, as well as the context of the MRI of the brain, the acquisition of big data sets and their combination with uniform acquisition settings, are all factors to consider. The disadvantage of the 3D SRCNN is that it takes longer to forecast the SR scan.	PSNR= 48.40 dB, SSIM= 0.975 .

VDSR

To build a high-resolution image, A extremely deep convolutional network was utilized, which was inspired by Simonyan & Zisserman (2014) work from 2015. The "D" number of layers was used by CNN. The first layer takes an LR image as input and reconstructs the high-resolution image in the final layer. Our main contribution is a rigorous analysis of rising-depth networks using an architecture with extremely small (3 3) convolution filters, which shows that raising the depth to 16–19 weight layers improves performance significantly over prior-art setups. The input network image is an interpolated version of a low-resolution image of the same size as the output image.

It has been discovered that employing residual images provides a number of advantages, including faster convergence and better performance, that is, it gives better PSNR. Because super-resolution requires surrounding pixels to correctly forecast the pixel, one of the key challenges that deep networks face during the prediction phase is that feature maps are reduced with each iteration of convolution.

In 2016, Kim et al, (2016) presented deeper structures, which was very deep convolutional network (VDSR) uses residual-learning and extremely as well gradient clipping to ensure the training stability with require processing to optimize accurate conclusion, skip connection and quick. And the results are PSNR=37.53, SSIM= 0.9587 and the results bicubic are PSNR= 33.66, SSIM= 0.9299. Yet, because of the slow convergence rate and the requirement for processing bicubic-upscaled LR images, training a very deep network is difficult.

In 2018, Zhang et al, (2018) the depth of a convolutional neural network is critical for visual super-resolution (SR). Deeper networks for image SR, on the other hand, it is additional hard to train. Low-resolution inputs and features contain a lot of low-frequency data, which is processed the same way across channels, limiting CNNs' representational ability. To address these issues, propose the use of very deep residual channel attention networks (RCAN) and the results PSNR= 34.83, SSIM= 0.9296.

DRCN

Kim et al (2016) He presented a net embedding, inference network, and reconstruction network are the three essential aspects of the proposed algorithm. The algorithm was a deep-recursive convolution network (DRCN), which contains a very deep recursive layer (up to 16 iterations). Despite the fact that the processing of the developed bicubic LR images may be complicated with the standard gradient approach due to burst/fade gradients and increased expulsion between parameters and layers, the results were PSNR = 37.63, SSIM = 0.9588 and PSNR = 33.66, SSIM = 0.9299. Yang & Lu, (2020) proposed in 2020 to improve model representations by increasing the grid's depth and width, which would result in greater image rebuilding quality.

However, a larger network requires more computer resources and memory, making training the network more difficult and increasing forecast time. Due to the aforementioned issues, a novel low- and high-frequency fuser network (DRFFN) is presented, which uses a parallel branching structure to extract low and high-frequency information from the image, respectively. And the results were PSNR=38.16, SSIM=0.9649.

SRCNN

The goal of SRCNN is to learn the bicubic LR picture to HR image end-to-end mapping function. Because the network includes all convolution layers, the output image is the same size as the input image (Yang et al., 2010). As shown in Figure 1.
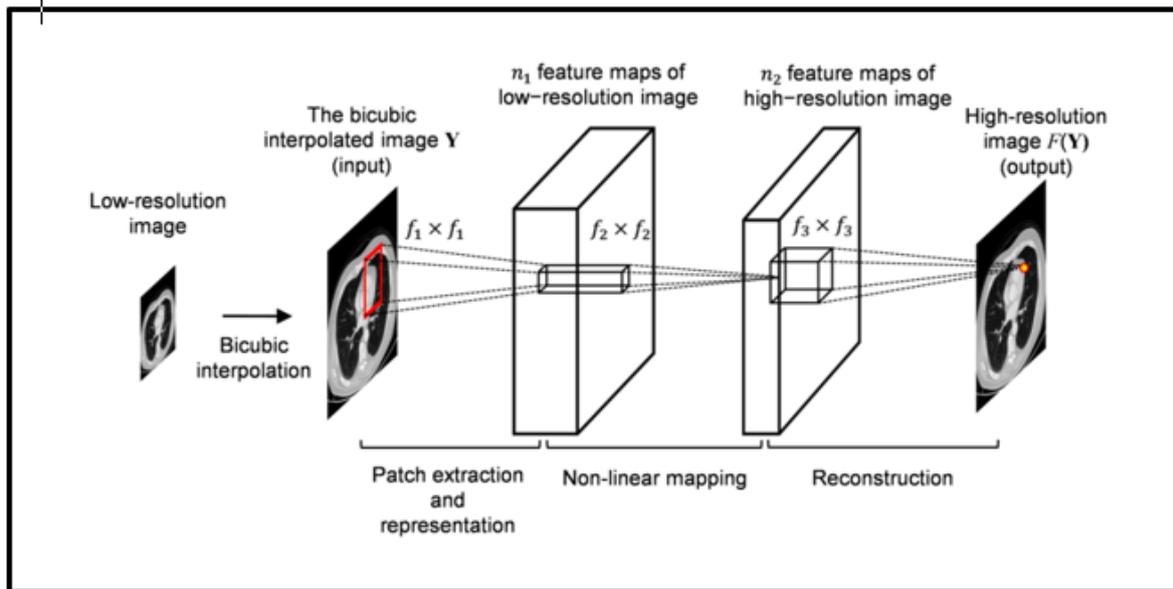
Figure 1. illustrate SRCNN architecture (Umehara et al., 2018)

In 2015, showed the research by Dong et al, (2015) which suggested that SRCNN vaulting superior performance of previous models of hand-made (advanced techniques), both regarding quality of recovery and speed, also the results were PSNR = 36.66, SSIM = 0.9542. Yet, the high-computational costs prevent them from practical use, which requires the performance of real-time (24fps). In 2019, Researcher Pham et al, (2017) developed the SRCNN3D to a focus on increasing the data for SR rotation and flipping to improve generalization of the algorithm and then used it to estimate a network parameter training data set consisting of pairs of HR and LR images by reducing the objective function to better performance and faster convergence with large data sets clustered. And the results were PSNR= 39.01. Critical issues are the size of the training correction and the subject number as well as the case of the MRI of the brain, acquisition of large data sets and their combination with homogeneous acquisition settings.

Li, (2017) implement SRCNN with used 2D non-clinical images then measured by peak signal-to-noise ratio (PSNR) although there advancer methods, but using this simpler model to improve its and implement a preformat super-resolution to model led to gain practical insight into how to better apply deep learning to the task of super-resolution where reached to a (PSNR) of 23.226 dB and 26.442 dB for bicubic up sampling baseline test and SRCNN respectively. It results in increased image resolution but it fails to produce clarity desired in the super-resolution task. The lack of availability of GPUs also limited us to a relatively shallow neural network. This application is more significant in MRI scans and satellite images in medical imaging. The PSNR values of SRCNN image have been increased, as higher PSNR values lead to better image quality. Lastly, the output is going to be the same as the original image, which will be better in terms of texture quality and high resolution. Because using bicubic-interpolation, which has a smoother surface and fewer interpolation tools, is a limitation, by select bicubic-interpolation as the input and send it into 3 warp layers for more processing.

FSRCNN

Feature extraction, shrinkage, mapping, expansion, and deconvolution are the five parts of FSRCNN. Wrap layers make up the first four sections, while a jaw layer makes up the final section. The convolution layer is referred to as Conv (fi, ni, ci) and the deconvolution layer is

referred to as DeConv (fi, ni, and ci), with the variables fi, ni, ci representing filter size, number of filters, and number of channels, respectively. Dong et al, (2016) proposed speeding up SRCNN in 2015 by the same authors with 3 aspects. Initially, a de-convolutional layer has been added to the network's end. Second, we shrink the mapping layer to reformulate it. The proposal used small filter sizes, yet more layers of mapping in the third step. With Superior restore quality, the model achieved a speed of over 40 times.

The results for fast super resolution convolution neural network FSRCNN are PSNR=37.00, SSIM=0.9558 and the results bicubic are PSNR=33.66, SSIM= 0.9299. Mane et al. (2020), in the 2020, propose an Neural Networks (NN) referred to as 3-D Faster Super-Resolution Convolutional Neural Networks (3DFSRCNN) based (FSRCNN), with the perfect trade-off between performance and complexity, which is significant for presenting a memory consumption challenges, massive expenses, making them less realistic. The increases in time complexity while network becomes deeper is one of the issues I've encountered. They trained the model 3D FSRCNN batch size of 32 due to the limited RAM. PSNR values appeared to be stable from the 500-th epoch after training for 750 For epoch. The result was PSNR= 48.40, SSIM= 0.975. From all the previous researches can be summarized their methods, datasets, and the results in Table (1.1) Determine the problem that display in this study depending on the CNN types that present in previous sections

Conclusion

A study of recent ultra-precision algorithms based on learning instances is included in the publication. Because there are more examples to match, traditional learning methods are more accurate than self-correcting approaches. The approaches described are based on convolutional neural networks, which are more accurate, and it was discovered that FSRCNN is the fastest and produces the best results.

Acknowledgments

References

Al-falluji, R. A. A., Youssif, A. A. H., & Guirguis, S. K. (2017). Single image super resolution algorithms: A survey and evaluation. *Int. J. Adv. Res. Comput. Eng. Technol*, *6*, 1445-1451.

Borman, S., & Stevenson, R. L. (1998, August). Super-resolution from image sequences-a review. In *1998 Midwest symposium on circuits and systems (Cat. No. 98CB36268)* (pp. 374-378). IEEE.

Dong, C., Loy, C. C., & Tang, X. (2016, October). Accelerating the super-resolution convolutional neural network. In *European conference on computer vision* (pp. 391-407). Springer, Cham.

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

Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1646-1654).

Kim, J., Lee, J. K., & Lee, K. M. (2016). Deeply-recursive convolutional network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1637-1645).

Li, Y., Hu, J., Zhao, X., Xie, W., & Li, J. (2017). Hyperspectral image super-resolution using deep convolutional neural network. *Neurocomputing*, *266*, 29-41.

Mancas-Thillou, C., & Mirmehdi, M. (2007). An introduction to super-resolution text. In *Digital document processing* (pp. 305-327). Springer, London.

Mane, V., Jadhav, S., & Lal, P. (2020). Image Super-Resolution for MRI Images using 3D Faster Super-Resolution Convolutional Neural Network architecture. In *ITM Web of Conferences* (Vol. 32). EDP Sciences.

Park, S. C., Park, M. K., & Kang, M. G. (2003). Super-resolution image reconstruction: a technical overview. *IEEE signal processing magazine*, *20*(3), 21-36.

Pham, C. H., Ducournau, A., Fablet, R., & Rousseau, F. (2017, April). Brain MRI super-resolution using deep 3D convolutional networks. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (pp. 197-200). IEEE.

Razzak, M. I., Naz, S., & Zaib, A. (2018). Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps*, 323-350.

Shukla, A., Merugu, S., & Jain, K. (2020). A Technical Review on Image Super-Resolution Techniques. *Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies*, 543-565.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Tian, J., & Ma, K. K. (2011). A survey on super-resolution imaging. *Signal, Image and Video Processing*, *5*(3), 329-342.

Umehara, K., Ota, J., & Ishida, T. (2018). Application of super-resolution convolutional neural network for enhancing image resolution in chest CT. *Journal of digital imaging*, *31*(4), 441-450.

Yang, C., & Lu, G. (2020). Deeply recursive low-and high-frequency fusing networks for single image super-resolution. *Sensors*, *20*(24), 7268.

Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE transactions on image processing*, *19*(11), 2861-2873.

Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., & Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 286-301).