Convolutional Neural Network for Smoke Image Super-Resolution

Maoshen Liu*
Beijing University of Technology
Beijing 100124
China
lmsbjut@163.com

Ke Gu
Beijing University of Technology
Beijing 100124
China
guke.doctor@gmail.com

Junfei Qiao
Beijing University of Technology
Beijing 100124
China
junfeiq@bjut.edu.cn

ABSTRACT

The[1] flare stack is the last line of defense in the safe production of large-scale chemical plants. Monitoring black smoke produced by the incomplete flare stack exhaust combustion can effectively reduce environmental pollution and production accident. In order to improve the ability to recognition and analyze the black smoke, high-resolution flare stack scene images are in urgent need. To this end, we in this paper propose a super-resolution algorithm based on convolutional network that focuses only on smoke area for the purpose of identifying the smoke of flare stack. With a lightweight convolutional neural network structure, our network specializes in learning smoke characteristics mapping between the low-resolution images and the associated high-resolution. To verify validity, our algorithm compares the super-resolution quality of the smoky region of the flare stack image with several classic super-resolution algorithms. The experimental results show that our algorithm is superior to the classical algorithms when applied to smoke images.

CCS CONCEPTS

• **Computing methodologies → Artificial intelligence**; *Computer vision*; Computer vision problems; Reconstruction

KEYWORDS

Flare stack, super-resolution, convolutional neural network, environmental pollution

1 INTRODUCTION

With the rapid development of industrial automation, high-resolution industrial surveillance images are becoming more and more urgent, such as flare stack monitoring. Flare stack is a gas combustion device used in industrial plants such as petroleum refineries, chemical plants, and natural gas processing plants. Insufficient flare stack combustion will produce a large number of toxic and harmful gases, causing accidents and pollution. The surface phenomenon of insufficient flare stack burning is the generation of black smoke. Therefore, the effective identification and analysis of the flare stack black smoke can avoid the above hazards. Due to the restriction of the factory environment and facilities, the low resolution of the collected images brings great difficulties to the identification and analysis of the black smoke. Unclear images may ignore a part of the smoke and cause a certain safety hazard. For this reason, an efficient super-resolution algorithm is a breakthrough in solving this problem.

Single image super resolution reconstruction (SISR) is an important technique for building bridges between input low-resolution (LR) images and output high-resolution (HR) images. In recent years, more attention has been paid to it. The challenges of SISR mainly include: identifying important visual clues, filling details, and as far as possible faithfully and beautifully presented [1]. At present, SISR can generally be divided into two categories: Reconstruction-based and Example Learning-based. The reconstruction-based method aims to reconstruct the high-frequency signals lost during the degrading process. Irani et al. [2] formulated the iterative back-projection (IBP) SR reconstruction approach that is similar to the back projection used in tomography. In this approach, the HR image is estimated by back projecting the error (difference) between simulated LR images via imaging blur and the observed LR images. Dai et al. [3] designed the regularization term of the difference weighted sum of the center pixel and the neighboring pixels by using the prior knowledge of continuity on both sides of the edge where the contrast is large and sharp. However, this method always produces ambiguous results for the texture area, since the texture image features do not satisfy the priori assumptions. Singh et al. [4] used pixel values at the edge algorithm learning a priori from the image block instead of the edge in order to cover various image structures. There are a number of other traditional super-resolution algorithms that have been proposed, such as [5,6,7,8].

Traditional super-resolution algorithm recovers image quality based on extracting image features envisioned by researchers. However, the feature extraction designed by the researchers has some limitations. Recently, more and more attention has been paid to the super-resolution algorithm based on deep convolutional neural networks. Convolutional neural networks

CSAE '18, October 22–24, 2018, Hohhot, China

ACM ISBN 978-1-4503-6512-3/18/10…$15.00
https://doi.org/10.1145/3207677.3277950

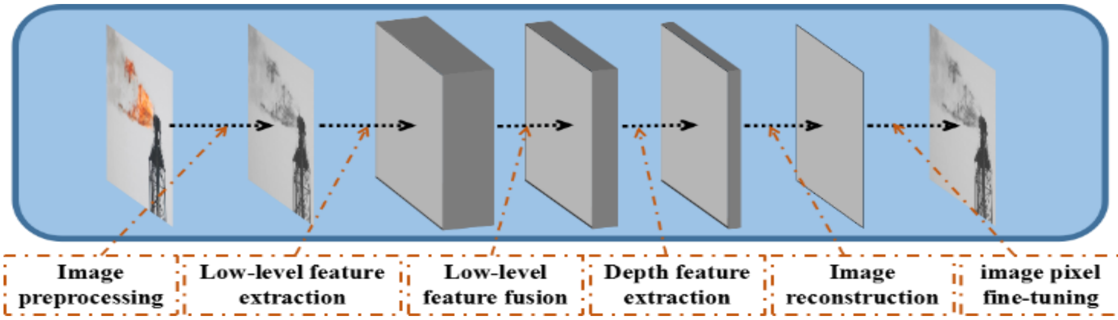

Figure 1: The algorithm flow of our proposed smoke image super-resolution method.

can learn useful features by themselves, and some of these features go beyond manual methods. Such as, Dong *et al.* first proposed a super-resolution algorithm based on deep learning, with a very simple network structure, using only three convolutional layers [9]. Kim *et al.* found that only the high-frequency partial residuals between high-resolution and low-resolution images need to be learned. Therefore, they proposed a residual-based super-resolution network and improved the number of network layers [10]. Lai *et al.* proposed a Laplacian Pyramid Super-Resolution Network to progressively reconstruct the sub-band residuals of high-resolution images [11]. Lim *et al.* proposed a super-resolution model based on optimized residual network, which removes the redundant modules [12]. Tai *et al.* designed a deeper network to further improve super-resolution performance by using recursive and weight sharing strategies [13]. Existing convolutional neural network-based super-resolution algorithms are constantly increasing the number of network layers to better restore complex and diverse details in natural images. However, the above algorithm is clearly redundant in structure for black smoke images. In order to solve the super-resolution algorithm for the flare stack black smoke problem, we use the advanced deep convolutional neural network method. We consider that the black smoke image contains only simple texture and edge information compared to the complex structural information in the natural image. Therefore, we designed a lightweight convolutional neural network, which performs super-resolution reconstruction after extracting simple features of black smoke. Experiments prove the superiority of our algorithm in smoky super resolution.

The remainder of this paper is organized as follows. In Section 2, we mainly introduce the principle of structural design and parameter selection of our convolutional neural network. In Section 3, we explain the experimental details, and compared with the existing traditional super-resolution algorithm to verify the advantages of our algorithm. And we summarize the experimental results. Some concluding remarks and future works are provided in Section 4.

2 ALGORITHM DETAILS

The existing classical super-resolution algorithm is mainly proposed for the characteristics of natural scene images, lacking to deal with special scene problems. For the problem of super-resolution of the torch black smoke image, we designed a convolutional neural network to specifically study the super-resolution mapping of the black smoke region. We name the proposed network as a smoky super-resolution convolutional neural network, abbreviated as 'SSRCN'. The algorithm flow is shown in the Fig. 1. And the details of the algorithm are as follows.

2.1 Image Preprocessing in Algorithm Flow

The purpose of this algorithm is mainly to improve the quality of the black smoke region in the super-resolution image of the flare stack image to better identify and analyze the black smoke. Therefore, we don't need to process color images for human feelings. We only consider the luminance channel (in YCrCb color space) in our proposed SSRCN. Next, we upscale the image to be processed by bicubic interpolation to get the desired size. The principle of our algorithm is that our SSRCN learns the details needed for super-resolution reconstruction to recover the details of the low-quality map obtained by the bicubic algorithm.

2.2 Design of Network Structure

Our SSRCN has five layers, and it's each layer is named based on the characteristics of the convolutional neural Network as illustrated in Fig. 1. In low-level feature extraction (LFE) layer, the features of the input image are extracted by a filter with a convolution kernel size of 5. In a convolutional neural network, the initial convolutional layer extracts simple features of the image, such as edges, corners, and textures. However, in the process of super-resolution reconstruction, these simple features are more important for image restoration. Therefore, the number of convolution kernels in the first layer is much larger compared to other layers to extract richer simple features. In low-level feature fusion (LFF) layer, the filter with a convolution kernel size of one does not change the geometry of the feature map extracted from the previous layer. It only fuses the feature map by pixels to extract a richer feature map, and reduces the number of transmitted feature map channels, thereby reducing the network parameter. In depth feature extraction (DFE) layer, filters with a convolution kernel size of three are used to nonlinearly transform the feature map of the previous layer and the characteristics are further refined to reduce the network parameters. In image reconstruction (IR) layer, the reconstruction of super-resolution images is performed using the feature maps previously presented. In image pixel fine-tuning (DFE) layer, the super-resolution image reconstructed from the previous layer is fine-tuned by pixels by a filter with a convolution kernel size of one. In our SSRCN, all activation functions are set to ReLU in order to ensure that the pixel value is positive, and the network uses padding operations to ensure that the network output size is the same as the input. For the convenience of the reader, we list the detailed parameters of the network in Table 1.

Table 1: Model Relevant Hyper-Parameters

Layer	Layer Type	Hyper-parameters
Input	Input	Image size: 96×96
LFE	convolution	Filters size: 5×5 Filter number: 128 Stride: 1, padding Activitionfunction: ReLU
LFF	convolution	Filters size: 1×1 Filter number: 64 Stride: 1, padding Activition function: ReLU
DFE	convolution	Filters size: 3×3 Filter number: 32 Stride: 1, padding Activition function: ReLU
IR	convolution	Filters size: 5×5 Filter number: 1 Stride: 1, padding Activition function: ReLU
DFE	convolution	Filters size: 1×1 Filter number: 1 Stride: 1, padding Activition function: ReLU

2.3 Network Training Configuration Details

The image data used by our SSRCN during training are patches of black smoke with a length and width of 96×96, as shown in Fig. 2. The image needs to be normalized in order to make the network easier to converge during training. The total number of pictures in the training process is 3200 pieces, and the pictures are equally distributed for weight update and performance verification. Training different magnification scales requires separate processing of the pictures. For example, to train a network that is magnified three times, we need to first reduce the luminance channel of the image to the original one-third by bicubic method and then enlarge it to the original size as the input of the network. The original image is used as the label of

Figure 2: Dataset for super-resolution model training.

the network. During training, the batch is set to 16 and the epoch is set to 150. We use Mean Squared Error (MSE) as the loss function and update the weights through Adam algorithm, in which the learning rate is set to 0.0003.

3 EXPERIMENTS AND DISCUSSION

3.1 Datasets and Comparison Algorithm

For a fair comparison with the state-of-the-art SR methods, we performed the same preprocessing on the used flare stack scene image. We transform the image from the RGB color space to the YCrCb color space and compare only the luminance channel in our experiments. In order to demonstrate the effect, the displayed picture is a color picture obtained by combining the other two channels processed by bicubic method.

We compare our SSRCN with the traditional SR methods: the Yang et al. [13], the Zeyde et al. [14], the ANR (Anchored Neighbourhood Regression) method [15], and the NE+LLE (Neighbor Embedding with Locally Linear Embedding) method [16]. We also compare it with the SRCNN algorithm based on the convolutional neural network. The SRCNN algorithm used for comparison is a model published by the website source code through natural image training. The implementation of other algorithms come from publicly available code provided online. All experiments were carried out in the software environment with Matlab2015a under Windows 10 operation system and run on a PC with two Inter(R) Xeon(R) E5-2683 v3 2.00GHz CPUs and two Nvidia GeForce GTX 1080 GPUs.

3.2 Comparing Results

In order to make the experiment more full and powerful. In this experiment, our network was trained to two magnifications, three times and four times. We use PSNR (Peak Signal to Noise Ratio) as a measure, which is a widely used metric for quantitatively evaluating image restoration quality, and is at least partially related to the perceptual quality. Firstly, one hundred randomly selected black smoke image patches similar to the training data set are used to validate the effectiveness of our proposed SSRCN. The results are listed in Table 2. It can be seen from Table 2 that our SSRCN restores the details of the black smoke super-resolution image obtained by the bicubic algorithm.

Table 2: The Average Psnr and Rmse Results For One Hundred Smoky Image Blocks

scale	**Bicubic**		**SRCNN**		**SSRCN (pro.)**	
	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR
3	1.87	43.59	1.62	44.71	1.56	**45.06**
4	2.55	40.86	2.29	41.79	2.16	**42.26**

Figure 3: Flare stack scene image with an upscaling factor 3. Black smoke block diagrams in the red frame area of the original image, and marked the PSNR of the corresponding block.

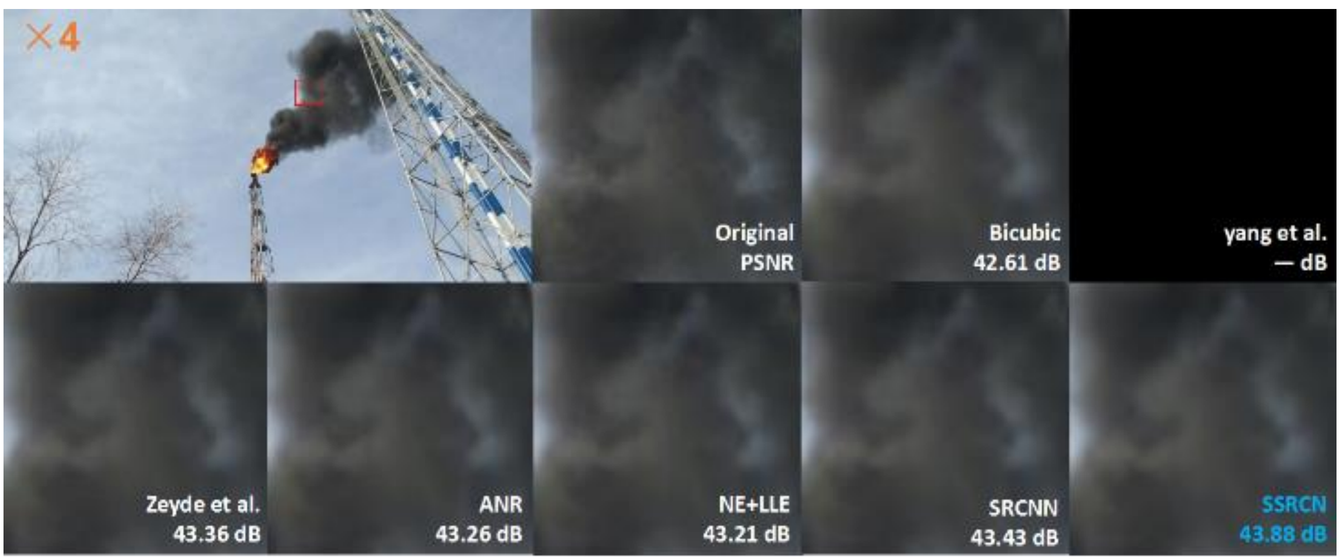

Figure 4: Flare stack scene image with an upscaling factor 4. Black smoke block diagrams in the red frame area of the original image, and marked the PSNR of the corresponding block. (The Yang et al. source program cannot be used for 4x upsampling.)

We photographed the actual image at the flare scene. By performing super-resolution processing on the field map, we compare the proposed SSRCN with the comparison algorithms in practical applications. Fig. 3 and Fig. 4 illustrate the 3x and 4x super-resolution effects of each algorithm on the image. Due to the lack of sharp edge details in black smoke images, the algorithm effect in the diagram cannot be clearly observed. However, careful observation can reveal that the greater the value of PSNR, the clearer visual experience of the black smoke image. Compared with other methods, our algorithm has obvious PSNR advantage in the black smoke area. At the upscaling factor of 3, the highest PSNR in the comparison algorithm, 44.65dB, is obtained by the SRCNN algorithms. However, the PSNR obtained by our proposed SSRCN is 45.05dB. At the upscaling factor of 4, the highest PSNR in the comparison algorithm, 43.43dB, is obtained by the SRCNN algorithms. However, the PSNR obtained by our proposed SSRCN is 43.88dB. Although it seems to the human visual system that these images do not give much visual difference, the small difference in the PSNR may have influence on the accuracy of the identification of black smoke in the flare stack images. Therefore, this algorithm does not focus on people's feelings but it is for the actual needs of the industry.

We further analyze the super resolution results of the entire flare stack images. We analyze the local PSNR of the black smoke region and the global PSNR of the whole flare stack image, and the experimental data is listed in Table 3. At the upscaling factor of 3, our proposed SSRCN algorithm is optimal in the local PSNR of the black smoke region and not outstanding in the entire map. Similarly, when the upsampling factor is 4, our proposed SSRCN algorithm is optimal in the local PSNR of the black smoke region, and there is no obvious advantage in the entire map. This shows that our algorithm is more specific to the super resolution of black smoke area in images. And it has certain robustness to the effect of the whole image super-resolution.

Table 3: Local Psnr and Global Psnr Results of All Methods

methods	scale	Local PSNR	Global PSNR
Bicubic	3	44.12 dB	25.11 dB
Yang et al.	3	40.92 dB	25.59 dB
Zeyde et al.	3	44.56 dB	25.51 dB
ANR	3	44.34 dB	25.59 dB
NE+LLE	3	44.41 dB	25.56 dB
SRCNN	3	44.65 dB	25.76 dB
SSRCN (pro.)	3	45.05 dB	25.70 dB
Bicubic	4	42.61 dB	24.37 dB
Zeyde et al.	4	43.36 dB	24.69 dB
ANR	4	43.26 dB	24.75 dB
NE+LLE	4	43.21 dB	24.73 dB
SRCNN	4	43.43 dB	24.82 dB
SSRCN (pro.)	4	43.88 dB	24.84 dB

3.3 Discussions

Through the above experiments, we can summarize the following three points for our study. Firstly, through deep convolutional neural network can learn the corresponding relationship of the LR/HR samples of flare stack black smoke images. Compared with the super resolution of the natural scene images, the characteristics of the black smoke images need to be learned is relatively simple. Therefore, we can achieve satisfactory results with only a small amount of learning samples. Secondly, because our training sample is only black smoke images, so in the experiment, the effect on the black smoke region is the best and the effect on other non-smoke regions is relatively poor. The SRCNN model trained by natural images is superior to our algorithm in the PSNR of the whole map, but worse than our algorithm in the black smoke region. Therefore, we can select specific training samples to learn a special-application super-resolution algorithm based on convolutional neural network. Thirdly, our network only has a better effect on the black smoke area, but it does not work well on the non-smoke area. On the contrary, it is more favorable for the identification and analysis of black smoke in the torch black smoke image.

4 CONCLUSIONS

The flare stack is very important for the safe production of the factory. Accurate monitoring of flare combustion is an important challenge for the current flare stack. Super-resolution processing of flare images can improve the identification of black smoke. However, the current general super-resolution algorithm does not perform well in super-resolution processing of black smoke images. For this reason, we have proposed a super-resolution algorithm based on convolutional neural network dedicated to black smoke images. By making the network only learn the restoration of black smoke image features, a super-resolution algorithm dedicated to black smoke images is obtained. Comparing with the classic super-resolution algorithm, our algorithm works best in the black smoke region of the image. Therefore, our algorithm can greatly improve the accuracy of flare stack image black smoke recognition. For the powerful image feature extraction capability of convolutional neural networks, applying a convolutional network to a specific industrial scene image is a new breakthrough in solving industrial problems. In practical applications, the reconstructed image of the SR algorithm does not have a high-definition reference image, so future work considers using more state-of-the-art blind image quality assessment methods [17-20] to evaluate the performance of the SR algorithm.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation of China (Grants 61703009), and the Young Elite Scientist Sponsorship Program by China Association for Science and Technology (Grant 2017QNRC001).

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang. 2003. Super-resolution image reconstruction: a technical overview. *IEEE SPM*, 20(3), 21–36, May 2003.
[2] M. Irani, and S. Peleg. "Improving resolution by image registration." *CVGIP*, 53(3), 231-239, May 1991.
[3] S. Dai, M. Han, W. Xu, Y. Wu, and Y. Gong. "Soft edge smoothness prior for alpha channel super resolution." *CVPR*, 1-8, 2007.
[4] A. Singh, and N. Ahuja. "Single image super-resolution using adaptive domain transformation." *ICIP*, 947 - 951, 2013.
[5] W. T. Freeman, T. R. Jones, and E. C. Pasztor. "Example-based super-resolution." *IEEE CGA*, 22(2), 56 - 65, Apr. 2002.
[6] Q. Wang, X. Tang, and H. Shum. Patch based blind image super resolution. *ICCV. 1, 709-716, 2005.*
[7] D. Glasner, S. Bagon, and M. Irani. "Super-resolution from a single image," *ICCV*, 349 – 356, 2009.
[8] J. Yang, J. Wright, T. S. Huang, and Y. Ma. 2010. Image super-resolution via sparse representation. *IEEE TIP*, 19(11), 2861-2873, Nov. 2010.
[9] C. Dong, C. C. Loy, K. He, and X. Tang. 2016. Image super-resolution using deep convolutional networks. *IEEE TPAMI*, 38(2), 295-307, Feb. 2016.
[10] J. Kim, J. K. Lee, and K. M. Lee. 2016. Accurate image super-resolution using very deep convolutional networks. *CVPR*, 1646-1654, Dec. 2016.
[11] W. S. Lai, J. B. Huang, N. Ahuja, and M. H. Yang. 2017. Deep laplacian pyramid networks for fast and accurate super-resolution. *CVPR*, 5835-5843, Nov. 2017.
[12] B. Lim, S. Son, H. Kim, S. Nah, and KM Lee. 2017. Enhanced deep residual networks for single image super-resolution. *CVPRW*, 1132-1140, Aug. 2017.
[13] Y. Tai, J. Yang, and X. Liu. 2017. Image super-resolution via deep recursive residual network. *CVPR*, 2790-2798, Nov. 2017.
[14] R. Zeyde, M. Elad, and M. Protter. 2010. On single image scale-up using sparse-representations. *International conference on curves and surfaces*, Springer, Berlin, Heidelberg, 2010.
[15] R. Timofte, V. De, and L. Van Gool. 2013. Anchored neighborhood regression for fast example-based super-resolution. In *Proc. IEEE Int. Conf. Comput. Vis.*,

1920-1927, 2013.
[16] H. Chang, D. Y. Yeung, and Y. Xiong. 2004. Super-resolution through neighbor embedding. Presented at *the IEEE Conf. Comput. Vis. Pattern Recog., Washington, DC, USA*, 2004.
[17] K. Gu, G. Zhai, X. Yang, and W. Zhang. 2015. Using free energy principle for blind image quality assessment. *IEEE TMM*, 17(1), 50–63, Jan. 2015.
[18] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang. 2015. No-reference image sharpness assessment in autoregressive parameter space. *IEEE TIP*, 24(10), 3218-3231, Oct. 2015.
[19] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen. 2017. Noreference quality metric of contrast-distorted images based on information maximization. *IEEE Trans. Cybern.*, 47(12), 4559-4565, Dec. 2017.
[20] K. Gu, D. Tao, J.-F. Qiao, and W. Lin. 2018. Learning a no-reference quality assessment model of enhanced images with big data. *IEEE TNNLS.*, 29(4), 1301-1313, Apr. 2018.