Deep Learning for Subsurface Penetrating Super-Resolution Imaging

Yan Zhang[1], Zelong Xiao[*1], Li Wu[1], Xuan Lu[2],Yuankai Wang[1]
1 School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
2 Department of Electronic Information Engineering, Shanxi University, Taiyuan 030006, China
E-mail: zelongxiao@mail.njust.edu.cn

***Abstract*—Image resolution enhancement for shallow buried small targets is a meaningful step in holographic subsurface penetrating (HSR) imaging process, due to the fact that image results are easily affected by the complex underground environment and the follow-up high-level vision task is hindered. In this paper, we employ super-resolution convolutional neural network (SRCNN) in HSR image resolution enhancement process. It can get nonlinear mapping from the low-resolution image patches to the high-resolution space and perform better than traditional hand-crafted methods. According to the features of degraded HSR images, we add a layer in SRCNN structure to ensure a noise-robust state as well as preserve more image details. In addition, we use the exponential linear unit (ELU) activation function to reduce computational complexity in training process. By training HSR image datasets and carrying PNSR criteria to evaluate the experiment performance, a considerable gain is achieved compared to classical restoration techniques.**

Keywords—Deep Learning; subsurface penetrating imaging; super-resolution; Convolutional Neural Network

I. INTRODUCTION

Holographic subsurface penetrating radar (HSR) plays an important role in various fields such as medicine, military and security inspection. Because of the high frequency of the transmitting signal, HSR can achieve high positioning accuracy, making it feasible in the detection of shallow buried objects. On a processed plan-view image, objects can often be identified directly according to their shape and texture [1]. However, compared with traditional optical imaging results, HSR images are prone to have more serious degradation problems. Hence the image contour obtained by subsurface penetrating imaging is not clear enough, which will leads to the fuzzy problem. In addition, inhomogeneous space distribution of underground medium always results in ghosts covering the targets [2]. As a result, the follow-up high-level vision task, such as image classification and object detection, will be hindered especially when the detected targets are small. Therefore, it is necessary to enhance image resolution to highlight the characteristics of imaging targets and reduce the degradation influence for the follow-up detection and recognition process.

In this paper, we focus on the image resolution issue for HSR results. It is a thorny problem of reconstructing a high-resolution (HR) observation from a low-resolution (LR) one. Apart from traditional histogram modification and interpolation technique [3], the most advanced approaches for subsurface image resolution enhancement include the statistical image priors [4] and the sparse-coding-based method [5]. Both of them learn the dictionaries and model the patch space to get LR and HR exemplar pairs. It is admitted that these methods are limited to the ill-posed problem, and overly rely on primitive image properties such as edges and segments [6]. Moreover, they cannot find a uniform solution to preserve targets features. Therefore they cannot keep a balance between the suppression for degradation and the preservation of the image details. Recent state-of-the-art methods for single image super-resolution are mostly based on the deep learning theory. In these methods, convolutional neural network (CNN) is regarded as an end-to -end mapping between LR and HR images [7]. CNN not only puts the patch extraction and aggregation steps into a whole structure, but also performs better than traditional approaches. This paper concerns with the problem of subsurface small targets detection by means of HSR imaging technique. The image enhancement model is based on the CNN theory and improved according to the characteristics of HSR images. Both the image super-resolution structure and the projection for HSR images are discussed in detail. In addition, the corresponding experimental results also demonstrate the feasibility and effectiveness of the method.

II. THEORY: RESOLUTION ENHANCEMENT FOR HSR IMAGING

The main process of microwave holographic imaging is recording the interference pattern on the receiving plane between the reference wave and the wave reflected from targets, which relies on the measurement of the magnitude and phase of the wave scattered from the imaged target. Referring to holographic optical imaging model, the HSR imaging degradation model can be written as equation(1)

$$y(i,j) = D(x(p,q) * h) + n(i,j) \qquad (1)$$

The formula describes the relationship between LR image y and HR image x, where h represents the PSF of imaging system. $D(\bullet)$ indicates down-sampling process, which means the additive noise during the imaging process.

In subsurface penetrating imaging process, we always cannot know specific information of the measured target and aim to provide images without HR observation, hence it is a challenging issue to enhance results resolution. Recently, convolutional neural networks were regarded as a form of

978-1-5386-2720-4/17/$31.00 ©2017 IEEE

generalization of sparse coding, which has ability of non-linearity and automatic network flexibility dictionary learning, improving upon past state-of-the-art performance[8].After learning a set of representations from original data, CNNS can automatically extract prominent high-level features through inter-connected layers [9]. In resolution enhancement area, it can learn a mapping from LR to HR image pairs and seems the HR image as output by predicting from the LR observation. In addition, it can capture more complete spatial information from HSR image data rather than existing measures [10]. Besides, it is fully feed-forward and has faster running speed than other SR methods. During whole super-resolution CNN (SRCNN) process, the LR image is first resized to fit the HR image through traditional bicubic interpolation to serve as network input, then after learning the weights of the different filters of the network and minimizing the error between the output produced by the network and original HR images, we get the final parameters for the resolution enhancement model. It is composed by three convolutional neural networks as follows:

A. patch extraction and representation :

We use first layer to extract patches from LR image and represent them by a set of pre-trained bases, the whole extracting process is seemed as using a set of filters to convolute images in CNNS. Compared with rectified linear units (ReLUS) in SRCNN activation function which be defined as $y = \max(0, Wx + b)$, we apply the exponential linear unit (ELU) activation function in this paper to bring the normal gradient closer to the unit natural gradient because of a reduced bias shift effect [11] and reduce computational complexity. In addition, ELU ensure a noise-robust deactivation state and decrease the forward propagated variation and information. Then the first layer is expressed as an operation F1:

$$F_1(Y) = \max(0, W_1 * Y + B_1) + \min(0, \alpha(\exp(X)\text{-}1)) \quad (2)$$

where W_1 is of a size $c \times f_1 \times f_1 \times n$, c is the number of channels in the input image, f_1 is the spatial size of a filter and n_1 is the number of filters, α =0.1

B. non-linear mapping

The second layer nonlinearly maps n_1 dimensional features extracted by the first layer into n_2 dimensional vectors which represent the HR patch, the mapping operation between LR features and HR features can be expressed as:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) + +\min(0, \alpha(\exp(X)\text{-}1)) \quad (3)$$

where W_2 is of a size $n_1 \times f_2 \times f_2$, $B_2$is deviation.

C. Reconstruction

The final HR image is formed by averaging HR patches from the second mapping layer, so the last layer combines predictions within a spatial neighborhood to produce the final HR image F(Y):

$$F(Y) = W_3 * F_2(Y) + B_3 \quad (4)$$

where W_3 is of a size $n_2 \times f_3 \times f_3 \times c$, B_3 is a c dimensional biases vector.

D. Training

We use the Mean Squared Error (MSE) to estimate the network parameters $\Theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ to learn the end-to-end mapping F. By using stochastic gradient descent (SGD), the weights W_1, W_2, W_3 are renewed. The loss function with inputs $\{X_i\}$ and labels $\{Y_i\}$ is used where the optimization objective is defined as:

$$L(\Theta) = \frac{1}{n}\sum_{i=1}^{n}\left\| F(X_i;\Theta) - Y_i \right\|^2 \quad (5)$$

n is the number of training samples.

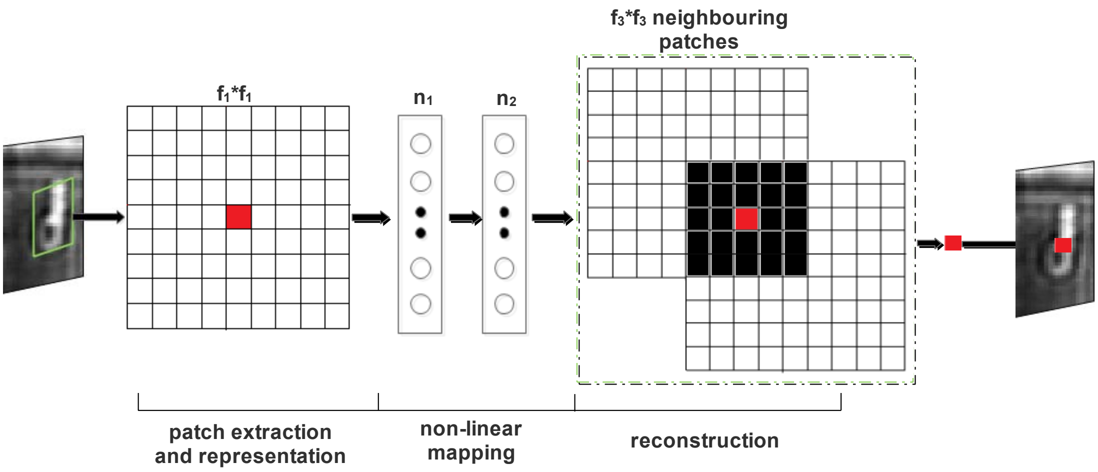

Fig4. Give a LR image, the first layer of the SRCNN extracts a set of feature maps. The second layer maps these features maps nonlinearly to HR patches representations. The last layer combines the predictions within a spatial neighborhood to produce the final HR image.

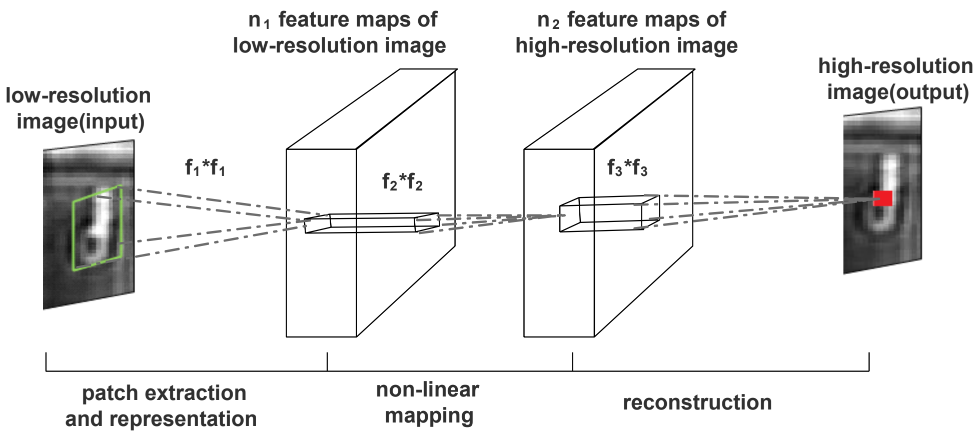

Fig.5 An illustration of sparse-coding-based method in the view of a convolutional neural network.

III. EXPERIMENTS

In this section, we applied the training model described upon to the HSR images, including the experimental setting and results. The platform of the whole training was GTX 960 GPU.

A. Experimental setting

1) Benchmarked models:

a) Step one: we use the model structure described in sectionⅡ, to be specific, we upscale training sets by bicubic interpolation method to get LR inputs, then segment them into 33×33 image patches used for the input layer, 64 filters of size 9×9 in the first layer, 32 filters in the second layer and filter in the 5×5 of third layer, the learning rate for the input layer and the mapping layer is 10^{-4},the thrid layer and the reconstruction layer is 10^{-5}.

b) Step two: We use a deeper neural network structure named NSRCNN to improve anti-noise performance by adding a layer to SRCNN between the mapping layer and the reconstruction layer to get the deeper one .Specific parameters are: 64 filters of size 9×9 in the first layer, 64 filters of size 7×7 in the second layer,32 filters of 7×7 in the third layer and 5×5 for the last layer, learning rates are same as the step one. LR models are all resized through bicubic interpolation as usual.

2) Evaluation setting: We evaluate enhancement results in terms of the PNSR between LR images and HR images.

$$PNSR = 10\log\left(\frac{1}{\sum_{i=1}^{n}\left|A(i)-B(i)\right|^2 / n}\right)$$

Where A denotes the original HSR imaging result, B is the reconstructed HR image and n is the number of pixels in the image. $|A(i)-B(i)|$ is the difference between the reconstruction and the original HR image.

B. Experimental results

We use SRCNN model and NSRCNN model to enhance the resolution for target images respectively. We also use bicubic interpolation, wavelet-based interpolation and phase-driven spatially variant regularization for the same picture as control. Fig.6 represents the reconstruction performance, in terms of mean PNSR (dB) gain of the different images tested in experiments. Over all, SRCNN model and NSRCNN model perform well in resolution enhancement against traditional interpolation methods according to the PNSR results .

TABLE I. THE RESULTS OF PNSR (DB) COMPARISON ON THE CAMERA(1),MICROPHONE(2),ROCKET(3),WRENCH(4),SHELL(5) HSR IMAGE RESPECTIVELY

Images / Methods	1	2	3	4	5	Average
Bicubic	30.79	28.57	27.69	32.25	30.01	29.86
PDM	30.95	29.25	27.98	32.17	30.14	30.10
WIM	31.53	29.72	27.84	33.19	30.17	30.49
SRCNN	31.75	30.34	28.06	33.22	30.83	30.84
NSRCNN	31.79	30.42	28.11	33.26	30.97	30.91

IV. CONCLUTION

This paper represents a resolution enhancement model used for holographic subsurface penetrating imaging, on the basis of an intensive description of resolution enhancement structure. We use the exponential linear unit (ELU) activation function to reduce computational complexity and add a layer between the mapping layer and reconstruction layer to ensure a noise-robust state. In addition, experiments have been conducted and the experiment results are presented, both of the subjective and objective evaluations represent the superior performance of convolutional neural network resolution enhancement than other methods. However, it is admitted that the processing speed for NSRCNN is still unsatisfactory owning to its time-consuming non-linear mapping step. Therefore, our further work will mainly concentrate on raising processing speed and paving the way for the follow-up detection and recognition tasks.

ACKNOWLEDGMENT

This paper is supported by National Natural Science Foundation of China Grant No: 61501234. Thanks are due to Dr. S. I. Ivashov from Bauman Moscow State Technical University (BMSTU) for assistance with the training and testing HSR images.

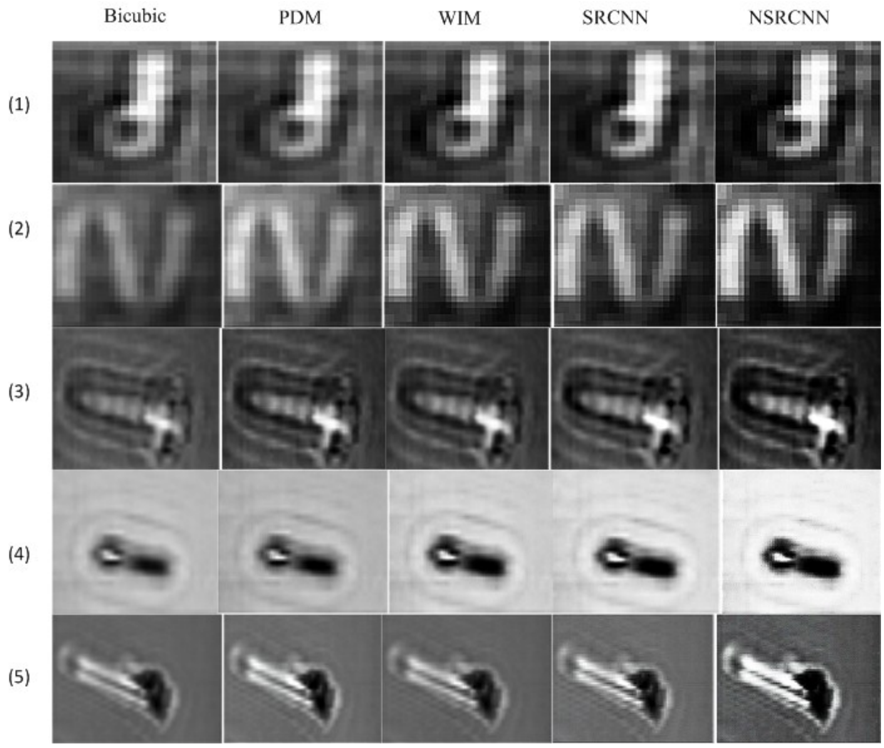

Fig.6 Corresponding result of .camera(1),microphone(2),rocket(3),wrench(4),shell(5) HSR image respectively.

REFERENCES

[1] S. I. Ivashov, V. V. Razevig, I. A. Vasiliev, A. V. Zhuravlev, T. D. Bechtel and L. Capineri, "Holographic Subsurface Radar of RASCAN Type: Development and Applications," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 4, pp. 763-778, Dec. 2011.

[2] S. I. Ivashov, V. V. Razevig, T. D. Bechtel, I. A. Vasiliev, L. Capineri and A. V. Zhuravlev, "Microwave holography for NDT of dielectric structures," *2015 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS)*, Tel-Aviv, 2015, pp. 1-5.

[3] L. Capineri, P. Falorni, V. Razevig, M. Inagaki, T. Bechtel and C. Windsor, "Quantitative interpretation of RASCAN holographic radar response from inclined plane reflectors by a theoretical model," *Proceedings of the XIII Internarional Conference on Ground Penetrating Radar*, Lecce, 2010, pp. 1-4.

[4] A. Quintero-Zea and M. Osorio-Cárdenas, "RASCAN type radar image resolution enhancement for non-metallic landmine detection," *Symposium of Signals, Images and Artificial Vision - 2013: STSIVA - 2013*, Bogota, 2013, pp. 1-4.

[5] A. C. Gurbuz, J. H. McClellan and W. R. Scott, "Compressive Sensing for GPR Imaging," *2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2007, pp. 2223-2227.

[6] R. Timofte, R. Rothe and L. V. Gool, "Seven Ways to Improve Example-Based Single Image Super Resolution," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 1865-1873.

[7] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, Feb. 1 2016.

[8] J. Sun, Wenfei Cao, Zongben Xu and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 769-777.

[9] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus,OH, 2014, pp. 580-587.

[10] M. Oquab, L. Bottou, I. Laptev and J. Sivic, "Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 2014, pp. 1717-1724.

[11] H. Ide and T. Kurita, "Improvement of learning for CNN with ReLU activation by sparse regularization," *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, USA, 2017, pp. 2684-2691.