DL-CSNet: Dictionary Learning based Compressed Sensing Neural Network

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

23-02-2022 / 01-03-2022

CITATION

Qiu, Yanzhen; Zhang, Chuangfeng; Xie, Yuhang; Huang, Ruishan; Tian, Haochen; Xiong, Chenkui; et al. (2022): DL-CSNet: Dictionary Learning based Compressed Sensing Neural Network. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.19222650.v1>

DOI

[10.36227/techrxiv.19222650.v1](https://doi.org/10.36227/techrxiv.19222650.v1)

DL-CSNet: Dictionary Learning based Compressed Sensing Neural Network

Yanzhen Qiu¹, Chuangfeng Zhang^{*1}, Yuhang Xie², Ruishan Huang¹, Haochen Tian¹, Chenkui Xiong¹, Shaolin Liao^{*1,3}

¹Sun Yat-sen University, No. 135, Xingang Xi Road, Guangzhou, 510275, P. R. China

²Texas A&M University, College Station, TX 77843, USA

³Illinois Institute of Technology, 10 West 35th Street, Chicago, IL 60616, USA

*Corresponding authors' e-mails:

liaoshlin@mail.sysu.edu.cn; zhangchf8@mail2.sysu.edu.cn

Abstract. In this paper, we propose a novel neural network for Compressed Sensing (CS) application: the Dictionary Learning based Compressed Sensing neural Network (DL-CSNet). It is fairly simple but highly effective, which consists of only three layers: 1) a DL layer for latent sparse features extraction; 2) a smoothing layer via Total Variation (TV) like constraint; and 3) a CS acquisition layer for neural network training. In particular, the TV-like smoothing layer is a perfect complement to the sparsity-oriented DL layer to achieve smooth images. The trained DL-CSNet can learn the optimal dictionary matrix so that images can be reconstructed in high quality. At last, extensive experiments have been carried out on binary images and compared to most classical CS algorithms, which shows the superior performance of the proposed DL-CSNet.

1. Introduction

After Nyquist-Shannon sampling theorem, Donoho and Candès proposed a new sensing modality Compressed Sensing (CS [1, 2, 3]), which compresses the signal being acquired at the time of sensing. Since then, CS has been used in many applications such as fast millimeter-wave wireless image transmission [4] and millimeter-wave imaging [5, 6, 7, 8]. CS not only breaks the sampling rate limitation of the Nyquist-Shannon sampling theorem, but also alleviates the storage demand. For sparse signals, CS can accurately reconstruct the original signals with a much smaller number of samples than Nyquist-Shannon theorem. Specially, the sparsity of signals means that the whole information contained in signals can be represented in some basis set and signals can be completely reconstructed in such basis set, which is the utmost important and necessary property for CS. Generally, some signals (for example, natural images) in their original basis sets are not sparse. However, these signals are implicitly redundant, which have sparse representations in some transforming basis set, such as the Fourier transform and the wavelet transform. CS theory shows that the sparser the signal, the better the signal can be reconstructed. However, not all signals are sparse in Fourier or wavelet basis set, especially when they change greatly with time. All of these call for Dictionary Learning (DL), which has stronger sparse basis set learning capability.

In this paper, we propose a novel DL based CS neural Network (DL-CSNet) with a Total Variation (TV) like constraint for better CS image reconstruction and our contributions are: 1) a DL layer is used for latent features learning; 2) a TV-like smoothing layer is used to improve the quality of the CS image reconstruction; 3) experiments have been carried out to verify the superiority of our DL-CSNet, compared to most classical CS algorithms, such as Orthogonal Matching Pursuit (OMP [9]),

Figure 1. The process of CS.

Least Absolute Shrinkage and Selection Operator (LASSO [10]), Iterative Hard Thresholding (IHT [11]), Subspace Pursuit (SP [12]) and so on.

2. Preliminary

2.1. Sparse Image

In real-world, most signals are non-sparse in their original representation basis set. However, these signals are implicitly sparse in some basis set, for example, in the Fourier basis set. In general, an implicit sparse image 1D vector $\bar{X}$ can be expressed in its sparsifying basis set $\bar{D}$ as follows

$$\bar{X} = \bar{D}\bar{Z}, \quad (1)$$

where $\bar{Z}$ is the K -sparse coefficient vectors with K non-zero elements.

2.2. Compressed Sensing

The process of CS is illustrated in Fig. 1, which consists of 4 main steps: 1) image sparse representation; 2) the CS acquisition by random measurement matrix; 3) CS reconstruction of sparse coefficients; 4) original image recovery.

With the help of Equation (1), the acquisition model of CS can be described as:

$$\bar{Y} = \bar{M}\bar{X} = \bar{M}\bar{D}\bar{Z} = \bar{\Phi}\bar{Z}; \quad \bar{\Phi} = \bar{M}\bar{D}, \quad (2)$$

where $\bar{M} \in \mathbb{R}^{M \times N}$ and $\bar{\Phi} \in \mathbb{R}^{M \times N}$ are the original measurement matrix and the effective measurement matrix respectively; and $\bar{Y} \in \mathbb{R}^M (M \ll N)$ is the measurement vector. The sparse coefficients vectors $\bar{Z}$ can be solved through classical algorithms like OMP [9], LASSO [10], IHT [11], and SP [12], as well as the L_1 norm minimization

$$\underset{\bar{Z}}{\operatorname{argmin}}\{\|\bar{Z}\|_1\}, \text{ s.t. } \bar{Y} = \bar{\Phi}\bar{Z}. \quad (3)$$

2.3. Total-Variation like Constraint

TV [13] is popular in noise removal which improves the visual quality of images. Here, the following TV like constraint is used for the reconstructed 2D CS image $\bar{\bar{X}} \in \mathbb{R}^{n \times m}$,

$$TV(\bar{\bar{X}}) = \sum_{i=1}^n \sum_{j=1}^m \left[\left(\nabla_{i,j}^h \bar{\bar{X}} \right)^2 + \left(\nabla_{i,j}^v \bar{\bar{X}} \right)^2 \right]; \quad \nabla_{i,j}^h \bar{\bar{X}} = \bar{\bar{X}}_{i+1,j} - \bar{\bar{X}}_{i,j}, \quad \nabla_{i,j}^v \bar{\bar{X}} = \bar{\bar{X}}_{i,j+1} - \bar{\bar{X}}_{i,j}. \quad (4)$$

3. Our Work

In this section, we propose the DL-CSNet, as illustrated in Fig. 2, which can not only learn the suitable dictionary basis of the original images, but also reconstruct the original images with high quality. The DL-CSNet consists of 3 layers: a DL layer, a TV-like smoothing layer, and a CS acquisition layer.

3.1. Dictionary Learning Layer

Figure 2. The Architecture of the proposed DL-CSNet.

Two kinds of sparsifying basis sets can be used to represent sparse images: 1) the orthogonal basis set, 2) the redundant dictionary basis set. Also, the redundant dictionary basis set consists of the parameter-based dictionary basis set and the sample training-based dictionary basis set. In our DL layer, the K-SVD (Kernel Singular Value Decomposition [14]) dictionary learning basis set is used, which has been proved to get better effect on image sparse representation [14]. Specifically, the DL layer with output $\hat{X}$ can be trained by solving the following optimization problem [15]:

$$\underset{\bar{D}, \bar{Z}}{\operatorname{argmin}} \left\| \hat{X} - \bar{D} \bar{Z} \right\|_2^2, \quad \text{s.t. } \forall i, \|z_i\|_1 \leq T_0; \quad \bar{Z} = [z_1, z_2, \dots, z_N]^T. \quad (5)$$

3.2. Smoothing Layer

Dictionary learning can effectively extract the latent structural features from the images, but it may generate unwanted high-frequency variation to the reconstructed images. To overcome this defect, a smoothing layer is added behind the DL layer. In smoothing layer, the images become smoother by minimizing the TV-like regularization term TV , where represents $\tilde{X}$ the images reconstructed from the smoothing layer. Finally, it should be noted that Rectified Linear Unit (ReLU) has been used to control the convergence rate of the DL-CSNet training.

3.3. CS Acquisition Layer

As showed in Fig. 2, the CS acquisition layer is the last layer of DL-CSNet for network training. With the help of Equations (2)-(5), the DL-CSNet is trained with the following loss function

$$L = \alpha \left\| \tilde{Y} - \bar{Y} \right\|_2^2 + \beta \left\| \tilde{Z} \right\|_1 + \gamma TV(\tilde{X}), \quad \text{s.t. } \forall i, \|z_i\|_1 \leq T_0,$$

where α, β, γ are the weights of the Mean Square Error (MSE), the L_1 norm in Equation (3) and the TV-like constraint in Equation (4) respectively.

4. Experiments

In this section, we evaluate the performance of the proposed DL-CSNet for CS reconstruction and compare it to other 5 classical CS algorithms, *i.e.*, OMP [9], LASSO [10], IHT [11], SP [12] and L_1 -min [16]. First, 28×28 binary images are used for all experiments, which contain patches and sparse points that are common in object classification of automotive sensors like millimeter-wave radars [17]; Second, the CS sampling matrix used is a Gaussian random matrix; Finally, the method described in [18] is used to initialize the weights, which is a sound procedure for neural networks utilizing ReLU.

Figure 3. Reconstructed images for each algorithm of Experiment 1.

Table 1. MSE, SSIM, PSNR of Experiment 1 with a CS Sampling Ratio of 6.25%

Algorithm	MSE	SSIM	PSNR (dB)
DL-CSNet	3.51×10^{-5}	0.9998	44.55
OMP	0.0305	0.4671	15.15
IHT	0.0319	0.4698	14.97
L_1 -min	0.0319	0.3706	14.96
LASSO	0.0319	0.3788	14.96
SP	0.0274	0.6403	15.63

A total of 3 experiments have been carried out: 1) in Experiment 1, we design a binary image with one 5×5 patch, with 6.25% CS sampling ratio 2) in Experiment 2, we design a binary image with one 5×5 patch and some random sparse points, with 11% CS sampling ratio; and 3) in Experiment 3, the binary image has one 5×5 patch and one 3×5 patch, with 6.25% CS sampling ratio.

5. Results and Discussion

We evaluate the performance of each algorithm based on the criteria of MSE, Structural Similarity Image Measurement (SSIM) and Peak Signal-to-Noise Ratio (PSNR). The results of Experiment 1 are shown in Fig. 3 and Table 1, showing that our DL-CSNet achieves the best SSIM (0.9998 vs. ≤ 0.4698) and PSNR (44.55 dB vs. ≤ 15.63 dB), as well as the smallest MSE value. Similarly, Fig. 4 and Table 2 show the reconstructed images and MSE, SSIM and PSNR of all algorithms for Experiment 2. Clearly, our DL-CSNet again achieves the highest SSIM (0.9999 vs. ≤ 0.3464) and PSNR (51.79 dB vs. ≤ 14.69 dB), and the smallest MSE among all algorithms. In particular, the PSNR of our DL-CSNet is > 37 dB over all other classical CS algorithms. Finally, Fig. 5 and Table 3 show that the results of Experiment 3. Even though the binary image becomes more complex and the sampling ratio is set to a very small value of 6.25%, our DL-CSNet can still recover the binary image with much better quality than other algorithms: a SSIM of 0.7323 vs. ≤ 0.3948 and a PSNR of 20.38 dB vs. ≤ 12.97 dB.

6. Conclusion

A novel Dictionary Learning based CS neural Network, *i.e.*, the DL-CSNet, has been proposed for efficient CS image reconstruction. Under the DL-CSNet, the K-SVD like DLNet has been used to obtain the latent features of the image. In addition, in order to overcome the shortcoming of the high-frequency variation induced by the DLNet, a TV-like smoothing layer has been added to the DL-CSNet. Extensive experiments demonstrate that the DL-CSNet shows clear superiority over most classical CS algorithms. At last, the DL-CSNet could be accelerated on hardware such as FPGA [19].

Figure 4. Reconstructed images for each algorithm of Experiment 2.

Table 2. MSE, SSIM, PSNR of Experiment 2 with a CS Sampling Ratio of 11%

Algorithm	MSE	SSIM	PSNR (dB)
DL-CSNet	6.63×10^{-6}	0.9999	51.79
OMP	0.0383	0.3464	14.17
IHT	0.0381	0.1973	14.20
L_1 -min	0.0340	0.2416	14.69
LASSO	0.0319	0.3226	14.69
SP	0.0363	0.2367	14.40

Figure 5. Reconstructed images for each algorithm of Experiment 3.

Table 3. MSE, SSIM, PSNR of Experiment 3 with a CS Sampling Ratio of 6.25%

Algorithm	MSE	SSIM	PSNR (dB)
DL-CSNet	0.0092	0.7323	20.38
OMP	0.0505	0.3790	12.97
IHT	0.0534	0.3663	12.73
L_1 -min	0.0525	0.3386	12.80
LASSO	0.0515	0.3948	12.88
SP	0.0510	0.4384	12.93

References

- [1] Donoho, D.L. (2006) Compressed sensing. *IEEE Transactions on Information Theory*, 52:489-509.
- [2] Candès, E., Romberg, J. and Tao, T. (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Transactions on Information Theory*, 52:489–509.
- [3] Candès, E. and Tao, T. (2006) Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies? *IEEE Transactions on Information Theory*, 52: 5406–5425.
- [4] Liao, S. and Ou, L. (2020) High-speed Millimeter-wave 5G/6G Image Transmission via Artificial Intelligence. In: 2020 IEEE Asia-Pacific Microwave Conference (APMC). HongKong. pp 655–657.
- [5] Gopalsami, N., Liao, S., Elmer, T., et al. (2011) Compressive sampling in active and passive millimeter-wave imaging. In: 2011 International Conference on Infrared, Millimeter, and Terahertz Waves. Houston, TX. pp 1–2.
- [6] Babacan, S.D., Luessi, M., Spinoulas, L., et al. (2011) Compressive passive millimeter-wave imaging. In: 2011 18th IEEE International Conference on Image Processing. Brussels. pp 2705–2708.
- [7] Gopalsami, N., Elmer, T.W., Liao, S., et al. (2011) Compressive sampling in passive millimeter wave imaging. In: SPIE 8022, Passive Millimeter-Wave Imaging Technology XIV. Orlando. pp 80220I–1–6.
- [8] Gopalsami, N., Liao, S., Elmer, T., et al. (2012) Passive millimeter-wave imaging with compressive sensing. *Optical Engineering*, 51: 1–9.
- [9] Pati, Y., Rezaeiifar, R. and Krishnaprasad, P. (1993) Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Proceedings of 27th Asilomar Conference on Signals, Systems and Computers. Pacific Grove. pp 40–44.
- [10] Tibshirani, R.J. (1996) Regression Shrinkage and Selection via the LASSO. *Journal of the Royal Statistical Society. Series B: Methodological*, 73: 273–282.
- [11] Blumensath T. and Davies M.E. (2009) Iterative hard thresholding for compressed sensing. *Applied and Computational Harmonic Analysis*, 27: 265–274.
- [12] Dai, W. and Milenkovic, O. (2009) Subspace Pursuit for Compressive Sensing Signal Reconstruction. *IEEE Transactions on Information Theory*, 55: 2230–2249.
- [13] Wen, Y.W., Ng, Michael, K. and Huang, Y.M. (2008) Efficient Total Variation Minimization Methods for Color Image Restoration. *IEEE Transactions on Image Processing*, 17: 2081–2088.
- [14] Aharon, M., Elad, M. and Bruckstein, A. (2006) K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54: 4311–4322.
- [15] Singh, S., Singhal, V. and Majumdar, A. (2017) Deep Blind Compressed Sensing. In: 2017 Data Compression Conference (DCC). Snowbird, Utah USA. pp. 459-459.
- [16] Donoho, D.L. (2006) For most large underdetermined systems of linear equations the minimal L_1 -norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59: 797–829.
- [17] Zhang W., Li, N., Zha H., et al., A Software-Adaptive 77GHz Radar Sensor for Traffic Applications, 2021 IEEE MTT-S International Wireless Symposium (IWS), 2021, pp. 1-3, doi: 10.1109/IWS52775.2021.9499682.
- [18] He, K., Zhang, X., Ren, S., et al. (2015) Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In: 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile. pp. 1026–1034.
- [19] Huang W., Wu H., Chen Q. et al., FPGA-Based High-Throughput CNN Hardware Accelerator with High Computing Resource Utilization Ratio, *IEEE Transactions on Neural Networks and Learning Systems*, doi: 10.1109/TNNLS.2021.3055814.