Deep learning-based denoising for magnetic resonance spectroscopy signals

Lei, Yang, Ji, Bing, Liu, Tian, Curran, Walter, Mao, Hui, et al.

Yang Lei, Bing Ji, Tian Liu, Walter J. Curran, Hui Mao, Xiaofeng Yang, "Deep learning-based denoising for magnetic resonance spectroscopy signals," Proc. SPIE 11600, Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, 1160006 (15 February 2021); doi: 10.1117/12.2580988

SPIE.

Event: SPIE Medical Imaging, 2021, Online Only

Deep Learning-based Denoising for Magnetic Resonance Spectroscopy Signals

Yang Lei[a,c#], Bing Ji[b#], Tian Liu[a,c], Walter J. Curran[a,c], Hui Mao[b,c*], Xiaofeng Yang[a,c*]

[a]Department of Radiation Oncology,
[b]Department of Radiology and Imaging Science,
[c]Winship Cancer Institute,
Emory University, Atlanta, GA 30322.

[#]Co-author

*Corresponding author: xiaofeng.yang@emory.edu and hmao@emory.edu

ABSTRACT

Magnetic resonance spectroscopy (MRS) is a non-invasive imaging tool for detecting and quantifying metabolites in vivo. However, the spectral data often suffer from poor signal-to-noise ratio (SNR) due to low concentrations of metabolites and limited acquisition time. In this work, we introduce a deep learning-based method for denoising MRS spectra. By identifying and characterizing the sparse representations in the feature vector domain of the noise in MRS data using the deep learning network, the non-signal components can be removed to improve the SNR of spectral data. We used a stack auto-encoder (SAE) network to train the deep learning-based model based on high SNR data collected from a brain phantom using a high number of signal average (NSA=192). To overcome overfitting, the SAE network is trained in a patch-based manner. The network and denoising method were then tested using the noisy data collected in the same sample volume but using low NSA (NSA=8) with much shorter acquisition time. The denoising performance of the reported method was then evaluated using statistical comparison analysis of spectra before and after being denoised and the paired "ground truth" high NSA spectra based on the measurements of SNR, and mean squared error (MSE) and known metabolite levels. for the results of testing and evaluating the reported method showed that the SNR improved by 40% and the MSE reduced by 72% in the data collected from the brain phantom, while , the SNR improved by 47% and the MSE reduced by 27% in the data collected from the human subjects. In summary, the reported deep-learning method demonstrated a model-free approach to enhance SNR of noisy MRS data that otherwise cannot be used for quantitative analysis. The potential applications of include using low NSA or short data collection time to accelerate MRS exams while maintaining adequate spectroscopic information for detection and quantification of the metabolites of interest when a little time is available for an MRS exam in the clinical setting.

Keywords: Magnetic resonance spectroscopy, spectrum, denoise, stack auto-encoder, sparse representation.

1. INTRODUCTION

Magnetic resonance spectroscopy (MRS) is capable of revealing important biochemical and metabolic information of tissues non-invasively [1]. However, the low concentrations of metabolites and limited acquisition time in studing live subjects or patients often lead to poor signal-to-noise ratio (SNR) of the spectra [2]. Reducing or even eliminating noise can improve SNR sufficiently to obtain high quality spectra instead of increasing the number of signal averaging (NSA) or the field strength, both of which may not be practical in clinical applications. While the noise can be mitigated by applying a smoothing filter, such as Gaussian filter, in either spatial or spectral domains, filtering causes loss of a portion of the signal as it is impossible to completely differentiate noise from the signal. Furthermore, filtering also leads to spectral linewidth broadening, therefore compromising spatial resolution [3].. Other denoising approaches that have been investigated include: the wavelet thresholding method which separates signal and noise wavelets based on a preset threshold value [4-6] and the wavelet shrinkage method which uses specific wavelet coefficients to recognize the signal from the noise.

Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging, edited by Barjor S. Gimi, Andrzej Krol, Proc. of SPIE Vol. 11600, 1160006 · © 2021 SPIE
CCC code: 1605-7422/21/$21 · doi: 10.1117/12.2580988

Here we present the stack auto-encoder (SAE) network approach to obtain the high SNR spectra from the original noisy spectra by minimizing the reconstruction error. Due to the noise peaks in MRS data have their sparse representation in the feature vector domain [7, 8], a stack auto-encoder (SAE) network is used to identify and characterize the noise components followed by denoising the spectroscopic data. SAE is used due to its ability of sparse representation without the need of prior knowledge as compared to compress sensing. A patch-based method was utilized to enlarge data variation. SAE was then used to make representing noisy data in a sparse feature vector filed to reduce noisy component. The performance of this denoising approach was evaluated by measuring and comparing SNRs of the low NSA spectra before and after denoising. The results demonstrated that SAE effectively improved SNR by removing noise from the noisy spectra. This paper is organized as follows. Section 2 presents how we structured a SAE network to perform denoising for MRS data. In Section 3, we described the validation studies. We then discussed the limitations and potentials of the proposed methods and concluded in Section 4

2. METHOD

A spectroscopy phantom (Model 2152220, GE Healthcare, Chicago, IL) containing selected brain metabolites with known concentrations, including N-acetylaspartate (NAA), creatine (Cr), phosphocreatine (PCr), choline (Cho), myo-inositol (MI), lactate (Lac), glutamate (Glu) and glutamine (Gln), was placed in a 20-channel phase-array head coil positioned at the iso-center of a 3T MR scanner (Prisma, Siemens Healthcare, Erlangen, Germany).

T1-weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) images were first acquired using parameters of TR of 500 ms; TE of 8.7 ms; flip angle of 90°, matrix of 180×240 mm^2, field of view (FOV) of 192×256 pixels, slice thickness of 2 mm, to assist localization of the sampling voxel. Single-voxel proton (1H) spectroscopy (SVS) data were collected in the volume of 20×20×20 mm^3 using a point resolved spectroscopy sequence (PRESS) with echo time (TE) of 30 ms, repetition time (TR) of 2000 ms, acquisition bandwidth of 1200 Hz and vector size of 1024. The standard shimming method implemented on the scanner was applied. Chemical shift selected pre-saturation was used for water suppression in all MRS acquisitions.

SVS data acquisitions were performed on the phantom at different sampling locations, i.e., one voxel at the iso-center of the phantom and then one each (same volume) shifted 30 mm away from the initial iso-center position to the anterior, right, and feet direction, respectively. In each session, two spectral datasets with NSA of 192 (henceforth referred to as the high NSA spectra) and 8 (henceforth referred to as the low NSA spectra) were acquired at the same voxel locations and volumes.

In addition, we collected brain SVS data from healthy subjects (N = 5, four females, age range: 29-41 years) with the same acquisition protocol used for the phantom scan. The voxels of 20×20×20 mm^3 were selected in parietal and temporal regions Similar to data collection from the phantom, two spectral datasets, one high NSA spectrum (i.e., NSA of 192) and one low NSA spectrum (i.e., NSA of 8), were acquired from the same voxel under the same shimming and water suppression conditions. The human subject study was done in accordance with the institutional review board and written consents were obtained from all of participants.

Figure 1 shows the workflow of our proposed SAE algorithm constituting the training (left) and testing (right). The training and testing follow the same feed forward path of the network. The difference between training and testing is, during training a backward path was performed to optimized the learnable parameters of the network via minimizing the mean square error (MSE) loss term calculated between input noisy and clean data SAE was used to build the denoising model because SAE can represent the input data into a deep sparse feature representation space without the need of prior knowledge, then via hidden layers and nonlinear activation function, the noisy component can be removed in feature space. For a pair of training original (noisy) spectra and its corresponding ground truth (clean) signal which was obtained from the high SNR dataset collected using the number of signal average (NSA) of 192. The clean signal was used as learning target of the proposed SAE model.. Among a total number of 192 individual spectra in the high NSA dataset, we randomly and equally separate the MRS spectra for the 96 signals as training and the rest 96 signals as testing. During training, patch-based vectors are extracted from training MRS spectra and the ground truth (high NSA spectra). The SAE reconstructs the clean signal from the original noisy signal by minimizing the reconstruction error. The reconstruction error was obtained by calculating the mean

absolute error between reconstructed signal and ground truth signal. The SAE is established by stacking a series of the encoder and decoder layers, and the output of the bottom layer is the input of the second layer. The learnable weights of depth structure are initialized layer by layer, in addition to fine-tuning all weights with Adam gradient descent algorithm.

After training the SAE, the original noisy spectra of a new arrival data are fed into the trained SAE model to obtain the denoised signal as shown in the left part of Figure 1. The framework is robust to the input as it can reconstruct the signal from the noisy input. In addition, it can capture robust features by punishing and inhibiting the larger changes in the hidden layer.

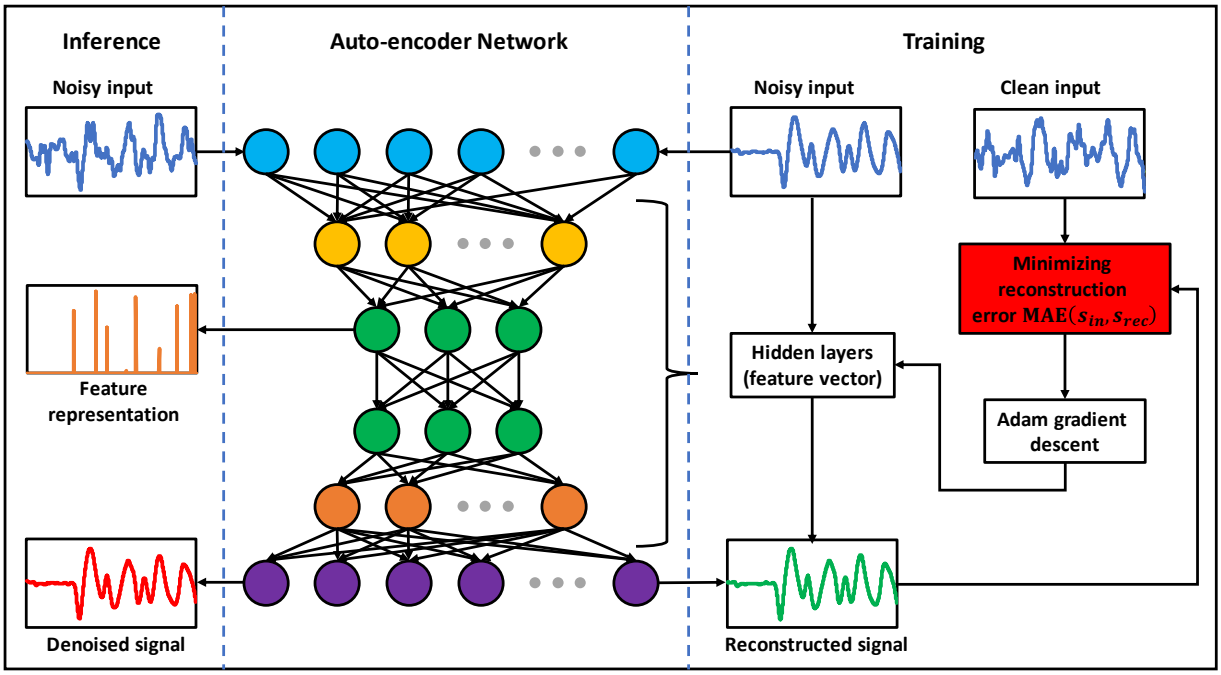

Figure 1. Schematic flow-chart of the proposed MRS signal denoising. The left column shows the inference of an arrival MRS signal. The middle column shows the network architecture of auto-encoder. The right column shows the training procedure.

2.1 Stack auto-encoder

Based on the architecture of SAE shown in Figure 1, the input signal v is corrupted to noisy input $\hat{v}$. Through batch normalization, the input noisy vector $\hat{v}$ is normalized into $[0,1]^d$, where d is the dimension of the vector. Via encoder layers, the input $\hat{v}$ is encoded to the hidden output h. The objective of encoder layers is to hide some of the information from the corrupted signals $\hat{v}$, and thus get rid of noisy information. Then, the partial information h is reconstructed via decoder layers to a full-length signal y. Thus, the aim of SAE is to make h a feature representation of the input vector $\hat{f}$ by minimizing the reconstruction error between v and y. Furthermore, it will train the hidden layer to reconstruct the full signals from the partial information of the original signals.

Denoting $\theta = \{A, b\}$ are weighting and bias parameters for encoder layers and $s(x) = 1/\big(1 + exp(-x)\big)$ is a sigmoid operator, which is a nonlinear activation function, the hidden output can be represented as

$$h = s(A\hat{v} + b) \tag{1}$$

Denoting $\theta' = \{A', b'\}$ are weighting and bias parameters for decoder layers, we define the decoder's output as

$$y = s(A'h + b') \tag{2}$$

For each denoised signal v, it is corrupted to a noisy signal $\hat{v}$, and get the feature vector h by encoding the noisy signal. Then it decodes the feature vector to the reconstruction vector of the denoised signals v. Thus, the learnable parameters are optimized by minimizing the reconstruction error:

$$\theta, \theta' = arg \min_{\theta,\theta'} \|v - y\|_2^2 + \mu \cdot \|h\|_1 \quad (3)$$

where μ is a balancing parameter, $\|v - y\|_2^2$ aims to minimize the difference between v and y, $\|h\|_1$ is a penalization term to force the feature vector of hidden layer be sparse.

2.2 Evaluation and Validation

The time domain free induction decay (FID) data were first converted to the frequency domain spectra by applying Fourier transformation using LCModel (version 6.3-1H) [9]. Phase correction was performed automatically in LCModel. After phase correction, the estimated baseline was obtained through fitting the spectra with a sextic polynominal equation [10] using an in-house program written in MATLAB (R2019a, The MathWorks Inc., Natick, MA). This estimated baseline was subsequently subtracted from the spectrum to remove the distortion used by most reports [11, 12] for baseline correction. Finally, the baseline-corrected spectra were normalized to help compare the performance of different denoising algorithm using NAA peak at the chemical shift of 2.02 ppm as the reference for the subsequent analysis[13].

A hold-out test was used to evaluate the performance of proposed method, namely, we randomly selected 96 from 192 high NSA spectra as training data and took the rest signals as testing data. In addition, additional 9 human noisy MRS signals were used for testing. To evaluate denoising performance of the SAE algorithm, we measure the SNR and MSE of the spectra before and after denoising processes.

Our algorithm was implemented in python 3.6 and Tensor flow with Adam gradient descent optimizer and were trained and tested on a NVIDIA Tesla V100 GPU with 32GB of memory. The learning rate for Adam optimizer was set to 2E-5. The batch size was set to 100. The number of epochs for training was set to 200. For each epoch, 100 iterations of training were performed. The patch size was set to 32. During training, 1.2 GB CPU memory and 7.6 GB GPU memory was used for each batch optimization. It took about 0.6 hour during training. In the testing, it took within 0.1 second to generate estimated full time-scale signals for the measurement.

3. RESULTS

The spectra in Figure 2 show significant SNR improvements in low NSA spectra from both phantom and the parietal and temporal lobes of human brains before and after being denoised with the SAE approach ($p < 0.05$). In addition, SAE approach yielded can preserve the majority of prominent metabolites (i.e., N-acetylaspartate, Creatine, Choline, myo-inositol and glutamine/glutamate) in the spectra of human spectra data.

Table 1 shows the numerical results calculated before denoising (i.e., a comparison between noisy MRS and clean MRS) and after denoising (i.e., a comparison between denoised MRS and clean MRS). In denosing data collected from the phantom, the SNR improved by 40% and the MSE reduced by 72%. When used to denosing the data from the human subjects, the SNR improved by 47% and the MSE reduced by 27%.

Table 1. Model performance results on 96 phantom MRS signals and 9 human MRS signals. Data are reported as mean ± STD.

	SNR Before (dB)	SNR After (dB)	MSE Before	MSE After
Phantom Data	22.38±4.28	31.34±3.35	45.7±12.6	12.8±4.6
Human Data	21.02±6.27	30.96±5.75	1.42E-3±0.94E-3	1.03E-3±0.83E-3

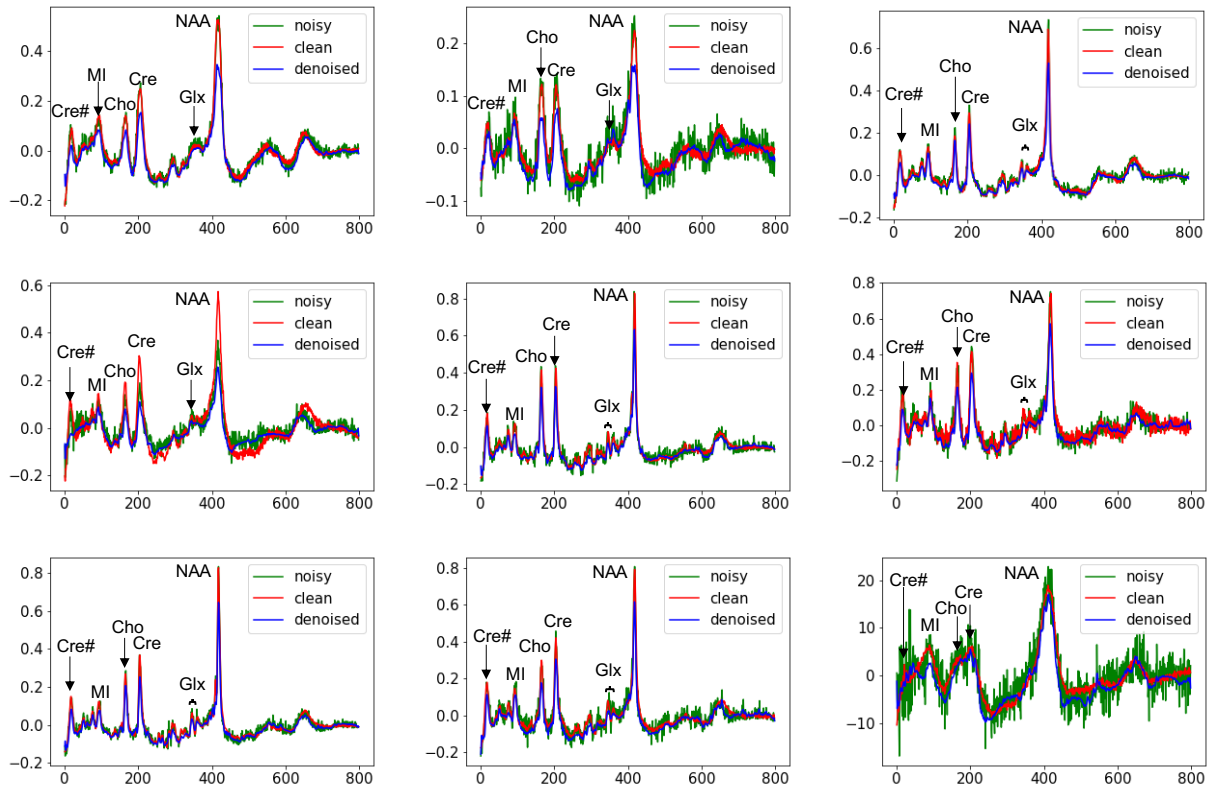

Figure 2. Comparison of the original and denoised 9 spectra from the human subjects. NAA: Nacetylaspartate; Cre, Cre#: Creatine Choline; MI: myo-inositol; Glx: glutamine/glutamate. NAA: N-acetylaspartate; Cre, Cre#: Creatine Choline; MI: myo-inositol; Glx: glutamine/glutamate.

4. DISCUSSION AND CONCLUSION

We demonstrated a novel deep-learning based spectral denoising method that integrates the patch-based stack auto-encoder model into MRS data denoising framework. This approach has two distinctive strengths: 1) to enlarge data variation, patch-based method was utilized to perform denoising. 2) to represent noisy data in a sparse feature vector filed and thus reduce noisy component, stack auto-encoder was used due to its ability of sparse representation without the need of prior knowledge as compared to compress sensing.

In summary, the presented SAE strategy offers an effective and efficient denoising method for gaining high quality spectra with high SNRs from the noisy spectra acquired with the low number of signal average and a short acquisition time. The excellent denoising performance without compromising the metabolite quantification by the SAE method is achieved by represent noisy data in a sparse feature vector filed and thus reduce noisy component. In the future, the reported SAE approach may assist MRS with the accurate and sensitive detection and measurement of metabolic abnormalities. The immediate impact of the SAE algorithm in clinical applications is to enable underutilized MRS to overcome major limitations of low sensitivity and long acquisition time. The potential applications of include using low NSA to accelerate MRS acquisition while maintaining adequate spectroscopic information for detection and quantification of the metabolites of interest when a little time is available for an MRS exam in the clinical setting.

ACKNOWLEDGEMENT

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award Number R01CA215718 (XY) and R01CA203388 (HM) and Emory Winship Cancer Institute pilot grant.

REFERENCES

[1] D. Soares, and M. Law, "Magnetic resonance spectroscopy of the brain: review of metabolites and clinical applications," Clinical radiology, 64(1), 12-21 (2009).

[2] R. Kreis, "Issues of spectral quality in clinical 1H-magnetic resonance spectroscopy and a gallery of artifacts," NMR in Biomedicine, 17(6), 361-381 (2004).

[3] Y. Zhang, J. Zhou, and P. A. Bottomley, "Minimizing lipid signal bleed in brain 1H chemical shift imaging by post-acquisition grid shifting," Magnetic resonance in medicine, 74(2), 320-329 (2015).

[4] W. Chunli, and Z. Chunlei, "Denoising algorithm based on wavelet adaptive threshold," Physics Procedia, 24, 678-685 (2012).

[5] B. Zhang, L. Sun, H. Yu *et al.*, "A method for improving wavelet threshold denoising in laser-induced breakdown spectroscopy," Spectrochimica Acta Part B: Atomic Spectroscopy, 107, 32-44 (2015).

[6] M.-s. Niu, P.-g. Han, L.-k. Song *et al.*, "Comparison and application of wavelet transform and Kalman filtering for denoising in δ 13 CO 2 measurement by tunable diode laser absorption spectroscopy at 2.008 μm," Optics Express, 25(20), A896-A905 (2017).

[7] Y. Lei, D. Xu, Z. Zhou *et al.*, "A denoising algorithm for CT image using low-rank sparse coding." 10574.

[8] H. Xie, T. Niu, S. Tang *et al.*, "Content-oriented sparse representation (COSR) for CT denoising with preservation of texture and edge," Med Phys, 45(11), 4942-4954 (2018).

[9] S. Provencher, [LCModel & LCMgui User's Manual. LCModel Version 62–4], (2012).

[10] H. Hu, J. Bai, G. Xia *et al.*, "Improved baseline correction method based on polynomial fitting for Raman spectroscopy," Photonic Sensors, 8(4), 332-340 (2018).

[11] Y. Xi, and D. M. Rocke, "Baseline correction for NMR spectroscopic metabolomics data analysis," BMC bioinformatics, 9(1), 324 (2008).

[12] T. E. Stanford, C. J. Bagley, and P. J. Solomon, "Informed baseline subtraction of proteomic mass spectrometry data aided by a novel sliding window algorithm," Proteome science, 14(1), 19 (2016).

[13] G. Öz, J. R. Alger, P. B. Barker *et al.*, "Clinical proton MR spectroscopy in central nervous system disorders," Radiology, 270(3), 658-679 (2014).