Denoising Single Voxel Magnetic Resonance Spectroscopy with Deep Learning on Repeatedly Sampled In Vivo Data

Wanqi Hu, Dicheng Chen, Tianyu Qiu, Hao Chen, Xi Chen, Lin Yang, Gen Yan, Di Guo, Xiaobo Qu*

Abstract—Objective: Magnetic Resonance Spectroscopy (MRS) is a noninvasive tool to reveal metabolic information. One challenge of MRS is the relatively low Signal-Noise Ratio (SNR) due to low concentrations of metabolites. To improve the SNR, the most common approach is to average signals that are acquired in multiple times. The data acquisition time, however, is increased by multiple times accordingly, resulting in the scanned objects uncomfortable or even unbearable. Methods: By exploring the multiple sampled data, a deep learning denoising approach is proposed to learn a mapping from the low SNR signal to the high SNR one. Results: Results on simulated and in vivo data show that the proposed method significantly reduces the data acquisition time with slightly compromised metabolic accuracy. Conclusion: A deep learning denoising method was proposed to significantly shorten the time of data acquisition, while maintaining signal accuracy and reliability. Significance: Provide a solution of the fundamental low SNR problem in MRS with artificial intelligence.

Index Terms—Magnetic resonance spectroscopy, denoising, fast acquisition, deep learning, artificial intelligence.

I. INTRODUCTION

Magnetic Resonance Spectroscopy (MRS) is an examination method to determine molecular composition. As a non-invasive technique, it provides a quantitative analysis of metabolites in brains [1-4]. However, for the in vivo brain spectrum, the Signal-Noise Ratio (SNR) is relatively low due to the low concentrations of metabolites, leading to the difficulty in the further metabolic quantification and analysis [1, 5]. To gain the high SNR, MRS is commonly sampled repeatedly in many times and then averaged [6-8]. However, repeated sampling will lengthen the total acquisition time as it increases linearly with the number of repeats. For example, one time sampling of the spectrum on a 3T human scanner costs 14 seconds and the common 128 repeated sampling requires 4 minutes 28 seconds. Thus, how to reduce the times of repeated sampling while maintaining high SNR become important for fast MRS.

Denoising is a basic approach to achieve this goal and many methods have made great achievements in recent years [9-13]. Among these approaches, the Low Rank method is evidenced powerful to denoise MRS [9, 10, 14, 15] as its hypothesis, the time domain signal of MRS are exponential functions, fits the physics of magnetic resonance very well [9, 16]. The exponential assumption, sometimes, may deviate from real data due to field inhomogeneity or other imaging system imperfections [17].

Very recently, Deep Learning (DL), a representative artificial intelligence technique, has been introduced to spectra processing [18], such as sparse sampling reconstruction [19] and spectrum interpretation [20, 21]. These DL approaches require a huge database to learning the nonlinear mapping from input to output. By using the simulated basis set of metabolites [22] and incorporating the advanced signal models [23], DL has been introduced into the MRS denoising and quantification [22, 23].

Training data are commonly required to learn the neural network parameters in DL. However, in clinical, collecting enough data is very time-consuming and may encounter ethical issues if the scanning time is too long for patients. Therefore, these DL MRS approaches generally adopt the computer-simulated signal [19, 22, 23], such as exponential functions. Thus, some potential difference, such as physical and psychologic conditions, may exist between the simulated training set and in vivo data, which may hinder the applications of DL in MRS [22].

In this work, we propose a DL MRS approach that solely uses in vivo MRS as the training data. By learning the mapping from the low-SNR input and the high-SNR output, neural
networks are optimized and then used to denoise the target MRS. The contributions of this work include:
1) Build in vivo training data through deeply excavating and utilizing its repeated samplings from healthy volunteers.
2) Propose a causal Long Short-Term Memory (LSTM) neural network to suit for MRS denoising.
3) Reduce 80% data acquisition time with slightly compromised metabolic accuracy.

II. DATA PREPARATION

In this section, the construction of simulated and in vivo MRS data set will be introduced.

A. Simulated data set

Our simulation data is generated by the basis set. The basis set parameters used in this part are as follows: Point REsolved Spectroscopy (PRESS) sequence, 3T Field strength = 123.246 Hz/ppm, TE = 30 ms, number of points = 2048, delta time = 0.5 ms. To make it easier to see the signals, we selected 10 metabolites that are present in higher concentrations in healthy brains. The basis spectra of the 10 metabolites in the basis set are shown in the Appendix-A. The default concentration of each metabolite is 1.

Here we have 50 virtual people. The concentration range of each metabolite in the brain was determined according to the literature [22, 24]. For a virtual person, the concentrations of the 10 metabolites were randomly selected independently from the range in TABLE A1, and we assume that the concentration of metabolites is same in the repeated sampling. The simulated noise-free spectrum is obtained by multiplying each metabolite's concentration by the basis spectrum and adding the spectra of each metabolite (Fig. 1(a) and (b)).

For each virtual person, 128 Gaussian noises are generated independently for each noise-free spectrum, and the noise generates with the distribution of mean 0 and standard deviation 0.1. They were superimposed with noise-free spectra respectively to simulate 128 repeated samples (Fig. 1(c) and (d)).

Eventually, we build a simulated data set consists of 50 virtual spectra. Each spectrum contains 128 repeated samplings with independent noise.

B. In vivo data set

The in vivo data come from eight healthy volunteers (for training and test) and one patient with Parkinson disease (for test). The former data acquisition was approved by the Institutional Review Board of the Shanghai Jiao Tong University and the latter was approved by Medical Ethics Committee of The Second Affiliated Hospital of Xiamen Medical College. Before the study, the written informed consents from all people have been obtained.

All measurements of the healthy volunteer were conducted on the United Imaging 3T human scanner with the PRESS sequence [25]. The Free Induction Decay (FID) acquisition was applied to obtain 1H spectra with a nominal 90° RF excitation pulse and the following parameters: Field strength = 128.378 Hz/ppm, TR/TE = 2000/30 ms, spectral width = 1 kHz, number of points = 2048, delta time = 0.5 ms, voxel size = 20 mm×20 mm×20 mm. For each person and each position, it takes 4 minutes 21 seconds for 128 repeated samplings.

Specifically, we scanned three healthy volunteers at three voxel locations (LA, LB and LC in Figs. 2(a)-(c)). The rest five healthy volunteers only took the spectra at voxel position LA. Consequently, there are 14 different spectra and each of them had 128 repeated samplings (see Table I for more details).

The Parkinson patient data was collected by 3T GE scanner in the Second Affiliated Hospital of Xiamen Medical College. The FID acquisition was applied to obtain 1H spectra with a nominal 90° RF excitation pulse and the following parameters: Field strength = 127.764 Hz/ppm, TR/TE = 2000/30 ms, spectral width = 1 kHz, number of points = 4096, delta time = 0.2 ms, voxel size = 20 mm×20 mm×20 mm. For each person and each position, it takes 7 min 30 seconds for 128 repeated samplings. The voxel location is marked as LD in Fig. 2 (d).
the huge of input and output spectra are obtained by taking volunteer A. parameter settings in network. proposes the denoising neural network, and finally provides the denoise results, is called as virtual FID

\[ \mathbf{s} = [s_1, s_2, \ldots, s_T] \in \mathbb{C}^{1 \times T}. \]

Each point \( x_t \) maps to the corresponding point \( s_t \) through a Block composed of multi-layers (Fig. 4. (b)). The Block’s input has five parts: the current point \( x_t \), the Low SNR Past Frames (LSPF) \( \mathbf{x}_t \), and the Virtual Past Frames (VPF) \( \mathbf{s}_t \), memory cells \( \mathbf{h}_t \), and state cells \( \mathbf{c}_t \). The utilization of \( \mathbf{h}_t \) and \( \mathbf{c}_t \) are the same as the general LSTM [29].

LSPF is a vector consisting of \( r \) points before \( x_t \):

\[ \mathbf{x}_t = [x_{t-r}, x_{t-r+1}, x_{t-r+2}, \ldots, x_{t-2}, x_{t-1}]. \]  

(1)

VPF is the vector corresponding to LSPF in virtual FID. Its content comes from the result of each Block:

\[ \mathbf{s}_t = [s_{t-r}, s_{t-r+1}, s_{t-r+2}, \ldots, s_{t-2}, s_{t-1}]. \]  

(2)

The output of the CLSTM Block is the denoise result \( s_t \).

The parameters are marked as \( \theta \). The mapping of the entire Block can be expressed as:

\[ s_t = f_{\text{Block}}(x_t, \mathbf{x}_t, \mathbf{s}_t, \mathbf{h}_t, \mathbf{c}_t | \theta) \]  

(3)

The output of each step \( s_t \) constitutes the brand virtual FID \( s_{t+1} \). And all \( s_t \) make up the final noise reduction sequence \( \mathbf{s} \).

**III. METHODS**

This section first introduces how to build training set, then proposes the denoising neural network, and finally provides parameter settings in network.

**A. Dataset designs**

For the in vivo data, the network is trained on the MRS obtained from three locations (LA, LB, LC) of 7 healthy volunteers (light blue grid in Table I). In training phase, the total number of spectra in network output is 11.

For each spectrum, from \( M \) repeated samplings, we take \( m_1 \) samplings and then average them as an input spectrum (orange Input in Fig. 3). The output spectrum (green Label in Fig. 3) is obtained by taking \( m_2 \) samplings from \( M \) repeated samplings. Then, a pair of (Input, Label) forms one item. As the number of input and output spectra are \( C_m^m \) and \( C_m^{m_2} \), which are too huge to machine learning, only \( K \) items are randomly selected from all combinations of (Input, Label), i.e. \( k = 1, 2, \ldots, K \). For the training on in vivo data, we set \( m_1 = 24 \), \( m_2 = 124 \) and \( K = 1000 \).

The training process on simulated data is the same as that of in vivo data, and we set \( m_1 = 24 \), \( m_2 = 124 \) and \( K = 1000 \).
where the new next memory module is designed to use LSPF and VPF to denoise the input of LSTM cell $x_t$ and $x'_t$, and $\theta_r$ represents the training parameters of the regression layers $\mathcal{M}$. Each Block contains four modules (See Fig. 4 (b)):

**Input Module:** Combines the LSPF and VPF through several linear layers and ReLU function. The output $x'_t$ is regarded as the input of the LSTM cell:

$$x'_t = f_s((x_t, s) | \theta')$$

(4)

where $\theta'$ represents the training parameters of the input module.

**LSTM cell:** Employ the classic LSTM strategy to calculate the new next memory cells $h_{t+1}$ and states cells $c_{t+1}$ as below:

$$(h_{t+1}, c_{t+1}) = f_{lstm}((x'_t, h_t, c_t) | \theta^h).$$

(5)

where $\theta^h$ represents the training parameters of the LSTM.

**Regression module:** Input the new memory cells $h_{t+1}$ to this model, and regression layers map $h_{t+1}$ to a point $\hat{x}_t$ as:

$$\hat{x}_t = f_r(h_{t+1} | \theta^r),$$

(6)

where $\hat{x}_t$ is regarded as the final experiment model. In addition, the network is scheduled to decay iteratively through back-propagation. Next, the signal is normalized in the time domain for all the tasks to enhance stability and robustness. The learning rate is scheduled to decay according to the validation loss by a factor of 0.1. Finally, the best checkpoint model with the smallest validation loss is saved as the final experiment model. In addition, the network is trained using one NVIDIA RTX2080Ti GPU for 50 epochs. During testing, the typical inference time for 512 sampled points is around 1 second in each batch.

### C. Network Implementation

The network is implemented using the PyTorch [30] library. Model parameters $\theta = \{\theta^s, \theta^h, \theta^r, \theta^P\}$ are trained by an Adam optimizer [31] with learning rate $1e-4$ and batch size 128 to minimize the Mean Absolute Error loss function at every sampling points. Then, the network parameters are updated iteratively through back-propagation. Next, the signal is normalized in the time domain for all the tasks to enhance stability and robustness. The learning rate is scheduled to decay according to the validation loss by a factor of 0.1. Finally, the best checkpoint model with the smallest validation loss is saved as the final experiment model. In addition, the network is trained using one NVIDIA RTX2080Ti GPU for 50 epochs. During testing, the typical inference time for 512 sampled points is around 1 second in each batch.

### IV. RESULTS

This section is devoted to verify the denoising performance. For comparisons, we choose Low Rank method [15], one of the state-of-the-art techniques in MRS denoising, and the most common approach that averages maximal times of repeated samplings.

The test set contains two parts: (a) 10 simulated spectra, and (b) 3 in vivo spectra, which have been introduced in Section III-A. In the evaluation, input signals are treated as low-SNR observations and spectra averaged by 128 repeated samplings are regarded as high-SNR references. All spectra, including observations, references and denoising results, were analyzed with LCModel, a commercial software whose quantitative results are widely used in MRS applications [11, 22]. We evaluate spectra from two aspects: Concentration (Conc.) and Standard Deviations Percentage (SDP). The Conc. is the absolute concentration with water scaling. The SDP is the estimated standard deviations (Cramer-Rao lower bounds) expressed in percentage of estimated concentrations [32]. The SDP ranges from 0 to 999 and a SDP $< 15$ has been used as a criterion of acceptable reliability [32, 33]. For each trial, the lower the SDP is, the more accurate the metabolite estimate is.

**TABLE II** Parameters of Network Layers

| Model            | Layers | Input size, output size |
|------------------|--------|-------------------------|
| Input model      | Linear $(r \times 2, r)$ + Linear $(r \times 32, 1)$ | Linear $(r \times 32, r \times 128)$ |
| LSTM cell model  | LSTM cell $(r \times 128, 1024)$ | Linear $(1024, 1)$ |
| Regression model | Linear | Linear $(2, 4) \times 1 + \text{ReLU} +$ Linear $(4, 1)$ |

Note: The layer description follows the standard in PyTorch 1.4. The size $r$ of LSPF and VPF is 10.

**Data validation:** Do the weighted $\theta'$ sum for fidelity term of $x_t$ and $x'_t$ as:

$$s_t = f_r(x_t, s_t | \theta')$$

(7)

where $s_t$ is the denoised value that updates to the virtual FID. Details of each model can be found in Fig. 4 (c).

### IV. RESULTS

This section is devoted to verify the denoising performance. For comparisons, we choose Low Rank method [15], one of the state-of-the-art techniques in MRS denoising, and the most common approach that averages maximal times of repeated samplings.

The test set contains two parts: (a) 10 simulated spectra, and (b) 3 in vivo spectra, which have been introduced in Section III-A. In the evaluation, input signals are treated as low-SNR observations and spectra averaged by 128 repeated samplings are regarded as high-SNR references.

All spectra, including observations, references and denoising results, were analyzed with LCModel, a commercial software whose quantitative results are widely used in MRS applications [11, 22]. We evaluate spectra from two aspects: Concentration (Conc.) and Standard Deviations Percentage (SDP). The Conc. is the absolute concentration with water scaling. The SDP is the estimated standard deviations (Cramer-Rao lower bounds) expressed in percentage of estimated concentrations [32]. The SDP ranges from 0 to 999 and a SDP $< 15$ has been used as a criterion of acceptable reliability [32, 33]. For each trial, the lower the SDP is, the more accurate the metabolite estimate is.

0 is the best while 999 corresponds to an occasion where some
metabolite cannot be detected or do not exist. Seven metabolites whose SDP are smaller than 15 in references are presented because these metabolites, such as NAA, Cr+PCr, and Glu, are treated as important biomedical indicators in the diagnosis \cite{24, 34}. To avoid the bias of random selection, we repeated 50 trials for each spectrum and counted the average and standard deviation.

The denoised results on simulated data are given in Fig. 5, showing that CLSTM outperforms the compared Low Rank method, no matter on the accuracy of concentration or reliability SDP. Compared with the input signal (average of 24 repeated samplings), although both Low Rank and CLSTM smoothen the spectra, CLSTM still preserves more details (Fig. 5(d)) than Low Rank (Fig. 5(c)). In quantification analysis, both methods provide reasonable estimates for most metabolic concentrations. However, Low Rank reduces concentrations of some metabolites, such as Asp and Glu (Fig. 5(e)). Moreover, for some metabolites which are difficult to be detected because of the low concentration, such as Asp, Low Rank fails to offer an acceptable reliability SDP. On the contrary, CLSTM is still capable of preserving this component during the denoising, even owns smaller SDP than that of References. Thus, by we taking 24 repeated samplings from 128 samplings, the proposed CLSTM achieves comparable concentration accuracy and reliability as the maximal times of averages, indicating saving 81.25% in the data acquisition time.

The denoising performance of in vivo $^1$H spectrum confirms our conclusion in the simulated data. Both Low Rank and CLSTM suppress the noise well (Figs. 6(c) and (d)). However, Low Rank over-smooths some signals (e. g. NAA at 2.7–2.8 ppm).
ppm), leading to lower concentration estimation than the Reference (Fig. 6(c)). By contrast, results of CLSTM are more comparable with Reference.

We further verify the performance on Parkinson patient data. As shown in Fig. 7, for Inputs with very low SNR, CLSTM (Fig. 7(d)) removes noise better than Low Rank (Fig. 7(c)). For quantification analysis, CLSTM provides more comparable estimate with References, such as GPC and Glu, and reduces SDPs compared with Inputs (24 averages). Furthermore, for some metabolite whose SDPs are with a large bar, CLSTM obviously decreases the average and the standard deviation of SDPs, indicating that CLSTM offers much more reliable and robust estimates than the inputs and Low Rank method. These observations imply that, although the network was trained on the data acquired from healthy volunteers, the proposed CLSTM still applicable to improve the patient MRS.

In summary, results on the simulated and in vivo data demonstrate that the proposed CLSTM can reduce the sampling times from 128 to 24, saving 81.25% data acquisition time, with comparable or slightly compromised metabolic concentrations.

V. DISCUSSIONS

In this section, we will test the denoising performance of CLSTM under different numbers of repeated samplings (Fig. 8) and different voxels (Fig. 9) to verifying the generality of method.

A. Different Number of Samplings

For Inputs with different sampling times in Fig. 8, CLSTM provides comparable estimates with Reference. For the high-concentration metabolites (Glu and NAA), CLSTM offers an estimate which is very close to Reference. For the

![Fig. 7. Denoised spectra and the quantitative analysis of the Parkinson patient (P9-LD). (a)-(d) are reference (128 averages), low-SNR input (24 averages), denoised spectra with low rank and the proposed CLSTM, respectively. (e) Estimated concentrations (above the zero line) and the reliability criteria, SDPs (under the zero line), of the metabolites. Note: In (a)-(d), the fit curves of the corresponding spectra (red, yellow, blue and green color lines) are presented with grey lines. In (e), only metabolites with the lowest seven SDPs are presented (Cr and PCr is shown as the sum Cr+PCr of the two metabolites). -CrCH2 is a correction term, rather than a metabolite. It is a simulated negative creatine CH2 singlet around 3.94 ppm), and the error bars denote standard deviations under 50 denoising trials. The variance of GPC’s SDP is far over the scope which is marked with triangle. The dashed solid line is SDP=15 and above this line implies acceptable reliability of estimation.

![Fig. 8. The quantitative analysis of the volunteer (P7-LA) with different number of samplings as input. (a) 16 average (b) 32 average. Both (a) and (b) shows the estimated concentrations (above the zero line) and the reliability criteria, SDPs (under the zero line), of the metabolites. Note: only metabolites with the lowest seven SDPs are presented (Cr and PCr is shown as the sum Cr+PCr of the two metabolites), and the error bars denote standard deviations under 50 denoising trials. The dashed solid line is SDP=15 and above this line implies acceptable reliability of estimation.}
low-concentration metabolites (GPC, GSH, and Asp), besides the accurate concentrations, CLSTM obviously reduces the SDPs compared with Inputs. Thus, the benefits of CLSTM on accurate concentrations and high reliability still hold.

B. Different Locations of Voxels

Voxel locations do not limit the benefit of CLSTM. In other locations (Fig. 9), CLSTM still offers reasonably estimated concentration compared with References. Besides, the averaged SDPs obtained with CLSTM are smaller than 15 for most metabolites, ensuring that concentration estimates are reliable enough.

C. Limitations on Accuracy of Low Concentration Metabolites

However, for some metabolites, CLSTM may provide an estimate with relatively insufficient accuracy (accuracy of Conc. >5%). For example, on the data acquired from healthy volunteers, the relative error of one the estimated concentration of Asp is 5.4%. It speculates that this may be because the metabolite has more peaks and lower peak height, and its concentration is greatly affected by noise.

VI. CONCLUSION

In this work, we proposed a deep learning denoising method, CLSTM, with limited measured data. The denoising result is comparable with the 128 averaged spectra, meaning that the proposed method saves about 81.25% acquisition time. In addition, the denoising result achieves the comparable quantitative accuracy and greatly reduces the SDPs, making the quantification more reliable. CLSTM may be applied to the MRS denoising on brain diseases.

APPENDIX

A. Simulated Dataset Detail

Fig. 10 is the basis spectra of the metabolites in the basis set shown at 4.0-0.2 ppm.

TABLE A1 is the reference range of the concentration of metabolites in healthy human brain, which is used in the of the simulated spectrum in this work.

B. Examples of Limitations of Traditional LSTM

Traditional LSTM encounters problems to preserve small peaks when the training set (Fig. 11(a)) differs from the target spectrum (Fig. 11(b)).

The results of Fig. 11 are trained by LSTM with 9000 simulated spectra with different concentration of NAAG or NAA. Hidden size of LSTM is 128. The weights are both trained by 100 epochs. Comparing NAAG with NAA, when the number of metabolite peaks and the height difference of the spectral peak increased, the LSTM for denoising got worse.

TABLE A1

| Metabolite (abbreviations) | Reference Range (mmol/L) | Concentration in Fig.1 (mmol/L) |
|----------------------------|--------------------------|--------------------------------|
|                            | Low          | High        | (a)     | (b)      |
| Aspartate (Asp)            | 1.00         | 1.40        | 1.15    | 1.38     |
| Creatine (Cr)              | 5.10         | 10.60       | 8.12    | 5.65     |
| Glutamine (Gln)            | 3.00         | 5.80        | 3.96    | 3.64     |
| Glutamate (Glu)            | 6.00         | 12.50       | 8.11    | 10.53    |
| Glutathione (GSH)          | 2.00         | 2.00        | 2.00    | 2.00     |
| Glycerophosphoryl-choline   | 1.00         | 1.00        | 1.00    | 1.00     |
| (GPC)                      |              |             |         |          |
| Myo-inositol (mI)          | 3.80         | 8.10        | 6.01    | 5.80     |
| N-acetyl-aspartate (NAA)   | 7.90         | 16.60       | 12.25   | 11.59    |
| N-acetyl-aspartyl-glutamate| 0.60         | 2.70        | 1.27    | 1.56     |
| (NAAG)                     |              |             |         |          |
| Phosphocreatine (PCr)      | 3.20         | 5.50        | 4.68    | 3.49     |

C. Accuracy of Concentration

TABLE A2 showed the accuracy of the metabolites' concentration relative to the reference spectrum.
Fig. 10. Baseline spectra of 10 metabolites in basis set. The default concentration of each metabolite is 1.

Fig. 11. The denoising result of the NAAG (a) and NAA (b) simulated spectra and the corresponding FID.

Table A2
Accuracy of Concentration

| Figure | /Metabolites | GPC | GSH | Asp | ml | Cr+PCr | Glu | NAA | Average absolute value |
|--------|--------------|-----|-----|-----|----|--------|-----|-----|------------------------|
|        | Method       |     |     |     |    |        |     |     |                        |
| Fig. 5 | Input        | 1.16% | 0   | 7.63% | 0 | 0.98% | -0.85% | 1.01% | 1.66% |
|        | Low Rank     | -5.38% | -8.61% | -62.37% | -1.95% | -1.03% | -8.02% | 0.45% | 12.54% |
|        | CLSTM        | 2.18% | 0 | -11.45% | -0.87% | -0.085% | 0.17% | 1.57% | 2.33% |
| Fig. 6 | Input        | 6.64% | 6.36% | -8.82% | -6.97% | 1.60% | 6.29% | 0.57% | 5.32% |
|        | Low Rank     | 18.49% | -9.59% | -9.31% | -3.28% | 0.82% | -9.33% | -7.44% | 8.32% |
|        | CLSTM        | -4.31% | -2.50% | 5.39% | 0.82% | 1.30% | -3.45% | -1.72% | 2.64% |
| Fig. 8(a) | Input    | -1.20% | 8.73% | 0.76% | 0 | -4.62% | -0.79% | 0% | 2.93% | 2.72% |
|        | Low Rank    | -16.59% | -12.14% | -34.20% | -4.87% | -0.165% | -17.26% | -3.61% | 12.69% |
|        | CLSTM       | -6.55% | -3.27% | -9.92% | -1.54% | -0.55% | -5.44% | -1.72% | 2.64% |
| Fig. 8(b) | Input    | 4.15% | 4.49% | 0 | -6.72% | 0.26% | 2.95% | 3.42% | 3.14% |
|        | Low Rank    | -12.44% | -11.24% | -32.85% | -0.56% | -0.84% | -13.15% | -3.18% | 10.61% |
|        | CLSTM       | -4.95% | -4.83% | -7.52% | -0.28% | -0.47% | -5.44% | -1.71% | 3.60% |
| Fig. 9(a) | Input    | 13.29% | 6.36% | -13.24% | -11.89% | 1.86% | 8.32% | 1.15% | 8.02% |
|        | Low Rank    | -25.49% | -15.75% | -21.57% | -4.51% | 0.19% | -13.19% | -2.10% | 11.99% |
|        | CLSTM       | -1.47% | -5.08% | 3.86% | 2.69% | 1.00% | -3.48% | 0.29% | 2.55% |
| Fig. 9(b) | Input    | 11.67% | 6.36% | -5.88% | -4.51% | 0.98% | 6.09% | 0.19% | 5.10% |
|        | Low Rank    | -19.39% | -10.74% | -12.25% | -2.46% | 0.61% | -11.16% | -16.22% | 10.40% |
|        | CLSTM       | 1.20% | -1.21% | 0.66% | 3.56% | 0.48% | -2.11% | -1.68% | 1.55% |

Note: Accuracy = (Conc. (Method) - Conc. (Reference)) / Conc. (Reference) ×100%;
Average absolute value is the average of the absolute accuracy of each metabolite.
ACKNOWLEDGEMENTS

The authors would like to thank Zi Wang and Xinlin Zhang for assisting in data acquisition and helpful discussions. The authors also thank the staff from Shanghai Jiao Tong University and the Second Affiliated Hospital of Xiamen Medical College for technical support.

REFERENCES

[1] J.-B. Poulet, D.M. Sima, and S. Van Huffel, "MRS signal quantitation: a review of biological, biochemical, and frequency-domain methods," J. Magn. Reson., vol. 195, no. 2, pp. 134-144, 2008.

[2] M. Bagory, F. Durand-Dubief, D. Ibarrola, J.-C. Comte, F. Cotton, C. Confavreux, and D. Sappey-Marinier, "Implementation of an absolute brain '1H-MRS quantification method to assess different tissue alterations in multiple sclerosis," IEEE Trans. Biomed. Eng., vol. 59, no. 10, pp. 2687-2694, 2011.

[3] S. Feddewa, and V. Abolghasemi, "Semiblind spectral factorization approach for magnetic resonance spectroscopy quantification," IEEE Trans. Biomed. Eng., vol. 65, no. 8, pp. 1717-1724, 2017.

[4] Y. Guo, S. Ruan, J. Landré, and J.M. Constans, "A sparse representation method for magnetic resonance spectroscopy quantification," IEEE Trans. Biomed. Eng., vol. 57, no. 7, pp. 1620-1627, 2010.

[5] S.W. Provencher, "Estimation of metabolite concentrations from localized in vivo proton NMR spectra," Magn. Reson. Med., vol. 30, no. 6, pp. 672-679, 1993.

[6] S.K. Valapara, E.M. Ripley, G.R. Boone, T.Q. Duong, and G.D. Clarke, "Evaluation of vastus lateralis muscle fat fraction measured by two-point Dixon water-fat Imaging and 1H-MRS," in Proc. Intl. Soc. Mag. Reson. Med. 2014, vol. 22, p. 1221.

[7] Y. Liu, X. Mei, J. Li, N. Lai, and X. Yu, "Mitochondrial function assessed by 31P MRS and BOLD MRI in non-obese type 2 diabetic rats," Physiological Reports, vol. 6, no. 15, p. e1290, 2016.

[8] C. Choi, S.K. Ganji, R.J. DeBerardinis, I.E. Dimitrov, J.M. Pascual, R. Bachoo, B.E. Mickey, C.R. Malloy, and E.A. Maher, "Measurement of glycohe in the human brain in vivo by 1H - MRS at 3 T: application in brain tumors," Magn. Reson. Med., vol. 66, no. 3, pp. 609-618, 2011.

[9] H.M. Nguyen, X. Peng, M.N. Do, and Z.-P. Liang, "Denoising MR spectroscopic imaging data with low-rank approximations," IEEE Trans. Biomed. Eng., vol. 60, no. 1, pp. 78-89, 2012.

[10] Y. Chen, Y. Li, and Z. Xu, "Improved low-rank filtering of MR spectroscopic imaging data with pre-learned subspace and spatial constraints," IEEE Trans. Biomed. Eng., vol. 67, no. 8, pp. 2381-2388, 2019.

[11] S.H. Joshi, A. Marquina, S. Njau, K.L. Narr, and R.P. Woods, "Denoising of MR spectroscopy signals using total variation and iterative Gauss-Seidel gradient updates," in 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), 2015: IEEE, pp. 576-579.

[12] H.F.C. de Greiff, R. Ramos-Garcia, and J.V. Lorenzo-Ginori, "Signal de-noising in magnetic resonance spectroscopy using wavelet transforms," Concepts Magn. Reson., vol. 14, no. 6, pp. 388-401, 2002.

[13] O.A. Ahmed, "New denoising scheme for magnetic resonance spectroscopy signals," IEEE Trans. Med. Imaging, vol. 24, no. 6, pp. 660-670, 2005.

[14] Y. Liu, C. Ma, B.A. Clifford, F. Lam, C.L. Johnson, and Z.-P. Liang, "Improved low-rank filtering of magnetic resonance spectroscopic imaging data corrupted by noise and B0 field inhomogeneity," IEEE Trans. Biomed. Eng., vol. 63, no. 4, pp. 841-849, 2015.

[15] T. Qiu, W. Liao, D. Guo, D. Liu, X. Wang, and X. Qu, "Gaussian noise removal with exponential functions and spectral norm of weighted Hankel matrices," arXiv preprint arXiv:2001.11815, 2020.

[16] X. Qu, M. Mayzel, J.F. Cai, Z. Chen, and V. Orekhov, "Accelerated NMR spectroscopy with low-rank reconstruction," Angew. Chem. Int. Ed., vol. 54, no. 3, pp. 852-854, 2015.

[17] I. Marshall, J. Higinbotham, S. Bruce, and A. Freise, "Use of Voigt lineshape for quantification of in vivo 'H spectra," Magn. Reson. Med., vol. 37, no. 5, pp. 651-657, 1997.

[18] D. Chen, Z. Wang, D. Guo, V. Orekhov, and X. Qu, "Review and prospect: deep learning in nuclear magnetic resonance spectroscopy," Chem.-A Eur. J., vol. 26, no. 46, pp. 10391-10401, 2020.

[19] X. Qu, Y. Huang, H. Lu, T. Qiu, D. Guo, T. Agback, V. Orekhov, and Z. Chen, "Accelerated nuclear magnetic resonance spectroscopy with deep learning," Angew. Chem. Int. Ed., vol. 59, no. 26, pp. 10297-10300, 2019.

[20] P. Klukowski, M. Augoff, M. Dietz, M. Drwal, A. Gonczarek, and M.J. Walczak, "NMRNet: a deep learning approach to automated peak picking of protein NMR spectra," Bioinformatics, vol. 34, no. 15, pp. 2590-2597, 2018.

[21] S. Liu, J. Li, K.C. Bennett, B. Gao, T. Staub, M. Head-Gordon, A. Hexemer, D. Ushizima, and T. Head-Gordon, "Multiresolution 3D-DenseNet for chemical shift prediction in NMR crystallography," J. Phys. Chem. Lett., vol. 10, no. 16, pp. 4558-4565, 2019.

[22] Y. H. Lee and H. Kim, "Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain," Magn. Reson. Med., vol. 82, no. 1, pp. 33-48, 2019.

[23] F. Lam, Y. Li, and X. Peng, "Constrained magnetic resonance spectroscopic imaging by learning nonlinear low-dimensional models," IEEE Trans. Med. Imaging, vol. 39, no. 3, pp. 546-559, 2019.

[24] V. Govindaraju, K. Young, and A.A. Maudsley, "Proton NMR chemical shifts and coupling constants for brain metabolites," NMR Biomed., vol. 13, no. 3, pp. 129-153, 2000.

[25] P.A. Bottomley, "Selective volume method for performing localized NMR spectroscopy," ed: Google Patents, 1984.

[26] J. Xie and Q. Wang, "Benchmarking machine learning algorithms on blood glucose prediction for type 1 diabetes in comparison with classical time-series models," IEEE Trans. Biomed. Eng., vol. 67, no. 11, pp. 3101-3124, 2020.

[27] M. Wang, C. Lian, D. Yao, D. Zhang, M. Liu, and D. Shen, "Spatial-temporal dependency modeling and network hub detection for functional MRI analysis via convolutional-recurrent network," IEEE Trans. Biomed. Eng., vol. 67, no. 8, pp. 2241-2252, 2019.

[28] Y. Wang, K. Lin, Y. Qi, Q. Lian, S. Feng, Z. Wu, and G. Pan, "Estimating brain connectivity with varying-length time lags using a recurrent neural network," IEEE Trans. Biomed. Eng., vol. 65, no. 9, pp. 1953-1963, 2018.

[29] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997.

[30] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[31] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[32] S.W. Provencher, "Automatic quantitation of localized in vivo '1H spectra with LCMModel," NMR Biomed., vol. 14, no. 4, pp. 260-264, 2001.

[33] K.S. Opstad, B.A. Bell, J.R. Griffiths, and F.A. How, "Toward accurate quantification of metabolites, lipids, and macromolecules in HRMAS spectra of human brain tumor biopsies using LCMModel," Magn. Reson. Med., vol. 60, no. 5, pp. 1237-1242, 2008.

[34] G. Öz, J.R. Alger, P.B. Barker, R. Bartha, A. Bizzi, C. Boesch, P.J. Bolan, K.M. Brindle, C. Cudalbu, and A. Dinçer, "Clinical proton MR spectroscopy in central nervous system disorders," Radiology, vol. 270, no. 3, pp. 658-679, 2014.