ISSN (Print) 1226-1750
ISSN (Online) 2233-6656
Journal of Magnetics 26(3), 356-362 (2021) https://doi.org/10.4283/JMAG.2021.26.3.356

Extraction of Individual Metabolite Spectrum in Proton Magnetic Resonance Spectroscopy of Mouse Brain Using Deep Learning

Yoon Ho Hwang[1†], Woo-Seung Kim[1†], Chang-Soo Yun[2], Jae-Hyung Yeon[2], Hyeon-Man Baek[3], Bong Soo Han[2], and Dong Youn Kim[1]*

[1]*Deparment of Biomedical Engineering, Yonsei University, Wonju 26493, Republic of Korea*
[2]*Deparment of Radiation Convergence Engineering, Yonsei University, Wonju 26493, Republic of Korea*
[3]*Department of Basic Medical Sciences, Lee Gil Ya Cancer & Diabetes Institute, Gachon University, Incheon 21999, Republic of Korea*

(Received 8 July 2021, Received in final form 28 September 2021, Accepted 28 September 2021)

The present study aims to develop a deep learning (DL) model to quantify metabolites. To apply DL to metabolite quantification using ^{1}H-MRS data, Convolutional autoencoder (CAE) were designed to extract line-narrowed, baseline-removed, and noise-free metabolite spectra for each metabolite. Fifty thousand simulation data were generated by varying the SNR (4-12), linewidth (6-22 Hz), phase shift (± 5°), and frequency shift (± 5 Hz) on phantom spectra. The data were divided into 45,000 simulation data for training and 5,000 test data, and the mean absolute percent errors (MAPEs) were used to evaluate the performance of the CAE. The average MAPE of the metabolites was 13.64 ± 11.38 %. Fourteen metabolites were within the reported concentration ranges. These findings showed that the proposed method had similar or improved performance than conventional methods. The proposed method using DL was the recent and up-to-date quantification one and has clinically potential applicability.

Keywords : Proton magnetic resonance spectroscopy (^{1}H-MRS), high magnetic field (9.4T), MR spectrum, metabolite quantification, deep learning, convolutional autoencoder

1. Introduction

Proton magnetic resonance spectroscopy (^{1}H-MRS) as a nondestructive and noninvasive tool based on magnetic field provides the biochemical information extracted from a desired region of an organism [1]. ^{1}H-MRS in the field of magnetic resonance imaging (MRI) is a value-added modality in preclinical and clinical studies of disease states [2]. While other MRI studies only offer the structural details, that is, morphological findings such as cysts, tumors, swelling, etc., ^{1}H-MRS provides functional information on the metabolic status of the brain by quantifying the metabolites and determining their types from a spectrum, which is a ^{1}H-MRS data [3, 4]. Therefore, it is important to quantify the metabolites precisely. However, despite tremendous efforts and numerous publications on this topic [5-7], accurate quantification of metabolites is still difficult due to the strongly overlapping metabolite peaks, relatively low signal-to-noise ratio (SNR), and a broad background originating primarily from macromolecules and lipids. Since metabolic changes in the brain related to research on various diseases, including psychiatric disorders, are analyzed based on the concentration of metabolites, a precise quantitative method is needed to obtain reliable results [8].

Deep learning (DL), a technique that attempts to model high-level abstractions in data with multiple processing layers, has achieved popularity in recent years with the availability of powerful GPUs [9, 10]. It consists of a model with several hidden layers of artificial neurons using the underlying structure of the input data [11]. Starting with the extraction of lower-level features from the initial layers and proceeding to the subsequent layers, higher-level features are extracted based on the features of the previous layers [12]. This feature extraction using multiple layers allows the network to model complex variations of input data as the weighted sum of all features in each layer and make accurate predictions [13]. In addition, DL has been successfully used in MRI and MRS

†These authors contributed equally to this work
*Corresponding author: Tel: +82-33-760-2434
Fax: +82-33-760-2852, e-mail: dongkim77@yonsei.ac.kr

for classification, object recognition, noise removal, and prediction. Kyathanahally *et al*. reported the application of CNNs for the detection and removal of spectroscopic artifacts [14]. Chandler *et al.* quantified the concentration of γ-aminobutyric acid using DL [15]. Hatami *et al*. built a CNN that was trained on a time-domain dataset and was capable of quantifying individual metabolites in a simulated spectrum [16]. Lee *et al*. trained a CNN to predict the entire metabolite spectrum of the human brain and then quantified the metabolites using a simple inverse problem [17]. The CAE is a DL model specialized for noise reduction that extracts noise-free, line-narrowed, baseline-removed, phase- and frequency-corrected metabolite spectra from the spectrum of the mouse brain.

In the present study, a convolutional autoencoder (CAE) model for each of the metabolites was proposed for accurate precise quantification of metabolites by reflecting their individual characteristics. Twenty-one metabolite spectrum extraction algorithms were developed by training DL models for each metabolite.

2. Materials and Methods

2.1. MRS Data Acquisition

The metabolite phantom spectra and *in vivo* spectra of mice were acquired with a 9.4 T Bruker MRI/MRS system. For simulation data, phantom spectra were acquired using point resolved spectroscopy (PRESS) as the following parameter: repetition time (TR) = 10,000 ms, echo time (TE) = 10 ms, number of signal averages = 128, bandwidth = 5 kHz, number of data points = 4,096) for twenty-one metabolites that were used as the experimental basis set. Fig. 1 shows the following spectra for metabolites: alanine (Ala), ascorbic acid (Asc), aspartate (Asp), choline (Cho), creatine (Cr), γ-aminobutyric acid (GABA), glucose (Glc), glutamine (Gln), glutamate (Glu), glycine (Glyc), glutathione (GSH), lactate (Lac), *myo*-Inositol (*m*-Ins), *scyllo*-Inositol (*s*-Ins), glycerophosphorylcholine (GPC), taurine (Tau), N-acetylaspartate (NAA), N-acetylaspartylglutamate (NAAG), phosphocreatine (PCr), phosphorylcholine (PCh), phosphorylethanolamine (PE). Twelve C57BL/6N mice were used for the *in vivo* MRS data (male, six-week-old, 18-25 g). The data were acquired from the brain of the mice using PRESS (TR = 4,000 ms, TE = 10 ms, number of signal averages = 512, bandwidth = 5 kHz, number of data points = 4,096). A voxel (1.8 × 3.4 × 1.8 mm^3) was placed on the left hippocampus.

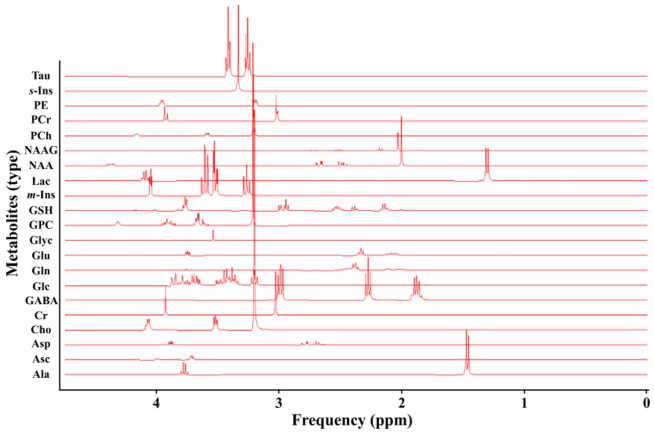

Fig. 1. (Color online) Neurochemical profiles of 21 metabolites in the experimental basis set made of phantom-based ^{1}H-MRS data acquired with a 9.4 T Bruker MRI/MRS system. Each phantom represents the metabolites.

2.2. MRS simulation data

To simulate the spectrum of the mouse brain, simulation data were generated as follows. This study basically referenced Lee's simulation method [17]: 1) Each simulation data set consists of a linear combination of twenty-one metabolite spectra. The ranges for the concentrations of each metabolite were confirmed by referring to previous studies [1, 18-21], 2) Line broadening was applied to the spectrum at 6-22 Hz and zero-order phase distortion was applied in the range of ± 5°, 3) Frequency shift was applied to the spectrum in the range of ± 5 Hz and the spectra of metabolites were merged into a spectrum, 4) The simulated spectrum was combined with a macromolecular baseline spectrum that was extracted from the metabolite-nulled spectra of the mice using jMRUI, 5) The SNR of total NAA (tNAA) was adjusted by adding random noise in the range of 4-12. Finally, 50,000 simulated spectra of the mouse brain were obtained and randomly assigned to the training data (n = 45,000) and the test data (n = 5,000) to minimize overfitting with the training data,.

2.3. CAE architecture

The CAE was designed and optimized using Keras (v.2.2.4; TensorFlow backend) on graphics processing units (GPU; GeForce GTX 980 Ti). The CAE is a type of convolutional neural network. The convolution operator allows filtering of the input signals from among nearby regions to extract some parts of its content. An autoencoder is an algorithm that extracts a noise-free signal from an input signal. The CAE consists of an input layer, convolution blocks, max-pooling layer, upsampling layer, flattening layer, dropout layer, and dense layer, as shown in Fig. 2. The data input to the network has 4,096 data points, which is the size of each simulation data. The Conv1D layer creates a convolution kernel that is con-

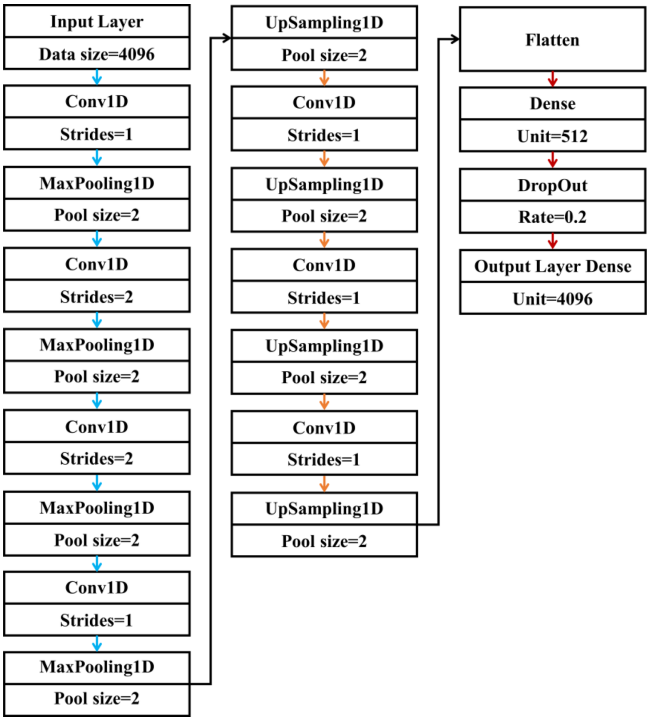

Fig. 2. (Color online) Architecture of a convolutional autoencoder (CAE) proposed in this study for metabolite quantification. The CAE is trained to predict the spectrum of each metabolite from *in vivo* ^{1}H-MRS data in the brain.

volved with the layer input over a single spatial (or temporal) dimension to produce a tensor of outputs. The convolutional filter sizes are 1 × 16, 1 × 32, 1 × 64, 1 × 128, 1 × 64, 1 × 32, and 1 × 16, respectively. The kernel size is 15-20 with experimental low mean squared error (MSE) for the model of each metabolite. The MaxPooling1D layer downsamples the input representation by taking the maximum value over the window defined by pool size. UpSampling1D layer repeats each pool size along the axis. The output data are noise-free, line-narrowed, baseline-removed, and phase and frequency corrected metabolite spectra of size 1 × 4,096 that are to be predicted. The minimum loss function was attained in the shortest time when using four pairs of convolutions and maxpooling in the encoder part and four pairs of upsampling and convolution in the decoder part. The pool size was set to 2 and the stride was set to either 1 or 2 through trial and error. A rectified linear unit (ReLU) activation function was used in the activation layer of Conv1D. An ADAM weight update was used to train the CAE. The CAE were trained with early stopping technique (monitor = "val custom metric", patience = 30, batch size = 256). The loss function was the MSE. In addition, the Keras custom metric was designed. This is the difference between the relative amplitude of the ground truth and the predicted metabolite spectrum. The relative amplitude indicates the scale factor relative to the basis set spectrum. Model learning was conducted in a manner such that the metric value was minimized. The package is designed for Python 3.6 and TensorFlow 1.12.

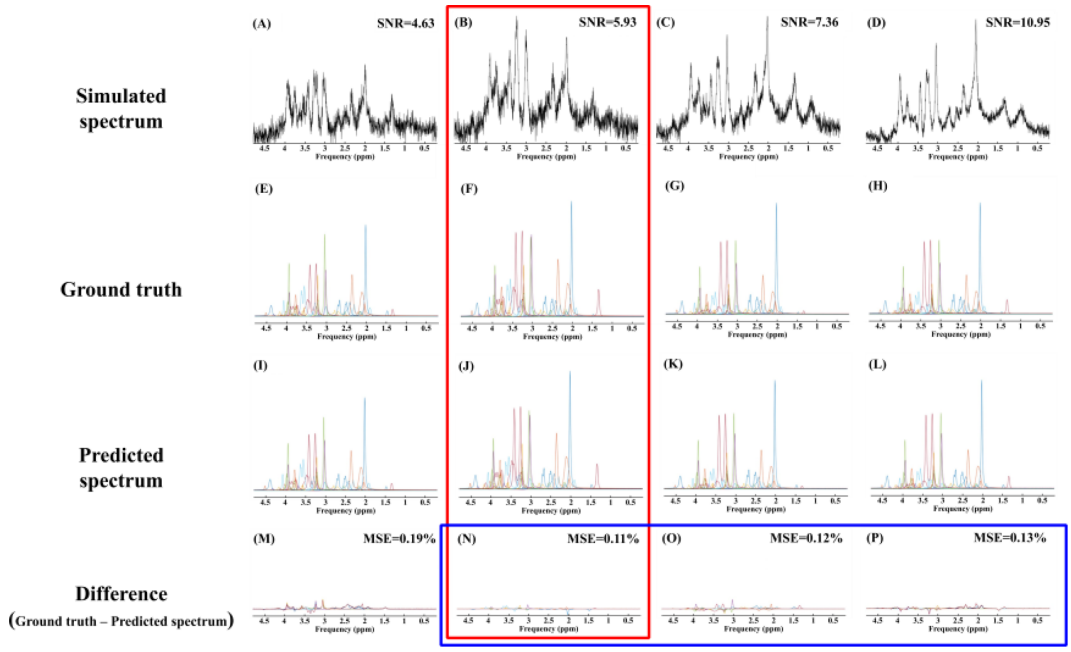

Fig. 3. (Color online) Representative simulated spectra in the test set (A-D), ground truth metabolite spectra (E-H), corresponding CAE-predicted metabolite spectra (I-L), and difference spectra obtained by subtracting the CAE-predicted spectra from the ground truth spectra (M-P).

MAPE	Ala	Asc	Asp	Cho	Cr	GABA	Glc	Gln	Glu	Glyc	GPC	GSH	*m*-Ins
Mean (SD)	59.10 (67.77)	10.95 (7.02)	17.33 (10.92)	22.06 (15.87)	7.62 (6.37)	10.43 (6.92)	9.72 (7.79)	14.02 (9.63)	4.44 (3.58)	31.14 (25.59)	7.58 (5.41)	8.57 (6.65)	7.23 (5.92)
MAPE	Lac	NAA	NAAG	PCh	PCr	PE	*s*-Ins	Tau	Glx	tCho	tCr	tNAA	All
Mean (SD)	10.20 (7.35)	3.19 (2.35)	34.72 (29.88)	10.44 (6.40)	8.63 (6.82)	9.64 (6.00)	31.60 (29.27)	3.82 (3.02)	3.48 (2.29)	9.15 (7.10)	2.43 (1.84)	3.56 (2.63)	13.64 (11.38)

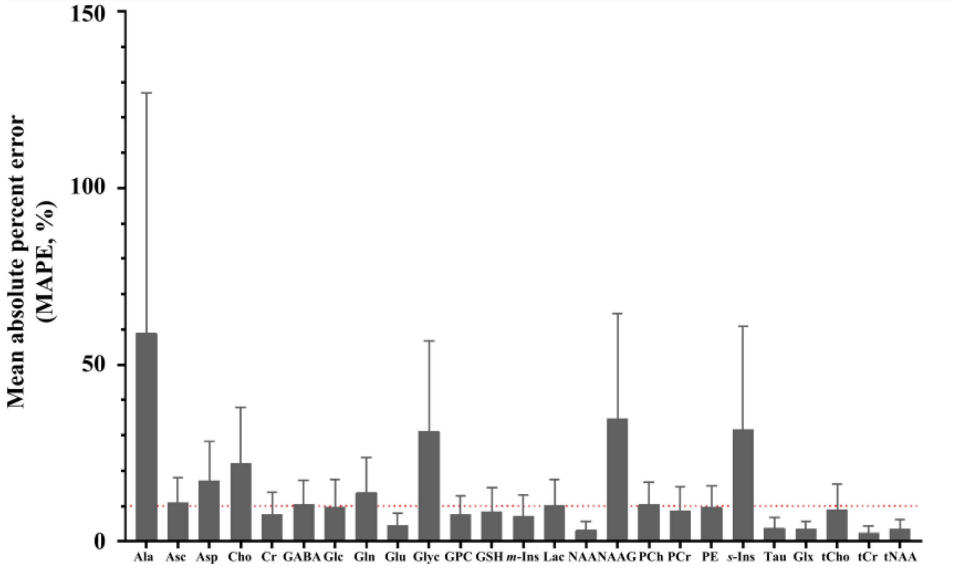

Fig. 4. (Color online) Mean absolute percent error (MAPE) of each metabolite's concentration for the simulated spectra in the test data. The red dotted line indicates that the MAPE is 10 %.

3. Results

Fig. 3 (A-D) shows the representative simulated spectra in the test data along with the ground truths (Fig. 3 (E-H)) and metabolite spectra predicted by the CAE (Fig. 3 (I-L)). Despite the different SNRs, linewidths, and macromolecular baselines, the CAE effectively extracts the metabolite spectra (Fig. 3 (I-L)). The linewidth at the range of 6-22 Hz in the simulated spectrum of the mouse brain was narrowed to 10 Hz in the spectrum predicted by the CAE, which corresponds to the linewidth of the ground truth. The differences between the ground truth and CAE-predicted spectra showed small when referring to the residual spectra that subtracted the CAE-predicted spectra from the ground truth spectra (Fig. 3 (M-P)). These differences were calculated by the mean squared errors (%).

The mean absolute percent error (MAPE) was calculated using the following relation.

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|a_i - p_i|}{a_i} \quad (1)$$

Here, a_i is the actual and p_i is the predicted relative concentration for each spectrum (i = 1, …, n).

Fig. 4 shows the MAPEs of the metabolites quantified from the simulated spectra in the test data. For all metabolites, the average MAPE of twenty-one metabolites was 13.64 ± 11.38 %. The metabolites with MAPE below 10 % (red dotted line) were Cr, Glc, Glu, GSH, Ins, NAA, PCr, Tau, Glx (Glu + Gln), total Cho (tCho), total Cr (tCr), and tNAA.

In the mammalian brain metabolites, the tCr reflects the presence of Cr and PCr, which are known to play an important role in energy metabolism. Because the concentration of tCr is relatively constant throughout the brain and resistant to change, one commonly used approach to quantifying metabolites from 1H-MRS data is the ratio to tCr as an internal standard for other metabolites [22].

In Fig. 5, the ratio of the metabolites to tCr is compared for *in vivo* data using the proposed method, LCModel, and jMRUI [15]. The permissible range of each metabolite concentration is the mean plus or minus 2 times its standard deviation (SD) [1, 14-17]. The number of metabolites within the concentration range of our method, LCModel, and jMRUI are 14, 15, and 6, respectively. Overall, the concentrations for the metabolites quantified using the proposed method were within or close to the expected ranges of the concentrations for the metabolites.

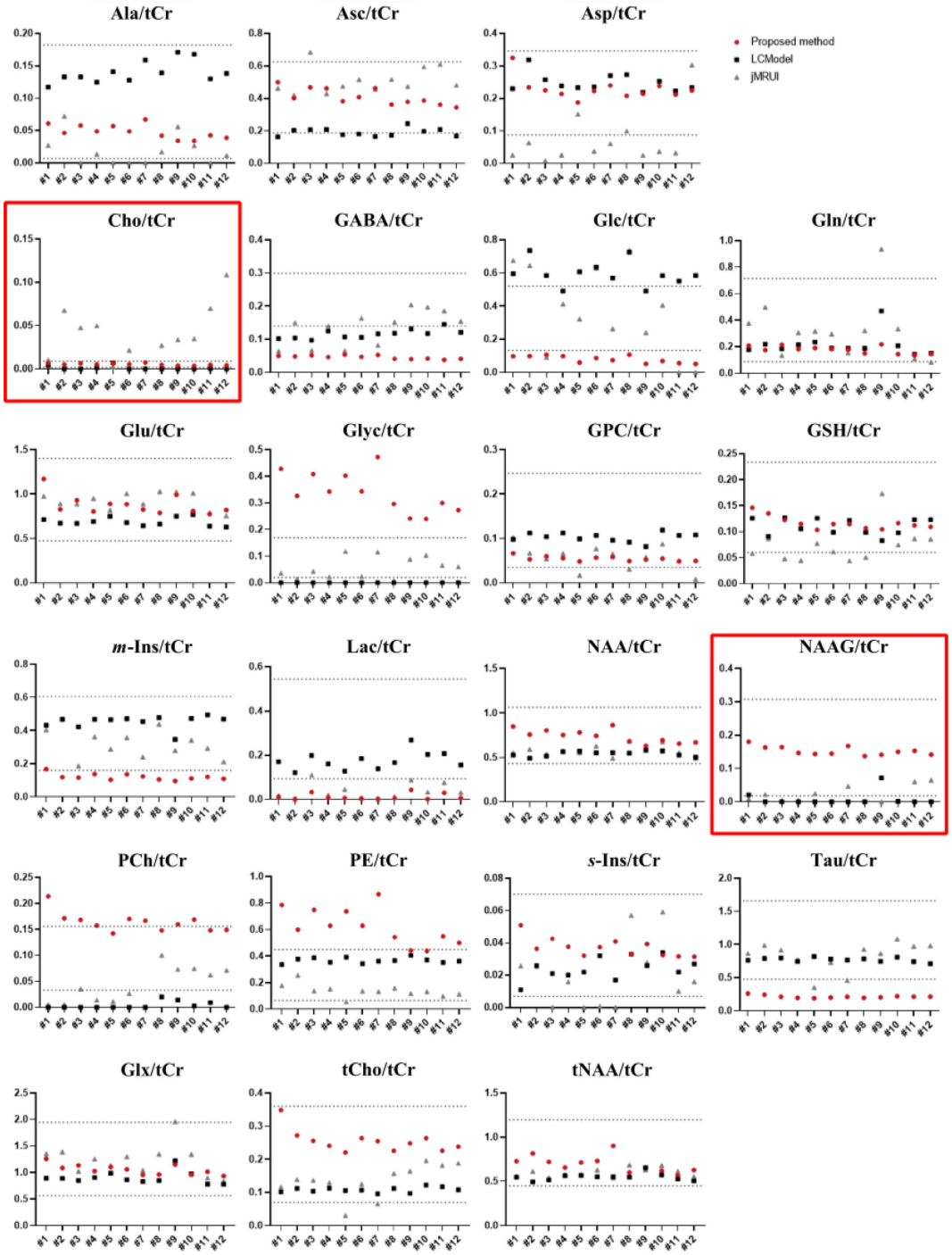

Fig. 5. (Color online) The ratio of concentration for each metabolite to that of total creatine (tCr) estimated by the CAE (red circles), LCModel (black squares), and jMRUI (gray triangles) from 12 *in vivo* ^{1}H-MRS data from the mouse brain (#1-#12). The expected range of the ratio for each metabolite are marked with the gray dotted lines.

4. Discussions

Generally, quantifying the metabolites through the spectra of the mouse brain obtained using ^{1}H-MRS is difficult for the following reasons: 1) strongly overlapped metabolite peaks, 2) low SNR, 3) macromolecules and lipids that overlap with the metabolite peaks, 4) non-analytical lineshapes of peaks due to high magnetic field strength, 5) residual water peaks [2]. Since the voxel size for a mouse is smaller than that for a human, the mouse

data is more degraded than the human one. Nonlinear least squares fitting methods have been widely used in the frequency domain and time domain to quantify the metabolite from the degraded brain spectra. However, this can be challenging even with the current state of the art software packages. Therefore, developing a robust method for brain metabolite quantification is an important issue in studies using ^{1}H-MRS. This study developed CAE for each metabolite to reflect their characteristics using simulated training data (n = 45,000).

For the simulated test data (n = 5,000), the average MAPE of the proposed method was 13.64 ± 11.38 %. Ala, Glyc, NAAG, and *s*-Ins were quantified with high MAPE for the simulated spectra of the mouse brain. These metabolites are known to present in mammalian brain at concentrations of < 1 mM [1, 23]. The concentrations of the metabolites were smaller and more difficult to quantify due to more overlap than that of other metabolites in the simulated spectra. Therefore, the high MAPEs of these metabolites are shown at low concentrations. This study had lower MAPEs than the CNN proposed by Lee [17], which were 20.67 ± 16.71 %. Hence, developing DL models for each metabolite produced better results than DL models trained for the entire metabolite spectrum.

For the *in vivo* spectra of mouse brain, the ratio of the metabolites to tCr was used to compare the performance of the proposed method with LCModel and jMRUI. The number of metabolites within the concentration range (mean ± 2 × SD) in our method, LCModel, and jMRUI were 14, 15, and 6, respectively. NAA (2.010 ppm), NAAG (2.042 ppm), Glu (3.746 ppm), and Gln (3.757 ppm) are difficult to quantify because of spectral overlaps. However, the proposed method extracted the spectra of the metabolites and quantified the metabolites better than the LCModel and jMRUI. The ratios of Asc, Cho, and NAAG to tCr were out of range for the LCModel and the ones of Asp, Cho, GPC, GSH, and NAAG to tCr were out of range for the jMRUI. In the proposed method, the above metabolites were within the expected range. DL has great potential for the quantification of brain metabolites using ^{1}H-MRS. The present study confirmed that DL with ^{1}H-MRS has potential applicability in both human and mouse data.

There are some limitations to the simulation in this study. Residual water signals, other spectroscopic artifacts, and first-order or higher phase distortions were not included in the simulation. Therefore, follow-up studies should be conducted to improve the robustness of the CAE on more realistically simulated *in vivo* spectra so that the CAE can extract metabolite spectra from a more contaminated spectrum.

5. Conclusions

Based on this finding in the present study, it can be concluded that the state-of-the-art quantification method using DL is like the result (*e.g*. similar quantified number of metabolites) of previous studies and quantifies the metabolites (*e.g*. ASC, Cho, and NAAG) that were not well quantified with the existing method. It was also proven to show improved performance (*e.g*. about 34 %) over other study (*e.g*. Lee *et al*. [17]) quantified using DL. This shows that DL has great potential for more precise quantification considering the many factors that can affect the one, and further clinical applications may be possible.

Acknowledgements

This research was supported by the Brain Research Program of the National Research Foundation (NRF) funded by the Korean Government (MSIT) (2016M3C7A1905385).

References

[1] R. A. De Graaf, In vivo NMR spectroscopy: Principles and techniques, John Wiley & Sons, New York (2007) pp. 43-51.
[2] D. Graveron-Demilly, MAGMA **27**, 113 (2014).
[3] J. M. Mateos-Pérez, M. Dadar, M. Lacalle-Aurioles, Y. Iturria-Medina, Y. Zeighami, and A. C. Evans, Neuroimage Clin. **20**, 506 (2018).
[4] L. Minati, M. Grisoli, and M. G. Bruzzone, J. Geriatr Psychiatry Neurol. **20**, 3 (2007).
[5] I. S. Haddadin, A. McIntosh, S. Meisamy, C. Corum, A. L. S. Snyder, N. J. Powell, M. T. Nelson, D. Yee, M. Garwood, and P. J. Bolan, NMR Biomed. **22**, 1 (2009).
[6] I. Tkáč, G. Öz, G. Adriany, K. Uğurbil, and R. Gruetter, Magn. Reson. Med. **62**, 4 (2009).
[7] R. Bartha, NMR Biomed. **20**, 5 (2007).
[8] K. Xu, S. Sigurdsson, V. Gudnason, T. Hue, A. Schwartz, and X. Li, Magn. Reson. Med. **79**, 1722 (2018).
[9] Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
[10] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, Comput. Intell. Neurosci. **2018**, (2018).
[11] G. Guney, B. O. Yigin, N. Guven, Y. H. Alici, B. Colak, G. Erzin, and G. Saygili, Clin. Psychopharmacol. Neurosci. **19**, 206 (2021).
[12] V. Andrearczyk and P. F. Whelan, Biomedical texture analysis, Academic Press, Cambridge (2017) pp 95-129.
[13] F. Khozeimeh, D. Sharifrazi, N. H. Izadi, J. H. Joloudari, A. Shoeibi, R. Alizadehsani, J. M. Gorriz, S. Hussain, Z.

A. Sani, H. Moosaei, A. Khosravi, S. Nahavandi, and S. M. S. Islam, Sci. Rep. **11**, 15343 (2021).
[14] S. P. Kyathanahally, A. Doring, and R. Kreis, Magn. Reson. Med. **80**, 851 (2018).
[15] M. Chandler, C. Jenkins, S. M. Shermer, and F. C. Langbein, arXiv:1909.03836 (2019).
[16] N. Hatami, M. Sdika, and H. Ratiney, arXiv:1806.07237 (2018).
[17] H. H. Lee and H. Kim, Magn. Reson. Med. **82**, 33 (2019).
[18] I. Tkáč, P. G. Henry, P. Andersen, C. D. Keene, W. C. Low, and R. Gruetter, Magn. Reson. Med. **52**, 478 (2004).
[19] M. D. Santin, R. Valabrègue, I. Rivals, R. Pénager, R. Paquin, L. Dauphinot, C. Albac, B. Delatour, and M. C. Potier, NMR Biomed. **27**, 1143 (2014).
[20] S. Y. Kim, Y. J. Lee, H. J. Kim, D. W. Lee, D. C. Woo, C. B. Choi, J. H. Chae, and B. Y. Choe, Brain Res. **1348**, 105 (2010).
[21] Y. Xuan, G. Yan, H. Peng, R. Wu, and H. Xu, Neurochem. Int. **69**, 20 (2014).
[22] Z. Caramanos, S. Narayanan, and D. L. Arnold, Brain **128**, 2483 (2005).
[23] J. K. Cho, I. Carreras, A. Dedeoglu, and B. G. Jenkins, Neuropharmacology **59**, 353 (2010).