DEEP LEARNING FOR MAGNETIC RESONANCE SPECTROSCOPY QUANTIFICATION: A TIME-FREQUENCY ANALYSIS APPROACH

Amirmohammad Shamaei

Doctoral Degree Programme (1st year at FEEC BUT), Czech Academy of Sciences, Institute of Scientific Instruments

E-mail: amirshamaei@isibrno.cz

Supervised by: Radovan Jiřík, Jana Starčuková

E-mail: jirik@feec.vutbr.cz; jana@isibrno.cz

Abstract: Magnetic resonance spectroscopy (MRS) is a technique capable of detecting chemical compounds from localized volumes in living tissues. Quantification of MRS signals is required for obtaining the metabolite concentrations of the tissue under investigation. However, reliable quantification of MRS is difficult. Recently deep learning (DL) has been used for metabolite quantification of MRS signals in the frequency domain. In another study, it was shown that DL in combination with time-frequency analysis could be used for artifact detection in MRS. In this study, we verify the hypothesis that DL in combination with time-frequency analysis can also be used for metabolite quantification and yields results more robust than DL trained with MR signals in the frequency domain. We used the complex matrix of absolute wavelet coefficients (WC) for the time-frequency representation of the signal, and convolutional neural network (CNN) implementation for DL. The comparison with DL used for quantification of data in the frequency domain is presented.

Keywords: magnetic resonance spectroscopy; quantification; deep learning; machine learning;

1 INTRODUCTION

Magnetic Resonance Spectroscopy (MRS) has attracted the MR community over the past 7 decades [1]. A significant part of the interest in biomedical MRS stems from the possibility of noninvasive measurements of metabolites. Information about tissue metabolites can help in clinical diagnostics. For example, detection of metabolic pathway changes may facilitate diagnosing disease in earlier stages before anatomy changes can be observed [1], [2], and thus enable more efficient treatment. E.g., in glioma, a decrease of N-acetylaspartate (NAA) and creatine concentrations of NAA and creatine and an increase of choline, lipids, and lactate predicts an increase of the glioma grade. To reach such a goal, at first, we need to quantify metabolic concentrations. Because there are many obstacles to reaching an accurate estimate of the metabolite concentrations, the use of MRS in daily clinical practice is still not common. The existing MRS quantitation methods are based on model fitting of a signal either in the time or the frequency domain [3]. Even though, in theory there is no difference in which domain is used for fitting, the reality in practice could be different.

Deep learning has achieved many accomplishments in a wide range of tasks, including the MRI field [4]. Due to the poor signal-to-noise ratio (SNR), chemical shift displacement, and overlapping of signal components of the MRS signal, deep learning can be a useful tool. Recentely, Hatami et al. showed the first step in this area by using the deep learning approach for MRS signal quantification [5]. Kim et al. conducted a comprehensive study on brain metabolite quantification using deep learning [6]. The input of both studies is a signal in the frequency domain (metabolite spectra), and their network is a 1D convolutional neural network (CNN). As we mentioned earlier, there are differences between time and frequency domain quantification in practice. Be a case in point, elimination of the first few distorted data points of a signal in the time domain does not significantly dis-

turb the time-domain analysis, whereas the missing time-domain data points can result in complicated modulations throughout the entire spectrum [2]. To overcome the difficulties of the signal analysis in a single domain, time-frequency analysis has been carried out for decades in other areas [4]. Nevertheless, finding an accurate tool for the time-frequency analysis is fraught with difficulty. Here is where deep learning comes to play. Thomas et al. constructed time-frequency images of a speech signal and used them as an input to a CNN for classification [7]. Kyathanahally et al. learned a CNN with time-frequency data to detect and remove ghosting artifacts in clinical magnetic resonance spectra of human brain [8]. Given the mentioned accomplishments of deep learning and time-frequency analysis in a variety of different areas, in particular in MRS for signal artifacts detection, this paper describes to our knowledge the first attempt to use this state-of-the-art technique to quantify MRS signal by deep learning and time-frequency analysis. First, we generate simulated MRS signals. Second, we transform the signals to the time-frequency representation. Third, we train a CNN with the new time-frequency representation. Finally, the result is compared with the previous study.

2 METHODS

A framework is created to generate MRS signals with different amplitudes, damping factors, and frequency shifts. Second, these one-dimensional signals are transformed into their two-dimensional time-frequency representation using wavelet transformation (WT). Finally, the data are split into two datasets, the training and testing datasets. The input of the CNN is the time-frequency representation of signals, and the output is 21 values, which are the concentration-related amplitudes of 20 metabolites and the amplitude of the background signal. The CNN is trained with a training dataset of signals of known amplitudes. Then, the trained CNN is used to estimate the metabolite amplitudes of the test dataset. Finally, the techniques for accuracy evaluation are used.

2.1 SIGNAL GENERATION

Deep learning approaches need a considerable amount of data. For this purpose, we need a basis set (metabolite signals with known concentrations) either simulated or acquired. To be able to compare our results with the previous studies [5], [9], we used the same simulated basis set as used in those studies, i.e., the basis set provided for the ISMRM challenge 2016 [10]. The MRS signal is defined as a combination of amplitude-scaled phase-shifted metabolite basis set signals, the baseline and noise (in this study we use a noisless signal). The mathematical model for the parametric part of the MRS signal is given by:

$$S[n] = \left[\sum_{m=1}^{M} A_m \cdot X_m[n] \cdot e^{(\Delta\alpha_m + 2i\pi\Delta f_m n\Delta T)}\right] + A_{MM} \cdot MM[n] \cdot e^{(\Delta\alpha_{MM} + 2i\pi\Delta f_{MM} n\Delta T)} \quad (1)$$

where $X_m[n]$ is the n-th sample of the m-th simulated metabolite, ΔT is a sampling period, A_m is the scaling factor of the metabolite ($A_m * X_m[0]$ is an indication of the metabolite concentration), $\Delta\alpha_m$ is the damping factor, Δf_m is the frequency shift of the m-th metabolite, and M is the number of metabolites. For our signal simulation the values of the amplitude, damping, and frequency shift are chosen randomly from a defined range with a uniform distribution ($A = [0,1], \Delta\alpha = [-10,10]$ *and* $\Delta f = [-10,10]$). The known background signal MM is considered as another metabolite, then is added to *S*[n] with a random scaling factor, damping, and frequency shift ($A_{MM}, \Delta\alpha_{MM}$ *and* Δf_{MM}). Ten thousand signals are generated, in which the process of value selection is entirely random, thus preventing any bias to our train dataset. The basis set used was simulated for sequence PRESS, magnetic field 3T, echo time TE= 30 ms, spectrum width SW = 4000 Hz, and 2048 time-domain samples.

2.2 SIGNAL PROCESSING

The time-frequency representation of the 1D signal shows a signal in both the time and frequency domain simultaneously. One of the forms of the time-frequency representation of the signal is a scalogram (a matrix of absolute values of the continuous wavelet coefficients (CWC) of a signal) that can be plot as a function of time and frequency. The scalogram is calculated using the Matlab Wavelet Toolbox (R2019a, Mathworks Inc.,Natick, MA, USA). We use Morse wavelet to compute the CWC. The last 512 points of the time signals are cut off to reduce the amount of computation for the CWC calculation. The selection of the number of points was decided by visual inspection of the signals to ensure that no significant information will be lost. The wavelet coefficients are computed. The minimum and maximum scales are determined automatically based on the energy spread of the wavelet in frequency and time by the toolbox. The coefficients matrix is a matrix where each row corresponds to one scale, and its column size is equal to the length of signal. Scalogram with 340 frequency bins and 1536 time points (340 × 1536 matrix) is created. Finally, the real and imaginary parts of 10000 matrices are stored in two channels.

2.3 CNN

A convolutional neural network is developed using the Matlab Deep Learning Toolbox (R2019a, Mathworks Inc.,Natick, MA, USA) on NVIDIA GTX 1050Ti graphics processing units. The architecture of the CNN is shown in Fig. 1 . This network includes one input layer with two channels, six convolutional layers, five max pool layers, and one regression layer. Rectified linear unit (ReLU) activation functions are used between CL and MP layers. The mean square error is implemented as the loss function. The output of regresression layer is 21 parameters which correspond to twenty metabolites and one background MM. Using these parameters and Eq (1), the estimated signal is reconstructed.

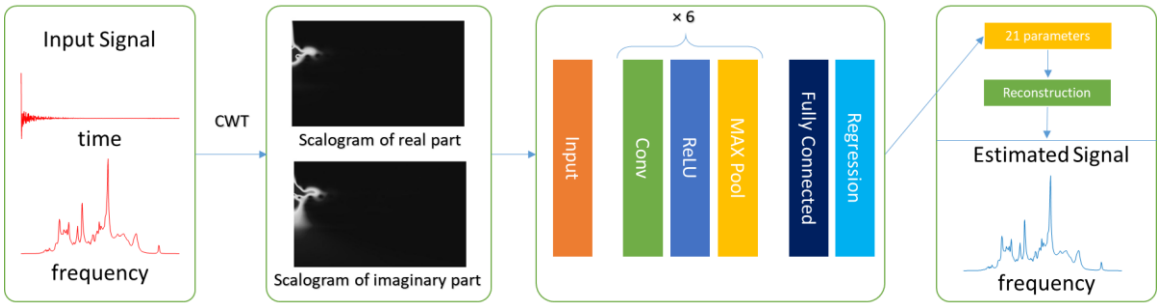

Figure 1: A schematic of the proposed approach. The generated signal based on a linear combination of metabolites basis sets is converted to two gray Scalogram images (real and imaginary). These images feed to CNN as inputs. The CNN includes 6 blocks, which comprise a Convolutional, Rectifier and Max-pooling layer. The last layers are a fully connected and a regression layer (which has 21 outputs). The estimated signal is reconstructed with the estimated parameters.

2.4 ACCURACY EVALUATION

Two methods are used to measure the accuracy of the model. First, the mean absolute error, which is the most straightforward regression error metric. MAE is defined as below for each metabolite:

$$\mathrm{MAE[m]} = \frac{1}{N}\sum_{n=1}^{N} |A_{mn} - A'_{mn}| \tag{2}$$

where m, N, A, and A' are the metabolite index, the number of test datasets, the ground truth, and the estimated amplitude, respectively. The second method is the Symmetric mean absolute percentage error (SMAPE) which is given by:

$$\mathrm{SMAPE[m]} = \frac{\sum_{n=1}^{N} |A_{mn} - A'_{mn}|}{\sum_{n=1}^{N} |A_{mn} + A'_{mn}|} \tag{3}$$

2.5 RESULTS

The dataset is separated into two datasets, namely a training dataset and a test dataset. The training dataset contains 80% of the data and the remaining 20% are the test dataset. CNNs with different hyperparameters such as minimum batch size, initial learning rate, and validation frequency are tested, and the CNN with the best result is chosen. The minimum batch size, initial learning rate, and validation frequency are 30, 1e-5, and 10, respectively. It has been shown that increasing the training sample would decrease the value of loss function [5]. Nonetheless, to be able to compare results obtained with our new approach (DL with time-frequency domain input) with the results of the Hatami et al approach (DL with frequency domain input) in a reasonable time, we decided to use only 10000 samples for CNN training and testing. Training and validation loss for the given dataset are 0.18 and 0.23, respectively.

Fig. 2 shows one of the tested (ground truth) signal, its estimate, and residual. The following conclusions may be drawn from this figure. First, the method used is able to estimate the tested signal. Second, residuals mainly occur when the signal shows rapid fluctuation.

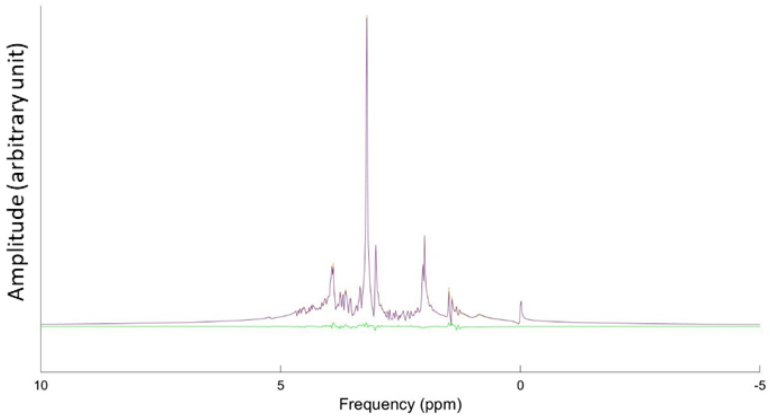

Figure 2: Example of the signal estimation – ground truth signal (orange), estimated signal (violet) and residual signal (green).

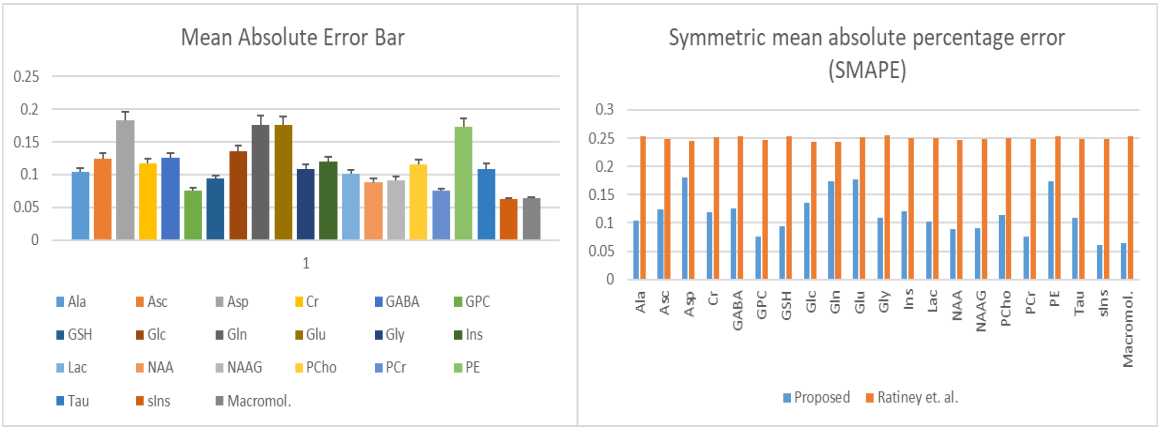

Figure 3: (left) Mean absolute error bar of every metabolite and its variance. (right) The symmetric mean absolute percentage error of each metabolite for (blue) our study (orange) Hatami et al.[5]

The mean absolute errors (MAE) of metabolites are shown in Fig. 3 (left). Even though the amount of error is not too low compared to the amplitude range ([0, 1]), the variance of the error is small. Fig. 3 (right) shows the comparison of the Symmetric mean absolute percentage error (SMAPE) between our approach and Hatami et al. approach. To avoid any bias in comparison, we used

the CNN described in this study and train it with our training set but in one case in the form of scalogram and in the other case (Hatami et al.) with the data in the frequency domain. The proposed approach shows less amount of error compared to their method.

3 CONCLUSION

Quantification of MRS is an important topic where a robust and universal panacea approach to quantify signals is needed. It was shown in this study that time-frequency deep learning quantification could outperform single domain quantification used in the previous studies [2, 5] and hopefully as a method using information from both MRS domains be successfully used also for quantification of signals with artifact patterns [8]. The next steps may be to verify the tested approach on 1) the simulated noisy MRS with different signal-to-noise-ratios and for different pulse sequences 2) on real MRS acquired from a phantom, 3) on MRS acquired from a rat, and 4) to implement this approach as a plugin in the jMRUI software package [11].

ACKNOWLEDGEMENT

This research was supported by European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 813120 (INSPiRE-MED).

REFERENCES

[1] M. Van Der Graaf, "In vivo magnetic resonance spectroscopy: Basic methodology and clinical applications," *Eur. Biophys. J.*, vol. 39, no. 4, pp. 527–540, 2010.

[2] Robin A. de Graaf, "In Vivo NMR Spectroscopy," *In Vivo NMR Spectroscopy*. pp. 1–42, 13-Feb-2019.

[3] J. B. Poullet, D. M. Sima, and S. Van Huffel, "MRS signal quantitation: A review of time- and frequency-domain methods," *J. Magn. Reson.*, vol. 195, no. 2, pp. 134–144, 2008.

[4] H. Alaskar, "Deep Learning-Based Model Architecture for Time-Frequency Images Analysis," no. January 2018, 2019.

[5] N. Hatami, B. Lyon, and U. Etienne, "Magnetic Resonance Spectroscopy Quantification using Deep Learning," 2018.

[6] H. H. Lee, "Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain," no. January, pp. 33–48, 2019.

[7] S. Thomas, S. Ganapathy, G. Saon, and H. Soltau, "Analyzing convolutional neural networks for speech activity detection in mismatched acoustic conditions," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2014, pp. 2519–2523.

[8] R. Kreis and S. P. Kyathanahally, "Deep Learning Approaches for Detection and Removal of Ghosting Artifacts in MR Spectroscopy," vol. 863, pp. 851–863, 2018.

[9] D. Das, E. Coello, R. F. Schulte, and B. H. Menze, "Quantification of Metabolites in Magnetic Resonance Spectroscopic Imaging Using Machine Learning," vol. 3, pp. 462–469, 2017.

[10] ISMRM, "MRS Fitting Challenge (ismrm.org/workshops/Spectroscopy16/mrs_fitting_challenge/)," 2016. [Online]. Available: ismrm.org/workshops/Spectroscopy16/mrs_fitting_challenge/.

[11] D. Stefan *et al.*, "Quantitation of magnetic resonance spectroscopy signals: The jMRUI software package," *Meas. Sci. Technol.*, vol. 20, no. 10, 2009.