2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

High-Resolution Magnetic Resonance Spectroscopic Imaging using a Multi-Encoder Attention U-Net with Structural and Adversarial Loss

Siyuan Dong[1], Gilbert Hangel[2], Wolfgang Bogner[2], Siegfried Trattnig[2], Karl Rössler[3], Georg Widhalm[3], Henk M. De Feyter[4], Robin A. de Graaf[4], and James S. Duncan[4]

***Abstract*— Common to most medical imaging techniques, the spatial resolution of Magnetic Resonance Spectroscopic Imaging (MRSI) is ultimately limited by the achievable SNR. This work presents a deep learning method for ^{1}H-MRSI spatial resolution enhancement, based on the observation that multi-parametric MRI images provide relevant spatial priors for MRSI enhancement. A Multi-encoder Attention U-Net (MAU-Net) architecture was constructed to process a MRSI metabolic map and three different MRI modalities through separate encoding paths. Spatial attention modules were incorporated to automatically learn spatial weights that highlight salient features for each MRI modality. MAU-Net was trained based on in vivo brain imaging data from patients with high-grade gliomas, using a combined loss function consisting of pixel, structural and adversarial loss. Experimental results showed that the proposed method is able to reconstruct high-quality metabolic maps with a high-resolution of 64×64 from a low-resolution of 16×16, with better performance compared to several baseline methods.**

***Clinical relevance*— This work shows that the spatial priors in multi-parametric MRI are useful for MRSI enhancement. After further validation, the proposed method can be used to accelerate spectroscopic imaging for clinical applications.**

I. INTRODUCTION

Magnetic Resonance Spectroscopic Imaging (MRSI) is a non-invasive imaging technique that allows the study of metabolism in various organs and was especially used for mapping metabolite levels for the human brain [1]. However, due to the low concentration of many important metabolites found in the brain, the application of MRSI is often limited to very coarse spatial resolutions [2]. High-resolution (HR) MRSI obtained from strong magnetic field or long acquisition time is often not affordable. Therefore, a post-processing approach for improving spatial resolution and SNR would benefit both clinical and preclinical MRSI applications.

Many post-processing techniques have been developed for HR MRSI [3]–[6]. Most of these methods are based on model-based regularization and often result in slow and unrealistic HR reconstructions. Recently deep learning has achieved significant success in medical imaging [7] and was applied to MRSI quantification [8]–[11]. However, very few deep learning methods utilized the spatial priors in MRI from the same subject to aid the MRSI enhancement. The main challenge is the lack of paired MRSI and MRI data publicly available. A recent work [12] proposed to incorporate native T1-weighted MRI for super-resolution MRSI by training a neural network with synthetic MRSI data. It circumvented the issue of no public MRSI data, but training entirely with synthetic data resulted in unrealistic HR metabolic maps. Moreover, incorporating spatial priors from other MRI modalities using data-driven methods remains unexplored, although some regularization-based approaches have been proposed [4], [13]. In this work, we summarize our contributions as follows:

- We propose to use in vivo training data for ^{1}H-MRSI spatial resolution enhancement.
- We demonstrate using a deep learning method the benefit of incorporating spatial priors in multi-parametric MRI when doing MRSI enhancement, specifically native T1-weighted (T1), contrast-enhanced T1-weighted (T1CE) and fluid-attenuated inversion recovery (FLAIR).
- We develop a Multi-encoder Attention U-Net (MAU-Net) architecture, in combination with pixel, structural and adversarial loss, to improve the resolution of 2D metabolic maps from low resolution (LR) of 16×16 to HR of 64×64.

II. METHODS

A. Data Acquisition and Processing

3D-HR-MRSI, MP2RAGE (T1) and FLAIR scans of high-grade glioma patients were acquired with a 7T whole-body-MRI (Magnetom, Siemens Healthcare, Erlangen, Germany) equipped with a 32-channel receive coil array (Nova Medical, Wilmington, MA, USA) [14]. Informed consent and IRB approval were successfully obtained. T1CE images were taken from clinical routine scans. The MRI images were skull-stripped and co-registered via FSL (FMRIB Software Library v5.0, Oxford, UK). The MRSI sequence used spatial-spectral encoding to acquire 3D metabolic maps with a 64×64×39 measurement matrix and 3.4×3.4×3.4 mm^3 nominal resolution in 15 min [14], [15]. Further parameters were with an acquisition delay of 1.3 ms and a TR of 450 ms and 2778 Hz spectral bandwidth. Voxel spectra were quantified using LCModel (v6.3-1, LCMODEL Inc, ONT, CA). The basis set included NAA, NAAG, Cr, phosphocreatine, phos-

The research is supported by NIH grant R01EB025840, R01CA206180 and R01NS035193.

[1]Department of Electrical Engineering, Yale University, New Haven, CT, USA s.dong@yale.edu

[2]Department of Biomedical Imaging and Image-guided Therapy, High-field MR Center, Medical University of Vienna, Vienna, Austria

[3]Department of Neurosurgery, Medical University of Vienna, Vienna, Austria

[4]Department of Radiology and Biomedical Imaging, Yale University School of Medicine, New Haven, CT, USA

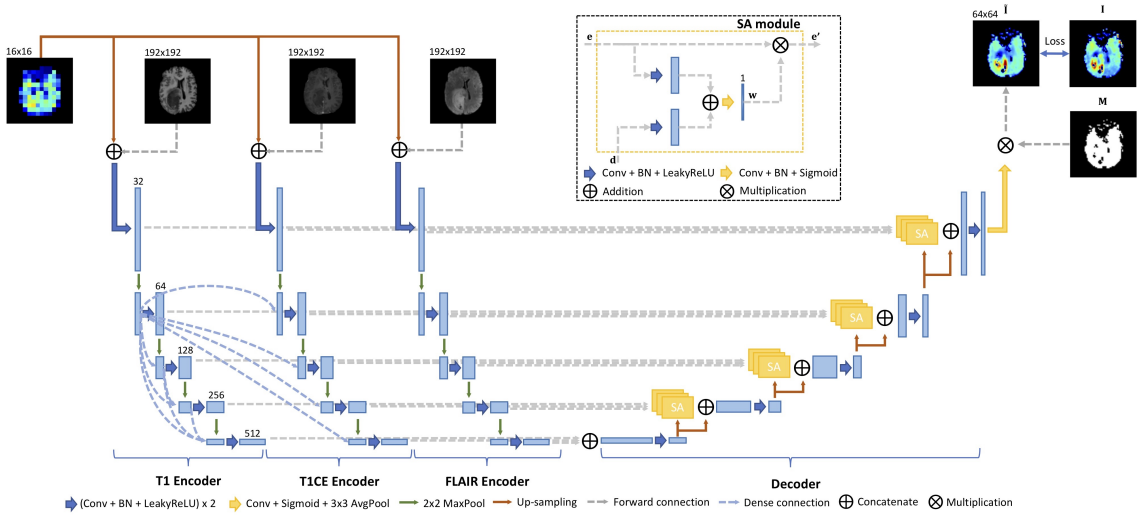

Fig. 1. Proposed architecture Multi-encoder Attention U-Net (MAU-Net). The detailed structure of the spatial attention (SA) modules is shown in the dotted box. Each light-blue box represents a feature map, above which is the number of channels. Note that only some of the dense connections are shown. Conv: Convolution, BN: batch-normalization, AvgPool: average-pooling, MaxPool: max-pooling.

phocholine, glyceryl-phosphorylcholine, mIns, scyllo-Ins, γ-aminobutyric acid (GABA), glutathione (GSH), Glu, Gln, Gly, Tau, cystathionine (Ctn), cysteine (Cys), serine (Ser), 2HG, and a macromolecular background, with an evaluation range of 0.2-1.2 ppm and 1.8-4.2 ppm (excluding the lipid resonance at 1.3 ppm) [14].

In this work, data from 3 different patients were used. From each patient, we selected 10˜13 axial slices, and from each slice 8˜10 metabolic maps were selected, namely total choline (tCho), total creatine (tCr), NAA, NAAG, Gly, Gln, Glu, Ins, GABA and Ser. The selections were based on the low voxel rejection rate in ROI [14]. We ultimately obtained 320 64×64 metabolic maps, of which the pixel values were normalized to $[0, 1]$. These maps were taken as the ground truth images and were down-sampled to 16×16 to obtain the LR inputs. The MRI slices at corresponding spatial locations were re-sampled to 192×192 and normalized to $[0, 1]$.

B. Network Architecture

U-Net [16] was originally proposed for image segmentation tasks. The encoder-decoder structure first down-samples the feature maps to capture high-level contextual information, which are then up-sampled to generate an output with the same resolution as the input. The skip connections from the encoder to the decoder serve to retain fine-grained details at each feature level. These characteristics of U-Net are also desirable for the resolution enhancement task in this work. The proposed architecture MAU-Net, shown in Fig.1, was built upon the U-Net, but it differs by having multiple encoders and having spatial attention modules.

Following a recent work on multi-modal image segmentation [17], we adopted a multi-encoder structure to incorporate each of the MRI modalities into MRSI reconstruction through an individual encoder. The goal of processing each modality separately is to disentangle complex information in different modalities that otherwise would be fused too early when all modalities are concatenated into a single input of a single-encoder U-Net [17]. The hyper-dense connections allow better flow of information and gradient within and between the modalities [17]. The building blocks consist of convolution layers with 3×3 kernels, batch-normalization [18], LeakyReLU, 2×2 max-pooling for down-sampling in the encoders and bicubic interpolation for up-sampling in the decoder. A 3×3 average-pooling was used in the last layer to produce an output image with a size of 64×64. Because there are many missing values in a ground truth image $\mathbf{I}$ due to voxel rejections in quality filtering [14], we applied a mask $\mathbf{M}$ that only keeps the pixels with valid ground truth values in order to have the valid computation of the loss function.

To further improve the learning capability of the network, inspired by recent segmentation networks Attention U-Net [19] and SCAU-Net [20], we incorporated spatial attention (SA) modules on the skip connection paths to automatically learn spatial weight maps that focus on salient features of each modality. Observing that different MRI modalities highlight distinct anatomical information (T1 shows nice image contrast between white matter and grey matter, T1CE highlights active tumor and FLAIR delineates the whole tumor), a SA module was used for each encoder at each feature level. Inside the SA module is a series of weight self-learning operations:

$$\mathbf{w} = \sigma(\mathbf{f_3}(\mathbf{f_1}(\mathbf{e}) + \mathbf{f_2}(\mathbf{d}))) \tag{1}$$

where $\mathbf{e}$ is the feature map from encoder, $\mathbf{d}$ is from decoder, $\mathbf{f}$ are 3×3 convolution layers and σ is the sigmoid activation function used to generate spatial weights $\mathbf{w} \in [0, 1]$ [20]. The spatial weight map is then multiplied with $\mathbf{e}$ to get the weighted feature map $\mathbf{e}'$:

$$\mathbf{e}' = \mathbf{e} \cdot \mathbf{w} \tag{2}$$

The discriminator used in adversarial loss (see next section) consists of 5 convolution building blocks, each has

a convolution layer with 3×3 kernels (number of kernels 32, 64, 128, 256, 512 respectively), instance normalization [21] and LeakyReLU. Two subsequent fully connected layers were used to compute a final discriminator score.

C. Loss Function

The mean-squared error imposes pixelwise accuracy between reconstructed image $\tilde{\mathbf{I}}$ and ground truth image $\mathbf{I}$

$$L_{pixel} = \frac{1}{HW}\sum_{i,j}^{H,W}(\tilde{\mathbf{I}}_{i,j} - \mathbf{I}_{i,j})^2 \tag{3}$$

where H and W are the image height and width respectively. However, the pixel loss alone does not capture the inter-pixel dependencies and usually results in over-smoothing.

In order to model inter-pixel correlations, the structural loss should be included as part of the loss function [22]. The structural similarity index (SSIM) [23] measures the similarity between images in terms of three perceptually motivated metrics luminance l, contrast c and structure s. As an advanced version of SSIM, the Multi-scale SSIM (MS-SSIM) [24] provides more flexibility by measuring the SSIM between two images $\mathbf{x}$ and $\mathbf{y}$ at multiple image scales:

$$\text{MS-SSIM}(\mathbf{x},\mathbf{y}) = [l_M(\mathbf{x},\mathbf{y})]^{\alpha_M}\prod_{j=1}^{M}[c_j(\mathbf{x},\mathbf{y})]^{\beta_j}[s_j(\mathbf{x},\mathbf{y})]^{\gamma_j} \tag{4}$$

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

where μ_x, μ_y, σ_x, σ_y, σ_{xy} represents mean, standard deviation (SD) and cross-covariance of the image pair $\mathbf{x}$ and $\mathbf{y}$. M is the number of scale levels and the rest of the parameters were set to constant values suggested in [24]. Therefore, the structural loss is expressed as

$$L_{structural} = 1 - \text{MS-SSIM}(\tilde{\mathbf{I}}, \mathbf{I}) \tag{5}$$

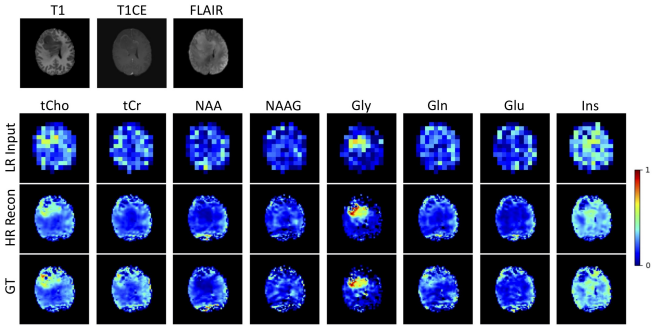

Fig. 2. High-resolution reconstruction results given by the MAU-Net trained with the proposed loss function in Eq.(8). From top to bottom: three MRI modalities, low-resolution input with noise SD 0.1 (LR Input), high-resolution reconstructions (HR Recon) and ground truth (GT).

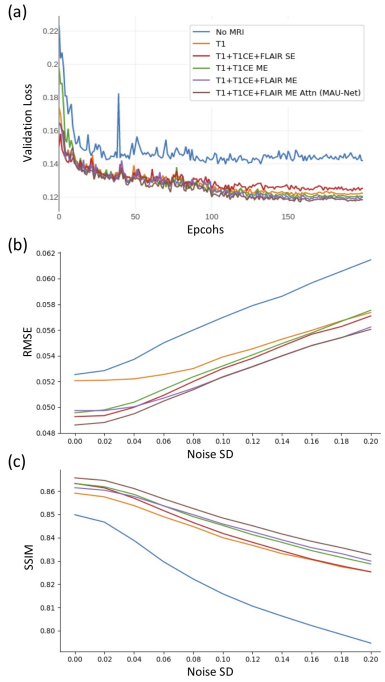

Fig. 3. Quantitative comparisons of 6 architectures defined in section II-D. (a) Validation loss versus the number of training epochs; (b) RMSE versus the input noise SD; (c) SSIM versus the input noise SD. SE: single-encoder, ME: multi-encoder, Attn: Attention.

However, using the structural loss can still miss high-frequency image details. Therefore the adversarial loss was incorporated to capture difficult-to-learn features and produce more visually pleasing HR metabolic maps [25]–[28]. Adversarial loss is the loss associated with training Generative Adversarial Networks (GAN) [29], which comprises two components: a generator G and a discriminator D. The role of the discriminator is to distinguish whether an image is from the real distribution or the generated distribution, and the generator tries to fool the discriminator by generating images as close to the real distribution as possible. In this work, the proposed network in Fig.1 was treated as the generator, and an extra discriminator was introduced to discriminate the generated image $\tilde{\mathbf{I}}$ from the ground truth $\mathbf{I}$. The original GAN uses Jenson-Shannon (JS) divergence to measure the similarity between distributions, but this usually results in unstable training [30]. An improved version of GAN, called the Wasserstein GAN [30], uses Wasserstein distance to alleviate the non-convergence problem. The training process involves alternatively updating the parameters in G and D based on the loss functions

$$L_{adversarial}(D) = -\mathbb{E}[D(\mathbf{I})] + \mathbb{E}[D(\tilde{\mathbf{I}})] + \lambda\mathbb{E}\left[\left(\left\|\nabla_{\hat{\mathbf{I}}}D(\hat{\mathbf{I}})\right\|_2 - 1\right)^2\right] \tag{6}$$

$$L_{adversarial}(G) = -\mathbb{E}[D(\tilde{\mathbf{I}})] \tag{7}$$

where the first two terms in Eq.(6) are for the Wasserstein distance, and the last term is a gradient penalty term for network regularization [31]. Note that $\tilde{\mathbf{I}}$ is generated by

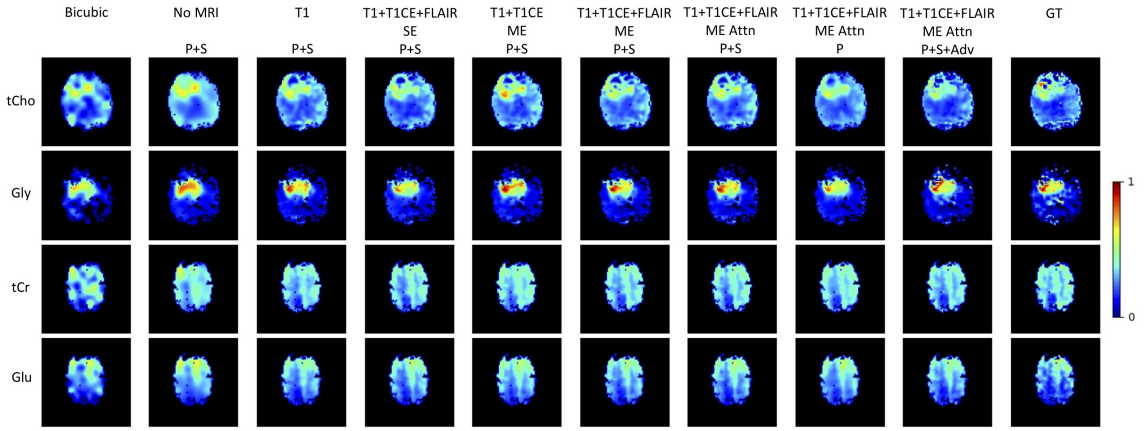

Fig. 4. Visual comparisons of the bicubic interpolation and several methods described in section II-D. tCho and Gly maps were obtained from the same patient at the same slice, and tCr and Glu maps were obtained from another patient at the same slice. Note that "T1+T1CE+FLAIR ME Attn" is equivalent to MAU-Net. P: pixel loss, S: structural loss, Adv: adversarial loss, GT: ground truth.

G, and $\hat{\mathbf{I}}$ is uniformly sampled along the straight lines connecting pairs of generated and ground truth images [31].

The final loss function for training the proposed MAU-Net is, therefore, a weighted combination of pixel, structural and adversarial loss:

$$L = L_{pixel} + \lambda_1 L_{structural} + \lambda_2 L_{adversarial}(G) \qquad (8)$$

D. Experiments and Comparisons

To justify the use of MAU-Net, we implemented 6 different architectures for ablation studies. Firstly, to demonstrate the effectiveness of incorporating multi-parametric MRI, i.e. T1, T1CE and FLAIR, into MRSI enhancement, we implemented the following combinations: (1) no MRI, a single-encoder U-Net takes only the LR metabolic map as the input, (2) T1 Encoder only, (3) T1 and T1CE Encoders only, and (4) all three Encoders. Secondly, to show the importance of the multi-encoder structure, we evaluated a single-encoder U-Net that takes the LR metabolic map and 3 MRI modalities as a single concatenated input. After that, we evaluated the contribution of the SA modules, which ultimately forms the proposed architecture MAU-NET. To reliably compare the learning capabilities of different architectures, we temporarily discarded the adversarial loss from Eq.(8) because of the probabilistic behavior (too much randomness) when training adversarial networks. Finally, to justify using the proposed loss function, we trained one MAU-Net with the pixel loss only and another one with the proposed loss in Eq.(8).

From 320 HR metabolic maps, we selected 272 for training, 28 for validation and 20 for testing. For data augmentation, the network was trained with random patches of 32×32 instead of the entire 64×64 images, which means the input MRI images and up-sampled LR metabolic map were cropped at 96×96 at corresponding positions. Further augmentations included random image flipping and rotation. We added random Gaussian noise to the LR inputs during training, and the noise SD was sampled uniformly from $[0, 0.2]$ to simulate a range of different SNR levels. The training was performed over 200 epochs for all experiments, amounting to 54400 training samples. The validation and test datasets were augmented to 308 and 220, respectively, by adding noise at 11 different levels to each LR input.

We set $\lambda = 10$ as suggested in [31]. λ_1 and λ_2 were manually tuned as 1 and 0.001 respectively, so the reconstructed maps achieve a good balance between accuracy and sharpness. All the networks were optimized with the Adam optimizer, and all experiments were implemented in PyTorch on a NVIDIA's GeForce GTX 1080 GPU.

III. RESULTS AND DISCUSSION

The ability of MAU-Net, trained with proposed loss function in Eq.(8), to reconstruct HR metabolic maps of tCho, tCr, NAA, NAAG, Gly, Gln, Glu and Ins at 64×64 from 16×16 is shown in Fig.2. The shown metabolic maps and MRI images were selected from a single slice of a particular patient. All reconstructed images show similar regional changes as the ground truth images. For example, tCho and Gly are more concentrated in the glioma tumor than in the normal brain tissue, and NAA is less concentrated. Additionally, Glu and Gln are more concentrated in gray matter than in white matter, consistent with the ground truth and the biological fact [32]. Note that the reconstructed HR images not only retained signals in LR inputs but also recovered lots of sharp details.

Fig.3 shows quantitative comparisons of 6 different architectures as described in section II-D. Fig.3(a) shows that our proposed MAU-Net achieved the lowest validation loss during training. Obviously, the one with no input from MRI stopped converging from very early epochs. Incorporating only T1 MRI improved the learning significantly, indicating the necessity of MRI spatial priors in MRSI enhancement. The validation loss converged to even lower levels progressively with the additional incorporation of priors in T1CE and FLAIR. It also shows that the single-encoder U-Net did not converge as well as the multi-encoder ones due to the early fusion problem. Fig.3(b) and (c) show the average root-mean-squared error (RMSE) and SSIM between reconstructed and ground truth images over the test dataset.

The metrics were calculated at a range of input noise levels, and at each level, we took the average of 100 experiments with different noise samples. MAU-Net outperformed the others for most noise levels in terms of both pixelwise accuracy (RMSE) and perceptual similarity (SSIM).

Fig.4 displays the reconstructed tCho, Gly, tCr and Glu maps given by the traditional bicubic interpolation, 6 different network architectures, and MAU-Net trained with different loss functions. The bicubic interpolation and the one with no MRI input gave very blurry images, not providing much HR information. While the incorporation of T1 could significantly improve the image quality, the additional incorporation of T1CE and FLAIR helped to get closer to the ground truth. Compared to MAU-Net trained with pixel loss plus structural loss, the one trained with only pixel loss generated over-smoothed metabolic maps with less spatial heterogeneity (mainly observed in the tCr and Glu maps), and the one with adversarial loss was able to recover some of the sharp details around the brain periphery (observed in the tCho map) and inside the tumor (observed in the Gly map). However, we observed that adversarial loss can introduce some artifacts, especially when λ_2 was set to higher levels. Therefore, the adversarial loss should be used with caution to balance the image sharpness and the hallucination risk.

IV. CONCLUSIONS

We proposed a novel deep learning approach to incorporate multi-parametric MRI into MRSI resolution enhancement based on in vivo training data. Experimental results indicated that using a multi-encoder structure and SA modules can boost the network learning capability for this problem. Results also showed that structural loss and adversarial loss recovered some of the HR features in the experimentally acquired HR metabolic maps. The proposed algorithm will be further evaluated after a more comprehensive in vivo dataset from a larger patient cohort is collected. After further validation, this method will be extended to accelerate spectroscopic imaging of other organs, e.g. liver, or of other nuclei, e.g. ^{2}H and ^{31}P.

REFERENCES

[1] D. P. Soares and M. Law, "Magnetic resonance spectroscopy of the brain: review of metabolites and clinical applications," Clinical Radiology, vol. 64, no. 1, pp. 12–21, Jan. 2009.

[2] R. A. de Graaf, In Vivo NMR Spectroscopy: Principles and Techniques. Chichester, UK: John Wiley & Sons, Ltd, 2019.

[3] F. Lam and Z.-P. Liang, "A subspace approach to high-resolution spectroscopic imaging: High-Resolution Spectroscopic Imaging," Magn. Reson. Med., vol. 71, no. 4, pp. 1349–1357, Apr. 2014.

[4] S. Jain et al., "Patch-Based Super-Resolution of MR Spectroscopic Images: Application to Multiple Sclerosis," Front. Neurosci., vol. 11, Jan. 2017.

[5] M. Jacob, X. Zhu, A. Ebel, N. Schuff, and Z.-P. Liang, "Improved Model-Based Magnetic Resonance Spectroscopic Imaging," IEEE Trans. Med. Imaging, vol. 26, no. 10, pp. 1305–1318, Oct. 2007.

[6] J. Kasten, A. Klauser, F. Lazeyras, and D. Van De Ville, "Magnetic resonance spectroscopic imaging at superresolution: Overview and perspectives," Journal of Magnetic Resonance, vol. 263, pp. 193–208, Feb. 2016.

[7] D. Shen, G. Wu, and H.-I. Suk, "Deep Learning in Medical Image Analysis," Annu. Rev. Biomed. Eng., vol. 19, no. 1, pp. 221–248, Jun. 2017.

[8] S. S. Gurbani, S. Sheriff, A. A. Maudsley, H. Shim, and L. A. D. Cooper, "Incorporation of a spectral model in a convolutional neural network for accelerated spectral fitting," Magn Reson Med, vol. 81, no. 5, pp. 3346–3357, May 2019.

[9] N. Hatami, M. Sdika, and H. Ratiney, "Magnetic Resonance Spectroscopy Quantification using Deep Learning," arXiv:1806.07237 [cs], Jun. 2018.

[10] F. Lam, Y. Li, and X. Peng, "Constrained Magnetic Resonance Spectroscopic Imaging by Learning Nonlinear Low-Dimensional Models," IEEE Trans. Med. Imaging, pp. 1–1, 2019.

[11] S. Dong, H. M. De Feyter, M. A. Thomas, R. A. de Graaf, and J. S. Duncan, "A Deep Learning Method for Sensitivity Enhancement in Deuterium Metabolic Imaging (DMI)," in Proceedings of the 28th Annual Meeting of ISMRM, 2020, No.0391.

[12] Z. Iqbal, D. Nguyen, G. Hangel, S. Motyka, W. Bogner, and S. Jiang, "Super-Resolution 1H Magnetic Resonance Spectroscopic Imaging Utilizing Deep Learning," Front. Oncol., vol. 9, p. 1010, Oct. 2019.

[13] J. P. Haldar, D. Hernando, S.-K. Song, and Z.-P. Liang, "Anatomically constrained reconstruction from noisy data," Magn. Reson. Med., vol. 59, no. 4, pp. 810–818, Apr. 2008.

[14] G. Hangel et al., "High-resolution metabolic imaging of high-grade gliomas using 7T-CRT-FID-MRSI," NeuroImage: Clinical, vol. 28, p. 102433, 2020.

[15] L. Hingerl et al., "Clinical High-Resolution 3D-MR Spectroscopic Imaging of the Human Brain at 7 T:," Investigative Radiology, vol. 55, no. 4, pp. 239–248, Apr. 2020.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," arXiv:1505.04597 [cs], May 2015.

[17] J. Dolz, I. B. Ayed, and C. Desrosiers, "Dense Multi-path U-Net for Ischemic Stroke Lesion Segmentation in Multiple Image Modalities," arXiv:1810.07003 [cs], Oct. 2018.

[18] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," arXiv:1502.03167 [cs], Mar. 2015.

[19] O. Oktay et al., "Attention U-Net: Learning Where to Look for the Pancreas," arXiv:1804.03999 [cs], May 2018.

[20] P. Zhao, J. Zhang, W. Fang, and S. Deng, "SCAU-Net: Spatial-Channel Attention U-Net for Gland Segmentation," Front. Bioeng. Biotechnol., vol. 8, p. 670, Jul. 2020.

[21] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance Normalization: The Missing Ingredient for Fast Stylization," arXiv:1607.08022 [cs], Nov. 2017.

[22] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss Functions for Neural Networks for Image Processing," arXiv:1511.08861 [cs], Apr. 2018.

[23] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Trans. on Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

[24] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, Pacific Grove, CA, USA, 2003, pp. 1398–1402.

[25] C. You et al., "Structurally-Sensitive Multi-Scale Deep Neural Network for Low-Dose CT Denoising," IEEE Access, vol. 6, pp. 41839–41855, 2018.

[26] M. Ran et al., "Denoising of 3D magnetic resonance images using a residual encoder–decoder Wasserstein generative adversarial network," Medical Image Analysis, vol. 55, pp. 165–180, Jul. 2019.

[27] Q. Yang et al., "Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss," IEEE Trans. Med. Imaging, vol. 37, no. 6, pp. 1348–1357, Jun. 2018.

[28] Z. Wang, J. Chen, S. C.H. Hoi, "Deep Learning for Image Super-resolution: A Survey," arXiv:1902.06068 [cs], Feb. 2020.

[29] I. J. Goodfellow et al., "Generative Adversarial Networks," arXiv:1406.2661 [cs, stat], Jun. 2014.

[30] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," arXiv:1701.07875 [cs, stat], Dec. 2017.

[31] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved Training of Wasserstein GANs," arXiv:1704.00028 [cs, stat], Dec. 2017.

[32] P. J. W. Pouwels and J. Frahm, "Regional metabolite concentrations in human brain as determined by quantitative localized proton MRS," Magn. Reson. Med., vol. 39, no. 1, pp. 53–60, Jan. 1998.