Çoklu-Kontrast MRG'de Kanal-Değişim-Ağı ile Görüntü Sentezi
Multi-Contrast MRI Synthesis with Channel-Exchanging-Network

Onat Dalmaz†*, Idil Aytekin†*, Salman Ul Hassan Dar†, Aykut Erdem‡, Erkut Erdem§ and Tolga Cukur†
†Dept. of Electrical and Electronics Engineering, Bilkent University, Ankara
‡Dept. of Computer Engineering, Koc University, Istanbul, Ankara
§Dept. of Computer Engineering, Hacettepe University, Ankara, Turkey

***Özetçe* —Manyetik rezonans görüntüleme (MRG), yüksek yumuşak doku kontrastına sahip ve müdahelesiz bir medikal görüntüleme yöntemi olması sebebiyle birçok hastalığın tanısında kullanılır. MR sinyal seviyeleri, dokuların kimyasal yapısına göre değişen T1, T2 ve PD parametrelerine göre farklılık gösterir. Bununla birlikte, uzun tarama süreleri, çoklu kontrastlardan görüntü almayı sınırlayabilir veya çoklu kontrastlardan görüntüler alınırsa, kontrastlar gürültülü olur. MRG'nin bu sınırlamasının üstesinden gelmek için çok kontrastlı sentez kullanılabilir. Bu çalışmada, çoklu kontrast MRG'de görüntü sentezi için Kanal-Değişim-Ağı'na (CEN) dayalı bir derin öğrenme yöntemi önerilmektedir. Deneyler için IXI veri seti kullanılmıştır. CEN'e dayalı önerilen model, CNN'lere ve GAN'lara dayalı alternatif yöntemlerle karşılaştırılmıştır. Sonuçlar, önerilen modelin rakip yöntemlere göre üstün performans gösterdiğini göstermektedir.**

***Anahtar Kelimeler*—multimodal füzyon, kanal-değişim-ağı, çoklu-kontrast görüntü sentezi, derin öğrenme.**

***Abstract*—Magnetic resonance imaging (MRI) is used in many diagnostic applications as it has a high soft-tissue contrast and is a non-invasive medical imaging method. MR signal levels differs according to the parameters T1, T2 and PD that change with respect to the chemical structure of the tissues. However, long scan times might limit acquiring images from multiple contrasts or if the multi-contrasts images are acquired, the contrasts are noisy. To overcome this limitation of MRI, multi-contrast synthesis can be utilized. In this paper, we propose a deep learning method based on Channel-Exchanging-Network (CEN) for multi-contrast image synthesis. Demonstrations are provided on IXI dataset. The proposed model based on CEN is compared against alternative methods based on CNNs and GANs. Our results show that the proposed model achieves superior performance to the competing methods.**

***Keywords*—multimodal fusion, channel-exchanging-network, multi-contrast image synthesis, deep learning.**

I. INTRODUCTION

Magnetic resonance imaging (MRI) is used in many diagnostic applications as it has a high soft-tissue contrast and is a non-invasive medical imaging method. MR signal levels differs according to the parameters T1, T2 and PD that change with respect to the chemical structure of the tissues. Thus, images of the same anatomy from multiple contrasts can be obtained via MRI. According to the anatomical differences, the acquired images can be T1-weighted, T2-weighted, or PD-weighted. For instance, T1-weighted brain scans can distinguish white and gray matter better while PD-weighted images distinguish cortical tissue from fluids better. Evaluating the images of the same tissue from different contrasts also increases the accuracy of the clinical diagnosis.

Although multi-contrast images provide more information for clinical diagnosis, the required scan durations are long. For patients at an advanced or very early age, the durations might even be longer. Thus, acquiring images from multiple contrasts might not be possible. Even if the images are acquired, they would be corrupted with noise and have low quality due to patient motion [1]. To overcome this limitation of multi-contrast imaging, multi-contrast image acquisitions should be accelerated without decreasing the quality of the images. A common approach is image reconstruction from under-sampled data to accelerate MR scans via compressed sensing (CS) [2]–[5]. CS enforces sparsity of images in a transform domain to recover from randomly sampled data. Another popular approach is image reconstruction via deep neural networks [6]–[9]. Since deep models require training on fully-sampled acquisitions that can be costly to collect, recent methods have aimed to lower reliance on large, paired training datasets. Domain-transferred models are firstly trained in a source domain where data is abundant, then transferred to the target domain for reconstruction [7].

A fundamental limitation of acceleration by reconstruction is that one must have undersampled acquisitions of the target image for recovery. In many cases, however, high-quality data from the target might not be available due to scan time limitations or artifacts that corrupt the scan. Synthesis is an alternative framework to cope with these cases, where missing or corrupted contrasts are recollected from the set of acquired contrasts in a multi-contrast MRI protocol. Multi-contrast MRI methods typically use one-to-one or many-to-one synthesis procedures according to the input when the target contrast is single. One-to-one approaches [10]–[13] use a single source

*Equal contribution

978-1-6654-5092-8/22/$31.00 ©2022 IEEE

contrast as input and develop a latent representation that is sensitive to the source's unique properties. Many-to-one approaches [13]–[17], on the other hand, accept several distinct sources and develop a shared latent representation that is more sensitive to common characteristics across sources [18]. Apart from many-to-one and one-to-one methods, a joint many-to-one and a combination of several one-to-one streams have been used [18].

In this work, we propose a deep learning method based on Channel-Exchanging-Network (CEN) for multi-contrast image synthesis. The proposed method enables aggregation of information from multiple different source contrasts during many-to-one mapping without introducing additional parameters related to fusion modules. Demonstrations are provided on IXI dataset containing T1-, T2-, and PD-weighted images. The proposed model based on CEN is compared against alternative methods based on CNNs and GANs. Our results show that the proposed model achieves superior performance to the competing methods.

II. Methods

A. Channel-Exchanging-Network (CEN)

CEN [19] is an adaptable, efficient, and parameter-free network. Previous methods were using aggregation and/or alignment for fusion whereas in CEN, the channels between the sub networks are exchanged adaptively. More specifically, the scaling factor ($\gamma_{k,l}$) of Batch Normalization is used to determine the relevance of each corresponding channel, and the channels with approximately zero factors of each modality are replaced by the mean of other modalities. Because it is dynamically regulated by scaling factors set by the training, this fashion of information exchanging is parameter-free and robust. Furthermore, to ensure intra-modal processing, only directed channel exchanging within a specific range of channels in each modality is enabled.

The overall optimization problem of the network is

$$\min_{f_{1:K}} \frac{1}{N}\sum_{i=1}^{N}\mathcal{L}\left(\sum_{k=1}^{K}\alpha_k f_k(x^i), y^i\right) + \lambda\sum_{k=1}^{K}\sum_{l=1}^{L}|\hat{\gamma}_{k,l}|, \quad (1)$$

such that

$$\sum_{k=1}^{K}\alpha_k = 1$$

where y^i is the target, K is the number of modalities, N is the batch size, $f_k(x^i)$ is the sub-network that exchanges channels and fuses multimodal message, $\hat{\gamma}_{k,l}$ is the portion of the scaling factor $\gamma_{k,l}$ that is l_1 norm penalized for sparsity, α_k is the decision score of the ensemble outputs which are trained via a softmax output. $\sum_{k=1}^{K}\alpha_k f_k(x^i)$ represents the ensemble of f_k with the decisions score α_k. Via Eq. 1, to describe the particular statistics of each modality, a parameter-free information fusion over modalities is performed while preserving the each subnetwork's propagation.

B. Relation between Batch Normalization Scaling Factor and Channel Exchanging

The scaling factor the $l^{th} layer$ feature maps of the k^{th} subnetwork and c^{th} channel, $\gamma_{k,l,c}$, assesses the relation between the input $x_{k,l,c}$ and the output. If $\gamma_{k,l,c}$ goes to zero, $x_{k,l,c}$ would lose its affect on the final output, and become redundant. Thus, replacing the channels with close-to-zero scaling factors by the channels of other subnetworks was applied. Hence,

$$x'_{k,l,c} = \begin{cases} \gamma_{k,l,c}\frac{x_{k,l,c}-\mu_{k,l,c}}{\sqrt{\sigma^2{}_{k,l,c}+\epsilon}} + \beta_{k,l,c}, if \gamma_{k,l,c} > \theta \\ \frac{1}{K-1}\sum_{k'\neq k}^{K}\gamma_{k',l,c}\frac{x_{k',l,c}-\mu_{k',l,c}}{\sqrt{\sigma^2{}_{k',l,c}+\epsilon}} + \beta_{k',l,c}, else \end{cases} \quad (2)$$

can be derived where if the scaling factor of the current channel is less than a specified threshold θ, it is replaced by the mean of other channels. Overall, one channel of one modality is substituted with the mean of other modalities if it has minimal influence on the final prediction. Each modality is fed into the nonlinear activation layer which is continued by the convolutions in the following layer, then Eq. 2 is applied to each of them. When a channel is replaced, its gradients are detached and the back-propagation is applied through the new channels.

During the implementation of CEN, each channel is divided into K equal parts. Channel exchanging is performed only for different sub parts of varying modalities. The scaling factors allowed to be replaced are represented by $\hat{\gamma}_{k,l}$. Furthermore, all subnetworks f_k share all of their parameters (including convolutional filters) except the Batch Normalization layers. This method of sharing is followed to reduce the complexity of the network and increase the power of the generalization.

C. Experimental Details

IXI (https://brain-development.org/ixi-dataset/) dataset was used. Specifically, T1-, T2- and PD-weighted images from 53 subjects were used. A total of 2780 images were used for training, and 2165 images were used for testing. During training and testing, approximately 100 axial cross-sections were utilized. Training images were normalized with a mean of 0.5 and a standard deviation of 0.5. Then, they are resized to 256x256. The number of epochs used for training was 80. For synthesizing T1 from T2- and PD-weighted images, learning rate for generator and discriminator was chosen as 0.0050. γ for $L1$ loss was 36.56. λ for L1 norm on Batch Normalization scales was set to 7.667, and threshold for slimming Batch Normalizations was set to 0.0298. For synthesizing T2 from T1- and PD-weighted images, learning rate for generator and discriminator was chosen as 0.001. γ for $L1$ loss was 116.37. λ for L1 norm on Batch Normalization scales was set to 0.0009, and threshold for slimming Batch Normalizations was set to 0.0061. For synthesizing PD from T1- and T2-weighted images, learning rate for generator and discriminator was chosen as 0.0019. γ for $L1$ loss was 111.73, λ for L1 norm on Batch Normalization scales was set to 0.0015, and threshold for slimming Batch Normalizations was set to 0.0026. For all modalities, Adam optimizer for the Generator and Discriminator was employed with $\beta_1 = 0.5$ and $\beta_2 = 0.999$.

III. Results

The proposed CEN method was compared with pGAN$_{many}$ method [10] and pix2pix [20] on the IXI dataset. 3 different tasks were used to evaluate the competing methods: synthesizing T1-weighted images from T2- and

PD-weighted images (T2, PD $\rightarrow$ T1), synthesizing T2-weighted images from T1- and PD-weighted images (T1, PD $\rightarrow$ T2), and synthesizing PD-weighted imaged from T1- and T2-weighted images (T1, T2 $\rightarrow$ PD).

The PSNR and SSIM values of the proposed method and the competing method are showed in Table 1. It can be seen that the proposed method performed better than the competing methods except the SSIM value of pix2pix in synthesizing PD-weighted images from T1- and T2-weighted images. The proposed method obtained 3.07 higher PSNR than PGAN and 3.52 higher PSNR than pix2pix, and 13% higher SSIM than PGAN and 4% higher SSIM than pix2pix in synthesizing T1-weighted images from T2- and PD-weighted images. In synthesizing T2-weighted images from T1- and PD-weighted images, the proposed technique obtained 1.27 higher PSNR than pGAN and 1.69 higher PSNR than pix2pix, 9% higher SSIM than pGAN approximately as well as pix2pix. For synthesizing PD-weighted images from T1- and T2-weighted images, the proposed model obtained 1.92 higher PSNR than pGAN and 2.76 higher PSNR than pix2pix, 6% higher SSIM than pGAN. However, pix2pix obtained 1% SSIM higher than the proposed model. Overall, the higher performance of the proposed method against pGAN is demonstrated in Figure 1.

TABLE I: QUALITY OF SYNTHESIS ON THE IXI DATASET

Models	T2, PD $\rightarrow$ T1		T1, PD $\rightarrow$ T2		T1, T2 $\rightarrow$ PD	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
pGAN many	28.80 $\pm$1.09	0.940 $\pm$0.013	34.04 $\pm$1.18	0.964 $\pm$0.006	33.09 $\pm$1.09	0.967 $\pm$0.005
pix2pix	28.35 1.24	0.949 0.016	33.62 1.31	0.973 0.009	32.25 1.24	**0.974 $\pm$0.006**
CEN	**31.870 $\pm$2.55**	**0.953 $\pm$0.026**	**35.314 $\pm$2.14**	**0.973 $\pm$0.730**	**35.01 $\pm$2.31**	0.973 0.011

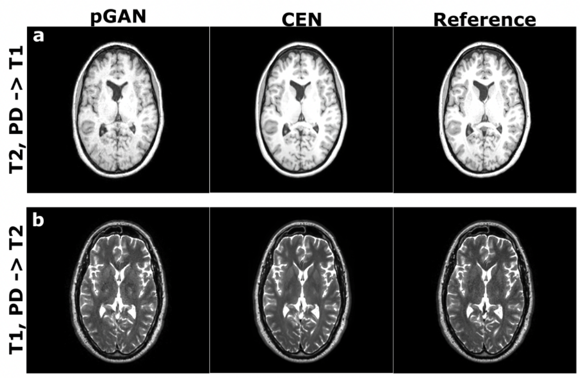

Figure 1: Proposed approach for three modalities. Synthesized images for (a) T1-weighted images from T2- and PD-weighted images, (b) T2-weighted images from T1- and PD-weighted images, and (c) PD-weighted imaged from T1- and T2-weighted images.

IV. DISCUSSION

In this study, a multiple-contrast MRI synthesis method was proposed. The proposed model was based on Channel-Exchanging-Networks (CEN) where the channels of each modality with the scaling factors of Batch Normalization that are close to zero are replaced by the mean of other modalities. This exchange provides a parameter-free and adaptable network. The proposed method was applied on the IXI dataset and the results were analyzed. As a result of the examinations, it was observed that the proposed method provides more successful synthesis performance compared to the competing methods. The proposed method can be further developed with new arrangements on the model in the future. The combination of multiple inputs can be optimized, as in the article mustGAN [18] developed for synthesizing multiple contrast images. It has been seen that the performance of deep learning methods on MR images improves with the transfer learning method, and the synthesis performance can be increased with a similar pre-training [7].

REFERENCES

[1] K. Krupa and M. Bekiesińska-Figatowska, "Artifacts in magnetic resonance imaging," *Polish journal of radiology*, vol. 80, p. 93, 2015.

[2] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial mri with multiple coils. iterative image reconstruction using a total variation constraint," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 57, no. 6, pp. 1086–1098, 2007.

[3] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse mri: The application of compressed sensing for rapid mr imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[4] T. Cukur, M. Lustig, and D. G. Nishimura, "Improving non-contrast-enhanced steady-state free precession angiography with compressed sensing," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 61, no. 5, pp. 1122–1131, 2009.

[5] M. Shahdloo, E. Ilicak, M. Tofighi, E. U. Saritas, A. E. Çetin, and T. Çukur, "Projection onto epigraph sets for rapid self-tuning compressed sensing mri," *IEEE transactions on medical imaging*, vol. 38, no. 7, pp. 1677–1689, 2018.

[6] S. U. H. Dar, M. Yurt, M. Shahdloo, M. E. Ildız, and T. Çukur, "Synergistic reconstruction and synthesis via generative adversarial networks for accelerated multi-contrast mri," *arXiv preprint arXiv:1805.10704*, 2018.

[7] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, "A transfer-learning approach for accelerated mri using deep neural networks," *Magnetic resonance in medicine*, vol. 84, no. 2, pp. 663–685, 2020.

[8] M. Yurt, M. Özbey, S. U. H. Dar, B. Tınaz, K. K. Oğuz, and T. Çukur, "Progressively volumetrized deep generative models for data-efficient contextual learning of mr image recovery," *arXiv preprint arXiv:2011.13913*, 2020.

[9] O. Dalmaz, M. Yurt, and T. Çukur, "Resvit: Residual vision transformers for multi-modal medical image synthesis," *arXiv preprint arXiv:2106.16031*, 2021.

[10] S. U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, and T. Çukur, "Image synthesis in multi-contrast mri with conditional generative adversarial networks," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2375–2388, 2019.

[11] D. Abramian and A. Eklund, "Refacing: reconstructing anonymized facial features using gans," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 1104–1108.

[12] C. Bowles, C. Qin, C. Ledig, R. Guerrero, R. Gunn, A. Hammers, E. Sakka, D. A. Dickie, M. V. Hernández, N. Royle *et al.*, "Pseudo-healthy image synthesis for white matter lesion segmentation," in *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, 2016, pp. 87–96.

[13] B. E. Dewey, C. Zhao, J. C. Reinhold, A. Carass, K. C. Fitzgerald, E. S. Sotirchos, S. Saidha, J. Oh, D. L. Pham, P. A. Calabresi *et al.*, "Deepharmony: A deep learning approach to contrast harmonization across scanner changes," *Magnetic resonance imaging*, vol. 64, pp. 160–170, 2019.

[14] S. U. Dar, M. Yurt, M. Shahdloo, M. E. Ildız, B. Tınaz, and T. Çukur, "Prior-guided image reconstruction for accelerated multi-contrast mri via generative adversarial networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1072–1087, 2020.

[15] S. Olut, Y. H. Sahin, U. Demir, and G. Unal, "Generative adversarial training for mra image synthesis using multi-contrast mri," in *International workshop on predictive intelligence in medicine*. Springer, 2018, pp. 147–154.

[16] B. E. Dewey, C. Zhao, A. Carass, J. Oh, P. A. Calabresi, P. Zijl, J. L. Prince *et al.*, "Deep harmonization of inconsistent mr data for consistent volume segmentation," in *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, 2018, pp. 20–30.

[17] A. Sharma and G. Hamarneh, "Missing mri pulse sequence synthesis using multi-modal generative adversarial network," *IEEE transactions on medical imaging*, vol. 39, no. 4, pp. 1170–1183, 2019.

[18] M. Yurt, S. U. Dar, A. Erdem, E. Erdem, K. K. Oguz, and T. Çukur, "mustgan: Multi-stream generative adversarial networks for mr image synthesis," *Medical image analysis*, vol. 70, p. 101944, 2021.

[19] Y. Wang, W. Huang, F. Sun, T. Xu, Y. Rong, and J. Huang, "Deep multimodal fusion by channel exchanging," *Advances in Neural Information Processing Systems*, vol. 33, pp. 4835–4845, 2020.

[20] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.