Üretken Çekişmeli Ağlar ile Medikal Görüntü Sentezi için Atlamalı Bağlantılar
Skip Connections for Medical Image Synthesis with Generative Adversarial Networks

Muhammad Usama Mirza[1], *Onat Dalmaz*[2], *Tolga Çukur*[3]

[1,2,3]Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey
mumirza@ee.bilkent.edu.tr, onat@ee.bilkent.edu.tr, cukur@ee.bilkent.edu.tr

2022 30th Signal Processing and Communications Applications Conference (SIU) | 978-1-6654-5092-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/SIU55565.2022.9864939

***Özetçe* —Manyetik Rezonans Görüntüleme (MRI), ayrıntılı anatomik görüntüler üretmek için kullanılan bir görüntüleme tekniğidir. Çoklu kontrastlı MRI görüntüleri elde etmek, hastayı hareketsiz kalmaya zorlayan uzun tarama süreleri gerektirebilir. Elde edilen kontrastlardan, elde edilmemiş kontrastların sentezlenmesiyle tarama süreleri azaltılabilir. Son yıllarda, bire bir eşleme yöntemleri kullanarak kontrastları sentezlemek için derin üretken çekişmeli ağlar kullanılmaktadır. Daha derin ağlar daha karmaşık işlevleri çözebilir, ancak aşırı öğrenme ve kaybolan gradyanlar gibi sorunlar nedeniyle performansları düşebilir. Bu çalışmada, artan karmaşıklıkla birlikte performanstaki düşüşün üstesinden gelmek için üretken modellere atlamalı bağlantılar eklemeyi öneriyoruz. Bu, ağın modeldeki gereksiz parametreleri atlamasına izin verecektir. Sonuçlarımız, atlama bağlantılarını entegre ederek bire bir görüntü sentezinde performansta bir artış olduğunu göstermektedir.**

Anahtar Kelimeler—Medikal görüntü sentezi, Manyetik Rezonans Görüntüleme (MRG), Çoklu kontrast MRG, Üretken Çekişmeli Ağlar, Atlamalı Bağlantılar

***Abstract*—Magnetic Resonance Imaging (MRI) is an imaging technique used to produce detailed anatomical images. Acquiring multiple contrast MRI images requires long scan times forcing the patient to remain still. Scan times can be reduced by synthesising unacquired contrasts from acquired contrasts. In recent years, deep generative adversarial networks have been used to synthesise contrasts using one-to-one mapping. Deeper networks can solve more complex functions, however, their performance can decline due to problems such as overfitting and vanishing gradients. In this study, we propose adding skip connections to generative models to overcome the decline in performance with increasing complexity. This will allow the network to bypass unnecessary parameters in the model. Our results show an increase in performance in one-to-one image synthesis by integrating skip connections.**

Keywords—Medical image synthesis, Magnetic resonance imaging (MRI), Multi-contrast MRI, Generative adversarial network, Skip connections

I. INTRODUCTION

Magnetic resonance imaging (MRI) can capture images of a given anatomy under various different contrasts by altering its pulse sequences. Different contrasts can then help tissue delineation and increase cumulated diagnostic information. For instance, T1-weighted images are more suited for distinguishing gray and white matter in the brain, whereas PD-weighted images are more suited for separating fluids and cortical tissues. However, acquiring images of multiple contrasts is challenging due to prolonged scan times and increases risk of motion artifacts.

A common approach for limiting scan durations during multi-contrast MRI is based on image synthesis [1], [2], [3], where images of one or more source contrasts are used to predict images of a target contrast. Depending on the number of sources available, two main categories are one-to-one methods [4], [5], [3], and many-to-one methods [6].

Earlier methods to solve one-to-one synthesis have approached it as a sparse dictionary reconstruction problem [7], [8]. Recent studies have instead used deep neural networks based on convolutional layers that process the entire image as opposed to patches. Encoder-decoder architectures have emerged as a gold-standard for synthesis models. Further improvement has been made by introducing an encoder-decoder setup [9]. The encoder is used to embed the image onto a latent representation and the decoder then recovers the target image from the latent representation. Recent models have introduced adversarial loss to create sharper images by learning high frequency features in the target images. One such adversarial method is pGAN [3] which uses pixel-wise and perceptual losses to improve performance.

Here, we propose a technique to improve the performance of one-to-one synthesis models by incorporating skip connections. Skip connections are connections which bypass certain layers in the network. When networks become too deep, problems such as overfitting, vanishing gradients and high computation cost arise. Deeper layers in the network are unable to learn simple features in the network and not only become useless, but cause a decline in the performance of the model [10]. Skip connections can overcome these problems by reducing the effect of layers that are deteriorating performance.

Magnetic resonance imaging (MRI) can capture images of a given anatomy under various different contrasts by altering its pulse sequences. Different contrasts can then help tissue delineation and increase cumulated diagnostic information. For instance, T1-weighted images are more suited for distinguishing gray and white matter in the brain, whereas PD-weighted images are more suited for separating fluids and cortical tissues. However, acquiring images of multiple contrasts is challenging due to prolonged scan times and increases risk of motion artifacts.

II. METHODS

A. Image Synthesis

Generative Adversarial Networks (GAN) [11] consist of two sub-networks; generator and discriminator. In image synthesis, the generator learns to generate the target image from the input image, while the discriminator learns to differentiate between the real image and the generated image. Both the generator and discriminator are trained simultaneously and upon convergence, the generator can generate realistic images. Recent studies have demonstrated the effectiveness of GANs in image-to-image mapping [12], [13].

In the pGAN model [3], a convolutional GAN model was used, where both the generator and discriminator are CNN based. The generator of the pGAN has a ResNet backbone with an encoder, a bottleneck and a decoder. The encoder and decoder are comprised of three convolutional layers each while the bottleneck consists of nine residual blocks [14]. The generator takes in an image from the source MRI contrast as the input and outputs the corresponding image in the target contrast.

In order to further improve the performance, we added skip connections between the residual blocks in the bottleneck of the generator. We tested different configurations of skip connections and observed better results upon adding skip connections only to the deeper layers of the network. The skip connections were added in a fashion similar to the DenseNet architecture [15]. In this architecture, every layer is connected to every other layer in a feed-forward fashion. However, adding skip connections to the earlier layers of the network resulted in a decline in performance, thus, feed-forward connections were only added to the layers at the end of the bottleneck. This configuration is shown in Figure 1.

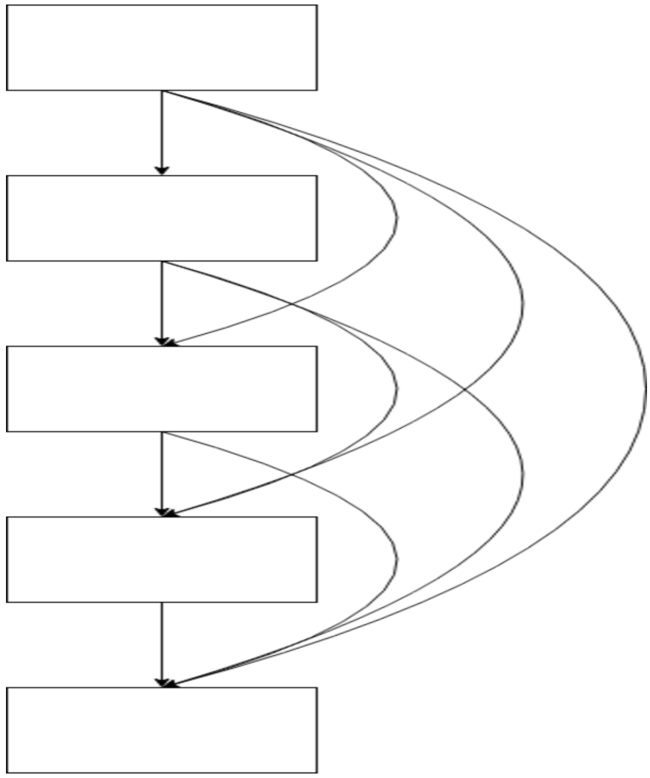

Figure 1: A schematic of dense skip connections among bottleneck layers. The boxes represent the layers and the arrows represent the data flow. The dense skip connections allow the model to finely control the affect of a layer on the model's output and allows unnecessary layers and connections to be bypassed if necessary.

B. Datasets

The model was trained on the IXI dataset (http://brain-development.org/ixi-dataset/) which contains data from healthy subjects. The dataset contains T1, T2 and PD weighted images. The model was trained to learn one-to-one mappings between two MRI contrasts. To optimize training and remove bias, the dataset was normalized to ensure comparable ranges of voxel intensities for all subjects. This normalization was performed separately for each contrast and each subject. The images were brought in the range of [-1, 1]. This specific range was selected to match the output range of the hyperbolic tangent activation function in the final layer of the model.

T1, T2 and PD weighted images taken from 53 subjects were used, 25 of which were used for training, 10 for validation and 18 for testing. The following parameters were used acquiring the images. T1-weighted images: TE = 4.603 ms, TR = 9.813 ms, flip angle = 8°, spatial resolution = 0.94 x 0.94 x 1.2 mm^3 matrix size = 256 x 256 x 150. T2-weighted images: TE = 100 ms, TR = 8178.34 ms, flip angle = 90°, spatial resolution = 0.94 × 0.94 × 1.2 mm^3, matrix size = 256 × 256 × 150. PD-weighted images: TE = 8 ms, TR = 8178.34 ms, flip angle = 90°, spatial resolution = 0.94 × 0.94 × 1.2 mm^3, matrix size = 256 × 256 × 150.

C. Model Training

The pGAN model [3] adopts its generator architecture from [16] and discriminator architecture from [13]. In complex networks such as GANs, it is difficult to tune hyperparameters, therefore, it is common to perform one-fold cross-validation or even directly adopt hyperparameters from published work. In pGAN, parameters were selected using the validation set. To test the effect of skip connections, the model was trained twice on each dataset, once with and once without skip connections. Each time the model was trained for 100 epochs. In the first 50 epochs the learning rate was set to 0.0001 and kept constant. In the last 50 epochs, the learning rate was linearly decreased from 0.0001 to 0. Training for more than 100 epochs resulted in a decline in performance of the network. The discriminator loss function was halved during each iteration to slow down the learning process of the discriminator. Adam Optimizer [17] was used and the decay rates for the first and second moments of gradient estimates were set as $\beta_1 = 0.5$ and

$\beta_2 = 0.999$ respectively. Instance normalization was applied [18]. All weights were initialized using normal distribution with 0 mean and 0.02 std. The model has a total of 14 million parameters.

The mean Peak Signal to Noise Ratio (PSNR) and L1 Average Loss were used as metrics to judge the performance of the model. PSNR is a direct and robust metric based solely on image quality, while, network loss includes an adversarial component is prone to introducing instabilities. After training the model with and without skip connections on different MR contrasts, the mean PSNR and L1 loss was compared to check the difference in performance.

III. RESULTS

We demonstrated the performance of the pGAN in learning synthesis models for multi-contrast MRI, with and without skip connections. We used T1, T2 and PD weighted images and trained our network to perform one-to-one mapping between separate source-target contrast pairs. For each synthesis task, the model was first trained without skip connections and then with skip connections. All hyperparameters and optimization procedures were kept identical across models. While training each model, we computed the mean PSNR value as well as the L1 average loss between the generated image and the real image. Our results are listed in Tables I and II.

It can be seen that adding skip connections to the network improves the performance of the model in most cases. The mean PSNR has improved in all the cases except T1 to T2 weighted images, while for L1 loss, the performance is better for all the cases except PD to T1.

Table I: MEAN PSNR FOR VARIOUS ONE-TO-ONE MAPPINGS BETWEEN SOURCE AND TARGET MRI CONTRASTS.

PSNR

	With Skip Connections	Without Skip Connections
PD→T1	**28.272**	28.168
T1→PD	**27.178**	27.022
T1→T2	27.867	**27.912**
T2→T1	**27.666**	27.659
PD→T2	**32.957**	32.899
T2→PD	**32.048**	31.997

Table II: L1 AVERAGE LOSS FOR VARIOUS ONE-TO-ONE MAPPINGS BETWEEN SOURCE AND TARGET MRI CONTRASTS.

L1 Average Loss

	With Skip Connections	Without Skip Connections
PD→T1	0.01183	**0.01168**
T1→PD	**0.01427**	0.01534
T1→T2	**0.01356**	0.01373
T2→T1	**0.01215**	0.01256
PD→T2	**0.00821**	0.00830
T2→PD	**0.01000**	0.01003

IV. DISCUSSION & CONCLUSION

Skip connections were introduced for multi-contrast MRI based on conditional generative adversarial networks. Previous methods, consisted of a single path for data flow through the generator network which caused them to suffer when the networks became too deep. Skip connections allow the multiple paths for the data to flow and enable certain layers to be bypassed based on the learning conditions of the network. This makes the network more versatile and enables it to reduce the effect of layers that become unnecessary to the performance of the network. The addition of skip connections has improved the results of the network and generated better images.

Synthesizing different contrasts from a single contrast will reduce the overall diagnosis time and make it easier for patients to get scanned, specifically for young children and elderly patients. It will also reduce the cost of exams due to repeated acquisitions. Therefore, improvement in image synthesis will greatly help improve the diagnostic information available in multi-contrast MRI.

REFERENCES

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

[2] M. Yurt, S. U. H. Dar, A. Erdem, E. Erdem, and T. Çukur, "mustgan: Multi-stream generative adversarial networks for mr image synthesis," 2019.

[3] S. U. H. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, and T. Çukur, "Image synthesis in multi-contrast mri with conditional generative adversarial networks," 2018.

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

[5] C. Bowles, C. Qin, C. Ledig, R. Guerrero, R. Gunn, A. Hammers, E. Sakka, D. A. Dickie, M. V. Hernández, N. Royle, J. Wardlaw, H. Rhodius-Meester, B. Tijms, A. W. Lemstra, W. van der Flier, F. Barkhof, P. Scheltens, and D. Rueckert, "Pseudo-healthy image synthesis for white matter lesion segmentation," in *Simulation and Synthesis in Medical Imaging* (S. A. Tsaftaris, A. Gooya, A. F. Frangi, and J. L. Prince, eds.), (Cham), pp. 87–96, Springer International Publishing, 2016.

[6] 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.

[7] Y. Huang, L. Shao, and A. F. Frangi, "Simultaneous super-resolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5787–5796, 2017.

[8] Y. Huang, L. Shao, and A. F. Frangi, "Cross-modality image synthesis via weakly coupled and geometry co-regularized joint dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 3, pp. 815–827, 2018.

[9] V. Sevetlidis, M. V. Giuffrida, and S. A. Tsaftaris, "Whole image synthesis using a deep encoder-decoder network," in *Simulation and Synthesis in Medical Imaging* (S. A. Tsaftaris, A. Gooya, A. F. Frangi, and J. L. Prince, eds.), (Cham), pp. 127–137, Springer International Publishing, 2016.

[10] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," 2015.

[11] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2018.

[13] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," 2020.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015.

[15] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” 2018.

[16] J. Johnson, A. Alahi, and L. Fei-Fei, “Perceptual losses for real-time style transfer and super-resolution,” 2016.

[17] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2017.

[18] D. Ulyanov, A. Vedaldi, and V. Lempitsky, “Instance normalization: The missing ingredient for fast stylization,” 2017.

[19] Y. Korkmaz, S. U. Dar, M. Yurt, M. Özbey, and T. Çukur, “Unsupervised mri reconstruction via zero-shot learned adversarial transformers,” 2022.

[20] S. U. H. Dar, M. Özbey, A. B. Çatlı, and T. Çukur, “A transfer-learning approach for accelerated mri using deep neural networks,” 2019.