



# AutoGAN-Synthesizer: Neural Architecture Search for Cross-Modality MRI Synthesis

Xiaobin Hu<sup>1,2,3</sup>(✉), Ruolin Shen<sup>1</sup>, Donghao Luo<sup>3</sup>, Ying Tai<sup>3</sup>, Chengjie Wang<sup>3</sup>,  
and Bjoern H. Menze<sup>1,2</sup>

<sup>1</sup> Department of Computer Science, Technical University of Munich,  
Munich, Germany  
xiaobin.hu@tum.de

<sup>2</sup> Department of Quantitative Biomedicine, University of Zurich, Zürich, Switzerland

<sup>3</sup> Tencent Youtu Lab, Shanghai, China

**Abstract.** Considering the difficulty to obtain complete multi-modality MRI scans in some real-world data acquisition situations, synthesizing MRI data is a highly relevant and important topic to complement diagnosis information in clinical practice. In this study, we present a novel MRI synthesizer, called *AutoGAN-Synthesizer*, which automatically discovers generative networks for cross-modality MRI synthesis. Our AutoGAN-Synthesizer adopts gradient-based search strategies to explore the generator architecture by determining how to fuse multi-resolution features and utilizes GAN-based perceptual searching losses to handle the trade-off between model complexity and performance. Our AutoGAN-Synthesizer can search for a remarkable and light-weight architecture with **6.31 Mb** parameters only occupying *12* GPU hours. Moreover, to incorporate richer prior knowledge for MRI synthesis, we derive K-space features containing the low- and high-spatial frequency information and incorporate such features into our model. To our best knowledge, this is the first work to explore AutoML for cross-modality MRI synthesis, and our approach is also capable of tailoring networks given either different multiple modalities or just a single modality as input. Extensive experiments show that our AutoGAN-Synthesizer outperforms the state-of-the-art MRI synthesis methods both quantitatively and qualitatively. The code are available at <https://github.com/HUuxiaobin/AutoGAN-Synthesizer>.

**Keywords:** Cross-modality MRI synthesis · Generative adversarial networks · Neural architecture search

---

X. Hu and R. Shen—Equal contribution.

---

**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/978-3-031-16446-0\\_38](https://doi.org/10.1007/978-3-031-16446-0_38).

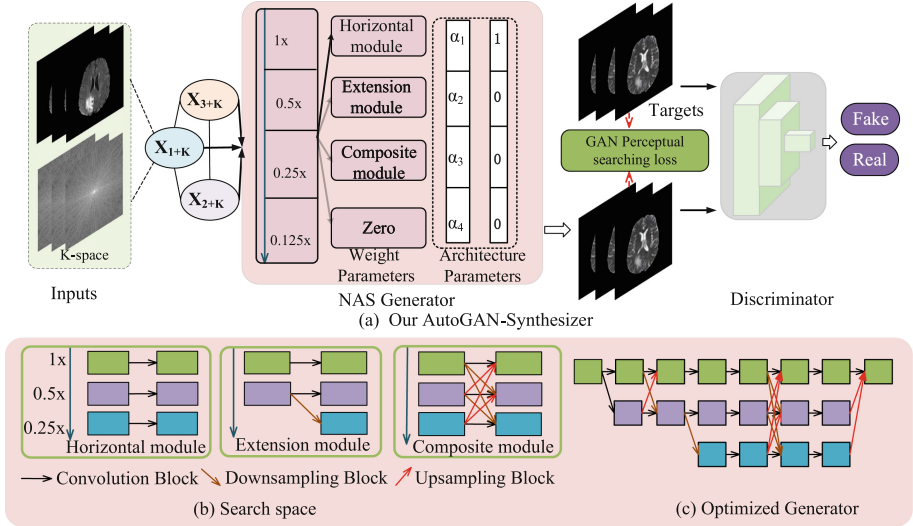
© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022  
L. Wang et al. (Eds.): MICCAI 2022, LNCS 13436, pp. 397–409, 2022.  
[https://doi.org/10.1007/978-3-031-16446-0\\_38](https://doi.org/10.1007/978-3-031-16446-0_38)

## 1 Introduction

Magnetic Resonance Imaging (MRI) has currently become prominent in neurology with the use of high-field scanners, given that MRI contrasting agent is less likely to cause an allergic reaction compared with X-ray or CT scan using iodine-based substances [24]. However, due to modality corruption, incorrect machine settings, allergies to specific contrast agents, and limited available time [7], it is often not guaranteed to obtain a complete set of MRI sequences to provide rich information for clinical diagnosis and therapy. In this regard, the development of cross-modality or cross-protocol MRI synthesis techniques is important to homogenize and “repair” such real-world data collections via efficient data infilling and re-synthesis, and make them accessible for algorithms that require complete data sets as input [2, 10, 23, 30].

Recently a large number of algorithms for medical image synthesis have been proposed with the rapid growth of deep learning techniques [6, 11, 27, 37]. Among them, generative adversarial networks (GANs) with the advantage of recovering an unprecedented level of image realism [15] have achieved significant advancement for medical image synthesis. For example, Ea-GANs [32] incorporated edge information and focused on the textural details of image content structure for cross-modality MRI synthesis. SA-GANS [33] added a sample-adaptive path additionally to learn the relationship of each sample with its neighboring training samples. MM-GAN [5] and mustGAN [34] were designed to deal with multi-modal MRI synthesis with structures capable of fusing latent representations of each input modality. However, these state-of-art methods [5, 7, 10, 13, 32–34, 36, 38] fixed the network architecture for various input modality combinations (e.g., T1, T1-weighted, Flair, T1+Flair, etc.) and ignored the mapping uniqueness between each source domain and target domain pair, and therefore could not reach the optimal solution for all situations using the same network structure.

Inspired by the great potential of neural architecture search (NAS) in computer vision field [8, 17–19, 25, 31, 35, 40], we explore NAS to automatically find an optimal network with fewer computation costs and parameters for different input modalities. Searching for a dedicated MRI synthesizer is essentially promising because the problem nature of using one network for many synthesis tasks caters to the NAS principle of construing one search architecture for multi-jobs. However, how to search the architecture of generative networks according to the different input modalities given to the synthesizer is still unexplored so far. In this paper, we aim to adaptively optimize and construct the neural architecture that is capable of understanding how to extract and merge features according to different input modalities. Specifically, we adopt a GAN-based structure as the backbone of NAS, where the generator of the GAN is searched by gradient-based NAS from the multi-scale module-based search space. Our main contributions of the AutoGAN-Synthesizer are as follows: (1). Aiming at the recovery of realistic texture while constraining model complexity, we propose a GAN-based perceptual searching loss that jointly incorporates the content loss and model complexity. (2). To incorporate richer priors for MRI synthesis, we exploit MRI *K-space* knowledge containing low-frequency (e.g., contrast and brightness) and



**Fig. 1.** (a). Overall architecture of our proposed AutoGAN-Synthesizer. The AutoGAN-Synthesizer contains two parts: 1) A NAS-based generator to adaptively build up an architecture based on input modalities ( $X_{1+k}$ ,  $X_{2+k}$  or  $X_{3+k}$ ).  $X_{i+k}$  represents the input with  $i$  modalities and the corresponding K-space features (denoted as  $k$ ). 2) A discriminator to distinguish between the synthesis and real modality. (b). Generator search space consisting of three modules to capture and fuse the detailed information and the global features from different multi-scale branches: horizontal module, extension module and the composite module. (c). An example of an optimized generator including the three proposed modules.

high-frequency information (e.g., edges and content details), to guide the NAS network to extract and merge features. (3). Considering that the low- and high-resolution of multi-scale networks can capture the global structure and local details respectively, we use a novel multi-scale module-based search space which is specifically designed for multi-resolution fusion. The module-based searching setting is also capable of reducing search time while maintaining the performance. (4). Finally, our searching strategy can produce a light-weight network with **6.31 Mb** parameters from module-based search space only taking *12* GPU hours and achieve state-of-the-art performance. From our knowledge, this is the first work to explore AutoML for cross-modality MRI synthesis tasks.

## 2 Proposed Method

**Motivation:** Most recent-used networks for MRI synthesis usually adopt an encoder-decoder structure [7, 14, 32, 38], which recovers the high-resolution features mainly based on the low-resolution representation received from successive convolutional blocks in the encoder. This latent representation contains only

high-level features and loses lots of detailed information, leading the recovered images neither semantically strong nor spatially precise. Inspired by the fact that the low- and high-resolution branches of multi-scale networks are capable of capturing global structure and preserving the local details respectively [8, 22, 28], we design a generator based on a multi-scale structure including three modules: *horizontal module* connects different resolution input in parallel without any fusion, *extension module* adds a downsampling block to extend a lower-resolution scale and *composite module* fuses cross-resolution representations to exchange information. An overview of our AutoGAN-Synthesizer is shown in Fig. 1. Specifically, the framework contains an adaptive generator constructed by neural architecture search according to input modalities and a typical discriminator to distinguish between predictions and ground truths.

## 2.1 NAS-Based Generator Search

**Generator Search Space:** There exists an open question on how to extract or fuse the features of modalities in a multi-scale generator. To solve this question, we propose three different modules (Fig. 1(b)) to give guidelines for extracting and merging multi-resolution features: namely *horizontal module*, *extension module* and *composite module*. These three modules behave differently to mimic the coarse-to-fine framework and exploit multiple possibilities of multi-scale fusion. Specifically, the horizontal module horizontally connects features via convolution block without feature fusion among multi-scales. As shown in Fig. 1(b), the feature resolution at the same scale keeps identical but is reduced by 0.5 when the scale goes deeper. The extension module extends a lower-resolution scale via a down-sampling block. This connection helps to exploit the high-level priors extracted at the low-resolution scale and simultaneously remain the unchanged resolution at the high-resolution scale. The composite module merges multi-resolution features by skip connection, stride convolution and up-sampling block, which can be summarized as:

$$F_g = \sum \mathcal{M}_{r \rightarrow g}(F_r) \quad (1)$$

where  $r$  is the resolution of the input feature maps while  $g$  is the resolution of the output features.  $F_r$  represents the input feature maps with resolution  $r$  and  $F_g$  denotes the output feature maps with resolution  $g$  after combing all the features from other resolution scales.  $\mathcal{M}_{r \rightarrow g}(\cdot)$  is the mapping function defined as follows:

$$\mathcal{M}_{r \rightarrow g}(F_r) = \begin{cases} F_r & r == g \\ \text{Upsampled } F_r & r < g \\ \text{Downsampled } F_r & r > g \end{cases} \quad (2)$$

Compared with the common fusion scheme [5, 26], that only fuses high-resolution features with the upsampled low-resolution features unidirectionally, this module aggregates the feature fusion via a bidirectional way among multi-resolution representations. Thus, this powerful multi-resolution fusion scheme

catches more spatial information extracted from all resolution scales and therefore is semantically stronger.

The combination of three modules constructs a superior neural architecture which gives guidance on how to extract and merge the features catering to the requirements of different input modalities. An example of an optimized generator could be found in Fig. 1(c). The input modalities are imported into a super-network with two fixed horizontal modules,  $S$  modules selected from horizontal, extension, and composite modules candidates, and a final composite module followed by a  $1 \times 1$  convolutional layer. During the searching process, the progressive structure can gradually add the multi-resolution modules and endow the output with multi-resolution knowledge.

## 2.2 GAN-Based Perceptual Loss Function

In order to recover a realistic and faithful image, we add both the perceptual and pixel-level loss into our generator loss function:

$$\mathcal{L}_{Generator} = \mathcal{L}_{content} + \lambda_{adv}\mathcal{L}_{adv} + \lambda_{complexity}\mathcal{L}_{complexity}, \quad (3)$$

where  $\mathcal{L}_{content}$  is the content loss consisting of pixel-level loss (mean square loss) and texture-level loss (perceptual loss) between the ground-truth and reconstructed images.  $\mathcal{L}_{adv}$  is the adversarial loss based on binary cross-entropy formulation to make the reconstructed image closer to the ground-truths.  $\mathcal{L}_{complexity}$  is the loss term for calculating the model complexity (e.g., FLOPS, consuming time, and model size).

## 2.3 K-space Learning

K-space is the spatial frequency representation of MRI images. Due to the long scan time acquiring MRI images, several MRI reconstruction methods based on under-sampled K-space learning are proposed for fast MRI [1, 9, 14]. Inspired by this, we embed K-space learning into our pipeline to introduce frequency priors of MRI images, which is defined as follows:

$$\hat{x}(k) = \mathcal{F}[x]\{k\} = \int_{R^2} e^{-jk \cdot r} x(r) dr, \quad (4)$$

where  $k \in \mathbb{R}^2$  represents the spatial frequency and  $j^2 = -1$ .  $x(r)$  is the pixel intensity in real space while  $\hat{x}(k)$  is the calculated intensity in frequency domain. K-space is computed according to the input modalities and is used as the input together with the MRI images in real space.

## 2.4 Implementation Details

**Searching Setting:** For each different input modality, we search for a new architecture to give guidance on how to extract and fuse the multi-modality features. First, we train the warm-up stage (ten epochs) to get desirable weights of

convolution layers and then train a searching stage with 200 epochs for optimizing the structure of architecture. For updating the parametric model, we adopt the standard SGD optimizer with the momentum of 0.9 and the learning rate decays from 0.025 to 0.001 by cosine annealing strategy [20]. Besides, to optimize the architecture parameters, Adam optimizer [16] is used with a learning rate of 0.0005. The batch size is 16 by randomly cropping and padding image size  $240 \times 240$ . Overall, the whole searching process consumes 12 h.

**Training Setting:** After finding an architecture, we train this for 500 epochs with batch size 16 and image size  $240 \times 240$ . The Adam optimizer with the learning rate of 0.0005 is adopted. All training experiments are implemented in Pytorch with a Tesla V100.

### 3 Experimental Results

#### 3.1 Experimental Settings

We evaluate the performance of AutoGAN-Synthesizer on one-to-one and multiple-to-one cross-modality MRI synthesis tasks using two public brain MRI datasets: BRATS2018 and IXI. BRATS2018 dataset [3, 4, 21] collects multi-modality MR image sets from patients with brain tumors including four different modalities: native (T1), T1-weighted and contrast-enhanced (T1ce), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR), where each scan has the size of  $240 \times 240 \times 155$ . In this paper, we conduct one-to-one and multi-to-one synthesis tasks on BRATS2018 dataset to show the effectiveness of our method. Following [7], we randomly select 50 low grade glioma (LGG) from total 75 LGG patients as the training set while another unseen 15 patients are selected as the test. Following [7, 32, 34], we also use the public IXI dataset<sup>1</sup> to verify the model generalization. IXI dataset collected multi-modality MR images from healthy subjects at three different hospitals. It is randomly divided into training (25 patients), validation (5 patients), and test patients (10 patients). For each subject, after removing some cases with major artifacts, approximately 100 axial cross sections that contained brain tissue are manually selected.

**Evaluation Metrics:** Following studies [32, 33], three metrics are used to evaluate the quantitative performance: normalized root mean-squared error (NRMSE), peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [29].

---

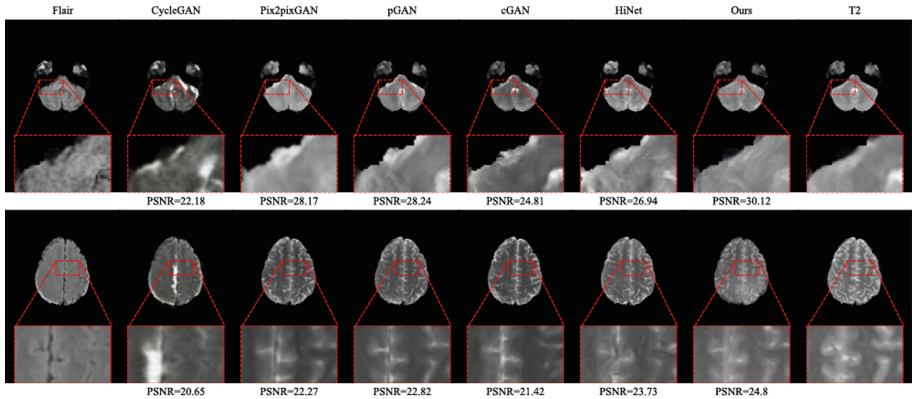
<sup>1</sup> <https://brain-development.org/ixi-dataset/>.

**Table 1.** Quantitative results of Flair-T2 (BRATS2018 dataset) and T1-T2 (IXI dataset) MRI cross-modality synthesis tasks.

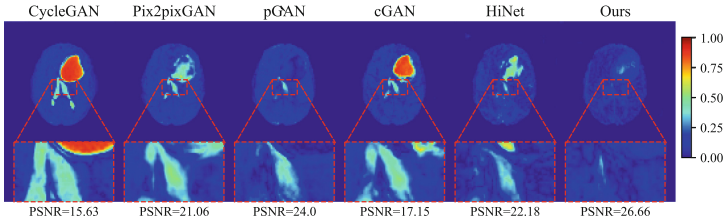
Methods	Flair-T2 (BRATS2018 dataset)			T1-T2 (IXI dataset)			Model size
	PSNR	SSIM	NRMSE	PSNR	SSIM	NRMSE	Params(M)
CycleGAN [39]	20.70 ± 2.37	0.83 ± 0.05	0.45 ± 0.14	23.03 ± 1.01	0.74 ± 0.05	0.48 ± 0.05	11.38
Pix2pix [12]	24.64 ± 3.97	0.88 ± 0.04	0.34 ± 0.19	25.70 ± 1.61	0.84 ± 0.05	0.32 ± 0.06	54.41
pGAN [7]	25.20 ± 4.34	0.89 ± 0.04	0.34 ± 0.22	26.62 ± 1.72	0.85 ± 0.05	0.29 ± 0.06	11.36
cGAN [7]	23.67 ± 3.99	0.87 ± 0.04	0.35 ± 0.16	23.77 ± 1.80	0.77 ± 0.05	0.31 ± 0.06	11.37
Hi-Net [38]	24.46 ± 3.56	0.87 ± 0.04	0.32 ± 0.15	25.63 ± 1.50	0.83 ± 0.05	0.33 ± 0.06	<b>3.87</b>
Ours	<b>25.54 ± 3.91</b>	<b>0.90 ± 0.03</b>	<b>0.30 ± 0.17</b>	<b>27.37 ± 1.81</b>	<b>0.86 ± 0.04</b>	<b>0.29 ± 0.05</b>	6.30

### 3.2 Comparisons with State-of-the-Art Methods

To verify the performance of our AutoGAN-Synthesizer, we compare it with five recent state-of-the-art methods: CycleGAN [39], Pix2pix [12], pGAN and cGAN [7], and Hi-Net [38]. To ensure a fair comparison with state-of-the-art methods, we train all networks on the same dataset by the open-source implementations as well as the recommended hyper-parameters from authors.



**Fig. 2.** Qualitative results of FLAIR  $\rightarrow$  T2 synthesis experiments on glioma patients in BRATS2018 dataset. Compared with other state-of-art results, our synthetic images recover favorable tissue contrast, tumor, and anatomy knowledge which have great potential in clinical diagnoses and treatments.



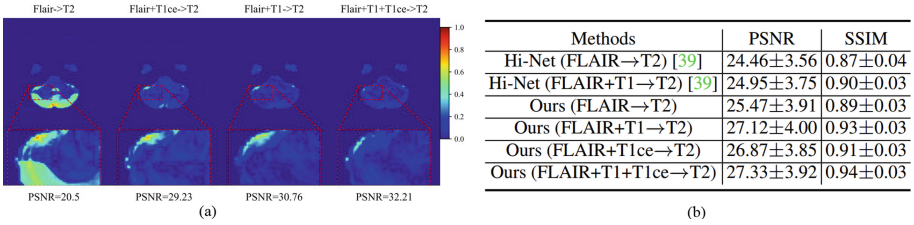
**Fig. 3.** Visual performance of synthetic T2 modality difference maps compared with other state-of-art methods on BRATS2018 dataset.

**One-to-One Cross-Modality MRI Synthesis Tasks:** We focus on synthesizing T2 contrasts that are complementary to T1 contrasts, and offer better information for investigating fluid-filled structures within tissue. The experimental results for one-to-one synthesis tasks are listed in Table 1. For FLAIR-T2 cross-modality synthesis task, Table 1 shows that our AutoGAN achieves better performance than other cutting-edge methods on three metrics. Figures 2 and 3 show the qualitative comparison between our proposed AutoGAN and other five state-of-art methods on BRATS2018 dataset. The difference maps are generated based on the pixel intensity and visualized in the type of heat maps. It can be seen that our synthetic images have clearer details in the zoomed rectangles, and also preserve favorable tissue contrast, tumor, and anatomy knowledge which have great potentials in clinical diagnoses and treatments. Overall, our methods could reach higher fidelity with the target images and our method can search for satisfactory synthesis networks which are better than manually designed architectures. The superiority of AutoGAN is mainly attributed to our module-based search space, which can well exploit the information fusion between the low- and high-resolution features. As shown in Table 1, the quantitative results on the IXI dataset also imply that our AutoGAN achieve better generalization than other methods.

**Model Complexity:** It can be seen from Table 1 that our AutoGAN achieves SOTA performance only using very light-weight network architecture with 6.30 Mb parameters, which is nearly half of the other manually-designed networks (around 11 Mb) [7, 39].

**Multiple-to-One Cross-Modality MRI Synthesis Tasks.** To verify the effectiveness of our method on multiple-to-one tasks, we conduct experiments with different combinations of input modalities on BRATS2018 dataset in Fig. 4(b). Compared with Hi-Net that is specifically designed for two modalities input, our AutoGAN demonstrates considerable improvements, with PSNR rising from 24.95 dB (Hi-Net) to 27.12 dB (ours) in the task of FLAIR+T1→T2. Figure 4(b) also verifies that our method can fuse multiple input modalities and provide a promising performance. In addition, it illustrates that more input modality knowledge can also boost the synthesis performance. Figure 4(a) shows qualitative results of different multiple modalities input tested on the searched





**Fig. 4.** Multiple-to-one MRI cross modality synthesis tasks on BRATS2018 dataset: (a). Qualitative comparison of difference maps. (b). Quantitative results.

models by our AutoGAN. The results of FLAIR+T1+T1ce→T2 task are much better than the other three configurations in visualization, which is consistent with the observation from quantitative evaluation. It also verifies that different modalities contains partly complementary knowledge, which can boost the synthesis performance.

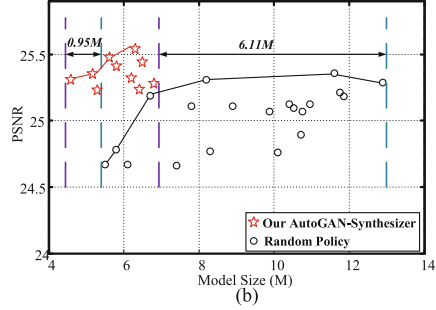
### 3.3 Ablation Study

**Study of Each Component:** We conduct an ablation study to demonstrate the effectiveness of each component, *i.e.*, the perceptual and adversarial part of our loss function, and the MRI K-space learning strategy. In Fig. 5(a), we list all results of different configurations on these three components in FLAIR →T2 synthesis tasks on BRATS2018 dataset. It indicates that the perceptual and adversarial loss can further improve quantitative performance. After adding the perceptual and adversarial loss, our algorithm can rehabilitate highly-realistic images with better structure similarity and peak signal-to-noise ratio. Furthermore, MRI K-space features embedded in the network can introduce additional information and therefore can also boost performance improvement. Figure 6 shows the qualitative results of the ablation study. We find that adding each component successively can obtain better synthetic images. In Fig. 6, the FLAIR image has poor quality and therefore is challenging to synthesize a reasonable T2 image. However, with the help of perceptual loss, adversarial loss, and K-space learning, the results are further improved and the missing part is gradually compensated.

**Effectiveness of Our Search Strategy:** To verify the effectiveness of our search strategy in AutoGAN, we compare our search strategy with random policy by randomly sampling 20 models from our search spaces. From Fig. 5(b), compared with random policy, our AutoGAN can search superior networks with less model size but better performance. More specifically, the networks from random policy have a wide range of model sizes from 6 Mb to around 12.5 Mb. But the search strategy of our AutoGAN is capable of constraining the model size of network within a much smaller interval by greatly reducing both the

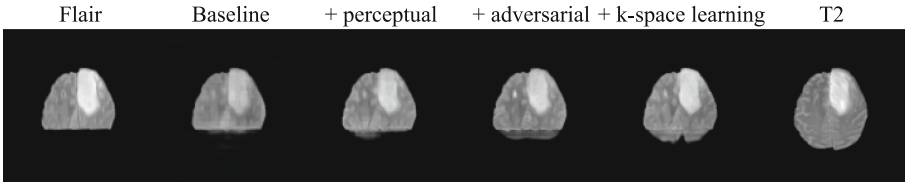
Configurations			Performance
Perceptual	Adversarial	K-space	PSNR/SSIM
×	×	×	25.23/0.87
✓	×	×	25.29/0.89
×	✓	×	25.37/0.90
✓	✓	×	25.47/0.89
×	✓	✓	25.49/0.90
✓	×	✓	25.35/0.89
✓	✓	✓	<b>25.54/0.90</b>

(a)



(b)

**Fig. 5.** (a). Ablation study of our GAN-based loss and MRI K-space features on BRATS2018 dataset (FLAIR →T2). (b). Comparison of our search strategy with random policy. Our AutoGAN can search light-weight networks with better performance.



**Fig. 6.** Visualization results of our ablation study, showing the effectiveness of three components in our pipeline: *i.e.*, perceptual loss, adversarial loss and K-space learning. The version of baseline represents the network without three components, *+perceptual* means the baseline with only perceptual loss, *+adversarial* denotes baseline with perceptual and adversarial loss and *+Kspace* represents our complete method with perceptual, adversarial loss and K-space learning.

lower-bound and the upper-bound of the model sizes without sacrificing the performance. This superiority makes it easier to deploy AI models in a variety of resource-constrained clinical scenarios.

## 4 Conclusion

We propose AutoGAN-Synthesizer to automatically design a generative network knowing how to extract and fuse features according to different input modalities for cross-modality MRI synthesis. A novel GAN-based perceptual searching loss incorporating specialized MRI K-space features is proposed to rehabilitate a highly-realistic image and to balance the trade-off between model complexity and performance. The proposed method outperforms other manually state-of-art synthesis algorithms and restores faithful tumor and anatomy information.

## References

1. Aggarwal, H.K., Mani, M.P., Jacob, M.: Multi-shot sensitivity-encoded diffusion MRI using model-based deep learning (MODL-MUSSELS). In: ISBI, pp. 1541–1544. IEEE (2019)
2. Armanious, K., et al.: MedGAN: medical image translation using GANs. *Comput. Med. Imaging Graph.* **79**, 101684 (2020)
3. Bakas, S., et al.: Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci. Data* **4**, 170117 (2017)
4. Bakas, S., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint [arXiv:1811.02629](https://arxiv.org/abs/1811.02629) (2018)
5. Chatsias, A., Joyce, T., Giuffrida, M.V., Tsaftaris, S.A.: Multimodal MR synthesis via modality-invariant latent representation. *IEEE Trans. Med. Imaging* **37**(3), 803–814 (2017)
6. Costa, P., et al.: End-to-end adversarial retinal image synthesis. *IEEE Trans. Med. Imaging* **37**(3), 781–791 (2017)
7. Dar, S.U., Yurt, M., Karacan, L., Erdem, A., Erdem, E., Çukur, T.: Image synthesis in multi-contrast MRI with conditional generative adversarial networks. *IEEE Trans. Med. Imaging* **38**(10), 2375–2388 (2019)
8. Gou, Y., Li, B., Liu, Z., Yang, S., Peng, X.: Clearer: multi-scale neural architecture search for image restoration. *NeurIPS*. **33**, 17129–17140 (2020)
9. Han, Y., Sunwoo, L., Ye, J.C.: k-space deep learning for accelerated MRI. *IEEE Trans. Med. Imaging* **39**(2), 377–386 (2019)
10. Hu, X.: Multi-texture GAN: exploring the multi-scale texture translation for brain MR images. arXiv preprint [arXiv:2102.07225](https://arxiv.org/abs/2102.07225) (2021)
11. Huang, Y., Shao, L., Frangi, A.F.: Simultaneous super-resolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding. In: CVPR, pp. 6070–6079 (2017)
12. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: CVPR, pp. 1125–1134 (2017)
13. Joyce, T., Chatsias, A., Tsaftaris, S.A.: Robust multi-modal MR image synthesis. In: Descoteaux, M., et al. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 347–355. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-66179-7\\_40](https://doi.org/10.1007/978-3-319-66179-7_40)
14. Jun, Y., Shin, H., Eo, T., Hwang, D.: Joint deep model-based MR image and coil sensitivity reconstruction network (Joint-ICNet) for fast MRI. In: CVPR, pp. 5270–5279 (2021)
15. Karras, T., Aila, T., Laine, S., Lehtinen, J.: Progressive growing of GANs for improved quality, stability, and variation. In: International Conference on Learning Representations (2018)
16. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014)
17. Liu, C., et al.: Progressive neural architecture search. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11205, pp. 19–35. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-01246-5\\_2](https://doi.org/10.1007/978-3-030-01246-5_2)
18. Liu, H., Simonyan, K., Yang, Y.: Darts: Differentiable architecture search. In: ICLR (2018)

19. Liu, Z., Wang, H., Zhang, S., Wang, G., Qi, J.: NAS-SCAM: neural architecture search-based spatial and channel joint attention module for nuclei semantic segmentation and classification. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 263–272. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-59710-8\\_26](https://doi.org/10.1007/978-3-030-59710-8_26)
20. Loshchilov, I., Hutter, F.: SGDR: stochastic gradient descent with warm restarts. arXiv preprint [arXiv:1608.03983](https://arxiv.org/abs/1608.03983) (2016)
21. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2014)
22. Nah, S., Hyun Kim, T., Mu Lee, K.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: CVPR, pp. 3883–3891 (2017)
23. Nie, D., et al.: Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans. Biomed. Eng.* **65**(12), 2720–2730 (2018)
24. Qianye, Y., Li, N., Zhao, Z., Xingyu, F., Eric, I., Chang, C., Xu, Y.: MRI cross-modality image-to-image translation. *Sci. Rep.* **10**(1), 1–8 (2020)
25. Real, E., Aggarwal, A., Huang, Y., Le, Q.V.: Regularized evolution for image classifier architecture search. In: AAI, vol. 33, pp. 4780–4789 (2019)
26. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
27. Sharma, A., Hamarneh, G.: Missing MRI pulse sequence synthesis using multi-modal generative adversarial network. *IEEE Trans. Med. Imaging* **39**(4), 1170–1183 (2019)
28. Wang, J., et al.: Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 3349–3364 (2020)
29. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
30. Welander, P., Karlsson, S., Eklund, A.: Generative adversarial networks for image-to-image translation on multi-contrast MR images—a comparison of cycleGAN and unit. arXiv preprint [arXiv:1806.07777](https://arxiv.org/abs/1806.07777) (2018)
31. Yan, X., Jiang, W., Shi, Y., Zhuo, C.: MS-NAS: multi-scale neural architecture search for medical image segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 388–397. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-59710-8\\_38](https://doi.org/10.1007/978-3-030-59710-8_38)
32. Yu, B., Zhou, L., Wang, L., Shi, Y., Fripp, J., Bourgeat, P.: EA-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis. *IEEE Trans. Med. Imaging* **38**(7), 1750–1762 (2019)
33. Yu, B., Zhou, L., Wang, L., Shi, Y., Fripp, J., Bourgeat, P.: Sample-adaptive GANs: linking global and local mappings for cross-modality MR image synthesis. *IEEE Trans. Med. Imaging* **39**(7), 2339–2350 (2020)
34. Yurt, M., Dar, S.U., Erdem, A., Erdem, E., Oguz, K.K., Çukur, T.: mustGAN: multi-stream generative adversarial networks for MR image synthesis. *Med. Image Anal.* **70**, 101944 (2021)
35. Zhang, H., Li, Y., Chen, H., Shen, C.: Memory-efficient hierarchical neural architecture search for image denoising. In: CVPR, pp. 3657–3666 (2020)
36. Zhang, R., Pfister, T., Li, J.: Harmonic unpaired image-to-image translation. arXiv preprint [arXiv:1902.09727](https://arxiv.org/abs/1902.09727) (2019)
37. Zhang, Y., Li, K., Li, K., Fu, Y.: MR image super-resolution with squeeze and excitation reasoning attention network. In: CVPR, pp. 13425–13434 (2021)

38. Zhou, T., Fu, H., Chen, G., Shen, J., Shao, L.: Hi-net: hybrid-fusion network for multi-modal MR image synthesis. *IEEE Trans. Med. Imaging* **39**(9), 2772–2781 (2020)
39. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *ICCV*, pp. 2223–2232 (2017)
40. Zhu, Z., Liu, C., Yang, D., Yuille, A., Xu, D.: V-NAS: neural architecture search for volumetric medical image segmentation. In: *3DV*, pp. 240–248. *IEEE* (2019)