

---

# MediSinGAN

---

**Amrit Kumar Jethi\***  
IIT Madras, India

**Anagha Zachariah**  
Vellore Institute of Technology, India

**Ekaterina Nepovinnikh**  
LUT University, Finland

**Rajkumar Vaghashiya**  
Pandit Deendayal Energy University, India

**Phuc Nguyen**  
University of Cincinnati, US

**Md.Sadek Hossain Asif**  
Notre Dame College, Dhaka, Bangladesh

**Madhu Mithra K K**  
IIT Madras, India

## Abstract

Deep learning-based medical image analysis has undergone substantial growth in recent times with the latest advancements being actively employed in a variety of applications, such as for segmentation, characterization, and reconstruction across domains like radiology, pathology, and ophthalmology. A major impediment in these advancements is the highly complicated process of medical data collection, owing to issues such as patient privacy and regulatory affairs, as well as the need for experts to annotate the data. GANs (Generative Adversarial Networks) seems to be a potential solution to address this issue due to their ability to generate plausible synthetic data. In this work, we study the utility of SinGAN—an unconditional generative model trained on a single image—for synthetic data generation across different imaging tasks, namely, multi-modal MRI (Magnetic Resonance Imaging) data generation, brain tumour data generation, and histopathology image segmentation. These applications, built on SinGAN, could become a significant remedy to the data-deprived medical imaging tasks.

## 1 Introduction

GANs have shown remarkable success in the generation of realistic, high-quality samples when trained on class-specific datasets. In the medical field, the collection of high-quality data is a huge challenge in the development of intelligent tools that can assist medical professionals.

To overcome the data-deficiency problem, many past works have explored different GAN architectures as a solution. Li et al. (8) proposed a novel GAN model called TumorGAN with two discriminators for a more realistic tumour image generation, wherein the model could synthesize virtual image pairs from real data pairs. The AutoSyncoder (10) network in a GAN and cyclicGAN setting provides multi-contrast MRI images. Van Eycke et al. (4) discussed various strategies to reduce the expert supervision via deep learning-based segmentation of histopathological images.

To address this problem, we analyze the applicability of SinGAN (11) for medical image synthesis. SinGAN generates samples after unconditionally learning from a single image by capturing the internal distribution of patches within the image. It can generate diverse samples, of arbitrary size and aspect ratio while preserving the visual contents of the training image. In this work, we experiment with SinGAN to study synthetic medical data generation, modality translation, and segmentation.

In our work, we study three main data generation tasks: 1) Multi-contrast MRI data generation - The combination of T1, T2, and FLAIR images consistent with each other; 2) Histopathology

---

\*All authors contributed equally to this work

image segmentation mask generation - The histology image with its corresponding segmentation map; and 3) Tumour data generation - The insertion of brain tumours in healthy MRI images via harmonization for more realistic results.

## 2 Methods

### 2.1 Multimodal MRI Data Generation

There are multiple modalities in MRI, each concerned with capturing certain characteristics in human anatomy. T1 images are favorable in the identification of fatty tissues and extraction of morphological information; T2 images are utilized for locating edema and inflammation; Flair images are used for the detection and localization of lesions with water suppression (13). The idea of generating numerous random sets of multimodal MRI images can be employed in applications such as MRI cross-modality image-to-image translation, MRI reconstruction (12), and others. We have used Multimodal Brain Tumour Image Segmentation Benchmark (BRATS) for the experiment (1; 9; 2). The SinGAN model accepts an RGB-channel image as an input for training. Since the MRI images are greyscale, we stacked the multimodal MRI images of T1, T2, and FLAIR into a single 3-channel image for training the model. The random samples are generated from the second scale of the trained model to maintain the global structure such as the brain outline while introducing local variations.

### 2.2 Segmentation Mask Generation

Here, we focused on the generation of segmentation maps of the nuclei in histopathology images. Histopathology involves microscopic examination of tissues for disease detection, wherein the arrangement and appearance of the nuclei essentially characterize the type of cancer present in the tissue (7). In this study, we used the PanNuke dataset (5), an open pan-cancer histology dataset for nuclei instance segmentation. Following the multimodal MRI generation method, the input-output pairs of the histopathology image and its corresponding segmentation mask are stacked to form a single input image for the model training. Since global perturbations on the image will produce varying positions of the nuclei, the random data generation scale is set to the lowest (coarsest) scale of the trained model to generate synthetic histopathology images and their corresponding segmentation maps.

### 2.3 Synthetic Tumour Generation

To generate tumour data, we train the model with a healthy brain MRI as the background image and pass a naively pasted composite of the tumour during test time, leveraging the harmonization image manipulation technique offered by SinGAN. In this case, the scale of injection plays a pivotal role in the resulting output. We injected the naively pasted tumour at varying scales to study the optimum result criteria. Compared to other existing methods (8), the output from SinGAN is more realistic and comparable to the real tumour images.

## 3 Tools and Framework

JAX is a framework that allows running tensor computation and differentiation code with NumPy-like interface on CPU, GPU, and TPU (3) and delivering appreciable computation speed-up. Flax is a high-performance neural network library for JAX, designed for flexibility and with all the necessary functions for deep learning (6). Since the original SinGAN implementation, in PyTorch, requires considerable compute time, we experimented with JAX and Flax to obtain a compute speed-gain in the training<sup>2</sup>. The training has been carried out online on Google Colab.

From the experiments, it was observed that despite being based on completely different principles, both Pytorch and JAX provide a convenient and flexible interface for implementing machine learning algorithms. The advantages of JAX include pure functions that make the code lucid and easy to debug, elimination of the need for manually tracking of the computation graphs, just-in-time compilation for computation speed up and the familiar NumPy-like interface for working with Tensors.

---

<sup>2</sup>The source code can be found [here](#)

Table 1: SinGAN training time comparison between Pytorch & Jax with Flax (in seconds)

	Scale 0	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7	Scale 8
PyTorch	150.76	166.37	184.15	200.41	222.73	259.21	936.90	1142.98	2003.36
JAX	48.82	151.50	251.57	356.5	484.12	595.64	739.03	914.50	1293.80

The results presented in Table 1 show that the training time in JAX is generally comparable to that of Pytorch, with JAX being faster in most cases.

## 4 Results and Discussion

The complete experimental results of the simulation are available [here](#).

### 4.1 Multimodal MRI Data Generation

The key observation in Multimodal MRI data generation is that the scale used to generate the images governs the amount of variability in the result. Starting the process on a lower scale imparts large and global variations to the output whereas a higher scale imparts lesser and more local variations, with the best results observed at scale 1. The output MRI images of different modalities, as shown in Fig. 1a, are consistent with each other. The random samples produced by the model are essentially a synthetic augmentation of the training image. The images produced by the model are not completely realistic, leaving scope for improvements.

### 4.2 Segmentation Mask Generation

The histopathology image and the corresponding segmentation mask of its nuclei generated from the model are shown in Fig. 1b. It is observed that the generated input-output pairs are very realistic in comparison with the training image, with generation scale 0 producing the most variations.

### 4.3 Synthetic Tumour Generation

The results of synthetic tumour generation in the healthy brain using harmonization are as shown in Fig. 2a, where it is observed that the generated fake output closely models the real tumour images. The best results were obtained on the scale 2 and 3. Figure 2b shows the effect of the different scales of injection on the final output. The model generates a realistically blended image of the tumour on the healthy MR image.

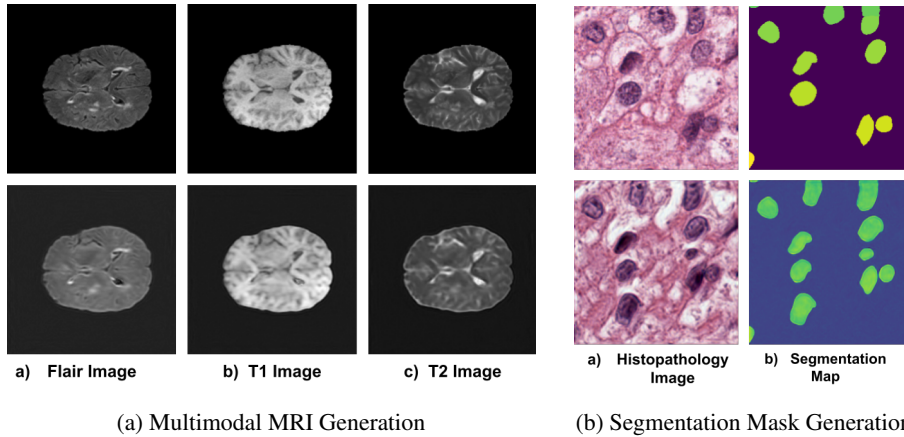


Figure 1: Results from multimodal MRI image translation and segmentation. The top row represents the training image while the bottom row is the model output

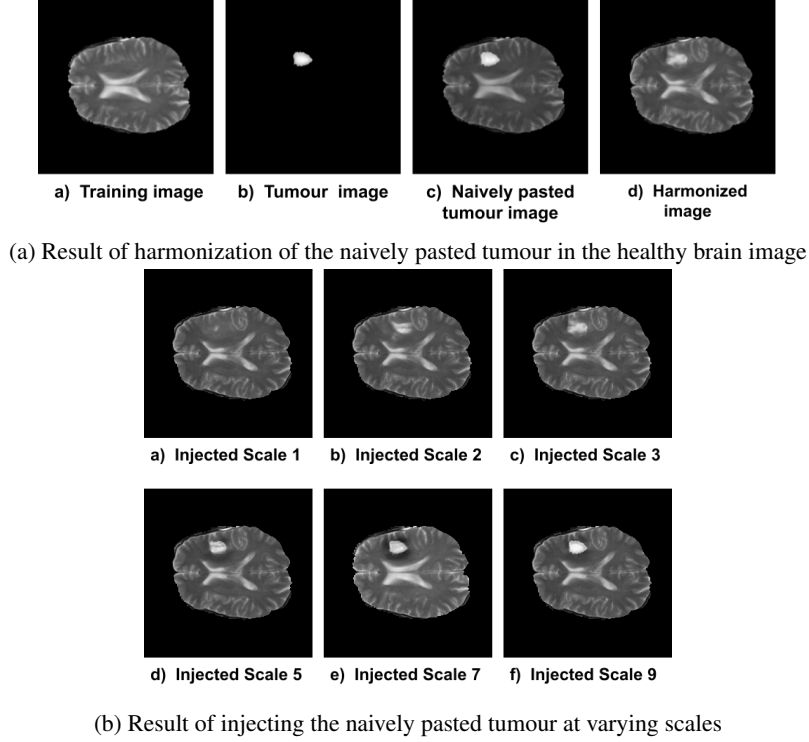


Figure 2: Synthetic tumour generation results

## 5 Conclusion

The experiments carried out in this preliminary study show promising results that clearly depict the potential utility of SinGAN in different medical imaging tasks. The results obtained from the model can be applied in various medical imaging-based deep models to solve the problem of data deficiency, such as for improving multi-modal brain image analysis, medical image segmentation, and generation of realistic samples for rare diseases. The results obtained from the model are not perfectly realistic and leave a lot of scope for further research on fine-tuning and improving the model to fit the specific applications.

## 6 Future works

The results obtained from the multimodal MR data generation do not look perfectly realistic. This shortcoming can be overcome by learning style from real-world images and injecting them along with the synthetic data. Also, the authenticity of the generated medical images should be properly verified using medical experts to quantify the obtained results. JAX implementation of SinGAN written from scratch will further help in improving the compute speed-up by leveraging just-in-time and vmap capabilities.

## Acknowledgments

The authors would like to thank the Eastern European Machine Learning Summer School (EEML'21) team, especially Viorica Patraucean, Razvan Pascanu, and Ferenc Huszar for the invaluable knowledge and the opportunity to work together on this project.

The authors would like to thank the advisers Margarete Kattau, The Institute of Cancer Research, London, and Fedor Zolotarev, LUT University, Finland, for their guidance and support.

## References

- [1] Ujjwal Baid, Satyam Ghodasara, Michel Bilello, and Suyash Mohan et al. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification, 2021.
- [2] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data*, 4(1):170117, 2017.
- [3] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.
- [4] Yves-Rémi Van Eycke, Adrien Foucart, and Christine Decaestecker. Strategies to reduce the expert supervision required for deep learning-based segmentation of histopathological images. *Frontiers in Medicine*, 6, Oct. 2019.
- [5] Jevgenij Gamper, Navid Alemi Koohbanani, Ksenija Benet, Ali Khuram, and Nasir Rajpoot. Pannuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In *European Congress on Digital Pathology*, pages 11–19. Springer, 2019.
- [6] Jonathan Heek, Anselm Levskaya, Avital Oliver, Marvin Ritter, Bertrand Rondepierre, Andreas Steiner, and Marc van Zee. Flax: A neural network library and ecosystem for JAX, 2020.
- [7] Le Hou, Rajarsi Gupta, John S. Van Arnam, Yuwei Zhang, Kaustubh Sivalenka, Dimitris Samaras, Tahsin M. Kurc, and Joel H. Saltz. Dataset of segmented nuclei in hematoxylin and eosin stained histopathology images of ten cancer types. *Scientific Data*, 7(1), June 2020.
- [8] Qingyun Li, Zhibin Yu, Yubo Wang, and Haiyong Zheng. TumorGAN: A multi-modal data augmentation framework for brain tumor segmentation. *Sensors*, 20(15):4203, July 2020.
- [9] Bjoern H. Menze, Andras Jakab, and Stefan et al. Bauer. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.
- [10] JayaChandra Raju, Balamurali Murugesan, Keerthi Ram, and Mohanasankar Sivaprakasam. AutoSyncoder: An adversarial AutoEncoder framework for multimodal MRI synthesis. In *Machine Learning for Medical Image Reconstruction*, pages 102–110. Springer International Publishing, 2020.
- [11] Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli. Singan: Learning a generative model from a single natural image. In *Computer Vision (ICCV), IEEE International Conference on*, 2019.
- [12] Lei Xiang, Yong Chen, Weitang Chang, Yiqiang Zhan, Weili Lin, Qian Wang, and Dinggang Shen. Ultra-fast t2-weighted MR reconstruction using complementary t1-weighted information. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 215–223. Springer International Publishing, 2018.
- [13] Qianye Yang, Nannan Li, Zixu Zhao, Xingyu Fan, E. Chang, and Yan Xu. Mri image-to-image translation for cross-modality image registration and segmentation. *ArXiv*, abs/1801.06940, 2018.