Tissue-Specific Color Encoding and GAN Synthesis for Enhanced Medical Image Generation

Yu Shi[1], Hannah Tang[1], Jianxin Sun[1], Xinyan Xie[1], Huijing Du[1], Dandan Zheng[2], Chi Zhang[1], Hongfeng Yu[1]
[1]*University of Nebraska-Lincoln, Lincoln, NE, USA*
[2]*University of Rochester Medical Center, Rochester, NY, USA*

***Abstract*—Medical image synthesis is important in diverse healthcare applications, such as computer-aided diagnosis, medical image analysis, and educational tools. While Generative Adversarial Networks (GANs) have shown remarkable success in generating natural images, their application to medical images often falls short in faithfully capturing essential anatomical features. In this paper, we introduce a new approach that focuses on tissue-specific color encoding to enhance medical image synthesis using GANs. Our method deviates from the conventional practice of directly training GANs on gray-scale medical images. Instead, we initiate the process by generating and encoding various gray-scale representations of distinct tissues into separate color channels within composite images. These tissue-specific color images are then utilized to train a GAN model. The GAN, once trained, excels in producing high-quality synthetic images for individual tissues, and when combined, these tissue images yield final synthesized images that better portray the intricate tissue characteristics found in medical data. We have conducted an experimental study to validate the effectiveness of our approach in comparison to alternative methods with both qualitative and quantitative assessments to evaluate the quality of synthesized individual tissues and their combined final results.**

***Index Terms*—GAN, medical image synthesis, tissue-specific color encoding**

I. Introduction

Medical imaging and analysis play a critical role in the realm of healthcare. They are crucial in identifying diseases at an early stage [1] and allowing for timely interventions and improving patient outcomes [2]. They are employed in precise diagnoses, understanding disease characteristics, and devising treatment plans. Furthermore, image analysis assists in guiding surgical processes and monitoring treatment effectiveness, minimizing invasiveness and complications [3]. Medical images are also indispensable in education, enhancing diagnostic skills and deepening our understanding of diseases [4]. They are integral to medical research, clinical trials, and the advancement of healthcare technologies [5].

With the rapid progression in computer vision and deep learning techniques, there has been significant progress in the field of medical image analysis. However, unlike the typical computer vision tasks involving natural images, medical images present greater diversity, are more costly to obtain, and raise additional ethical concerns, particularly regarding patient privacy and consent. As a consequence, the constrained availability of data stands as a significant obstacle hindering the progression of medical imaging research and education [6].

The availability of accessible datasets is crucial for the advancements in the medical domain. Generative Adversarial Networks (GANs) are a common deep-learning approach for generating synthetic images. GANs involve a generator and a discriminator engaged in a game-theoretic competition to produce authentic synthetic data [7]. GANs are well-suited for generating synthetic data as they can learn the data distribution and generate new samples that closely resemble the original data from a latent space.

While GANs have achieved remarkable success in synthesizing natural images, their application to synthetic medical images presents a unique challenge. Even though synthetic medical images can often achieve top-tier scores in image quality metrics such as FID (Fréchet Inception Distance), there is a critical concern related to the presence of artifacts that violate fundamental anatomical features. These anomalies, which compromise the anatomical accuracy of the generated images, can significantly limit the usability and reliability of synthetic medical imagery in crucial applications such as diagnostics, treatment planning, medical research, and education. Addressing and mitigating these artifacts is critical for harnessing the full potential of GANs in generating medically relevant imagery.

We present a new strategy to enhance the synthesis of medical images using GANs. Rather than directly training GANs on traditional gray-scale medical images, we propose a tissue-specific GAN approach, which divides original gray-scale medical images into distinct tissue-specific representations. These representations are then encoded into dedicated color channels of composite images, which are subsequently used to train GANs. Our premise is that this approach allows GANs to learn and replicate the specific characteristics of various tissues and their relationship more effectively. Our experimental investigations have shown that this method leverages the capabilities of GANs to generate synthetic images that can more accurately capture the subtle features of medical tissues.

II. Related Work

In recent years, significant progress has been made in optimizing GANs, there are variants of GANs that achieve significant improvements on general image synthesis tasks by varying the objective of discriminator [8], generator [9], [10] or architecture [11] from the vanilla GAN [7]. In the medical imaging domain, modalities like MRI, CT scan, ultrasound, and radiography have different natures in image acquisition, it is necessary to reevaluate the usability or potential of the deep

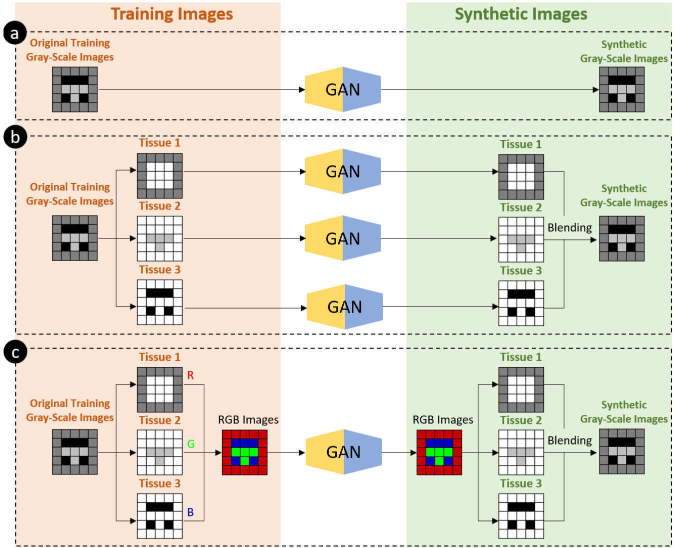

Fig. 1: Comparison of (a) a traditional GAN approach, (b) an alternative non-color encoding approach, and (c) our tissue-specific color encoding and GAN Approach.

generative models, which are usually developed with natural image datasets.

Several works have demonstrated the generation of realistic synthetic medical imaging data. In brain MRI data generation, Shin et al. used a pix2pix conditional GAN [12] to generate brain tumor MRI images with segmentation masks [13]. The application of a PGGAN-based model was also explored in generating brain MRI images [14]. In the realm of lung cancer nodule generation, a DCGAN-based method demonstrated convincing results that were positively received by radiologists [15], [16]. Subsequently, a later model was developed to generate de-identified public radiography datasets [17]. D2FE-GAN [18] utilized decoupled dual feature representations to synthesize cross-modality MRI. FedMed-GAN [19] enhanced unsupervised cross-modality synthesis of brain images in a federated manner. 3DGAUnet [20] incorporated a 3D U-Net architecture into the generator, improving 3D shape and texture synthesis, particularly for pancreatic ductal adenocarcinoma (PDAC) tumors and pancreatic tissues.

StyleGAN [21], by incorporating modified latent vectors across various network layers at different resolutions, effectively managed features at diverse scales. This approach significantly improved stability, visual fidelity, and maneuvering capabilities. Its successor, styleGAN2-ada, exhibited superior performance, especially when dealing with relatively small training datasets [22].

Despite the visual realism and high scores in image quality metrics, such as FID, synthetic medical images face challenges due to artifacts that deviate from fundamental anatomical features [23]. Presently, there is a lack of GAN approaches that specifically address this issue by encoding tissue-type information into the training data. To prioritize the learning of tissue-specific features by GANs, we propose incorporating tissue-specific channels as an encoding mechanism for tissue-type information in the original training data.

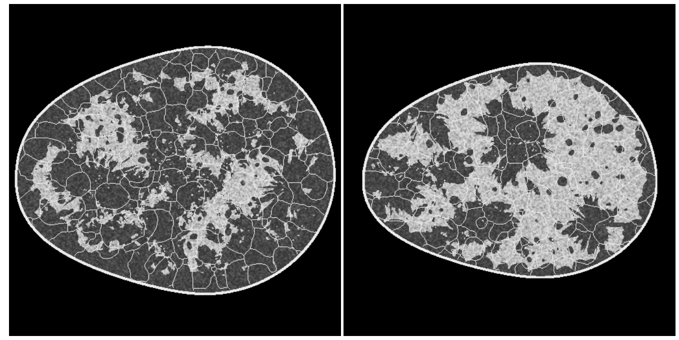

Fig. 2: Two ground truth breast coronal slices from the dataset used in our experiment.

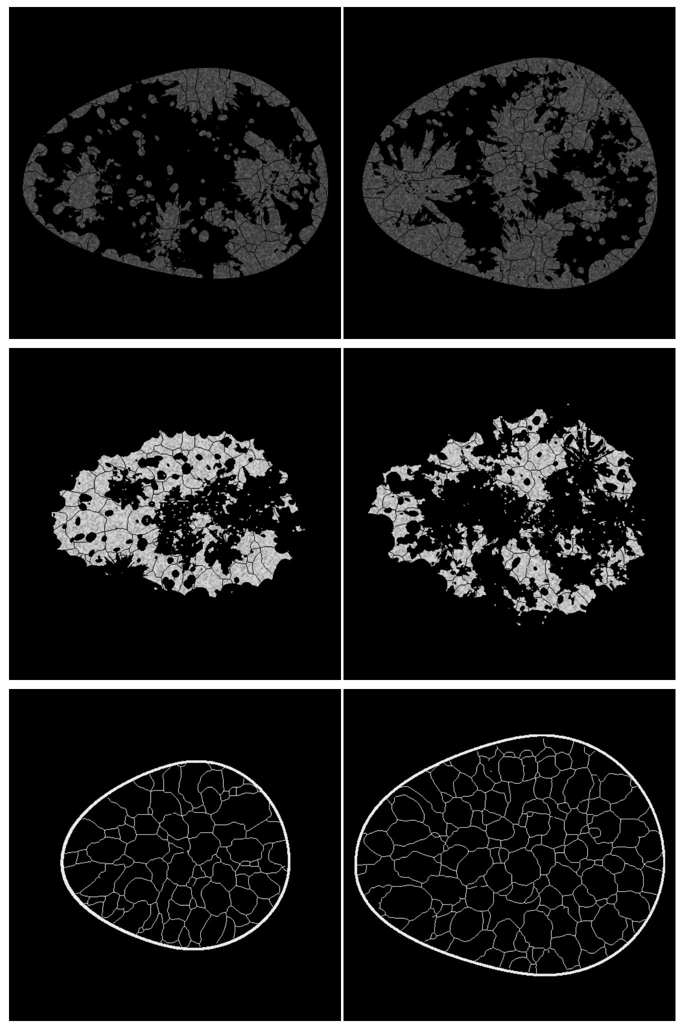

Fig. 3: Examples of fat, glandular, and dense tissues (from the top to bottom rows) separated from two ground truth breast coronal slices.

III. Methodology

The traditional method involves training a GAN model directly on the original medical images and using the trained GAN to generate synthetic images, where both the training and synthetic images are gray-scale medical images, as depicted in Figure 1(a). In contrast, our approach enables precise manipulation of tissue-specific image attributes within the original gray-scale medical images, facilitating the stable construction of realistic images with the desired tissue characteristics, as illustrated in Figure 1(c). Specifically, our approach can be summarized into the following key steps:

Data Preprocessing. We start with a collection of gray-scale medical images. For each image, we segment and isolate different tissues or structures, each having its own unique gray-scale range. We then generate separate gray-scale images for these tissues based on their respective ranges.

Tissue-Specific Color Encoding. We create color images with red, green, and blue channels. Each channel is reserved for encoding one of the gray-scale tissue images. The gray-scale images of different tissues are encoded into their respective channels, creating a multi-channel color image.

GAN Training. The multi-channel color images are used as the training dataset for a GAN. The GAN is trained to generate synthetic color images that capture the diversity of tissue features.

Post-Processing. We use the trained GAN to generate synthetic color images, from where we extract the gray-scale images from the red, green, and blue channels. These gray-scale images represent the individual tissues or structures. We blend these gray-scale images to create a final synthetic gray-scale medical image and capture tissue features.

Our intuition is that the effectiveness of our approach can be attributed to tissue-specific color encoding, which not only enhances the differentiation of various tissues but also improves the depiction of their spatial relationships within medical images. As a result, this approach enhances the representation of tissue features and results in the generation of more realistic synthetic medical images. We have conducted an empirical study to verify our hypothesis.

IV. Experimental Results

We have tested our approach using high-resolution breast coronal images. Our method allows the generative model to stably construct realistic breast coronal slices with desired fatty tissue and glandular tissue ratio. The quality of the generated images has been assessed in both qualitative manners and quantitative metrics compared to alternative approaches.

A. Dataset

Our model is trained on the 2023 AAPM Grand Challenge on Deep Generative Modeling for Learning Medical Image Statistics dataset [24]. The dataset contains 108,530 8-bit images with a size of 512 pixels $\times$512 pixels. Images are originally from VICTRE breast phantom creation software that emulates coronal slices from anthropomorphic breast phantoms. The data comprises four breast tissue composition categories: extremely dense, heterogeneously dense, fibrous and glandular density, and almost entirely fatty. Breast density is classified according to the BI-RADS system [25]. Figure 2 shows two images of coronal slices from this dataset.

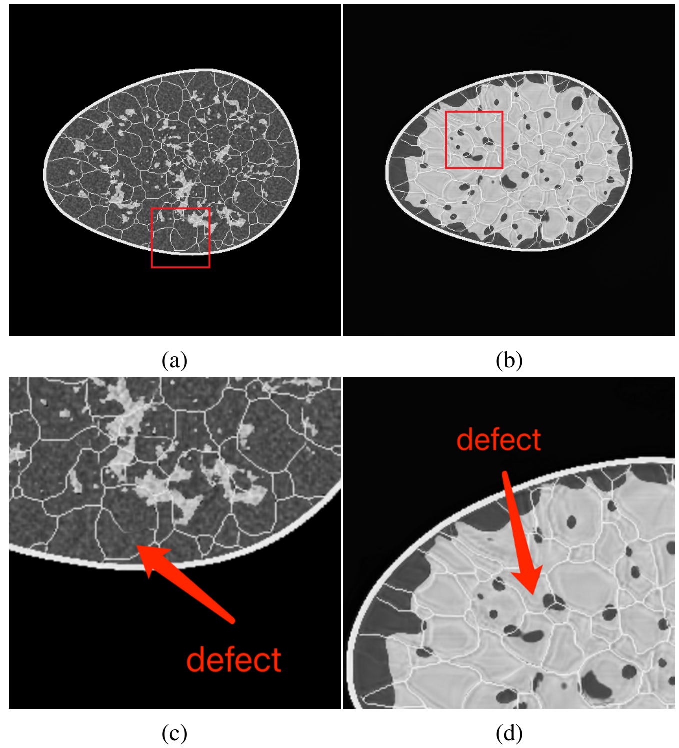

Fig. 4: Example results using the traditional GAN approach (top row) and their defects detected upon closer examination (bottom row).

B. Tissue-Specific Channels

We convert the original gray-scale image with a size of 512 pixels $\times$512 pixels into three channels based on its tissue-specific intensity range. The fat tissue channel includes all pixels in the range of [45,120), the glandular tissue channel includes all pixels in the range of [120,226), and the dense tissue channel includes all pixels in the range of [226,255). For each original image, a new color image with a size of 512 pixels $\times$512 pixels is created, and the fat, glandular, and dense tissue channels are encoded into the red, green, and blue channels of the color image, respectively. The threshold of the range is based on X-ray attenuation coefficients assigned to the various tissues on existing breast computed tomography. Figure 3 shows examples of fat, glandular, and dense tissues separated from two ground truth coronal slices.

C. Generative Model

To counter the constraint of the limited amount of data, augmentations are widely used, our generative model is based on the StyleGAN2-ada [22] since it has been reported to have top performance when dealing with limited data with a controllable augmentation pipeline. The training set has

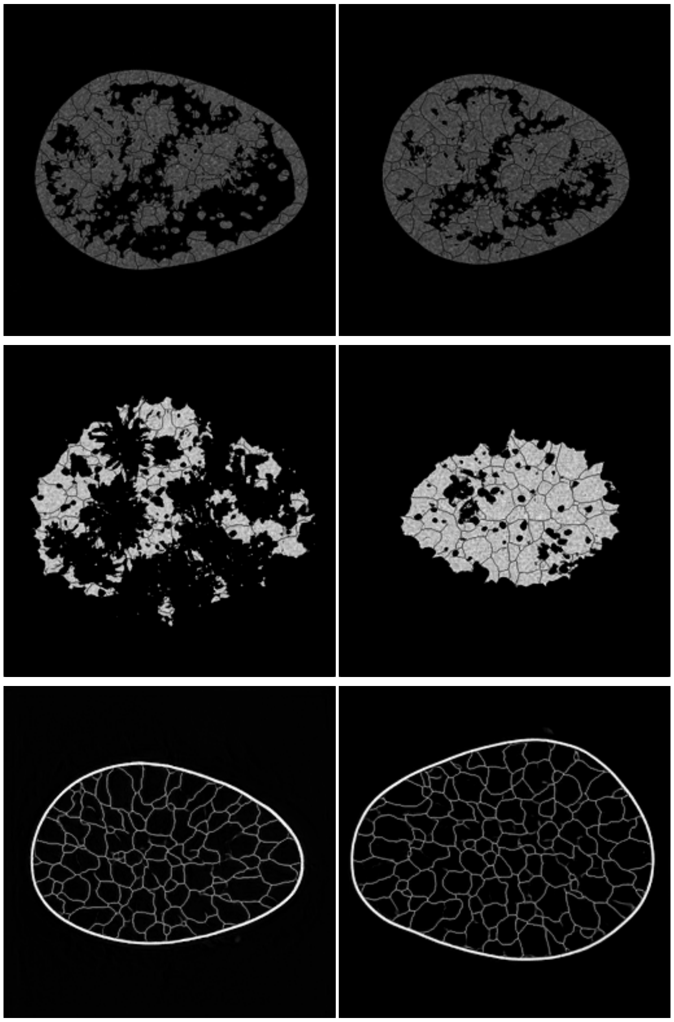

Fig. 5: Examples of synthetic fat, glandular, and dense tissue images (from the top to bottom rows) generated by the alternative non-color encoding method.

a total number of 108,530 images. The parameters are as follows: training duration = 25,000 kimg, R1-gamma = 6.6, batch size = 8, learning rate = 0.0025. The augmentation factor = 0.6, with blit, geom, color, filter, noise, and cutout image augmentation available. The best model achieved at 7600 kimg. 10,000 images were generated with the network weights for further tests. The original output of the images has three channels, each channel goes through a background noise reduction pipeline. Low-value pixels were removed, and the sum of three matrices was used to generate final gray-scale images.

D. Alternative Non-Color Encoding Method

For comparison, in addition to the traditional GAN approach illustrated in Figure 1(a), we have developed an alternative method, as depicted in Figure 1(b), which does not utilize color encoding. In this alternative approach, tissues are segregated into distinct gray-scale images based on their unique intensity ranges, similar to our method. However, in contrast to our approach, these individual gray-scale tissue images are utilized to train separate GAN models. Subsequently, these GANs are applied to generate synthetic images for the respective tissues. The synthetic images of the individual tissues are then blended to produce a final image.

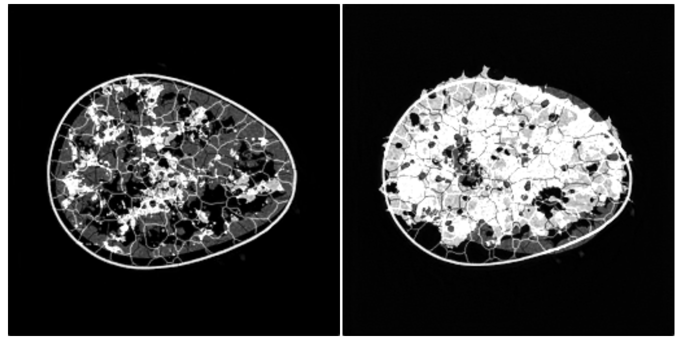

Fig. 6: Final breast coronal slice results, generated by blending fat, glandular, and dense tissue images synthesized using the alternative non-color encoding method, fail to capture the cohesive relationship among these tissues.

E. Results

Figure 4 shows two instances of results produced by applying the traditional GAN approach as depicted in Figure 1(a), with the breast images serving as the training dataset. In general, the traditional GAN technique is capable of generating visually pleasing results, as shown in Figure 4(a) and (b), that closely resemble ground truth images, e.g., Figure 2. However, upon closer examination, we can observe that these images have certain imperfections, failing to retain subtle tissue characteristics. For instance, when we zoom in, we notice discrepancies such as the discontinuity within the dense tissue (Figure 4(c)), and the presence of shadows within the glandular tissue (Figure 4(d)). These anomalies do not correspond with the expected anatomical features of these tissues.

Figure 5 shows examples of synthetic fat, glandular, and dense tissues generated by the alternative non-color encoding method. We can see that these results exhibit a remarkable similarity to ground truth tissue images, such as the ones shown in Figure 3. Furthermore, this approach can potentially reduce the imperfections produced by the conventional GAN method. Our hypothesis is that this is because of the means of isolating and employing individual tissues for training specific GANs, yielding better synthetic outcomes for each tissue. However, a mere combination of these tissue images does not ensure the production of high-quality final synthesized results. As revealed in Figure 6, the blended images fail to capture the coherent relationship among these tissues, as seen in the ground truth images, such as Figure 2.

Figure 7 shows examples of synthetic fat, glandular, and dense tissue images generated by our method. We can see that these results also exhibit a remarkable similarity to ground

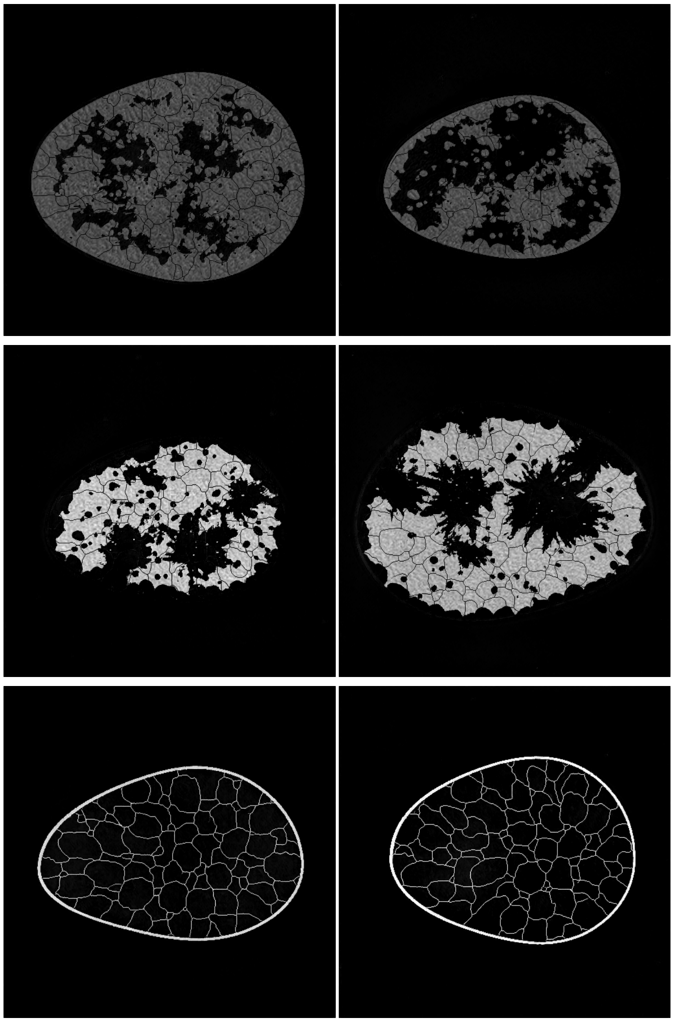

Fig. 7: Examples of synthetic fat, glandular, and dense tissue images (from the top to bottom rows) generated by our method.

truth tissue images, such as the ones shown in Figure 3, and effectively reduce the unwanted anomalies typically produced by the conventional GAN method. We postulate that this improvement is attributed to the separation of individual tissues and their encoding within distinct color channels, enabling the trained GANs to capture the unique characteristics of each tissue more precisely. In addition, a combination of these tissue images can produce high-quality final synthesized results. As revealed in Figure 8, the blended images can capture the coherent relationships among these tissues and are close to ground truth breast coronal slices, such as Figure 2. This suggests that our method can potentially preserve the spatial connections between different tissues, and encoding each tissue image into a separate color channel can help a trained GAN model capture and retain these inter-tissue spatial relationships.

Therefore, based on these qualitative experimental findings, we can clearly see that our method can not only synthesize realistic fat, glandular, and dense tissue images but also capture their anatomical relationship in the final images by combining these individual tissue images.

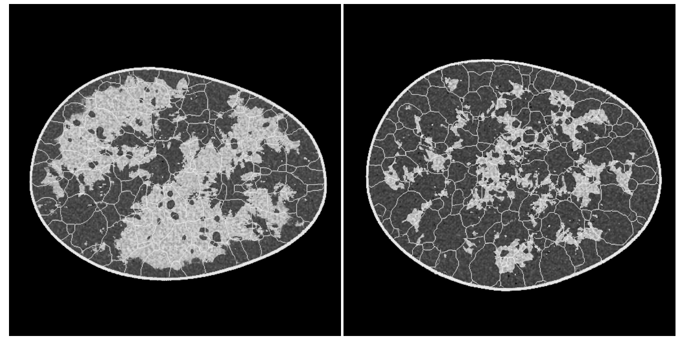

Fig. 8: Final breast coronal slice results, generated by blending fat, glandular, and dense tissue images synthesized using our method, can effectively capture the coherent relationships among these tissues.

We also utilize the FID metric [26] to conduct a quantitative evaluation of the results generated by the different approaches, as summarized in Table I. It is observed that the traditional GAN approach can yield synthetic outcomes with a reasonable FID score of 13.04. However, it falls short in preserving anatomical details, as demonstrated in Figure 4.

The alternative non-encoding approach excels in faithfully synthesizing individual fat, glandular, and dense tissue images, achieved through the training of GANs with separate tissue representations extracted from the ground truth. However, when these synthetic tissue images are combined to produce final coronal slice results, they fail to attain a satisfactory FID score. The FID scores for this approach, as presented in Table I, affirm the findings observed in Figures 5 and 6.

In contrast, our method delivers satisfactory FID scores for both individual tissue images and final breath coronal slice images with high qualities. Both the qualitative (Figures 7 and 8) and quantitative (Table I) findings demonstrate the effectiveness and superiority of our approach compared to the alternative methods.

V. Conclusion

We have presented a simple yet highly effective technique for medical image generation, featuring tissue-specific color encoding and GAN-based synthesis. By incorporating gray-scale tissue images into the red, green, and blue channels of color images, our method significantly improves the representation of tissue characteristics and their anatomical relationships, resulting in the production of synthetic medical images with enhanced realism. This innovative approach holds promise for various medical applications, offering access to high-quality, diverse, and easily interpretable synthetic medical data.

Although our preliminary study focuses on encoding intensities of three tissues into the three channels of color images, it is feasible to expand our method to accommodate multiple tissue images and encode them into multiple channels within high-dimensional vector representations for GAN training.

TABLE I: FID scores of the results generated by different methods.

Method	Traditional GAN	Non-color Encoding Approach				Our Approach			
Synthetic Tissue	breast coronal slice	fat	glandular	dense	breast coronal slice	fat	glandular	dense	breast coronal slice
FID	13.04	16.28	17.29	23.14	152.76	15.05	18.47	16.15	3.22

In the future, we aim to extend our approach further to encompass a broader spectrum of tissues, gaining a more comprehensive understanding of mechanisms involved in tissue encoding. This will enable us to capture better the fundamental anatomical features and interrelationships among tissues in synthetic models, leading to the generation of more realistic medical images. In addition, we plan to engage domain experts in the assessment of synthesized outcomes, which will help us to develop a deeper understanding of the practical implications of GAN models.

Acknowledgment

This research has been sponsored in part by Nebraska Collaboration Initiative grant 19-20 and 20-21, National Institutes of Health grant 5U54GM115458-03, National Science Foundation grant DMS-1853636, and the Layman Fund held at the University of Nebraska Foundation. This work was completed utilizing the Holland Computing Center of the University of Nebraska, which receives support from the UNL Office of Research and Economic Development and the Nebraska Research Initiative.

References

[1] S. P. Pereira, L. Oldfield, A. Ney, P. A. Hart, M. G. Keane, S. J. Pandol, D. Li, W. Greenhalf, C. Y. Jeon, E. J. Koay *et al.*, "Early detection of pancreatic cancer," *The lancet Gastroenterology & hepatology*, vol. 5, no. 7, pp. 698–710, 2020.

[2] B. Qin and D. Li, "Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19," *Sensors*, vol. 20, no. 18, p. 5236, 2020.

[3] T. M. Peters, "Image-guidance for surgical procedures," *Physics in Medicine & Biology*, vol. 51, no. 14, p. R505, 2006.

[4] B. Grignon, G. Oldrini, and F. Walter, "Teaching medical anatomy: what is the role of imaging today?" *Surgical and Radiologic Anatomy*, vol. 38, pp. 253–260, 2016.

[5] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. Aerts, "Artificial intelligence in radiology," *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, 2018.

[6] H.-M. Zhang and B. Dong, "A review on deep learning in medical image reconstruction," *Journal of the Operations Research Society of China*, vol. 8, pp. 311–340, 2020.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[8] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *International conference on machine learning*. PMLR, 2017, pp. 214–223.

[9] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," in *International conference on machine learning*. PMLR, 2016, pp. 1558–1566.

[10] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[11] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.

[12] 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.

[13] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in *Simulation and Synthesis in Medical Imaging: Third International Workshop, SASHIMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 3*. Springer, 2018, pp. 1–11.

[14] C. Han, L. Rundo, R. Araki, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, "Infinite brain MR images: PGGAN-based data augmentation for tumor detection," in *Neural approaches to dynamics of signal exchanges*. Springer, 2019, pp. 291–303.

[15] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[16] M. J. Chuquicusma, S. Hussein, J. Burt, and U. Bagci, "How to fool radiologists with generative adversarial networks? a visual turing test for lung cancer diagnosis," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 240–244.

[17] T. Han, S. Nebelung, C. Haarburger, N. Horst, S. Reinartz, D. Merhof, F. Kiessling, V. Schulz, and D. Truhn, "Breaking medical data sharing boundaries by using synthesized radiographs," *Science advances*, vol. 6, no. 49, p. eabb7973, 2020.

[18] B. Zhan, L. Zhou, Z. Li, X. Wu, Y. Pu, J. Zhou, Y. Wang, and D. Shen, "D2FE-GAN: Decoupled dual feature extraction based GAN for MRI image synthesis," *Knowledge-Based Systems*, vol. 252, p. 109362, 2022.

[19] J. Wang, G. Xie, Y. Huang, J. Lyu, F. Zheng, Y. Zheng, and Y. Jin, "FedMed-GAN: Federated domain translation on unsupervised cross-modality brain image synthesis," *Neurocomputing*, vol. 546, p. 126282, 2023.

[20] Y. Shi, H. Tang, M. J. Baine, M. A. Hollingsworth, H. Du, D. Zheng, C. Zhang, and H. Yu, "3DGAUnet: 3D generative adversarial networks with a 3D U-Net based generator to achieve the accurate and effective synthesis of clinical tumor image data for pancreatic cancer," *Cancers*, vol. 15, no. 23, 2023. [Online]. Available: https://www.mdpi.com/2072-6694/15/23/5496

[21] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4401–4410.

[22] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," *Advances in neural information processing systems*, vol. 33, pp. 12 104–12 114, 2020.

[23] V. A. Kelkar, D. S. Gotsis, F. J. Brooks, K. Prabhat, K. J. Myers, R. Zeng, and M. A. Anastasio, "Assessing the ability of generative adversarial networks to learn canonical medical image statistics," *IEEE transactions on medical imaging*, 2023.

[24] D. Gotsis, V. Kelkar, R. Deshpande, F. Brooks, P. KC, K. Myers, R. Zeng, and M. Anastasio, "Data for the 2023 AAPM grand challenge on deep generative modeling for learning medical image statistics," https://doi.org/10.13012/B2IDB-2773204_V1, 2023.

[25] C. D'Orsi, L. Bassett, S. Feig *et al.*, "Breast imaging reporting and data system (BI-RADS)," *Breast imaging atlas, 4th edn. American College of Radiology, Reston*, 2018.

[26] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," *Advances in neural information processing systems*, vol. 30, 2017.