CycleGAN-Driven Data Augmentation for Improved Neural Network Disease Detection

Annette Irungu[1], Kotaro Sonoda[2], Makoto Kawamoto[3], Kris Lami[3], Junya Fukuoka[3], and Senya Kiyasu[2]

[1] Graduate School of Engineering, Nagasaki University, Nagasaki 852-8521, Japan
[2] School of Information and Data Sciences, Nagasaki University, Nagasaki 852-8521, Japan
[3] Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8523, Japan

Abstract. Inadequate training data and class imbalances will often affect the generalizability of many deep learning models. Our work proposes a solution leveraging generative image-to-image translation as a data augmentation tool. We train a CycleGAN and utilize a semantic binary mask for controllable synthesis of pathology onto healthy cases. We validate our approach using the publicly available BReAst Carcinoma Subtyping (BRACS) dataset comprising breast histology images. Relying on binary masks means we retain the original features while introducing pathology without producing unrealistic synthetic imagery. When enhanced with classical data augmentation, our augmented dataset increases breast lesion detection capabilities. The model trained with the combined data has its area under the curve (AUC) closest to one, implying a minimal risk of missing potential positive diagnoses and the chance to identify potential breast cancer cases early. Our code is available at `https://github.com/Annette29/data-augmentation-cycleGAN.git`.

Keywords: Medical image synthesis · Adversarial learning · Image-to-image translation · Data augmentation

1 Introduction

Most datasets comprising whole slide images (WSIs) lack sufficient data to allow deep learning models to learn effectively. Many of these histology images are exceptionally large and contain complex information, making it time- and resource-consuming for professionals with medical degrees to analyze and annotate. Moreover, sharing medical data between institutions is complicated since a democratized dataset would require informed consent from healthcare professionals, patients, hospital administrators, and subsequent users on when and how the data could be used [3]. Also, since different countries often use unique protocols to code patient information, such a detailed guideline might still not be enough to allow information sharing to create a uniform large-scale dataset for various diseases.

However, using WSIs remains one of the best ways to provide cancer diagnosis and prognosis, implying most diagnostic models need to be trained using these images. But, doing so without addressing the issue of limited training sets that fail to fully represent underlying distributions risks creating diagnostic solutions that unexpectedly fail when introduced to unfamiliar data [6]. Therefore, the most straightforward solution, for now, is adding synthetic samples to the original data created using generative methods with photorealistic outputs.

Generative Adversarial Networks (GANs) [9] have recently been adopted as a data augmentation approach for medical images. GANs can create new data instances by learning from patterns in real data and producing synthetic examples that reflect those in the original dataset. These unsupervised models have achieved remarkable success in generating high-quality training data, leading to widespread application in furthering the study of rare diseases. An important caveat is that relying on random noise, as the original GAN model did, to synthesize high-resolution outputs, still requires large datasets for efficient training.

As a result, we train a cycle-consistent generative adversarial network (CycleGAN) to synthesize pathology onto images of healthy breast tissue using binary semantic masks to control the shape and location of breast lesions. Introducing these masks as additional input ensures that pathology is only in-painted in areas annotated by trained pathologists. Afterward, we augment the original dataset with the synthetic samples and train a DenseNet classification model to distinguish between samples with and without lesions to check if our approach improves generalization ability. Although breast cancer is unfortunately not a rare illness, most existing datasets designate patient cases as either malignant or benign, ignoring the large spectrum of lesions encountered during diagnosis and clinical artifacts that often present in non-standardized hospital images. Consequently, we aim to contribute to creating a cohort of WSIs encompassing a large variability of breast lesions, including precancerous ones that pathologists would encounter when advising patients.

Our paper is structured as follows: Section 2 discusses other works that have utilized GANs for medical image synthesis. Section 3 outlines our procedure for mask-guided image-to-image translation, including model architectural details and how to modify image patches sampled from the real data to fit the target output. Additionally, Section 4 specifies the training details, including the dataset used, while Section 5 shows slide-level qualitative results and elaborates on the domain adaptation approach to test the impact of the generated pathological images. Finally, Section 6 discusses the experimental findings and future improvements.

2 Related Work

Researchers have adopted GANs as an augmentation tool to create high-fidelity images to train supervised diagnostic models with demonstrated promising results. Notably, in [31], a multi-stream GAN that improves synthesis performance for brain images of glioma patients is introduced, while Madani *et al.* [21] com-

pare traditional augmentation methods to GAN-based ones for chest X-ray images and confirm improved accuracy when GANs are used. In [27], a two-stage GAN is used to simulate realistic ultrasound images, and a domain-aware framework has also been adopted to synthesize stroke lesions and kidney tumors [33].

Specifically, CycleGANs have been used to synthesize cerebral microbleeds in brain injury scans [8] and to automatically generate pathology in normal radiographs to improve bone lesion classification [11]. They have also helped generate CT images using a training dataset of MR images [19] and synthetic colorectal polyp images for less common polyp classes [28]. Additionally, in [6], CycleGAN is extended by introducing a region of interest discriminator for stain transfer and modified to create artificially stained renal tissue images [5].

Alternatively, adversarial training has been co-opted to create healthy images from pathological pairs without identity loss [30], including pix2pix to create anonymized MRI images with brain tumors [26] and progressive growing GANs to aggregate information from skin lesion features and improve melanoma detection [1]. Other instances of GANs favorably augmenting medical datasets were present when a GAN with fully convolutional networks estimated desired CT images given source MRI data [23], MI-GAN enhanced retinal image synthesis [16], and segmentation masks of microscopy images were applied to synthesize red blood cell images [3].

Unlike their diffusion counterparts that require long sampling times or Variational Auto-Encoders (VAEs) that may produce fuzzy, non-detailed images, GANs have consistently generated high-quality outputs with excellent details, making them ideal for medical image augmentation [18]. Although many works target X-ray, CT, and MRI data, whole slide images of breast tissue have received minimal attention. Therefore, we explore how the CycleGAN architecture can be modified for breast lesion synthesis and how such an augmented dataset impacts classification task performance.

3 Methodology

We describe a pathology factorization technique that relies on adversarial and cycle-consistent learning to increase the number of samples used to train a classification model, with the understanding that the method can be extended to images of any modality with any abnormality, even without additional annotations of the surrounding tissue.

3.1 Network Architecture

Zhu *et al.* [32] introduced CycleGAN to tackle an unpaired setting during image-to-image translation and alter the style of the input images without changing their content. CycleGAN comprises two generators, $G_{XY} : X \to Y$ and $G_{YX} : Y \to X$, with corresponding discriminators, D_Y and D_X. In our case, we have generators, $G_{HP} : H \to P$ and $G_{PH} : P \to H$ and discriminators, D_P and

D_H, where H and P are images of breast tissue in the healthy and pathological domains.

Our generators are UNetResNet34 models with a pre-trained ResNet34 network as the encoder inspired by [15], who combine a UNet and Residual Networks (ResNet) model for a segmentation task and observe improved performance. ResNet models employ residual blocks with skip connections that enhance gradient flow during backpropagation, enabling deeper networks to learn desired features better through identity mappings [14]. In addition to supporting the learning of complex transformations, ResNet's residual connections also help mitigate vanishing and exploding gradients by allowing direct gradient flow. On the other hand, UNet models are designed for biomedical image analysis tasks, making them ideal when detailed, pixel-wise maps of input images are required [12]. Our decoder comprises transposed convolutional layers connected to the corresponding encoder layers through skip connections, which is a key characteristic of UNet models [25]. Skip connections ensure that the decoder retains any spatial information that may have been lost during downsampling.

The discriminators follow the PatchGAN architecture in [17] as they output a classification matrix of probabilities determining how realistic patches of the generated image are, using values ranging from 0 to 1. PatchGAN divides the generator's output image into 70×70 patches and slides its field of view across the entire region while creating its output. Therefore, a PatchGAN discriminator will enforce photo-realism in synthetic images through:

$$D_{\text{patch}} : H \to \{0,1\}^{h \times w}. \tag{1}$$

by producing a $h \times w$ grid of outputs that is then compared with a corresponding grid of truth labels [6]. Although we train both generators, we focus on only one direction of synthesis ($H \to P$), as it is often rarer to obtain large numbers of samples in the pathological domain.

3.2 Mask-Based Translation

We use a semantic mask as additional input to the CycleGAN generators to guide the translation between the pathological and healthy domains. The grayscale image directs the data generation process, introducing some control on output modes, as suggested in [22]. We create a corresponding binary mask for each sample in the pathological data distribution as described in Section 4, but each healthy sample is assigned a pathology mask semi-randomly when fed into G_{HP}.

The mask is applied by multiplying it elementwise with the feature maps in the decoder to emphasize the shape and location of the lesions. By scaling the feature maps guided by the mask, we can localize areas on the binary mask with a value of 1 and suppress those with a value of 0. As a result, the values of the feature maps are directly influenced by those of the mask in a more aggressive attempt to highlight valuable regions. In the end, we are mindful that maintaining the structure of the original images is critical even when enhancing them to include pathological features. Additionally, feature-map elementwise multiplication addresses the one-to-many problem present in healthy-to-pathological

synthesis tasks since numerous ways exist to introduce pathology to a healthy sample. Notably, the masks provide explicit information about the shape and location of the desired pathological features.

3.3 Loss Functions

The learning of a typical CycleGAN is guided by three loss terms: main adversarial loss, cycle consistency loss, and an optional identity loss [32]. Cycle consistency and adversarial losses are necessary because a CycleGAN with the former alone does not produce realistic samples with discernible features. In contrast, one with adversarial loss only suffers from mode collapse [32]. In addition, inspired by [8], we introduce an abnormality mask loss in the second cycle to preserve the features of the synthetic healthy samples in regions not emphasized by the pathology mask used to create it. Therefore, we perform CycleGAN training according to:

$$\min_{G_{HP},G_{PH}} \max_{D_P,D_H} \mathcal{L}_{CG} = \mathcal{L}_{GAN}(G_{HP}, D_P, H, P) + \mathcal{L}_{GAN}(G_{PH}, D_H, H, P) \quad (2)$$

$$+ \lambda_{CYC}\mathcal{L}_{CYC}(G_{HP}, G_{PH}, H, P) \quad (3)$$

$$+ \lambda_{ID}\mathcal{L}_{ID}(G_{HP}, G_{PH}, H, P) \quad (4)$$

$$+ \lambda_{AM}\mathcal{L}_{AM}(G_{PH}, H, P). \quad (5)$$

where L_{GAN} is the main adversarial loss. In our case, L_{GAN} is the Wasserstein loss with gradient penalty [10] used by the PatchGAN discriminators, D_H and D_P. Gulrajani *et al.* [10] extended the original Wasserstein loss [2] by introducing a gradient penalty to encourage the discriminator's weights to have infinite values without raising its gradient norm above 1 to enforce 1-Lipshitz continuity [29].

In addition, L_{CYC} is the cycle consistency loss that encourages similarity between real and synthetic images with the same content and L_{ID} is the identity loss for color preservation in the synthetic outputs. Finally, L_{AM} the additional abnormality mask loss is used only in the second training cycle to guarantee that changes are restricted to pathological regions when removing pathology. $\lambda_{\mathrm{CYC}}, \lambda_{\mathrm{ID}}$, and λ_{AM} are manually adjusted weights that control the contribution of each loss function to the overall L_{CG}.

3.4 Synthesis Framework

The proposed approach has two training cycles: forward healthy-to-pathological-to-healthy (HPH) and backward pathological-to-healthy-to-pathological (PHP).

During the HPH cycle, G_{HP} receives a sample from the healthy data distribution and a pathology mask assigned semi-randomly as input. G_{HP} synthesizes pathology onto the input sample in the areas specified by the binary mask without altering the remaining features to create a synthetic pathological image as output. D_P receives this generator's output and attempts to distinguish it from real pathological images using L_{GAN}. L_{GAN} also informs G_{HP} how well it is doing in fooling D_P, to guide its learning.

Table 1: New BRACS data distribution according to lesion type.

	Without Lesions	With Lesions
Train	114	281
Validation	28	37
Test	18	69

Additionally, the synthetic image is concatenated with the semi-random mask that helped create it and fed into G_{PH} as input so it can use them to create synthetic healthy samples. This step is necessary to ensure cycle consistency, and so a pair of real and synthetic healthy images are compared using L_{CYC} by calculating the pixel distance between them [32]. Lastly, the healthy input sample is combined with an empty mask whose values are all zero and fed into G_{PH} which attempts to recreate a healthy image. L_{ID} helps analyze this second pair to confirm that no distortions are present and that G_{PH} maps any input to itself and does not modify any healthy samples.

The reverse occurs during the PHP cycle with a corresponding pair of a pathological sample and semantic mask fed into G_{PH}, and the output evaluated by D_H. In this case, D_H uses L_{GAN} to discriminate between real and synthetic healthy images and G_{HP} is the generator used to create samples to be compared when determining cycle consistency and identity preservation. But L_{AM} is also employed to confirm that all other image features and tissue structures are preserved when pathology is removed from the image, and a synthetic healthy sample is created by comparing the real pathological data to synthetic healthy data.

4 Experiments

4.1 Data

We validate our method on the Breast Carcinoma Subtyping (BRACS) dataset, which comprises Hematoxylin and Eosin (H&E)-stained WSIs of breast tissue [7]. BRACS has 547 images from 151 patients annotated through the consensus of three board-certified pathologists into three lesion types: 89 as atypical, 265 as benign, and 193 as malignant [7]. The authors further subdivide the data into seven categories: atypical into atypical ductal hyperplasia (ADH) and flat epithelial atypia (FEA), benign into normal tissue (N), pathological benign (PB), and usual ductal hyperplasia (UDH), and finally malignant into ductal carcinoma in situ (DCIS), and invasive carcinoma (IC) depending on the features of the tissue subtypes. A detailed explanation and a representative sample of the subtypes is provided in [7].

Since we focus on synthesizing pathology onto healthy cells, we reclassify the images into two groups: 387 in With Lesions and 160 in Without Lesions. The With Lesions category comprises images from the ADH, PB, DCIS, and IC

Table 2: Patch-level split according to lesion type.

	Without Lesions	With Lesions
Train	1630	3658
Validation	347	312
Test	187	571

categories, while Without Lesions has data from the FEA, N, and UDH types. Brancati *et al*. [7] provide pre-defined splits into training, validation, and test sets; we maintain those splits even after reclassification. For example, the file named `BRACS_1238.svs` belonged to the benign type and PB sub-category in the training dataset, and so even if we reclassify it into the With Lesions class, it remains in the training set. Table 1 reports the number of WSIs according to the new lesion types after reclassification.

Patch Sampling. The BRACS dataset utilizes sub-region annotations that outline regions of interest (ROIs), ranging over variable dimensions to include, in their entirety, every diagnostic lesion [7]. Hence, the authors include annotation files in qpdata format based on the QuPath software [4] for visualizing each ROI inside their respective WSI.

We loaded the Tiled TIFF image (SVS) files and corresponding qpdata annotations into QuPath for each image in the With Lesions class. Then, we extracted GeoJSON files that included coordinates describing the pathology's shape and location on the WSI. Next, we used the Rasterio Python library to read these geospatial data and create TIFF binary masks of the same size, distributed into the same training, validation, and test sets as their corresponding SVS images. However, the SVS WSI and TIFF binary masks proved too large to be imported directly during CycleGAN training, so we sampled relevant patches from each file. Guided by the GeoJSON coordinates to avoid the loss of diagnostically useful information, we extracted patches from the WSIs and binary masks. Therefore, each image and mask had different patches depending on how many ROIs were in each WSI.

Guided by [13], we divided each slide in the Without Lesions class into a random subset of non-background but recognizable 1024×1024 patches. A sample was considered a background image if more than 97.5% of its pixels exceeded an 85% intensity on all three RGB channels [13]. Additionally, we used the variance of the Laplacian to check focus quality and guarantee that all samples were of high enough quality and had no non-tissue objects. Based on these two conditions, certain SVS images yielded as many as 110 patches, while others had as few as five patches (with an average of 12). Finally, we conditionally applied random sampling for images with more than 20 patches, keeping at least 10 in such cases.

After extraction, our dataset comprised 6,705 patches, with 2,164 in the Without Lesions class and 4,541 in the With Lesions category, all distributed

into test, validation, and training sets that mirror their parent WSIs, as shown in Table 2. The patch files are in Portable Network Graphics (PNG) format, where the file `BRACS_1003732_x=47434_y=22955.png` describes a patch extracted from the WSI named `BRACS_1003732.svs` at the spatial coordinates (47434, 22955).

4.2 Training Details

We train our CycleGAN using the NVIDIA A100 GPU provided as part of the Google Colab Pro+ subscription package, with a batch size of eight randomly sampled image and mask pairs. Models use the Adam optimizer with a learning rate of 2e-4, $(\beta_0, \beta_1) = (0.5, 0.999)$, and randomly initialized weights.

The two generators comprise a UNet-style decoder and an encoder that reuses weights from a pre-trained ResNet34 model whose first layer is modified to accept four channels since the inputs combine an RGB image with a grayscale mask. The downsampling path comprises four residual blocks and a self-attention layer to help the model better understand the input pairs' overall structure. The UNet decoder has four ConvTranspose2D layers, which, together with increasing spatial resolution, perform element-wise multiplication to concatenate learned encoder features with upsampled masks. The final output is, thus, a 1×1 convolution that has been multiplied by the original mask to emphasize relevant pathology areas.

On the other hand, the PatchGAN discriminators comprise five convolutional layers that extract features at different scales (64, 128, 256, and 512) with a fixed 4×4 kernel size and stride of 2, except for the final two layers. This way, the discriminators can access input images patch-wise as spatial dimensions gradually reduce. We add LeakyReLU activations with a 0.2 negative slope to capture all subtle details without discarding negative values and InstanceNorm2D normalization, so our model is robust to contrast changes. Dropout with a probability of 0.1 is also applied throughout the generators and discriminators to address overfitting.

Inspired by [24], we introduce a stopping criterion during training instead of a fixed number of epochs by monitoring validation loss on previously unseen data every 20 epochs. Training stops if the validation loss does not significantly improve after five consecutive validation checks, and model checkpoints are saved whenever the loss improves.

5 Evaluation

5.1 Patch-Level Qualitative Results

Since the BRACS dataset contains expert annotations for each image, we can evaluate model performance by approximating the location of the lesions and their expected shape once pathology synthesis has been performed. Our primary focus is a healthy-to-pathological translation, so we first produce synthetic pathological images for each of the 187 without-lesion patches in the test set. We

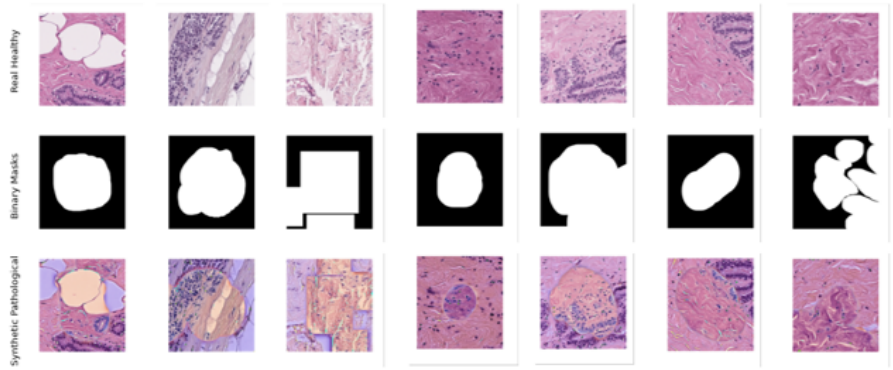
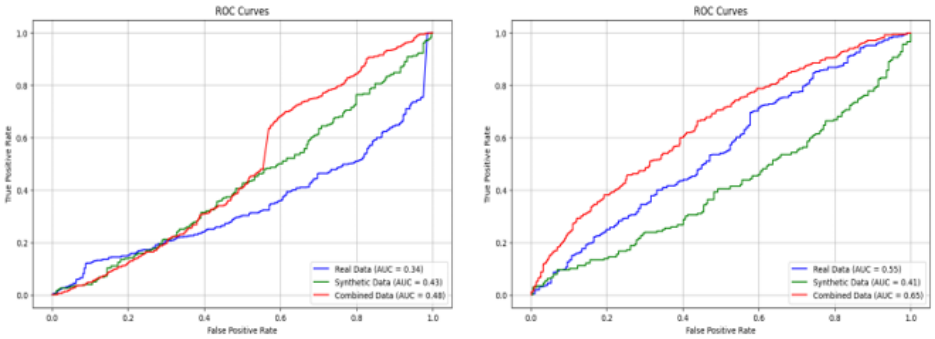


Fig. 1: Sample patches from the BRACS dataset with random binary masks and the resulting synthetic images created using G_{HP}.



(a) Without classical augmentation **(b)** With classical augmentation

Fig. 2: ROC performance on six training instances with original, generated, and combined data.

visualize samples in Fig. 1 for the proposed method alongside the real healthy sample and binary mask. One may observe that the network accurately localizes the pathology location and shape because it learns the shape and boundaries from the semantic mask. However, some of the generated data has different intensity values than the original image, for instance, the output on the second and third columns in Fig. 1. We hypothesize that the model profits from the one-to-one translation with the synthesis framework centered around the binary masks, as many of the output images contain realistic tissue structures.

5.2 Domain Adaptation

We ran an additional experiment designed as a classification task explicitly focusing on how using synthetic images to train the model affected network performance. We extend the DenseNet-121 model introduced by [12] by adding fully

connected layers after the original DenseNet blocks. First, we freeze the model's initial layers and leverage pre-trained ImageNet weights for feature extraction. Next, we regularize the network with dilated blocks of linear transformations, batch normalization, ReLU activations, and dropout layers to reduce dimensionality and ensure the model can recognize lesions varying in size and location. The primary loss function is the focal loss $(\alpha, \gamma) = (2, 3)$ since we intend for our model to focus on less common lesions and learn better from often-misclassified samples to avoid bias towards healthy images. Proposed in [20], the authors define focal loss as:

$$FL(p_t) = -(1 - p_t)^{\gamma} \log(p_t). \tag{6}$$

where p_t is the model's estimate for the true class label, meaning $(1 - p_t)$ reduces the contributions that easy-to-classify examples make to the overall loss value. Lin *et al.* [20] reshape the more common cross-entropy loss to address class imbalances and emphasize the impact of less common samples on gradient updates.

We train three different model instances using real data only, synthetic data only, and a combined set that samples images from the real and synthetic data using a 1:1 ratio. We also train the same number of model instances, but this time, each data distribution is augmented using classical methods, specifically horizontal and vertical flipping, rotation, shifting, scaling, shearing, Gaussian noise implemented as a custom transformation, and randomly applied brightness and contrast.

The classification performance of the six model instances is evaluated by plotting a Receiver Operating Characteristics (ROC) curve and calculating the area under the curve (AUC). The ROC curves in Fig. 2 visually represent how well each model can distinguish between images with and without lesions and how the different instances tradeoff between true and false positive rates.

A slight but consistent improvement is evident when using the combined dataset augmented with traditional approaches, as this model instance attains the highest AUC score. However, when synthetic data is used independently, the network has a significantly lower score, and we speculate this could be because some of the augmentations alter the appearance of the pathology introduced by the binary masks, confusing the model. This conflict between synthetic data and classical augmentations meant to change image features is unexpected and merits further exploration in a later study.

Training without classical augmentation methods reduces the AUC score for all models, but the instance trained with real and synthetic data maintains a higher score. Interestingly, model performance is worst when using real data without any augmentations, possibly because of the limited diversity and lack of edge cases in the original samples.

We confirm that adopting generated images as an augmentation technique will improve the performance of classification models, particularly their generalization ability. The slight benefit implies that adopting a translation approach when creating synthetic samples and using them alongside real data will maximize model capability for detection tasks.

6 Discussion

This paper discusses an approach to synthesize pathology onto healthy tissue samples guided by binary masks as a conditional element introduced to a modified CycleGAN framework. The generated data is then combined with real images to increase an otherwise small dataset meaningfully. The impact of such an augmentation strategy is then studied through a lesion classification task, resulting in a marginal improvement, particularly in how effective the model is at differentiating between classes, making it a more reliable diagnostic tool.

In summation, based on the experiments we have described, using GANs for data generation guided by a binary mask to localize the appearance of pathology and further combining the synthetic samples with the original images to train diagnostic models benefits generalization performance. Nonetheless, the additional overhead, particularly the heavy computation resources required and the model's unstable training process, may not justify the slight improvement depending on the specific medical task. Although our technique still relies on expert annotations describing the pathology's location on the image, our findings have revealed a possibility for further automation, for instance, when shifting between different imaging protocols.

Currently, our approach is limited since the model can only differentiate between samples with and without lesions. We intend to introduce an advanced dataset split that will include healthy tissues and pathological subtypes to fine-tune the solution and equip it with more detailed control. We also plan to synthesize binary masks and introduce different pathology sizes and shapes that address the variability encountered in real-world data and focus on less common subtypes underrepresented in breast cancer datasets.

References

1. Abdelhalim, I.S.A., Mohamed, M.F., Mahdy, Y.B.: Data augmentation for skin lesion using self-attention based progressive generative adversarial network **165**, 113922. https://doi.org/10.1016/j.eswa.2020.113922
2. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: Precup, D., Teh, Y.W. (eds.) Proceedings of the 34th International Conference on Machine Learning. vol. 70, pp. 214–223. PMLR
3. Bailo, O., Ham, D., Shin, Y.M.: Red blood cell image generation for data augmentation using conditional generative adversarial networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 1039–1048. IEEE. https://doi.org/10.1109/CVPRW.2019.00136
4. Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., James, J.A., Salto-Tellez, M., Hamilton, P.W.: QuPath: Open source software for digital pathology image analysis **7**(1), 16878. https://doi.org/10.1038/s41598-017-17204-5
5. de Bel, T., Hermsen, M., Kers, J., van der Laak, J., Litjens, G.: Stain-transforming cycle-consistent generative adversarial networks for improved segmentation of renal histopathology. In: Medical Imaging with Deep Learning. vol. 102, pp. 151–163 (8-10 July 2019)

6. Boyd, J., Villa, I., Mathieu, M.C., Deutsch, E., Paragios, N., Vakalopoulou, M., Christodoulidis, S.: Region-guided CycleGANs for stain transfer in whole slide images. In: 25th International Conference on Medical Image Computing and Computer Assisted Intervention. pp. 356–365. Springer Nature
7. Brancati, N., Anniciello, A.M., Pati, P., Riccio, D., Scognamiglio, G., Jaume, G., De Pietro, G., Di Bonito, M., Foncubierta, A., Botti, G., Gabrani, M., Feroce, F., Frucci, M.: BRACS: A Dataset for BReAst Carcinoma Subtyping in H&E Histology Images **2022**, baac093. https://doi.org/10.1093/database/baac093
8. Faryna, K., Koschmieder, K., Paul, M.M., van den Heuvel, T., van der Eerden, A., Manniesing, R., van Ginneken, B.: Adversarial cycle-consistent synthesis of cerebral microbleeds for data augmentation. In: 34th Annual Conference on Neural Information Processing Systems
9. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A.C., Bengio, Y.: Generative adversarial nets. In: NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems. vol. 2, pp. 2672–2680. MIT Press. https://doi.org/10.1145/3422622
10. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.: Improved training of Wasserstein GANs. In: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. pp. 5769–5779. Curran Associates, Inc.
11. Gupta, A., Venkatesh, S., Chopra, S., Ledig, C.: Generative image translation for data augmentation of bone lesion pathology **102**, 225–235
12. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778. IEEE. https://doi.org/10.1109/CVPR.2016.90
13. Hosseini, M.S., Chan, L., Tse, G., Tang, M., Deng, J., Norouzi, S., Rowsell, C., Plataniotis, K.N., Damaskinos, S.: Atlas of Digital Pathology: A Generalized Hierarchical Histological Tissue Type-Annotated Database for Deep Learning. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11739–11748. IEEE. https://doi.org/10.1109/CVPR.2019.01202
14. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely Connected Convolutional Networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4700–4708. IEEE. https://doi.org/10.1109/CVPR.2017.243
15. Huang, Z., Qu, E., Meng, Y., Zhang, M., Wei, Q., Bai, X., Zhang, X.: Deep learning-based pelvic levator hiatus segmentation from ultrasound images **9**(10), 100412. https://doi.org/10.1016/j.ejro.2022.100412
16. Iqbal, T., Ali, H.: Generative Adversarial Network for Medical Images (MI-GAN) **42**(11), 231. https://doi.org/10.1007/s10916-018-1072-9
17. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5967–5976. IEEE. https://doi.org/10.1109/CVPR.2017.632
18. Kebaili, A., Lapuyade-Lahorgue, J., Ruan, S.: Deep learning approaches for data augmentation in medical imaging: A review **9**(81). https://doi.org/10.3390/jimaging9040081
19. Li, W., Kazemifar, S., Bai, T., Nguyen, D., Weng, Y., Li, Y., Xia, J., Xiong, J., Xie, Y., Owrangi, A., Jiang, S.: Synthesizing CT images from MR images with deep learning: Model generalization for different datasets through transfer learning **7**(2), 025020. https://doi.org/10.1088/2057-1976/abe3a7

20. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2999–3007. IEEE. https://doi.org/10.1109/ICCV.2017.324
21. Madani, A., Moradi, M., Karargyris, A., Syeda-Mahmood, T.: Chest x-ray generation and data augmentation for cardiovascular abnormality classification. In: Angelini, E.D., Landman, B.A. (eds.) Medical Imaging 2018: Image Processing. vol. 10574. SPIE. https://doi.org/10.1117/12.2293971
22. Mirza, M., Osindero, S.: Conditional generative adversarial nets. https://doi.org/10.48550/arXiv.1411.1784
23. Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q., Shen, D.: Medical image synthesis with deep convolutional adversarial networks **65**(12), 2720–2730. https://doi.org/10.1109/TBME.2018.2814538
24. Prechelt, L.: Early stopping — but when? In: Montavon, G., Orr, G.B., Müller, K.R. (eds.) Neural Networks: Tricks of the Trade: Second Edition. pp. 53–67. Springer. https://doi.org/10.1007/978-3-642-35289-8_5
25. 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.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. pp. 234–241. Springer International Publishing
26. Shin, H.C., Tenenholtz, N.A., Rogers, J.K., Schwarz, C.G., Senjem, M.L., Gunter, J.L., Andriole, K.P., Michalski, M.: Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In: Gooya, A., Goksel, O., Oguz, I., Burgos, N. (eds.) SASHIMI 2018. pp. 1–11. Lecture Notes in Computer Science, Springer International Publishing
27. Tom, F., Sheet, D.: Simulating patho-realistic ultrasound images using deep generative networks with adversarial learning. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). pp. 1174–1177. IEEE. https://doi.org/10.1109/ISBI.2018.8363780
28. Wei, J.W., Suriawinata, A.A., Vaickus, L.J., Ren, B., Liu, X., Wei, J., Hassanpour, S.: Generative image translation for data augmentation in colorectal histopathology images **116**, 10–24
29. Wei, X., Gong, B., Liu, Z., Lu, W., Wang, L.: Improving the Improved Training of Wasserstein GANs: A Consistency Term and Its Dual Effect. In: 6th International Conference on Learning Representations
30. Xia, T., Chartsias, A., Tsaftaris, S.A.: Pseudo-healthy synthesis with pathology disentanglement and adversarial learning **64**, 101719. https://doi.org/10.1016/j.media.2020.101719
31. Yurt, M., Dar, S.U., Erdem, A., Erdem, E., Oguz, K.K., Çukur, T.: mustGAN: Multi-stream generative adversarial networks for MR image synthesis **70**, 101944. https://doi.org/10.1016/j.media.2020.101944
32. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2242–2251. IEEE. https://doi.org/10.1109/ICCV.2017.244
33. Zhu, Q., Yin, L., Tang, Q., Wang, Y., Cheng, Y., Li, S.: DCAug: Domain-aware and content-consistent cross-cycle framework for tumor augmentation. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) 26th International Conference of Medical Image Computing and Computer Assisted Intervention. Lecture Notes in Computer Science, vol. 14224, pp. 338–347. Springer Nature Switzerland (2023 8-12 October). https://doi.org/10.1007/978-3-031-43904-9_33