SYNTHETIC HISTOPATHOLOGY LYMPHOCYTE IMAGES AUGMENTATION USING CONDITIONAL GENERATIVE ADVERSARIAL NETWORKS

Alexandra-Georgiana ANDREI[1], Bogdan IONESCU[2]

Rezumat. *Analiza imaginilor histopatologice este utilizată pe scară largă și este o metodă esențială pentru diagnosticarea și clasificarea cancerului, inclusiv a cancerului colorectal. Cu toate acestea, din cauza lipsei datelor și a procedurilor laborioase de adnotare, obținerea unei colecții semnificative și variate de imagini histopatologice pentru antrenarea modelelor de învățare automată continuă să fie dificilă. Pentru a rezolva această problemă, sugerăm o metodă Latent-to-Image care creează imagini histopatologice sintetice pentru cancerul colorectal, utilizând rețele generative adversariale condiționale (cGAN-uri). În acest articol, investigăm utilizarea cGAN-urilor, specific pentru generarea de imagini pentru țesutul limfocitar, folosind șapte clase diferite de țesut pentru antrenare. Rezultatele experimentale arată că imaginile sintetice generate nu pot fi diferențiate de imaginile histopatologice reale, capturând caracteristicile texturale și structurale distinctive ale țesutului. Metoda propusă în acest articol dovedește abilitatea de a genera imagini de înaltă calitate, cu un Fréchet Inception Distance de 21,2. Imaginile generate au fost evaluate, de asemenea, de patru patologi și nu s-a constatat nicio diferență semnificativă între imaginile reale și cele generate.*

Abstract. *Histopathology image analysis is widely used and is essential for diagnosis and cancer grading, including colorectal cancer. However, due to a lack of availability and labor-intensive annotation procedures, getting a significant and varied collection of histopathology images for training machine learning models continues to be difficult. To solve this problem, we suggest a Latent-to-Image method that creates synthetic colorectal histopathology images using Conditional Generative Adversarial Networks (cGANs). In this article, we investigate the use of cGANs specifically for generating images for lymphocytes tissue, using seven different tissue classes for training. The results show that the generated synthetic images are indistinguishable from the real histopathological images, capturing the distinctive textural and structural characteristics of the tissue. We show that our method generates high quality images with a Fréchet Inception Distance of 21.2. The generated images were also assessed by four pathologists and no significant difference between the real and generated images were found.*

Keywords: histopathology, data augmentation, synthetic data, Generative Adversarial Networks.

DOI **10.56082/annalsarsciinfo.2025.1.36**

[1] Title: PhD student, AI Multimedia Lab, National University of Science and Technology Politehnica Bucharest, Bucharest, Romania (Email: alexandra.andrei@upb.ro).
[2] Title: Professor, AI Multimedia Lab, National University of Science and Technology Politehnica Bucharest, Bucharest, Romania, associate member of Academy of Romanian Scientists (Email: bogdan.ionescu@upb.ro).

1. Introduction

Histopathology is a branch of pathology dedicated to the study of diseases through the microscopic examination of tissue samples. This discipline plays a crucial role in medical diagnosis by enabling clinicians and researchers to analyze cellular structures and identify abnormalities that can indicate specific disease processes. In the context of colorectal cancer—a malignancy that originates in the colon, the final part of the digestive tract—accurate histopathological analysis is especially vital. Colorectal cancer is not only one of the most common cancers worldwide but also demonstrates significant heterogeneity in cellular composition and morphology, which directly impacts disease staging and subsequent treatment planning [1]. While it predominately impacts older individuals, it has the protentional to manifest at any stage of life. According to World Health Organization, it is the third most common cancer worldwide, accounting for approximately 10% of all cancer cases and it is the second leading cause of cancer-related deaths worldwide [2].

Deep learning techniques [3] have increasingly been integrated into histopathology image processing as they offer new capabilities for automating and refining the diagnostic process. However, successful implementation of these approaches relies heavily on large volumes of annotated images from various tissue subtypes to train robust models. Unfortunately, acquiring such datasets poses unique challenges due to the sensitivity of medical data, confidentiality constraints, legal concerns, and the inherent complexity and labor-intensive nature of precise annotation. Data augmentation [4] has therefore become an essential strategy to enhance the quantity and diversity of available training samples, ultimately supporting improved model generalization without the need for additional new data acquisition.

Generative Adversarial Networks (GANs) [5] have emerged as a particularly effective means of data augmentation within this domain. In this paper, we present a method employed to generate colorectal cancer histopathology images, patches of lymphocytes (LYM) tissue using a Conditional Generative Adversarial Network (cGAN) that was trained on multiple tissue classes. The accurate depiction of lymphocytes is critical in histopathological evaluation since these immune cells serve as important indicators of disease progression and patient prognosis, particularly in the tumour microenvironment [6]. Our approach leverages the strengths of the latent-to-image paradigm to generate high-fidelity images that capture the subtle morphological characteristics of lymphocytes. By augmenting existing datasets with synthetic but realistic lymphocyte images, our method aims to address the data scarcity challenge, thereby enhancing subsequent tasks such as cell segmentation and disease classification.

The rest of this paper is structured as follows: Section 2 reviews the current state-of-the-art GANs application in the histopathology field; Section 3 presents the methods used for generating and evaluating the generated images and the dataset used in this study. The results achieved in this paper are presented in Section 4 and the paper concludes with Section 5.

2. Related work

One key use of GANs in digital histopathology is stain color normalization [7][8]. Tissue slide images often look very different because of how labs stain samples, how thick the sections are, the scanner used, and how the image is saved. GANs learn to shift the colors of each slide so they all match a chosen reference slide, without changing the actual tissue details. Doing this before analysis makes the images more uniform, which helps machine-learning models see true biological patterns rather than just color differences. As a result, predictions become more reliable and diagnostic accuracy can improve.

Another important use of GANs in medical imaging is image enhancement [9]. Scans and microscope images can suffer from noise, low resolution or uneven contrast, which makes it harder to see fine details. GANs learn to translate poor-quality inputs into clearer, higher-resolution outputs—reducing blur and speckle without inventing false structures. As a result, subtle features become easier to detect, aiding both human review and automated analysis.

GANs also enable virtual staining, where raw or label-free images are digitally colored to mimic traditional chemical stains [10]. Instead of physically applying dyes—an often costly, time-consuming and sometimes destructive step—GANs map label-free scans (like autofluorescence or brightfield) onto the appearance of standard stains (e.g. H&E (hematoxylin and eosin) or immunofluorescence). This preserves the same tissue morphology while giving pathologists familiar visual cues, speeding workflows and avoiding reagent use.

Another task that can be performed by GANs is ink and mark removal on histology slides [11]. Pathologists often annotate glass slides or digital scans with pens, markers or stamps, which can obscure tissue details and confuse analysis pipelines. A GAN trained on paired “marked” and “clean” images learns to erase these annotations cleanly, restoring the underlying tissue patterns. The corrected images then feed into diagnostic models or archival systems without losing any true histologic information.

To address the limited availability of annotated data, data augmentation techniques have been developed to expand the datasets. In particular, GANs have shown to be a suitable solution for generating realistic synthetic images across various domains. GANs, developed in 2014 by Goodfellow et al. [12], are a class

of deep learning models comprising two neural networks: the generator and the discriminator. The generator takes the input (noise, image) and generates fake samples that resemble the real data that was used for training. The discriminator is a binary classifier that has the role to distinguish between real and synthetic data generated by the generator. The two models are trained together, and they are competing against each other playing an adversarial game where the generator tries to fool the discriminator. Based on the input type applied to the generator, two main approaches have been developed: the latent to image approach and the image-to-image approach.

Latent to image approach implies that the generator's input consists of a latent vector or random noise. Image to image approach implies that the GAN produces output images conditioned on a specific input image. The generator's input consists of images from one domain and aims to generate images in a different domain (domain transfer) [13].

Image synthesis methods using GANs have been studied and developed to mitigate this problem in many fields, including healthcare. Different types of medical images were obtained using GANs. Chen et al. [14] used texture-embedded GANs to synthesize 3D pulmonary nodules computer tomography (CT) images, Prezja et al. [15] proposed a CycleGAN to generate knee osteoarthritis radiographs, Groceri et al. [16] proposed a Deep GAN-based augmentation method for dermoscopy images and Fakuda et al. [17] explore the use of StyleGAN to generate synthetic panoramic radiographs. Ho et al. [18] show that diffusion probabilistic models can generate high-fidelity images comparable to those generated by GANs. Puria Azadi et al. [19] use diffusion probabilistic models to synthesize high quality histopathology images of brain cancer.

Over the past years several studies have been conducted using GANs in the field of digital histopathology for image augmentation. Different GAN architectures were used, for example Quiros et al. [20] proposed a framework called PathologyGAN to generate H&E colorectal and breast cancer tissues from a structured latent space. The model combines BigGAN, StyleGAN and Relativistic Average Discriminator. Deshpande et al. [21] proposed a method to generate realistic and annotated colorectal cancer histology images from glandular layout inputs and Wei et al. [22] proposed a method to generate synthetic colorectal polyp images from normal colonic mucosa images using CycleGANs.

Considering these advances, we therefore introduce a latent-to-image augmentation framework that synthesizes high-fidelity lymphocyte images to enrich scarce datasets. In our approach, a 1,500-dimensional noise vector is simply concatenated with a tissue label and projected into a small multi-channel

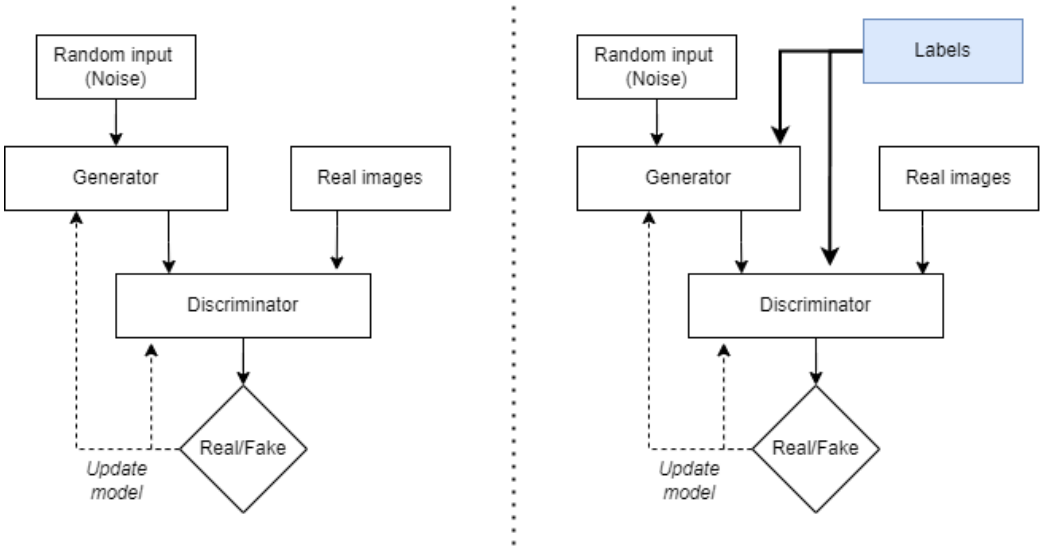

Fig. 1. Schematic representation of the difference between Deep Generative Adversarial Networks and Conditional Generative Adversarial Networks.

feature map, which is then upsampled through a cascade of transposed-convolutional layers to yield realistic 128×128 patches. Compared to the traditional use of pretrained style encoders, rigid feature-map injections or paired image requirements, the generator is free to discover the full spectrum of lymphocyte morphologies directly from data, while a symmetrically conditioned discriminator ensures that only class-consistent textures and structures are learned. This unstructured latent design not only delivers greater sample diversity but also scales effortlessly to additional tissue classes.

3. Methods

3.1. Proposed Generative Adversarial Network

Generative Adversarial Networks [10] are used to generate new data by learning an implicit representation of the dataset distribution. As presented in the Introduction, GANs consist of a pair of networks: a generator and discriminator. Noise is used as input to the generator (G) and outputs an image that represents the input for the discriminator (D) which determines if it is fake or real. Both neural networks are trained in a competitive manner, until G learns the data distribution and D is not able to differentiate between fake and real data anymore. There are different GAN architectures, and the one proposed in this paper is a

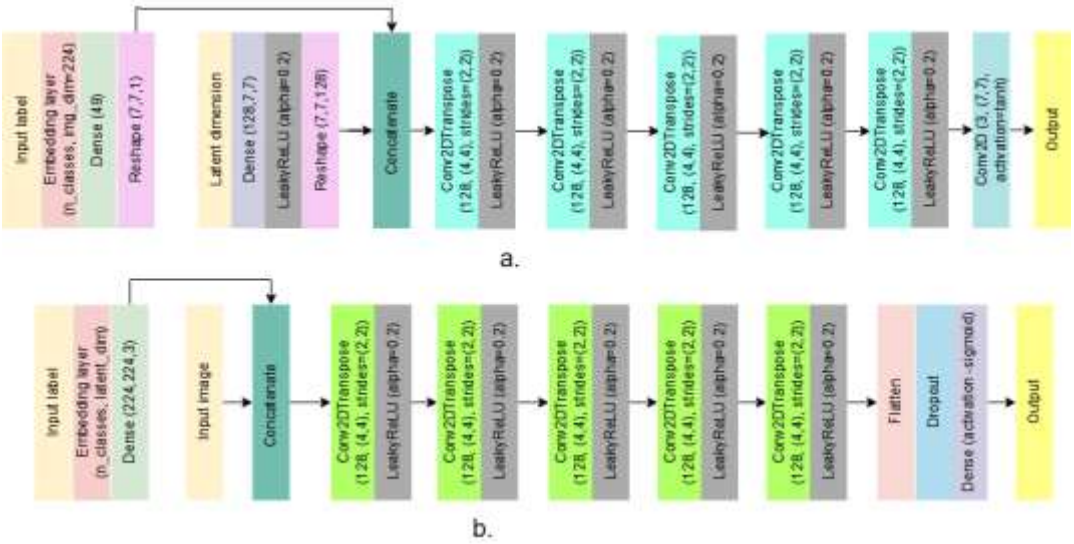

Fig. 2. Overview of proposed GAN architecture: a) Generator, b) Discriminator.

conditional GAN [23]. Conditional GANs, are an extension of a GAN, in which a conditional setting is applied so both the generator and discriminator are conditioned by auxiliary information such as class label or data, as exemplified in Fig. 1. One of the advantages of cGANs is that the output of the generator can be controlled by giving the label to the desired class.

We proposed a Latent-to-Image approach for generating data. This implies that our network generates images based on unstructured latent space (noise). The model is depicted in Fig. 2. The conditional generator is designed to contain 5 deconvolutional layers with a kernel size (4,4), 128 filters and stride 2. A concatenate layer is used for the latent vector of size 1,500 and the labels, and the result is reshaped and used as input for the first deconvolutional layer. Each deconvolutional layer is followed by a LeakyRelu [24] activation function with a slope of 0.2 and the last convolutional layer use hyperbolic tangent function. The discriminator contains 5 convolutional blocks.

Each convolutional block contains a 4 x 4 convolutional layer with kernel size (4,4) and stride 2 and followed by a LeakyRelu activation layer. The last convolution layer of the discriminator is flattened and then fed into a sigmoid output. Our model is trained using a batch size of 125 by Adam optimizer with learning rate 2e-4 (beta1 = 0.5) and cross-entropy loss. 89.434 images from 8 different classes [adipose (ADI), debris (DEB), lymphocytes (LYM), mucus

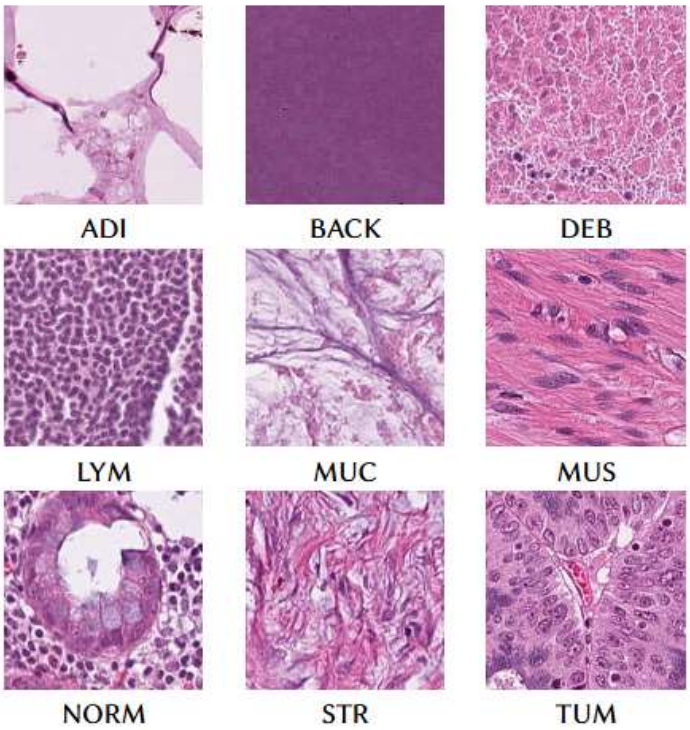

Fig. 3. Examples of real images for each tissue type contained in the"NCT-CRC-HE-100K" dataset [21].

(MUC), smooth muscle (MUS), normal colon mucosa (NORM), cancer-associated stroma (STR), colorectal adenocarcinoma epithelium (TUM)] were used for training to generate 2,534 images for LYM tissue.

3.2. Dataset

The dataset used in this paper is the "NCT-CRC-HE-100K" dataset [25]. It consists of 100,000 non-overlapping image patches extracted from 86 H&E stained human cancer tissue slides and normal tissue from the NCT biobank (National Center for Tumor Diseases) and the UMM pathology archive (University Medical Center Mannheim). It was created by pathologists by manually delineating tissue regions in whole slide images into the following nine tissue classes: Adipose (ADI), background (BACK), debris (DEB), lymphocytes (LYM), mucus (MUC), smooth muscle (MUS), normal colon mucosa (NORM), cancer-associated stroma (STR), colorectal adenocarcinoma epithelium (TUM). Examples of real images for each tissue type are depicted in Figure 3. Excluding BACK and ADI classes which do not contain relevant information for the diagnosis, all the other classes were used in this study.

3.3. Evaluation

3.3.1. Fréchet Inception Distance

We assessed the fidelity and variety of our synthetic lymphocyte images using the Fréchet Inception Distance (FID) [26], a widely adopted quantitative metric for generative modeling. FID computes the distance between the multivariate Gaussian distributions fitted to Inception-v3 feature embeddings of real versus generated image sets, thereby capturing differences in both visual quality and sample diversity. In practice, a lower FID score indicates that the synthetic images more closely match the real data distribution in terms of texture, structure, and overall variability. By reporting FID alongside our qualitative results, we provide an objective measure of how well our latent-to-image framework reproduces the complex morphological patterns found in authentic lymphocyte specimens.

3.3.2. Perceptual Experiments

We conducted a Turing test to validate whether the synthesized images are realistic and withstand clinical examination. The test included 4 specialists who performed the test with no time limit. The test included other tasks, but the one relevant to this paper is the following: each expert was presented with a mixed set of real and generated lymphocyte patches and asked simply to identify which ones were artificial.

4. Results

Fig. 4 shows a selection of our generated lymphocyte images alongside their real counterparts: in each row the left-most patch is drawn from the original dataset, while the three adjacent patches are synthesized by our latent-to-image model. Visually, the synthetic images capture not only the overall staining intensity and color distribution of the real samples, but also the fine-grained textural patterns – such as chromatin granularity and cytoplasmic granularity – as well as the characteristic nuclear shapes, sizes, and clustering behaviour of lymphocytes. Importantly, these images show the same variety of morphological phenotypes found in real data, demonstrating that our approach can produce high-fidelity augmentations that are virtually indistinguishable from authentic histopathological images.

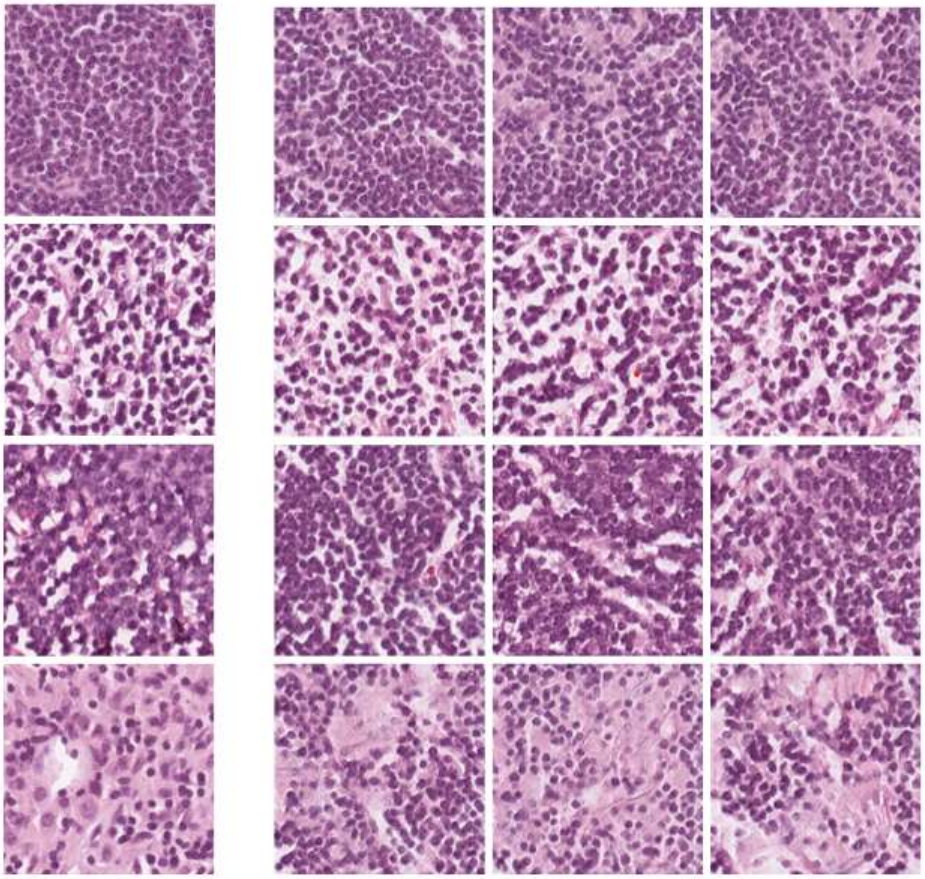

Fig. 4. Examples of real and synthetic generated images of LYM: (a) real images, (b) synthetic generated images with the method presented in this paper.

As a quantitative measure, we computed the FID score between all generated images and an equal number of real images from the dataset. Our method achieved an FID of 21.20, substantially outperforming PathologyGAN's [20] score of 32.05 on mixed tissue classes (Table 1). This reduction in FID indicates that the synthetic images align more closely with the real data distribution in terms of both low- and high-level image statistics, reflecting improvements in texture realism and inter-sample variability.

Table 1. Evaluation using FID score.

Method	*Tissue type*	*FID score*
Our method	Lymphocytes	21.20
PathologyGAN [20]	All classes	32.05

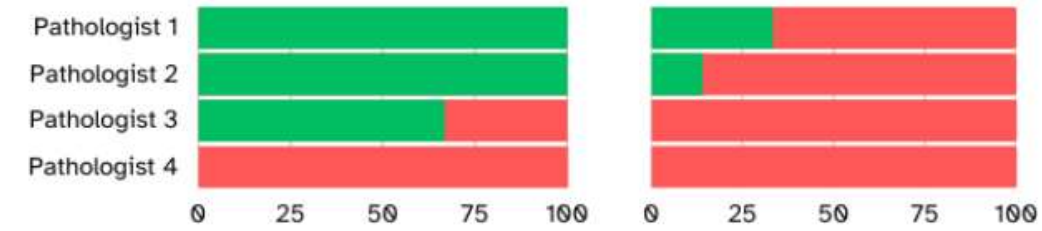

Fig. 5. Results for the Turing test – evaluation of real (left) and synthetic (right) images. Green – percentage of correctly classified images; red - percentage of incorrectly classified images.

To assess perceptual indistinguishability, we conducted a Turing test validation with four pathologists who were asked to label a shuffled mix of real and synthetic lymphocyte images without time constraints. The results are depicted in Fig. 5. The specialists identified real images correctly 55% of the time and labelled synthetic patches as fake only 33 % of the time. Pathologists' ability to identify real images as being real is marginally above 50% expected by random guessing. This suggests that even real images in isolation are not easy to be identified as "real" when mixed with high-quality generated images. On the other hand, specialists labelled just 33% of the generated patches as being synthetic, meaning that two out of three generated images were identified as real. These findings confirm that our latent-to-image framework produces lymphocyte images that withstands clinical examination, effectively closing the gap between synthetic augmentation and authentic histopathological.

Conclusions

Synthetic images serve as valuable tools to improve the generalization capabilities of machine learning algorithms. In this paper we proposed a conditional generative model that uses a Z2I approach to generate images out of noise. We demonstrated the ability to generate histological images of human colorectal cancer for lymphocytes tissue type using a cGAN that was trained on 8 different tissue classes. The contributions of our work can be summarized as follows: we expand the size and variety of the available dataset by generating synthetic images, which enables more thorough and robust training of machine learning models.

Furthermore, we can extend the study by incorporating image generation methods from other classes present in the dataset. This expansion allows us to investigate the capabilities of the trained model in generating images beyond lymphocyte tissues and assess its generalization abilities.

R E F E R E N C E S

[1] PDQ® Adult Treatment Editorial Board. PDQ Colon Cancer Treatment. Bethesda, MD: National Cancer Institute. Updated <10/03/2025>. Available at: https://www.cancer.gov/types/colorectal/patient/colon-treatment-pdq. Accessed 10/04/22025. [PMID: 26389319]

[2] World Health Organization: WHO and World Health Organization: WHO (2023) Colorectal cancer. https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer. Accessed 10/04/22025

[3] Van der Laak, J., Litjens, G. and Ciompi, F., 2021. Deep learning in histopathology: the path to the clinic. Nature medicine, 27(5), pp.775-784

[4] Shorten, C. and Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of big data, 6(1), pp.1-48.

[5] Andrei, A.G., Constantin, M.G., Graziani, M., Müller, H. and Ionescu, B., 2025. Privacy preserving histopathological image augmentation with Conditional Generative Adversarial Networks. Pattern Recognition Letters, 188, pp.185-192.

[6] Hendry, S., Salgado, R., Gevaert, T., Russell, P.A., John, T., Thapa, B., Christie, M., Van De Vijver, K., Estrada, M.V., Gonzalez-Ericsson, P.I. and Sanders, M., 2017. Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the international immunooncology biomarkers working group: part 1: assessing the host immune response, TILs in invasive breast carcinoma and ductal carcinoma in situ, metastatic tumor deposits and areas for further research. Advances in anatomic pathology, 24(5), pp.235-251.

[7] Shaban MT, Baur C, Navab N, Albarqouni S. Staingan: Stain style transfer for digital histological images. In2019 Ieee 16th international symposium on biomedical imaging (Isbi 2019) 2019 Apr 8 (pp. 953-956). IEEE.

[8] Du Z, Zhang P, Huang X, Hu Z, Yang G, Xi M, Liu D. Deeply supervised two stage generative adversarial network for stain normalization. Scientific Reports. 2025 Feb 27;15(1):7068.

[9] Rong R, Wang S, Zhang X, Wen Z, Cheng X, Jia L, Yang DM, Xie Y, Zhan X, Xiao G. Enhanced pathology image quality with restore–generative adversarial network. The American Journal of Pathology. 2023 Apr 1;193(4):404-16.

[10] Bai B, Yang X, Li Y, Zhang Y, Pillar N, Ozcan A. Deep learning-enabled virtual histological staining of biological samples. Light: Science & Applications. 2023 Mar 3;12(1):57

[11] Jiang J, Prodduturi N, Chen D, Gu Q, Flotte T, Feng Q, Hart S. Image-to-image translation for automatic ink removal in whole slide images. Journal of Medical Imaging. 2020 Sep 1;7(5):057502-.

[12]I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Advances in neural information processing systems 27 (2014).

[13] Dalva, Y., Altındiş, S.F. and Dundar, A., 2022, October. Vecgan: Image-to-image translation with interpretable latent directions. In European conference on computer vision (pp. 153-169). Cham: Springer Nature Switzerland.

[14] Chen YC, Chiu LY, Lee CE, Huang WC, Chen LW, Lin MW, Yang AS, Ye YZ, Ou DX, Chang YC, Chen CM. Texture-embedded Generative Adversarial Nets for the synthesis of 3D pulmonary nodules computed tomography images. Expert Systems with Applications. 2025 May 15;274:126860.

[15] Prezja F, Annala L, Kiiskinen S, Lahtinen S, Ojala T, Nieminen P. Generating synthetic past and future states of Knee Osteoarthritis radiographs using Cycle-Consistent Generative Adversarial Neural Networks. Computers in Biology and Medicine. 2025 Mar 1;187:109785.

[16] Goceri E. GAN based augmentation using a hybrid loss function for dermoscopy images. Artificial Intelligence Review. 2024 Aug 7;57(9):234.
[17] Fukuda M, Kotaki S, Nozawa M, Tsuji K, Watanabe M, Akiyama H, Ariji Y. An attempt to generate panoramic radiographs including jaw cysts using StyleGAN3. Dentomaxillofacial Radiology. 2024 Nov;53(8):535-41.
[18] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. Advances in neural information processing systems. 2020;33:6840-51.
[19] Moghadam PA, Van Dalen S, Martin KC, Lennerz J, Yip S, Farahani H, Bashashati A. A morphology focused diffusion probabilistic model for synthesis of histopathology images. InProceedings of the IEEE/CVF winter conference on applications of computer vision 2023 (pp. 2000-2009).
[20] Quiros AC, Murray-Smith R, Yuan K. Pathologygan: Learning deep representations of cancer tissue. arXiv preprint arXiv:1907.02644. 2019 Jul 4.
[21] Deshpande S, Minhas F, Rajpoot N. Synthesis of annotated colon cancer tissue images from gland layout. InMedical Imaging 2024: Digital and Computational Pathology 2024 Apr 3 (Vol. 12933, pp. 101-113). SPIE.
[22] Wei J, Suriawinata A, Vaickus L, Ren B, Liu X, Wei J, Hassanpour S. Generative image translation for data augmentation in colorectal histopathology images. Proceedings of machine learning research. 2019 Dec;116:10.9
[23] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784. 2014 Nov 6.
[24] A.L. Maas, A.Y. Hannun, A.Y. Ng, Rectifier nonlinearities improve neural network acoustic models, in: International Conference on Machine Learning, Vol. 30, 2013, p. 3
[25] Kather, J. Nikolas, Halama, Niels, M. Alexander, 100,000 histological images of human colorectal cancer and healthy tissue, Zenodo, 2018. URL: https://doi.org/10.5281/zenodo.1214456.
[26] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, Gans trained by a two time-scale update rule converge to a local nash equilibrium, Advances in neural information processing systems 30 (2017).