Image Synthesis for Prostate Cancer Biopsies Using Conditional Deep Convolutional Generative Adversarial Network

Paul Wahome Kariuki, *Member, IEEE*, Patrick Kinyua Gikunda, and John Mwangi Wandeto

***Abstract*— Deep learning models have shown promising results in complementing computer-aided diagnosis. However, the availability of quality medical image data sets is inhibited by the limited and imbalanced data sets available, privacy issues, and the high cost of generating labeled medical imaging data sets. When a deep learning model is trained on such data, its performance is greatly impeded, with a high chance of overfitting, and in turn, limiting the model's ability to scale and generalize well to previously unseen data. More recently, empirical studies have set the stage for image synthesis as a solution to these problems using generative adversarial neural networks. This paper presents a novel conditional DCGAN capable of growing synthetic prostate cancer biopsy whole slide images augmented from the PANDA histopathology data set. From the results, the novel Conditional Deep Convolutional Generative Adversarial Network model achieved an average Fréchet Inception Distance score of 1.3 with high quality synthetic images. The results validated the model's exceptional ability to generate realistic images with high discriminatory accuracy and low generator loss.**

***Index Terms*— Conditional Deep Convolutional Generative Adversarial Network, Data Augmentation, Image Synthesis, Prostate Cancer.**

I. Introduction

Prostate cancer (PCa) is among the most prevalent male cancers worldwide, accounting for over 1.2 million new diagnoses and over 350,000 deaths annually [1]. Previous empirical reports have shown that Black men have a two- to four-times higher risk of acquiring prostate cancer compared to all other races and ethnic groups [2] [7].

The path to diagnosing prostate cancer can involve several steps, with the biopsy serving as a crucial confirmation method. Two primary approaches are used for initial screening: the Prostate-Specific Antigen (PSA) and the Digital Rectal Exam (DRE). PSA is a blood test which measures the protein levels produced by the prostate; elevated levels can indicate a potential need for further investigation, but not necessarily cancer. On the other hand, DRE involves the doctor manually feeling the prostate for abnormalities in size or texture through the rectum. If the initial screening raises suspicion, a biopsy is often recommended.

A biopsy is a minimally invasive procedure that involves taking tiny tissue samples from the prostate for microscopic examination. Tissue samples are extracted with a needle, coated with a protective layer, dyed with hematoxylin and eosin (H&E), and preserved in glass containers. Diagnosis and prognosis require a skilled pathologist to examine the tissue sample under a microscope to identify malignant patterns using the Gleason classification system. However, this process is time-consuming and characterized by inter and intra-observer variability [8]. The appalling ratio of pathologists to patients, especially in developing nations, makes this situation even worse.

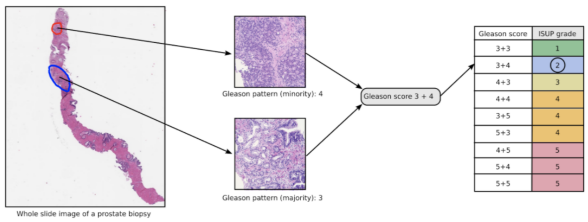

Fig. 1. Gleason grading procedure for a sample biopsy with prostate cancer. The predominant cancer growth patterns observed in the biopsy are outlined in blue, representing Gleason pattern 3, and the red outline shows Gleason pattern 4. These patterns determine the Gleason score for this biopsy, which is 3+4.

Figure 1 shows a representation of the Gleason grading procedure for a sample biopsy that has prostate cancer. The Gleason score is then converted into an ISUP grade, which is 2 for this biopsy sample, according to the guidelines set by the International Society of Urological Pathology (ISUP). Biopsies that do not contain cancer are assigned an ISUP grade of 0 [9].

Machine learning (ML) has shown great precision in the diagnosis of prostate lesion and prediction of patient outcomes, including survival and response to treatment, when used in diagnostic imaging [3] [10]. It has the potential to complement medical professionals in detecting diseases at an early stage therefore reducing the spread of diseases and high mortality rates [4]. In the case of prostate cancer, for instance, ML classifiers accurately identified Gleason patterns with an even higher level of accuracy compared to experienced radiologists [11]. Despite the impressive performance, medical imaging

First A. Author is a PhD student at Dedan Kimathi University of Technology, Kenya (e-mail: kariuki.paul21@students.dkut.ac.ke).
Second B. Author is a Lecturer at Dedan Kimathi University of Technology, Kenya (e-mail: patrick.gikunda@dkut.ac.ke).
Third C. Author is a Senior Lecturer at Dedan Kimathi University of Technology, Kenya (e-mail: john.wandeto@dkut.ac.ke).

datasets suffer from insufficiency, resulting to potential overfitting during ML model training. To counter this, Generative Adversarial Networks (GANs) have been used in medical image synthesis for PET, CT, MRI, ultrasound, and X-ray imaging [5] [6]. Several studies have leveraged GANs to aid in the diagnosis and prognosis of brain, abdomen, and chest-related diseases, as shown in [12]. A recent study on colorectal cancer using synthetic MRI found that there was no clear difference between regular and synthetic T2-weighted images in terms of signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), or other measures of image quality. Furthermore, GANs have shown the ability to generate authentic artificial brain MRIs [13]. All these collective facts support our hypothesis that deep learning image synthesis may be used to generate synthetic PCa images that are applicable in clinical settings, therefore alleviating the challenges associated with the availability of large medical image data sets for PCa computer-aided diagnosis.

In this study, we demonstrate that ML projects can employ CNN to generate synthetic PCa images using improved discriminator and generator models in the form of Conditional Deep Convolutional GANs (CDCGANs). Consequently, this work serves as the first endeavour to showcase the capability of CDCGANs in generating synthetic images that can accurately simulate traditional prostate biopsy images using the PANDA challenge dataset [9]. We propose a novel conditional Deep Convolutional Generative Adversarial Network model (Med-CDCGAN) with a global-local generator to augment the PCa biopsy images from the dataset. Both the generator and discriminator are designed as dual networks, which enables them to learn both global and local information independently from the biopsy images. This is a significant advantage, as it allows the networks to capture a more comprehensive understanding of the imaging data. By learning global information, the networks can identify broad patterns and features, while the ability to learn local information allows for the detection of more precise details.

II. Related Work

A. Generative Models for Medical Image Augmentation

In 2014, Goodfellow et al. [14] invented Generative Adversarial Networks (GANs). These networks are comprised of two separate neural networks, the generator G and the discriminator D. G takes a random noise vector $z \in p_z$ as input and outputs synthetic data $G(z)$; D takes as input, the output of the generator $G(z)$ and real images $x \in p_data$ to classify them as real or synthetic. The goal of training D is to maximize the probability of assigning the correct label to real $D(x)$ and synthetic data $D(G(z))$. Simultaneously, G is trained to minimize $\log(1 - D(G(z)))$. Hence, G and D play a two-player minimax game with value function $V(G, D)$:

$$\begin{aligned}\min_G \max_D \text{GAN}(G, D) = E_{x\sim p_{\text{data}}(x)}[\log D(x)] \\ + E_{z\sim p_z(z)}[\log(1 - D(G(z)))]\end{aligned} \tag{1}$$

where, x represents real data sampled from the data distribution $pdata$ and z is a d-dimensional noise vector sampled from a Gaussian distribution pz.

Following the first introduction of the first GAN framework by [14], the possibility of enhancing this architecture to generate even better and more realistic synthetic images based on a specific input label was left open. Conditional GANs generate images by factoring auxiliary variables such as labels [23], text data [24, 25], or images [26, 27, 28]. In conditional GANs, both the discriminator and the generator networks are provided with the conditioning variable, denoted as c in Equation (2). During image synthesis, the generator incorporates this extra information, while the discriminator bases its choice on both the conditioning variable and the received image. The input to the discriminator consists of a real pair, which includes an authentic image from a certain distribution and its matching label. In contrast, the false pair comprises a synthetic image and its label. The objective function of a conditional GAN is

$$\begin{aligned}\min_G \max_D cGAN(G, D) = E_{x,c\sim p_{\text{data}}(x,c)}[\log D(x, c)] \\ + E_{x',c\sim p_{\text{data}}(x',c)}[\log(1 - D(G(x', c)))],\end{aligned} \tag{2}$$

where x' is the image generated from latent noise.

Mirza and Osindero [15] developed a conditional GAN as an improvement of the original GAN Framework. This Conditional Generative Adversarial Network (GAN) used a supervised learning methodology, enabling precise control over the class of the produced outputs. Additionally, it offered the benefit of producing superior representations for generating multimodal images. Conditional models may be derived from GANs by augmenting both the generator and the discriminator with auxiliary or supplementary information.

During the training of a conditional GAN, it is imperative to reduce the cost function so that the generator can produce as close as possible images to the ground truth. Previous studies (e.g., [22] [23] [24]) have attempted to reduce the L1 or L2 distances between actual and produced image pairings, in addition to the GAN loss. This stage helps the generator in producing visuals that closely resemble real images. Minimizing the L1 distance results in sharper pictures compared to minimizing the L2 distance. Using the L1 distance enhances image sharpness in generation activities. Consequently, in this study, we adopt the L1 distance in our approach. The expression to minimize the L1 distance is:

$$\min_G L_{L1}(G) = E_{x,x'\sim p_{\text{data}}(x,x')}\left[\|x - x'\|_1\right], \tag{3}$$

B. Histopathology Image Synthesis

GANs have several applications in the interpretation of histology images [17]. Notably, important uses of this technology include correcting for staining variations, segmenting images using models that need supervision, generating synthetic data to enable learning with less supervision or without supervision, and generating more data to improve classification accuracy. This article primarily focuses on the use of enhanced GANs for image production and augmentation. The following are some notable studies referenced in the literature on this topic. Runz et al. [24] modified the CycleGAN framework to enhance data augmentation for lymph nodes. Instead of engaging in domain adaptation, they taught their network to seamlessly generate

and distinguish between normal and abnormal tissue images. From their results, they effectively constructed a generative model that could create equivalent samples from one class when given samples from the other class.

Two instances of using the DCGAN framework for histological image synthesis can be seen in the works of Xue et al. [25] and Breen et al. [26]. In the first study, the authors aimed to reduce the annotation effort required for developing accurate segmentation models by utilizing generative modelling techniques. They demonstrated that models trained using datasetGAN significantly reduced the amount of annotation required compared to traditional methods by achieving a correlation coefficient of 0.68, and further improved to 0.78 with transfer learning, indicating its effectiveness in reducing human annotation effort while maintaining high performance. In the second paper, the authors sought to use GANs for stain normalization in histopathology and compare traditional methods with modern GAN-based approaches. The study found that GAN-based methods significantly enhance model performance by reducing variability and improving diagnostic accuracy in histopathology

The work by Cornish [27] used a DCGAN framework to apply conditional GAN and DCGAN to generate HPV carcinoma patches. The study successfully demonstrated that using GANs for image-to-image translation could effectively generate segmentation masks from H&E-stained images. In Moghadam et al., [28], they looked into how diffusion probabilistic models can be used to make high-quality histopathology images. Specifically, they focused on making images that accurately show the morphological features of histopathological data. The performance of the model was evaluated through Inception Score (IS), Frechet Inception Distance (FID), and sFID. The results of the model were compared with ProGAN as follows; Inception Score: 2.08 (Diffusion Model) vs. 1.67 (ProGAN), Frechet Inception Distance (FID): 20.11 (Diffusion Model) vs. 53.85 (ProGAN), sFID: 6.32 (Diffusion Model) vs. 24.37 (ProGAN). They concluded that diffusion probabilistic models significantly outperformed ProGAN (a state-of-the-art GAN model) in generating high-quality histopathology images.

Teramoto et al. [26] reported another study using a similar architecture. A ProGAN architecture was used to generate lung cancer histological pictures in this instance. Their objective was to enhance the precision of their classification models in order to accurately categorize their samples as either benign or cancerous. They particularly said that there was a 4.3% improvement in the accuracy of their models. It is important to mention that the previous frameworks lacked the ability to do conditional synthesis. In this study, the authors introduced a multi-scale conditional GAN to generate and segment high-resolution, large-scale histopathology images. This approach consists of a hierarchical organization of GAN structures, where each level is specifically designed to generate and segment images at a different size.

In addition to the studies above, [29] investigated the use of conditional generative adversarial networks (CGANs), to detect microsatellite instability (MSI) from routine histology images in colorectal cancer (CRC). The study aimed to determine whether synthetic histology images generated by CGANs could retain information about genetic alterations and be used effectively to train deep learning models for MSI detection. The CGAN generated 10,000 synthetic MSI and non-MSI images, which were evaluated by trained observers and deemed realistic. The performance of MSI detection by the model achieved AUROC 0.777 on both real and synthetic data. The study therefore, concluded that synthetic histology images generated by CGANs retain critical information about genetic alterations in colorectal cancer (CRC).

C. The proposed Conditional DCGAN

This study presents a novel method that integrates progressive training and conditional synthesis to generate histopathology images for PCa. The model is guided by two main algorithms: Algorithm 1, which prepares the available images for synthesis, and Algorithm 2, which depicts the image synthesis process leading to augmentation.

1) How Algorithm 1 and 2 Work: Algorithm 1. aims at creating a dataset for training generator and discriminator models in the Conditional GAN. From [9], images whose size is less than 10MB are selected, resulting to at least 389 images, as referenced in Table I with each class (ISUP grades 0 - 5) well represented. Due to disparities in sizes of the images, all images are resized to a common size of (256×256). Consequently, to increase the diversity of the images for training the Conditional GAN models, each biopsy image for the ISUP classes is augmented 10 times through rotation, shearing, zooming and flipping. The algorithm concludes by combining all the augmented images for each class to be part of the training data, as summarized in Table I.

Description	Shape
Original train images	(389, 256, 256, 3)
Augmented train images	(3890, 256, 256, 3)
Combined train images	(4279, 256, 256, 3)

TABLE I
SHAPE OF TRAINING IMAGES GENERATOR AND DISCRIMINATOR NETWORK

The main objective of Algorithm 2 is to develop CDCGAN for generating synthetic images to be used in augmenting few sample images obtained from Algorithm 1. Like in a heuristic minimax game, the algorithm creates generator and discriminator models that try to win against each other. Over a specified number of iterations, a generator model creates a synthetic image from noise, with an aim of generating a realistic image indistinguishable with a real image. The model aims to minimize its loss function, while on the other hand, the discriminator model tries to maximize its loss function by comparing the generated image with the real images obtained from Algorithm 1. While trying to win against each other, the generator produces images that the discriminator cannot differentiate as real or fake.

III. MATERIALS AND METHODS

Algorithm 1 : Augmenting the Subset of Extracted PANDA Dataset

Require: Max File Size S, Target Size a, Max Num of Images Per Class B, Total classes C, Multiplier K (number of times each image is augmented), Max Total Augmented Images $M \leftarrow B \times K \times C$,

Ensure: Train dataset $\mathcal{T}$

Initialize: rotation_range, width_shift_range, height_shift_range, shear_range, zoom_range, horizontal_flip, vertical_flip, fill_mode as $\mathcal{D}$
Set *total_augmented* $\leftarrow 0$
while ImageDataGenerator is active **do**
  for $c \leftarrow 1$ **to** C **do**
    Set *total_augmented_per_class* $\leftarrow 0$
    for $i \leftarrow 1$ **to** B **do**
      for $j \leftarrow 1$ **to** K **do**
        Set rotation_range in $B(i)$;
        Set width_shift_range in $B(i)$;
        Set height_shift_range in $B(i)$;
        Set shear_range in $B(i)$;
        Set zoom_range in $B(i)$;
        Set horizontal_flip in $B(i)$;
        Set vertical_flip in $B(i)$;
        Update *total_augmented* $\leftarrow$ total_augmented +1
        Update *total_augmented_per_class* $\leftarrow$ total_augmented_per_class +1
      end for
    end for
    if total_augmented_per_class $\geq$ B $\times$ K **then**
      break
  end for
  if total_augmented$\geq$ M **then**
    break
  Update $\mathcal{T} \leftarrow \mathcal{D}(B)$
end while

Algorithm 2 Conditional DCGAN for Image Generation

Require: λ Learning rate, α Latent noise dim, β channels, ω monitors, set of synthetic images $x'_i \in X$ and real images $y_i \in Y$, set of labels $c \in C$, number of epochs T, generator network update per step (K_g), discriminator network update per step (K_d), Generator model parameters θ_G, Discriminator model parameters θ_D, Generator loss L_G, Discriminator loss L_D

Ensure: Output, Conditional DCGAN model θ

Initialize: ω and θ
while θ has not converged **do**
  for $t = 1, \ldots, T$ **do**
    for $k = 1, \ldots, K_g$ **do**
      Sample minibatch of latent noise z_i and corresponding labels c_i;
      Generate synthetic image $x'_i = G(z_i, c_i; \theta_G)$;
      Update generator θ_G: $\theta_G \leftarrow \theta_G - \lambda \nabla_{\theta_G} L_G(\theta_G, \theta_D)$;
    end for
    for $k = 1, \ldots, K_d$ **do**
      Sample minibatch of real images y_i and corresponding labels c_i;
      Sample minibatch of synthetic images x'_i and corresponding labels c_i;
      Update discriminator θ_D: $\theta_D \leftarrow \theta_D - \lambda \nabla_{\theta_D} L_D(\theta_D, y_i, x'_i)$;
    end for
  end for
  Compute monitors ω
end while

A. The Data set

The study used a data set from [9], which remains the largest publicly available dataset of prostate biopsies to date. A total of 10,616 whole slide images (WSIs) comprising 5160 and 5456 WSIs collected from the Radboud University Medical Center and the Karolinska Institute, respectively [16].

The two universities used unique scanners with somewhat varying maximum microscopic resolutions with varying annotation procedures. The Radboud University Medical Center obtained labels for the 5160 Whole Slide Images (WSIs) from pathology reports, which included the initial diagnosis. A skilled pathologist interpreted the 5456 WSIs at the Karolinska Institute. Hence, the labels of 5160 Whole Slide Images (WSIs) obtained from Radboud University Medical Center may exhibit a greater level of inconsistency compared to the labels from the Karolinska Institute.

B. Experiment Design and Setup

In this paper, we extended the epistemology of the original GAN network by [14], which laid the foundation for creating realistic data samples (image, video and audio) which are difficult to distinguish from real samples. Using the original network topology as a baseline, image preprocessing techniques, loss function, and other relevant components, the study created a conditional deep convolution generative adversarial network called Med-CDCGAN. First, because the PCa images from the [9] were characterized by high dimensionality, presence of non-tissue lesions in the original WSIs and huge in size, the steps in Algorithm 1 were followed for data engineering. The images were loaded with filtering based on maximum file size, maximum number of images to load, and resized to a target size of (32×32) pixels. To increase the diversity of biopsy images for GAN, augmentation techniques such as rotation, width shift, height shift, shear, zoom, horizontal flip, vertical flip were used. During augmentation, each WSI was augmented 10 times and expanded to match the input shape expected by the generator (1, 32, 32, 3). Due to the large image sizes, tiling was used to divide the large images into smaller, more manageable sections of WSIs to generate tiled images as shown in Table I. Original and augmented images were concatenated, resulting in a training set that contained both original and augmented images. For the use and faster training of Algorithm 2, all combined images from Table 1 were resized to a shape of (32×32) pixels. The generator architecture was designed to generate realistic images by

combining random noise vectors with class labels. The input layer consisted of two components: one for random noise with a specified shape and the other for class labels. The class labels were embedded using an embedding layer and flattened to match the noise input shape. These inputs were then concatenated to form a combined input tensor. This model employed Conv2DTranspose layers for spatial upsampling, progressively increasing the resolution from 4×4 to 32×32 pixels while reducing channel complexity. Batch normalization and ReLU activation, and momentum, were used to maintain stability and introduce nonlinearity. In the output layer, we utilized tanh activation to generate high-quality images while preserving the desired spatial dimensions.

On the other hand, the discriminator model was structured to evaluate the authenticity of input images alongside class labels. It started with convolutional layers that progressively extract features while downsampling the spatial dimensions. For stability and non-linearity, leaky ReLU activation and batch normalization were used. Subsequently, flattened features underwent dropout regularization before being concatenated with embedded class labels. This concatenated representation went through a dense layer with sigmoid activation, which gave a binary classification output that tells us whether the input image-label pairs are real or not.

During the training phase of the Med-CDCGAN model, we continuously monitored the generator gradients and subsequently updated the weights to ensure the optimization of the generator network by adjusting the weights based on the gradients computed during backpropagation. In addition to gradient calculations, we implemented performance metrics throughout the training process to evaluate the effectiveness of the model. These metrics included the Fréchet Inception Distance (FID), the Sharpness of the generated images, the mean square error (MSE) and the peak signal-to-noise ratio (PSNR).

Med-CDCGAN training spanned 250 epochs, with specific hyperparameters aimed at optimizing model training. These hyperparameters included a discriminator learning rate (DLR) of 2×10^{-5}, a generator learning rate (GLR) of 2×10^{-4}, momentum of 0.9 for smoothing images in the batch normalization layers, and a batch size of 128 for efficient mini-batch training. These hyperparameters were carefully chosen to strike a balance between model convergence speed and stability, resulting in a well-trained model capable of generating high-quality and diverse PCa images.

IV. Results

The performance of the Med-CDCGAN model was evaluated based on Fréchet Inception Distance (FID), Sharpness, Peak Signal-to-Noise Ratio (PSNR), Generator and Discriminator losses.

A. Fréchet Inception Distance (FID)

The Fréchet Inception Distance (FID) measures the similarity between generated and real images. It quantifies the distance between two multivariate Gaussian distributions, one representing the features of real images and the other representing the features of generated images. The FID was calculated using the formula below:

$$\text{FID}(X, Y) = ||\mu_X - \mu_Y||^2 + \text{Tr}(C_X + C_Y - 2(C_X C_Y)^{1/2}) \tag{4}$$

- X and Y represent the feature representations of real and generated images, respectively.
- μ denotes the mean of the feature representations.
- C represents the covariance matrix of the feature representations.
- Tr denotes the trace of a matrix.
- $^{1/2}$ represents the square root.

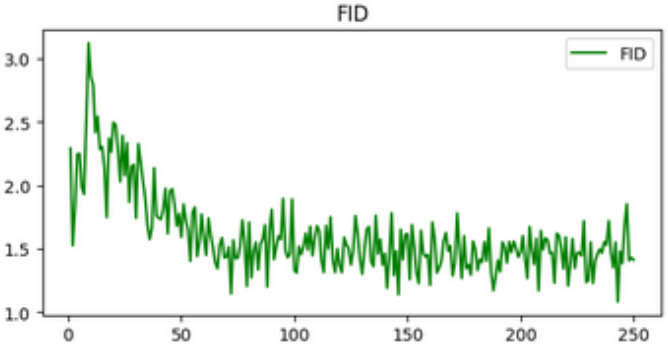

Fig. 2. Fréchet Inception Distance Score

A lower FID score indicates a closer match between the distributions, which means better image quality and similarity to real images.

B. Sharpness

Sharpness evaluates the clarity and detail levels of an image. In this experiment, it was used to indicate the ability of the generated image to preserve fine details and edges, which are often critical in medical diagnoses. This was calculated as follows:

$$\text{Sharpness}(I) = \frac{1}{N}\sum_{i=1}^{N}(\Delta I_i - \mu_{\Delta I})^2 \tag{5}$$

- I represents the image.
- N is the number of pixels in the image.
- ΔI_i represents the Laplacian of the pixel i.
- $\mu_{\Delta I}$ is the mean of the Laplacian values of all pixels in the image.

The Laplacian ΔI of an image I is calculated as:

$$\Delta I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{6}$$

A high sharpness value depicts a synthetic image with well-defined edges and fine details. These are essential attributes for accurate medical analysis and interpretation.

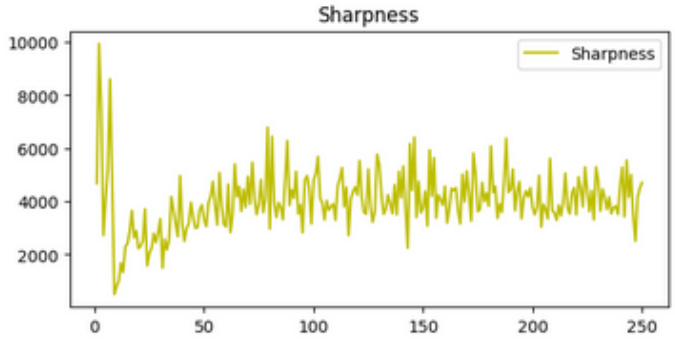

Fig. 3. Sharpness

C. Peak Signal-to-Noise Ratio (PSNR)

PSNR measures the ratio between the maximum possible power of a signal (the original image) and the power of the noise (the difference between the original and generated images). In essence, PSNR provides valuable insights into the noise level present in generated medical images, offering a quantitative measure of image quality and noise reduction efficacy. Central to PSNR is Mean Squared Error (MSE). The MSE quantifies the average squared difference between pixel values in the original and generated images.

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(X_i - Y_i)^2 \tag{7}$$

- N is the number of pixels in the image.
- X_i represents the pixel value of the original image at position i.
- Y_i represents the pixel value of the generated image at position i.

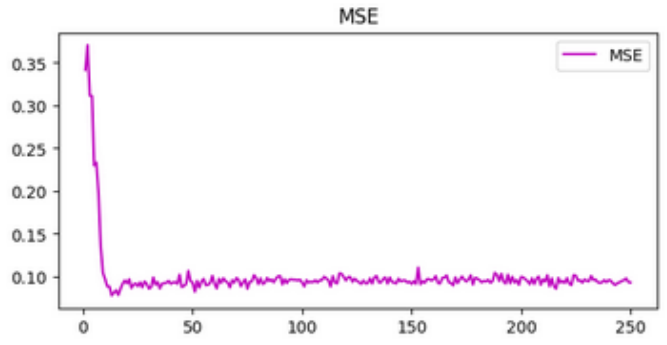

Fig. 4. Mean Squared Error

A lower MSE indicates a closer match between the images from the generator and discriminator, reflecting higher image fidelity and similarity. This has also been evidenced in our work in reference to Figure. 4 where the MSE error curve over the training epochs showed a downward trend signifying a continued improvement of generated images.

With MSE, PSNR is given as:

$$\text{PSNR} = 20 \cdot \log_{10}\left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}}\right) \tag{8}$$

- MAX_I is the maximum possible pixel value of the image. For an 8-bit image, this is typically 255.

With this in mind, our computed PSNR over the 250 epochs is illustrated in Figure 5. with an average PSNR of 10.64 decibels

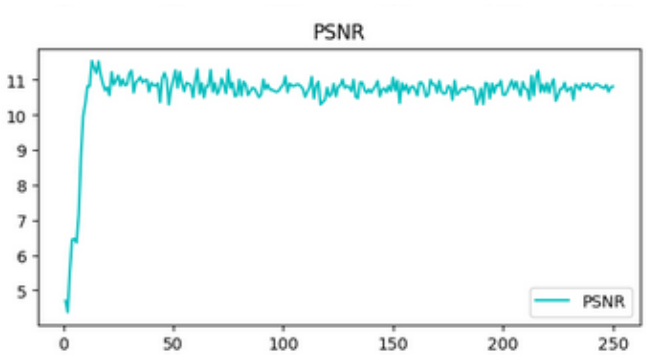

Fig. 5. Peak Signal-to-Noise Ratio

(dB). A higher PSNR values indicates better image quality and reduced noise.

D. Generator and Discriminator losses

Generator loss measures how well a generator model is able to fool a discriminator. On the other hand, the discriminator loss measures how well the discriminator model is able to distinguish a real image from a fake one. As training continues, a low generator loss indicates the ability of the generator model to fool the discriminator successfully. At the same time, a low discriminator loss signifies a stronger generator model, and the discriminator is struggling to differentiate a fake image from a real one.

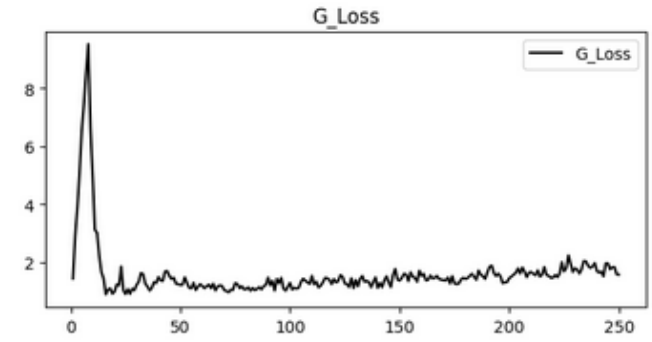

Fig. 6. Generator loss

As shown in Figure 6, our generator model is able to create fake images that the discriminator cannot discern whether it is real or fake, as the training continues. In Figure. 7, the loss of the discriminator showed a downward trend with fluctuations over time.

In both Figures. 6 and 7, each model improved over time and showed evidence of stabilization when neither the generator nor the discriminator could fool the other. At this point, the models had reached an equilibrium point, and the synthetic images generated looked almost similar to the real images for the proposed Med-CDCGAN.

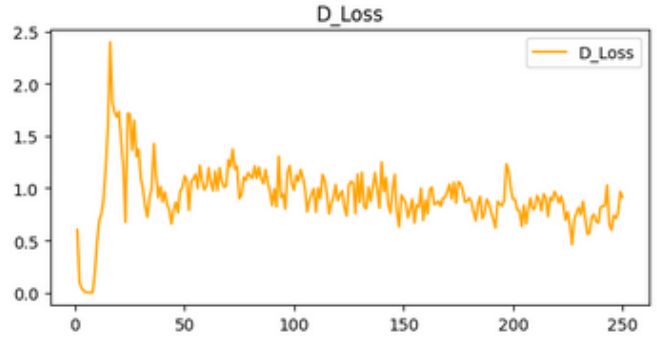

Fig. 7. Discriminator loss

V. Discussion

In this paper, we developed and evaluated a novel conditional deep convolutional generative adversarial network (Med−CDCGAN) for the synthesis of prostate cancer images. Our objective was to generate high-quality synthetic images that can effectively augment limited medical imaging datasets. The performance of Conditional DCGAN was assessed using a variety of evaluation metrics described in the above section. Training was examined through the generator and discriminator losses (plotted on the same scale), as shown in Figure 8. The generator loss, though initially high, exhibited a significant reduction and subsequent stabilization during the training epochs. This performance indicated that the generator had made a progressive improvement in producing realistic images that could increasingly fool the discriminator. In contrast, the discriminator loss, which started at a relatively low value, also showed a stabilization trend, demonstrating that the discriminator struggles in differentiating between real and generated images. The initial fluctuations in both losses are indicative of the adversarial training process, where both models improve by learning from each other's mistakes through a Min-Max-inspired game theory. The eventual stabilization of these losses suggests that the model has reached a balanced state, where the generator produces images of high quality, and the discriminator struggles to distinguish between real and synthetic images. The Fréchet Inception Distance depicted a generally decreasing trend, although with some fluctuations. A lower FID score signifies that the generated images were becoming increasingly similar to the real images in terms of feature representation. The trend decrease in FID score over the training period indicated a progressive improvement in the generator's performance. In addition, the Peak Signal-to-Noise Ratio (PSNR), also stabilizes at higher values after initial fluctuations, indicating that the generated images have higher quality with less noise, suggesting refined image generation over time.

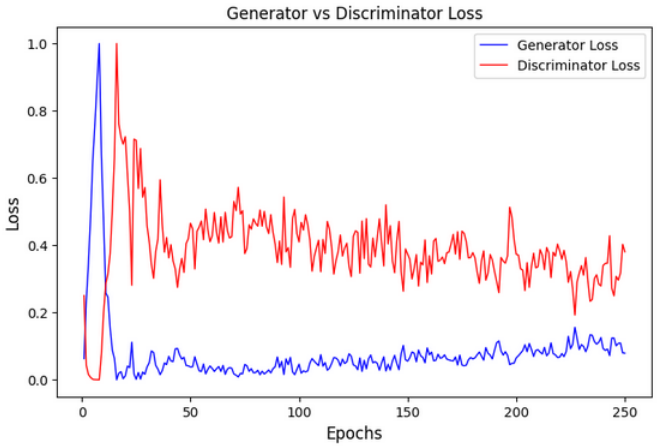

Fig. 8. Generator Loss and Discriminator Loss

After training, several synthetic images of prostate biopsies were generated, resized to the shape of, (256×256) and plotted as illustrated in Figure 9 and Figure 10. showing the proof of concept for Med−CDCGAN.

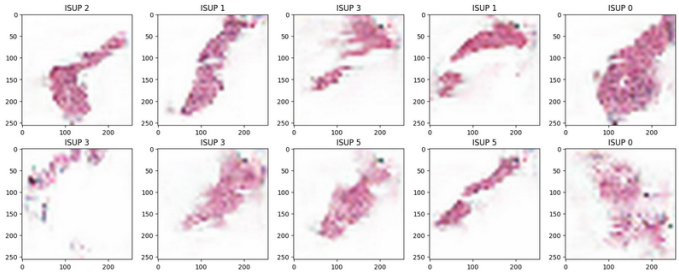

Fig. 9. Generated Images from Med-CDCGAN

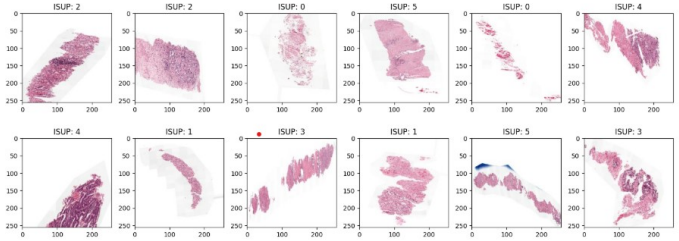

Fig. 10. Real Images from PANDA Dataset

VI. Results comparison of Med − CDCGAN

FID scores remain the most predominant metrics used in the evaluation of GAN. According to [30], a low FID result indicates a closer match of the synthetic image compared to the real image and can range between 0 to infinity. Similarly, it indicates that the quality and divergence of the generated images are high. The results of our GAN model was compared with the works of [31] [32] and [33], who used FID to evaluate their GAN networks in generating synthetic images.

Model	Med−CDCGAN	GAN [31]	TilGAN [32]	GAN [33]
FID Score	1.3	21.03	1.44	75.73
PSNR Score	10.64	–	–	22.77

TABLE II
Comparison Performance of Med−CDCGAN

From the results, our proposed model produced better results than [31] [32] and [33] with an FID score of 1.3 and a lower PSNR of 10.64 *db* lower than that of [33]. Although the FID

results were slightly lower than those of [32] used in cancer tumour cell synthesis, the results of our Med−CDCGAN model highlighted the potential to generate synthetic PCa images that can be used further in transfer learning training models for the classification of PCa-related ISUP grades.

VII. Conclusion

The goal of this paper was to develop, train, and thoroughly test a new conditional deep convolutional generative adversarial network, Med−CDCGAN, to enhance images of prostate cancer. Med−CDCGAN proved to be an effective and robust tool for generating realistic and diverse prostate cancer images. The extensive evaluation metrics demonstrate its superior performance, highlighting its potential utility in enhancing medical imaging datasets and enhancing diagnostic model training. Despite the higher computational cost, the benefits in terms of image quality and diversity justify its use in high-stakes medical applications, which are characterized by the absence of large training data sets. Through qualitative inspection of the images, it was realized that our generated images were at the same levels of accuracy with the original images from the PANDA dataset, indicating a high confidence in our Med−CDCGAN performance.

Future work will involve developing a transfer learning-based classification algorithm trained on a synthetic data set generated by Med−CDCGAN for the prediction of prostate cancer. Further, to pave the way for broader adoption and integration into clinical practice, a panel of professional pathologists will be chosen to qualitatively assess the synthetic biopsy images generated by the model as well as the predictions.

[1] Zhang, E., Chen, Z., Liu, W., Lin, L., Wu, L., Guan, J., Wang, J., Kong, C., Bi, J. and Zhang, M., 2024. NCAPG2 promotes prostate cancer malignancy and stemness via STAT3/c-MYC signalling. Journal of Translational Medicine, 22(1), pp.1-19.

[2] Siegel, R.L., Miller, K.D., Wagle, N.S. and Jemal, A., 2023. Cancer statistics, 2023. Ca Cancer J Clin, 73(1), pp.17-48.

[3] K. Yan, Y. Peng, V. Sandfort, M. Bagheri, Z. Lu, and R. M. Summers, "Holistic and comprehensive annotation of clinically significant findings on diverse CT images: Learning from radiology reports and label ontology," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 8523–8532.

[4] Z.Cui,C.Li, and W.Wang, "Toothnet: Automatic tooth instance segmentation and identification from cone beam CT images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 6368–6377.

[5] Z.Zhang,A.Romero,M.J.Muckley,P.Vincent,L.Yang,andM.Drozdzal, "Reducing uncertainty in undersampled MRI reconstruction with active acquisition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 2049–2058.

[6] X. Ying, H. Guo, K. Ma, J. Wu, Z. Weng, and Y. Zheng, "X2CT-GAN: Reconstructing CT from biplanar X-rays with generative adversarial networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 10 619–10 628.

[7] Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024 Jan-Feb;74(1):12-49. doi: 10.3322/caac.21820. Epub 2024 Jan 17. Erratum in: CA Cancer J Clin. 2024 Mar-Apr;74(2):203. PMID: 38230766.

[8] K. Yan, Y. Peng, V. Sandfort, M. Bagheri, Z. Lu, and R. M. Summers, "Holistic and comprehensive annotation of clinically significant find- ings on diverse CT images: Learning from radiology reports and label ontology," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 8523–8532.

[9] GeertLitjens, Hans Pinckaers, Kimmo Kartasalo, Maggie, Martin Eklund, PekkaRuusuvuori, PeterStröm, Sohier Dane, Wouter Bulten. (2020). Prostate cANcer graDe Assessment (PANDA) Challenge. Kaggle. https://kaggle.com/competitions/prostate-cancer-grade-assessment.

[10] Tataru, O.S.; Vartolomei, M.D.; Rassweiler, J.J.; Virgil, O.; Lucarelli, G.; Porpiglia, F.; Amparore, D.; Manfredi, M.; Carrieri, G.; Falagario, U.; et al. Artificial Intelligence and Machine Learning in Prostate Cancer Patient Management-Current Trends and Future Perspectives. Diagnostics 2021, 11, 354.

[11] Zhao, H.B.; Liu, C.; Ye, J.; Chang, L.F.; Xu, Q.; Shi, B.W.; Liu, L.L.; Yin, Y.L.; Shi, B.B. A comparison between deep learning convolutional neural networks and radiologists in the differentiation of benign and malignant thyroid nodules on CT images. Endokrynol. Pol. 2021, 72, 217–225.

[12] Yu, B.; Wang, Y.; Wang, L.; Shen, D.; Zhou, L. Medical Image Synthesis via Deep Learning. Adv. Exp. Med. Biol. 2020, 1213, 23–44.

[13] Kazuhiro, K.; Werner, R.A.; Toriumi, F.; Javadi, M.S.; Pomper, M.G.; Solnes, L.B.; Verde, F.; Higuchi, T.; Rowe, S.P. Generative Adversarial Networks for the Creation of Realistic Artificial Brain Magnetic Resonance Images. Tomography 2018, 4, 159–163.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Adv. Neural Inf. Process. Syst. 27 (2014).

[15] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784(2014).

[16] Bulten, W., Kartasalo, K., Chen, PH.C. et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. Nat Med (2022). https://doi.org/10.1038/s41591-021-01620-2

[17] Bulten, W., Kartasalo, K., Chen, PH.C. et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. Nat Med (2022). https://doi.org/10.1038/s41591-021-01620-2

[18] Yeom T, Gu C, Lee M. DuDGAN: improving class-conditional GANs via dual-diffusion. IEEE Access. 2024.

[19] Wang Y, Wu C, Herranz L. Transferring GANs: generating images from limited data. Proceedings of the European Conference on Computer Vision (ECCV). 2018:218-34.

[20] Jaiswal A, AbdAlmageed W, Wu Y. Bidirectional conditional generative adversarial networks. Asian Conference on Computer Vision. 2019:123-38.

[21] Wołczyk M, Proszewska M, Maziarka Ł. Plugen: Multi-label conditional generation from pre-trained models. Proceedings of the AAAI Conference on Artificial Intelligence. 2022;36(10):11238-46.

[22] Wang Y, Wu C, Herranz L. Transferring GANs: generating images from limited data. Proceedings of the European Conference on Computer Vision (ECCV). 2018:218-34.

[23] Jaiswal A, AbdAlmageed W, Wu Y. Bidirectional conditional generative adversarial networks. Asian Conference on Computer Vision. 2019:123-38.

[24] Runz, M., Rusche, D., Schmidt, S. et al. Normalization of HE-stained histological images using cycle consistent generative adversarial networks. Diagn Pathol 16, 71 (2021).

[25] Brendon Lutnick, Pinaki Sarder. Generative modeling of histology tissue reduces human annotation effort for segmentation model development.

[26] Breen J, Zucker K, Allen K, Ravikumar N, Orsi NM. Generative Adversarial Networks for Stain Normalisation in Histopathology. InApplications of Generative AI 2024 Mar 6 (pp. 227-247). Cham: Springer International Publishing.

[27] Cornish TC. Artificial intelligence for automating the measurement of histologic image biomarkers. The Journal of Clinical Investigation. 2021 Apr 15;131(8).

[28] 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. In Proceedings of the IEEE/CVF winter conference on applications of computer vision 2023 (pp. 2000-2009).

[29] Krause J, Grabsch HI, Kloor M, Jendrusch M, Echle A, Buelow RD, Boor P, Luedde T, Brinker TJ, Trautwein C, Pearson AT. Deep learning detects genetic alterations in cancer histology generated by adversarial networks. The Journal of pathology. 2021 May;254(1):70-9.

[30] Kynkäänniemi, T., Karras, T., Aittala, M., Aila, T., Lehtinen, J. (2022). The role of imagenet classes in fréchet inception distance. arXiv preprint arXiv:2203.06026.
[31] Haarburger, C., Horst, N., Truhn, D., Broeckmann, M., Schrading, S., Kuhl, C., Merhof, D. (2019, December). Multiparametric Magnetic Resonance Image Synthesis using Generative Adversarial Networks. In VCBM (pp. 11-15).
[32] Saha, M., Guo, X., Sharma, A. (2021). Tilgan: Gan for facilitating tumor-infiltrating lymphocyte pathology image synthesis with improved image classification. IEEE Access, 9, 79829-79840.
[33] Zheng, Z., Wang, M., Fan, C., Wang, C., He, X., & He, X. (2024). Light&fast generative adversarial network for high-fidelity CT image synthesis of liver tumor. Computer Methods and Programs in Biomedicine, 254, 108252.

First A. Author Paul Kariuki is a PhD candidate at Dedan Kimathi University of Technology. His research interests are in the area of advancing medical imaging through deep learning, with a specific focus on data augmentation for medical imaging. He is passionate about addressing the limitations of scarce and imbalanced datasets often encountered in medical diagnostics, especially in Africa. He earned his Master of Science (MSc) in Data Communications from KCA University,Kenya and a Bachelor of Science (BSc) in Information Technology from Ndejje University, Uganda.

Second B. Author, Kinyua Gikunda is a Computer Science Lecturer at the Dedan Kimathi University of Technology. He received his BSc (Computer Science) in 2011 from Kenyatta University, MSc (Computer Science) in 2015 from the University of Nairobi, and Ph.D. (Computer Science) from the University of Paris8. His research area includes Machine Learning with a focus on deep optimization methods. His specific expertise is in Active Learning, Transfer Learning, and data-driven analytics solutions.

Third C. Author Jr. John Mwangi Wandeto is a senior lecturer at the Department of Information Technology, Dedan Kimathi University of Technology in Nyeri, Kenya. He teaches units that relate to solving problems through computer system development. His research area is in computer vision, dealing with the extraction and analysis of information from medical images. Currently, he is engaged in a urinalysis project that studies microscopic urine images for infections and early detection of diseases. .