Heliyon 10 (2024) e37902

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

HcGAN: Harmonic conditional generative adversarial network for efficiently generating high-quality IHC images from H&E

Shuying Wu, Shiwei Xu

School of Artificial Intelligence, Wenzhou Polytechnic, Wenzhou City, 325035, Zhejiang Province, China

ARTICLE INFO

Keywords:
Image generation
Deep learning
Generative adversarial network
IHC

ABSTRACT

Generating high quality histopathology images like immunohistochemistry (IHC) stained images is essential for precise diagnosis and the advancement of computer-aided diagnostic (CAD) systems. Producing IHC images in laboratory is quite expensive and time consuming. Recently, some attempts have been made based on artificial intelligence techniques (particularly, deep learning) to generate IHC images. Existing IHC stained image generation methods, still have a limited performance due to the complex structures and variations in cells shapes, potential nonspecific staining and variable antibody sensitivity. This paper proposes a novel technique known as harmonic conditional generative adversarial network (HcGAN) for generating high quality IHC-stained images. To generate the IHC images, the HcGAN model is fed with the widely available hematoxylin and eosin (H&E) images that contain cellular and morphological underlying structures of diverse cancer tissues. Such approach helps generate high quality IHC images mimicking the real ones that highlight the positive cells. The proposed HcGAN model is based generative adversarial learning with generator and discriminator networks. In HcGAN, harmonic convolution based on discrete cosine transform filter banks is employed in the generator and discriminator networks instead of the standard convolution in order to improve visual quality of the generated images and address the issue of overfitting. Our qualitative and quantitative results demonstrate that the proposed HcGAN achieved the highest performance over state-of-the-art methods using two publicly available datasets.

1. Introduction

The generation of high-quality histopathology images holds paramount significance in advancing both the accuracy of diagnosis and the capabilities of computer-aided diagnosis (CAD) systems. In the realm of medical imaging, particularly in histopathology, the quality of visual data is pivotal for pathologists to discern subtle nuances and abnormalities that might be indicative of diseases such as cancer [1]. Notably, there are two common staining techniques used to generate histopathology images comprises hematoxylin & eosin ($H\&E$) and immunohistochemistry (IHC). $H\&E$ is the gold standard staining technique to permits visualization of the structure, distribution of cells and morphological changes from tissue, whereas IHC highlight the specific protein markers that can assist with accurate tumor classification and diagnosis [2].

High-fidelity images, whether real or synthetically generated, contribute to a more comprehensive and detailed representation of cellular structures, enabling pathologists to make more precise diagnoses. Additionally, in the context of CAD systems, which

E-mail address: wushuying0905@sina.com (S. Wu).

https://doi.org/10.1016/j.heliyon.2024.e37902
Received 15 February 2024; Received in revised form 9 September 2024; Accepted 12 September 2024

Available online 1 October 2024

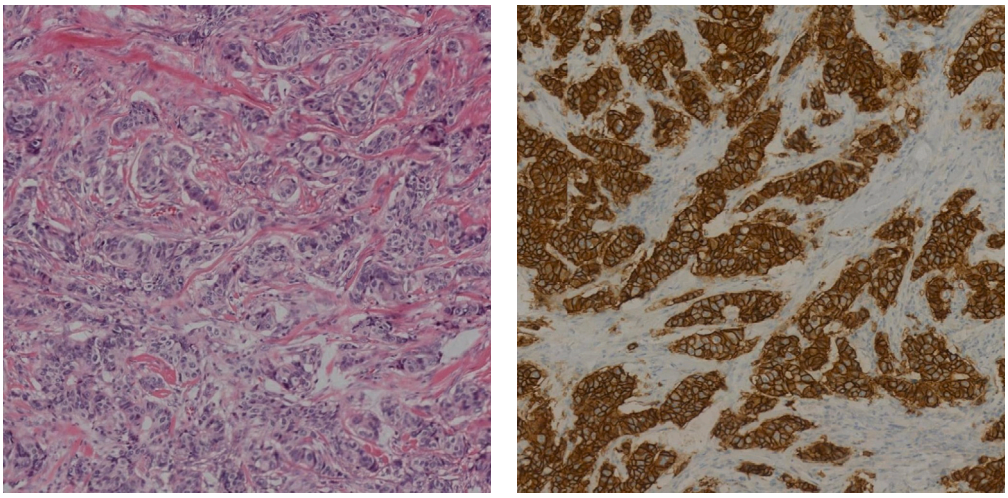

Fig. 1. The difference between the appearance of a patch stained with the H&E and the corresponding patch of the same tissue stained with the IHC method is shown. Here, (a) shows the H&E patch images extracted from the WSI, and (b) corresponds to the IHC slide patch.

leverage machine learning algorithms for automated analysis, the generation of realistic and diverse histopathology images becomes instrumental for training robust models [3]. These high-quality images facilitate the development and enhancement of CAD systems, allowing them to learn and generalize from a richer dataset [4]. In essence, the generation of superior histopathology images not only aids pathologists in their diagnostic endeavors but also propels the capabilities of automated systems, ultimately leading to more accurate and efficient medical diagnoses [5].

It should be noted that the generation of IHC images in laboratory is quite expensive and time consuming. Artificial intelligence technologies, especially deep learning, have recently played a pivotal role in generating high-quality IHC images [6,7]. Deep learning methods for generating high-quality IHC images primarily fall into two categories: GANs and variational autoencoders (VAEs) [8]. GANs employ adversarial training to create realistic images, while VAEs focus on learning latent representations for generating diverse samples. Conditional generative models extend GANs and VAEs to generate images based on specific attributes or classes, aiding targeted synthesis. Some models adopt a progressive generation approach, refining images in sequential steps for intricate details. Attention mechanisms enhance focus on diagnostically significant regions in generated images. These methods collectively address data scarcity challenges and improve dataset diversity for training diagnostic models.

In the context of histopathology image generation, GANs have demonstrated remarkable success in synthesizing realistic and diverse images that closely resemble actual tissue samples [5,9]. Such deep GAN models learn intricate patterns and textures from existing datasets, capturing the complexity of cellular structures, and then generate new images that exhibit similar characteristics. Notably, the capability of GANs is particularly valuable in scenarios where there is a scarcity of annotated data for training, as GANs can augment datasets significantly [10]. In the last years, researchers have extended GAN architectures with conditional settings, allowing for the generation of images based on specific characteristics or classes, enhancing their utility in medical applications [11]. The ability of deep learning models, including GANs, to generate high-quality histopathology images has not only advanced the field of medical image synthesis but has also contributed to improving the training data for diagnostic models, aiding pathologists, and enhancing the overall accuracy of computer-aided diagnosis systems [12].

However many deep learning based IHC image generation methods have been proposed in the last years, their performance is limited due to the intricate structures and variations in cell shapes, potential nonspecific staining, and variable antibody sensitivity. Unlike IHC stained images, H&E stained images are widely available which obtain after the biopsy of tissues. Fig. 1 (a, b) presents the difference between the appearance of a patch stained with the H&E method and corresponding patch of the same tissue stained with IHC method.

In an attempt to handle these issues, in this paper we propose a novel approach called the harmonic conditional generative adversarial network (HcGAN) to overcome these limitations and produce high-quality IHC-stained images. The HcGAN model is trained using widely available H&E images that contain diverse cellular and morphological structures found in cancer tissues. By leveraging generative adversarial learning with generator and discriminator networks, the proposed HcGAN model utilizes harmonic convolution based on discrete cosine transform filter banks instead of standard convolution. This modification aims to boost the visual quality of generated images and mitigate the issue of overfitting, resulting in high-quality IHC images that faithfully represent real staining patterns, particularly emphasizing positive cells.

The subsequent sections of this paper are structured as follows: Section 2 delves into the related work concerning histopathology image generation. Following that, Section 3 elucidates the proposed method for histopathology image generation based on deep learning. Section 4 outlines the outcomes of the proposed method, including comparisons with state-of-the-art approaches. Finally, Section 5 concludes the paper and puts forth potential avenues for future research.

2. Related work

Here, we provide a brief review of prior research findings and developments in the field of histopathology image generation. Histopathology image generation methods primarily fall into two major categories: supervised approaches, involving expert-annotated data for training; and unsupervised approaches, utilizing techniques such as GANs to generate images without explicit annotation guidance.

Different unsupervised approaches have been proposed for generating high resolution histopathology images. For instance, Hou et al. [13] introduces an unsupervised methodology for histopathology image segmentation, wherein diverse sets of training patches was synthesized to represent each tissue type. This pioneering approach incorporates a re-weighting mechanism for the training loss over synthetic data, minimizing the unbiased generalization loss across the true data distribution. This enabled the use of a random polygon generator to synthesize approximate cellular structures, specifically nuclear masks, in instances where real examples are lacking across various tissue types, making GAN-based methods impractical. Moreover, they proposed a hybrid synthesis pipeline that integrates textures from real histopathology patches and GAN models to address tissue texture heterogeneity. This approach exhibited superior generalization, particularly in cancer types with limited training data.

In [14], the authors adapted and evaluated various high-resolution, state-of-the-art generative learning based models primarly designed for face synthesis, highlighting their superiority over existing approaches in the intricate field of digital pathology. The results demonstrated a manifold enhancement in image synthesis, underscored by improved quality and resolution in the generated images, validated against supervised models. Besides, Ma et al. [15] proposed employing these high-resolution RGB images as guidance for the super-resolution reconstruction of high-resolution hyperspectral images. They introduced a straightforward yet effective unsupervised super-resolution reconstruction network that integrates spatial information from high-resolution RGB images with spectral information from low-resolution hyperspectral images. This method not only has the potential to reduce acquisition time and storage space for hyperspectral images but also addresses the low quality of certain bands in hyperspectral images, thereby facilitating the application of hyperspectral imaging for whole slide imaging (WSI) and automated histopathological cancer detection.

In turn, numerous supervised histopathology image generations have been presented in the last years. For instance, the study presented in [16] introduced a dataset comprising matched RGB digitized histology and histological hyperspectral imaging data for breast cancer. Additionally, a conditional generative adversarial network (cGAN) has been designed for the artificial synthesis of HSI from conventional RGB images depicting normal and cancer cells. The outcomes of the GAN-synthesized HSI exhibit promise, with a structural similarity (SSIM) reaching approximately 80% and a mean absolute error (MAE) ranging between 6% and 11%.

The study presented in [17], the histopathology DatasetGAN (HDGAN) framework was introduced, an extension of the DatasetGAN semi-supervised approach for image generation and segmentation tailored for large-resolution histopathology images. Several modifications have been implemented from the original framework, including the enhancement of the generative backbone, selective extraction of latent features from the generator, and a transition to memory-mapped arrays. These adjustments significantly reduce the framework's memory consumption, enhancing its suitability for medical imaging applications. We assess the performance of HDGAN using a thrombotic microangiopathy high-resolution tile dataset, showcasing robust capabilities in the high-resolution image-annotation generation task. In [9], Li et al. presented a multi-scale GAN designed for the generation and segmentation of high-resolution, large-scale histopathology images. The model is structured as a pyramid of GAN architectures, with each module specifically tasked with generating and segmenting images at varying scales. Utilizing semantic masks, the generative component of this model demonstrated high performance in synthesizing histopathology images that exhibit a visually realistic quality. Recently, a novel coarse-to-fine sampling strategy was presented in [18] to overcome the task of generating high-resolution WSIs. Within this framework, we augment the resolution of an initial low-resolution image to produce a high-resolution WSI. Notably, a diffusion model is utilized to systematically incorporate fine details into the images, incrementally enhancing their resolution.

Additionally, [19] the authors held a challenge to generate breast cancer IHC images with the help of deep learning approaches. Participants were given $H\&E$ and IHC-stained image pairs to train a model that generates IHC-stained images from corresponding $H\&E$-stained images. The five highest-ranking methods were reviewed based on their peak signal-to-noise-ratio (PSNR) and SSIM metrics, and their corresponding pipelines and implementations were provided.

Despite numerous proposals for deep learning-based IHC image generation methods in recent years, their efficacy is constrained by intricate structures, variations in cell shapes, potential nonspecific staining, and variable antibody sensitivity. To handle these issues, in this paper we introduce the HcGAN model, utilizing generative adversarial learning and harmonic convolution. HcGAN leverages DCT filters for harmonic extraction from lower-level data, leading to improved image quality and mitigated overfitting.

3. Proposed method

Fig. 2 presents the schematic diagram of the proposed IHC image generation method. As shown in Fig. 2, the proposed HcGAN model includes generator (G) and discriminator (D) networks. The generator comprises encoder and decoding layers that extract the clinically relevant meaningful features and synthesize them into images like real. In this model, the traditional convolution is replaced with harmonic convolution to mitigate overfitting and produce improved IHC image quality. More details about the model are explained below.

3.1. Extracting features using harmonic convolution

Following the methodology proposed by [20], we utilized harmonic convolution instead of normal convolution to mitigate overfitting. This was achieved by extracting harmonics from lower-level data. The harmonic convolution refers to the summation of

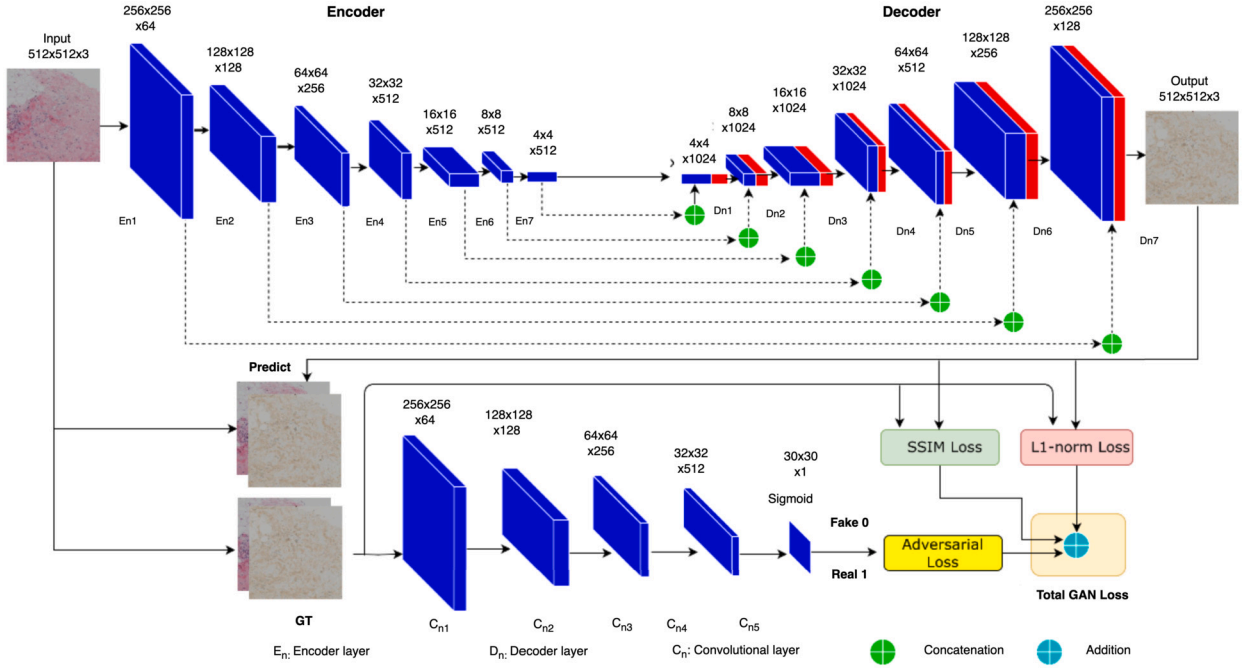

Fig. 2. Illustration of the proposed HcGAN model.

responses from a discrete cosine transform (DCT) filter bank, with each response being weighted. The DCT is a mathematical operation that transfers a signal from the time domain to the frequency domain in a way that can be easily separated into independent components. In Eq. (1), the DCT can be mathematically expressed for a two-dimensional image I with dimensions $H \times W$ and a pixel discretization step of one, as described in the reference [20].

$$\mathbb{T}_{u,v} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \sqrt{\frac{\beta_u}{H}} \sqrt{\frac{\beta_v}{W}} I_{(i,j)} \times \cos\left[\frac{\pi}{H}\left(i+\frac{1}{2}\right)u\right] \cos\left[\frac{\pi}{W}\left(j+\frac{1}{2}\right)v\right]. \tag{1}$$

The coefficient $\mathbb{T}_{u,v}$ represents the sinusoidal frequency with values of u and v in two perpendicular directions. The scaling coefficients β_0 and β_u, with values of 1 and 2 respectively, are employed to normalize the basis function's value.

To perform the convolution procedure, the input image features undergo decomposition using the Discrete Cosine Transform (DCT). The filter bank has a kernel size of $f \times f$ and a depth of f^2. The frequency selective DCT filter with kernel size $f \times f$ for frequencies u and v is represented by the notation $\gamma_{(u,v)}$. As in Eq. (2), the feature map $\mathbb{F}_m$ at depth d is then expressed as a weighted linear combination of discrete cosine transform (DCT) coefficients over all input channels C.

$$\mathbb{F}_m^d = \sum_{n=0}^{C-1} \sum_{u=0}^{f-1} \sum_{v=0}^{f-1} w_{n,u,v}^d \gamma_{(u,v)} ** \mathbb{F}_{m(n)}^{d-1}. \tag{2}$$

The variable $w_{n,u,v}^d$ represents the weight learned for the n^{th} feature at the frequency u and v. The symbol $**$ denotes the 2-D convolution operator. The transformation procedure facilitates the decomposition of the input feature into harmonics, enabling the utilization of learnt weights for the fusion of the changed signals.

3.2. Harmonic conditional generative adversarial network (HcGAN)

In the context of the Harmonic Conditional Generative Adversarial Network (HcGAN), the generator network is referred to as the G network. This network employs encoding and decoding layers to grasp the intrinsic attributes of provided input images. Leveraging these acquired features, the network produces images of superior quality.

Conversely, the discriminator network, labeled as D, assesses the probability that the generated images exhibit realistic high quality. Through the integration of the adversarial score in the computation of the generator loss, HcGAN amplifies its capability to produce precise, high-quality images.

Through the amalgamation of the G and D networks, the learning process gains increased robustness, particularly in scenarios with a restricted number of training samples. Utilizing IHC images as conditioning inputs for both the generator (G) and the discriminator (D) contributes to a superior outcome, aligning more precisely with the visual characteristics of the stained IHC.

In accordance with Fig. 2, the generator (G) and discriminator (D) structures are depicted. The generator comprises seven encoding and decoding layers. The encoding layers feature a series of DCT kernels, followed by batch normalization and a non-linear leaky

ReLU activation function with a slope of 0.2. Conversely, the decoding layers consist of deconvolutional filters, batch normalization, dropout, and ReLU activation.

The DCT kernels used in this model consist of both convolutional and deconvolutional layers with a 4×4 kernel and a 2×2 stride. These layers are responsible for decreasing and increasing the size of the activation maps by a factor of 2. It is worth noting that the first and last convolutional filters (En_1 and En_7) do not employ batch normalization. Additionally, the leaky ReLU is replaced by the ReLU activation function beyond En_7. Dropout is only present in the initial three decoding layers (Dn_1, Dn_2, and Dn_3), while the last decoding layer (Dn_7) lacks a skip connection. Finally, the $TanH$ activation function is used to generate high-quality IHC-stained images from the H&E input provided.

As shown in Fig. 2-down, the configuration of discriminator comprises five encoding layers. The first three layers use DCT-based convolutional filters with a 4×4 kernel and a 2×2 stride, whereas the 4^{th} and 5^{th} layers have a stride of 1×1. Batch normalization is applied after layers Cn_2, Cn_3, and Cn_4. Also, a leaky ReLU activation function with a slope of 0.2 is applied after each layer, except for the final layer. Finally, the sigmoid activation function is employed after the last DCT filter (Cn_5). The output generated by this is a 30×30 matrix with values ranging from 0.0 (implying that it is entirely fake) to 1.0 (indicating that it is entirely plausible and authentic). These output values represent the probability of generating a precise, high-quality IHC image, with each value corresponding to a receptive field of 70×70.

3.3. Loss function

The loss function of the generator in Eq. (3) (ℓ_{Gen}) incorporated three main terms: adversarial loss (binary cross-entropy loss), $L1$- norm to promote similarity between generated image and ground-truth image, and SSIM loss [21] to enhance the shape of the boundaries of high quality IHC-stained images:

$$\ell_{Gen}(G, D) = \mathbb{E}_{x,y,z}(-\log(D(x, G(x, z)))) + \lambda \mathbb{E}_{x,y,z}(\ell_{L1}(y, G(x, z))) + \alpha \mathbb{E}_{x,y,z}(\ell_{SSIM}(y, G(x, z))) \tag{3}$$

Both components of the loss function play a crucial role. The $L1$ term overcome to the generation of blurred IHC images, as the distance metric computes the average of all pixel differences. Integration of the adversarial term equips the generator with the capacity to manipulate input H&E images with intricate features (high frequencies), enabling the creation of accurate high-quality IHC images. Note that we used the λ and α values of 100 and 75, respectively.

The computation of the loss for the discriminator network can be represented as in Eq. (4):

$$\zeta_{Discriminator}(G, D) = \mathbb{E}_{x,y}\big(-log(D(x, y))\big) + \mathbb{E}_{x,y,z}\big(-log(1 - D(x, G(x, z)))\big). \tag{4}$$

As a result, the optimizer will adjust the discriminator network to increase the likelihood of precisely identifying the real high-quality IHC images (by minimizing $-log(D(x, y))$) while decreasing the possibility of misidentifying generated IHC images (by minimizing $-log(1 - D(x, G(x, z)))$).

4. Experimental results and discussion

4.1. Dataset

In this study, we used the two publicly available datasets that dataset one (DS_1) comprises 4,872 pairs of aligned $H\&E$-IHC pathology image patches for epidermal growth factor receptor-2 (HER2) biomarker, all sourced from the WSIs of over 300 patients [19]. The dataset used in this study randomly is split into the 3396 pairs of training, 200 pairs of validation, and 300 testing set images. Additionally, the MIST dataset (DS_2) [22] contains four breast cancer biomarkers: estrogen receptor (ER), progesterone receptor (PR), Ki67, and HER2. The dataset includes 4,153, 4,139, 4,361, and 4,642 training paired patches for each biomarker, respectively, and 1,000 patches for testing from 64 WSIs. All the paired patches are non-overlapping with 1024×1024 pixels. We independently trained and evaluated the proposed model on each individual biomarker. Fig. 3 shows the six example images including both datasets with H&E stained images alongside their corresponding patches stained with IHC.

4.2. Implementation details

For the proposed model, we first used the input image and resized it to 512×512 pixels. Next, we normalized the resized images from $0 - 255$ pixel values to a range of 0-1. We employed an ADAM optimizer with an initial learning rate of 0.0001 to optimize the model. The model was trained from scratch for 500 epochs using a batch size of 2 images. To increase the sample size and provide more variability of features in the training network, we performed data augmentation consisting of rotation at 90 degrees, horizontal flipping, and scaling with a probability of 0.5. This selected strategy helped mitigate overfitting. To quantitatively evaluate the model performance, we used three important metrices including PSNR in dB, SSIM and Fréchet inception distance (FID). We used the PyTorch neural network library to develop the model and trained it on a RAM with 32 GB memory and CUDA version 11.2. The model was trained and evaluated with a NVIDIA RTX2080Ti GPU with 11 GB video RAM.

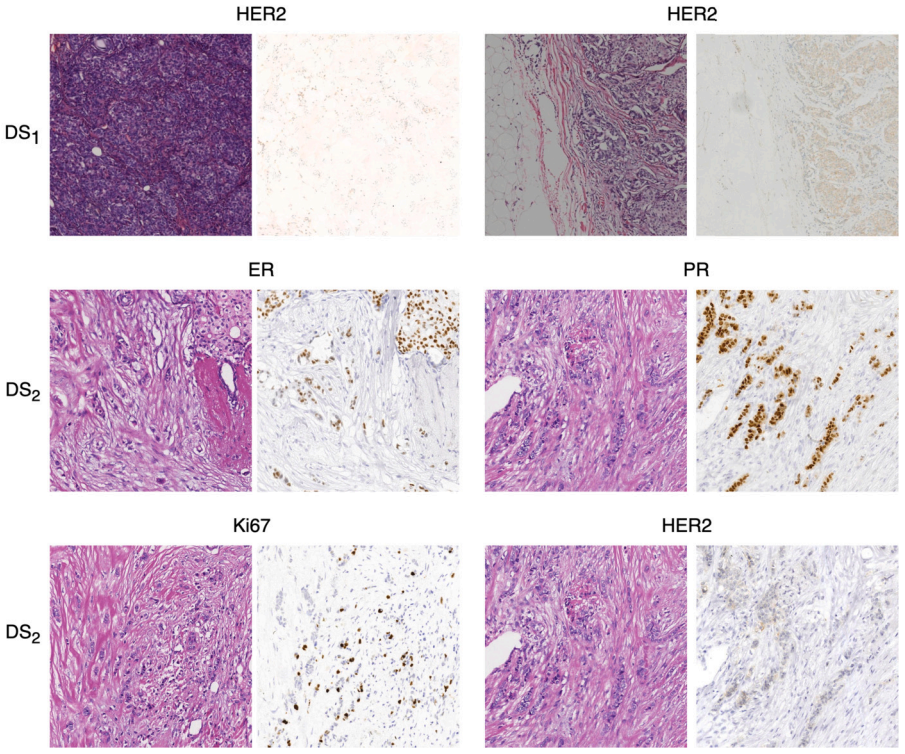

Fig. 3. Illustration of six example images from the both the datasets (DS_1 and DS_2) showing H&E stained images alongside their corresponding patches stained with IHC.

Table 1
Comparing the PSNR, SSIM and FID scores of the proposed method with Pix2Pix, DCGAN, unsupervised CycleGAN, and the teams results participated in the challenge.

Method	PSNR (dB) ↑	SSIM ↑	FID ↓
Pix2Pix [23]	14.11	0.487	67.14
DCGAN [24]	17.56	0.553	58.45
CycleGAN [25]	14.89	0.436	73.20
arpitdec5 [26]	19.73	0.574	-
Just4Fun [26]	22.92	0.559	-
lifangda02 [26]	17.92	0.555	-
stan9 [26]	17.95	0.543	-
guanxianchao [26]	19.56	0.497	-
vivek23 [26]	15.27	0.493	-
HcGAN (Proposed)	**21.77**	**0.613**	**47.91**

4.3. Results

Table 1 shows that our proposed model outperforms state-of-the-art methods such as Pix2Pix, DCGAN, CycleGAN and the teams results participated in the challenge. Based on the experimental results, HcGAN demonstrated its better capabilities by outperforming other approaches with very high SSIM (0.613), lowest FID (47.91), and highest PSNR (21.77). The proposed convolutional architecture of the model allows the network to get hierarchical representations of features, which helps in capturing complex details and retaining the overall structure of the image. This helps with the generation of IHC images that accurately capture the textures, shapes, and structures identified in the provided H&E images. The DCGAN had the second highest result in comparison to the recommended model. While DCGAN allows for the creation of various outputs by sampling from a latent space, recognizing and handling the latent space representation may be challenging. While Pix2Pix can learn the input-output image association using a conditional GAN, it may struggle to effectively represent complex cell structures and fine details compared to HcGAN. It achieved relatively poor performance on PSNR, SSIM, and FID. Nevertheless, of all the approaches evaluated, CycleGAN had the lowest performance. A potential reason for this might be that virtual staining requiring detailed information and exact cell structural correctness, such as generating IHC images, is not well-suited to CycleGAN's architecture, which was developed for image translation without paired examples. Perhaps

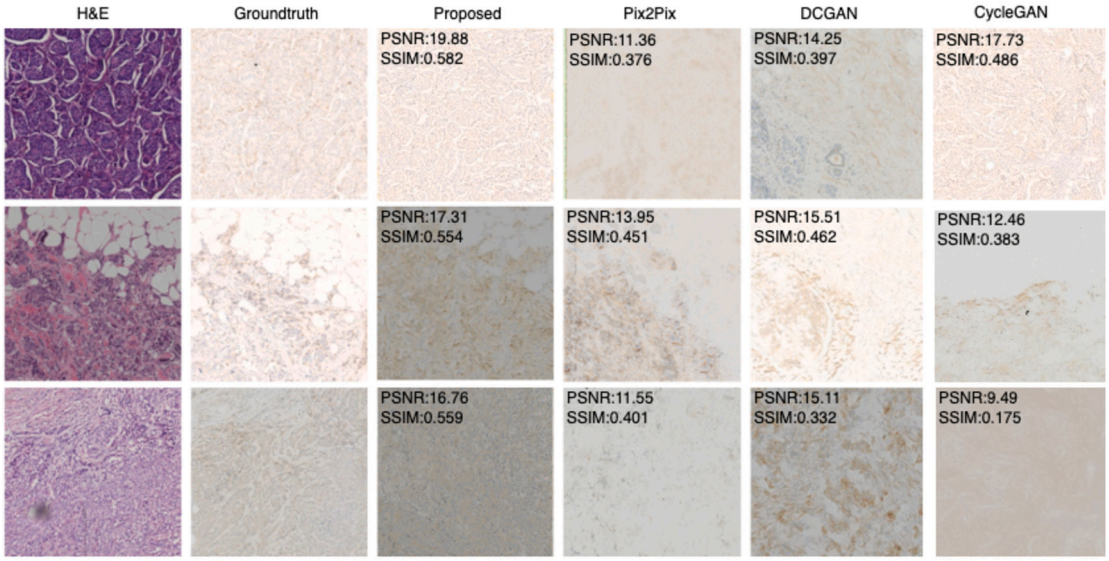

Fig. 4. Illustration of the high-quality IHC-stained images generated by the proposed model compared with Pix2pix, DCGAN and CycleGAN.

Table 2
Analyzing the effect of input resolution using the Hc-GAN on DS_1 samples.

Input Size	PSNR ↑ (dB)	SSIM ↑	FID ↓
128×128	10.81	0.404	77.8
256×256	19.84	0.571	55.31
384×384	16.47	0.508	64.60
512×512	21.77	0.613	47.91

understanding its poor performance in this instance is that CycleGAN utilizes cycle consistency loss, which can lead to less precise reconstructions in some examples.

Fig. 4 compares the proposed model to several applied methods. The proposed HcGAN approach excels at creating IHC-stained images that preserve the cellular structure and highlight positive cells. The model can successfully capture the detailed associations between the input H&E images and their corresponding IHC-stained images using harmonic convolution, which is the approach that achieves this impressive performance. The model's ability to understand spatial patterns and contextual information in histological images is enhanced by harmonic convolution that leads to a more precise and detailed conversion from H&E to IHC staining.

4.4. Ablation study

4.4.1. Effect of image resolution

This ablation study examines how a model's performance depends on the input resolution. Table 2 shows the results, with PSNR (dB), SSIM, and FID scores given for various input resolutions. The maximum PSNR value of 21.77 dB, the highest SSIM score of 0.613 and lowest FID 47.91 were produced by the input resolution of 512×512, as shown clearly in Table 2. On the other hand, the model's performance is considerably decreased when using a lowered input resolution of 128×128, as seen by the PSNR score of a poor 10.81 dB and an SSIM score of 0.404. The increased amount of complex data provided to the model is likely contributing to the higher accuracy at a resolution of 512×512. The model can capture more complex features and textures with higher input resolution, leading to more accurate reconstructions and higher-quality outputs. This leads to better PSNR and SSIM scores, which in turn indicate better image quality and more structural similarity. To optimize the model's performance, it is most beneficial to use an input resolution of 512×512 pixels.

4.4.2. Effect of loss function

Table 3 shows a detailed evaluation of several loss functions and how they impact the efficiency of generating IHC-stained images using the proposed method. The results show that the proposed loss function outperforms the other evaluated loss functions, with scores of 21.77 dB for PSNR, 0.613 for SSIM, and 47.91 for FID. It effectively combines BCE, L1 loss, and SSIM. In addition to evaluating the model's ability to generate high-quality virtual stained images, these metrics additionally assess how well it maintains structural accuracy to the real images. The importance of selecting an appropriate loss function while developing models to generate IHC-stained images is demonstrated by the results. The outcomes show a significant performance boost when BCE, $L1$, and SSIM are

Table 3
Analyzing the effect of loss function using the HcGAN on DS_1 samples.

Loss Function	PSNR (dB) ↑	SSIM ↑	FID ↓
BCE	11.47	0.428	70.29
$L1$	14.25	0.495	66.31
SSIM	16.79	0.548	57.77
BCE+ $L1$+SSIM (Proposed)	**21.77**	**0.613**	47.91

Table 4
Performance assessment of the proposed model using the MIST dataset.

Biomarker	Proposed			Li et al. [22]		
	PSNR (dB) ↑	SSIM ↑	FID ↓	PSNR (dB) ↑	SSIM ↑	FID ↓
ER	21.23	0.225	40.80	-	0.221	43.7
PR	22.07	0.248	43.11	-	0.240	44.8
Ki67	24.66	0.255	48.27	-	0.241	51.0
HER2	20.59	0.219	44.28	-	0.215	45.2

used. Overall, this ablation study emphasizes how crucial it is to choose the loss function effectively for improving image generating method.

4.5. Evaluating the HcGAN on MIST dataset

Table 4 presents an independent comparison of the results from the proposed HcGAN model. This model is designed to generate IHC images from H&E stained images in the MIST dataset. We generated IHC images for four specific biomarkers: ER, PR, Ki67, and HER2. The outcomes from the proposed HcGAN model are compared with results generated by the existing method developed by Li et al. [22] for each biomarker. The HcGAN model showed good results for ER, achieving a PSNR of 21.23 dB, an SSIM of 0.225, and an FID of 40.80. Although limited information is provided by Li et al. [22], their available SSIM value of 0.221 and FID of 43.7 suggest that the proposed model achieves comparable or even better results in both metrics. For PR, the proposed model generated a PSNR of 22.07 dB, an SSIM of 0.248, and an FID of 43.11. Li et al. [22] SSIM was 0.240, indicating slightly lower accuracy, while their FID was 44.8, revealing a marginally larger quality difference. These results suggest that the proposed model offers a slight improvement in capturing structural details and overall accuracy in the generated images for PR. The proposed model performed very well for Ki67, attaining a PSNR of 24.66 dB, an SSIM of 0.255, and an FID of 48.27. These results outperform the SSIM of 0.241 and FID of 51.0 reported by Li et al. [22] demonstrating improvements in both visual quality and similarity between the generated images and the original IHC images.

However, the results for HER2 indicate potential for further advancement, with a PSNR of 20.59 dB, an SSIM of 0.219, and an FID of 44.28. These results are somewhat lower than the SSIM of 0.215 and FID of 45.2 reported by Li et al. [22]. This suggests that while the proposed model is effective, there is room for improvement, particularly for this biomarker. The proposed HcGAN demonstrates robust potential for generating high-quality IHC images from H&E stained images for the evaluated biomarkers. Its performance is comparable to or exceeds that of existing methods in terms of PSNR, SSIM, and FID, highlighting its efficiency and potential for enhancing IHC image generation.

Fig. 5 shows high-resolution IHC images generated by the proposed HcGAN model. It highlights the original H&E stained images alongside the corresponding IHC ground truth images and model predictions for the biomarkers ER, PR, Ki67, and HER2. Each example demonstrates the model's capability to accurately translate H&E stained images into detailed and visually precise IHC images. The generated IHC images reveal distinct and accurate staining patterns that demonstrate the model's ability to produce high-quality images that closely replicate the predicted IHC staining results.

5. Conclusion

This paper addresses an important challenge of generating high-quality IHC stained images, crucial for precise medical diagnoses and the advancement of CAD systems. To achieve this, we have proposed the HcGAN model based on generative adversarial learning and harmonic convolution. HcGAN can overcome limitations observed in existing IHC stained image generation methods, thanks to the DCT filters that extract harmonics from lower-level data, resulting in enhanced image quality and reduced overfitting. The comparative results demonstrated that the proposed HcGAN model outperformed state-of-the-art methods on two publicly available datasets, as evidenced by higher PSNR, SSIM and lower FID scores. Future work could explore the robustness of the proposed HcGAN model across multiple tumor types and staining conditions would enhance its applicability in real-world scenarios.

Code availability statement

Due to the project guidelines, the source code cannot be made available.

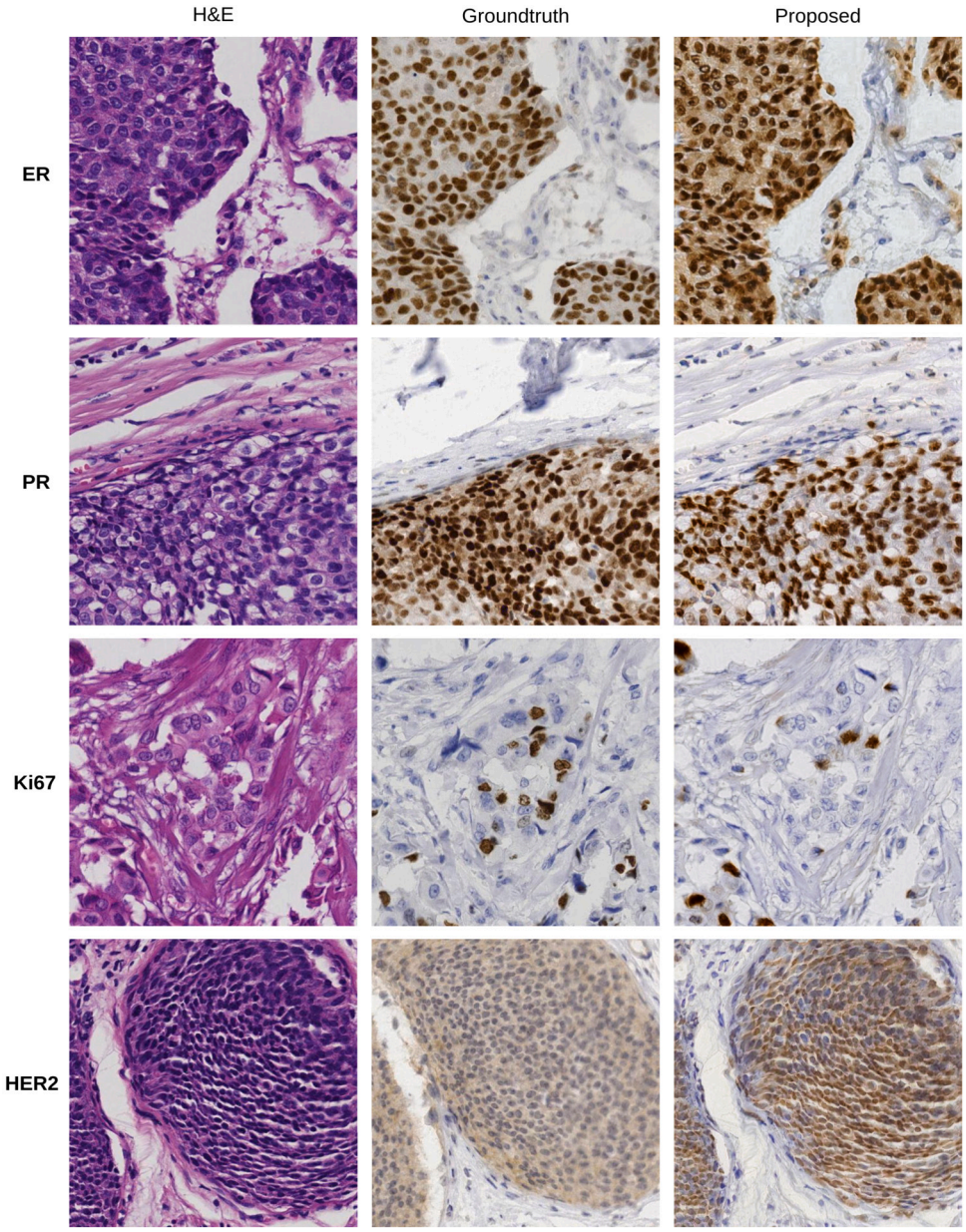

Fig. 5. Illustration of the high-quality IHC-stained images generated by the proposed model.

CRediT authorship contribution statement

Shuying Wu: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shiwei Xu:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

[1] E.D. McAlpine, P. Michelow, T. Celik, The utility of unsupervised machine learning in anatomic pathology, Am. J. Clin. Pathol. 157 (1) (2022) 5–14.

[2] A.-A. Grosset, K. Loayza-Vega, É. Adam-Granger, M. Birlea, B. Gilks, B. Nguyen, G. Soucy, D. Tran-Thanh, R. Albadine, D. Trudel, Hematoxylin and eosin counterstaining protocol for immunohistochemistry interpretation and diagnosis, Appl. Immunohistochem. Mol. Morphology 27 (7) (2019) 558–563.

[3] G. Apou, F. Feuerhake, G. Forestier, B. Naegel, C. Wemmert, Synthesizing whole slide images, in: 2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE, 2015, pp. 154–159.

[4] A. Davri, E. Birbas, T. Kanavos, G. Ntritsos, N. Giannakeas, A.T. Tzallas, A. Batistatou, Deep learning for lung cancer diagnosis, prognosis and prediction using histological and cytological images: a systematic review, Cancers 15 (15) (2023) 3981.

[5] L. Jose, S. Liu, C. Russo, A. Nadort, A. Di Ieva, Generative adversarial networks in digital pathology and histopathological image processing: a review, J. Pathol. Inform. 12 (1) (2021) 43.

[6] S.J. Wagner, C. Matek, S.S. Boushehri, M. Boxberg, L. Lamm, A. Sadafi, D.J. Winter, C. Marr, T. Peng, Built to last? Reproducibility and reusability of deep learning algorithms in computational pathology, Mod. Pathol. 37 (1) (2024) 100350.

[7] K. Al-Thelaya, N.U. Gilal, M. Alzubaidi, F. Majeed, M. Agus, J. Schneider, M. Househ, Applications of discriminative and deep learning feature extraction methods for whole slide image analysis: a survey, J. Pathol. Inform. 14 (2023) 100335.

[8] S. Banerji, S. Mitra, Deep learning in histopathology: a review, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 12 (1) (2022) e1439.

[9] W. Li, J. Li, J. Polson, Z. Wang, W. Speier, C. Arnold, High resolution histopathology image generation and segmentation through adversarial training, Med. Image Anal. 75 (2022) 102251.

[10] J. Ye, Y. Xue, P. Liu, R. Zaino, K.C. Cheng, X. Huang, A multi-attribute controllable generative model for histopathology image synthesis, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, in: Part VIII, vol. 24, Springer, 2021, pp. 613–623.

[11] R. Rong, S. Wang, X. Zhang, Z. Wen, X. Cheng, L. Jia, D.M. Yang, Y. Xie, X. Zhan, G. Xiao, Enhanced pathology image quality with restore–generative adversarial network, Am. J. Pathol. 193 (4) (2023) 404–416.

[12] J.M. Dolezal, R. Wolk, H.M. Hieromnimon, F.M. Howard, A. Srisuwananukorn, D. Karpeyev, S. Ramesh, S. Kochanny, J.W. Kwon, M. Agni, et al., Deep learning generates synthetic cancer histology for explainability and education, NPJ Precis. Oncol. 7 (1) (2023) 49.

[13] L. Hou, A. Agarwal, D. Samaras, T.M. Kurc, R.R. Gupta, J.H. Saltz, Robust histopathology image analysis: to label or to synthesize?, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 8533–8542.

[14] D. Štepec, D. Skočaj, Image synthesis as a pretext for unsupervised histopathological diagnosis, in: Simulation and Synthesis in Medical Imaging: 5th International Workshop, SASHIMI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, in: Proceedings, vol. 5, Springer, 2020, pp. 174–183.

[15] L. Ma, A. Rathgeb, H. Mubarak, M. Tran, B. Fei, Unsupervised super-resolution reconstruction of hyperspectral histology images for whole-slide imaging, J. Biomed. Opt. 27 (5) (2022) 056502.

[16] M. Halicek, S. Ortega, H. Fabelo, C. Lopez, M. Lejeune, G.M. Callico, B. Fei, Conditional Generative Adversarial Network for Synthesizing Hyperspectral Images of Breast Cancer Cells from Digitized Histology, Medical Imaging 2020: Digital Pathology, vol. 11320, SPIE, 2020, pp. 198–205.

[17] S. Rizvi, P. Cicalese, S. Seshan, S. Sciascia, J. Becke, H. Nguyen, Histopathology datasetgan: synthesizing large-resolution histopathology datasets, in: 2022 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), IEEE, 2022, pp. 1–4.

[18] R. Harb, T. Pock, H. Müller, Diffusion-based generation of histopathological whole slide images at a gigapixel scale, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024, pp. 5131–5140.

[19] S. Liu, C. Zhu, F. Xu, X. Jia, Z. Shi, M. Jin, Bci: breast cancer immunohistochemical image generation through pyramid pix2pix, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2022, pp. 1815–1824.

[20] M. Ulicny, V.A. Krylov, R. Dahyot, Harmonic convolutional networks based on discrete cosine transform, arXiv preprint, arXiv:2001.06570, 2020.

[21] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, et al., Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

[22] F. Li, Z. Hu, W. Chen, A. Kak, Adaptive supervised patchnce loss for learning h&e-to-ihc stain translation with inconsistent groundtruth image pairs, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2023, pp. 632–641.

[23] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1125–1134.

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

[25] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.

[26] C. Zhu, S. Liu, Z. Yu, F. Xu, A. Aggarwal, G. Corredor, A. Madabhushi, Q. Qu, H. Fan, F. Li, et al., Breast cancer immunohistochemical image generation: a benchmark dataset and challenge review, arXiv preprint, arXiv:2305.03546, 2023.