



Original Article

MiNuGAN: Dual Segmentation of Mitoses and Nuclei Using Conditional GANs on Multi-center Breast H&E Images

Salar Razavi^{a,*}, Fariba D. Khameneh^a, Hana Nouri^a, Dimitrios Androutsos^a, Susan J. Done^b, April Khademi^a^a Electrical, Computer and Biomedical Engineering, Ryerson University, Toronto, Ontario, Canada^b Laboratory Medicine Program, University Health Network, Toronto, Ontario, Canada

ARTICLE INFO

Article history:

Received 13 September 2021

Accepted 30 November 2021

Available online 20 January 2022

Keywords:

Mitosis detection

Semantic segmentation

Generative adversarial network

Breast cancer

Focal loss

Computer-aided detection

ABSTRACT

Breast cancer is the second most commonly diagnosed type of cancer among women as of 2021. Grading of histopathological images is used to guide breast cancer treatment decisions and a critical component of this is a mitotic score, which is related to tumor aggressiveness. Manual mitosis counting is an extremely tedious manual task, but automated approaches can be used to overcome inefficiency and subjectivity. In this paper, we propose an automatic mitosis and nuclear segmentation method for a diverse set of H&E breast cancer pathology images. The method is based on a conditional generative adversarial network to segment both mitoses and nuclei at the same time. Architecture optimizations are investigated, including hyper parameters and the addition of a focal loss. The accuracy of the proposed method is investigated using images from multiple centers and scanners, including TUPAC16, ICPR14 and ICPR12 datasets. In TUPAC16, we use 618 carefully annotated images of size 256×256 scanned at $40 \times$. TUPAC16 is used to train the model, and segmentation performance is measured on the test set for both nuclei and mitoses. Results on 200 held-out testing images from the TUPAC16 dataset were mean DSC = 0.784 and 0.721 for nuclear and mitosis, respectively. On 202 ICPR12 images, mitosis segmentation accuracy had a mean DSC = 0.782, indicating the model generalizes well to unseen datasets. For datasets that had mitosis centroid annotations, which included 200 TUPAC16, 202 ICPR12 and 524 ICPR14, a mean F1-score of 0.854 was found indicating high mitosis detection accuracy.

Key Messages

Segmentation of Mitoses and Nuclei using Conditional GANs

Introduction

Mitosis is the process of cell duplication, in which one cell divides into two genetically identical daughter cells.¹ In breast tumors, the number of mitotic figures is related to tumor aggressiveness and proliferation² and is used for tumor grading and determining treatment options. Experienced pathologists determine mitotic counts through visual examination of hematoxylin and eosin (H&E) stained tissue sections by counting the number of mitotic figures in 10 high-power fields (HPF). Due to large variations in mitosis appearance, the laborious and subjective nature of the task, as well as variability in preparation protocols (staining, scanning), manual mitosis counts have up to 25% discordance between pathologists.³ Computer-aided diagnostic (CAD) tools hold promise to overcome the laborious and subjective nature of mitotic grading⁴ by providing robust, consistent, and efficient analysis. The public mitosis detection challenges at ICPR12,⁵ AMIDA13,⁶ ICPR14,⁷

and TUPAC16⁸ have facilitated development of various artificial intelligence (AI) solutions to this problem. These methods employed a variety of tools from handcrafted features,⁹ deep learning,¹⁰ or a combination of both machine learning (ML) and deep learning (DL) systems.^{11,12} Recently, DL systems for automated mitosis detection in breast cancer images have gained in popularity due to the ability of these systems to adapt to diverse data with often superior performance compared to traditional machine learning methods.^{9,13} The DL methods employed on TUPAC16, ICPR12, and ICPR14 datasets have largely considered convolutional neural networks (CNNs) to classify candidate regions¹⁴ or to detect mitotic figures.¹ State-of-the-art methods apply a multistage methodology to locate ROI with candidate mitoses and then detect cells in mitosis in those regions.^{1,10,15–19} To improve the performance and generalization, pre-processing methods are commonly employed such as wavelet decomposition,²⁰ color deconvolution,^{21,22} and stain normalization.^{14,23,24} While most methods focus on obtaining mitotic counts through mitosis detection, there could be value in segmenting the mitosis as a whole, which can easily facilitate mitosis counts, and can be used to analyze properties of the mitoses themselves (such as mitotic stage). Mitosis segmentation is the focus of this

* Corresponding author.

E-mail addresses: salar.razavi@ryerson.ca (S. Razavi), fdambandkhameneh@ryerson.ca (F.D. Khameneh), hana.nouri@ryerson.ca (H. Nouri), dimitri@ryerson.ca (D. Androutsos), susan.done@uhn.ca (S.J. Done), akhademi@ryerson.ca (A. Khademi).

work. Generative adversarial networks (GANs) have recently been applied to translate histopathology images between domains. Rana et al utilized a GAN-based CNN network to virtually stain H&E specimens. Hou et al²⁵ synthesize histopathology images based on nuclei masks which are then used for task-specific CNN training. Jerry Wei et al²⁶ presented an image translation methodology to generate augmented data to mitigate data imbalance issues in colorectal histopathology images. Bentaib et al uses a discriminator with stain normalization that transfers stains across different datasets with various staining appearance. Tellez et al²⁷ use an unsupervised image-to-image translation for color normalization and augmentation and the effects of these preprocessing steps are evaluated on different classification tasks. In, Quiros et al,²⁸ pathologyGAN is proposed to segment nuclei and learn pathological features within cancer tissue images to correlate breast cancer disease with molecular information. Mahmood et al²⁹ use a conditional GAN to generate H&E training images with labels and both synthetic and real images are utilized to train an unpaired GAN framework to segment nuclei. Inspired by the challenges of mitosis segmentation and previous works in image translation using GANs, this paper proposes the use of conditional generative adversarial networks (cGANs) to segment both mitoses and nuclei at the same time. cGANs may hold potential by constraining solutions to be more "realistic" through discriminators which could mitigate inconsistency and generalization challenges. Although the main application is mitosis segmentation and detection, dual segmentation of nuclei serves two purposes. Firstly, providing the classifier with annotations of non-dividing nuclei may help the system learn the mitosis class more effectively. Secondly, segmenting nuclei can help to facilitate automated nuclear grading. The cGAN model utilized in this work consists of a conditional encoder decoder network for a generator which offers a coarse to fine representation of imaging data, and a multiscale discriminator network that uses both local and global scales to differentiate between real and fake images. The result is a translated image that represents the segmented mitoses and nuclei. Ground-truth delineations for nuclei and mitoses provide conditional labels for the generator and a focal loss is investigated to emphasize the mitosis (imbalanced) class. Adversarial and feature matching losses from multiple scales are utilized. Training on multiple scales enables the architecture to generate images with finer details. Mitosis and nuclear annotations were developed for 618 images from the TUPAC16 dataset (3 centers) where 418 were used to train and validate the segmentation model. This dataset was acquired from three different centers and has wide color variability. Segmentation accuracy for both nuclei and mitoses is measured using the DSC through 5-fold cross-validation, and is compared to baseline U-Net and U-Net with focal loss. To examine mitosis detection in unseen datasets, the model was tested on images from ICPR12 (202), ICPR14 (524), and held out TUPAC16 (200). In multicenter breast cancer pathology images, data variability can arise from different staining protocols, scanner vendors, and patients with different tissue morphology

which can create generalization challenges for deep learning methods. Therefore, experimenting with large and diverse datasets is needed to optimize such tools for clinical translation.

Dataset

In this work, we consider three open source datasets for mitosis detection and segmentation: TUPAC16, ICPR12, and ICPR14. From each dataset, random patches of size 256×256 were extracted from the HPF from each dataset such that patches had at least one mitosis that was not on the boundary. TUPAC16 has a total of 656 HPF for 73 patients, scanned at 3 different centers. Annotations for both nuclei and mitoses were generated for TUPAC16 since it contains more variability (three centers with wide stain variations). Annotations were generated using Pathcore Sedeen³⁰ by a biomedical student trained by a breast pathologist. Pixel-level annotations were completed for 48 of the 73 patients resulting in 418 patches for training and validation. This dataset is used to fine-tune the segmentation model and to determine validation performance over all 418 patches through 5-fold cross-validation. Other model comparisons also use this dataset to compare segmentation performance. Example annotations are shown in Fig. 1. During testing, patches from 25 patients from TUPAC16 are used to determine held-out mitosis and nuclear segmentation performance. Images from these patients are not used during training. The number of patches from the 25 patients for testing was 200. In the ICPR12 dataset, there are images with pixel-level annotations for mitosis from five patients, acquired using two scanners (Aperio and Hamamatsu) for a total of 70 HPF. In this work, 202 test images of size 256×256 were sampled from the HPF to create the ICPR12 test set. In ICPR14, there are 11 HPF scanned with Aperio and Hamamatsu scanners and 524 test images of size 256×256 were sampled from all HPF. In this dataset, only mitosis centroid annotations are available. ICPR12 and ICPR14 datasets are used to measure generalization performance. The total number of patches generated for each dataset is described in Table 1. All the images are scanned at $40 \times$ magnifications.

Methods

The proposed cGAN architecture called MiNuGAN for dual mitosis and nuclear segmentation is adapted from the original pix2pixHD cGAN architecture³¹ and is shown in Fig. 2. An encoder-decoder architecture with ResNet³² blocks are used for the generator network, consisting of five convolutional layers and nine residual blocks in the encoding arm, four convolutional upsampling layers in the decoder, followed by three residual blocks after deconvolution, and two more convolution layers. The output is a generated segmentation mask for both mitosis and nuclear classes. After each convolution layer, instance normalization³³ is employed,

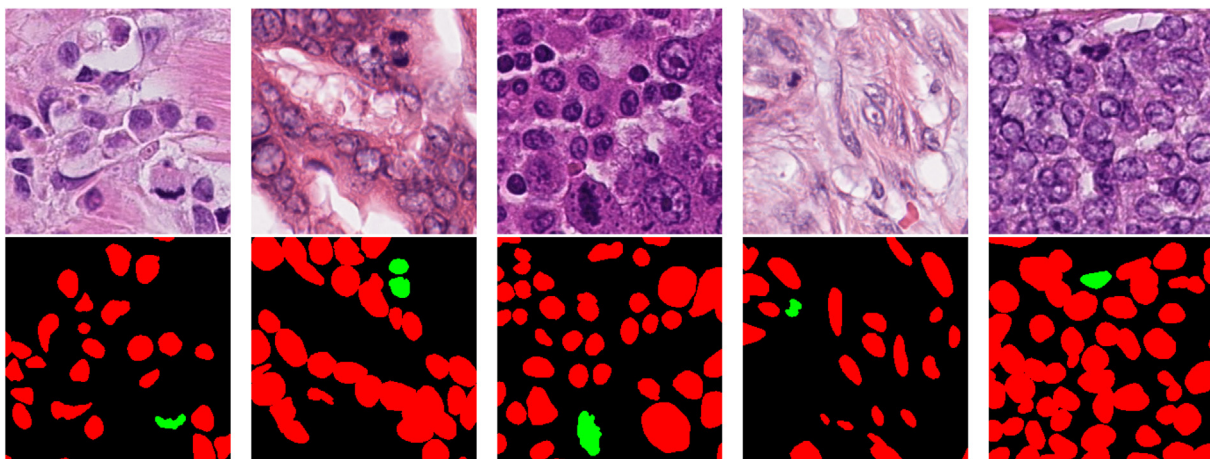


Figure 1. Mitosis (green) and nuclear (red) annotations from TUPAC16.

Table 1
Experimental datasets.

Dataset	Patches	Centres	Scanners	Mitoses	Labels
TUPAC16	618	3	Leica, Aperio	730	Pixel-level (mitosis, nuclear), centroid (mitosis)
ICPR14	524	1	Aperio, Hamamatsu	550	Centroid (mitosis)
ICPR12	202	1	Aperio, Hamamatsu	226	Pixel-level (mitosis), centroid (mitosis)

followed by rectified linear unit (ReLU) activation. The coarse-to-fine generator encodes features from multiple scales, which permits for different sized-details to be detected. Skip connections are added between the local enhancer (G_1) and global (G_2) enhancer which aims to maintain and transfer global information from G_1 to G_2 . G_1 and G_2 generators are jointly trained to generate masks from input images and the inputs are 128×128 and 256×256 in size, respectively. The last convolution layer in the architecture is followed by *Tanh*. The discriminator in Fig. 2 predicts whether the generated segmentation masks are real or fake. The discriminators use images of size 256×256 , 128×128 , and 64×64 for scales 1, 2, and 3. The baseline discriminator has three multiscale discriminators (D_1 , D_2 , D_3) which use images of different resolutions to predict image realism based on multiple resolutions. For each scale, the discriminator adopts the PatchGAN³⁴ architecture with four convolutional layers, followed by a sigmoid output layer for classification. The multiscale PatchGAN discriminator classifies patches of various resolutions. The multiscale D_1 , D_2 , and D_3 discriminators are trained simultaneously to distinguish the generated segmentation masks from ground truth images.

In this work, a focal loss³⁵ is implemented in the generator and deployed after the last hidden layer to improve mitosis and nuclear segmentation performance. Unlike general classification tasks that assume an equal occurrence from all of the classes, medical applications suffer small sample sizes and resulting class imbalances. In mitosis segmentation, there is a heavy imbalance between the classes (there are much more nuclei than mitosis pixels) which can hamper the detection of mitotic cells.⁶ Training a network on imbalanced datasets can make the network biased towards

learning more representations of the data-dominated class (i.e., nuclei) while other classes (i.e., mitosis) will receive less importance. A focal loss can be used to reduce the contribution of the more commonly occurring nuclear pixels and focus the classifier on the harder to predict (mitosis) pixels and is incorporated into the total loss function for cGANs in this work. The total loss function for cGANs is a combination of the adversarial loss L_{GAN} and the feature matching (FM) loss L_{FM} as in

$$\min_G \max_{D_1, D_2} \sum_{k=1,2} L(D_k, G) = \min_G \max_{D_1, D_2} \sum_{k=1,2} L_{GAN}(D_k, G) + \lambda \sum_{k=1,2} L_{FM}(D_k, G)$$

where λ balances the influence of the feature matching loss L_{FM} with the adversarial loss L_{GAN} to prevent overfitting to specific distributions. The feature matching loss considers dissimilarities between feature representations in local and global levels. The adversarial loss tries to maximize the probability of real images and minimize the probability of fake inputs to the discriminator. In this work, L_{GAN} is computed by considering a focal loss parameter γ that reduces or increases the relative loss for well-classified or poorly classified pixels, as in

$$L_{GAN}(D_k, G)_{FL} = \mathbb{E}[(1-D)^\gamma \log(D)] + \mathbb{E}[(D)^\gamma \log(1-D)]$$

The output of the model is a multiclass image with pixel-level labels for each of the predicted mitosis and nuclear objects. To measure segmentation performance, pixel-level predictions are compared to ground truth delineations on a per-pixel basis. For mitosis detection, the mitosis

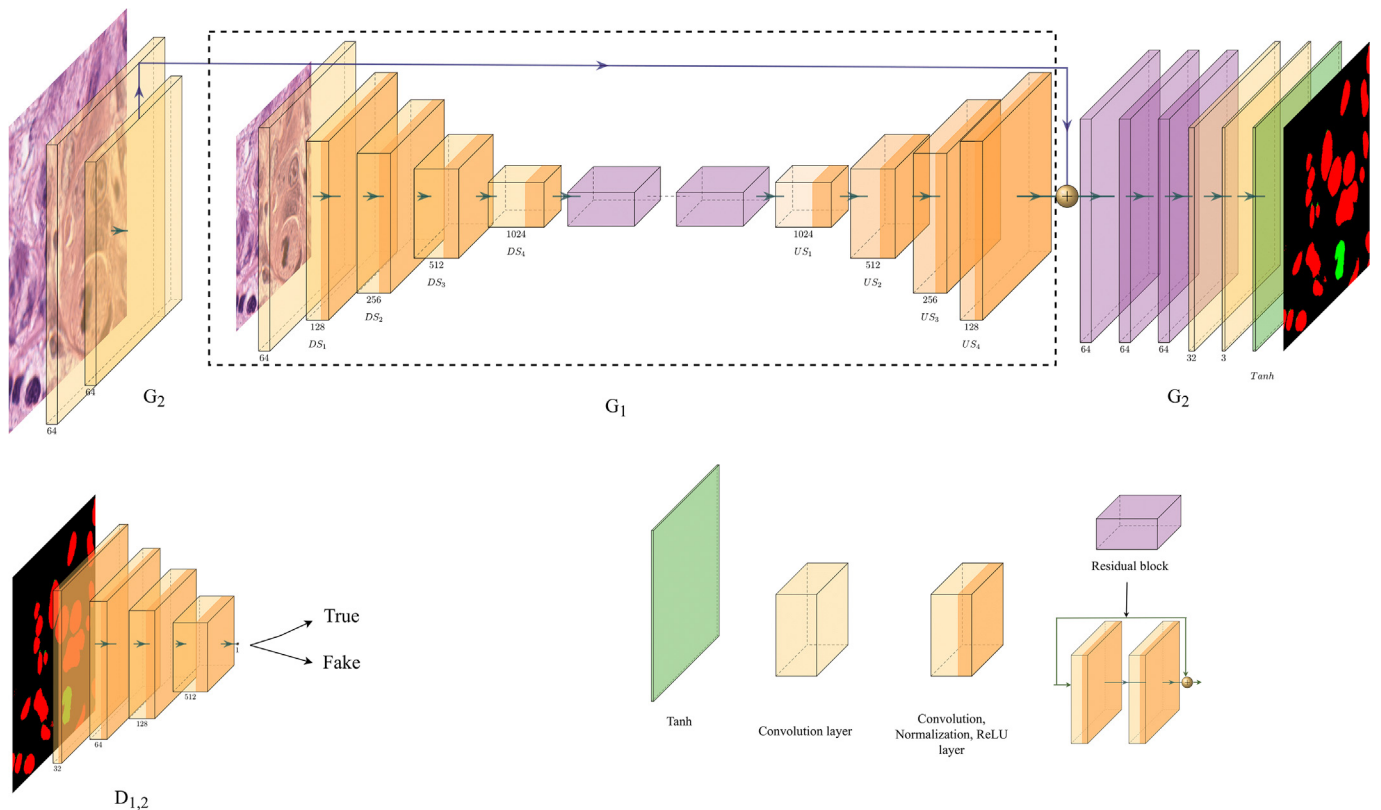


Figure 2. MiNuGAN for dual mitosis and nuclear segmentation in H&E breast cancer images.

Table 2
cGAN configurations and average DSC of mitotic and nuclear classes.

Configuration	Discrim	Resblocks	Lambda	DSC
1	2	9	0	0.739 ± 0:208
2	2	9	0.1	0.742 ± 0:200
3	2	9	1	0.741 ± 0:196
4	2	9	2	0.747 ± 0:194
5	2	9	5	0.738 ± 0:208
6	3	9	0	0.747 ± 0:189
7	3	9	0.1	0.747 ± 0:199
8	3	9	1	0.737 ± 0:210
9	3	9	2	0.758 ± 0:192
10	3	9	5	0.752 ± 0:185

segmentation masks are used to automatically determine the centroids of each mitosis object.

Metrics

We evaluate segmentation performance of the cGAN model using the dice similarity coefficient (DSC) which compares automated segmentations to manual delineations for the mitosis and nuclear classes separately. To measure mitosis detection performance on datasets with ground-truth mitosis

centroids, the F1-score is utilized. Precision and sensitivity are determined by the number of correctly predicted mitosis centroids within a 25 pixel-size window centered on the ground truth centroid. This corresponds to a distance of 7 μM and was chosen since it is the average of the mitosis competitions.

Results and Discussion

All models are developed using TUPAC16 data with pixel-level annotations of nuclei and mitoses (334/84 training/validation split and 5-fold cross-validation). Average segmentation performance is reported for all 418 patches for nuclei and mitoses separately. Model fine-tuning is done using the detailed annotation dataset and the final dual nuclear and mitosis segmentation model termed "MiNuGAN" is carried forward for further analysis. To measure generalization performance of the proposed system, two held-out test sets with pixel-level annotations were used. There were 200 patches from TUPAC16 that had mitosis and nuclear annotations, and 202 patches from ICPR12 with mitosis delineations. The segmentation results of MiNuGAN are compared to baseline models, such as U-Net and U-Net with a focal loss, which all were trained using the same 5-fold cross-validation datasets. In addition to measuring segmentation performance, additional experiments were conducted to measure mitosis detection accuracy through the F1-scores. These are reported all held out datasets for mitosis centroid

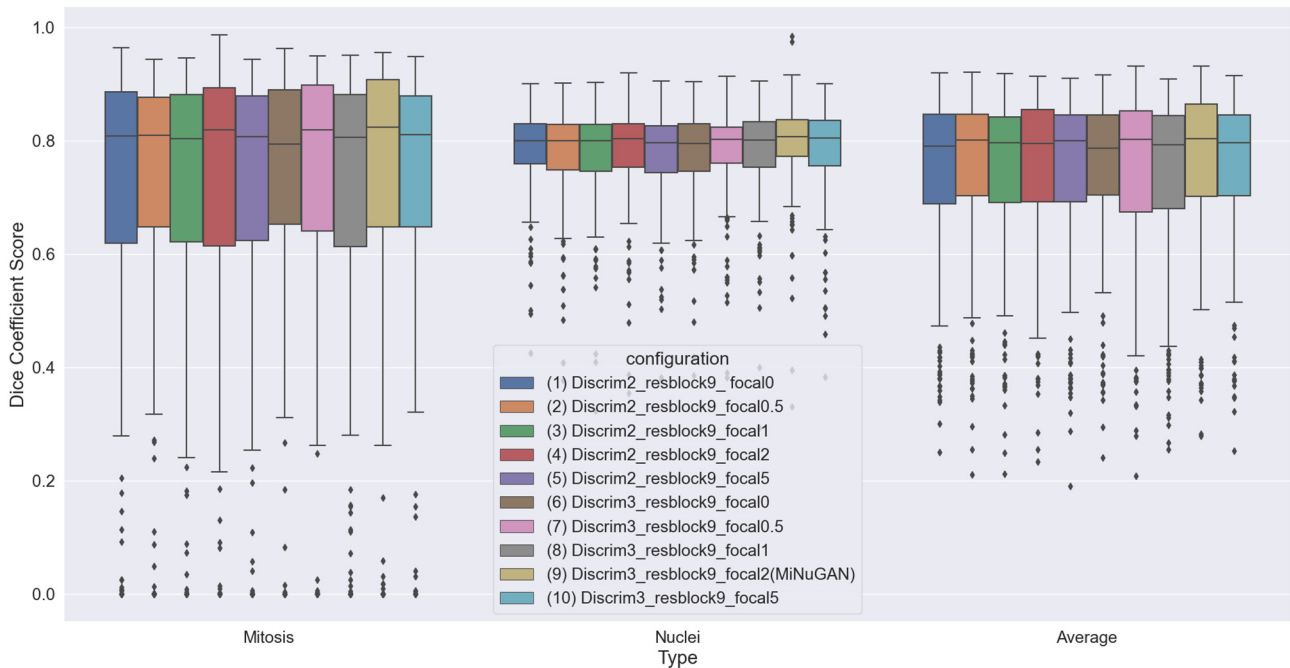


Figure 3. DSC distributions for cGAN configurations in Table 2.

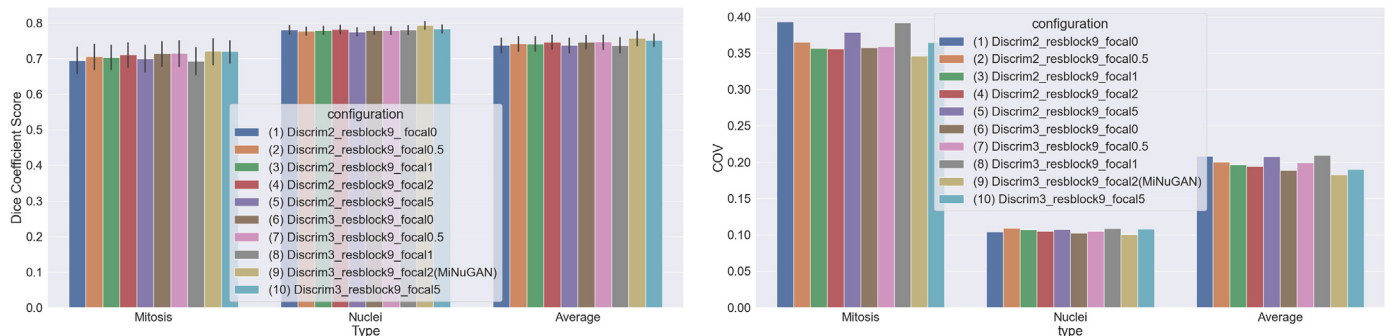


Figure 4. Mean DSC and DSC coefficient of variation (CoV) for cGAN configurations.

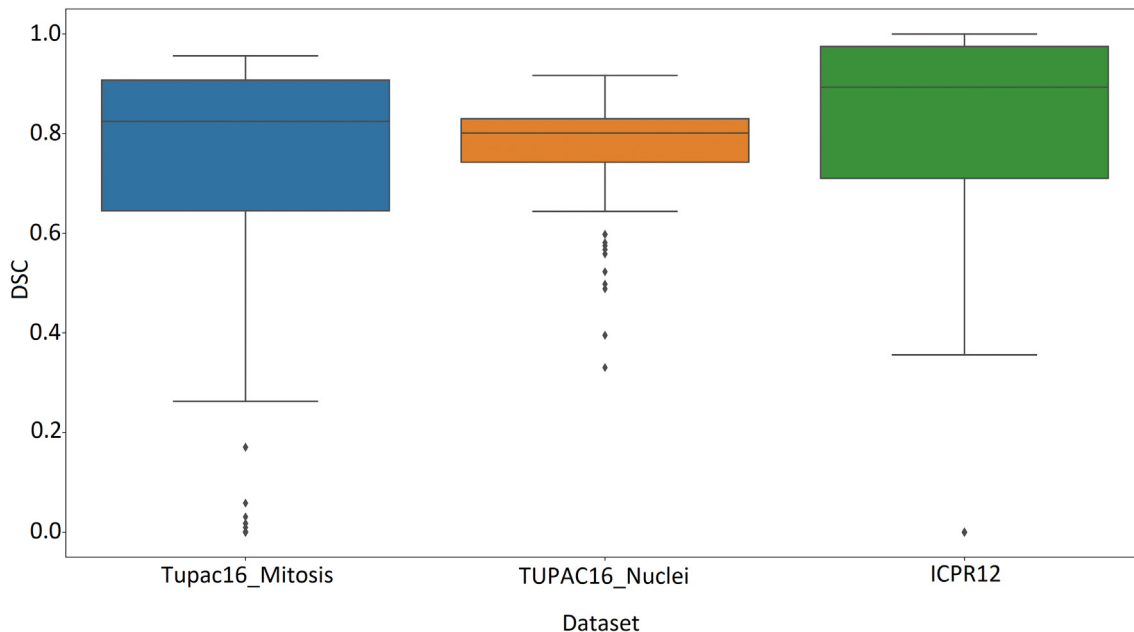


Figure 5. Mitosis segmentation performance for TUPAC16 and ICPR12.

annotations, which includes: ICPR12 (202 patches), ICPR14 (524 patches), and for the held-out TUPAC16 set (200 patches). All models are trained on Nvidia 2080TI GPU with 11 GB memory. The training images consist of original images along with RGB label maps (with each color representing either the nuclei, boundary or background class). All images are of size 256×256 .

Mitosis and Nuclear Segmentation

Performance of the proposed mitosis and nuclear segmentation model is reported in this section. In total, 418 patches (334 training/84 validation, 5-fold cross-validation) from TUPAC16 are used for training and model fine-

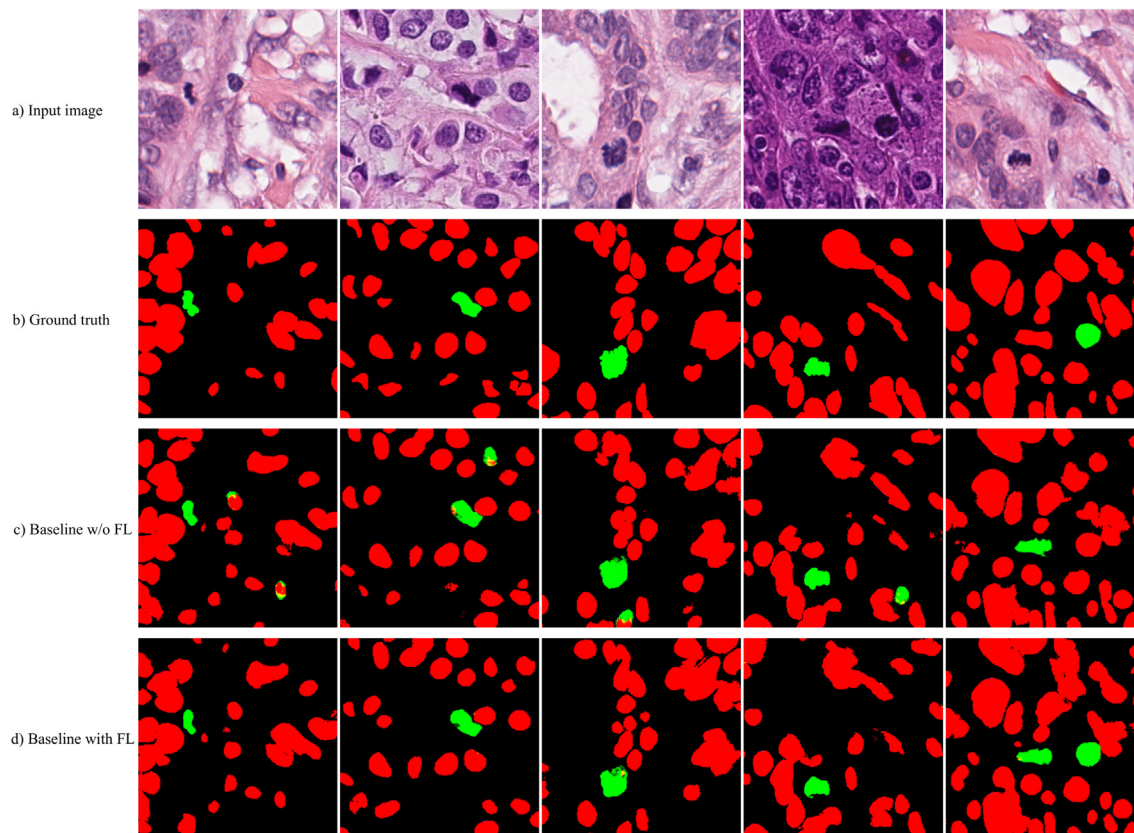


Figure 6. cGAN segmentation results with and without focal loss. a) original images, b) ground truth, c) configuration 1 (baseline without focal loss) and d) configuration 4 (baseline with focal loss).

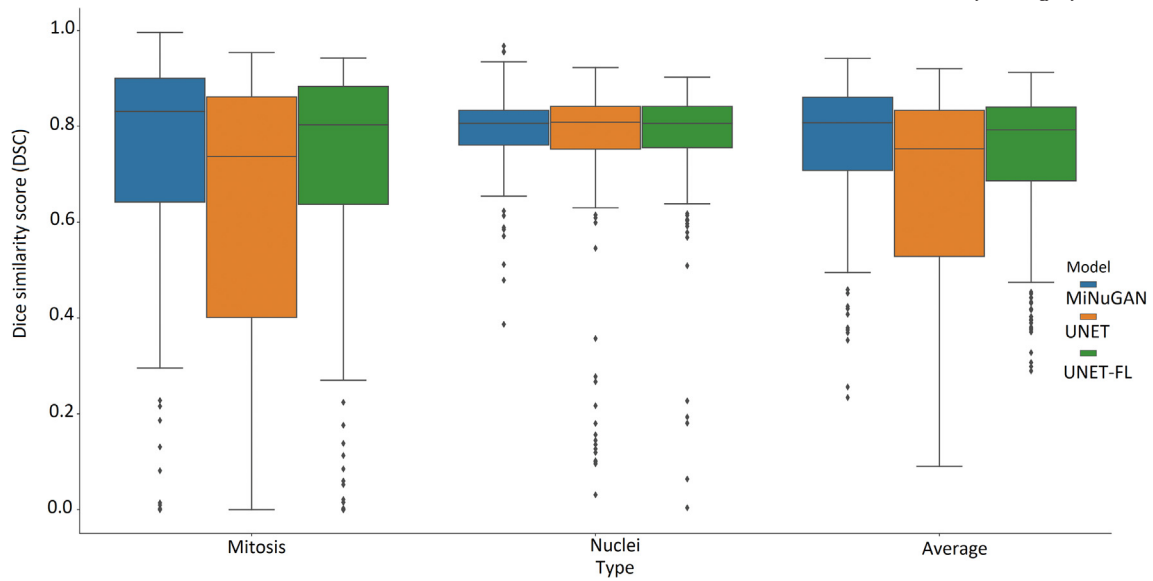


Figure 7. DSC distributions for mitoses and nuclei for MiNuGAN, UNET, UNET/focal.

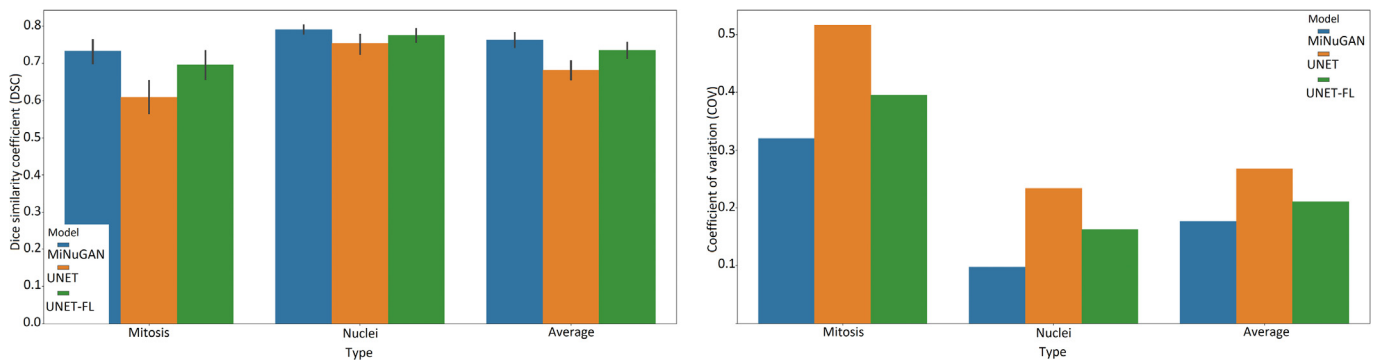


Figure 8. Mean DSC and DSC coefficient of variation for MiNuGAN, UNET, UNET/focal.

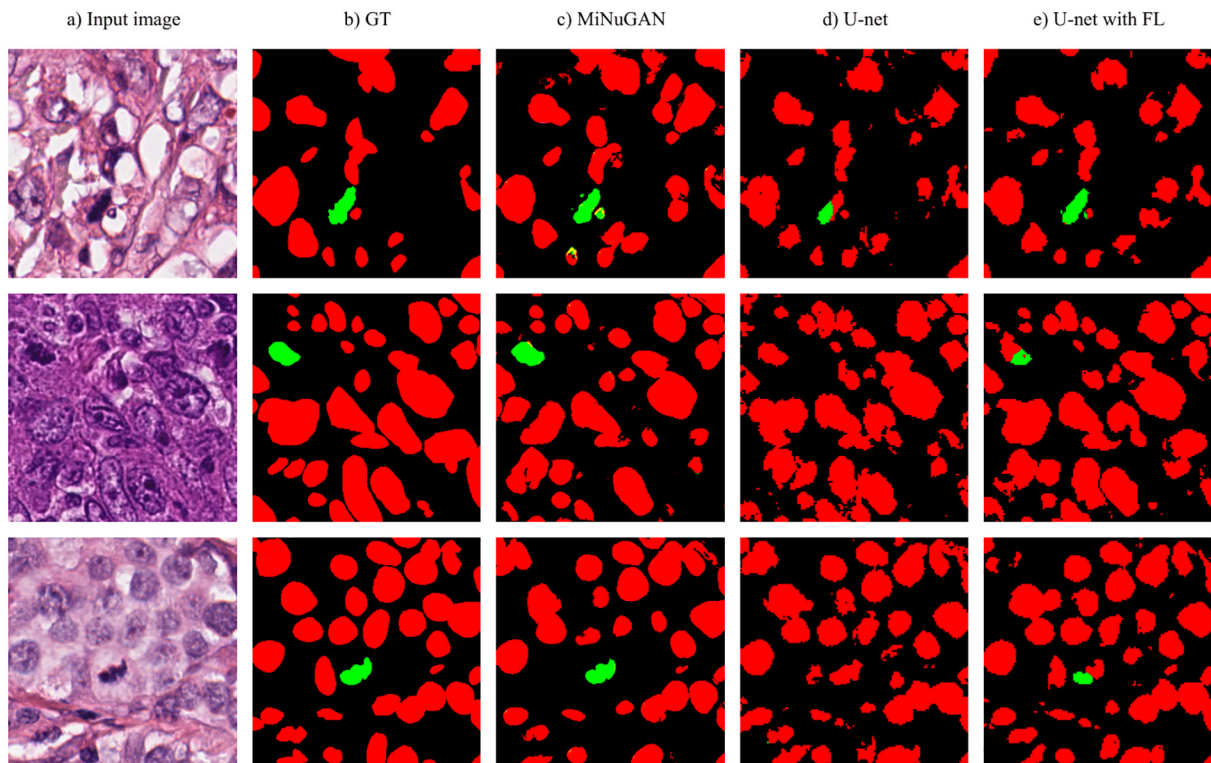


Figure 9. Segmentation results using three different models. a) input image, b) ground truth, c) MiNuGAN d) U-Net e) U-Net with focal loss.

tuning without any preprocessing such as stain normalization. The dataset is split in a 5-fold stratified fashion (mixing the validation and training sets) and average performance on the validation data over 5-folds is reported. Each model is trained for 100 epochs with a learning rate of 0.5. The baseline pix2pixHD model has three multiscale discriminators, nine residual blocks in the generator and uses $\lambda = 5$. In addition to these parameters, we also investigated a focal loss with $\gamma = 0; 0.1; 1; 2; 5$ to increase the importance of mitosis when classifying pixels. We varied the number of discriminators, residual blocks, λ , and the focal loss and report some of the top-performing configurations for mitosis and nuclear segmentation. In total, this resulted in 10 configurations as outlined in Table 2. The average DSC over both mitosis and nuclear classes in the validation set, for 5-folds for each configuration is shown in the last column of Table 2, and graphically in Figs 3 and 4. The highest DSC for the mitosis class is obtained for configurations that use a focal loss and baseline parameters: three multiscale discriminators, nine residual blocks, and $\gamma = 2$. This configuration also

results in the lowest coefficient of variation, indicating that this model segments mitoses with the highest consistency. Visual results for the top-performing method (baseline + focal loss) and the same architecture without focal loss are shown Fig. 6. Without focal loss, there are lots of small false-positive mitoses that would decrease mitosis detection accuracy. Predictions based on a focal loss have more accurately segmented the mitotic figures.

When considering the DSC for the nuclear and mitosis class separately, it is apparent that lower performance is generally obtained for the cells in mitosis. This could be a result of the complex nature and variable appearance of mitotic cells and could also be due to a relative class imbalance. However, as shown in Fig. 6, the focal loss improves the imbalance issue and segments mitoses with improved accuracy. Furthermore, in Figs. 3 and 4, using fewer residual blocks in the generator decreases overall DSC performance and nine ResNet blocks (as per baseline Pix2pixHD) seems appropriate for this task. Results also show that increasing the number of the

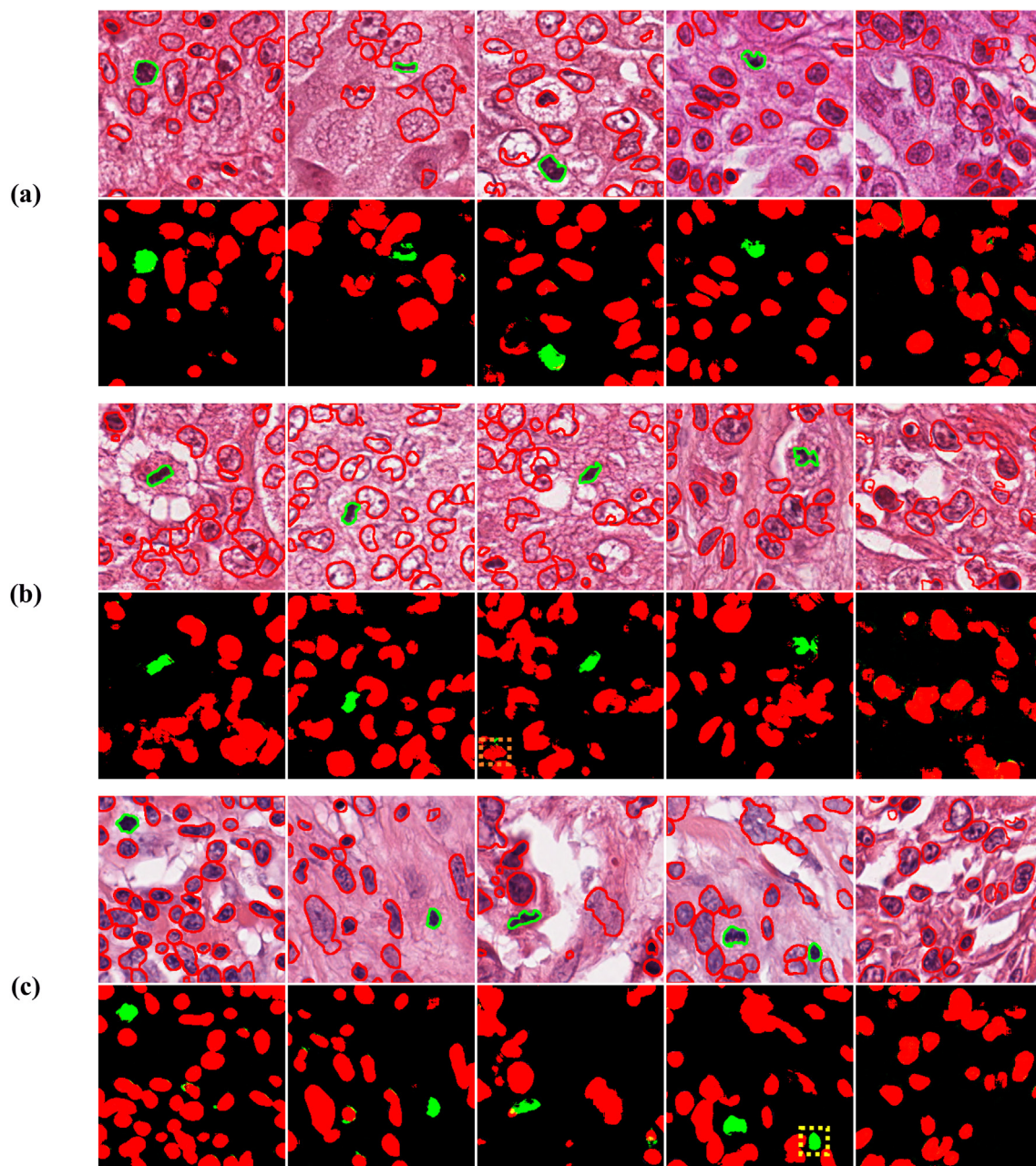


Figure 10. Segmentation and detection of mitosis using MiNuGAN for a) ICPR12, b) ICPR14, c) TUPAC16. Segmented mitoses are shown in green and nuclei in red. False positives are denoted by a yellow box, and false negatives by an orange box.

discriminators from two to three increases the DSC, which may indicate that mitotic cells are better detected at lower resolutions. For the remainder of this work, we use the baseline cGAN (three multiscale discriminators, nine residual blocks, and $\gamma = 2$) with focal loss and call this configuration MiNuGAN.

Using the finalized model MiNuGAN just described, the mitosis segmentation testing performance on the held-out TUPAC16 and ICPR12 is computed. In total, there were 402 patches, and the results are summarized in Fig. 5. The mean mitosis and nuclear segmentation performance on TUPAC16 were: 72.1% and 78.4%, respectively. On ICPR12, the mean mitosis segmentation accuracy was 78.2%, which is slightly higher than the TUPAC16 dataset (on which the model was trained). This may be due to a number of factors. Firstly, the TUPAC16 dataset has large variability and may be allowing the model to learn more diverse features which permits for better generalization to unseen data.

Second, the ICPR12 dataset is of relatively high contrast, which may permit for better mitosis segmentation. Third, the discriminator function allows the classifier to learn realistic mitosis patterns. The high segmentation performance across both datasets is a good indication that these features are beneficial for generalization to new data. To compare the performance of the MiNuGAN to baseline systems, we compared segmentation results over the validation set with U-Net and U-Net with a focal loss. These systems were trained for 100 epochs using Adam optimizer, 0.001 learning rate, batch size of 16, 4 layers and with/without focal loss. The same training and validation datasets were used in a 5-fold cross-validation manner. Figure 7 shows the DSC distribution and Fig. 8 shows the mean DSC and the DSC coefficient of variation (CoV) of both U-Net models versus MiNuGAN. The mean DSC segmentation performance is lowest for U-Net without focal loss, which is slightly improved when a focal loss is used. The highest mean DSC performance is achieved by MiNuGAN for both nuclear and mitosis datasets. Additionally, the coefficient of variation is significantly lower for the proposed method, indicating results are more consistent in this multicenter dataset. Visual results are shown in Fig. 9. U-Net misses mitoses and also generates more false positives, even with focal loss. In contrast, the proposed method has less false positives and negatives, and the segmentation masks are smoother, which may be attributed to the multiscale discriminator which forces the output segmentation masks to be more realistic (i.e., similar to ground truths).

Mitosis Detection Performance using MiNuGAN

Using the baseline cGAN system with focal loss (MiNuGAN), the detection performance of the proposed method is evaluated on images from multiple datasets and scanners (held-out TUPAC16 samples: 200 images, ICPR12: 202 images and ICPR14: 524 images). As the model predicts both nuclei and mitoses on a per-pixel basis, only mitosis labels are retained for detection. Morphological processing is applied to the mitosis masks (small erosion, hole closing) and the centroid of every object is determined. The ground-truth centroid is compared to the predicted centroids through the F1. See Table 3 for the average F1 performance for each dataset. The average F1 over all three datasets was found to be 0.854 ± 0.268 and considering that over 80% of the data comes from sources that were not seen by the classifier during training this is a very promising result. F1-scores on ICPR12 and ICPR14 were lower compared with the TUPAC16 dataset (roughly 10% performance drop). See Fig. 10 for example mitosis detection results for MiNuGAN over all three datasets. The variation between dataset distributions (color, stain, patient, scanner) could be the cause for some of the challenges in ICPR12 and ICPR14, although over all the model is performing well. The last

Table 3

Mitosis detection performance on different datasets and scanners.

	Dataset				Scanner
	TUPAC16	ICPR14	ICPR12	Aperio	Hamamatsu
F1-score	0.872	0.847	0.851	0.832	0.881

column in Fig. 10 shows example segmentations in patches with no mitoses. As can be seen, only nuclei were detected in these images. In future works, we will apply this method on whole slide images and determine the effect of negative patches on performance. We may consider an RCNN to find candidate regions first, and then use MiNuGAN to segment and detect mitoses from those patches. To improve generalization, we will investigate color normalization as a means to mitigate stain and color variability in multicenter datasets.

Conclusion

This work proposes a cGAN-based model called MiNuGAN that segments both mitoses and nuclei at the same time from H&E stained breast cancer images. It is based on the baseline pix2pixHD model, except a focal loss has been added to bring attention to the imbalanced mitosis class. Testing results show that segmentation performance is high for both mitoses and nuclei in TUPAC16 with mean DSC = 0.721 and 0.784 for mitoses and nuclei, respectively, over 200 image patches and mean DSC = 0.782 for 202 ICPR12 images. Results show the proposed method generates more realistic and accurate segmentation maps compared to baseline U-Net models, even U-Net with focal loss, indicating the discriminator forces continuity in the segmentation mask that is beneficial for segmentation. Using the mitosis segmentations, mitosis centroids were detected and compared to datasets with centroid annotations. The F1-scores were computed over three multicenter H&E datasets (TUPAC16, ICPR12, ICPR14) and an average F1-score of 0.854 was achieved. Ablation studies show the effectiveness of the focal loss to increase the performance of mitosis segmentation and detection. In the future, stain normalization will be considered, as well as augmentation to see if performance across scanners and centers can be improved. In this work, we focused mainly on mitosis segmentation and detection in breast cancer H&E images. In the future, it is possible to evaluate MiNuGAN on different tissue types and datasets.

Acknowledgements

We thank the Canadian Cancer Society for funding this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpi.2022.100002>.

References

- Sebai M, Wang X, Wang T. Maskmitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images. *Med Biol Eng Comput* 2020;58:1603–1623. <https://doi.org/10.1007/s11517-020-02175-z>.
- Paul JVD, van der Wall Elsen, Baak JPA. Prognostic value of proliferation in invasive breast cancer: a review. *J Clin Pathol* 2004;57(7):675–681.
- Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *Jama* 2015;313(11):1122–1132.
- Mercan E, Mehta S, Bartlett J, Shapiro LG, Weaver DL, Elmore JG. Assessment of machine learning of breast pathology structures for automated differentiation of breast cancer and high-risk proliferative lesions. *JAMA Network Open* 2019;2(8):e198777.
- Roux L, Racoceanu D, Lomenie N, et al. Mitosis detection in breast cancer histological images an icpr 2012 contest. *J Pathol Inform* 2013;4.
- Veta M, Van Diest PJ, Willems SM, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal* 2015;20(1):237–248.
- Roux. MITOS-ATYPIA-14. <https://mitos-atypia-14.grand-challenge.org/>.
- Veta M, Heng YJ, Stathonikos N, et al. Predicting breast tumor proliferation from whole-slide images: the tupac16 challenge. *Med Image Anal* 2019;54:111–121.
- Irshad H, Roux L, Racoceanu D. Multi-channels statistical and morphological features based mitosis detection in breast cancer histopathology. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2013. p. 6091–6094.
- Li C, Wang X, Liu W, Latecki LJ, eds. *Med Image Anal* 2018;45:121–133.
- Malon CD, Cosatto E. Classification of mitotic gues with convolutional neural networks and seeded blob features. *J Pathol Inform* 2013;4.
- Mahmood T, Arsalan M, Owais M, Lee MB, Park KR. Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster r-cnn and deep cns. *J Clin Med* 2020;9(3):749.

13. Irshad H, Gouaillard A, Roux L, Raco-cEANU D. Multispectral band selection and spatial characterization: Application to mitosis detection in breast cancer histopathology. *Comput Med Imag Graphics* 2014;38(5):390–402.
14. Das DK, Dutta PK. Efficient automated detection of mitotic cells from breast histological images using deep convolution neutral network with wavelet decomposed patches. *Comput Biol Med* 2019;104:29–42.
15. Ma M, Shi Y, Li W, Gao Y, Xu J. A novel two-stage deep method for mitosis detection in breast cancer histology images. 2018 24th International Conference on Pattern Recognition (ICPR). IEEE; 2018. p. 3892–3897.
16. Paeng K, Hwang S, Park S, Kim M. A unified framework for tumor proliferation score prediction in breast histopathology. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer; 2017. p. 231–239.
17. Li Y, Mercan E, Knezevitch S, Elmore JG, Shapiro LG. Efficient and accurate mitosis detection - a lightweight rcnn approach. *Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM. INSTICC, SciTePress*; 2018. p. 69–77.
18. Aubreville M, Bertram CA, Klopfeisch R, Maier A. Augmented mitotic cell count using eld of interest proposal. *Bildverarbeitung fur die Medizin* 2019. Springer; 2019. p. 321–326.
19. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J Pathol Inform* 2016;7.
20. Wan T, Zhang W, Zhu M, Chen J, Achim A, Qin Z. Automated mitosis detection in histopathology based on non-gaussian modeling of complex wavelet coefficients. *Neurocomputing* 2017;237:291–303.
21. Tellez D, Balkenhol M, Otte-Holler I, et al. Whole-slide mitosis detection in H&E breast histology using phh3 as a reference to train distilled stain-invariant convolutional networks. *IEEE Trans Med Imaging* 2018;37(9):2126–2136.
22. Irshad H. Automated mitosis detection in histopathology using morphological and multi-channel statistics features. *J Pathol Inform* 2013;4.
23. Lafarge MW, Pluim JPW, Eppenhof KAJ, Moeskops P, Veta M. Domain-adversarial neural networks to address the appearance variability of histopathology images. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer; 2017. p. 83–91.
24. Ma Y, Sun J, Zhou Q, Cheng K, Chen X, Zhao Y. Chs-net: a cascaded neural network with semi-focal loss for mitosis detection. *Asian Conference on Machine Learning*; 2018. p. 161–175.
25. Hou L, Agarwal A, Samaras D, Kurc TM, Gupta RR, Saltz JH. Unsupervised histopathology image synthesis. *arXiv preprint* 2021;abs/1712.05021:arXiv-1712:arXiv:1712.05021.
26. Wei J, Suriawinata A, Vaickus L, et al. Generative image translation for data augmentation in colorectal histopathology images. *Proceedings of machine learning research* 2019;116:10.
27. Tellez D, Litjens G, Bandi P, et al. Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. *Med Image Anal* 2019;58, 101544.
28. Quiros AC, Murray-Smith R, Yuan K. Pathology gan: learning deep representations of cancer tissue. *Machine Learning for Biomedical Imaging* 2019;1:1–10.
29. Mahmood F, Borders D, Chen R, et al. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE Trans Med Imaging* 2019;39:3257–3267.
30. Martel AL, Hosseinzadeh D, Senaras C, et al. An image analysis resource for cancer research: Piip|pathology image informatics platform for visualization, analysis, and management. *Cancer Res* 2017;77(21):e83–e86.
31. High-resolution image synthesis and semantic manipulation with conditional gans. In: Wang T-C, Liu M-Y, Zhu J-Y, Tao A, Kautz J, Catanzaro B, eds. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2018. p. 8798–8807.
32. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016. p. 770–778.
33. Pan X, Luo P, Shi J, Tang X. Two at once: Enhancing learning and generalization capacities via ibn-net. *Proceedings of the European Conference on Computer Vision (ECCV)*; 2018. p. 464–479.
34. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2017. p. 1125–1134.
35. Lin T-Y, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. *Proceedings of the IEEE international conference on computer vision*; 2017. p. 2980–2988.