ACCURATE CELL SEGMENTATION BASED ON GENERATIVE ADVERSARIAL NETWORKS AND NUCLEI GUIDE FACTORS

Kostantsa Lavntaniti[†] *Marina E. Plissiti*[†] *Michalis Vrigkas*[⋆] *Christophoros Nikou*[†]

[†] Dept. Computer Science and Engineering, University of Ioannina, Greece.
[⋆]Dept. Communication and Digital Media, University of Western Macedonia, 52100 Kastoria, Greece

ABSTRACT

The accurate segmentation of cells in cervical images is crucial for the recognition of pathological situations and the estimation of their severity. In this work, we investigate the segmentation of both the nucleus and the cytoplasm of each cell based on two generative adversarial networks (GANs). First, we detect the location of the nucleus with the extraction of the nucleus boundaries in each cell, which is obtained by the training of the Nucleus-GAN. The segmented nucleus area serves as a guide factor for the definition of the cell boundary, and it is used as input in the Cell-GAN, for the segmentation of the cell boundaries. As it is verified by the experimental results, the proposed method is efficient and leads to accurate nucleus and cell boundaries, presenting high performance.

Index Terms— Generative Adversarial Networks, Cytological Image Segmentation, Nuclei Detection, Cell Segmentation.

1. INTRODUCTION

The accurate diagnosis of cervical diseases rely on the exhaustive examination of the Pap smear images obtained by the optical microscope. This procedure is time consuming and error-prone when it is performed manually. For this reason, several automated methods have been proposed for the segmentation of the structural parts of the cell, that is the nucleus and the cytoplasm.

Many methods are based on traditional segmentation techniques. These methods exploit the properties of the well established segmentation algorithms, such as iterative thresholding [1], level sets [2], mathematical morphology [3] and physical models [4] in order to obtain the boundaries of the region of the interest in the image.

Furthermore, methods that are based on deep learning are constantly developing, providing new powerful techniques which overcome the complexity of cytological images and provide accurate segmentation and classification of the cells. Thus, convolutional neural networks (CNNs) have been applied for the segmentation of single [5] or overlapping [6] cells. Furthermore, fully convolutional networks (FCNs) techniques [7] and methods that combine traditional techniques, such as watersheds with CNNs have been also proposed [8][9].

Another notable category of deep learning models are the well-known generative adversarial networks (GANs), which are extensively used in many applications. In this framework the Cell-GAN [10] has been proposed, which provides reliable results. However, a prerequisite for the correct execution of this method is the prior definition of the guide factor, which is the location of the nucleus of each cell. In this work, we automatically define the guide factor based on the Nucleus-GAN, which is then used as input in the Cell-GAN for the segmentation of the cells. Experimental results validate the efficiency of the proposed method, as it is described in the following paragraphs.

2. METHOD

The segmentation of cervical cell images comprises three stages: (i) the generation of the guide factors obtained by the Nucleus-GAN, (ii) the segmentation of the cell produced by the Cell-GAN, and (iii) a post-processing step in order to avoid undesirable noise effects in the final cell boundaries. The overview of the method is depicted in Fig. 1.

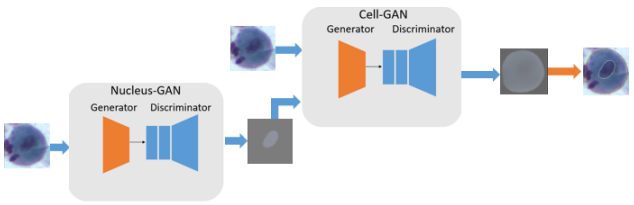

Fig. 1: The combination of Nucleus-Gan and Cell-Gan in the segmentation process. The output of the Nucleus-Gan is used as input at the Cell-GAN.

2.1. Nucleus-GAN

The first stage is to generate nucleus images that approximate the boundaries of the original nucleus in the input images.

For this purpose, the Nucleus-GAN is designed based on generative adversarial networks [11]. Specifically, the Nucleus-GAN generator produces images of the nucleus while the discriminator tries to distinguish the generated from the real images, allowing the generator to obtain a probability distribution of nucleus morphology.

2.2. Cell-GAN

The next step is to extract the boundary of the cell that closely resembles its shape in the original image. For this purpose, we adopt the Cell-GAN [10] approach, which results in reliable segmentation. The Cell-GAN is trained to acquire the probability distribution of cell morphology through the assessment of dissimilarities between the single-cell images it generates, and the ones that are annotated. For the training procedure, it requires guide factors [12] which indicate the location of the nucleus in the image. In our work, the guide factors are the nuclei images generated by the Nucleus-GAN, which contain the accurate area of the nucleus in each image.

2.3. GANs Architecture

The architectures of Nucleus-GAN and Cell-GAN are depicted in Fig. 2 and Fig. 3 respectively. Both networks aim to produce single-cell images $\hat{x}$ from cell images x, namely $G : x \rightarrow \hat{x}$. To achieve this, the generator in both networks is redesigned [10], using the principles of the auto-encoder [13] and the inception model [14]. However, the architecture of the discriminator is the same as in the original DCGAN [15].

2.3.1. Generator

The generator architecture of both networks is based on the autoencoder architecture [13]. It consists of an encoder network and a decoder network. The main structure of the encoder is a four-layer down-sampling network with Leaky-ReLU activation, and each layer adopts the inception module [14].

The use of the inception module aims to capture multi-scale features at different levels of abstraction within the network. This is achieved by using a combination of different filter sizes and operations. The sizes of the filter groups are 1×1, 3×3 and 5×5 in order to promote the generalization ability of the encoder network. The resulting feature maps are then merged and filtered by a final convolutional layer. Finally, batch normalization is applied and the resulting feature map is passed to the next downsampling layer.

Normally, autoencoders use fully connected layers to pass the information from the encoder to the decoder. Due to the complexity of the proposed method, this procedure can lead to an explosive number of parameters. This is why channel-wise fully connected layers are preferred instead.

The input of the Nucleus-GAN is different from the input of the Cell-GAN. The first network is fed with an RGB image showing a cervical cell with its background (203x203x3). It is then processed by a convolutional layer before being received by the first layer of the encoder.

On the other hand, the Cell-GAN is fed with two images, the original cell image, as it is described above, and the corresponding nucleus image produced by the Nucleus-GAN, which we refer to as the guide factor. The purpose of the guide factor is to help the network identify the cell from its noisy background in the received image. The guide factor is processed by two convolutional layers, and the resulting feature map from each layer is merged with the initial image using additive operations.

The decoder network contains 6-layer upsampling networks, each consisting of 3×3 deconvolutions [16] and ReLU activation. The Sigmoid function is used to replace the ReLU activation of the last layer network to generate images. To reduce the complexity of the networks, the decoder layers are simple upsampling layers without using the inception model architecture.

2.3.2. Discriminator

The architecture of the discriminator in both networks is similar to DCGAN [15]. It consists of an input layer, five convolutional layers followed by batch normalization and Leaky-ReLU activation functions, and one output layer that provides the probability score for the input image to be real or fake. The output layer uses the Sigmoid activation function instead of the Leaky-ReLU.

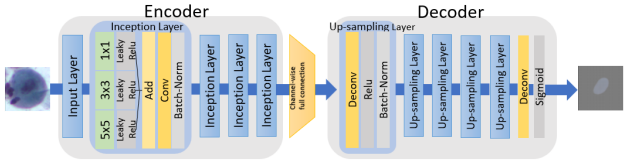

Fig. 2: The architecture of Nucleus-GAN.

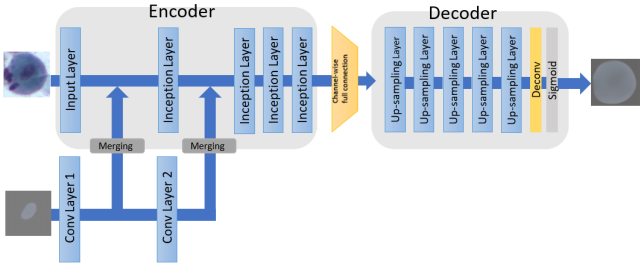

Fig. 3: The architecture of Cell-GAN.

2.4. Joint Loss Function

The joint loss function for each network consists of the adversarial loss function and the L_2 loss function [17].

The adversarial loss function L_{adv} expresses the game-like mechanism used in GANs. However, in this case, the input to both networks is an image rather than random noise. The adversarial loss function can be defined as:

$$L_{adv} = \min_G \max_D E_{\hat{x}\in\hat{X}}[\log(D(\hat{x}))] + E_{x\in X}[\log(1 - D(G(x)))] \quad (1)$$

where D is the output of the discriminator and G is the output of the generator. For the Nucleus-GAN x represents the cervical cell image and $\hat{x}$ the annotated nucleus image, whereas for the Cell-GAN x represents the combination of the cervical cell image with the corresponding guide factor and $\hat{x}$ represents the annotated cell image.

The L_2 loss function focuses more on capturing the overall structure of the image and can be expressed as:

$$L_{\text{smi}} = \alpha\|M\cdot(X-G(x))\|_2+\beta\|(1-M)\cdot(X-G(x))\|_2 \quad (2)$$

where G is the result of the generator, x is the proper input for each network, X is the corresponding annotated image and M is the binary mask of the annotated image. The parameters α and β are used to adjust the generating scale of each network for uncertain cell or nucleus area, retained or discarded. Finally, the joint loss function is expressed as:

$$L_{\text{tot}} = \gamma_{\text{smi}} \cdot L_{\text{smi}} + \gamma_{\text{adv}} \cdot L_{\text{adv}} \quad (3)$$

2.5. Image Post-Processing

The final stage aims to segment the original cell from the input image using the generated cell image. To effectively segment the cell from the original image, it is convenient to create a binary mask of the generated image. The background noise is limited by image blurring with a bilateral filter. The images are then converted to grayscale, and a thresholding procedure is applied, in order to generate the binary image. Finally, all remaining gaps are filled.

3. EXPERIMENTAL RESULTS

We performed several experiments using the SIPaKMeD database [18]. The dataset contains 5 types of cervical cells with a total of 4049 images as shown in Table 1. We included a data augmentation scheme and each cell image is rotated in a clockwise manner in four directions, resulting in a total of 16,196 images. Finally, the dataset is divided into a training set, a validation set, and a test set, with each set corresponding to 80%, 10%, and 10% of the entire dataset respectively. Indicative results are depicted in Fig. 4

The Pytorch framework is used to construct the Nucleus-GAN and the Cell-GAN. The Adam optimizer is applied to train both networks with momentum of 0.5. The number of training iterations for the Nucleus-GAN and the Cell-GAN is 196 and 137 respectively, and the learning rate for both networks is 5×10^{-4}. For the L_2 loss function, both α and β are assigned a value of 1. The joint loss function sets γ_{smi} to 0.1 and γ_{adv} to 0.9, encouraging a better overall structure of the image.

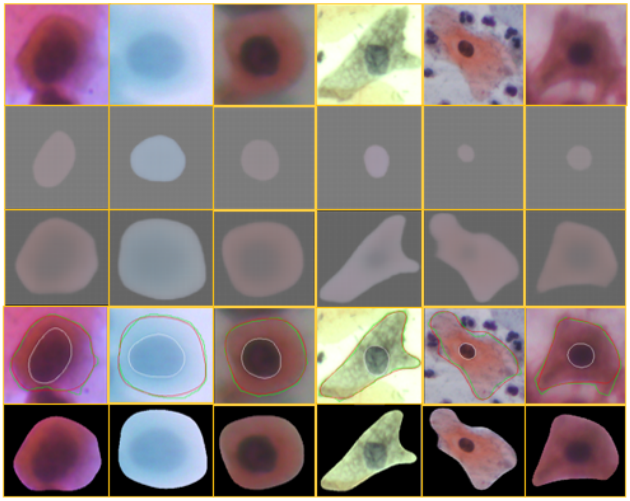

Fig. 4: Segmentation Results. The first row includes original cell images. In the second and third row the extracted nuclei and cells areas are depicted. The fourth row contains the obtained boundaries with the ground truth on the initial images (green:ground truth, red:cell, white:nucleus). In the fifth row, the final segmentation of each cell is depicted.

For efficient training, mini-batches consisting of 64 images were used. In order to avoid overfitting, the output of each layer network is normalized, and a dropout with a probability of 0.5 is added to the generator. Finally, to prevent mode collapse and ensure stable training, the technique of label smoothing is implemented.

To evaluate the segmentation accuracy of our method, we compared it with the method proposed in [10]. We used the same parameters, however the input in our Cell-GAN is different, since we used as guide factor the output of the Nucleus-GAN. This improves the performance of the cell segmentation method and at the same time, it provides a reliable nucleus boundary in each cell image. Furthermore, our trained models do not require the existence of previously annotated nucleus images as guide factors, for the segmentation of the cell. Thus, from a single image, both the nucleus and the cell boundaries are extracted.

Metrics including True Positive Rate (TPR), False Positive Rate (FPR), Dice Coefficient (DC) and Intersection over Union IoU) are calculated, in order to evaluate the efficiency of our method. Table 2 contains the results of these metrics for the Original Cell-GAN [10] and our method that combines the Nucleus-GAN and the proposed improved Cell-GAN. As we can observed, both methods have good performances as they have high TPR, DC and IoU and low FPR. Nonetheless, it is noteworthy that our method has higher TPR, DC and IoU met-

Cell Category	#Images	Augmentation
Superficial-Intermediate	831	3324
Parabasal	787	3148
Koilocytotic	825	3300
Dysketarotic	813	3252
Metaplastic	793	3172

Table 1: Dataset Composition

Method	TPR ↑	FPR ↓	DC ↑	IoU ↑
Original Cell-GAN [10]	0.946	0.031	0.957	0.919
Proposed Cell-GAN	0.953	0.037	0.958	0.921
Proposed Nucleus-GAN	0.970	0.012	0.938	0.895

Table 2: Comparative results of the different methods.

rics obtained by the Cell-GAN. This indicates that our method exhibits better performance and the segmentation of the nucleus obtained from the Nucleus-GAN, enhances the identification and generalization ability of the Cell-GAN. Furthermore, the Nucleus-GAN can identify the nuclei boundaries with high accuracy, as it presents remarkable performance regarding to all metrics.

4. CONCLUSIONS

In this paper, we proposed an improved version of the Cell-GAN which is based on guide factors obtained by the accurate determination of the nucleus in each image. This is feasible with the implementation of the Nucleus-GAN, which aims to produce nucleus images that approximate the original nucleus in the image, by evaluating their integrity via the probability distribution of the nucleus morphology. The generated images of this network serve as indicators of cell positioning within the initial image, thereby providing reliable input in Cell-GAN for cell identification.

The combination of these two networks provides a model that is able to learn the distinct characteristics of the cell and identify it among the noisy background interference that is present in many cell images. Additionally, the nucleus and the cell boundaries are extracted at the same time, providing a complete segmentation of the structural parts of the cell in each image. The experimental results indicate that our method enhances the Cell-GAN performance, providing accurate cell segmentation in an efficient manner.

5. ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan V GPU used for this research.

6. REFERENCES

[1] Hady Ahmady Phoulady, Dmitry B. Goldgof, Lawrence O. Hall, and Peter R. Mouton, "A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 201–204.

[2] Zhi Lu, Gustavo Carneiro, and Andrew P. Bradley, "An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells," *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1261–1272, 2015.

[3] Marina E. Plissiti, Christophoros Nikou, and Antonia Charchanti, "Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 2, pp. 233–241, 2011.

[4] M. E. Plissiti and C. Nikou, "Overlapping cell nuclei segmentation using a spatially adaptive active physical model," *IEEE Transactions on Image Processing*, vol. 21, no. 22, pp. 4568–4580, 2012.

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, Eds., Cham, 2015, pp. 234–241, Springer International Publishing.

[6] Youyi Song, Ling Zhang, Siping Chen, Dong Ni, Baopu Li, Yongjin Zhou, Baiying Lei, and Tianfu Wang, "A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei," *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2903–2906, 2014.

[7] Tianyi Zhao and Zhaozheng Yin, "Pyramid-based fully convolutional networks for cell segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, Eds., Cham, 2018, pp. 677–685, Springer International Publishing.

[8] Weikang Wang, David A. Taft, Yi-Jiun Chen, Jingyu Zhang, Callen T. Wallace, Min Xu, Simon C. Watkins, and Jianhua Xing, "Learn to segment single cells with deep distance estimator and deep cell detector," 2019.

[9] Filip Lux and Petr Matula, "Cell segmentation by combining marker-controlled watershed and deep learning," 2020.

[10] Jinjie Huang, Guihua Yang, Biao Li, Yongjun He, and Yani Liang, "Segmentation of cervical cell images based on generative adversarial networks," *IEEE Access*, vol. 9, pp. 115415–115428, 2021.

[11] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial networks," 2014.

[12] Tao Wang, Jinjie Huang, Dequan Zheng, and Yongjun He, "Nucleus segmentation of cervical cytology images based on depth information," *IEEE Access*, vol. 8, pp. 75846–75859, 2020.

[13] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning*, New York, NY, USA, 2008, ICML '08, p. 1096–1103, Association for Computing Machinery.

[14] Tianjun Xiao, Yichong Xu, Kuiyuan Yang, Jiaxing Zhang, Yuxin Peng, and Zheng Zhang, "The application of two-level attention models in deep convolutional neural network for fine-grained image classification," 2014.

[15] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2016.

[16] Li Xu, Jimmy SJ Ren, Ce Liu, and Jiaya Jia, "Deep convolutional neural network for image deconvolution," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, Eds. 2014, vol. 27, Curran Associates, Inc.

[17] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi, "Photo-realistic single image super-resolution using a generative adversarial network," 2017.

[18] Marina E. Plissiti, P. Dimitrakopoulos, G. Sfikas, Christophoros Nikou, O. Krikoni, and A. Charchanti, "SIPaKMeD: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 3144–3148.