

# Generic Isolated Cell Image Generator

Marin Scalbert,\* Florent Couzinie-Devy, Riadh Fezzani

Department of Research & Development, VitaDX, Paris, France

Received 21 June 2019; Revised 30 August 2019; Accepted 10 September 2019

Additional Supporting Information may be found in the online version of this article.

\*Correspondence to: Marin Scalbert, Research & Development, VitaDX, Rennes, France. Email: m.scalbert@vitadx.com

Published online 8 October 2019 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.23899

© 2019 The Authors. *Cytometry Part A* published by Wiley Periodicals, Inc. on behalf of International Society for Advancement of Cytometry.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

## • Abstract

Building automated cancer screening systems based on image analysis is currently a hot topic in computer vision and medical imaging community. One of the biggest challenges of such systems, especially those using state-of-the-art deep learning techniques, is that they usually require a large amount of training data to be accurate. However, in the medical field, the confidentiality of the data and the need for medical expertise to label them significantly reduce the amount of training data available. A common practice to overcome this problem is to apply data set augmentation techniques to artificially increase the size of the training data set. Classical data set augmentation methods such as geometrical or color transformations are efficient but still produce a limited amount of new data. Hence, there has been interest in data set augmentation methods using generative models able to synthesize a wider variety of new data. VitaDX is actually developing an automated bladder cancer screening system based on the analysis of cell images contained in urinary cytology digital slides. Currently, the number of available labeled cell images is limited and therefore exploitation of the full potential of deep learning techniques is not possible. In an attempt to increase the number of labeled cell images, a new generic generator for 2D cell images has been developed and is described in this article. This framework combines previous works on cell image generation and a recent style transfer method referred to as doodle-style transfer in this article. To the best of our knowledge, we are the first to use a doodle-style transfer method for synthetic cell image generation. This framework is quite modular and could be applied to other cell image generation problems. A statistical evaluation has shown that features of real and synthetic cell images followed roughly the same distribution. Finally, the realism of the synthetic cell images has been assessed through a visual evaluation performed with the help of medical experts. © 2019 The Authors. *Cytometry Part A* published by Wiley Periodicals, Inc. on behalf of International Society for Advancement of Cytometry.

## • Key terms

bladder cancer; urinary cytology; bright-field microscopy; synthetic cell images; deep learning; style transfer

Research for building automated cancer screening systems based on image analysis has been of increasing interest in medical imaging and computer vision community (1–3).

Before providing a final diagnosis, these systems usually need to solve preliminary tasks such as detection, segmentation, and classification. Deep learning has set the new state-of-the-art algorithms for such computer vision tasks. The main drawback of these algorithms is that they usually require large training data sets to outperform classical machine learning algorithms.

In the medical field, the acquisition of large data sets can be difficult due to data confidentiality constraints. Furthermore, the amount of labeled data might be limited by financial restrictions when medical expertise is required for the labeling process.

For training deep learning models in computer vision tasks, it is a common practice to apply data set augmentation techniques in order to artificially increase the size of the data set. Basic data set augmentation methods such as geometrical or color transformations are efficient but still are limited in the variety of the new images they can produce. Hence, the recent trend of data set augmentation methods is toward the use of generative models that can generate a wider diversity of new images (4–8).

Several previous studies have built frameworks for synthetic cell image generation for cytology slides, usually specialized for specific illumination technique (bright field, fluorescence) and data format (2D images, 3D images, and 2D/3D images + time).

For 2D/3D cell image generation, these works range from isolated cell image generation (9–11), grouped cells image generation (12) to whole cellular population image generation (13,14).

In these studies, the most common strategy is to combine a biological shape model to describe and generate the cell components morphology with textures models describing cell components textures, lighting conditions, and signal transmission of the optical system (9,12,14). Aberrations at various scales are usually considered, from local deformations (such as optical blur or sensors noise) to whole-slide aberrations (such as uneven illumination).

Different approaches for biological shape modeling exist, such as differential geometric based methods (10), shape initialization combined with random deformations model (12), or shapes generation by sampling and inverting invertible shape descriptors (14).

Regarding texture modeling, nonparametric and procedural methods are more often used. For example, Ref. (12) uses a stochastic method with Perlin noise to generate new cell texture samples, whereas Ref. (14) uses respectively patch-based texture synthesis and a deformations model to generate nucleus and cytoplasm textures.

More recently, Ref. (11) has proposed a different kind of approaches for fluorescent cell image generation based on generative adversarial networks (GANs). Compared to the methods sequentially generating cell morphology, cell textures, light conditions, and signal transmission of the optical system, GANs are able to jointly generate shape and textures to obtain synthetic cell images similar to real cell images. Nevertheless, GANs are hard to train (unstable training, unbalance between discriminator and generator, etc.) and usually require a large amount of data to produce realistic synthetic images. This makes these approaches not practical for cells generation.

Shape models combining a shape initialization with random deformations might not be representative of the real deformations undergone by the cell. Therefore, considering this shape modeling could lead to unrealistic generated cell morphology.

To generate a new cell morphology, represented by a segmentation mask with nucleus and cytoplasm shapes, the framework uses an approach in the same spirit of Ref. (14). First, the method consists in approximating a joint probability density function of invertible nucleus and cytoplasm shape descriptors describing statistically the whole cell morphology. Then, from the approximated joint probability density function, new nucleus and cytoplasm shape descriptors are sampled and inverted to generate a new cell segmentation mask with dependent nucleus and cytoplasm shapes.

Concerning textures modeling, Perlin-noise-based textures synthesis seems to be more adapted for simple and homogeneous textures synthesis. Patch-based textures synthesis requires

a large set of patches from the same cell to reproduce realistic cell textures which is not practical for small cell images.

To transfer real cell textures onto a synthetic cell segmentation mask, the style transfer method of Ref. (15) has been applied. This method, referred to as doodle-style transfer in this work, allows to transfer different textures from real segmented cell examples to the corresponding parts of a generated synthetic cell segmentation mask. This approach does not need an explicit modeling step to mimic local aberrations of the optical system as the doodle-style transfer method transfers real cell textures that already contain this visual information.

For a given cell class, the proposed framework consists in applying the cell segmentation masks generation procedure to obtain a new cell segmentation mask containing morphological specificities of the cell class. Then, textures of a real example cell from this class are transferred onto the generated segmentation mask via the doodle-style transfer method. With this procedure, the presented framework is able to generate synthetic cell images for any cell class with any illumination condition.

First, the data set used for synthetic cell image generation is described. Second, this article introduces the procedures to generate synthetic cell segmentation masks and how to transfer real cell textures onto generated cell segmentation masks. Finally, to assess the realism of the generated cell images using the developed framework and their similarity with real cell images, both statistical and visual evaluations have been conducted on a set of generated cell images.

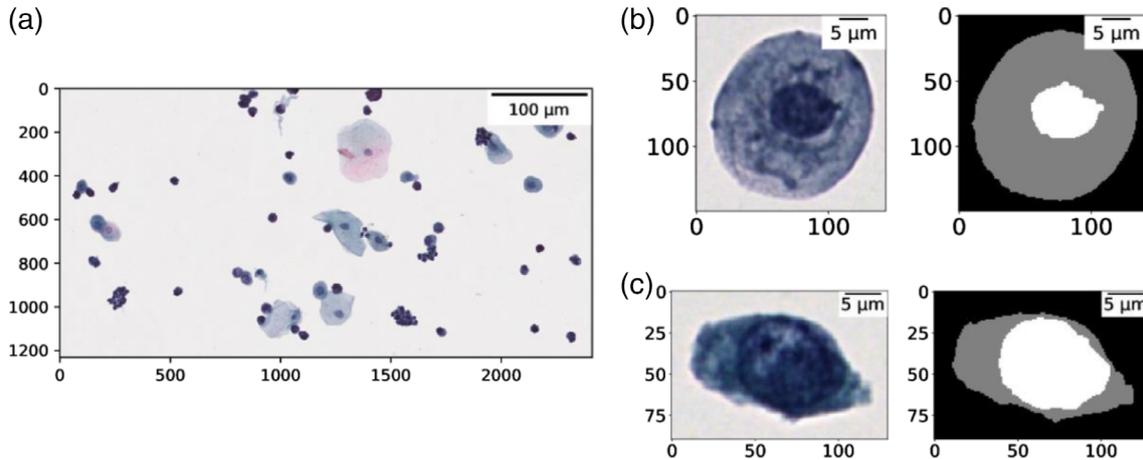
## MATERIAL AND METHODS

### Data Set Description

This work is made in the context of developing an automated bladder cancer diagnosis system based on the analysis of urinary cytology digital slides. Given a digital slide, the diagnosis produced by the screening system depends mainly on the detection of cells coming from the urothelium called urothelial cells and on the distinction between healthy and atypical ones. A particular focus will be put on this class of cells in this article.

The exploited urinary cytology digital slides are produced via the following protocol: first, urine sample is obtained from a patient, filtered, and spread onto a glass slide. Then, the glass slide is colored with a Papanicolaou stain (16) by immersing it into several dyes baths. The Papanicolaou staining enhances contrast between nucleus and cytoplasm and accentuates morphological and textural differences between healthy and atypical urothelial cells. Finally, the physical slide is scanned with a digital slide scanner Hamamatsu NanoZoomer S60 (Hamamatsu Photonics, Massy, France) able to scan a single slide on several focal plans with 40× magnification corresponding to a 230 nm/pixel scale.

Once the digital slide is obtained, anatomical cytopathologists annotate different objects considering different labels: object type, position on the slide, cell status (healthy or atypical), etc. Several different biological objects are likely to appear on a digital slide such as urothelial cells, Malpighian cells,



**Figure 1.** Crop of urinary cytology digital slide (a), examples of one healthy urothelial cell (b), and one atypical urothelial cell (c) with their segmentation masks. [Color figure can be viewed at wileyonlinelibrary.com]

polymorphonuclear neutrophils, etc. Consequently, the anatomical cytopathologist annotations can be very diversified.

In this work, only labeled urothelial cell images have been considered. Additionally, to learn how to generate the shape of urothelial cells, some of the labeled urothelial cell images have been manually segmented. From this process, we have obtained a data set made of:

- 500 segmented healthy urothelial cell images with their corresponding segmentation masks.
- 500 segmented atypical urothelial cell images with their corresponding segmentation masks.

Some examples of healthy/atypical urothelial cells with their corresponding segmentation masks can be found in Figure 1.

The use of the urinary cytology digital slides was authorized by the *Agence Nationale de Sécurité Médicament et des Produits de Santé (ANSM)* and the *Comité de Protection des Personnes* as part of the clinical trial VisioCyt1.

**Cell Segmentation Masks Generation**

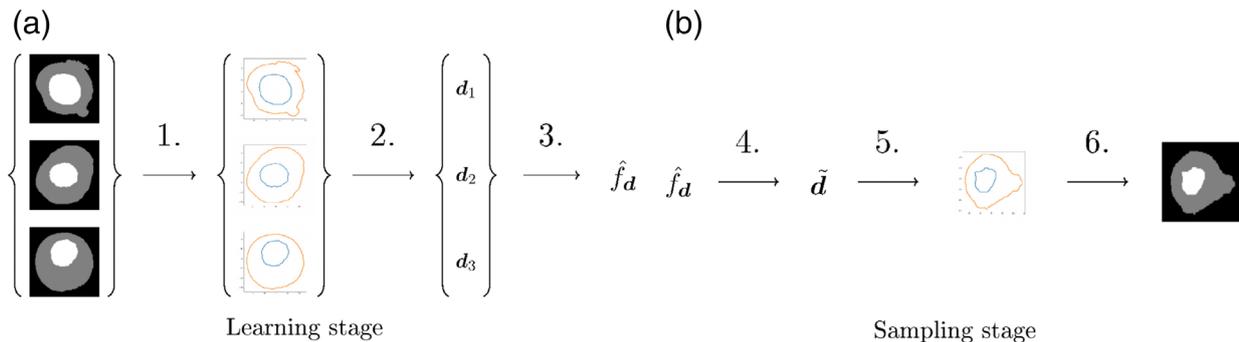
In the nucleus shape modeling proposed by Ref. (14), nucleus segmentation masks are converted into discretized contours by

sampling points evenly spaced on the masks boundary. From each nucleus discretized contour, Fourier shape descriptors are computed to summarize the nucleus shape information. This statistical population of nucleus shape descriptors is used to approximate the probability density function of training nucleus shape descriptors  $f_d$  with a probability density function  $\hat{f}_d$ . Finally, new nucleus shape descriptors are drawn from  $\hat{f}_d$  and inverted to get nucleus contours and thus new nucleus segmentation masks.

The proposed cell segmentation masks generation method follows the same idea as Ref. (14), but it generates correlated nucleus and cytoplasm contours with a realistic nucleus position into the cell. To achieve this, a joint probability density function of the nucleus and cytoplasm elliptical Fourier shape descriptors (17) is approximated. Figure 2 summarizes the whole process for cell segmentation masks generation proposed in this article.

The cell segmentation masks generation method can be decomposed into a learning stage:

1. Given a cell class, for each training cell segmentation mask, a pair of nucleus and cytoplasm discretized contours is retrieved with the algorithm marching squares (18).



**Figure 2.** Summary diagram illustrating the process for cell segmentation masks generation. During the learning stage (a), training nucleus and cytoplasm shape descriptors are computed from the training cell segmentation masks and the joint probability density function of these descriptors is approximated with a density  $\hat{f}_d$ . During the sampling stage (b),  $\hat{f}_d$  is used to sample new shape descriptors  $\tilde{d}$  that are inverted to get new cell segmentations masks. [Color figure can be viewed at wileyonlinelibrary.com]

2. For each pair of nucleus and cytoplasm discretized contours, pair of elliptical Fourier shape descriptors is computed and concatenated. The concatenated nucleus and cytoplasm shape descriptors denoted  $\mathbf{d}$  describe a whole cell morphology.
3. The joint probability density function of the nucleus and cytoplasm shape descriptors  $f_{\mathbf{d}}$  is approximated with a Gaussian mixture model (GMM) whose probability density function is denoted by  $\hat{f}_{\mathbf{d}}$ .  $\hat{f}_{\mathbf{d}}$  describes statistically the whole cell morphology of the given class.

and a sampling stage:

4.  $\hat{f}_{\mathbf{d}}$  is used to sample new concatenated nucleus and cytoplasm shape descriptors denoted  $\tilde{\mathbf{d}}$ .
5. For each concatenated nucleus and cytoplasm shape descriptors  $\tilde{\mathbf{d}}$ , nucleus and cytoplasm shape descriptors are inverted separately with the truncated Fourier series expansion formula detailed in Supplementary Eq. (6) to obtain a new pair of nucleus and cytoplasm discretized contours.
6. Each new pair of nucleus and cytoplasm discretized contours is converted into a cell segmentation mask.

The computation and inversion principle of elliptical Fourier shape descriptors are fully detailed in Supplementary Section A.1.

### Texture Transfer onto a Synthetic Cell Segmentation Mask

#### Texture synthesis

Li and Wand (19) have proposed a texture synthesis method to synthesize a texture  $\mathbf{x}_t \in \mathbb{R}^{h_t \times w_t \times 3}$  visually similar to a reference texture  $\mathbf{x}_s \in \mathbb{R}^{h_s \times w_s \times 3}$  using the feature maps of  $\mathbf{x}_s$  and  $\mathbf{x}_t$  computed on a pretrained convolutional neural network  $\phi$ .

Their method assumes that any texture can be described by  $L$  sets of local statistics computed respectively on the feature maps of  $L$  different layers of  $\phi$ . Therefore, as the reference texture  $\mathbf{x}_s$  can be described by its  $L$  sets of local statistics, they propose to generate the texture  $\mathbf{x}_t$  by optimizing  $\mathbf{x}_t$  pixels so as to match the  $L$  sets of local statistics of  $\mathbf{x}_t$  with those of  $\mathbf{x}_s$ .

In their method, each set of local statistics is made of all possible fixed-size volume patches extracted from the feature maps at a specific layer of  $\phi$ . From now on, these patches will be referred as neural patches to avoid confusion with patches extracted directly on the image.

Then, to generate the new texture  $\mathbf{x}_t$ , the pixels of  $\mathbf{x}_t$  are randomly initialized and optimized so as to match the  $L$  sets of neural patches of  $\mathbf{x}_t$  with those of  $\mathbf{x}_s$ . The matching process between a set of neural patches of  $\mathbf{x}_t$  and a set of neural patches of  $\mathbf{x}_s$  is done through the minimization of a distance between the two sets. The distance between a set of neural patches of  $\mathbf{x}_t$  and a set of neural patches of  $\mathbf{x}_s$  corresponds to a sum of distances between each neural patch of  $\mathbf{x}_t$  with its nearest neural patch of  $\mathbf{x}_s$ .

The drawback of this method is that the use of local information with the neural patches of  $\mathbf{x}_s$  is not sufficient to capture global arrangement of the textures in  $\mathbf{x}_s$ . Consequently, the different textures of  $\mathbf{x}_s$  and their relative positions might be unintentionally mixed in the generated image  $\mathbf{x}_t$ .

To better control the spatial layout of the synthesized textures, Ref. (15) has extended the method of Ref. (19). This new method referred as doodle-style transfer will be presented in the next section.

#### Layout aware texture synthesis: Doodle-style transfer

The method proposed by Ref. (15) aims at synthesizing the image  $\mathbf{x}_t$  by transferring the different textures of the image  $\mathbf{x}_s$  delimited by its segmentation mask  $\mathbf{m}_s$  onto the corresponding parts of the segmentation mask of the image  $\mathbf{x}_t$  denoted  $\mathbf{m}_t$ . Therefore, in the method of Ref. (15), the inputs are the reference texture  $\mathbf{x}_s$ , its segmentation mask  $\mathbf{m}_s$ , and the segmentation mask  $\mathbf{m}_t$  of the image to be synthesized  $\mathbf{x}_t$ . The textures from each region of the source image will be transferred to the corresponding region of the target image. In our case, segmentation masks  $\mathbf{m}_s$  and  $\mathbf{m}_t$  are made of three regions: background, cytoplasm, and nucleus.

To control the layout of the synthesized textures, Ref. (15) introduces semantic information contained in the segmentation masks  $\mathbf{m}_s$  and  $\mathbf{m}_t$  into the neural patches of  $\mathbf{x}_s$  and  $\mathbf{x}_t$ . Including  $\mathbf{m}_s$  and  $\mathbf{m}_t$  provides information on the localization of the different textures in  $\mathbf{x}_s$  and where they have to be synthesized on the image  $\mathbf{x}_t$ .

To introduce semantic information into the neural patches of  $\mathbf{x}_s$  and  $\mathbf{x}_t$ , Ref. (15) proposes to concatenate the binary masks with the feature maps before extracting neural patches. The problem is that the deeper you go into the trained neural network  $\phi$ , the smaller the feature maps are and therefore they cannot be concatenated with the binary masks due to dimensions mismatch.

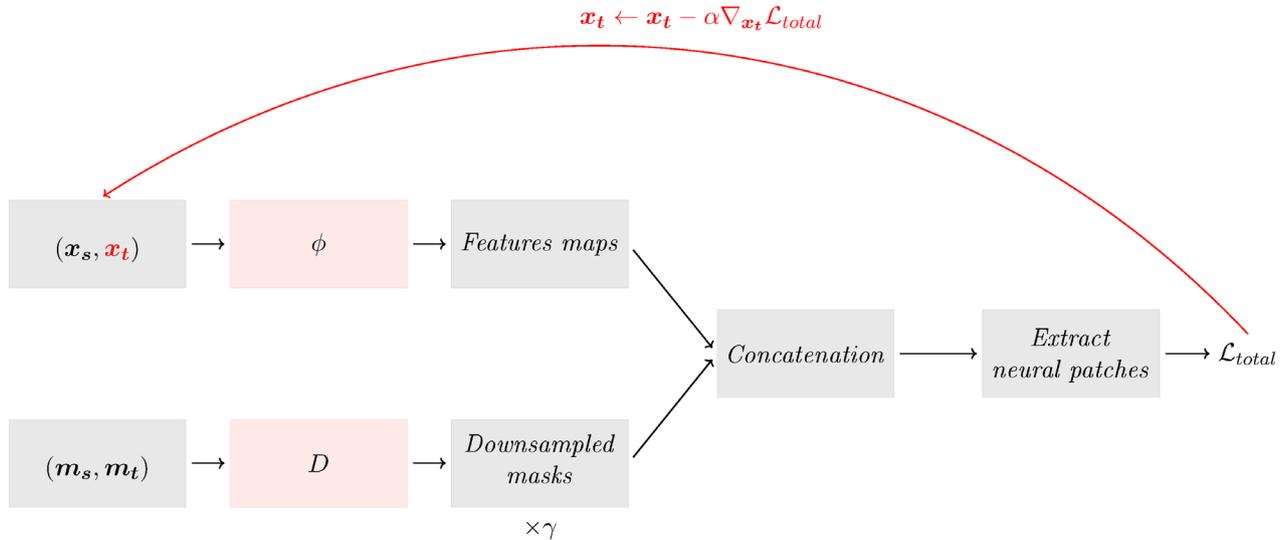
To bypass this problem, Ref. (15) suggests to down-sample the binary masks as the downsampling performed in  $\phi$ . The binary masks downsampling is performed through a model  $D$  made of successive average poolings. Thus, at a specific layer and for a given image, feature maps and downsampled masks can be concatenated. But before concatenating them and extracting neural patches, the downsampled masks are weighted by multiplying them by a hyperparameter  $\gamma$ . This hyperparameter allows balancing between spatial constraints and features similarity during neural patches matching of  $\mathbf{x}_s$  and  $\mathbf{x}_t$ .

Finally, like Ref. (19), once the sets of neural patches of  $\mathbf{x}_s$  have been computed,  $\mathbf{x}_t$  is randomly initialized and optimized to match its sets of neural patches with those of  $\mathbf{x}_s$ . The matching of the sets of neural patches of  $\mathbf{x}_t$  with the sets of neural patches of  $\mathbf{x}_s$  is achieved via the minimization of a cost function  $\mathcal{L}_{\text{total}}$  with respect to  $\mathbf{x}_t$ .

A summary diagram of the doodle-style transfer technique is presented in Figure 3. For further technical details on the loss function  $\mathcal{L}_{\text{total}}$ , the regularization term and the involved hyperparameters  $(\alpha_1, \dots, \alpha_L)$ ,  $\lambda_{\text{TV}}$ , the reader may refer to Supplementary Section A.2.

#### Full Pipeline for Synthetic Cell Image Generation

Cell textures and cell morphology are correlated. Two cells with significantly different morphologies are likely to have different textures. Hence, we should not transfer textures of



**Figure 3.** Summary diagram of the doodle-style transfer technique. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.com)]

any real cell image on any segmentation mask, otherwise transferred textures might not be compatible with the morphology of the target segmentation mask.

To ensure that the textures of the source cell image are plausible regarding the morphology of the target mask on which they are going to be synthesized, the target mask is chosen in order to be geometrically similar to the source segmentation masks.

For a given cell class, once the joint probability density function of the nucleus and cytoplasm shape descriptors  $\hat{f}_d$  was approximated as described in Section 2.2, a synthetic cell can be generated as follows:

1. select a source real image and its corresponding segmentation mask  $(\mathbf{x}_s, \mathbf{m}_s)$ ,
2. compute the shape descriptor  $\mathbf{d}_s$  of the source segmentation mask,
3. draw shape descriptors  $\mathbf{d}_t$  from  $\hat{f}_d$  until  $\|\mathbf{d}_s - \mathbf{d}_t\|_2$  is lower than a given tolerance  $\epsilon$ ,
4. invert the drawn shape descriptors to obtain the target synthetic segmentation mask  $\mathbf{m}_t$ , and
5. transfer the source texture  $\mathbf{x}_s$  on  $\mathbf{m}_t$  using the method described in Section 2.3.2.

## RESULTS

### Synthetic Urothelial Cell Segmentation Masks

For each cell class (healthy urothelial cell, atypical urothelial cell), the cell segmentation masks generation method described in Section 2.2 has been applied. The number of considered harmonics in the elliptical Fourier shape descriptors and the number of components of the GMM were empirically and respectively set to 10 and 30. In our experiments, the tolerance  $\epsilon$  of Step 4 of the synthetic cell image generation pipeline

described in Section 2.4 is set to  $10^{-1}$ . Examples of synthetic segmentation masks are reported in Figure 4.

We notice in Figure 4 that synthetic healthy urothelial segmentation masks present smooth and circular nucleus. Cytoplasm shapes are more irregular but, in practice, this is observed on real healthy urothelial cells. The synthetic healthy urothelial cell segmentation masks seem to present nuclear cytoplasmic ratios quite low which is a common characteristic of real healthy urothelial cells.

In Figure 4, nucleus shapes of synthetic atypical urothelial segmentation masks are more irregular than those observed on synthetic healthy urothelial cell segmentation masks. Furthermore, nuclear cytoplasmic ratios are significantly higher for synthetic atypical urothelial cell segmentation masks than synthetic healthy urothelial cell segmentation masks. In practice, this difference of nuclear cytoplasmic ratios is observed and is one of the Paris system criteria (20) used by anatomical cytopathologists to differentiate an atypical urothelial cell from a healthy urothelial cell.

### Synthetic Urothelial Cell Images

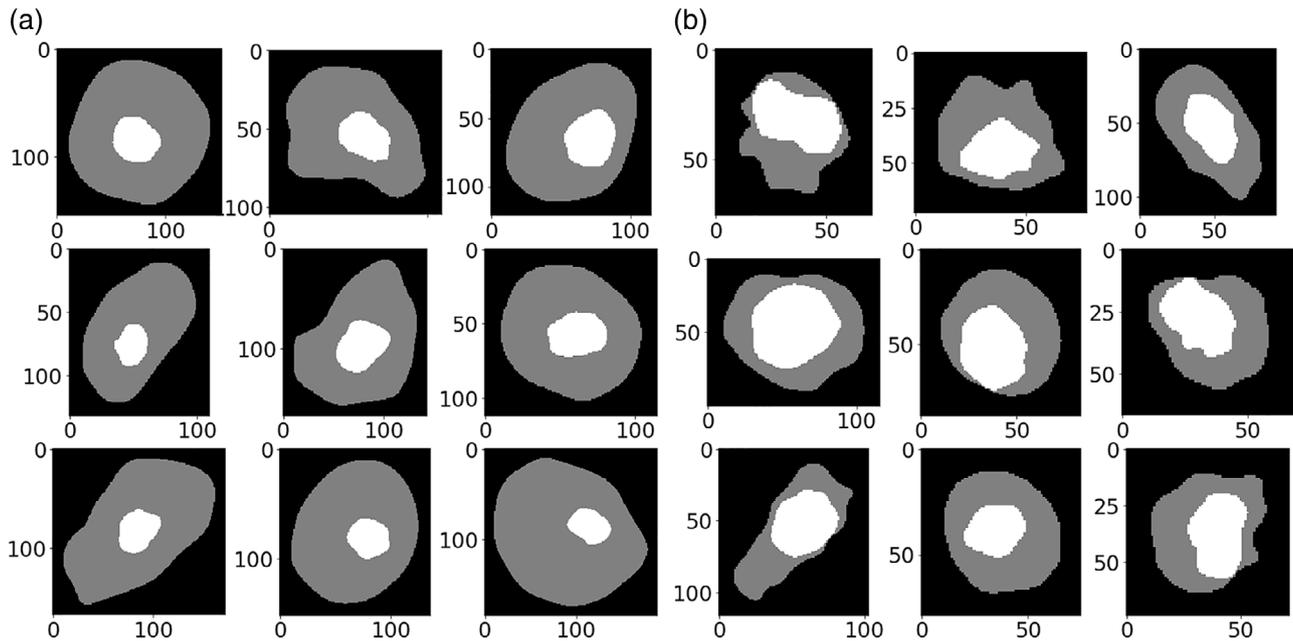
#### Hyperparameters

For each cell class (healthy urothelial cell, atypical urothelial cell), the doodle-style transfer method described in Section 2.3.2 was applied on a set of synthetic urothelial cell segmentation masks with the hyperparameters detailed in Figure 5.

#### Average generation time

All the experiments have been conducted with a Nvidia GeForce RTX 2080 Ti GPU (Nvidia, Santa Clara, CA). The generation time mainly depends on generated image size. For an average image size of 10,032 pixels, the measured average generation time is 24.63 s.

This is the result of the costly optimization used in the doodle-style transfer method. Indeed, each optimization step



**Figure 4.** Examples of generated healthy urothelial segmentation masks (a) and generated atypical urothelial segmentation masks (b) by sampling and inverting elliptical Fourier shape descriptors as described in Section 2.2.

to minimize the cost function  $\mathcal{L}_{\text{total}}$  with respect to  $\mathbf{x}_t$  implies to perform a forward and a backward pass into the trained neural network  $\phi$  (VGG19 (21) in our case) to compute respectively the feature maps and the gradients. These heavy computations due to the large size of the trained neural network  $\phi$  make the optimization time consuming.

A possible solution for this problem would be to use offline doodle-style transfer methods inspired from (22) work that separates learning and generation stages. These methods include an additional neural network  $\Phi_W$  of parameters  $W$  whose role is to generate the image  $\mathbf{x}_t$  given a triplet  $(\mathbf{x}_s, \mathbf{m}_s, \mathbf{m}_t)$ . In fact, during the learning stage, the total loss is no longer minimized with respect to  $\mathbf{x}_t$  but with respect to the weights  $W$ . Then, once  $\Phi_W$  is trained and given an input triplet  $(\mathbf{x}_s, \mathbf{m}_s, \mathbf{m}_t)$ , it can transfer the textures of  $\mathbf{x}_s$  onto  $\mathbf{m}_t$  only with inference.

Nonetheless, the architectures of  $\Phi_W$ , presented so far in the literature, can only transfer textures from a single real cell image  $\mathbf{x}_s$ . Consequently, several models must be trained to transfer textures from several real cell images  $\mathbf{x}_s$  which can be constraining.

$\phi$	VGG19
$(1, \dots, L)$	(relu_1_2, relu_2_1, relu_3_1)
$\gamma$	5000
$(\alpha_1, \dots, \alpha_L)$	$(\frac{1}{4}, \frac{2}{4}, \frac{1}{4})$
$\lambda_{TV}$	20
Optimisation	LBFGS
Maximal number of iterations	50

**Figure 5.** Hyperparameters of the doodle-style transfer technique used for both cell classes.

#### Examples of synthetic cell images

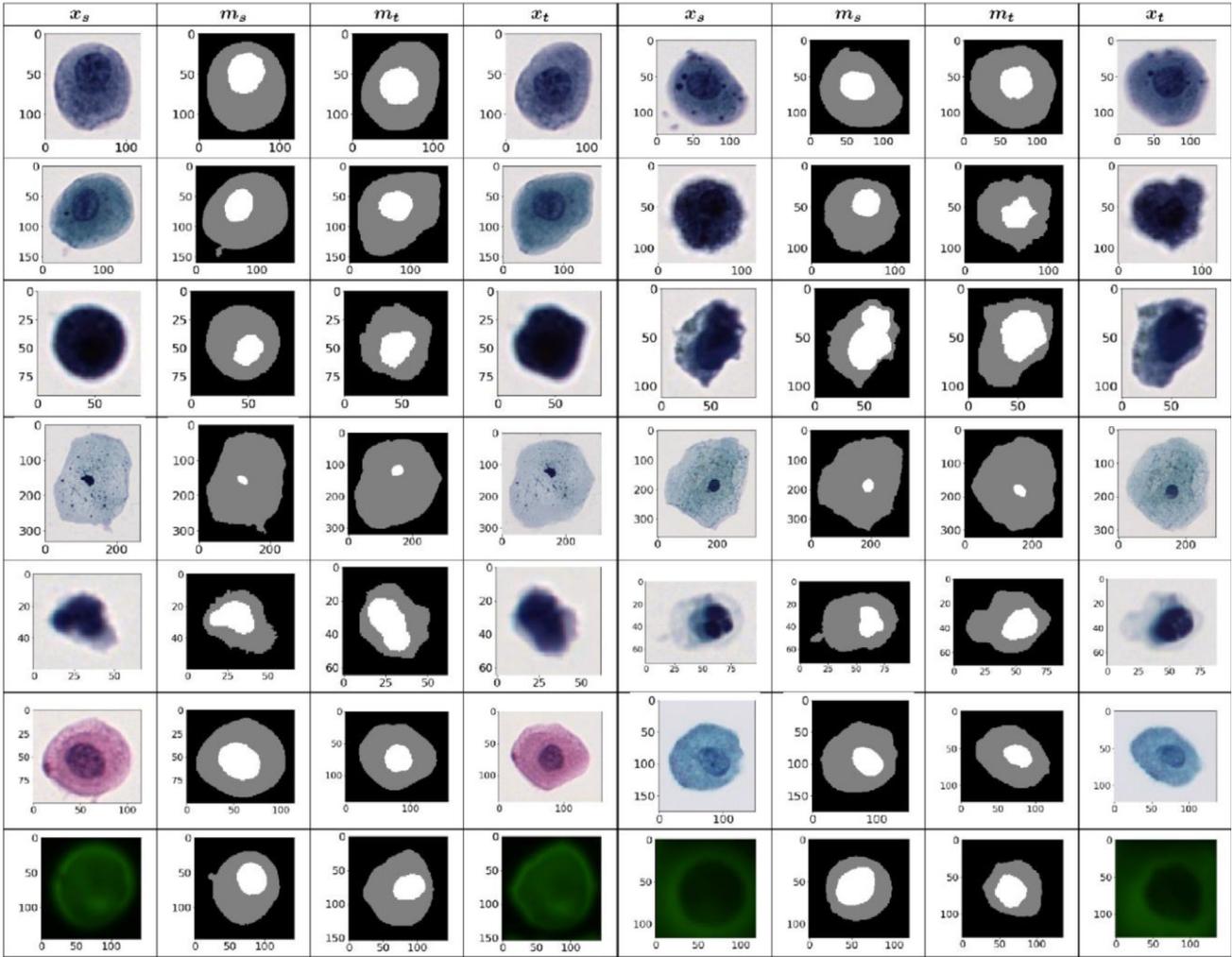
For each cell class (healthy urothelial cell, atypical urothelial cell), synthetic urothelial cell images have been generated by applying the doodle-style transfer method on triplets  $(\mathbf{x}_s, \mathbf{m}_s, \mathbf{m}_t)$  generated as described in Section 2.4. Some of them are presented on Figure 6 (1st to 3rd line).

In Figure 6, other types of generated objects are also presented. These generated objects include Malpighian cells (4th line), polymorphonuclear neutrophils (5th line), cells with different staining (Hematoxylin Eosin Safran [HES]) / Papanicolaou staining protocol (6th line) or cells scanned with different illumination technique such as FITC fluorescence (7th line).

From the synthetic urothelial cells of the three first lines, we can infer that the doodle-style transfer method is capable of transferring each texture of the real cell image  $\mathbf{x}_s$  delimited by its segmentation mask  $\mathbf{m}_s$  onto a synthetic cell segmentation mask  $\mathbf{m}_t$ . The textures in the synthetic cell images  $\mathbf{x}_t$  are not mixed as we are able to clearly distinguish each cell component.

By visually inspecting the synthetic urothelial cell images  $\mathbf{x}_t$ , the method produces quite realistic cell images. It ensures smooth colors and textures transitions between components boundaries. Indeed, for all synthesized images  $\mathbf{x}_t$ , the color transitions from the nucleus to the cytoplasm and from the cytoplasm to the background are comparable to those from real cell images  $\mathbf{x}_s$ .

From one real cell image  $\mathbf{x}_s$  to another, we clearly see that the varying depth of field and the optical system aberrations (blur, uneven illumination, noise, etc.) can cause the cell boundaries to be more or less sharp with color variations. For instance, the image  $\mathbf{x}_s$  in the (2nd line, 1st column) shows a sharp cell boundary with a uniform white light halo whereas



**Figure 6.** Examples of synthetic healthy urothelial cell (1st → 3rd line, 1st → 4th column) and synthetic atypical urothelial cells (1st → 3rd line, 5th → 8th column) generated from the presented framework. From left to right: real cell image  $x_s$ , real cell segmentation mask  $m_s$ , synthetic cell segmentation mask  $m_t$  generated with the method described in Section 2.2 and synthetic cell  $x_t$  generated with the doodle-style transfer method described in Section 2.3.2. From line 4th to line 7th, other types of objects are presented such as Malpighian cells (4th line), polynuclear neutrophils (5th line), cells with different staining (HES) / Papanicolaou staining protocol (6th line) and cells scanned with a different illumination technique (FITC fluorescence) (7th line). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

the image  $x_s$  in the (2nd line, 5th column) reveals a more blurred cell boundary with color variations ranging from light blue to pink. Usually, so as to make the synthetic cell image generation framework capable of handling these variations, varying light conditions and signal transmission of the used optical system are modeled (12,14). With the presented framework, as it can be noticed on the generated images  $x_t$ , these modeling steps are no longer required as, by transferring the textures of the real cell image  $x_s$  onto the generated segmentation mask  $m_t$ , the true light conditions and signal transmission of the optical system are directly replicated.

Furthermore, the framework is quite generic and flexible for synthetic cell image generation as long as we are dealing with 2D cell images that can be roughly segmented. Indeed, as it can be seen in Figure 6 from line 4th to 7th, the presented framework is able to generate other cellular objects

(Malpighian cells, polynuclear neutrophils), to handle different staining (HES)/staining protocols and illumination technique (FITC fluorescence). It can be noticed on the last line of the Figure 6 that the doodle-style transfer technique manages to produce realistic cell images even when the real and synthetic cell segmentation masks are non-well-defined due to the nature of the illumination technique.

**Synthetic Cell Images Evaluation Methods**

To assess the quality of the synthetic urothelial cell images, two types of image evaluation have been carried out. First, to test the similarity between real and synthetic urothelial cell images, multiple statistical equivalence tests have been carried out on features extracted on sets of synthetic urothelial cell images and the sets of real urothelial cell images used to generate them. This evaluation has been done so as to show that

features extracted on real and synthetic cell images followed roughly the same distributions. Then, a visual evaluation has been performed to test if the generated urothelial cell images could fool the eyes of medical experts.

### Statistical evaluation of synthetic urothelial cells

To assess the similarity between real and synthetic urothelial cell images with a statistical test, relevant features that describe well cell images should be used.

For this statistical evaluation, we have decided to choose some of expert features often used in cytological images (23,24) and local binary patterns features (25) that have been proven to be efficient as texture features for medical image analysis (26). These features can be split into three categories and are the following:

- Geometrical: nucleus and cell areas, nuclear cytoplasmic ratio, lengths of semi-major axis and semi-minor axis of ellipses fitted on the nucleus and cytoplasm contours, elongations of nucleus, and cytoplasm.
- Color: mean color on each image channel.
- Textures: median and standard deviation of local binary patterns computed with radius  $r \in \{3, 4\}$  and a number of points  $n_p = 8$ .

Regarding the choice of the statistical test, when one wants to prove that two samples of features come from different distributions, parametric/nonparametric two sample tests, such as  $t$ -test (27) or Kolmogorov–Smirnov test (28), are the relevant statistical tests to use. In this case, the objective is to prove that the null hypothesis, assuming that the distributions are the same, can be rejected.

In our case, as we are trying to prove the contrary (similarity between the real features distribution and the synthetic features distribution), hypotheses should be reversed in order to still control Type-I error. These statistical tests are referred to as equivalence tests. Furthermore, because each synthetic cell image can be paired with the real cell image that has been used for the synthetic cell morphology generation and the textures transfer, a paired equivalence statistical test is more appropriate.

For both cell classes (healthy urothelial, atypical urothelial), by visually inspecting the histograms of extracted features on a set of 500 real and 500 synthetic cell images, it seems reasonable to assume that each feature is distributed according a normal distribution with same variance.

From this assumption, we can carry out a paired two one-sided  $t$ -test (TOST) (29) for each feature to test whether or not they come from the same distribution. Because the number of performed TOST statistical tests will be as many as the number of distinct features, a Bonferonni correction on the  $P$ -values (30) will be applied.

In the TOST statistical test, we have two samples  $(y_1, \dots, y_n)$  and  $(z_1, \dots, z_n)$  that contain respectively independent and identically distributed random variables according  $\mathcal{N}(\mu_1, \sigma)$  and  $\mathcal{N}(\mu_2, \sigma)$  distributions. In our case, the first sample and the second sample will respectively correspond to real and synthetic features of the 500 real and the 500

synthetic urothelial cell images. The TOST hypotheses are defined as following:

$$\begin{cases} H_0 : \mu_1 - \mu_2 < -\varepsilon_2 \text{ and } \varepsilon_1 < \mu_1 - \mu_2 \\ H_1 : -\varepsilon_2 \leq \mu_1 - \mu_2 \leq \varepsilon_1 \end{cases}$$

For all the statistical tests, a reasonable margin around the mean of real features  $\mu_1$  is set to:

$$[-\varepsilon_1, \varepsilon_2] = \left[ q_{40\%}^{\mathcal{N}(0,1)} \sigma, q_{60\%}^{\mathcal{N}(0,1)} \sigma \right]$$

where  $q_x^{\mathcal{N}(0,1)}$  corresponds to the  $x$ th percentile of the standard normal distribution and  $\sigma$  is recomputed for each feature.

For each cell type (healthy urothelial cell, atypical urothelial cell), all the TOST statistical tests were carried out independently on the 16 distinct features computed on the set of 500 real cell images and the set of 500 synthetic cell images. The corrected  $P$ -values were all significant which prove to some extent that real and synthetic urothelial cells share common features. The  $P$ -values of each test are available on Supplementary Tables S1 and S2.

### Visual evaluation of synthetic urothelial cells

To test the realism of the synthetic urothelial cell images, six human evaluators have been asked to carry out a visual evaluation. Among these evaluators, three were biologists specifically accustomed to analyze urothelial cell images.

Each evaluator had to answer to 100 evaluations for each cell class (healthy urothelial cell, atypical urothelial cell), thus a total number of 200 evaluations. During each of these evaluations, the participant was seeing simultaneously four urothelial cell images (one synthetic image and three real images) displayed in a random order and had a maximum of 10 s to identify the synthesized image. In the case where the timer finished, a random answer was automatically selected.

In case of perfect synthetic images, a person answering randomly during the evaluations would have one chance out of four to select the synthetic image so an average identification rate of 25% on all the evaluations. Thus, the closer from 25% the average identification rate is the more realistic the synthetic images. All these evaluations have been made through a graphical user interface that can be seen in Supplementary Figure S3. After all the evaluations were completed, the results were aggregated by computing the average identification rates for each cell class:

- synthetic healthy urothelial cell: 37.6% and
- synthetic atypical urothelial cell: 38.4%.

Overall, both average identification rates are not so far from 25% which indicates that synthetic urothelial cell images are often hardly distinguishable from real urothelial cell images and can fool the eyes of medical experts.

### DISCUSSION

In this article, a new framework for synthetic bright-field urothelial cell image generation is presented.

For different cell classes, the framework can generate synthetic cell segmentation masks with correlated nucleus and cytoplasm shape. Furthermore, it can transfer the different textures of a real cell image delimited by its segmentation mask onto a generated segmentation mask. This new framework does not need to explicitly model the optical system aberrations. The local aberrations are replicated directly through the texture transfer and the effects of the global aberrations are not significant since only small areas are generated.

Applied to the urothelial cell image generation problem, the framework has managed to synthesize realistic urothelial cell segmentation masks. Morphological characteristics (shape, size, nuclear cytoplasmic ratio, and boundaries smoothness), usually observed and used to differentiate real healthy and real atypical urothelial cells, are visible in the synthetic segmentation masks. Regarding the textures of synthetic urothelial cells, the doodle-style transfer method ensures smooth textures and colors transitions between nucleus/cytoplasm and cytoplasm/background boundaries.

For the evaluation of this approach, a visual evaluation has shown that synthetic urothelial cell images are hardly distinguishable from real urothelial cell images. A statistical evaluation shows that geometrical and textural features are shared between real and synthetic cell images.

Doodle-style transfer requires a heavy optimization step at each image generation. This implies a relatively slow generation time. This could be reduced by applying offline doodle-style transfer methods that separate learning and generation stages.

The synthetic cell image generator presented in this article is demonstrated for unique isolated cells, but the proposed approach could be generalized to the generation of small groups of cells. The texture transfer step of the generation pipeline should not be modified. Only the mask generation step would have to be adapted to the synthetic conglomerate generation task.

A major future work will also be to evaluate the impact of data set augmentation using this approach on various classification and segmentation tasks for cytology images.

## AVAILABILITY

The Python code for the cell segmentation mask generator and the doodle-style transfer method described in this article are available at [https://gitlab.com/vitadx/articles/generic\\_isolated\\_cell\\_images\\_generator](https://gitlab.com/vitadx/articles/generic_isolated_cell_images_generator).

## ACKNOWLEDGMENT

The authors would like to thank the participants of the visual evaluation study and all people who have contributed or helped for this research project and this manuscript.

## CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

## LITERATURE CITED

- Lee Y, Hara T, Fujita H, Itoh S, Ishigaki T. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Trans Med Imaging* 2001;20:595–604.
- Demir C, Yener B. Automated cancer diagnosis based on histopathological images: A systematic survey. Rensselaer Polytechnic Institute, Technical Report; 2005.
- Cheng H-D, Shan J, Ju W, Guo Y, Zhang L. Automated breast cancer detection and classification using ultrasound images: A survey. *Pattern Recognit* 2010;43:299–317.
- Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*; 2017.
- Antoniou A, Storkey A, Edwards H. Data augmentation generative adversarial networks. *arXiv preprint arXiv:1711.04340*; 2017.
- Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018;321:321–331.
- Jackson PT, Atapour-Abarghouei A, Bonner S, Breckon T, Obara B. Style augmentation: Data augmentation via style randomization. *arXiv preprint arXiv:1809.05375*; 2018.
- Mikołajczyk A, Grochowski M. Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPHDW); 2018.
- Zhao T, Murphy RF. Automated learning of generative models for subcellular location: Building blocks for systems biology. *Cytometry A* 2007;71:978–990.
- Peng T, Wang W, Rohde GK, Murphy RF. Instance-based generative biological shape modeling. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009 (ISBI'09)*; 2009.
- Osokin A, Chessel A, Carazo Salas RE, Vaggi F. Gans for biological image synthesis. *Proceedings of the IEEE International Conference on Computer Vision*; 2017.
- Svoboda D, Kozubek M, Stejskal S. Generation of digital phantoms of cell nuclei and simulation of image formation in 3D image cytometry. *Cytometry A* 2009;75:494–509.
- Lehmussola A, Ruusuvaara P, Selinummi J, Huttunen H, Yli-Harja O. Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Trans Med Imaging* 2007;26:1010–1016.
- Malm P, Brun A, Bengtsson E. Simulation of bright-field microscopy images depicting pap-smear specimen. *Cytometry A* 2015;87:212–226.
- Champandard AJ. Semantic style transfer and turning two-bit doodles into fine artworks. *arXiv preprint arXiv:1603.01768*; 2016.
- Papanicolaou GN, Traut HF. *Diagnosis of Uterine Cancer by the Vaginal Smear*. Vol 46. New York: The Commonwealth Fund, 1943.
- Kuhl FP, Giardina CR. Elliptic Fourier features of a closed contour. *Comput Graphics Image Process* 1982;18:236–258.
- Lorenzen WE, Cline HE. Marching cubes: A high resolution 3D surface construction algorithm. *SIGGRAPH Comput Graph* 1987;21:163–169.
- Li C, Wand M. Combining markov random fields and convolutional neural networks for image synthesis. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016.
- Barkan, GA., Wojcik, EM., Nayar, R., Savic-Prince, S., Quek, ML., Kurtycz, DF., & Rosenthal, DL. (2016). The Paris system for reporting urinary cytology: the quest to develop a standardized terminology. *Acta Cytologica*, 60(3), 185–197.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*; 2014.
- Ulyanov D, Lebedev V, Vedaldi A, Lempitsky VS. Texture Networks: Feed-forward Synthesis of Textures and Stylized Images. *ICML*; 2016.
- Zhang L, Kong H, Ting Chin C, Liu S, Fan X, Wang T, Chen S. Automation-assisted cervical cancer screening in manual liquid-based cytology with hematoxylin and eosin staining. *Cytometry A* 2014;85:214–230.
- Duangate C, Uyyanonvara B, Koanantakul T. A review of image analysis and pattern classification techniques for automatic pap smear screening process. *International Conference on Embedded Systems and Intelligent Technology*; 2008.
- Ojala T, Pietikainen M, Harwood D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *Proceedings of 12th International Conference on Pattern Recognition*; 1994.
- Nanni L, Lumini A, Brahnam S. Local binary patterns variants as texture descriptors for medical image analysis. *Artif Intell Med* 2010;49:117–125.
- Student. The probable error of a mean. *Biometrika* 1908;6:1–25.
- Massey FJ Jr. The Kolmogorov-Smirnov test for goodness of fit. *J Am Stat Assoc* 1951;46:68–78.
- Schirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinetics Biopharm* 1987;15:657–680.
- Bonferroni C. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 1936;8:3–62.