Learning to segment cell nuclei in phase-contrast microscopy from fluorescence images for drug discovery

Hana Mertanová[a], Jan Kybic[a], Jaroslava Staňková[b], Petr Džubák[b], and Marián Hajdúch[b]

[a]Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic
[b]Institute of Molecular and Translational Medicine, Olomouc, Czech Republic

ABSTRACT

We describe a method for analyzing geometrical properties of cell nuclei from phase contrast microscopy images. This is useful in drug discovery for quantifying the effect of candidate chemical compounds, bypassing the need for fluorescence imaging. Fluorescence images are then only used for training our nuclei segmentation, avoiding the need for the time consuming expert annotations. Geometry based descriptors are calculated and aggregated and fed into a classifier to distinguish the different types of chemical treatments. The drug treatment can be distinguished from no treatment with accuracy better than 95% from fluorescence images and better than 77% from phase contrast images.

Keywords: segmentation, cell, nuclei, microscopy, image translation, drug discovery, deep learning, phase contrast microscopy

1. INTRODUCTION

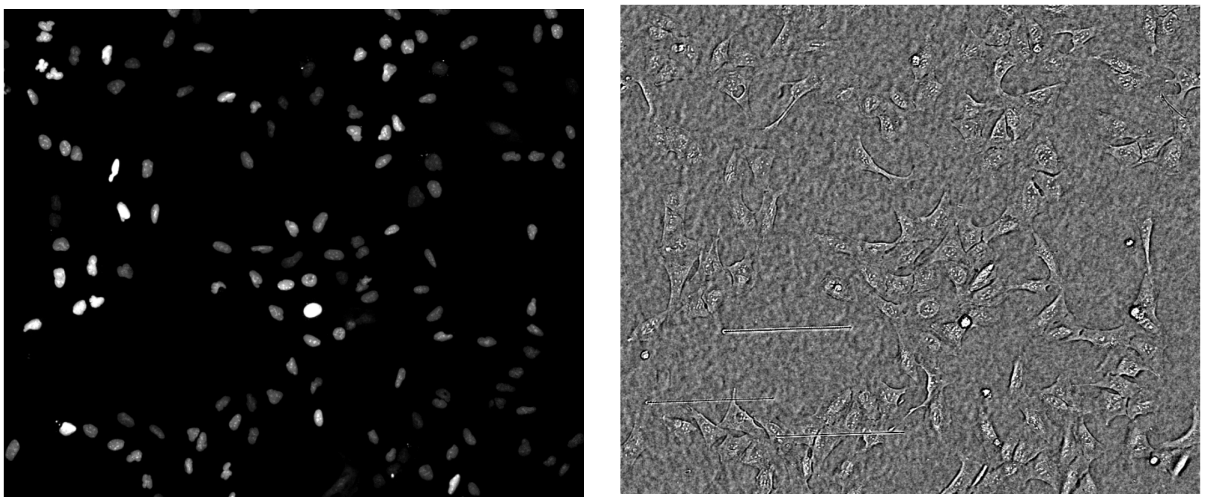
Current drug discovery procedures[1,2] require high-throughput quantification of the responses of the living cells by the chemical compounds being considered. Large number of combinations of the cell type and chemical compounds need to be evaluated quickly. Various properties of the cells can be measured, depending on the expected cellular response.

We shall focus on the analysis of cell nuclei, their number, size and shape. This is usually performed using fluorescence microscopy (see Fig. 1a) which provides very clear images of the individual nuclei that are easy to segment and evaluate automatically. On the other hand, the staining required by fluorescence microscopy increases the cost and processing time. Other disadvantages include bleaching and phototoxicity, which makes live-cell imaging difficult or impossible. The staining intensity is also not homogeneous and some cells may not be stained at all (see bottom part of Fig. 1a).

For these reasons, we have attempted to replace fluorescence images by phase contrast microscopy (Fig. 1b) which are much faster to acquire and do not damage the cells. However, these images are also much more difficult to analyze—they are noisy with cluttered background and contain imaging artifacts. The structure of the nuclei is much more complicated and the nucleus boundary is often not clearly delineated. To make the problem even more challenging, we do not have ground truth nuclei segmentations for the phase contrast images, so we cannot apply standard supervised segmentation methods, such as U-Net.[3] Instead, we need to learn to segment the nuclei in phase contrast images from the easy to segment fluorescence images, which is one of the contributions of this proof of concept work.

Fig. 2 shows in green the currently used legacy pipeline, which consists of a semi-automatic method for nuclei segmentation from fluorescence images, calculating a set of hand-crafted features, and performing a statistical hypothesis testing to see whether sets of features corresponding to chemical compounds being tested are statistically significantly different or not. Our proposed new image processing pipeline (in Fig. 2 blue) begins by a segmentation method for phase-contrast images (Section 3.2) based on the image to image translation using the *pix2pix* adversarial learning technique.[4] We continue by identifying individual nuclei, extracting shape-based descriptors (Section 3.3), aggregating them, and then attempting to distinguish the chemical compound being tested from these features by a classifier (Section 3.4). The advantages of the new approach are: (i) faster, cheaper, and less damaging image acquisition, (ii) no manual interaction, (iii) more general non-parametric statistical modeling.

Send correspondence to Jan Kybic, e-mail: kybic@fel.cvut.cz



(a) (b)

Figure 1. Example of a fluorescence microscopy image with stained nuclei *(a)* and a corresponding phase contrast microscopy image *(b)*.

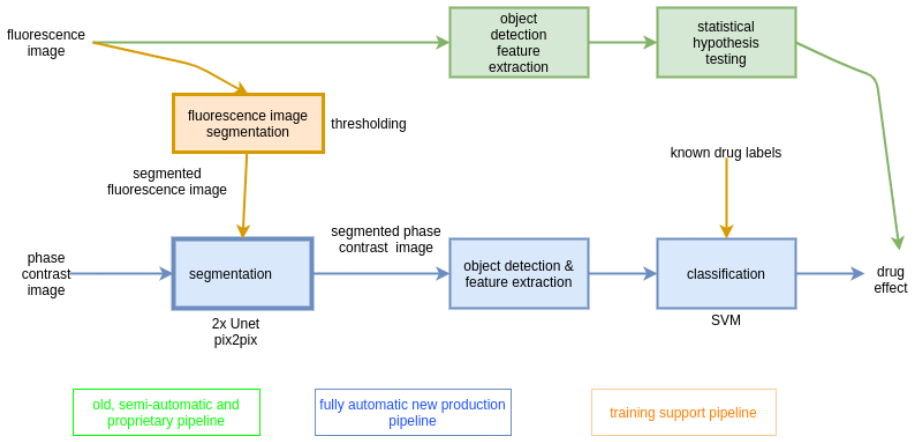
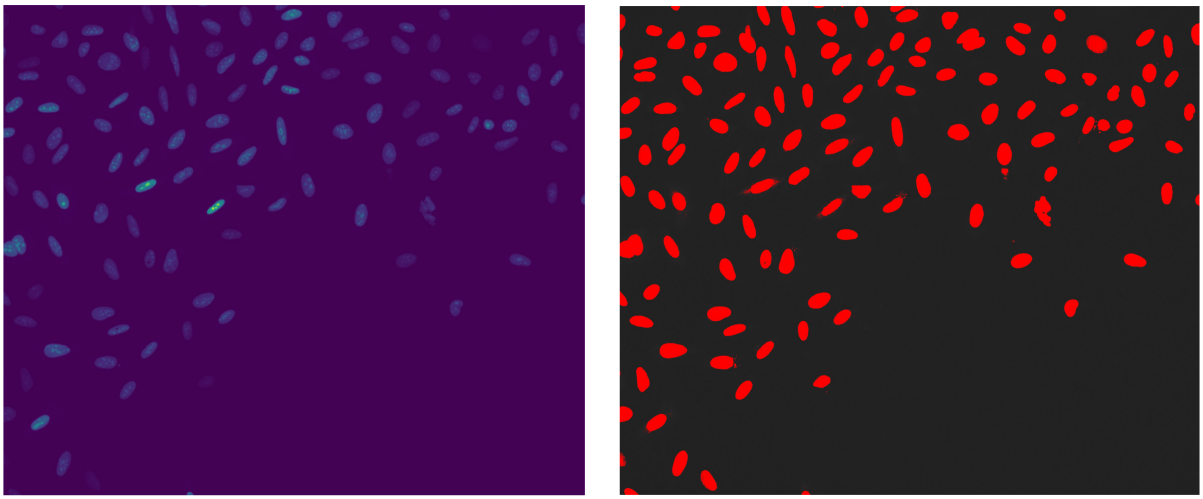


Figure 2. We replace a semi-automatic legacy pipeline *(green)* based on fluorescence images by a new fully automatic pipeline *(blue)* based on phase-contrast images. Fluorescence image segmentation and known drug labels *(orange)* are only used for training.



(a) (b)

Figure 3. Example of a fluorescence microscopy image with stained nuclei *(a)* and the corresponding segmentation using global thresholding *(b)*.

2. DATA

U-2 OS cells transducted by Vectalis for organelle labeling were seeded 24 hours before the experiment into a CellCarier 384 well plate, 1500 cells per well. Nuclei were stained by red fluorescent mCherry-NLS (nuclear localization signal). Images (2560×2160 pixels, 16 bits) were acquired using CV7000 microscope by Yokogawa at $20\times$ magnification for both phase-contrast and fluorescent imaging. The phase-contrast and fluorescent images are geometrically aligned, we observed geometric alignment error of at most a few pixels. Each well was scanned at 5 acquisition points, yielding 5 non-overlapping images for each modality. For dataset I, we used 44 wells without any treatments, i.e. $44 \cdot 5 = 220$ images for each modality. For dataset II, we added 3 different chemotherapeutic agents used in cancer treatment — a 0.5 μM solution of Topotecan, Daunorubicin and Etoposide — each applied to 8 wells. As controls, we have applied a 0.05% solution of DMSO to 8 wells and no treatment to 12 additional wells. Images were acquired before the treatment and then after 24 and 72 hours after the treatment. In total for dataset II, there are 44 wells with 5 images per well at 3 time points, yielding $44 \cdot 5 \cdot 4 = 660$ images for each of the two modalities.

3. METHODS

3.1 Nuclei segmentation in fluorescence images

For training, we take advantage of having pairs of corresponding phase contrast and fluorescence images. We consider the variation of intensity of the nuclei to be a nuisance parameter and we therefore chose to segment the nuclei from these image before using them as a reference in the subsequent step. Several methods were tested but it turned out that simple global thresholding was sufficient (Fig. 3).

3.2 Nuclei segmentation in phase contrast images

We formulate the task of nuclei segmentation in phase contrast images as an *image translation* task with the automatic segmentation from the fluorescence images (Section 3.1) considered as the target. We use the conditional adversarial network architecture *pix2pix*,[4] with a number of changes. It consists of two competing sub-networks: A *generator* G attempts to generate a segmentation $\hat{S}$ given an input phase contrast image I. A *discriminator* D is given an input phase contrast image I and a segmentation S and its task is to determine whether the segmentation comes from the training data (is "real", represented by 1) or if it was calculated by the generator (is "fake", represented by 0). In our implementation, the decoder provides one value from the interval $[0, 1]$ for each pixel, which is fed into an L_{GAN} loss function based on the ℓ_2 norm (instead of the original cross-entropy). The other part of the loss, denoted L_{L1}, facilitates training by encouraging the generator output $G(I)$ to be close to

the reference segmentation S. More formally, the optimal generator G^* minimizes the following combined loss function

$$G^* = \arg\min_G \max_D L_{\text{GAN}}(G, D) + \lambda L_{\text{L1}}(G) \tag{1}$$

$$\text{where} \quad L_{\text{GAN}}(G, D) = \mathrm{E}\left[\|D(I, S)\|_2\right] + \mathrm{E}\left[\|1 - D\big(I, G(I)\big)\|_2\right] \tag{2}$$

$$L_{\text{L1}}(G) = \mathrm{E}\left[\|S - G(I)\|_1\right], \tag{3}$$

the norms are taken over all pixels in the image and the expectations over the training images. As a generator, we used the U-Net architecture with four downsampling blocks (instead of 8), max-pooling and no dropout. The discriminator architecture used 4 blocks with 4×4 kernels and 64, 128, 256 and 512 channels, respectively.* We used $\lambda = 10$ but this value did not have a significant influence.

Because of the size of the input images, we divided them into 560×560 tiles with overlap of 400 pixels and processed them separately. The final segmentation is a mean of all overlapping tiles with spatial Gaussian weights.

Figs. 4,5 show that the nuclei are segmented reliably, even in high density situations, and that our proposed method produces less artifacts than both U-Net and original pix2pix. This is confirmed by numerical result where the median Dice loss on the test dataset was 0.36 for U-Net, 0.33 for original pix2pix, and 0.24 for our ℓ_1/ℓ_2 pix2pix. The segmentation works much better for the images 24h after treatment (Dice score 0.16) than 72h after treatment (Dice score 0.35).

3.3 Feature extraction

The segmentation is first cleaned by morphological closing, individual nuclei are found by connected component analysis, and very small components are eliminated. Touching nuclei are separated by the watershed method[5] — fortunately they are not very common in our data. For each object (nuclei), the following descriptors are calculated:

1. Area A in pixels.
2. Granularity $G = P/A$, where P is the number of boundary (perimeter) pixels.
3. Hu rotation invariants[6]
$$L_{23} = \eta_{20} + \eta_{02}, \quad \text{and} \quad L_{24} = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{4}$$
where η are the centered moments.
4. Major and minor approximation ellipse axis lengths $C_1 = 2\sqrt{\lambda_1}$, $C_2 = 2\sqrt{\lambda_2}$, where $\lambda_1 \geq \lambda_2$ are eigenvalues of the object covariance matrix.
5. Aspect ratio $R = C_1/C_2$.

We aggregate these features over all objects in the image by calculating their mean, median and standard deviation. Moreover, we calculate histograms with 10 bins for features A, P and R. The last feature is the total area $\sum A_i$. This yields a 52 element feature vector per image.

3.4 Classification

The extracted image feature vectors are used to differentiate between the 5 groups (Topotecan, Daunorubicin, Etoposide, DMSO, and no treatment). Feature selection on the training set (60% of available images) using correlation analysis found the following 15 most important features:

1. Mean, median and standard deviation of the nuclei area A.

*Implementation by Erik Linder-Norén, `https://github.com/eriklindernoren/PyTorch-GAN/blob/master/implementations/pix2pix`

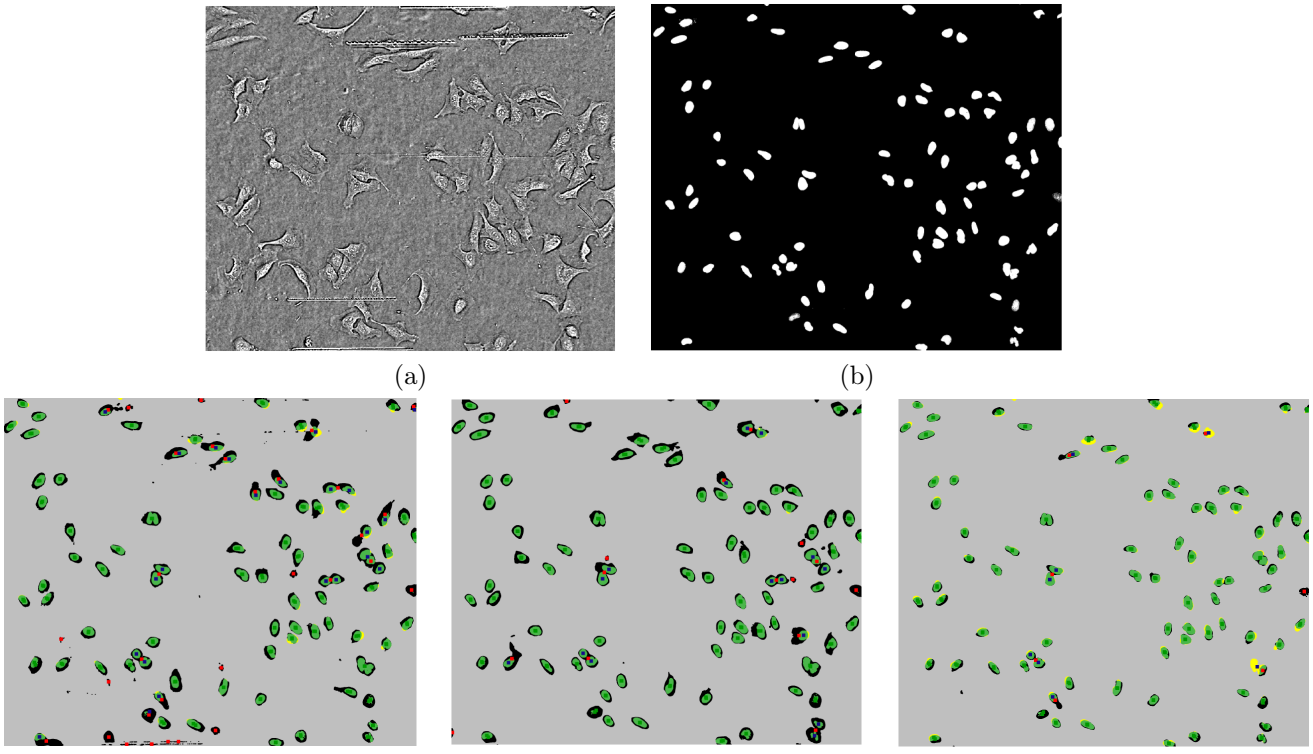


Figure 4. Comparison of the segmentation methods. Input phase contrast image *(a)* and the reference segmentation *(b)* extracted from a fluorescence image. Segmentation results of the U-Net *(c)*, original pix2pix with cross-entropy loss *(d)* and our pix2pix with the ℓ_2/ℓ_1 loss *(e)*. In the last row, green pixels indicates true positives, black false-positives, yellow false negatives and gray true negatives. Matched nuclei centers are marked with green dots, while false positives are red and false negatives blue.

2. Mean and median of the granularity G.
3. Mean, median and standard deviation of the ellipse lengths C_1, C_2.
4. Mean and median of the aspect ratio R.
5. Total area of the nuclei $\sum A_i$.
6. 8^{th} element of the histogram of the perimeter length P.

The pruned feature vector is used for an SVM classifier with an RBF kernel. The number of selected features and the parameters of the SVM classifier were found using cross-validation.

4. RESULTS

The pairwise classification accuracy was evaluated on the test set, consisting of the remaining 40% of available images. The evaluation was performed both with the segmentations extracted from the fluorescence images (Section 3.1) and the segmentations calculated from the phase contrast images (Section 3.2).

Table 1 shows that with a good segmentation, the accuracy is very high ($> 95\%$) for most class pairs, i.e. that the effect of the different treatment can be well distinguished from the nuclei shapes and sizes, with two exceptions. The inability to distinguish between DMSO and "no treatment" was expected, as DMSO is not toxic for the cells at this concentration. The inability to distinguish between Topotecan and Daunorubicin may be caused by a similar action mechanism but further investigation is necessary.

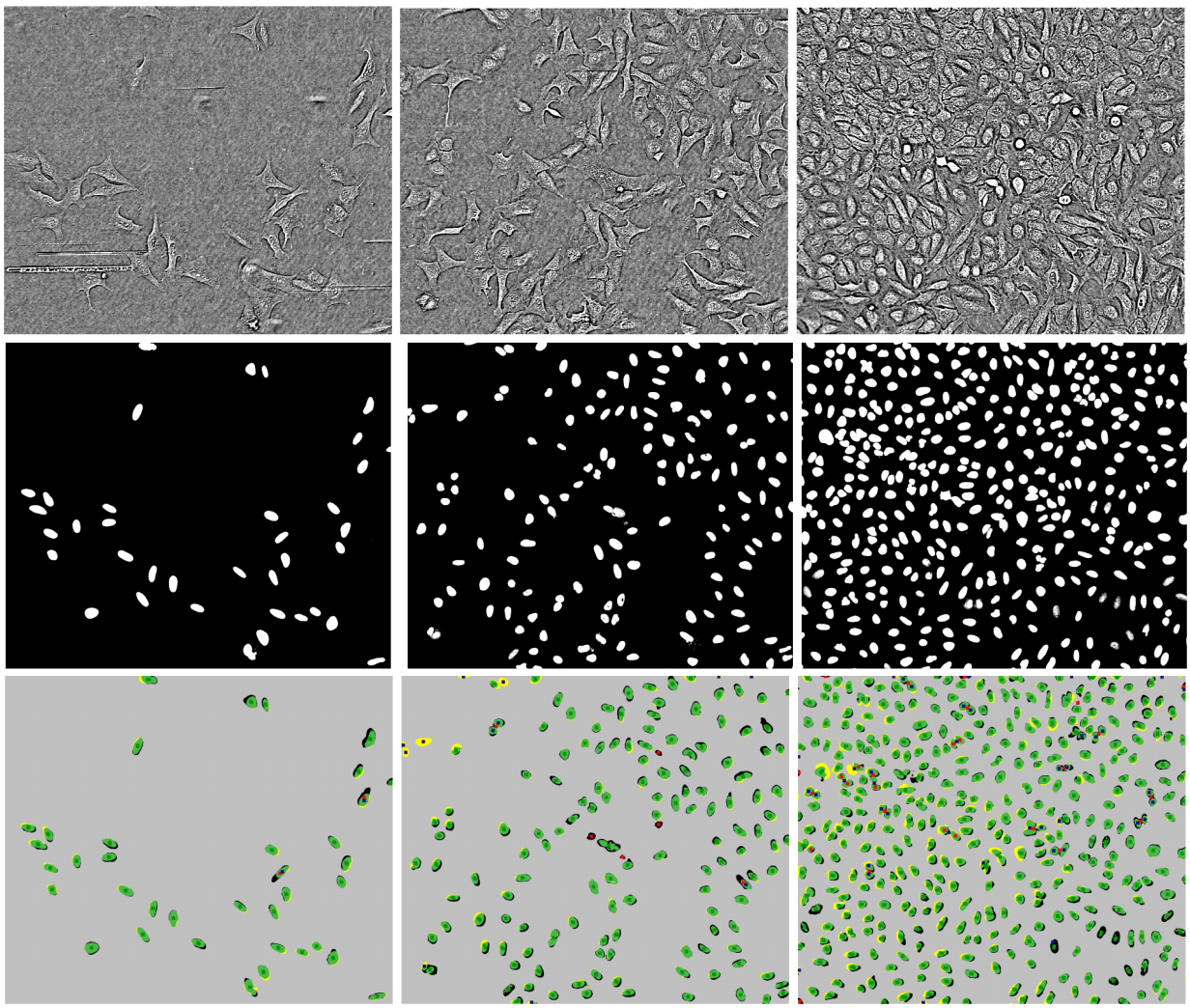


Figure 5. Input phase contrast images *(top row)*, reference segmentations extracted from the fluorescence image *(middle row)* and resulting nuclei segmentations derived from the phase-contrast images *(bottom row)*. The color code is the same as in Fig. 4.

Table 1. The classification accuracy for different treatment pairs on test data using the segmentation from either fluorescence or phase contrast images.

treatments			fluorescence	phase contrast
Topotecan	vs	Daunorubicin	0.52	0.41
Topotecan	vs	Etoposide	0.85	0.78
Topotecan	vs	DMSO	0.98	0.79
Topotecan	vs	no treatment	0.96	0.77
Daunorubicin	vs	Etoposide	0.95	0.79
Daunorubicin	vs	DMSO	1.00	0.91
Daunorubicin	vs	no treatment	1.00	0.95
Etoposide	vs	DMSO	0.97	0.72
Etoposide	vs	no treatment	0.98	0.77
DMSO	vs	no treatment	0.64	0.54

For phase contrast microscopy image segmentations, the accuracy remains sufficiently high to allow to detect the differences between the treatments. However, the drop in performance is important.

5. CONCLUSIONS

We have shown a method for segmenting cell nuclei from phase contrast microscopy images which does not need expert segmentation for training but can be trained from pairs of fluorescence and phase contrast microscopy images. Moreover, this segmentation works better than classical image segmentation techniques. We have integrated the segmentation into a complete image processing pipeline, which calculates aggregated geometrical descriptors for the cells in the image and uses this information to detect the effect of the drugs being tested. We found that it is possible to distinguish the treatment from no treatment very reliably (with accuracy better than 95%) using fluorescence images. Our method allows using more convenient phase-contrast images, although the accuracy than drops to $77 \sim 95\%$ (depending on the drug).

ACKNOWLEDGMENTS

The authors acknowledge the support of the OP VVV funded project "CZ.02.1.01/0.0/0.0/16_019/0000765 Research Center for Informatics".

REFERENCES

[1] Smith, A., "Screening for drug discovery: the leading question," *Nature* **418** (2002).

[2] Land, P., Yoew, K., Nichols, A., and Scheer, A., "Cellular imaging in drug discovery," *Nature Reviews Drug Discovery* (5), 343–356 (2006).

[3] Ronneberger, O., Fischer, P., and Brox, T., "U-Net: Convolutional networks for biomedical image segmentation," in [*MICCAI: Medical Image Computing and Computer-Assisted Intervention*], (May 2015).

[4] Isola, P., Zhu, J., Zhou, T., and Efros, A. A., "Image-to-image translation with conditional adversarial networks," in [*CVPR: IEEE Conference on Computer Vision and Pattern Recognition*], 5967–5976, IEEE Computer Society (2017).

[5] Shafarenko, L., Petrou, M., and Kittler, J., "Automatic watershed segmentation of randomly textured color images," *IEEE Transactions on Image Processing* **11**(6), 1530–1544 (1997).

[6] Hu, M. K., "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory* **8**(2), 179–187 (1962).