https://doi.org/10.2352/ISSN.2470-1173.2020.14.COIMG-341

2D Label Free Microscopy Imaging Analysis Using Machine Learning

Han Hu; School of Electrical and Computer Engineering; Purdue University; West Lafayette, Indiana, USA
Yang Lei; HP Labs, HP Inc; Palo Alto, California, USA
Daisy Xin; HP Labs, HP Inc; Palo Alto, California, USA
Viktor Shkolnikov; HP Labs, HP Inc; Palo Alto, California, USA
Steven Barcelo; HP Labs, HP Inc; Palo Alto, California, USA
Jan Allebach; School of Electrical and Computer Engineering; Purdue University; West Lafayette, Indiana, USA
Edward J. Delp; School of Electrical and Computer Engineering; Purdue University; West Lafayette, Indiana, USA

Abstract

Separation and isolation of living cells plays an important role in the fields of medicine and biology with label-free imaging often used for isolating cells. The analysis of label-free cell images has many challenges when examining the behavior of cells. This paper presents methods to analyze label-free cells. Many of the tools we describe are based on machine learning approaches. We also investigate ways of augmenting limited availability of training data. Our results demonstrate that our proposed methods are capable of successfully segmenting and classifying label-free cells.

Introduction

Separation and isolation of particles such as cells is important in medicine and biology. This can be challenging, especially in the case of rare cells that are present only as a small fraction of the total cell population [1, 2]. One example of this are circulating tumor cells (CTCs), which are metastasized cancer cells in the bloodstream [3]. Extracting CTCs from blood can be used for early cancer screening in a non-invasive way and can provide a mechanism for testing potential therapies on an individual basis.

Rare cells [1] are extremely low-abundance cells within a larger population of background cells. Rare cell isolation (RCI) devices [4] physically isolate and classify rare cells using microfluidics. The use of microfluidics has many advantages such as handling small sample volumes and multiplexing capabilities for high-throughput processing [1]. This makes microfluidics one of the best platforms to deal with the isolation and analysis of rare cells. A RCI device [4] is shown in Figure 1.

Here, we will briefly review how the RCI device works. First, cells are hydrodynamically focused into a narrow stream. The RCI device then uses a uniform dielectrophoretic force field to nudge the particles from streamlines and direct them into appropriate outlets [5, 6]. The cells experiencing positive dielectrophoresis are deflected into one outlet channel and the cells experiencing negative dielectrophoresis are deflected into the other. For simplicity, the above RCI microfluidics device will be referred to as "the chip" in this paper. Various microscopy imaging techniques have been used to observe label-free cells. These include phase-contrast (PC) microscopy and differential interference contrast (DIC) microscopy [7]. For the work we present here, phase contrast microscopy is used. In order to evaluate the performance of the chip, we need to quantify the precision and recall [1, 8, 9, 10] of images captured from the chip. The precision of the chip is the ratio between the number of correctly isolated cells at the exit of the chip and the number of all isolated cells. The recall of the chip is the ratio between the number of correctly isolated cells at the exit of the chip and the number of known targeted cells introduced at the inlet of the chip . The circulating tumor cells (CTC) have low concentration in the blood sample, ranging from 1-10 cells per mL in whole blood. If separated and collected, the CTCs can be further analyzed, for example, through genomic sequencing, and can provide valuable diagnostic insights. If some blood cells are identified as cancer cells, additional steps, including manual steps, can be used to remove them. Our goal is to collect as many cancer cells as possible. Therefore, recall of cancer cells is more important than precision. To evaluate the performance of the RCI device, we need to be able to identify and track cells as they move through the device. This requires that we segment and classify the cells.

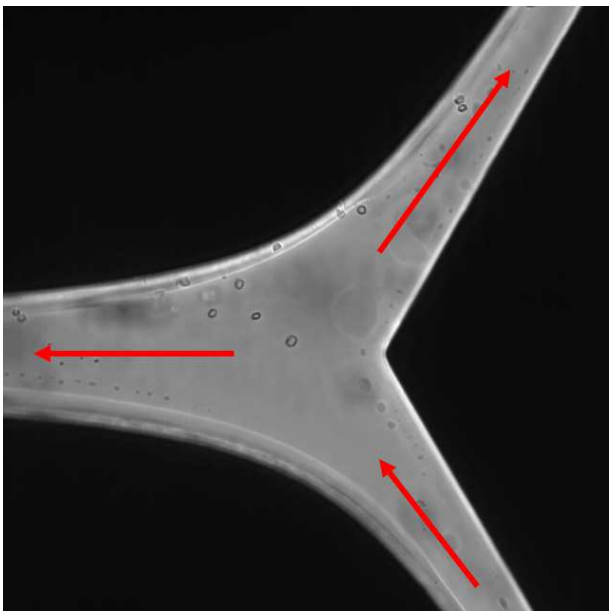


Figure 1. *A rare cell isolation(RCI) device, i.e., "the chip".*

Image segmentation plays a crucial role in cell image pro-

cessing and analysis. Many cell segmentation methods have been proposed [11, 12, 13, 14]. These methods range from relatively simple thresholding methods to more complicated statistical and machine learning approaches. Thresholding-based methods [11] assume that there is a relatively good separation in the distribution between the cells in the background. One widely used approach is Otsu's method [15] which uses a 2 class hypothesis testing approach to find the threshold by examining the histograms of "objects/cells" and "background." Connected component analysis is then used to select the largest image object. Active contour models [16] utilize deformable contours for edge detection and segmentation. These methods have had a good success in segmentation of cell images in diverse backgrounds and have been extended to 3D microscopy [17, 18, 19]. Watershed [20] models a grayscale cell image as a topographic map. This topographic map can be separated into adjacent flooded basins with watersheds lines dividing the basins. Watershed often suffers from over segmentation [21]. Post-processing, such as merging separate regions that belong to the same structure, is often used to reduce over-segmentation [12].

Machine learning approaches, particularly deep learning, have proven to be very powerful in image segmentation [22]. Of the deep learning methods, convolutional neural networks (CNN) have had the most impact with respect to challenging image segmentation tasks [23, 24]. CNNs have shown excellent results in cell image segmentation [25, 14, 13]. In [26], a new CNN architecture known as U-Net, which used an encoder-decoder structure was proposed. The encoder of U-Net is a stack of convolutional and max pooling layers which is used to capture the context in the image. The decoder is the symmetric expanding path, which is used to enable precise localization. This model demonstrates good performance when addressing the problems in 2D and 3D biomedical image segmentation [27]. In this paper, we will use U-Net to segment the label-free breast cancer cells moving through the RCI device.

Deep learning networks are also widely used in cell classification problems [28, 29, 30, 31]. In [30], a new deep learning approach known as transferring of pre-trained generative adversarial network(TOP-GAN) is proposed to cope with the problem of a small training set. TOP-GAN classifies healthy and cancer cells acquired by quantitative phase imaging. This network structure can be used for many other medical image classification problems that have a small training set. In [31] a novel deep learning architecture is introduced to achieve label-free cell classification. A five-fold cross-validation is used to split the dataset into training, validation, and test subsets. The paper shows good classification results for white blood T-cells relative to colon cancer cells.

Deep learning methods often require 1000s of annotated samples per class [22] to train the networks to have good performance. This can be a problem in many situations where only one or few images of the object that we want to detect are available. In this paper, we used an application known as "ITK-SNAP" [32] to annotate images captured by our chip system. This process was very tedious and it is very difficult to generate a large set of images[1]. One way to address this problem is through the use of data augmentation methods. Traditional methods include linear and nonlinear transforms done on the training data to create "new" or synthetic training images [33, 34]. Typical transformations include spatial flipping, warping and other deformations. An important concept of data augmentation is that the deformations applied to the labeled training images do not change the semantic meaning of the labels. One of the disadvantages of traditional data augmentation approaches is that they produce highly correlated image training data [35, 36].

One solution to this problem is to generate realistic synthetic images using deep learning approaches. In [35], a generative adversarial network(GAN) is proposed to solve image-to-image translation problems. The GAN uses two adversarial networks, a generative network and a discriminative network. The discriminative network learns a loss function to distinguish whether the output image is real or synthetic, whereas the generative network tries to minimize this loss function. In recent years, there have appeared various extensions [34, 35, 36, 37, 38, 39] to GANs, such as Pix2Pix, cycle-consistent adversarial networks (CycleGAN) and spatially constrained cycle-consistent adversarial networks (sp-CycleGAN). Pix2Pix uses conditional adversarial networks to address image-to-image translation problems [35]. One limitation to Pix2Pix is that paired images are required to train the network. Another extension of GAN is cycle-consistent adversarial networks (CycleGAN), which can generate images without paired training data [35]. CycleGAN uses a cycle consistent loss term in the adversarial loss function to generate synthetic images. However, although CycleGAN can generate synthetic images without paired training data, the output images of CycleGAN are unpaired. Sp-CycleGAN is a potential solution to this problem. This network model adds a spatially constrained term to the loss function of CycleGAN so that paired synthetic images can be generated [36, 40, 38].

In the paper, we describe methods to analyze label-free cell images acquired with "the chip" device to automatically detect and recognize the cells. Our approach enables automatic counting of different classes of cells independently, and quantifying the chip's efficiency. In the paper, we describe methods to analyze label-free cell images acquired with "the chip" device to automatically detect and recognize the cells. Our approach enables automatic counting of different classes of cells independently, and quantifying the chip's efficiency.

Cell Segmentation

Our cell segmentation process is summarized in Figure 2. We first get original cell images from phase contrast microscopy [41]. Then we augment the images using traditional and synthetic data augmentation methods. We train the U-Net model [26] with these augmente images and then put the testing images into the trained U-Net to get the prediction results. Finally, we evaluate the U-Net segmentation results.

[1]One needs to be careful of the nomenclature here. In the microscopy community when one says an image is "labeled", this means that fluorescent imaging techniques are being used. In the data science/machine learning community, when one says that an image is "labeled", this means that the images have been annotated (often manually) to indicate where certain structures are located. For example, all the nuclei in the image have been annotated with circles or squares. This process is also known as having "ground truth" information about the image. Here we will use the term "ground truth" or "annotation" in this paper to indicate that the images are "labeled" or annotated.

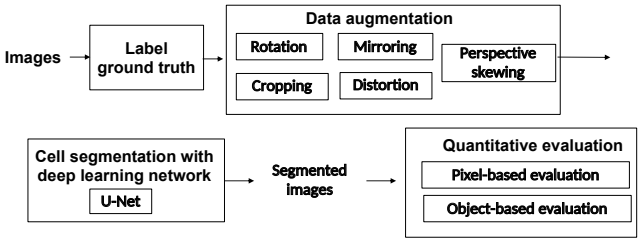


Figure 2. *The block diagram of our cell segmentation.*

Dataset Description

The breast cancer cells were injected into the input channel of the chip [4]. Then the motion of the cells on the chip are observed through phase contrast microscopy [41], and are recorded as a video through a high speed camera. After obtaining the video, we randomly select 20 individual frames from the video for training and 10 individual frames for testing. There is no overlap between the 20 frames for training and the 10 frames for testing. All the pixels that represent the breast cancer cells are manually annotated under the help of our biologists. The application known as "ITK-SNAP" [42], as we described in the introduction section, is used for annotating. Since the depth of the channel is much deeper than the size of the cell, the cells will get in and out of the microscope's focus plane when they travel inside the channel. We are only interested in segmenting in-focus cells, thus only the in-focus cells are annotated.

Traditional Data Augmentation

The 20 labeled frames for training are augmented using traditional data augmentation methods [33]. The main operations used are described in the rest of this section. The operation of perspective skewing transforms the image so that it appears as if you are looking at it from a different angle in three dimensional space. The operation of elastic distortions tries to make distortions to an image while maintaining the image's aspect ratio. The operation of cropping crops the images and resizes them to the original size. In the operation of rotation, we rotate the image to a certain angle. When the angle of rotation is an arbitrary, non-modulo 90 degree angle, the rotation will pad the image in each corner. To alleviate the effect of padding, the largest possible crop is applied to the rotated image while maintaining the image's aspect ratio. The operation of tilting tilts an image along one of its sides. To maintain the original size of the image, we need to crop the image after tilting. In the mirroring operation, the images can be mirrored from top to bottom, or from left to right, or mirrored at a random angle.

After traditional data augmentation, we finally get 1000 augmented training images and their corresponding ground truth images.

Pix2Pix Data Augmentation

Although traditional data augmentation could enlarge the dataset quickly, the images in the augmented dataset are highly correlated since new data is generated through a combination of transforms applied to the training data. Generative Adversarial Nets (GANs) [34] are a promising tool to generate synthetic images for training purposes. GANs can generate a realistic output image from a random noise vector. In an unconditional GAN, only the discriminator observes the input images. In contrast, conditional GANs [35] can generate an output image based on the information of an input image. Both the generator and the discriminator observe the input images. We explored one of the conditional GANs - pix2ix [35] in order to translate ground truth images to synthetic original images. The process of generating realistic synthetic images using pix2pix is summarized in Figure 3.

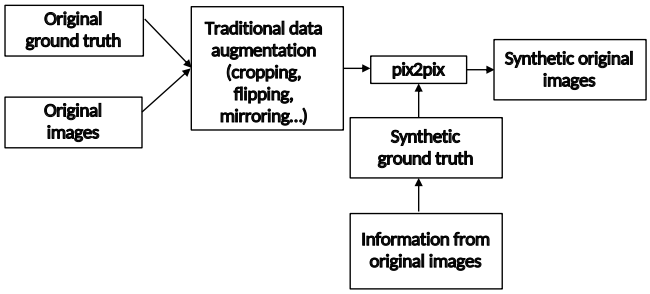


Figure 3. *The block diagram of pix2pix.*

We decided to train a pix2pix network [35] to generate synthetic original images from synthetic ground truth images, since it is easier to generate synthetic ground truth images through traditional image processing methods [33]. We first fed 1000 pairs of original images and corresponding ground truth into the pix2pix model to train the network. The generator is a U-Net structure that contains an encoder-decoder with skip connections. The discriminator is a PatchGAN [35]. During the training, the generator G and the discriminator D are trained simultaneously. The generator G maps ground truth to synthetic original images. Also, the discriminator D learns to distinguish between the synthetic pair composed of ground truth and synthetic images and the real pair composed of ground truth and original images.

After the pix2pix training is finished, and before it is used for inference, we need to generate synthetic ground truth as the input of inference. We analyzed the cells from 20 manually annotated training images by measuring their size and counting their numbers. Then we set the following rules to generate the synthetic ground truth by the information extracted from the annotated training images. The radius of the cells ranges between 7 to 20 pixels. The number of cells on the chip varies between 5 to 20. The ratio of the short axis and the long axis is between 0.7 to 0.95. The above three parameters of the cells are uniformly distributed within their own range. The shapes of the cells are all elliptical. All the cells are distributed with equal probability within the channels of the chip. We also plotted the shape of the chip on our synthetic ground truth to help generate the chip in the synthetic images. We generated 1000 synthetic ground truth images, and put them into the pix2pix network. Then we ran the generator G to get 1000 corresponding synthetic original images as the output. The ground truth of each output synthetic image is exactly the input synthetic ground truth. In this way, synthetic original images and corresponding ground truth are generated.

Figure 4 shows one synthetic ground truth image and the corresponding original image generated by pix2pix. We can find that the synthetic image looks realistic compared to the real original image in Figure 1. Background noise and one out-of-focus cell are also generated in Figure 4.

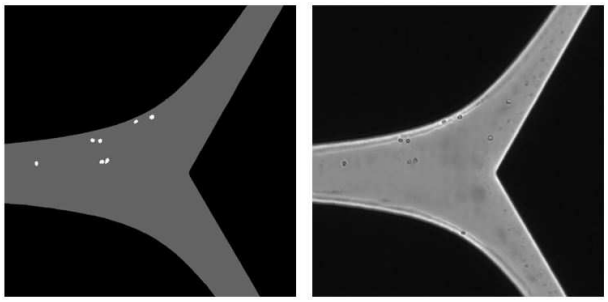


Figure 4. *One example of the synthetic ground truth generated by pix2pix and its corresponding original image.*

Our Approach

We used the standard U-Net network to segment our images since the U-Net architecture showed promising performance in medical image processing [26]. Figure 5 shows the architecture of the U-Net. The filter size of each convolution is 3×3. To maintain the same size of pixel during convolution, a pixel padding of 1×1 is used in each convolution. A batch normalization and a Rectified Linear Unit (ReLU) function are employed after each convolution. In the downsampling path, a max pooling operation with a 2×2 window and a stride of 2 is used. In the upsampling path, feature information is retrieved using transpose convolutions. Our training loss function is the Dice function [43], which is calculated by dividing the union pixels between ground truth and prediction result by the total number of pixels.

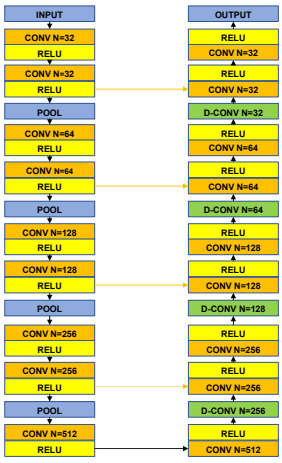


Figure 5. *The structure of U-Net.*

We trained two U-Net models using two different training datasets, one generated by the traditional data augmentation method, the other generated by our synthetic data augmentation method pix2pix. We first put 1000 traditional augmented images into the U-Net model to get the trained network. Then 1000 synthetic images were fed into U-Net and another network was trained. During the inference, we put 10 test images into the above two trained U-Net models separately and got two different sets of segmentation results.

Quantitative Evaluation

We evaluated the U-Net segmentation results at the image pixel level [43]. The pixel accuracy (PA), $Type-I$ error, $Type-II$ error and $F1$ score were obtained by comparing the U-Net output images and the manually annotated ground truth images.

PA, $Type-I$ error and $Type-II$ error are defined as follows:

$$PA = \frac{N^p_{tp} + N^p_{tn}}{N^p_{total}} \tag{1}$$

$$Type-I = \frac{N^p_{fp}}{N^p_{total}} \tag{2}$$

$$Type-II = \frac{N^p_{fn}}{N^p_{total}} \tag{3}$$

where N^p_{tp} is defined as the number of cell pixels that were segmented as cell pixels, N^p_{tn} is the number of background pixels that were segmented as background pixels, N^p_{fp} is the number of background pixels that were segmented as cell pixels, N^p_{fn} is defined be to the number of cell pixels that were segmented as background pixels, and N^p_{total} denotes the total number of cell and background pixels in an image.

Anther pixel level measure we used is the $F1$ score metric. It is the harmonic mean of the precision P and the recall R.

$$P = \frac{N^p_{tp}}{N^p_{tp} + N^p_{fp}} \tag{4}$$

$$R = \frac{N^p_{tp}}{N^p_{tp} + N^p_{fn}} \tag{5}$$

Given the values of precision P and recall R, the $F1$ score is:

$$F1 = \frac{2PR}{P+R} \tag{6}$$

Table 1 summarizes and compares the pixel level segmentation performance of U-Net models trained with two different sets of training data. We can find that they both show very promising results. The method using synthetic images as training data performs a little better than that using traditional-augmented images as training data.

Cell Classification

Data Description

A mixture of blood cells with Hela cells [44, 45] was injected into the chip. The blood cells were enriched with white blood cells with some residual red blood cells. There could also have been cellular debris in the mixture. The motion of the cells on the chip was recorded as a video. We only analyzed white blood cells and Hela cells in the video and ignored all other kinds of

Table 1 Quantitative evaluation of segmenting breast cancer cells between traditional and synthetic data augmentation methods

Data Augmentation Method	Traditional	Synthetic
Pixel Accuracy	99.94%	99.94%
$Type-I$ Error	0.063%	0.060%
$Type-II$ Error	0.001%	0.001%
$F1$ score	84.59%	85.20%

cells. We selected 20 individual frames from the video for training and 5 individual frames from the video for testing. The cells on these frames were manually labeled using "ITK-SNAP"[42]. White blood cells and Hela cells were labeled as different colors. Figure 6 shows one frame and the corresponding manually annotated ground truth. White blood cells are small and usually have irregular shape, while Hela cells are big and usually have regular shape. In the ground truth, green-labeled cells represent Hela cells and red-labeled cells represent white blood cells.

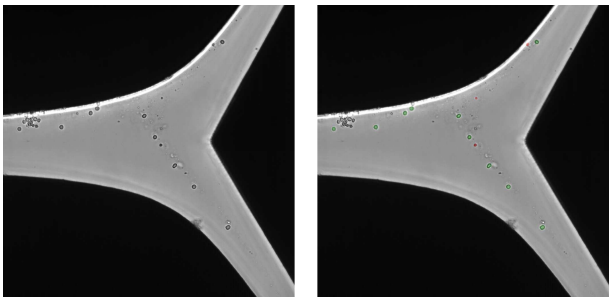


Figure 6. *The original image and the corresponding ground truth.*

Traditional Data Augmentation

Similar to the data augmentation used in cell segmentation, we also used traditional data augmentation methods to augment the training dataset. The following main operations were used: In the perspective skewing operation, the image is transformed so that it appears that one is looking at the image from a different angle. "Elastic distortions" makes distortions to an image while maintaining the image's aspect ratio. "Cropping" tries to crop the images and resize it to the original size. "Rotation" rotates the image to a certain angle. When the image is rotated, an arbitrary, non-modulo 90 degree, rotation will result in the image being padded in each corner. To alleviate this effect, the image is cropped and the largest possible crop is retained while maintaining the image's aspect ratio. "Shearing" tilts an image along one of its sides. In the mirroring operation, we can mirror the image from top to bottom, or from left to right, or mirror it at a random angle. After traditional data augmentation, we finally get 1000 augmented training images and the corresponding ground truth.

Pix2Pix Data Augmentation

There are two main drawbacks of traditional data augmentation in cell classification. One drawback is that the augmented dataset are highly correlated. The main differences between the two types of cells are their shape and the size. Traditional data augmentation distorts the shape and changes the size so that the difference between the two types of cells becomes less obvious. Similar to what we did for cell segmentation, we used pix2pix to generate synthetic images as shown in Figure 3.

We first input the traditional-augmented paired original images and corresponding manually annotated ground truth into the pix2pix model to train the network. Then, we generated synthetic ground truth images with two classes of cells. Two classes of cells are independently labeled with different values. We retain the shape of the chip on our synthetic ground truth, since we plan to generate both the chip and the cells in the synthetic images. After sufficient synthetic ground truth images are generated, we put the synthetic ground truth into the pix2pix network and obtain the synthetic images as the output. The ground truth of each output synthetic image is exactly the input synthetic ground truth. In this way, synthetic original images and corresponding ground truth are generated.

Cell Classification

We used two independent standard U-Net networks to segment each class of cells. The filter size of each convolution is 3×3. To maintain the same size of pixel during convolution, a pixel padding of 1×1 is used in each convolution. A batch normalization and a ReLU function are employed after each convolution. In the downsampling path, a max pooling operation with a 2×2 window and a stride of 2 is used. In the upsampling path, feature information is retrieved using transpose convolutions. Our training loss function is the Dice function.

The classification process is described as follows: 1000 augmented images and corresponding ground truth of Hela cells are used to train the first U-Net network. The same 1000 augmented images and corresponding ground truth of white blood cells are used to train the second U-Net network. Then, we put the test images into both trained U-Net networks and get the segmentation results separately for both Hela cells and white blood cells . Finally, we overlay the segmentation results of the Hela cells and the segmentation result of white blood cells. Figure 7 shows one test image and the corresponding ground truth and classification result. The left figure is the ground truth, the middle figure is the original image overlaid by the ground truth, the right figure is the original image overlaid by the classification result. The green-labeled cells represent Hela cells and red-labeled cells represent white blood cells.

Classification Evaluation

We evaluate the segmentation results at both the pixel-based level and the object-based level. For the pixel-based metrics, the pixel accuracy (PA), $Type-I$ error, and $Type-II$ error of pixel segmentation were obtained based on the manually annotated ground truth images. PA, $Type-I$ error, and $Type-II$ error are defined in Eqs. (1)-(3):

The following three metrics are used to evaluate the results at the object level: the $F1$ score, the Dice index, the Hausdorff

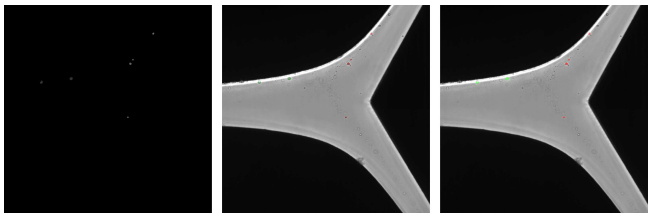


Figure 7. *The ground truth, the overlaid ground truth and the overlaid classification result. The reader is advised to zoom in to see the differently colored classification of the two cell types*

distance. The $F1$ score metric is a measure of the segmentation/detection accuracy of individual objects. The evaluation of the $F1$ score metric is based on two sub-metrics, precision P and recall R.

$$P = \frac{N_{tp}^o}{N_{tp}^o + N_{fp}^o} \tag{7}$$

$$R = \frac{N_{tp}^o}{N_{tp}^o + N_{fn}^o} \tag{8}$$

Given the values of precision P and recall R, the $F1$ score is calculated by the following formula:

$$F1 = \frac{2PR}{P+R} \tag{9}$$

where N_{tp}^o is defined to be the number of cells segmented by the proposed method that overlap at least 50% with their corresponding manually annotated cells. Otherwise, the cell is considered as a false positive and is added to the count of the false positives N_{fp}^o. Similarly, a manually annotated cell that has no corresponding segmented cell or overlaps less than 50% with the segmented cell regions is considered to be a false negative and IS added to the count of the false negatives N_{fn}^o.

The second object-level metric is the Dice Index. The Dice Index measures the similarity between two sets of samples. In our case, the first set is the set of cells belonging to the manually annotated cells demoted by G. And the second set is the set of cells belonging to the segmented cells denoted by S. The Dice Index between G and S is calculated by the following formula:

$$D(G,S) = \frac{2|G| \cap |S|}{|G| + |S|} \tag{10}$$

The third metric, the Hausdorff Distance, focuses on evaluating shape similarity. The Hausdorff Distance is given by [46]:

$$H(G,S) = \frac{1}{2}\left[\sum_{i=1}^{n_S} w_i \widetilde{H}(G_i,S_i) + \sum_{j=1}^{n_G} \widetilde{w}_j \widetilde{H}\left(\widetilde{G}_j,\widetilde{S}_j\right)\right] \tag{11}$$

where,

$$w_i = \frac{|S_i|}{\sum_{p=1}^{n_S}|S_p|}, \widetilde{w}_i = \frac{\left|\widetilde{G}_i\right|}{\sum_{q=1}^{n_G}\left|\widetilde{G}_q\right|} \tag{12}$$

$\widetilde{H}$ is defined to be:

$$\widetilde{H}(G,S) = max\left\{\underset{x\in G}{sup}\,\underset{y\in S}{inf}||x-y||_2, \underset{y\in S}{sup}\,\underset{x\in G}{inf}||x-y||_2\right\} \tag{13}$$

In Eq (11), S_i denotes the i^{th} cell ($i \in \{1,....,n_S\}$) obtained by our segmentation method and G_i denotes a manually annotated cell that is maximally matched with S_i. $\widetilde{G}_j$ denotes the j^{th} cell ($j \in \{1,....,n_G\}$) anotated in the ground truth and $\widetilde{S}_j$ denotes a segmented cell that is maximally matched with $\widetilde{G}_j$. n_S and n_G denote the total number of segmented and manually annotated cells. In Eq (13), $||x-y||_2$ denotes the Euclidean distance between a part of pixels x and y. $\sum_{i=1}^{n_S} w_i \widetilde{H}(G_i,S_i)$ in Eq (11) presents how well each ground truth overlaps with its segmentation, whereas $\sum_{j=1}^{n_G} \widetilde{w}_j \widetilde{H}\left(\widetilde{G}_j,\widetilde{S}_j\right)$ in Eq (11) presents how well each segmented cell overlaps with its ground truth.

Table 2 shows the segmentation quantitative evaluation results.

Then, we evaluate the performance of classification in Table 3. The recognition accuracy is defined as:

$$A = \frac{N_{tp}}{N_{total}} \tag{14}$$

where N_{tp} is defined to be the number of cells that are correctly classified into a certain class and overlap at least 50% with its corresponding manually annotated cells, and N_{total} denotes the total number of cells classified as this class.

Conclusions

In this paper we presented an overview of the work we are doing in cell image analysis. In order to segment label-free cells, we designed a U-Net CNN architecture and demonstrated promising results for the rare cell isolation data. We also investigated synthetic data augmentation to enlarge the training data and compared the results with traditional data augmentation methods Finally, we designed a U-Net CNN architecture and demonstrated initial results for rare cell classification.

In the future, we will examine using temporal information to increase the accuracy of segmentation and classification. We will also investigate other CNN architectures for cell segmentation and classification.

Acknowledgments

This work is sponsored by HP Labs, Palo Alto. Address all correspondence to Edward J. Delp, ace@ecn.purdue.edu

References

[1] Y. Chen, P. Li, P. Huang, Y. Xie, J. D. Mai, Lin Wang, N. Nguyen, and T. J. Huang, "Rare cell isolation and analysis in microfluidics," *Lab Chip*, vol. 14, no. 4, pp. 626–645, February 2014.

[2] M. Watanabe, Y. Uehara, N. Yamashita, Y. Fujimura, K. Nishio, T. Kazuo, F. Koizumi, and Y. Koh, "Multicolor detection of rare tumor cells in blood using a novel flow cytometry-based system," *Cytometry Part A*, vol. 85, no. 3, pp. 206–213, December 2013.

[3] V. Plaks, C. D. Koopman, and Z. Werb, "Circulating tumor cells," *Science*, vol. 341, no. 6151, pp. 1186–1188, September 2013.

Table 2 Performance evaluation of segmenting both types of cells from the background

Metrics	Value
Pixel Accuracy	99.95%
$Type-I$ Error	0.05239%
$Type-II$ Error	0.00217%
$F1$ score	81.15%
Dice Index	86.01%
Hausdorff Distance	101.23

Table 3 Performance evaluation of classification

Metrics	Value
Hela cells recognition accuracy	97.15%
White blood cells recognition accuracy	92.38%

[4] V. Shkolnikov, D. Xin, and C. Chen, "Continuous dielectrophoretic particle separation via isomotive dielectrophoresis with bifurcating stagnation flow," *Electrophoresis*, vol. 40, no. 22, pp. 2988–2995, November 2019.

[5] C. M. Feeley and F. McGovern, "Dielectrophoresis of bubbles in isomotive electric fields," *Journal of Physics D: Applied Physics*, vol. 21, no. 8, pp. 1251–1254, August 1988.

[6] S. Tada, Y. Omi, and M. Eguchi, "Analysis of the dielectrophoretic properties of cells using the isomotive AC electric field," *Biomicrofluidics*, vol. 12, no. 4, pp. 1–12, July 2018.

[7] B. Matthias, W. Theresa, H. Felix, T. Christian, T. Frank, T. Gabriele, S. Sibille, and W. Ralf, "Isolation and time lapse microscopy of highly pure hepatic stellate cells," *Analytical Cellular Pathology: Cellular Oncology*, vol. 2015, no. 417023, pp. 1–14, March 2015.

[8] M. Zborowski and J. Chalmers, "Rare cell separation and analysis by magnetic sorting," *Analytical Chemistry*, vol. 83, no. 21, pp. 8050–8056, November 2011.

[9] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness and Correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, December 2011.

[10] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, June 2006.

[11] P. Bamford, "Empirical comparison of cell segmentation algorithms using an annotated dataset," *Proceedings of the IEEE International Conference on Image Processing*, pp. 1073–1076, September 2003, Barcelona, Spain.

[12] D. Pham, C. Xu, and Jerry L. Prince, "Current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–337, August 2000.

[13] A. Gupta, P.J. Harrison, H. Wieslander, N. Pielawski K., Kartasalo, G. Partel, L. Solorzano, A. Suveer, A. H. Klemm, O. Spjuth, I. Sintorn, and C. Wählby, "Deep learning in image cytometry: A review," *Cytometry Part A*, vol. 95, no. 4, pp. 366–380, 2019.

[14] F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, "Deep learning in microscopy image analysis: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4550–4568, October 2018.

[15] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, And Cybernetics*, vol. 9, no. 1, pp. 62–66, January 1979.

[16] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321–331, January 1988.

[17] S. Lee, P. Salama, K. W. Dunn, and E. J. Delp, "Segmentation of fluorescence microscopy images using three dimensional active contours with inhomogeneity correction," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 709–713, April 2017, Melbourne, Australia.

[18] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours," *Signal Processing*, vol. 71, no. 2, pp. 203–213, December 1998.

[19] K. S. Lorenz, P. Salama, K. W. Dunn, and E. J. Delp, "Three dimensional segmentation of fluorescence microscopy images using active surfaces," *Proceedings of the IEEE International Conference on Image Processing*, pp. 1153–1157, September 2013, Melbourne, Australia.

[20] P. Kornilov and I. Safonov, "An overview of watershed algorithm implementations in open source libraries," *Journal of Imaging*, vol. 4, no. 123, pp. 1–15, October 2018.

[21] T. Atta-Fosu, W. Guo, D. Jeter, C. Mizutani, N. Stopczynski, and R. Sousa-Neves, "3D clumped cell segmentation using curvature based seeded watershed," *Journal of Imaging*, vol. 2, no. 31, pp. 1–17, November 2016.

[22] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, December 2012, Stateline, NV.

[24] C.-C. J. Kuo, "Understanding convolutional neural networks with a mathematical model," *Visual Communication and Image Representation*, vol. 41, pp. 406–413, November 2016.

[25] K. W. Dunn, C. Fu, D. J. Ho, S. Lee, S. Han, P. Salama, and E. J. Delp, "DeepSynth: Three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data," *Scientific Reports*, vol. 9, no. 18295, pp. 1–15, December 2019.

[26] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, October 2015, Munich, German.

[27] D. J. Ho, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Nuclei

detection and segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 418–422, April 2018, Washington, DC.

[28] Y. Ozaki, H. Yamada, H. Kikuchi, A. Hirotsu, T. Murakami, T. Matsumoto, T. Kawabata, Y. Hiramatsu, K. Kamiya, T. Yamauchi, K. Goto, Y. Ueda, S. Okazaki, M. Kitagawa, H. Takeuchi, and H. Konno, "Label-free classification of cells based on supervised machine learning of subcellular structures," *PLOS ONE*, vol. 14, no. 1, pp. 1–20, January 2019.

[29] R. W. Oei, G. Hou, F. Liu, J. Zhong, J. Zhang, Z. An, L. Xu, and Y. Yang, "Convolutional neural network for cell classification using microscope images of intracellular actin networks," *PLOS ONE*, vol. 14, no. 3, pp. 1–13, March 2019.

[30] M. Rubin, S. Omer, T. Nir, N. Yoav, and R. Darina, "TOP-GAN: Label-free cancer cell classification using deep learning with a small training set," *Medical Image Analysis*, vol. 57, pp. 176–185, October 2019.

[31] A. Mahjoubfar, C. L. Chen, and B. Jalali, *Artificial Intelligence in Label-free Microscopy*, Springer, 2017.

[32] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[33] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, July 2019.

[34] H. Shin, N. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Sen-jem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," *Proceedings of the International Workshop on Simulation and Synthesis in Medical Imaging*, pp. 1–11, September 2018, Granada, Spain.

[35] P. Isola, J. Zhu, T. Zhou, and A. A. Efros., "Image-to-image translation with conditional adversarial networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, July 2017, Honolulu, HI.

[36] C. Fu, S. Han, D. J. Ho, P. Salam, K. W. Dunn, and E. J. Delp, "Nuclei segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," *Proceedings of the Computer Vision and Pattern Recognition Workshops*, July 2017, Honolulu, HI.

[37] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, June 2014, Columbus, OH.

[38] I. Goodfellow, Y. Bengio, and A. Courville, "Introduction," in *Deep Learning*, vol. 1, pp. 1–26. MIT Press, Cambridge, MA, 2016.

[39] Ross Girshick, "Fast R-CNN," *Proceedings of the International Conference on Computer Vision*, pp. 1440–1448, December 2015, Santiago, Chile.

[40] A. Rozantsev, V. Lepetit, and P. Fua, "On rendering synthetic images for training an object detector," *Computer Vision and Image Understanding*, vol. 137, pp. 24–37, August 2015.

[41] R. Kasprowicz, R. Sumanab, and P. Tooleb, "Characterising live cell behaviour: Traditional label-free and quantitative phase imaging approaches," *International Journal of Biochemistry and Cell Biology*, vol. 84, pp. 89–95, March 2017.

[42] P. Yushkevich, J. Piven, C. H. Heather, G.Rachel, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, July 2006.

[43] S. Lee, C. Fu, P. Salama, K. W. Dunn, and E. J. Delp, "Tubule segmentation of fluorescence microscopy images based on convolutional neural networks with inhomogeneity correction," *Proceedings of the IS&T International Symposium on Electronic Imaging*, vol. 2018, no. 15, pp. 199–1–199–8, January 2018.

[44] R. Skloot, *The Immortal Life of Henrietta Lacks*, Broadway Paperbacks, 2011.

[45] J. Landry, P. Theodor, T. Rausch, T. Zichner, M. Tekkedil, A. Stütz, A. Jauch, R. Aiyar, G. Pau, N. Delhomme, J. Gagneur, J. O. Korbel, W. Huber, and L. M. Steinmetz, "The genomic and transcriptomic landscape of a hela cell line," *G3: Genes, Genomes, Genetics*, vol. 3, no. 8, pp. 1213–1224, 2013.

[46] K. Sirinukunwattana, J. P. W. Pluim, X. Qi H. Chen, P. A. Heng, Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez, A. Bohm, O. Ronneberger, B. B. Cheikh, D. Racoceanu, P. Kainz, M. Pfeiffer, M. Urschler, D. R. J. Snead, and N. M. Rajpoot, "Gland segmentation in colon histology images: The glas challenge contest," *Medical Image Analysis*, vol. 35, no. 1, pp. 489–502, January 2017.

Author Biography

Han Hu received her B.S. degree from Zhejiang University in China in 2018. She is currently attending graduate school at Purdue University and is working in the Video and Image Processing Laboratory at Purdue University. Her main areas of research are deep learning and image processing.

Yang Lei is a Senior Research Scientist at HP Labs, leading the efforts to develop computer vision solutions in multiple life science projects. Her research interests are computer vision algorithm development and its applications in life science, including 3D cell imaging, cell detection, recognition, and tracking. Dr. Lei has published and presented more than 15 articles in journals and international conferences, in the field of computer vision and image processing. Dr. Lei has been the Chair of IEEE Signal Processing Society Santa Clara Valley chapter since 2019. She received her B.S. degree from Sichuan University and Ph.D. from Purdue University, in Electrical and Computer Engineering.

Daisy Xin is currently a research scientist at HP Labs working to develop a variety of healthcare and research technologies aimed to improve diagnostic and research efficiency. Her background and expertise is in developmental and cell biology, and her professional interests include connecting fundamental biology knowledge and techniques to modern technology. Her experience in biology research helps highlight and define the areas in which technology can make a big impact to enable better and faster research both academically as well as clinically. Daisy received her B.S. in Biological Sciences from UC Irvine and her Ph.D. in Genetics from Yale University.

Viktor Shkolnikov leads a team of scientists and engineers at HP Labs that develop novel fluidic and electrokinetic technologies for solving valuable medical diagnostics problems, such as minimally invasive diagnosis of cancer. His interests and expertise lie at the interface between medicine and fluid mechanics (and the associated physics,

chemistry, and engineering) and he constantly seeks ways of how these can be leveraged to alleviate pain and improve health. Shkolnikov is also the author of "Principles of Microfluidics," a fundamental text in the field. Shkolnikov received his B.S., M.S., and Ph.D. degrees all in Mechanical Engineering and all from Stanford University.

Steven Barcelo manages the Life Sciences research teams at HP Labs. His research has covered the application of low cost nanofabrication techniques to a range of topics including color reflective displays, SERS sensors and memristors. He is most excited about the current opportunity to apply HP's expertise in microfluidics and related technologies to the field of life sciences. Steven holds more than 15 patents and has authored more than 10 journal articles in the fields of SERS, nanoimprint lithography and hydrogen storage materials. He received his B.S. and M.Eng. from Cornell and Ph.D. from UC Berkeley, all in mechanical engineering.

Jan P. Allebach is Hewlett-Packard Distinguished Professor of Electrical and Computer Engineering at Purdue University. Allebach is a Fellow of the IEEE, the Society for Imaging Science and Technology (IS&T), and SPIE. He was named Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named Honorary Member of IS&T, the highest award that IS&T bestows. He has received the IEEE Daniel E. Noble Award, the IS&T/OSA Edwin Land Medal, is a member of the National Academy of Engineering, and a Fellow of the National Academy of Inventors.

Edward J. Delp was born in Cincinnati, Ohio. He is currently The Charles William Harrison Distinguished Professor of Electrical and Computer Engineering and Professor of Biomedical Engineering at Purdue University. In 2004 he received the Technical Achievement Award from the IEEE Signal Processing Society, in 2008 the Society Award from IEEE Signal Processing Society, and in 2017 the SPIE Technology Achievement Award. In 2015 he was named Electronic Imaging Scientist of the Year by IS&T and SPIE. Dr. Delp is a Life Fellow of the IEEE, a Fellow of the SPIE, a Fellow of IS&T, and a Fellow of the American Institute of Medical and Biological Engineering.

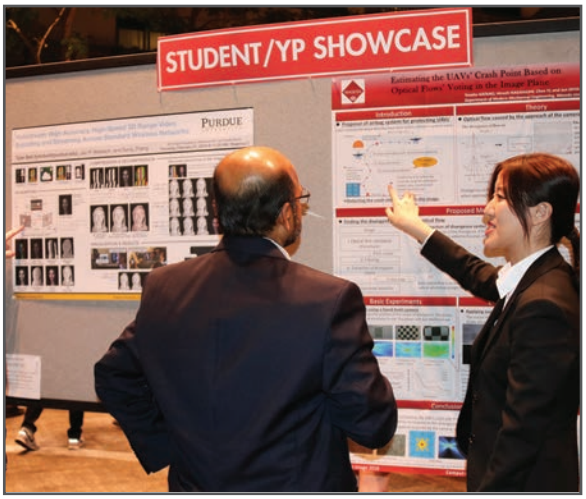
JOIN US AT THE NEXT EI!

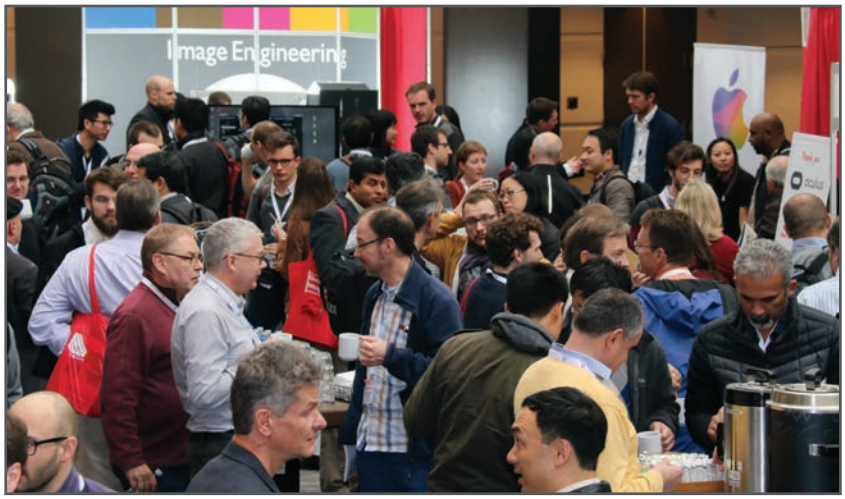
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org