

METHODOLOGY ARTICLE

Open Access



# Iterative unsupervised domain adaptation for generalized cell detection from brightfield z-stacks

Kaisa Liimatainen<sup>1</sup>, Lauri Kananen<sup>1</sup>, Leena Latonen<sup>1,2</sup> and Pekka Ruusuvuori<sup>1\*</sup>

## Abstract

**Background:** Cell counting from cell cultures is required in multiple biological and biomedical research applications. Especially, accurate brightfield-based cell counting methods are needed for cell growth analysis. With deep learning, cells can be detected with high accuracy, but manually annotated training data is required. We propose a method for cell detection that requires annotated training data for one cell line only, and generalizes to other, unseen cell lines.

**Results:** Training a deep learning model with one cell line only can provide accurate detections for similar unseen cell lines (domains). However, if the new domain is very dissimilar from training domain, high precision but lower recall is achieved. Generalization capabilities of the model can be improved with training data transformations, but only to a certain degree. To further improve the detection accuracy of unseen domains, we propose iterative unsupervised domain adaptation method. Predictions of unseen cell lines with high precision enable automatic generation of training data, which is used to train the model together with parts of the previously used annotated training data. We used U-Net-based model, and three consecutive focal planes from brightfield image z-stacks. We trained the model initially with PC-3 cell line, and used LNCaP, BT-474 and 22Rv1 cell lines as target domains for domain adaptation. Highest improvement in accuracy was achieved for 22Rv1 cells. F<sub>1</sub>-score after supervised training was only 0.65, but after unsupervised domain adaptation we achieved a score of 0.84. Mean accuracy for target domains was 0.87, with mean improvement of 16 percent.

**Conclusions:** With our method for generalized cell detection, we can train a model that accurately detects different cell lines from brightfield images. A new cell line can be introduced to the model without a single manual annotation, and after iterative domain adaptation the model is ready to detect these cells with high accuracy.

**Keywords:** Cell detection, Brightfield, Deep learning, Semi-supervised learning, Unsupervised domain adaptation

## Background

Identifying and counting individual cells from cell cultures form the basis of numerous biological and biomedical research applications [1, 2]. Determining numbers of cells reflecting the growth, survival, and death of cell populations form the foundations of e.g. basic cancer research and early drug development. Currently, the most commonly used methods for counting cells in cultures are based on either biochemical measurements, or on fluorescent stainings or markers. These methods are

often either far from optimal in accuracy, costly, or time-consuming. For example, biochemical measurements are indirect measurements in terms of cell numbers. With fluorescent-based imaging, accurate cell numbers can be obtained with well-established image analysis solutions [3]. The fluorescent methods are, however, often problematic, as they require either 1) fixation and staining of cells, being costly and also limiting the number of data obtained per assay and culture, 2) live stains that are toxic to cells, limiting the time-frame of experiments [4], or 3) are based on expression of fluorescent markers in cells, severely limiting the number of cell lines available for use. In addition, the use of fluorescence requires specified imaging equipment and facilities, not at hand in all laboratories.

\*Correspondence: [pekka.ruusuvuori@tuni.fi](mailto:pekka.ruusuvuori@tuni.fi)

<sup>1</sup>Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Full list of author information is available at the end of the article



To avoid the need for fluorescence-based imaging, methods for brightfield imaging are used. Imaging with brightfield microscopy is straightforward with standard facilities available in almost any laboratory, and requires no labeling, making it an efficient and affordable choice. Also the drawbacks from the use of fluorophores on living cells are avoided. However, these benefits come at the cost of inferior contrast compared to fluorescence microscopy.

Most of the current brightfield-based methods rely on feature extraction from single in-focus images, or calculating the area which the cells have covered from the imaged surface. While the former works well for sparse cultures where the cells have individual profiles clearly separated from their background, these methods often do not perform well with dense cultures or cell lines with growth patterns of low contrast. Calculating the area, on the other hand, is once again an indirect estimate for cell count, and also performs more poorly the denser the cultures get. Thus, more accurate brightfield-based methods are desired for cell identification and cell number determination. Especially, improvement in identification of individual cells in dense cell clusters, as well as of cell lines with low contrast growth patterns, are required.

Various cell detection methods for brightfield images in focus have been developed in recent years [5–8]. Unfocused brightfield images or whole brightfield z-stacks have also been applied to cell detection. In our previous study, a z-stack with 25 focal planes was used as input to count PC-3 prostate cancer cells [9]. The method was based solely on the intensity values in images combined with logistic regression classifier. Selinummi et al. used z-stack for creating contrast-enhanced two-dimensional images that provided segmentation results comparable to fluorescence based segmentation [10]. Z-stacks were found to provide useful information for especially cell boundary detection. With a pinhole aperture, one can acquire unfocused images with bright spots marking the cells, which also provides results matching fluorescence based methods [11]. Ali et al. utilized unfocused images for cell and nucleus boundary detection for robust automatic segmentation procedure [12]. Their method is based on differences between two out-of-focused images from opposite directions in z-stack. A similar method with two unfocused, opposite images was proposed by Dehlinger et al. [13]. Z-stacks can also be used to handle slight variations in focal depth, which can be very useful when using autofocus algorithms. This was shown in the research performed by Sadanandan et al., where three consecutive focal planes were used as an input to deep CNN [14].

Convolutional neural networks (CNNs) are the state-of-the-art in machine learning research. After the success of CNNs in ImageNet competition [15], they have been

adopted for classification tasks in biomedical imaging [16, 17]. By discarding the fully connected layers of CNN, the network becomes fully convolutional (FCN) which outputs a heatmap instead of single class value [18]. FCNs have previously been successfully used in cell detection tasks [19]. In a recent study, FCNs have been applied to class-agnostic counting using only a single training example from new domain [20]. A more sophisticated version of basic FCN is the U-Net, especially designed for biomedical image segmentation where localization has high importance [21, 22]. Usually, neural networks are trained in a supervised manner, meaning that large amounts of annotated training data is required.

Domain adaptation can solve machine learning problems where high amount of labeled training data from source domain is present, but there is only little or no labeled data for target domain [23]. Many domain adaptation methods are based on creating a transformation between source and target domains [24, 25]. Domain adaptation can also be performed by learning feature representations shared by both the source and target domains [26, 27].

We propose a method for generalized cell detection from brightfield z-stacks using single annotated cell line (PC-3) for supervised training step. U-Net is chosen as the deep learning model due to its exceptional performance in related tasks, and also due to its fully convolutional and resolution preserving nature. We test the generalization capabilities with LNCaP, BT-474 and 22Rv1 cell lines. Each of these cell lines has their own unique appearance in brightfield images, and these cell lines can be considered as target domains which are somewhat similar to the source domain. The model is trained first with annotated PC-3 samples, which results in high precision but sometimes low recall for other cell lines. We use the predictions from the pre-trained model for generating targets for unseen domains (cell lines) in unsupervised domain adaptation step. In contrast to many transfer learning approaches, we do not use any manually annotated training data for the target domain. However, we preserve some of the original training data to prevent excessive influence of the imperfections in predictions. Thus, training is performed in semi-supervised manner while domain adaptation is unsupervised.

## Methods

In this study, we explore methods for improving generalization in cell detection. We use brightfield z-stacks of PC-3 cell line for supervised training of U-Net-based model for cell detection, and apply an unsupervised domain adaptation step to improve the detection accuracy of cell lines lacking annotated training data. Implementation was programmed with Python language and Keras and TensorFlow modules for deep learning.

### Brightfield data

The data consists of brightfield focus stacks (a.k.a. z-stacks) of monolayer cultures of cancer cell lines. Images were acquired with QImaging Retiga-2000R camera using Olympus IX71 microscope and Objective Imaging Surveyor scanning and imaging software. The z-stack range was  $240\ \mu\text{m}$ , and distance between adjacent focal planes was  $10\ \mu\text{m}$ . Images have a pixel resolution of  $1596 \times 1196$  pixels, corresponding to an area of  $1190.8\ \mu\text{m} \times 891.4\ \mu\text{m}$ . Autofocus method was used to detect most focused image, above and below of which 12 focal planes were imaged. Thus, each stack consists of 25 focal planes in which the 13th focal plane is in focus.

Four cancer cell lines were used in this research: prostate cancer cell lines PC-3, 22Rv1 and LNCaP, and breast cancer cell line BT-474. All cell lines were obtained from American Type Culture Collection (ATCC, Rockville, MD, U.S.A.) and cultured under the recommended conditions. Example images from these cells are shown in Fig. 1. These cell lines were chosen due to their differential appearance in z-stacks, varying from separately growing, high contrast PC-3 and LNCaP to dense and low contrast populations of BT-474 and 22Rv1. The networks were trained with PC-3 using the same data that was used in our previous study [9]. This data was acquired from one cell cultivation, where the cells grew for six days and were imaged daily. For training, we used two images from each day. Thus, twelve images of size  $1196 \times 1596$ , including 5878 annotated cells in total were used.

Four images from each cell line were annotated for the purpose of validating results. The annotated cell count in testing data is 1975 for PC-3 cells, 2183 for LNCaP cells, 1022 for BT-474 cells and 2883 for 22Rv1 cells. Annotations for validating PC-3 cells were not used in training, and the images of PC-3 cells for validation are from a

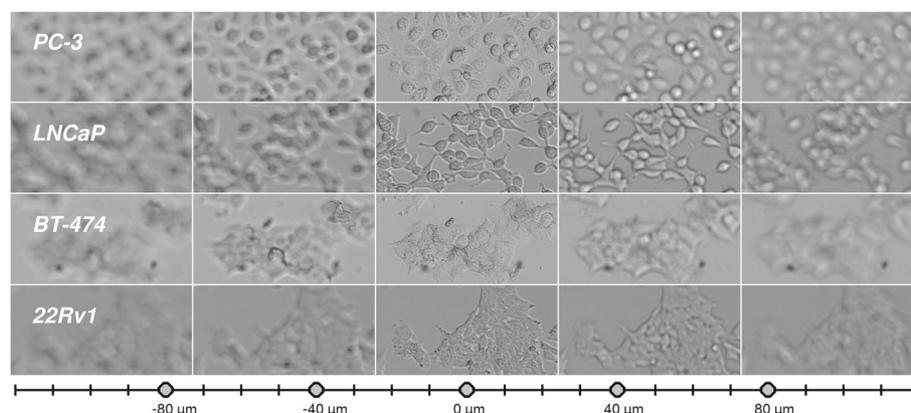
separate cultivation than the training data. In the domain adaptation step, none of the annotated images of other cell lines were used to prevent any distortion of results from over-fitting.

### Method selection and comparison

When selecting the best method for this task, the first criterion was that the resolution of prediction has to be similar to input image resolution, to ensure best possible separation of the cells. Second criterion was that the prediction should be performed to the whole image at once, since pixelwise prediction would require excessive computational resources when fulfilling the resolution requirement. Thus, many deep learning architectures were discarded. Deep learning architectures similar to U-Net fulfill these requirements. Also other, more basic fully convolutional networks without maxpooling layers were taken into consideration, however, the results were inferior to U-Net-like architectures. Two chosen methods from literature were included for comparison: subtraction between opposite z-stack planes [12], and image processing based method from one unfocused image [7]. Some experiments were also performed with pixel intensity based logistic regression classification [9], but the results for other cell lines than PC-3 were not comparable to other methods.

In addition to the actual method, the most suitable focal planes from z-stack were defined. Each method was tested with various degrees of unfocusing and, if possible, various amount of input focal planes. The results of these experiments are shown in Table 1.

Best overall results were acquired with U-Net architecture taking focal planes 13, 14 and 15 as input. Example images of these input focal planes for each cell line are shown in Additional file 1. Most focused focal plane in z-stack has index 13 according to autofocus algorithm



**Fig. 1** Examples of each studied cell line in z-stack with 25 focal planes. The grid represents focal planes, circles marking the planes in the figure in corresponding order

**Table 1** Method and input comparison. For methods by Ali [12] and Buggenthin [7] the results from focal plane producing best overall score are given

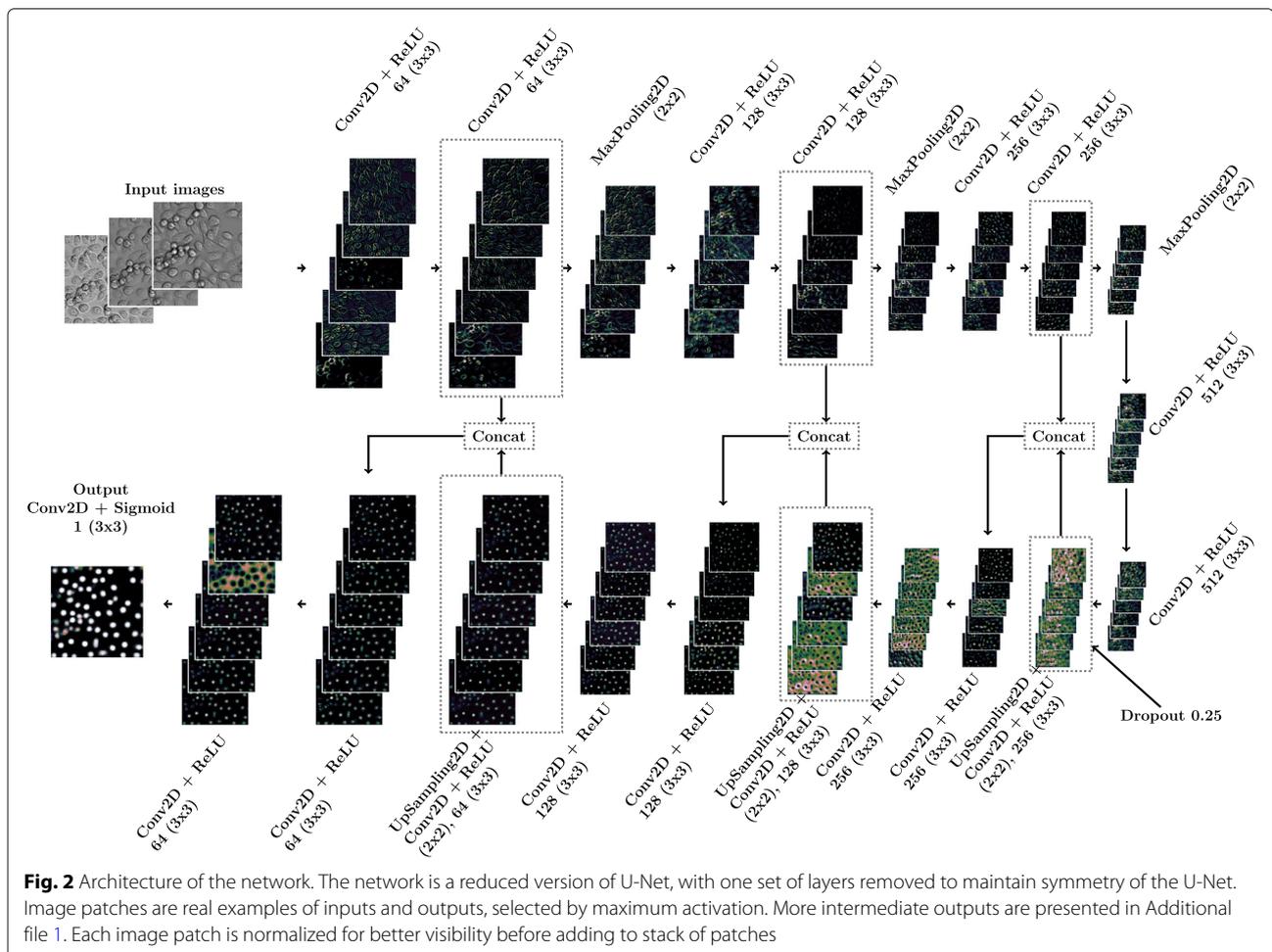
	Ali	Buggenthin	U-Net	U-Net	U-Net	U-Net	U-Net (smaller)
Focal planes	8,18	9	13	11-15	12-14	13-15	13-15
PC-3	0.90	0.8	0.933	0.949	0.944	<b>0.955</b>	<b>0.955</b>
LNCaP	0.683	0.673	0.885	0.889	0.88	0.893	<b>0.895</b>
BT-474	0.576	0.671	0.76	0.724	<b>0.777</b>	0.76	0.757
22Rv1	0.511	0.606	<b>0.69</b>	0.612	0.619	0.673	0.672
Total	0.67	0.688	0.817	0.794	0.805	<b>0.82</b>	<b>0.82</b>

Best result for each cell line is marked with boldface

used in imaging. However, when looking at the planes in question, one can argue that actually the most focused focal plane is the one with index 14. Then, our conclusion for best suited planes would be the same as in [14].

When studying the intermediate outputs of U-Net model, we noticed that cell detections were already present in multiple intermediate layers before the last. Thus, the question rose whether some of the layers could be discarded without losing accuracy. One residual layer

set was removed from the architecture while keeping the symmetry, and in the last column of Table 1 we can see that the results are as good as with whole U-Net. Thus, to reduce computational burden and memory requirements of the model, this smaller architecture was chosen as the method. This network architecture, with some example layer outputs, is shown in Fig. 2. More detailed description of layers is presented in Additional file 1. After selecting the best method, the training pipeline was further improved for better accuracy.



### Training details

Initially, we trained the deep learning model with PC-3 exclusively. For training target, a binary mask was created. Each cell is presented with a disk shaped structuring element with a radius of 8, but if these circles touch each other, the radius is reduced for better separation of the cells. The maximum radius of 8 is a good trade-off between cell separation capability of the model and class balance in training. With this radius, we get approximately 20 to 1 relation between background and cell pixels, which is still a manageable class balance. We use a total of 12 1596 × 1196 images for training the network. One quarter of the training set was set aside for validation during training.

The most difficult cell line from cell detection point-of-view is 22Rv1. These cells are much smaller than the cells from other cell lines. Based on this knowledge, we augmented the training data by resizing PC-3 images to 75 percent of their original size, which doubled the amount of training data. To match the resized images, maximum radius for circles in training targets was reduced to 6 pixels, using the same proportion.

Since training target is a binary mask, binary cross-entropy was chosen as a loss function. We used stochastic gradient descent (SGD) optimizer with a Nesterov momentum of 0.8, and batch size was 5 samples. We set the learning rate to 0.1 at the beginning of training, and after every 10 epochs reduced it to half. Since convolution discards a small proportion of information from the border of the image, and the fully convolutional nature of U-Net architecture allows changing input size, the size of input patches was switched after each set of 10 epochs for better utilization of the training data. After loading each training set, image transformations were applied randomly to each patch of training data. These transformations include rotation, translation, small intensity shifts and adding noise. The model was trained for a total of 60 epochs. However, the weights were saved only when

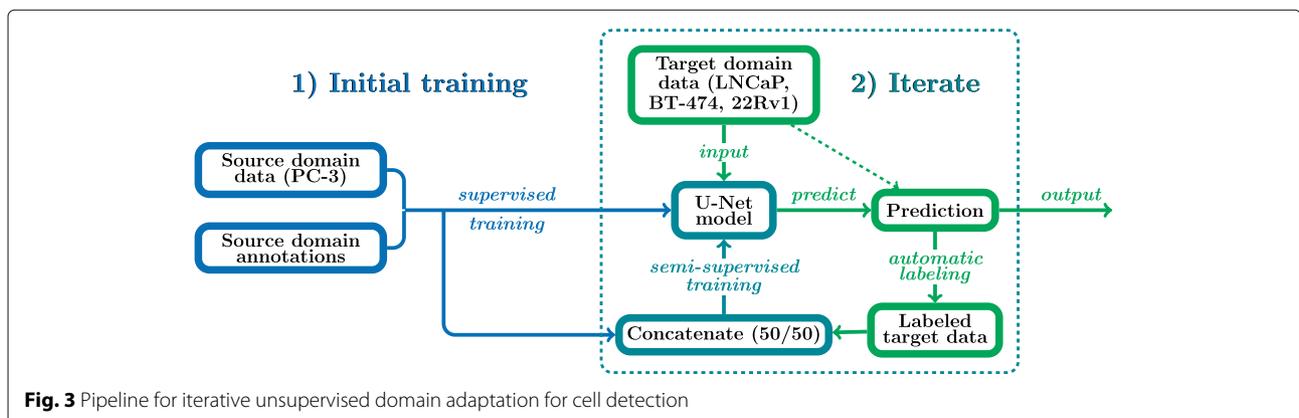
validation loss decreased, resulting in an actual amount of epochs less than 60.

### Domain adaptation

In this study, domain adaptation is used to fit the model trained with PC-3 cell line (source domain) to other cell lines (target domains). Not a single cell from target domains LNCaP, BT-474 and 22Rv1 was annotated for training purposes.

After achieving reasonable recall and, more importantly, high precision for all cell lines via training with only PC-3, the domain adaptation step was applied. In Fig. 3, the pipeline of the method is presented. Domain adaptation is marked with a blue dashed square and it is repeated six times in total. Domain adaptation was performed in an unsupervised manner; half of the training data was generated by auto-labeling from the target cell line, while the other half represented randomly selected patches of previously used PC-3 training data. Since the domain adaptation step is based on predictions of unseen cell lines, annotated PC-3 training data is required to prevent the model from fitting to false positives or negatives. Without the annotated data, the false predictions can get amplified during training since auto-labeling is performed iteratively during training.

Auto-labeling was performed in the following manner. First, prediction was calculated for four randomly selected images of the target cell line. The images that were annotated for testing purposes were not available for the selection. Then, local peaks were detected from the predicted heat map with a threshold of 0.2 and a minimum distance of 5 pixels between peaks. Threshold was set low to include also weak predictions to training data. Each peak was marked as a cell, and these cell points were transformed to training targets with binary dilation with a disk shaped structuring element with a radius of 6 pixels. The model was trained for another 60 epochs in total. New prediction-based training data was calculated after each



set of 10 epochs, simultaneously decreasing the learning rate to half and changing patch size, as was performed in supervised training.

### Validation metrics

In the task of cell counting, true negatives are ambiguous. Since true negatives are not included in  $F_1$ -score, it is a suitable metric for validating our results. To find matching cells between prediction and ground truth, detected cell coordinates were compared to ground truth coordinates with Euclidean distance, and distance below 20 was considered as detection. If multiple coordinates were found within the threshold distance, only the closest of these was accepted as detection.

We count true positives ( $TP$ ), false positives ( $FP$ ) and false negatives ( $FN$ ) for each image in the test set. With these values, we calculate precision (positive predicting value),  $\frac{TP}{TP+FP}$ , and recall (sensitivity),  $\frac{TP}{TP+FN}$ . Finally,  $F_1$ -score is calculated via equation  $2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$ .

In density-based accuracy calculations, a density map was created using kernel density estimation with normal kernel ( $\sigma = 50$  pix) for smoothing discrete cell locations into local neighborhood. This density map was then divided into five areas covering equal density range. Note that this does not correspond to dividing the areas based on equal area coverage, nor will there be equal number of cells within the density areas.

### Results

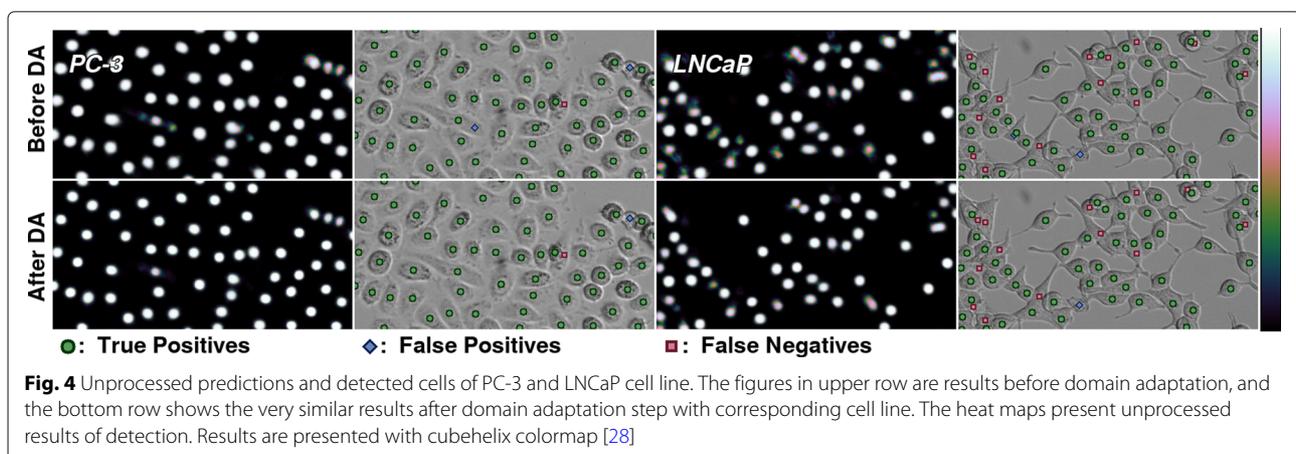
To demonstrate the performance of the proposed cell detection methodology, we present results for different experimental setups. First, we show how deep learning masters the challenge of label-free cell detection from bright field focus stacks in a very accurate manner. Second, we show how precision remains high for cell lines never seen by the classifier. Finally, we present how the high precision can be used for iterative unsupervised domain adaptation. Furthermore, we present detection

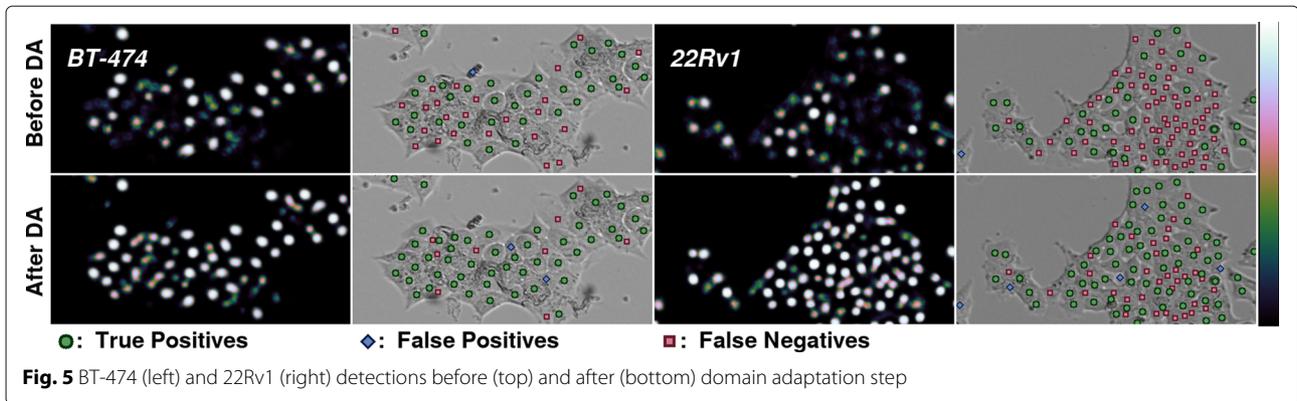
accuracy in relation to cell growth density. While small example images are presented in result figures, examples of whole image level detections are given in Additional file 1.

We performed convolutional neural network based label-free cell detection of PC-3 cancer cells. We acquire  $F_1$ -score of 0.95 for this cell line. An example prediction is shown in Fig. 4, on the left half. Even though the prediction is often close to perfect, the stacking cells are not always well separated. This slightly reduces the overall accuracy, but with a score of 0.95, our method can still be used to e.g. accurately count the growth curve of PC-3 cell line.

Then, we tested how the cell detector generalizes from PC-3 to multiple cancer cell lines. The scores acquired with the model trained with PC-3 cell line only are shown in first halves of plots in Fig. 6. We achieve a high accuracy with 0.89  $F_1$ -score for LNCaP cell line. An example detection is shown in Fig. 4. When moving to cell lines of dense populations and low contrast, namely BT-474 and 22Rv1, the scores are decreased. Precision still remains high for both of these cell lines, but recall drops drastically when compared to LNCaP and PC-3. Also, recall fluctuates a lot between each set of 10 epochs. We could say that the model does not actually fit to BT-474 and 22Rv1 as it does to PC-3 and LNCaP. Indeed, the best score is acquired after only 20 epochs of training, which implies that the more the model is fitted towards PC-3, the farther it goes from fitting to BT-474 and 22Rv1. At this stage, heavy data augmentation had already been applied and this improved especially the 22Rv1 detection. For this cell line, the  $F_1$ -score is about 0.1 higher when we double the training data by resizing it to 75 percent of the original size. 22Rv1 cells are the smallest in our set, so smaller cells in training data naturally improve the accuracy.

Next, we applied unsupervised domain adaptation to improve generalization to unlabeled data from unseen cell lines. In Fig. 5, we show how domain adaptation step





improves the results.  $F_1$ -score for 22Rv1 rises from 0.65 to 0.84, and the score for BT-474 rises from 0.74 to 0.87.

In Fig. 7, we show detection accuracies divided into groups of different cell densities, and cell amount within those areas. Density areas are illustrated as contour overlays in Additional file 1. Accuracies are calculated for models adapted to each unseen domain, and also for the model before domain adaptation. In Fig. 7, left panel, the dashed lines represent accuracy before domain adaptation. The accuracy decreases considerably when moving to denser areas, but after domain adaptation (solid lines), the accuracies of the densest areas are comparable to sparser areas. It should be noted that in the second to last densest group, there are only 4 LNCaP cells, which results in sudden drop in accuracy.

## Discussion

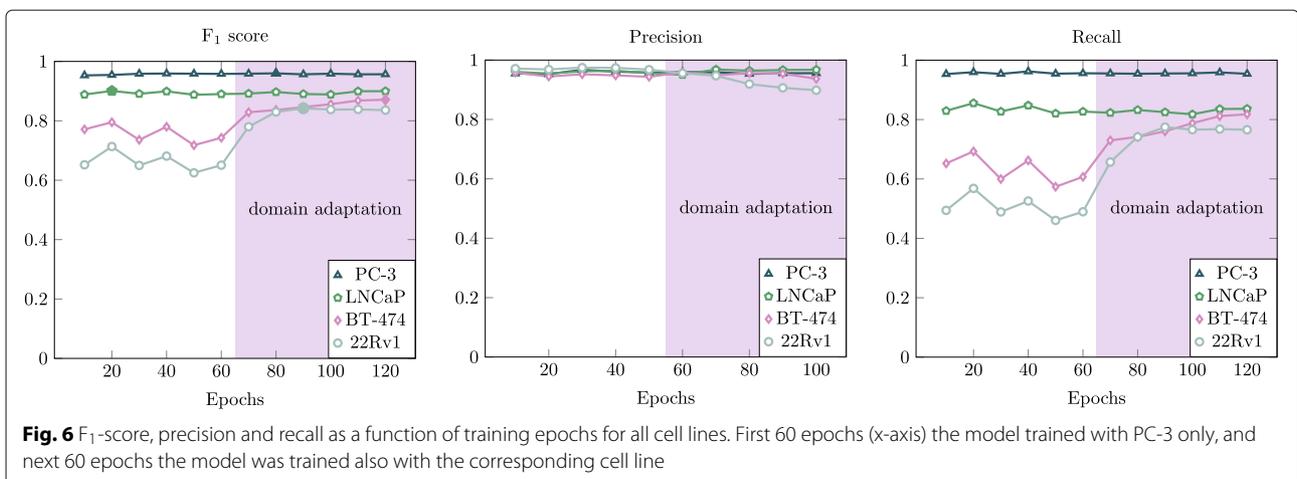
### Convolutional neural network based label-free cell detection of PC-3 cancer cells

The convolutional neural network based label-free cell detection was applied to data from PC-3 cancer cell line with the accuracy of  $F_1$ -score 0.95. PC-3 cell line has a clear profile in brightfield focus stacks and high detection

accuracy is acquired with as few as 10 to 20 epochs of training of a deep learning model, as shown in Fig. 6. With a U-Net-like model, we obtain a clean and sharp heat map as an output, where each cell is represented with a circle. It should be noted that we aim at cell detection, not segmentation, and the model outputs circular detections since it was trained with a binary mask where circles represent the cells. A high accuracy can be achieved also with a single image in focus, but with z-stacks we can improve especially the precision of detections. Non-cell objects are often not as similar to cells in images out-of-focus as they are in only focused image. For example, the artifacts caused by impurities in camera lens do not change their appearance when going out of focus.

### Generalization from PC-3 to multiple cancer cell lines

Generalization to other cell lines was analyzed by applying the cell detection model trained with PC-3 cell line to data from other cancer cell lines, which were LNCaP, BT-474 and 22Rv1. In brief, the results obtained for LNCaP were of high accuracy, while the accuracy dropped for BT-474 and 22Rv1. The LNCaP cell line somewhat resembles



PC-3 since the cells tend to grow separately. Especially the cell lines that grow in dense populations do not receive very high accuracy with the network trained with PC-3 only. One reason for this might be that there often is no background around cells that grow close to each other, while each non-stacked cell of PC-3 has at least some background around them. Also the height and shape of the cells in  $z$  may affect their contrast properties. In Fig. 5, on top row, we see detections of BT-474 and 22Rv1 after supervised training. BT-474 is detected with reasonable accuracy, achieving an  $F_1$ -score of 0.74. For the dense population of 22Rv1 cells, most cells in the center of the example image have not been detected, and the  $F_1$ -score is only 0.64. However, the score is high enough for successful domain adaptation.

### Improved generalization through iterative unsupervised domain adaptation

With domain adaptation, domain being another cell line, we can greatly increase the accuracy of the cell detection for the unseen cell lines, especially those growing in dense populations. The worse the accuracy is before domain adaptation, the more it is improved with domain adaptation. In Fig. 5, we show how multiple previously undetected cells get clear detections after domain adaptation (bottom row). Especially in 22Rv1 (Fig. 5 on the right), the improvement is drastic. Even though some of the cells in the dense center have non-zero confidence that is not registered to the score before domain adaptation, there are several clear detections for these cells after domain adaptation (compare top and bottom row).

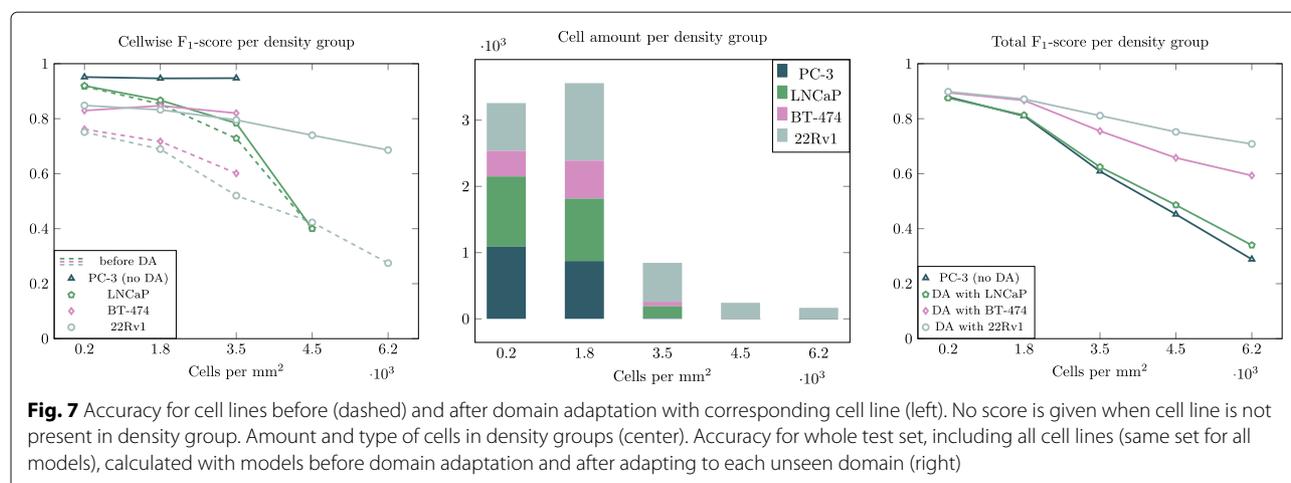
For cell lines that already get very accurate predictions with  $F_1$ -scores around 0.9 after initial training with PC-3, the domain adaptation step does not result in a significant change. Thus, we can apply the domain adaptation step to any cell line with reasonable confidence of not reducing the detection accuracy. In Fig. 4, we can see very

little difference in the detections for PC-3 and LNCaP cells, even though in the predicted heat map, the detection signals are visibly more distinct.

### Relation between accuracy and cell growth density

In order to gain more insight into the effect of cell density on the detection accuracy, we created a pooled test set by using data from all cell lines for determining the detection accuracy. It should be noted that this time, the absolute number of cells and the relative fraction of cells from each cell line varies from density area to another (see Fig. 7, center panel for the cell type distributions among areas). From results for the whole test set (Fig. 7, right panel), we see that with the model trained with PC-3 only, the accuracy is low on dense areas. In addition, when adapting to LNCaP domain, the score remains low although slight improvement is apparent. When training with the densely growing cell lines, BT-474 and 22Rv1, the scores are considerably improved. Even though both of these cell lines grow in dense populations, the BT-474 cells are bigger than 22Rv1 cells. Thus, 22Rv1 cell line is able to grow more dense than the other cell lines, resulting in that they are the only cell line present in densest of areas (Fig. 7, center panel). Yet, the improvement in scores for BT-474 is comparable to 22Rv1-trained model. This implies that the size of cells is not a property that greatly differentiates the cells in the model's point of view. However, the height of cells affects the contrast of cells in  $z$ -stacks. This is a property that also affects the similarity between cell lines. In addition, the cells within dense populations do not have any visible background surrounding them, which is a joint property of BT-474 and 22Rv1 cell lines.

According to these results, the overall accuracy never decreases when adapting to a new domain. Thus, the new features learned during domain adaptation cannot be just cell line specific. In addition, since PC-3 data is also



used when adapting to a new domain, the PC-3 detection accuracy does not decrease.

## Conclusions

Many applications of biological and biomedical research require accurate cell detection and counting. Our results show that with deep learning we can accurately detect PC-3 cells from brightfield z-stacks without the need for fluorescence imaging. Furthermore, the model generalizes well for cell lines similar to PC-3. In case of densely growing cells with low contrast, properties that differentiate these cells from PC-3, we achieve lower recall but high precision. High precision enables automated generation of suitable training targets for domain adaptation. With iterative unsupervised domain adaptation, we can increase the accuracy of previously poorly detected cell lines considerably. The higher the dissimilarity is between the source and the target cell lines, the more improvement can be achieved via domain adaptation.

Our contribution to research fields depending on cell counting is a framework for unsupervised domain adaptation, including a pre-trained model, for accurate detection of various cell lines unseen by the classifier. Manual annotation for these cell lines is not required due to automated labeling of new training data. Since our method is based on brightfield images, it is available for all laboratories with just basic imaging equipment.

## Additional file

**Additional file 1:** Supplementary **Figures S1–13** and Supplementary **Table S1.** (PDF 18200 kb)

## Abbreviations

CNN: Convolutional neural networks; FCN: Fully convolutional network; FN: False negative; FP: False positive; TP: True positive

## Funding

The authors gratefully acknowledge Academy of Finland projects #314558 & #313921 (PR) and #317871 (LL) for funding.

## Availability of data and materials

Python implementation is available at <https://github.com/Bioimageinformatics/Tampere/domain-adaptation-cell-detection>. All data and annotations are available via the same link.

## Authors' contributions

KL implemented the image processing and classification pipeline, analyzed the data, and wrote the bulk of the manuscript. LK participated in manual annotation of the validation data. LL carried out the laboratory experiments and imaging, and participated in manual annotation and writing of the manuscript. PR designed the study, supervised the computational work and participated in writing of the manuscript. All authors read and approved the final version of the manuscript.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. <sup>2</sup>Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland.

Received: 2 October 2018 Accepted: 4 January 2019

Published online: 15 February 2019

## References

1. Meijering E. Cell segmentation: 50 years down the road [life sciences]. *IEEE Signal Proc Mag.* 2012;29(5):140–5.
2. Meijering E, Carpenter AE, Peng H, Hamprecht FA, Olivo-Marín J-C. Imagining the future of bioimage analysis. *Nat Biotechnol.* 2016;34(12):1250–5.
3. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, et al. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006;7(10):100.
4. Laissue PP, Alghamdi RA, Tomancak P, Reynaud EG, Shroff H. Assessing phototoxicity in live fluorescence imaging. *Nat Methods.* 2017;14(7):657–61.
5. Tscherepanov M, Zöllner F, Hillebrand M, Kummert F. Automatic segmentation of unstained living cells in bright-field microscope images. *Advances in mass data analysis of images and signals in medicine, biotechnology, chemistry and food industry.* Springer; 2008. p. 158–72.
6. Adiga U, Taylor D, Bell B, Ponomareva L, Kanzlemar S, Kramer R, Saldanha R, Nelson S, Lamkin TJ. Automated analysis and classification of infected macrophages using bright-field amplitude contrast data. *J Biomol Screen.* 2012;17(3):401–8.
7. Buggenthin F, Marr C, Schwarzfischer M, Hoppe PS, Hilsenbeck O, Schroeder T, Theis FJ. An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinformatics.* 2013;14(1):297.
8. Tikkanen T, Ruusuvaari P, Latonen L, Huttunen H. Training based cell detection from bright-field microscope images. In: *Image and Signal Processing and Analysis (ISPA), 2015 9th International Symposium On.* IEEE; 2015. p. 160–4.
9. Liimatainen K, Ruusuvaari P, Latonen L, Huttunen H. Supervised method for cell counting from bright field focus stacks. In: *IEEE 13th International Symposium on Biomedical Imaging (ISBI).* Prague: IEEE; 2016. p. 391–4.
10. Selinummi J, Ruusuvaari P, Podolsky I, Ozinsky A, Gold E, Yli-Harja O, Aderem A, Shmulevich I. Bright field microscopy as an alternative to whole cell fluorescence in automated analysis of macrophage images. *PLoS ONE.* 2009;4(10).
11. Drey LL, Graber MC, Bieschke J. Counting unstained, confluent cells by modified bright-field microscopy. *Biotechniques.* 2013;55(1):28–33.
12. Ali R, Gooding M, Szilágyi T, Vojnovic B, Christlieb M, Brady M. Automatic segmentation of adherent biological cell boundaries and nuclei from brightfield microscopy images. *Mach Vis Appl.* 2012;23:607–21.
13. Dehlinger D, Suer L, Elsheikh M, Peña J, Naraghi-Arani P. Dye free automated cell counting and analysis. *Biotech Bioeng.* 2013;110(3):838–47.
14. Sadanandan SK, Ranefall P, Le Guyader S, Wählby C. Automated training of deep convolutional neural networks for cell segmentation. *Sci Rep.* 2017;7(1):2045–322.
15. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. *Advances in Neural Information Processing Systems*; 2012. p. 1097–1105.

16. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199–210. <https://doi.org/10.1001/jama.2017.14585>.
17. Sirinukunwattana K, Raza SEA, Tsang Y, Snead DRJ, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Trans Med Imaging*. 2016;35(5):1196–206. <https://doi.org/10.1109/TMI.2016.2525803>.
18. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2015. p. 3431–40.
19. Xie W, Noble JA, Zisserman A. Microscopy cell counting and detection with fully convolutional regression networks. *Comput Methods Biomech Biomed Eng Imaging Vis*. 2016;0(0):1–10. <https://doi.org/10.1080/21681163.2016.1149104>.
20. Lu E, Xie W, Zisserman A. Class-agnostic counting. 2018. [arXiv:1811.00472](https://arxiv.org/abs/1811.00472).
21. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *CoRR*. 2015;abs/1505.04597. [1505.04597](https://arxiv.org/abs/1505.04597).
22. Dong H, Yang G, Liu F, Mo Y, Guo Y. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In: Valdés Hernández M, González-Castro V, editors. *Medical Image Understanding and Analysis*. Cham: Springer; 2017. p. 506–517.
23. Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vaughan JW. A theory of learning from different domains. *Mach Learn*. 2010;79(1): 151–75. <https://doi.org/10.1007/s10994-009-5152-4>.
24. Sun B, Feng J, Saenko K. Return of frustratingly easy domain adaptation. *CoRR*. 2015;abs/1511.05547. [1511.05547](https://arxiv.org/abs/1511.05547).
25. Gopalan R, Li R, Chellappa R. Domain adaptation for object recognition: An unsupervised approach. In: *2011 International Conference on Computer Vision*; 2011. p. 999–1006. <https://doi.org/10.1109/ICCV.2011.6126344>.
26. Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, Marchand M, Lempitsky V. Domain-adversarial training of neural networks. *J Mach Learn Res*. 2016;17(59):1–35.
27. Bousmalis K, Trigeorgis G, Silberman N, Krishnan D, Erhan D. Domain separation networks. *CoRR*. 2016;abs/1608.06019. [1608.06019](https://arxiv.org/abs/1608.06019).
28. Green DA. A colour scheme for the display of astronomical intensity images. *Bull Astron Soc India*. 2011;39:289–95.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

