PROJ-H402

Computing Project

Convolutional deep neural network for cell segmentation

Author:
László ALMÁSY

Supervisor:
Olivier DEBEIR

Université Libre de Bruxelles

March 29, 2019

Contents

1 Description of the project

1.1 Aim

Segmenting cells on sequences of images is a complex and repetitive task that would be a waste of a researcher's time. However these images are not always easy to segment automatically using the traditional methods of segmentation, in particular phase contrast images that have a low contrast between the cells and the background. The use of convolutional neural networks is on the rise in the field of image processing, as the current computer performances allow for the training of powerful networks in reasonable time. Neural networks may bring a new reliable solution for automated and unsupervised cell segmentation.

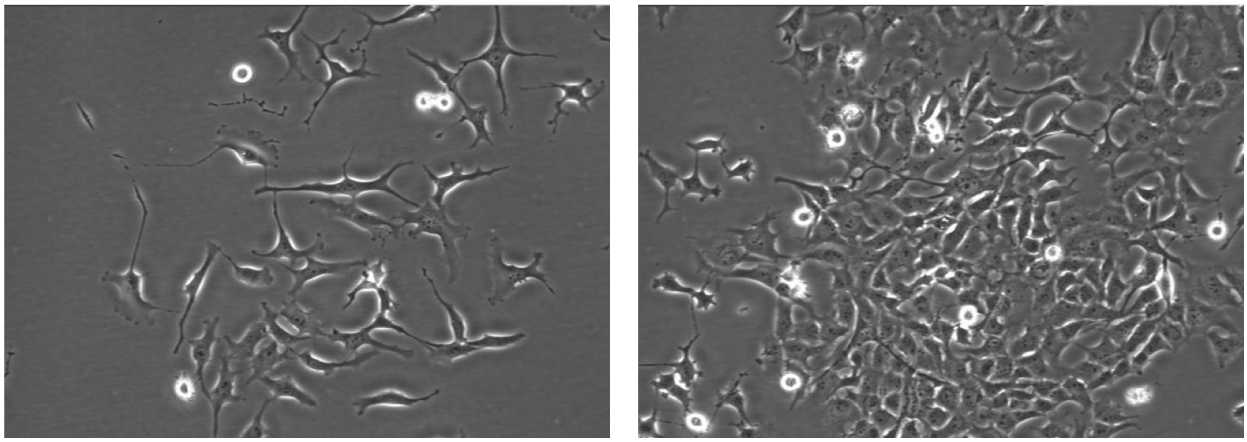
As a student in first year of a computer engineering master, I was given the task to investigate this segmentation method, and attempt using it on a dataset of unsegmented phase contrast microscopy (PCM) images.

1.2 Dataset

The dataset consists of multiple sequences of phase contrast images of cancer cells. This dataset has some characteristics that make it difficult to segment.

1. The cells change a lot throughout the sequences. The number of cell increases and goes from sparse (figure 1a) to dense (figure 1b) due to the mitosis.
2. The cells can present a lot of outgrowth, that may overlap with other cells (figure 2).
3. The sequences come from different experiments, therefore the images have varying quality, sizes. Some are more zoomed than others and the relative size of the cells to the image varies (figure 3).
4. Some sequences present very densely clustered cells that are very hard to separate.

This dataset also only contains the raw images of the experiments. By not having a set of corresponding masks, a neural network cannot use the dataset to learn the correct segmentation.



(a) Beginning of the sequence with sparse cells (b) End of the sequence with a high density of cells

Figure 1: Example of a sequence with an increasing density of cell

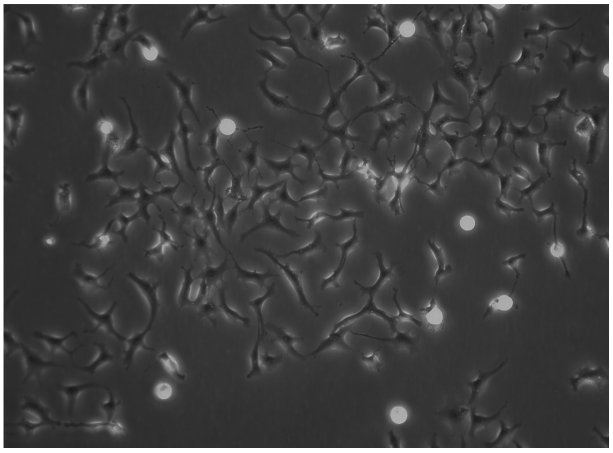
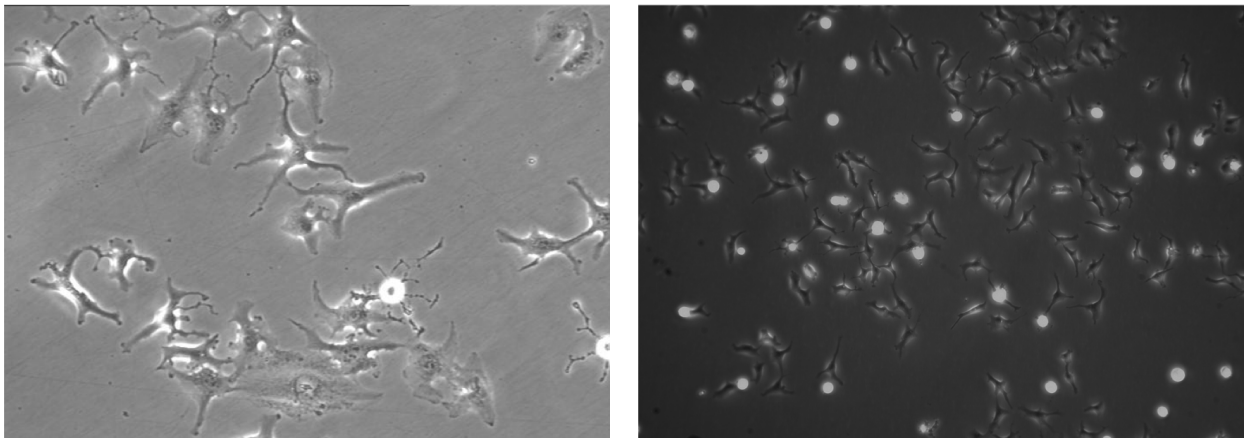


Figure 2: Example of an image presenting a lot of outgrowths and cell overlaps



(a) Restricted field and lower resolution

(b) Large field and high resolution

Figure 3: Example of a two sequences with different characteristics

2 State of the art

The first part of this project was to find out what has already been done on the topic of cell segmentation using convolutional neural networks and in particular on phase contrast microscopy. The main sources used for this project are detailed below.

DeepCell [1]

This article was the starting point of the project, DeepCell gave very promising segmentation results on multiple cell imaging methods. However the neural network was not trained for phase contrast images like those in our dataset. DeepCell has been re-implemented by multiple teams but none to accommodate this type of cell imaging. The maintained implementation can be found at `https://github.com/vanvalenlab/deepcell-tf`.

Mask-RCNN [2]

This is a powerful CNN structure, implemented by Matterport, that was trained using the MS COCO dataset to segment images. It can be modified to be applied to a variety of problems to extract regions of interest and create segmentation masks.

Usiigaci [3]

This Japanese team provides a cell segmentation and tracking software for phase contrast images, knowing the difficulties of migrating cells. Usiigaci has a very similar aim to this project and therefore was the main source used. We will also use the weights of their trained CNN to make our own predictions. The Usiigaci implementation is available on GitHub: `https://github.com/oist/Usiigaci`. The CNN uses the Mask-RCNN structure.

3 Implementation

3.1 Neural network

Because there were no corresponding masks to the images in the dataset, training a neural network from scratch based only on the data available was not a desirable option. Thus the Usiigaci project's trained network was a golden opportunity. The network had been trained on similar images and would hopefully make good predictions. However the Usiigaci implementation did not apply perfectly to the requirements of our dataset. Consequently the structure of the network was re-implemented, starting from the Mask-RCNN implementation, and custom scripts were created to make predictions.

The neural network uses the Mask-RCNN [2] structure, which uses a combination of neural networks to successively find regions of interest, bounding boxes, and masks for the specified output classes. More precisely it uses a Feature Pyramid Network (FPN) to extract features from the image and a ResNet101 to create the masks. The output layer is defined to only detect a single class of objects: the cells. The output image extracted from the prediction is a labelled mask of the cells on the input image.

The network was pre-trained using the MS COCO dataset, then trained additionally by the Usiigaci team on their own set of phase contrast images of cells.

Scripts were created to automatically load images, make predictions with the neural network and store the resulting masks. The encoding of the input images is normalised to be identical to the training set.

The padding system provided by the Mask-RCNN implementation is used to accept images of different sizes.

3.2 Post-processing

Three trained networks are made available by the Usiigaci team and the resulting prediction can be combined to create a more refined output. This uses *intersection over union* (IoU) to merge cells from different masks.

The resulting mask were satisfactory but often contained single cells that were divided between two or more labels in the mask. To merge neighbouring labels, a dilution of the mask of a cell is made and the intersection between the diluted mask and the other cells is computed. If the intersection is greater than a threshold, the labels are merged.The figure 4 shows an example of this label merge.

(a) Multiple labels for a single cell (b) Merged neighbouring label

Figure 4: Merging of multiple labels for a single cell

Because the mask of each image of a sequence is predicted independently by the CNN, the label assigned to each cell changes between each frame of the sequence. To track the cells between masks we compare cells between consecutive frames and assign the same label value to the best matching cells.

4 Results

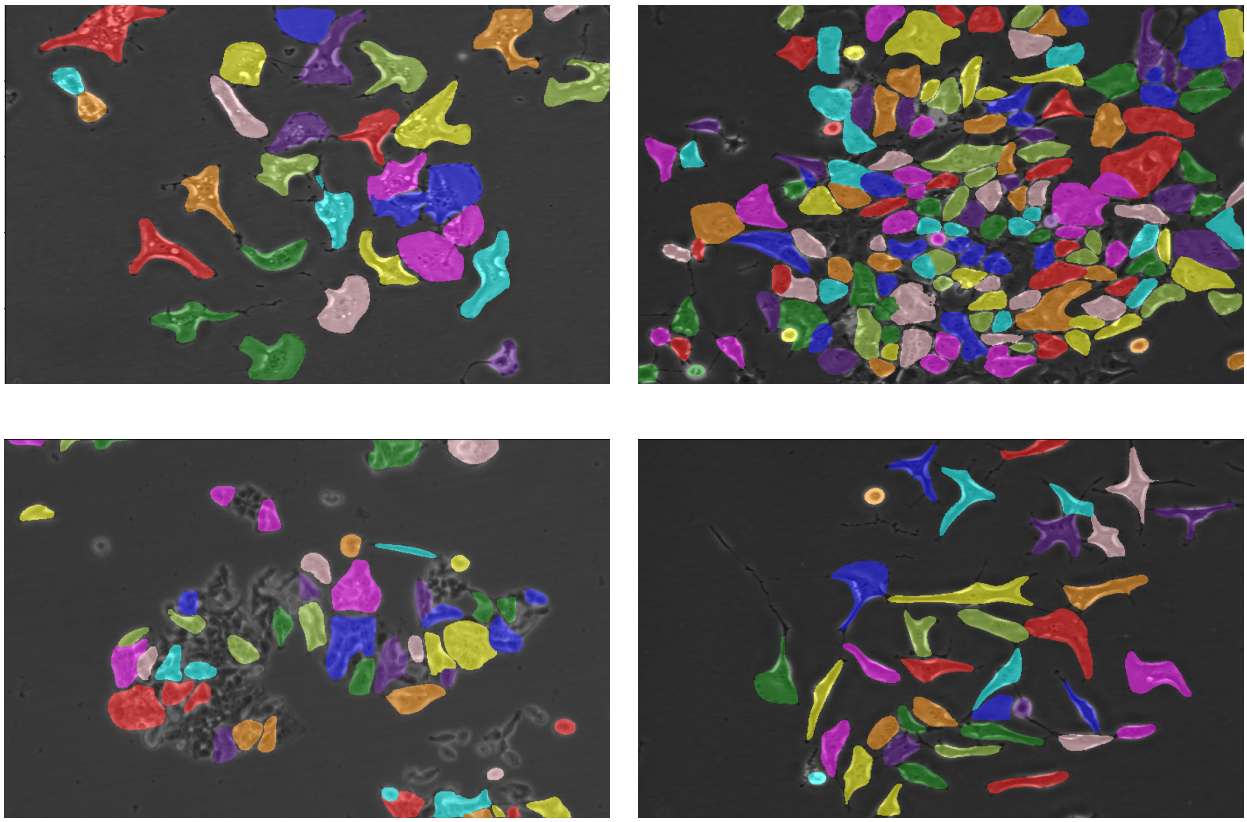


Figure 5: Resulting masks superimposed on the original images

Figures 5 and 6 display some of the results obtained for a diversity of sequences. The segmentation is of good quality unless the images are blurry or contain very densely clustered cells.
Some cells are undetected by the neural network, in particular cells that are in mitosis and cells that got out of focus. It should also be noted that the thin outgrowths are not always part of the mask of the cell. When the cell density is high, neighbouring cells are often merged together as a single label by the CNN, while others are evaluated as background. Images with large clusters like the bottom left images of figures 5 and 6 are not well segmented.

Completely segmented and tracked sequences can be viewed on this Google drive[1]. Viewing entire sequences shows that when a cell is not detected in a frame of the sequence, its tracking stops, which causes a change of the label value in the rest of the sequence.

[1] https://drive.google.com/drive/folders/1D--_sGyH7pUPGUmQ6D2EVig2j2tbrJsf?usp=sharing

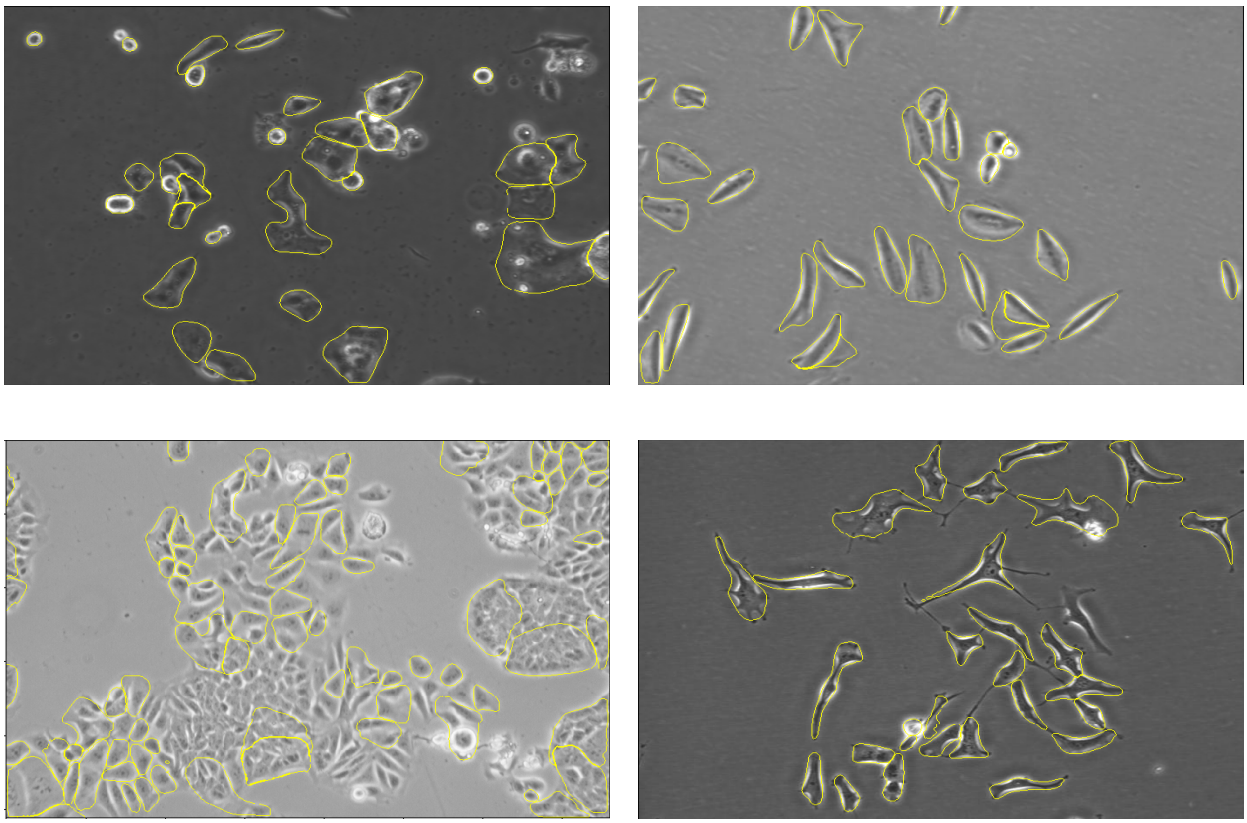


Figure 6: Resulting mask borders on the original images

5 Performances

The performances of the segmentation and tracking pipeline are good enough to make predictions on phase contrast microscopy sequences within reasonable time. In table 1 are some standard execution times when making predictions on images of size 700x500 pixels, on a computer of the LISA lab.

Step	**Time per image**	**Time for a sequence**[1]
Mask prediction	1s to 3s	15 to 45 minutes
Mask merge	1.5s to 6s	25 to 90 minutes
Post-processing	1s to 4s	15 to 60 minutes

Table 1: Running speed of the program

However when making predictions on images that have a larger size and contain a lot of cells (e.g. large field images), choke points appear during the mask merge step and the post-processing steps. However

[1] A sequence of around 1000 images

the sources of these slow downs (up to 30s per images) can be identified and potentially optimised.

Merging the outputs from the three trained networks is a very costly process for large images containing a lot of cells because it uses intersection over union (IoU) between every cell of the masks to evaluate if they should be merged.

1. The number of loops is the number of cell comparison which is about the number of cells on the image squared.

2. The cost of computing each loop increases with image size (more pixels).

A possible optimisation would be to use the cell positions to limit the number of comparisons. Alternatively we could make the sum of the three binary masks and use a watershed or an edge detection method to re-segment the cells.

The post-processing steps[2] face identical problems, however the number of comparisons made during these steps is smaller than during the mask merging. Similarly, a possible optimisation would be to reduce the number of cell comparisons by exploiting their positions.

6 Conclusion

Convolutional neural networks definitely have great potential to automate the segmentation of phase contrast images. The implementation of this project[3] gives some promising results that would be greatly improved by some additional fine tuning and possibly more training of the neural network. The time required to make prediction is reasonable for most sequences and some optimisations may allow more computationally heavy sequences to be segmented.

[2]Reforming divided cells by merging neighbouring masks, and tracking cells between frames

[3]Available at `https://github.com/LAlmasy/Phase`

References

[1] David A. Van Valen, Takamasa Kudo, Keara M. Lane, Derek N. Macklin, Nicolas T. Quach, Mialy M. DeFelice, Inbal Maayan, Yu Tanouchi, Euan A. Ashley, and Markus W. Covert. Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLOS Computational Biology*, 12(11):1–24, 11 2016.

[2] Waleed Abdulla. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. `https://github.com/matterport/Mask_RCNN`, 2017.

[3] Hsieh-Fu Tsai, Joanna Gajda, Tyler F.W. Sloan, Andrei Rares, and Amy Q. Shen. Usiigaci: Instance-aware cell tracking in stain-free phase contrast microscopy enabled by machine learning. *SoftwareX*, 9:230 – 237, 2019.