Stem Cells Segmentation and Transfer Learning

Fan Yang
Student ID: 931031

Project Supervisor
Dr. Xingjun Ma

COMP90019
Distributed Computing Project
Conventional research project
Total Credit: 25 Points

1 Abstract

There is large consent that cell segmentation can be successfully implemented by training deep neural network, especially by convolutional neural network (CNN). In this project, one kind of CNNs, U-Net, is trained and used to mask the cell on another photo, so the outline of the cell can be drawn. The training data sets fitted in this project are photos of Human lung carcinoma cells and they are trained by different popular gradient-based optimizers. As the accuracy of testing result of lung carcinoma cells are higher than 99 per cent, transfer learning is applied to the stem cells to examine whether the model can be also utilized in segment the target domain, the stem cell photos. It is found that in this pretrained model, the accuracy of stem cells segmentation is not as high as assumed before the experiment. Therefore, this paper will demonstrate the method of training U-Net and transfer learning, and then analyse the reason that the results of transfer learning was produced.

2 Acknowledgements

I would like to express my deepest appreciation for the immaculate support and guidance given by Dr. Xingjun Ma and my teammates, Yinghao Zhu and Dong Jia.

3 Contents

Contents

1 Abstract **2**

2 Acknowledgements **3**

3 Contents **4**

4 Introduction **6**
4.1 Problem 6
4.2 Project Objectives 6

5 Background **6**
5.1 Related work 7
5.2 Data set 8

6 Methodology **8**
6.1 Prepare for training set 8
6.1.1 CellProfiler 8
6.1.2 Generate ground truth 9
6.1.3 Enhance Contrast 10
6.1.4 Normalization 10
6.2 Design 11
6.2.1 U-Net 11
6.2.2 Transfer learning 11
6.3 Implementation 12
6.3.1 Data Augmentation 12
6.3.2 Training U-Net 13
6.3.3 Fine-tuning 13
6.4 Testing 13

7 Result **15**
7.1 U-Net 15
7.2 Testing performance 15

8 Discussion **16**
8.1 Optimizer comparison 16
8.2 Transfer Learning with Stem Cells 17

9 Conclusion and Future Work **19**

4 Introduction

4.1 Problem

In medical field, how to describe the features of stem cells cultured by the laboratory is a critical issue. It is not only related to the judgement of quality of the stem cells, which will affect the economic benefits, but also has great impact on safety when they are applied in medical fields. To extract and select the features of the stem, the first step is to segment the stem cells on the bright field photos. However, due to the limited training data set of stem cells segmentation, the model of other types of cells is trained firstly, and then diverse transfer learning methods are utilized to identify and segment the stem cells on the stem cells data sets.

4.2 Project Objectives

This project aims to segment stem cells on photos provided by laboratory and considering the limited data sets, there are two objectives on each stage.

1. Train a model with Human lung carcinoma cells to get the high accuracy of the cells segmentation result.

2. Transfer learning applying the model trained above on the stem cells photos to segment the photos.

5 Background

Recently Deep Convolutional Neural Network (DCNN) has great performance on various image recognition tasks. Bright field images are easily captured, but they are much more difficult to segmented compared to fluorescent images because of low contrast between the outline of cells and background. [9] Therefore, there are several studies are analysing diverse cells using bright field images.

5.1 Related work

In a study which does not utilise DCNN, Wang et al. designed an algorithm to segment bright field yeast cell images. The reason why non-DCNN is utilised is that training data set is required. In their paper, they first pre-processed the images by Histogram equalization and Gaussian filtering to enhance the contrast and reduce noise. Then they tried to get cleaner edges and conducted morphology operation for the following seed points detection. After screening of the seed points, they designed an edge tracing algorithm to connect seed points so that the outline of yeast can be drawn. Although the algorithm shown in this study is robust, i.e. they considered the crowded yeast cells and can draw the boundary between them, the algorithm is still sensitive to noise on images. Moreover, the yeast cell is almost oval while stem cells are irregular shape, so this algorithm is hard to applied in this situation. However, the method they used to pre-process images can be reference in the course of pre-processing.

Kimmel et al. [3] utilised CNN to classify skeletal muscle stem cells based on motility behaviours. The structure of CNN they applied is that the images passed through 4 convolutional layers with max pooling, and then were input into 2 fully connected layers along with dropout, which relied less on individual units and reduced the probability of overfitting. Finally, to classify the images, a fully connected layer with softmax activation is used. This model is trained by stochastic gradient descent (SGD) with momentum and decay. In the experiment, the authors fit 3000 samples per class into the model and achieved 84 per cent accuracy while achieved 96 accuracy with 13500 samples per class. Although in the project, CNN was capable of classifying different types of cells, it required a great number of training data sets, which is restricted in this paper. Nevertheless, the authors proved that the CNN models can be trained effectively, i.e. after 30 epochs they can return a high-accuracy result, which means applying CNN on analysing images are feasible.

Another supervised classifier was constructed by Ciresan et al. [1] The architecture of their Deep Neural Network (DNN) was based on a sliding window cantered on a pixel. The window should be a square whose edge should be an odd number for its symmetry. Then the probability that a pixel belongs to a class was computed so the features of the cell can be localized. When the model was trained, training data set was augmented by mirroring or ro-

tating. It is evaluated that the result of their experiments were quite high at that time but the model was a bit slow.

There is another special CNN named U-Net created by Ronneberger et al. [6] The biggest advantage of this CNN is that it only requires a few training data set but achieve a higher accuracy. Moreover, in their experiment, they segmented "PhC-U373" data set containing Glioblastoma-astrocytom U373 cells and "DIC-HeLa" data set containing HeLa cells with high performance. This two data sets are totally different. Therefore, U-Net can be applied in various kinds of cells.

5.2 Data set

In this paper, the data sets consist of two different kinds of cell images. The first data set is Human lung carcinoma cells come from the project used by Sajith et al. [7]while the other is stem cells captured from laboratory in China.

6 Methodology

The whole architecture of the model used in this project is based on U-Net. Additionally, transfer learning methods are also applied to fit the target images, the stem cell images.

6.1 Prepare for training set

The ground truth of the stem cell images is limited as they are difficult to fetch. Therefore, the model is firstly pretrained by Human lung carcinoma cells.

6.1.1 CellProfiler

CellProfiler is a useful open-source software which allows the medical images analyser to generate a large quantities of measure phenotypes from the original images automatically. It provides an interface for users to form a pipeline

so that the features in biological images can be identified and the objects, such as cells, can also be measured.

6.1.2 Generate ground truth

According to Sajith et al. [7]three different types of images are placed into pipeline of CellProfiler to create a high-quality and clear ground truth. The three kinds of images are bright-field images, fluorescent images of nuclear and that of cytoplasmic captured at the same time and same place. By comparing these three kinds of images, CellProfiler can classify the pixels whether belong to a cell. After that, the overlay and segmentation images of the cell are generated, and the segmentation images are what this project need. The procedure will be shown at figure 1.

In this project, 22 of each kind of images of Human lung carcinoma cells are used. After they are processed by CellProfiler, 22 original images and ground truth images are gained. Finally, the sizes of the images are 2034x2034 but the size needed in the project is 512x512. As the resolution ratios of these images are too large to be fed to the model described below, the images are cropped into 16 images whose sizes are 512x512. Therefore, totally 352 original images and ground truth images are generated after this step. The result is shown in figure 2.

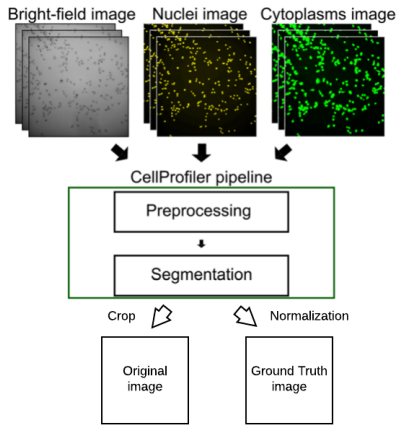

Figure 1: Pre-processing procedure

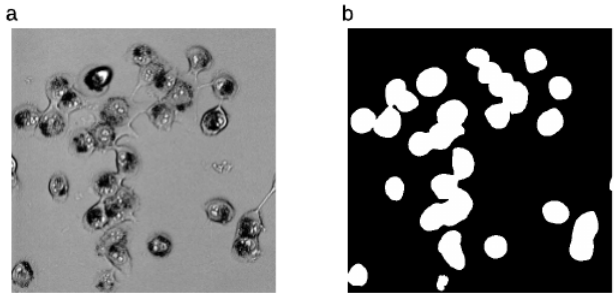

Figure 2: Example image and ground truth image after pre-processing

6.1.3 Enhance Contrast

As described above, Kimmel et al. [3] pre-processed their images by enhancing contrast and thresholding. However, through several experiments, the images enhanced contrast are good enough to conduct the experiment afterwards.

To enhance contrast, the method in Pillow, which is a powerful Python imaging library, is used. Due to extremely low contrast of the original images, the contrasts of these images are enhanced by eight times.

Additionally, contrast of the images of stem cells are also enhanced for the following test.

6.1.4 Normalization

The ground truth images gained before are not only black and white but include grey pixels which are also identified as cells in these images. Therefore, normalization, i.e. transferring all non-zero pixel in the grey ground truth image into 255 which represents totally white, is required to generate the mask set.

6.2 Design

As the goal of this project is to segment the stem cells, the two areas of the project are the architecture of U-Net and the transfer learning.

6.2.1 U-Net

The CNN utilised in this project is based on U-Net. The architecture is is illustrated in Figure 3. As shown in the figure, the network architecture includes a path from top left to bottom and the other path from bottom to top right. The left path is called contracting path while the right one is named expansive path. The contracting path consists of two main parts, two 3x3 convolutions layers followed by a 2x2 max pooling operation with stride 2 for down-sampling. After down-sampling step, the number of feature channels are doubled. The expansive path is similar to mirror of the contracting path. Each step in expansive path does a 2x2 convolution called up-convolution which halves the number of feature channels. After that, the feature channels are concatenated with corresponding channels from left. Then two 3x3 convolutions layers are followed. Finally a 1x1 convolution is applied to classify the 64 feature vectors into two classes, i.e. "0" represents background while "1" represents cells. Totally there are 23 convolutional layers and four max pooling layers.

The main difference between U-Net explained by Ronneberger et al [6]. and this network is that their convolutions are unpadded whereas this network use convolutions where output size is the same as input size. This avoid feature vector loss during convolution.

6.2.2 Transfer learning

Although the number of stem cell images are large, it is hard to generate ground truth of these images. This is not only because the noise in these images is difficult to remove, but stem cell images of nuclear and that of cytoplasmic are lacked. Therefore, the training set of stem cells is limited. According to Yosinski et al. [10], For those deep neural networks, there are some common phenomenon that as the layers become deeper, the more general features can be extracted in a range of datasets. Hence, in the case of U-Net, which is also a deep neural network, it is possible to use U-Net to segment stem cells in images.

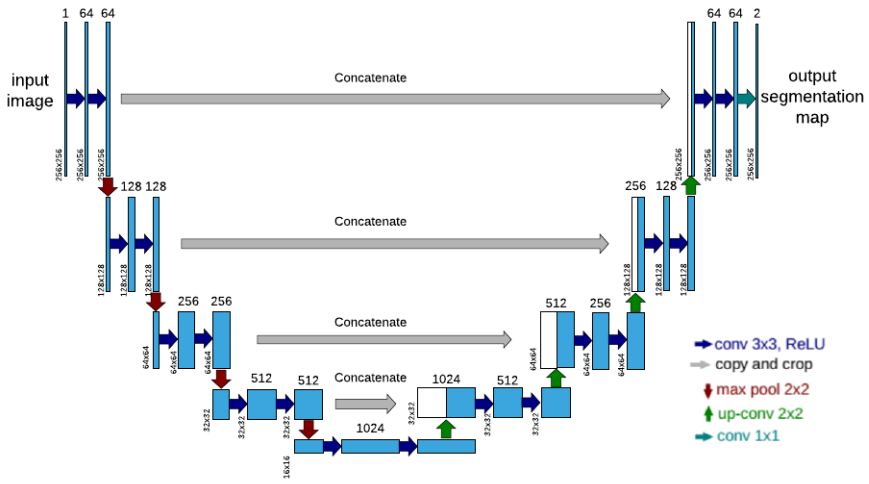

Figure 3: Architecture of U-Net

Summarized by Pan [5], as the images of stem cell are similar to images of lung carcinoma cells, mapping-based approach can be utilised. However, this project lacks ground truth of stem cell image, it is more like unsupervised transfer learning. Hence, the approach used in this project combines tried to combine the two types of images to mix the feature set. If the two types of images can be identified at the same time, the pre-trained U-Net may be able to segment the images.

6.3 Implementation

6.3.1 Data Augmentation

Data augmentation is a critical method to generate a large training set when only few images exist in training set. Moreover, it can assist the network learn to be robust.

In this project, the images are firstly transferred to grey image and the sizes are adjusted to 256x256. Then data augmentation of the images includes the operations, rotation, shift and flip, which are all randomly performed in a

range to improve the robustness of network.

6.3.2 Training U-Net

The U-Net model illustrated in figure 3 is built through Keras. As explained before, there are totally are 23 convolutional layers where 22 convolutional layers are activated by ReLU while the last layer is activated by sigmoid to do the classification and four max pooling layers. Besides, drop-out layers are added in the bottom two layers to reduce the chance of overfitting.

As described in 6.1.2, the training set contains 352 images. However, due to some low-quality and blank images existing in the data set, the whole set would not be used. Moreover, with data augmentation, only a part of images in the original training set are sufficient to train this model. Therefore, 30 images and corresponding segmentation mask maps are randomly selected from the training set and each image is augmented 10 times.

The loss function in this model is binary cross-entropy as the outputs are binary. To reduce the loss, two optimizers are applied in this model, namely Stochastic gradient descent (SGD) with Momentum and Nesterov Acceleration and Adam.

6.3.3 Fine-tuning

Fine-tuning is the process in which parameters of a model are adjusted slightly and precisely based on the parameters in pre-trained model. It is widely used for image segmentation tasks which can not only fit the training data rapidly, but also prevent ovet-fitting [2]. In this project, fine-tuning is used to improve the performance of U-Net.

6.4 Testing

The testing sets include three type of data. The first data set is human lung carcinoma cells from the data set described in 6.1.2 except the training set. The second data set is the stem cell images which are pre-processed by enhancing contrast. The last data set is based on transfer learning where two types of images are concatenated with diverse weights [8]. In this project, the weights are selected as 0.2, 0.5 and 0.8.

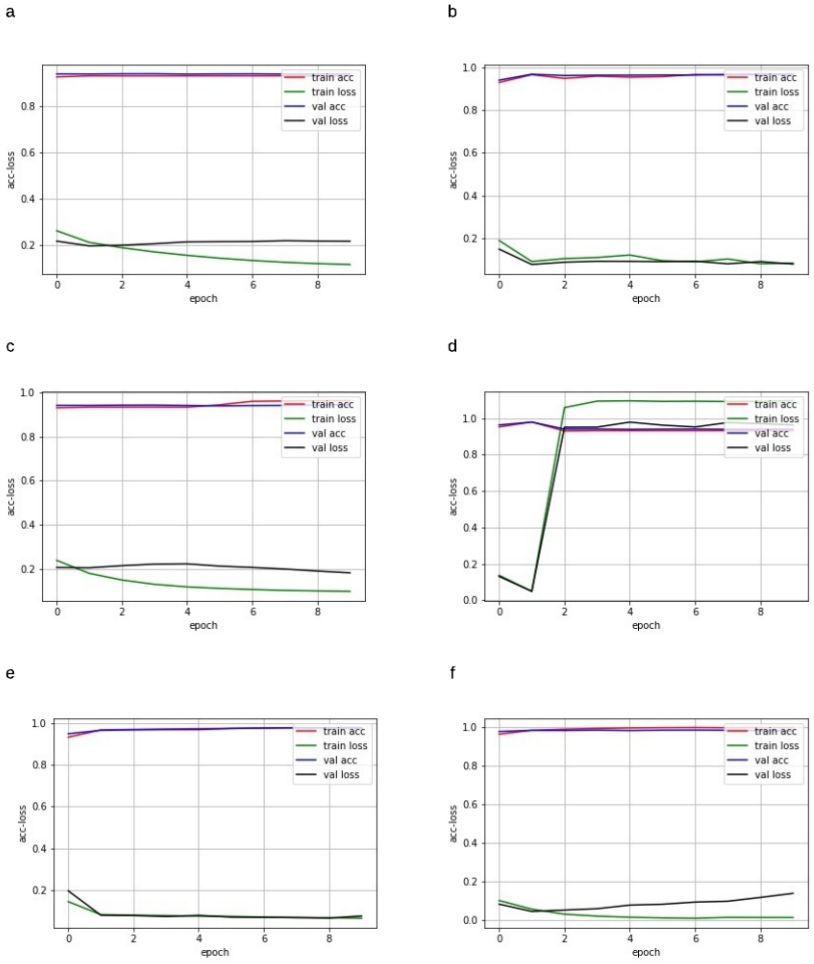

Figure 4: Training result of U-Net. Figure 4.a SGD optimizer with learning rate 0.0001, momentum 0.0, decay 0.0. Figure 4.b SGD optimizer with learning rate 0.01, momentum 0.0, decay 0.0. Figure 4.c SGD optimizer with learning rate 0.0001, momentum 0.5, decay 0.0. Figure 4.d SGD optimizer with learning rate 0.0001, momentum 0.99, decay 0.0. Figure 4.e SGD optimizer with learning rate 0.0001, momentum 0.99, decay 1e-7. Figure 4.f Adam optimizer with learning rate 0.0001

7 Result

7.1 U-Net

The model is trained for 10 epochs, as the loss after 10 epochs is extremely low and the accuracy is near 1. The batch size is eight which is limited by the memory in GPU.

In this project, the first optimizer, SGD, with different parameters is used to find the best performance. Besides, the second optimizer, Adam, is only set the learning rate as 0.0001. The results are shown in figure 4.

The result of fine-tuning has a high performance at first where the accuracy is around 98 per cent. Just after a few epochs, the accuracy increases to 99.5 percent, which is faster than the previous training.

7.2 Testing performance

By using model having the best performance in 7.1, which is Adam, the three types of data sets are tested. The results are shown in figure 5, figure 6 and figure 7.

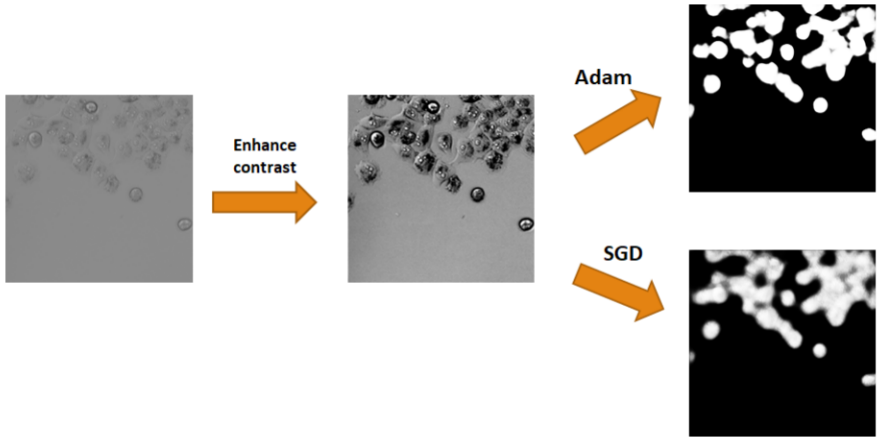

Figure 5: Test results of Human lung carcinoma cells by SGD and Adam

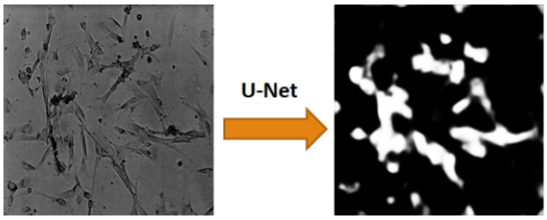

Figure 6: Test results of stem cells by SGD and Adam

8 Discussion

8.1 Optimizer comparison

There are four values used to measure the results of the model, namely training loss, validation loss, training accuracy and validation accuracy. The goal of training is to select a model where both types of loss decrease smoothly and rapidly and both kinds of accuracy are gradually closed to 1.

First of all, compared with SGD with lower learning rate, SGD with the learning rate which is 0.01 decrease the loss more rapidly (figure 4.a). However, as the number of epochs increases, the training loss fluctuates around a value, which means that the higher learning rate may cause over-fitting. As each epoch the network trains 300 images, which is sufficient, a lower learning rate can be applied in the SGD. (figure 4.b)

Secondly, two different momentums 0.5 and 0.99. As shown in figure 4.c and figure 4.d, result of SGD with momentums 0.5 is much better than that of 0.99. Obviously, on one hand, after the first epoch the loss in 0.99 momentums suddenly increases to 0.9, leading to extreme adjustments of parameters. On the other hand, the curve of train loss in 0.5-momentums-SGD decrease more smoothly. However, the validation losses of both SGD are increasing after several epochs. To solve the problem of over-fitting, the decay of learning rate is added in 0.99-momentums-SGD. From the figure 4.e, It can be seen that both training loss and validation loss are reduced quickly at the first few epochs and after that they are closed to zero. Meanwhile, both training accuracy and validation are around 99 per cent. However, as shown

in figure 5, the result of SGD is not clear, although the loss is low. This is because after a large number of training epoch, the learning rate is reduced to extremely low, which is not enough to adjust the parameters.

Hence, it is difficult to adjust learning rate manually. According to Kingma et al. [4] the more computationally efficient algorithms Adam is utilised. It combines the ability of AdaGrad that tackles sparse gradients, and the ability of RMSProp which deals with non-stationary objectives. As a result, although the validation loss increases a few after 10 epochs are trained, the segmentation maps are clearer. Therefore, if Adam optimizer is applied here, the model will be trained more efficiently. (figure 4.f)

Finally, with the method of fine-tuning, the model can be trained faster, which means pre-trained network does not only substantially reduce training time, but also assists to prevent over-fitting.

8.2 Transfer Learning with Stem Cells

As the model trained by Adam optimizer successfully segments the human lung carcinoma cells, this model is applied to stem cells.

However, as shown in figure 6, the result of segmentation of stem cells is not as perfect as that of human lung carcinoma cells. It only masks outline of cells roughly. This is mainly because the noise cannot be removed after pre-processed. Even the image shown in figure 6 has less noise, the result of segmentation is quite poor.

Table 1 shows the loss and accuracy of testing with concatenation of human lung carcinoma cell image and stem cells. As the weight of human lung carcinoma cells decreases, the accuracy decreases.

When the weight of 0.8 of Human lung carcinoma cell images is applied, these images are similar to the original images added a background of stem cell images. The accuracy of this image is around 98 per cent, which means in this case the noise of the stem cell has little effect on the result. Similar situation occurs in weight of 0.5. Although the accuracy decreases a bit, the images can also be segmented. This proves that U-Net is general enough to conduct segmentation.

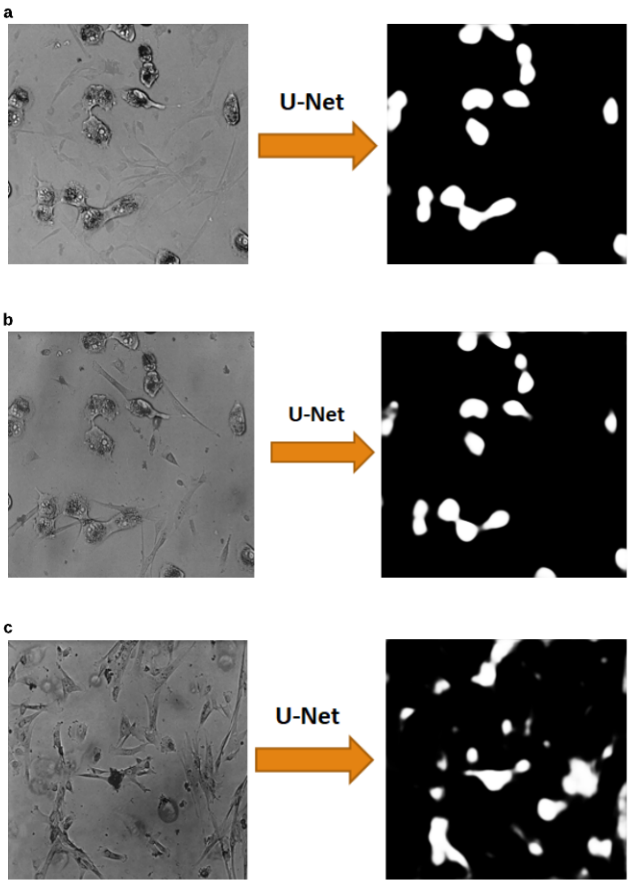

Figure 7: Test results of concatenation of Human lung carcinoma cell image and stem cells. Figure 7.a Image in which weight of lung carcinoma cell is 0.8. Figure 7.b Image in which weight of lung carcinoma cell is 0.5. Figure 7.c Image in which weight of lung carcinoma cell is 0.2

However, when the weight reduces to 0.2, few human lung carcinoma cells can be segmented while some stem cells can also be segmented. The accuracy is around 90 percent, but the loss is quite high as over 0.5. This means that the network cannot suppress interference of noise in stem cell image but on the other hand, the network can deal with the diverse cells images if the noise is removed.

Overall, the transfer learning does not succeed in this project. The first reason is that unlike the oral lung carcinoma cells, the stem cells have irregular shape. From figure 6, the sharp part of cells is segmented into smooth boundary so U-Net still needs several ground truth images to fine tune the network. Another reason is that these bright-field stem-cell images include noise generated by shadows. In this case, when the contrasts of images are enhanced, these shadows would form black field on images, which would be identified as cells.

Weight	Loss	Accuracy
0.8	0.103	0.980
0.5	0.146	0.970
0.2	0.528	0.907

Table 1: Test results of concatenation of Human Lung Carcinoma Cell image and stem cells

9 Conclusion and Future Work

The U-Net is a perfect CNN which is efficient. It required small number of data set and the loss is reduced rapidly so after a reasonable training time, this network can deal with the segmentation tasks much more easily. In this project, it also shows its ability that extract general features in diverse kinds of medical cells images, though it is better to fine tune the network with ground truth images of specific cell, i.e. stem cells in this project.

The reason why the transfer learning does not succeed is that not only the training set of stem cells is limited, but the noise of images is mixed with the cells.

Hence, in the future, removing noise on images will be the first tasks to be finished in order to generate ground truth of the cells. Then with the assist of various transfer learning methods, such as Domain-Adversarial Neural Networks or mapping-based approach, these stem cell images are possible to be segmented as perfect as human lung carcinoma cells. Besides, based on the result of segmentation, several features of cells will be measured and extracted to assist cultivation of cells in medical fields.

Reference

[1] Dan Ciresan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pages 2843–2851, 2012.

[2] Vladimir Iglovikov and Alexey Shvets. Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. *arXiv preprint arXiv:1801.05746*, 2018.

[3] Jacob C Kimmel, Andrew S Brack, and Wallace F Marshall. Deep convolutional neural networks allow analysis of cell motility during stem cell differentiation and neoplastic transformation. *bioRxiv*, page 159202, 2017.

[4] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[5] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[7] Sajith Kecheril Sadanandan, Petter Ranefall, Sylvie Le Guyader, and Carolina Wählby. Automated training of deep convolutional neural networks for cell segmentation. *Scientific reports*, 7(1):7860, 2017.

[8] Annegreet van Opbroek. Transfer learning for medical image segmentation. 2018.

[9] Linbo Wang, Simin Li, Zhenglong Sun, Gang Wen, Fan Zheng, Chuanhai Fu, and Hui Li. Segmentation of yeast cell's bright-field image with an edge-tracing algorithm. *Journal of biomedical optics*, 23(11):116503, 2018.

[10] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328, 2014.