

Assignment 1 Report

Chenqiu Zhao(zhao.chenqiu@ualberta.ca)

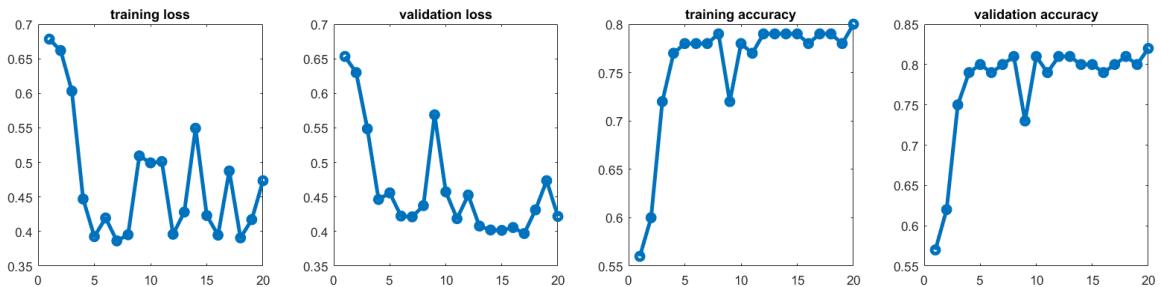
Question 1



There are 113 images in Microscopy dataset, which is download from the link provided in the assignment document. In particular, 50 images are labelled as positive and 63 images are classified as negative. Positive images and negative images are randomly split into 3 parts, training(70%), validation(20%) and test(10%) sets respectively, to guarantee that the proportion of positive and negative samples keep invariant in these three sets. For a particular image, patches with size of 60×60 are extracted, in which the position of patches are randomly generated. With the consideration of memory cost, 200 patches are extracted from one image, as shown in the left figure. Totally, 15800 patches are utilized for training.

Since, the assignment is a typical classification problem, the network architecture is devised as a classification network, which is similar to LeNet. The comprehensive details of the network architecture is shown the in the right table. During training, 20 epochs are set as the maximum and 0.001 is used as the learning rate. It should be noted that our program is ran on a remote server instead of a local PC, so we extract the training information and draw the plots by ourself. The plots of training loss, validation loss, training accuracy and validation accuracy are shown as follows:

| Type | Filters | Layer size | Data size |
|-----------------------|---------|-------------------------------|--------------------------|
| Input | | | $60 \times 60 \times 3$ |
| Convolution | 20 | $10 \times 10 \times 3$ | $51 \times 51 \times 20$ |
| Rectified linear unit | | | $51 \times 51 \times 20$ |
| Max pooling | | 2×2 with stride of 2 | $25 \times 25 \times 20$ |
| Convolution | 50 | $10 \times 10 \times 20$ | $16 \times 16 \times 50$ |
| Rectified linear unit | | | $16 \times 16 \times 50$ |
| Max pooling | | 2×2 with stride of 2 | $8 \times 8 \times 50$ |
| Convolution | 500 | $8 \times 8 \times 50$ | $1 \times 1 \times 500$ |
| Rectified linear unit | | | $1 \times 1 \times 500$ |
| Convolution | 2 | $1 \times 1 \times 500$ | $1 \times 1 \times 2$ |
| Softmax | | | |



After training, the test set is utilized for evaluating the proposed network. There are 11 images in the test set. In particular, 5 images are labelled as positive and 6 images are labelled as negative. During testing, 8 images are classified correctly, the classification accuracy is $\frac{8}{11} = 0.7273$. However, in a real application, the significance of positive and negative samples are different. In order to propose a fair and comprehensie evaluation of the proposed network,

Re , Pr and Fm metrics are utilized for evaluation. For the completeness of this report, the definition of these metrics are briefly introduced. Re and Pr are measures of completeness and accurateness, respectively. Fm is a combination of Re and Pr . These metrics are defined as follows:

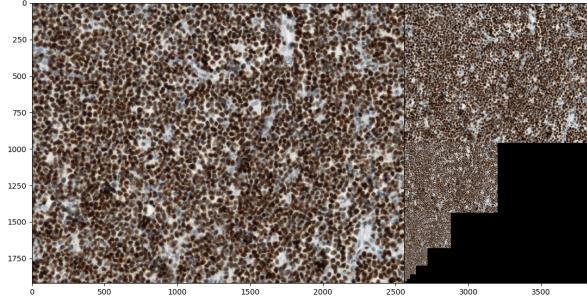
$$Re = \frac{TP}{TP+FN}, Pr = \frac{TP}{TP+FP}, Fm = \frac{2 \times Pr \times Re}{Pr + Re},$$

where TP and FP are True Positive and False Positive. True denotes that the result of this classification is correct, while False means otherwise. Thus, TP means that the result of the detection is positive as well as being the groundtruth.

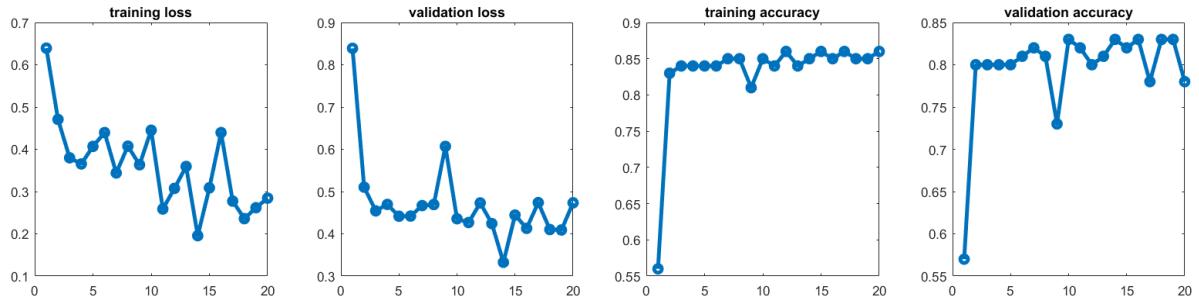
Moreover, since the training set and test set are randomly split, it is possible that the evaluation results are different for different repeating experiments. In order to propose a fair evaluation, we repeat the experiment 10 times, and the average performance is proposed as the final performance of the proposed network. The results is shown as follows:

| Repeating Experiments | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Average |
|-----------------------|----|------|-------|-----|-------|-------|-------|-------|-----|-------|-------|---------|
| Single Resolution | Re | 0.8 | 0.6 | 1.0 | 0.6 | 0.8 | 1.0 | 0.8 | 0.8 | 1.0 | 0.8 | 0.8190 |
| | Pr | 1.0 | 1.0 | 1.0 | 1.0 | 0.667 | 0.833 | 1.0 | 0.8 | 0.833 | 1.0 | 0.9001 |
| | Fm | 0.89 | 0.749 | 1.0 | 0.749 | 0.727 | 0.909 | 0.889 | 0.8 | 0.909 | 0.889 | 0.8513 |

Question 2



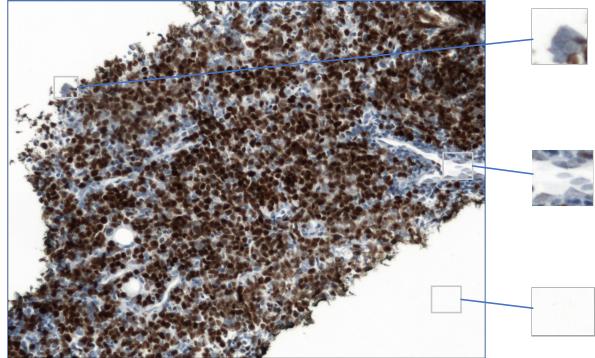
same parameters with the single resolution experiments. The plots of training loss, validation loss, training accuracy and validation accuracy are shown as follows:



Similar with the evaluation experiment proposed in last section, the evaluation of network with multiple resolutions input is also repeated for 10 times. The comparison between the network with single resolution input and multiple resolutions input is shown in the above table. All the source codes and running results are available online <https://github.com/zhaochenqiu/courses/tree/master CMPUT617/Assignment1/program>.

| Repeating Experiments | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Average |
|-----------------------|----|------|-------|-------|-------|-------|-------|-------|-----|-------|---------------|
| Single Resolution | Re | 0.8 | 0.6 | 1.0 | 0.6 | 0.8 | 1.0 | 0.8 | 1.0 | 0.8 | 0.8190 |
| | Pr | 1.0 | 1.0 | 1.0 | 1.0 | 0.667 | 0.833 | 1.0 | 0.8 | 0.833 | 1.0 |
| | Fm | 0.89 | 0.749 | 1.0 | 0.749 | 0.727 | 0.909 | 0.889 | 0.8 | 0.909 | 0.8513 |
| Multiple Resolutions | Re | 0.8 | 0.8 | 1.0 | 1.0 | 0.8 | 0.8 | 0.8 | 1.0 | 1.0 | 0.9000 |
| | Pr | 1.0 | 1.0 | 0.833 | 1.0 | 0.8 | 0.8 | 0.8 | 1.0 | 1.0 | 0.9233 |
| | Fm | 0.89 | 0.89 | 0.896 | 1.0 | 0.8 | 0.8 | 0.8 | 1.0 | 1.0 | 0.9076 |

Question 3



In my opinion, the generation of training data is the main challenge of applying deep learning to microscopy over the traditional method. Since the resolution of microscopy image is too large to be input into convolutional neural directly, patches captured from images are utilized as a compromise to be fed into the network for training. Moreover, the size of patches has significant influence on the performance of network, since there may be no completed cell in a patch when the size is small, while the patches with large size will cost a lot of memory. Hence, because the labels of patches are decided according to the name of microscopy image, it is possible that

some patches without any cells are labeled as positive like patches shown in the left figure. These patches play the role of distraction during training network. In contrast, since the method proposed in the paper [1] is a dimensionless method, with the utilization coefficient of variation, which is a dimensionless feature. The cells are detected by local maximum algorithms, and there is no need to consider the generation of training data.

Of course, there are also advantages for applying deep learning to microscopy images classification. As shown in the table above, the performance of deep learning is one of the biggest advantages. Unlike previous traditional methods, the hierarchy architecture of deep learning network makes the model has the ability to learn high-level features which is better than these manual features. Moreover, the deep learning network can automatically learn the features which are suitable for the applications, such as microscopy image classification. From this perceptive, the generality of a deep learning network is better than the traditional method.

Finally, I'd like to discuss another disadvantage of applying deep learning network to microscopy image classification. As we can see, there is not very high similarity between the visual patterns in different microscopy images, since it seems that the size and the positions of cells include randomness. In this condition, it is possible that the patterns in the test set may never appear in training set, and the performance of deep learning network is not promising. Personally, I think it is the main reason that the performance of the proposed network varies during repeating experiments.

References

- [1] Gilbert Bigras, Wei-Feng Dong, Sarah Canil, Raymond Lai, Didier Morel, Paul E Swanson, and Iyare Izevbaye. New myc ihc classifier integrating quantitative architecture parameters to predict myc gene translocation in diffuse large b-cell lymphoma. *Applied Immunohistochemistry & Molecular Morphology*, 26(1):54, 2018.