Detection of proliferation indices from microscopic image for tumour progression analysis

Introduction

Glioblastoma (GBM) is the most aggressive and common type of brain tumour in adults. In classic clinical practices, histopathology images are manually analysed by medical experts or pathologists for diagnosis of disease stage [1]. However, manual analysis can cause some problems:

- Experiences and subjectivity of medical professionals or pathologist can impact evaluation criteria significantly [2].
- pathologists visually navigate and review glass slides or whole slide images (WSI) to detect and analyse malformations in daily work [3]. After evaluating the brain graphics, if tumour existence is disbelieved, the patient's brain biopsy will be activated, which may even consume a month to determine an answer [4].

Objectives

The main goals of this research are:

1. To establish a computer-based system that can detect and classify different types of tumour cells.
2. Finding out the percentage of positive cells to all cells shown in the image.
3. Reducing the costs in viewing time, examination and interference from human factors.

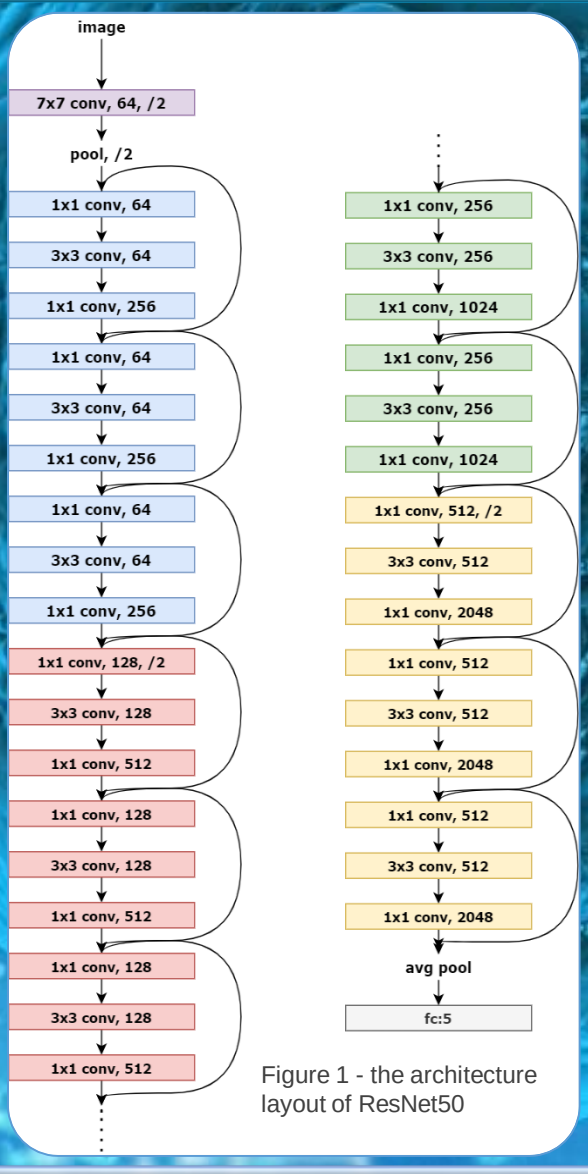

Figure 1 - the architecture layout of ResNet50

Methodology

Manual extraction of training and validation dataset

- Testing samples: 3 whole slide images.
- Training samples: 19 whole slide images. All cells in these images were extracted and divided into five classes: single positive cells (SP), single negative cells (SN), connected positive cells (CP) and connected negative cells (CN). Also, the background (BG) information was collected as training dataset.

Deep Residual Network (ResNet) classification

ResNet50 [5] was chosen to train on the extracted dataset with stochastic gradient descent (SGD) optimizer, momentum of 0.9, learning rate of 0.002, and cross entropy as loss function.

Cell detection and segmentation

While testing images, cells are identified by using sliding window with softer non-maximum suppression (NMS) [6]. In addition, the size of bounding boxes has been configurated in advance. All of connected cells classified by trained model were been split into a single cell based on the watershed algorithm. Finally, the detected cells were counted based on their predicted label.

Evaluation criteria

To assess the performance of our classifier, we used precision, recall and F-score as performance measurement matrices derived from true positive (TP), true negatives (TN), false negatives (FN) and false positive (FP) values. Also, we compare the predicted proliferation index that obtained from detection to ground truth and show the difference as error rate.

Precision, recall, F-score and error rate are defined as follows:

$$Precision = \frac{TP}{TP + FP} \qquad F = \frac{2(Precision \times Recall)}{Precision + Recall}$$

$$Recall = \frac{TP}{TP + FN} \qquad Error\ rate = \frac{DL - Munually}{Munually}$$

Results

To make full use of limited datasets and prevent the over-fit, K-fold cross validation was used for training the model. The cells that extracted from original images were split into five subsets before training. Every subset contains 20% of the total data. We performed 50 epochs with 5-fold cross validation to train the model.

Figure 2 shows the accuracy and loss values in training and validating phase for each 5-fold cross validation. The highest validated accuracy of model reached 93.3%, which was selected as the classifier to recognise cells.

For evaluating performance of the model, we prepared three original-size images (1392×1040 for the size) that are unseen for the model. The result of the assessment has been demonstrated in Table 1.

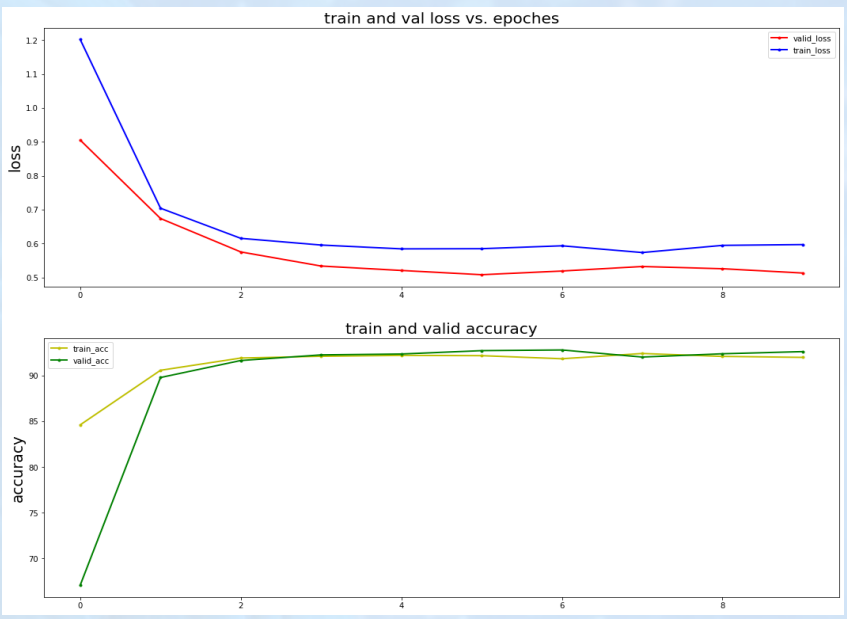

Figure 2 – loss and accuracy vs. the number of epochs

Table 1 – evaluation measure on three test images

Test image no.	Class name	Precision	Recall	F-score	Avg. Precision	Avg. Recall	Avg. F-score
Test image1	SN	0.87	0.90	0.89	0.93	0.78	0.83
	CN	0.90	0.60	0.72			
	SP	0.93	0.93	0.93			
	CP	1.00	0.67	0.80			
Test image2	SN	0.92	0.82	0.87	0.90	0.91	0.90
	CN	0.93	0.87	0.90			
	SP	0.96	0.96	0.96			
	CP	0.80	1.00	0.89			
Test image3	SN	0.80	0.80	0.80	0.91	0.88	0.89
	CN	0.92	0.89	0.90			
	SP	0.93	0.81	0.87			
	CP	1.00	1.00	1.00			

Table 2 – proliferation score concluded by different measures and comparison

Test image no.	Manually	DL	Error rate
Test image1	0.09	0.07	-0.17
Test image2	0.13	0.11	-0.10
Test image3	0.17	0.14	-0.18
Average	0.13	0.11	-0.15

Furthermore, a comparison of calculating proliferation manually and based on our deep learning (DL) method has been indicated in the Table 2. In our experiment, it took an average of 13 minutes for identifying each image.

Conclusion

In this study, we propose an approach combined with deep learning and image segmentation to detect, categorize and count cells in the GBM histological images. Compared with manual annotations, the proposed model shows an acceptable error rate of cells recognition.

However, this model still has some shortages.

- It is time-consuming to detect cells based on sliding window algorithm.
- Also, configured bounding boxes constrict the size of cells that can be detected. If the size of cells exceeds these bounding boxes, the detector may fail.

Further studies should concentrate more on optimization of object detection and segmentation. Some new methods, such as CBNet and YOLO v4 should be considered in future research to improve the performance.

Reference

1. Chen, R., Smith-Cohn, M., Cohen, A.L. and Colman, H., 2017. Glioma subclassifications and their clinical significance. *Neurotherapeutics*, *14*(2), pp.284-297.
2. Yonekura, A., Kawanaka, H., Prasath, V.S., Aronow, B.J. and Takase, H., 2017, September. Glioblastoma multiforme tissue histopathology images-based disease stage classification with deep CNN. In *2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT)* (pp. 1-5). IEEE.
3. Sharma, H., Zerbe, N., Klempert, I., Hellwich, O. and Hufnagl, P., 2017. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Computerized Medical Imaging and Graphics*, *61*, pp.2-13.
4. Anaraki, A.K., Ayati, M. and Kazemi, F., 2019. Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybernetics and Biomedical Engineering, 39(1), pp.63-74.
5. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
6. He, Y., Zhang, X., Savvides, M. and Kitani, K., 2018. Softer-nms: Rethinking bounding box regression for accurate object detection. arXiv preprint arXiv:1809.08545, 2.

Student: Qing Xu

Supervisor: Dr Wenting Duan & Prof Xujiong Ye