Application of Deep Learning in Immunotherapy

Nannan Guo [a], Lichen Zhang [b]

School of Guangdong University of Technology, Guangzhou 510000, China.

[a]nn9110305@163.com, [b]zhanglichen1962@163.com.

Abstract. In recent years, the study of deep learning in medicine has developed rapidly. This article uses a variety of deep neural networks to explore the relationship between DNA-mutated tumors and immunotherapy. This article uses a variety of methods to image preprocess images. Such as: interest area (ROI) extraction, color standardization, cleaning dirty data, and so on. A variety of deep learning models (eg caffe, keras), a variety of neural network models Xception, ResNet50 for training. Finally, the greater the value of TMB associated with DNA mutations, the better the effect of immunotherapy.

Keywords: Immunity; therapy; Deep learning.

1. Introduction

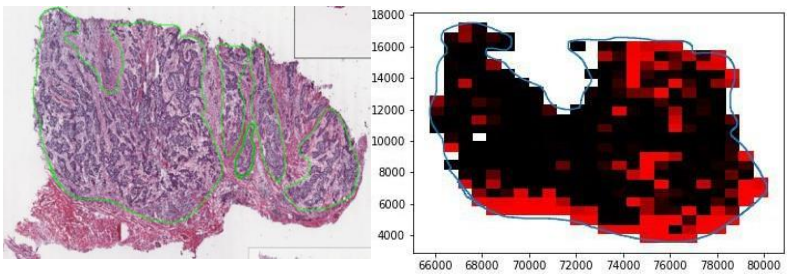
Deep learning has found a place in almost all walks of life, especially in medical field. "AI+Medical" [1] is making a strong revolutionary storm for the medical field. The fact that Convolutional Neural Networks (CNNs) is able to extract subtle and topological features which has made deep learning a handful tool for clinicians and pathologists to detect and classify mutations. The TMB mentioned here is actually the number of mutations in the coding region per megabase of genes in tumor tissues [2]. The greater the number of cancer cell mutations, the greater the possible phase identified by immune cells [3]. Moreover, the protein expression on the cell surface after canceration is different from that before carcinogenesis. This work proved the possibility of the prediction of immunotherapy treatment effectiveness, based solely on pathological images, by means of examining tumor region feature changes.

2. Region of Interest (ROI) Extraction

Since histopathological images tend to be very large (usually over 100,000 pixels), which makes it implausible to be directly used in today' s CNNs. We extracted ROIs from what pathologists labeled on WSIs [4], sliced them into patches, like 32x32 pixels [5], and kept RGB channels information. Different slicing strategies were used according to the needs of feature extraction and CNN model inputs.

2.1 Grid Cut

Labeled regions are generally larger than the size of sample patches, first we should cut the regions step by step, in order to grasp a complete sense of tumor areas. By varying the patch size, cut step size and overlapping threshold when the patch comes to the border of labeled regions, we can get different batch of samples in different numbers and feature dimensions.



(a) (b)

Figure 1. a Pathology large slice; b Pathology large slice.

As shown in Figure. Left: scan view of WSI. Right: grid cut illustration of the left most labeled region from the image on left panel.

2.2 Random Cut

The number of negative (immunotherapy ineffective) and positive (immunotherapy effective) WSIs are hardly balanced due to clinical imbalance, which will lead to imbalanced training data, and in turn will affect model prediction. The more the training data belonging to certain class, the easier the classifier will predict test data into that class. In order to eliminate this effect, we have tried to sample evenly between negative and positive WSIs. That is, given the number of wanted negative and positive patches, we keep slicing patches from both classes of WSIs, until target is reached. Since different WSIs have different number of labeled regions with different dimensions, we examined two strategies during random sampling [6]: A) each WSI was allocated evenly in contributing patches. Inside specific WSI, patches were distributed based on labeled sizes. B) The number of patches one WSI was took weight by summarizing one's labeled size among all WSIs.

3. Image Proprocessing

Pathological image data often has batch-related difference due to staining methods, sample preparation environment and image scanner. There will be some changes between different batches of images or even between images within the same batch. ROIs extracted from the pathological image data which will also be interfered with non-diagnostic related features. These useless or even negative image differences can affect feature extraction and algorithmic results; Therefore, image preprocessing is highly demanded. The purpose of image preprocessing, on one hand is to eliminate the non-feature correlation between different images, on the other hand is to eliminate interfering background, interstitial and other information inside the image [7]. The advantage of this article is using cells as an evaluation criterion, but the disadvantage is that a single case image is used as an index image, which is highly random. This article is a standardization base on 100 WSIs selected by the pathologists.

3.1 RGB Color Standardization

Considering that the color space of the patches we trained on is widely distributed, there may be outliers and noise. After standardization, we can reduce and avoid the influence of outliers and extreme values on the training results through the centralized method [8]. which have balanced the contribution of various characteristics to the result.

First step: Randomly selected standard pictures. A total of 100 WSIs were randomly selected, from each WSI randomly selected 100 small patches for target sampled. Patches with relatively large blank spaces and interstitials were removed. Second step: calculate average pixel values in RGB channels, individually, and then averaged over all patches. The three RGB values, or more precisely, the ratio between them, served as the reference that all the other patches will align with. Third step: calculate the channel color ratio. For each image, average RGB values were firstly computed, B channel value unchanged, the other two values were varied accordingly reference ratio. Pixels values larger than 255 were ceiled to 255. and last, standardize the color space vector, the flow chart as followed :

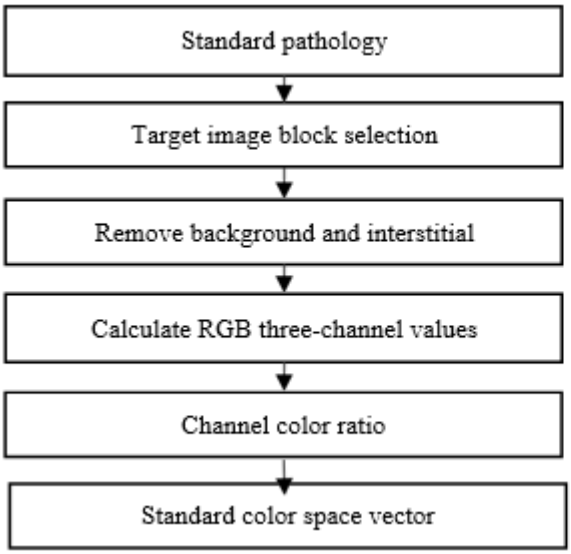


Figure 2 Standard color space flow chart

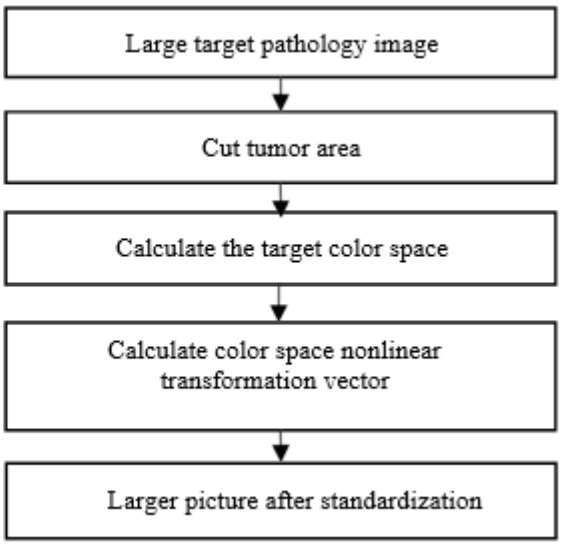


Figure 3 Standard target image

After Establishing a standard color space, we will standardize cases of images.

First step: slice the pathological image to patches then choose the patches which include the tumor region [9]. Second step: calculate the Nonlinear Transform vector of the color space, according to the standard color space which we have calculated at step 1. Third step: standardize the patches.

Figure 4 shows the distribution of RGB channels. After standardization, as we can see, the color of patches become dimming and the red, green channel become more consistent (PS: standardization is not to make the RGB channels become consistent).

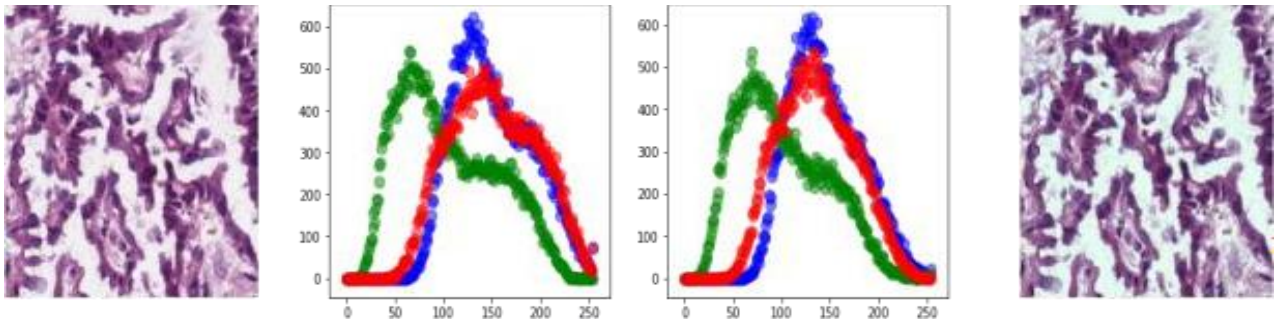


Figure 4 RGB channel distribution before and after standardization

3.2 Clean up Dirty Data and Increase the Brightness of Pictures

The stroma between cells tend to become blank in the image, and those blanks contribute little to the training, on the contrary, they do opposite influence to the training results, Of course including those images which contain less useful information, so we need to delete those images ,retain those we interested,while slicing the case of images. we used cell proportion to judge the images which contains less useful information, we considered those images which the cell proportion less than 0.3 as less useful, and we applied Laplace Algorithm, OTSU Algorithm to detect and delete those images. Before using Laplace Algorithm and OTSU Algorithm to detect and delete image, we need to binarize the case images. Then use Gaussian Blur algorithm to remove noise. The threshold method of opencv needs to adjust parameters, The OTSU Algorithm effectively solved this problem. The Otsu method (OTSU) [10] is based on the maximum class variance method, which can be adaptively thresholded and uses Laplacian operators to filter blurred images, as figure 5 shows. After deleting the images which contain large blank and less useful information, then we used Xception to train, and the accuracy rate of the training is up to 100%.

Increasing the brightness of the image will contribute to the accuracy rate of the training model, through increasing the brightness of the image [11], we can remove the lower pixel stroma of cell and

decrease the noise. Help to improve the accuracy of the model. First we need to transform the color space, RGB transform to HLS, and then adjust the brightness saturation.

The selected images above still have some interstitium, in order to remove background noise, We set the outside the cell's background pixel value to 255.filled the background of the images to white, in this operation ,We still use OTSU Algorithm, The OTSU Algorithm is not only used to remove the blank, it can also be used to make the nonwhite background to white background.

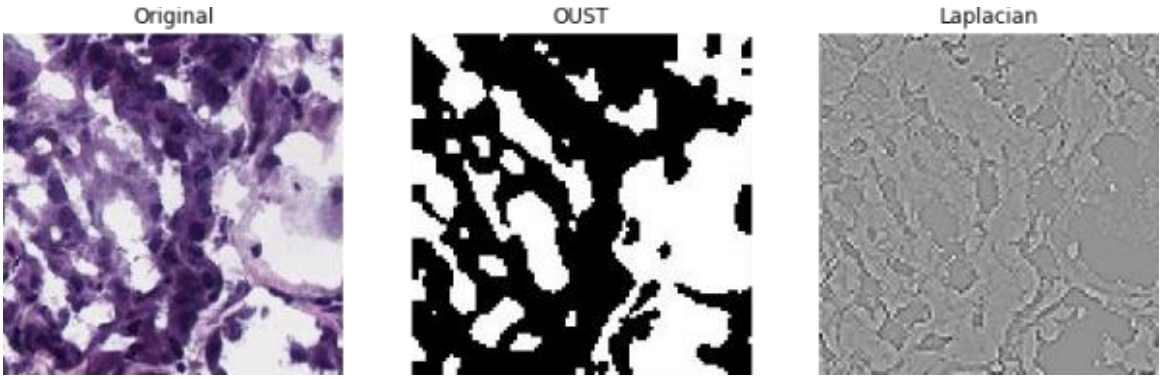


Figure 5 OTSU, Laplace transform

3.3 Data Enhancement

In most situation, the number of negative Data is far more than the positive data, which make the accuracy rate of positive data low.so we enhanced the positive data, by rotating 90 °, 180 °, 270 °, to the ratio of the positive data and negative data tend to 1:1.

4. Model Training

4.1 Algorithms Trial

We use Xcepiton model and ResNet50 model to detect the LUSC, the results were as follows.

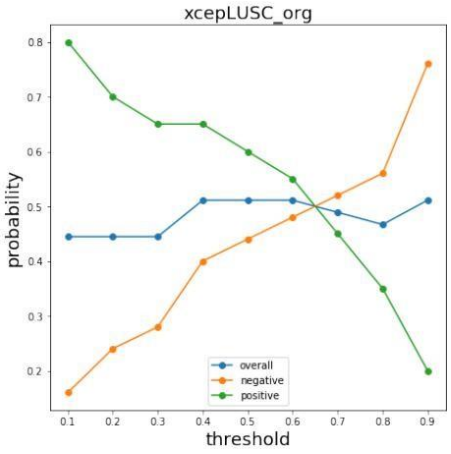


Figure 6(a) Replacement algorithm

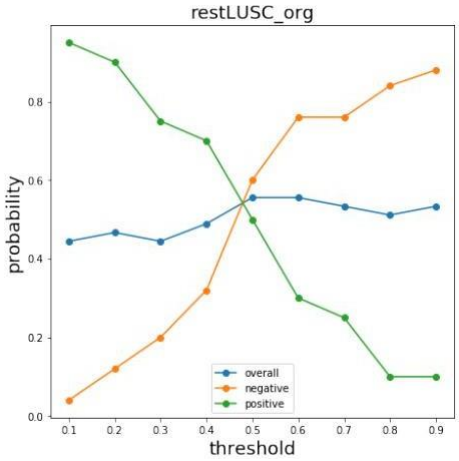


Figure 6(b) Replacement algorithm

As shown above, Figure 6-a shows the training results of Xception, Figure 6-b shows the training results of ResNet50[12] ,and the vertical axis represents accuracy rate, The horizontal axis represents positive proportion. Blue line represents the accuracy rate of the training samples. Green line represents the accuracy rate of the positive samples, Orange line represents the accuracy rate of the negative samples. Normally, we use 0.5 as a threshold segmentation points to deal with the binary classification problems, When the positive number is more than the negative quantity, the positive proportion is inevitably more than 0.5, and the negative proportion is necessarily less than 0.5. But in real life, it is not often used, for example, in our real life, Students with a score higher than 80 are identified as outstanding students. So here we try to change the positive proportion. The larger the x-

coordinate value is, the larger the proportion of positive data is, and the smaller the negative proportion is. However, even if the threshold value is changed, the results with high accuracy of all samples is also not obtained. From the results, we can see that the two training models have no significant difference.

4.2 Image Cut Trial

Due to the imbalance of negative and positive patches, we used random sample to cut out the same number of patches before training. Observed the difference between the two algorithms is still not too much.

4.3 Classification Training based on TMB Number

The above trainings all use 210 TMB[13]to divide the positive patches and negative patches, we selected 100 lowest TMB patches and 100 TMB highest patches (TMB < 147 ,TMB >325) as group one ,then we also selected 26 lowest TMB patches and 26 highest TMB(TMB < 55,TMB > 550) as group two, we have used two group as training, the results were as follows.

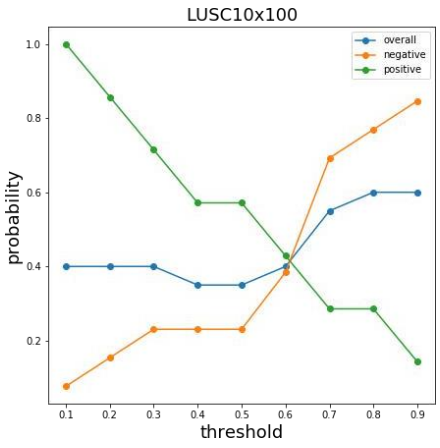


Figure 7(a) Change the cut mode

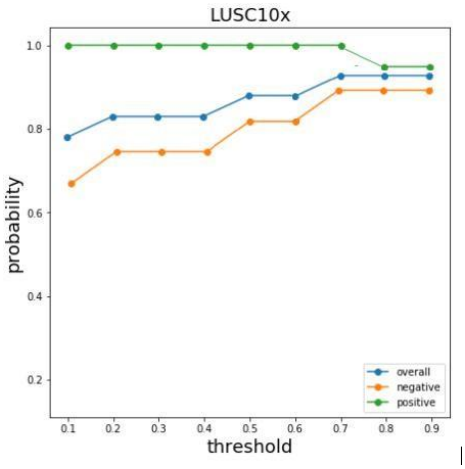


Figure 7(b) Change the cut mode

Figure 7-a shows the training results when we take 100 samples separately at both ends of the TMB. and as we can see, the figure does not exist three points where accuracy is relatively high. Figure 7-b shows the training results when we take 26 samples separately at both ends of the TMB. when we take 26 samples separately at both ends of the TMB, and set the positive threshold to

0.7 to train, the training achieved better results. Figure 7-b shows the accuracy rate of all patches reaches to 89%, the negative patches 'accuracy rate reaches to 80%, and the positive patches' accuracy rate reaches to 100%. Figure 8 shows the loss curve of this training. The horizontal axis represents the number of iterations. The vertical axis represents loss value. Figure 9 shows the AUC curve, from this, we drawn a conclusion that the high value of TMB or the low value of TMB can promote the immunotherapy.

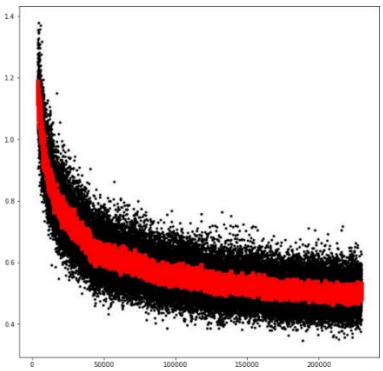


Figure 8 loss curve

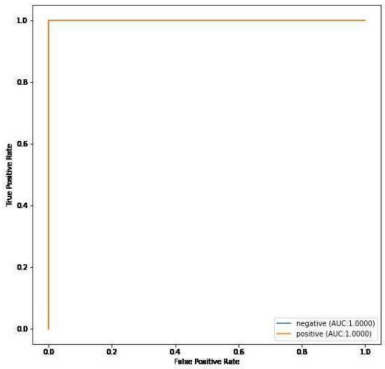


Figure 9 AUC curve

On this basis, we also have tried other tumors, bladder cancer BLCA, lung adenocarcinoma LUAD and cutaneous melanoma SKCM. From these cancer cells, we chosen those which TMB is less than 50, more than 500 to train, The accuracy rate all above 85%. During the training period, many other attempts have been made. For example, when the sample label of the small patches is uncertain, the samples with higher accuracy are selected for training, including the use of traditional methods to extract features and then use the classic SVM for classification. their effects are not particularly ideal. The results of high and low TMB training are ideal, and it also shows that TMB plays a certain role in promoting immunotherapy.

5. Summary

This paper proposes a tumor immunotherapy model based on deep neural network, and through different preprocessing, feature extraction and model attempts, finally, using TMB value training to obtain better results, improve the model's Accuracy.

There are still many problems to be studied in the study of deep learning in immunotherapy. This article has achieved some results only when the TMB takes the extreme value. If it rises to the application level, it needs more comprehensive research. The next step we will try to find a better training model from more aspects, such as using more accurate labeled data, more models, and distinguishing the effectiveness and ineffectiveness of immunotherapy across the entire TMB range.

References

[1]. Holt G. LETTER TO THE EDITOR: Clinical benchmarking for the validation of AI medical diagnostic classifiers[J].

[2]. Jr T F C, Sabel M S, Sugano M, et al. Growth of human tumor xenografts in SCID mice quantified using an immunoassay for tumor marker protein in serum[J]. Journal of Immunological Methods, 2000, 233(1):57-65.

[3]. Blanck G. Mutations and regulatory anomalies effecting tumor cell immune functions.[J]. Cancer Immunology Immunotherapy, 2004, 53(1):1-16.

[4]. Zheng Y, Jiang Z, Zhang H, et al. Size-scalable content-based histopathological image retrieval from database that consists of WSIs[J]. IEEE J Biomed Health Inform, 2017, PP(99):1-1.

[5]. Koller O, Forster J, Ney H. Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers.[J]. Computer Vision & Image Understanding, 2015, 141:108-125.

[6]. Barraquand J, Kavraki L, Latombe J C, et al. A Random Sampling Scheme for Path Planning[J]. International Journal of Robotics Research, 1997, 16(6):759-774.

[7]. Bejnordi B E, Litjens G, Timofeeva N, et al. Stain Specific Standardization of Whole-Slide Histopathological Images[J]. IEEE Transactions on Medical Imaging, 2016, 35(2):404-415.

[8]. Monaco J, Hipp J, Lucas D, et al. Image Segmentation with Implicit Color Standardization Using Spatially Constrained Expectation Maximization: Detection of Nuclei[M]// Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012. 2012.

[9]. Monaco J, Raess P, Chawla R, et al. Image segmentation with implicit color standardization using cascaded EM: Detection of myelodysplastic syndromes[C]// IEEE International Symposium on Biomedical Imaging. 2012.

[10]. Wang H Y, Pan D L, Xia D S. A Fast Algorithm for Two-dimensional Otsu Adaptive Threshold Algorithm[J]. Journal of Image & Graphics, 2005, 33(9):968-971.

[11]. Ribli D, Pataki B Á, Csabai I. Learning from deep learning: better cosmological parameter inference from weak lensing maps[J]. 2018.

[12]. Amendola V, Meneghetti M. Exploring How to Increase the Brightness of Surface-Enhanced Raman Spectroscopy Nanolabels: The Effect of the Raman-Active Molecules and of the Label Size[J]. Advanced Functional Materials, 2012, 22(2):353-360.

[13]. Chollet, François. Xception: Deep Learning with Depthwise Separable Convolutions[J]. 2016.