A deep learning-based computational prediction model for characterizing cellular biomarker distribution in tumor microenvironment

Zhengyao Peng[1,2], Chang Bian[1,2], Yang Du*[,1,2], Jie Tian*[,1,2,3,4]

1. CAS Key Laboratory of Molecular Imaging, the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

2. School of Artificial Intelligence, The University of Chinese Academy of Sciences, Beijing, 100049, China

3. Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine Science and Engineering, Beihang University, Beijing 100191, China.

4. School of Life Science and Technology, Xidian University, Xi'an 710071, China.

*Corresponding authors, E-mail: yang.du@ia.ac.cn, tian@ieee.org

Abstract

Evaluation of cancer cell and immune cell distribution in tumor microenvironment (TME) is one of the most important factors for guiding cancer immunotherapy and assessing therapeutic response. Multiplexed immunohistochemistry (mIHC) is often used to obtain the different types of cellular biomarker expression and distribution information in TME, but mIHC is limited by time-consuming and cost-intensive, and pathologists' objectives etc. In this work, we proposed a deep learning-based modified U-Net (m-Unet), by replacing the original convolution sub-module with a modified block to predict the distribution of several typical cellular biomarkers' expression and distribution information in TME. We have demonstrated that our model can be trained in both fully supervised and semi-supervised manners. The model can extract segmentation information from Hematoxylin and Eosin (H&E) images, and predict the cellular biomarker distributions including panCK for colon cancer cells, CD3 and CD20 for tumor infiltrating lymphocytes (TILs) and DAPI for nucleus. We have demonstrated that our model can be trained in both fully supervised and semi-supervised manners and. the performance of the m-Unet is better than the U-Net in this work. The optimal prediction accuracy of m-Unet is 88.3% on the test dataset. In general, this model possesses the potential to assist the clinical TME analysis.

Keywords: tumor microenvironment, cellular biomarker, deep learning, semi-supervised training

I.INTRODUCTION

In recent years, immunotherapy has progressed rapidly in theory and clinical practice and achieved dramatic therapeutic efficacy[1][2][3].Recent works have proven that understanding the immune cell expression and distribution in tumor microenvironment (TME) has significant benefit on evaluating the response of immunotherapy [4][5]. Although multiplexed immunohistochemistry (mIHC) can be used to obtain cellular biomarker distribution information in TME, it suffers from drawbacks such as time-consuming and cost-intensive, and it is also subjected to pathologists' objectives [6]. Developing a computational prediction method is of significant importance to avoid the disadvantages of mIHC technology and to facilitate the evaluation of immune cell distribution for the guidance of therapeutic optimization and patient's selection.

Deep learning has been applied on many image analysis tasks in digital pathology [7][8]. And recent works have revealed that deep learning-based methods have shown great potential for predicting TME biomarker distribution [7][9][10]. In this work, we aim to realize the pixel-level prediction of the cellular biomarker distribution in TME by semantic segmentation methods to better aid clinical analysis. Due to the heavy work to acquire pixel-level annotated data [8], the semi-supervised training strategy is of significant value for pathological image analysis tasks to reduce annotation burden.

In this work, our semi-supervised training strategy is based on conditional generative adversarial network (CGAN)

Medical Imaging 2022: Digital and Computational Pathology, edited by John E. Tomaszewski, Aaron D. Ward, Proc. of SPIE Vol. 12039, 120391H · © 2022 SPIE
1605-7422 · doi: 10.1117/12.2610640

[12]. CGAN is proposed on the basis of generative adversarial network (GAN) [13]. Unlike GAN, in which the generator and discriminator adopt noise signals as the initial input, CGAN adds conditional information to generator and discriminator. Therefore, CGAN allows us to control the output of the generator. Our semi-supervised learning method is similar to CGAN, except that we use the segmentation network to replace the generator.

In order to evaluate the distribution of cellar biomarkers efficiently and precisely, we propose a deep learning-based modified U-Net (m-Unet). The m-Unet can be trained in both supervised and semi-supervised manners using the dataset from the Cancer Genome Atlas (TCGA). We utilized m-Unet to translate H&E images to the feature maps, which can show the cellular biomarker distributions at the pixel-level including panCK for cancer cells, CD3 and CD20 for tumor infiltrating lymphocytes (TILs) and DAPI for cellular nucleus. These biomarkers play an important role in TME and immunotherapy[14]. In this work, when using the data for supervised training, the accuracy of m-Unet can reach 88.2%, which is an improvement compared with the results of 84.6% in[15], and the performance of the model has been further improved by adding unlabeled data for semi-supervised training. Among all models in this work, the m-Unet model showed the best performance in all the indicators we used, and the result shows that we provide an efficient and low-cost method for obtaining cellular biomarker distribution, which can facilitate analysis of TME for pathologists and provide guidance for the clinical immunotherapy.

II. MATERIALS AND METHODS

2.1 Dataset

In order to get better results and verify the feasibility of m-Unet in both fully supervised and semi-supervised manners in this study, we prepared labeled and unlabeled data.

Labeled data. The original labeled data was got from[15]， and 9420 labeled data were obtained after augmentation in this work. We further divided the data into training dataset, validation dataset, and test dataset at a ratio of 8:1:1.

Unlabeled data. The unlabeled dataset was got from three colon cancer patients in the TCGA-COAD project. After selection， the image data was cropped into 512 × 512-pixel 848 patches. And then we normalized these unlabeled image data using the annotation data to keep the image color consistent [16]. Finally, 8480 unlabeled image data were obtained after the same augmentation step as annotation data.

2.2 M-Unet

M-Unet is modified on the basis of the U-Net network [17]. The structure of m-Unet is shown in Figure 1(a), and it can be seen that the m-Unet is organized by replacing the original convolution submodule with modified block, which is shown in Figure 1(b). The modified block is inspired by Inception-v4 [18], and it uses the 3×3 small convolution kernels, and obtains the deeper receptive field by increasing the number of convolutions. Results of different convolutions show different levels of semantic information, and the neural network can better fuse semantic information and automatically learn its proportions with the parallel connection.

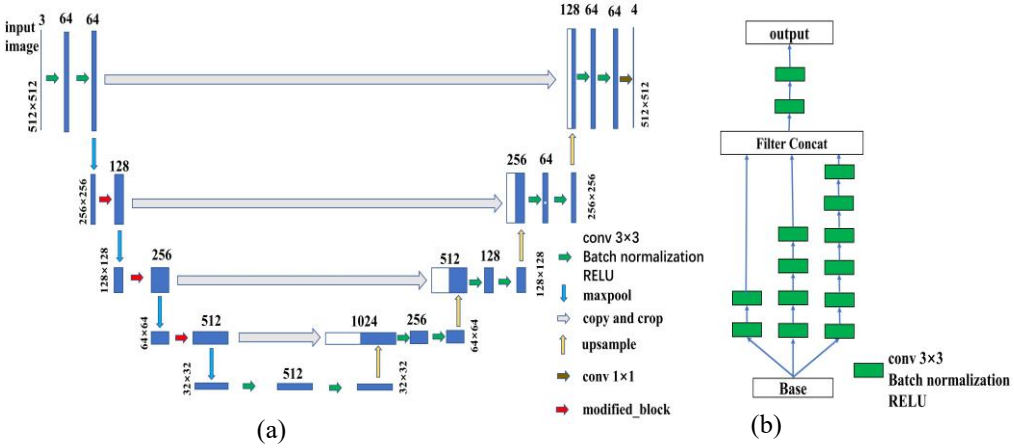

Fig.1(a) The structure of the m-Unet network, (b) the internal structure of the modified block

2.3 Training

Supervised training. The optimizer is stochastic gradient descent (SGD) included in PyTorch, and the loss function is linearly weighted sum of cross-entropy loss and dice loss between the prediction results of the segmentation network and the label. The cross-entropy loss aims at improve the accuracy of the segmentation network, and the dice loss is added for better F1 score [19]. The loss function can be described in formula (1).

$$L = \omega_1 L_{mce}(p_n, y) + \omega_2 L_{Dice}(p_n, y) \quad (1)$$

Where ω denotes different weights. y and p_n denote the ideal output and the corresponding real output of segmentation network. L_{mce} are multi-class cross-entropy loss and L_{Dice} is dice loss.

Semi-Supervised training. Due to the high cost of obtaining pixel-level labeled image data, semi-supervised training is of great importance for pathological image analysis tasks. We used a semi-supervised training method based on CGAN [12] to reduce the annotation burden. The semi-supervised training process is shown in the figure 3. the prediction network is the segmentation model, whose input is H&E images, and output is segmentation results. Compared with the supervised training, a classification network is used as the discriminator, whose input is the segmentation results and the original images, and output is the probability of the input from the labeled data [20]. In this work, the first stage of semi-supervised training is the supervised training of the predictive model, and the loss is calculated as formula (1). The second is the supervised training of the classification model, the supervised loss is the binary cross-entropy between the classification result and the real origin. And the discriminator produces consistency loss through the recognition of unlabeled data and segmentation results in last stage.

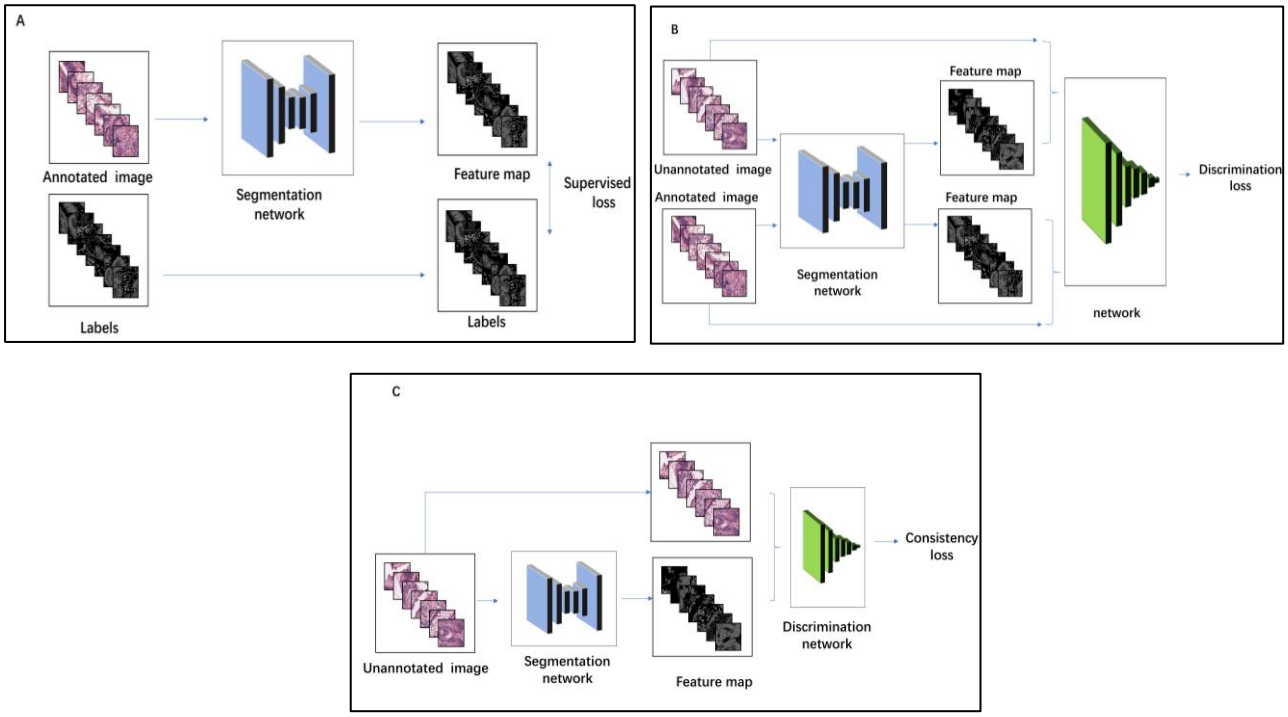

Fig.2 the processing of semi-supervised training method used in this work:(a) supervised training phase of segmentation model, (b) supervised training phase of discriminator, (c) consistency loss for the segmentation model's semi-supervised training.

When the prediction network is overfitting, the discriminator can distinguish the source of the data more easily to produce a larger consistency loss, which is used to optimize the parameters of prediction model, and hence the segmentation model can get performance improvements in the adversarial process with the identification network. The trained optimizer is SGD included in PyTorch and the loss function is as following.

$$L = \omega_1 L_{mce}(p_n, y) + \omega_2 L_{Dice}(p_n, y) + \omega_3 L_{bce}(d(z_m, p_m), m) \quad (2)$$

Where z_m and p_m denote the unlabeled data and the corresponding output of the segmentation network. L_{Dice} is dice loss. $d(z_m, p_m)$ denotes the output of the discriminator, and the m m denotes the source of data. L_{mce} and L_{bce} are multi-class and binary cross-entropy losses.

III. RESULTS AND DISCUSSION

In order to demonstrate the feasibility of m-Unet, in this work, we trained the U-Net network and our proposed m-Unet network in both supervised and semi-supervised manners. Table 1 shows the results of the network on the test dataset, and Figure 3 illustrates the prediction outputs.

It can be found in the Table 1 that the m-Unet provides a better result compared with U-Net in pixel accuracy (PA), Frequency Weighted Intersection over Union (FWIoU), precision, recall, F1-score in both supervised and semi-supervised manners, which proves that m-Unet is more suitable for predicting cellular biomarker distribution in TME. And the results show that the semi-supervised training based on CGAN further improves the performance of m-Unet. The experimental results demonstrate the superiority of the m-UNet structure in analyzing cellular biomarker distribution in TME.

Table 1 Experimental results of different networks

Network	PA	FWIoU	Precision	Recall	F1-score
U-Net (supervised)	0.879	0.789	0.752	0.741	0.731
m-Unet (supervised)	0.882	0.794	0.761	0.754	0.743
U-Net (semi-supervised)	0.879	0.791	0.743	0.747	0.731
m-Unet (semi-supervised)	**0.883**	**0.795**	**0.767**	**0.758**	**0.747**

Figure 2 shows the result of our models, and it can be seen from Table 1 and Figure 3 that our prediction is closed to the annotation made by pathologists, which suggests that our model is capable to make prediction of the distribution of biomarkers as panCK, DAPI, CD3 and CD20 in TME.

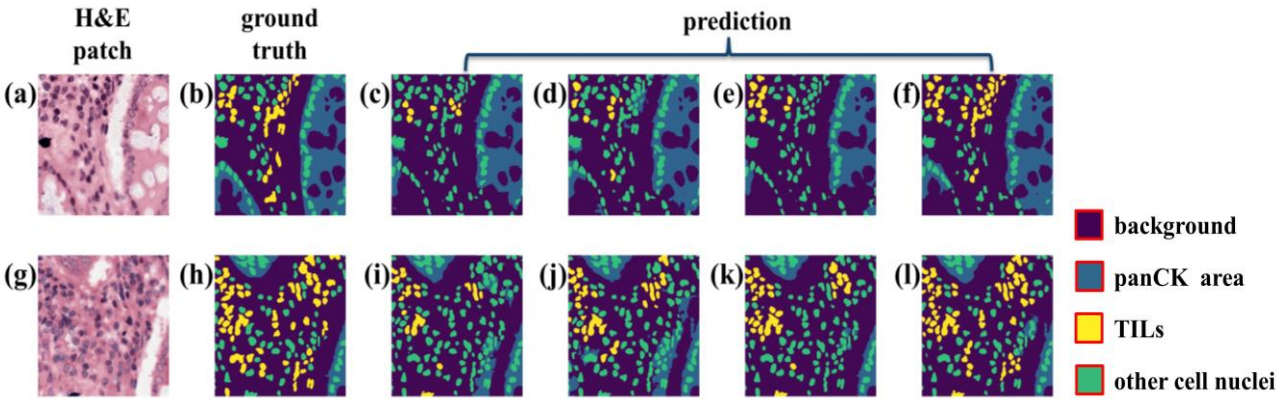

Fig.3 Prediction results: (a) & (g) the original H&E image, (b) & (h) the corresponding ground truth, (c) & (i) the prediction of U-Net trained with supervised method, (d) & (j) the prediction of U-Net trained with semi-supervised method, (e) & (k) the prediction of m-Unet trained with supervised method, (f) & (l) the prediction of m-Unet trained with semi-supervised method.

IV. CONCLUSION

In this work, we propose m-Unet to make predictions of cellular biomarker distribution in TME. The m-Unet is organized by replacing the original convolution submodule with modified block, and we have shown the network can be trained using supervised training method and semi-supervised training method based on CGAN. The m-UNet model is capable of making predictions of cellular biomarker distribution including panCK, DAPI, CD3 and CD20 after training, and we illustrate that the performance of the m-Unet is better than the original U-Net in all indexes we adopt, and semi-supervised training based on CGAN bring further improvements in model performance of m-Unet. Results showed that our m-Unet can accurately predict the cellular biomarker distribution with an optimal accuracy of 88.3%. Our work provides an efficient and low-cost computational solution for obtaining cellular biomarker distribution information in TME.

V. ACKNOWLEDGEMENTS

This work was supported by Ministry of Science and Technology of China under Grant No. 2017YFA0205; National Natural Science Foundation of China under Grant Nos. 81871514, 81227901, 81470083, 91859119, 61671449, 81527805, 61901472; National Public Welfare Basic Scientific Research Program of Chinese Academy of Medical Sciences under Grant No.2018PT32003, 2017PT32004, National Key R&D Program of China under Grant Nos. 2018YFC0910602, 2017YFA0205200, 2017YFA0700401, 2016YFA0100902, 2016YFC0103702. National Natural Science Foundation of shaanxi Province under Grant No.2019JM-459.

REFERENCES

[1] Khalil D N, Smith E L, Brentjens R J, *et al*, "The future of cancer treatment: immunomodulation," CARs and combination immunotherapy,"Nature reviews Clinical oncology, 13(5), 273 (2016).
[2] Du Y, Jin Y, Sun W, *et al*, "Advances in molecular imaging of immune checkpoint targets in malignancies: current and future prospect," European radiology, 29(8), 4294-4302 (2019).
[3] Taube J M, Klein A, Brahmer J R, *et al*, "Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti–PD-1 therapy," Clinical cancer research, 20(19), 5064-5074 (2014).
[4] Budhu A, Forgues M, Ye Q H, *et al*, "Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment," Cancer cell, 10(2), 99-111 (2006).
[5] Du Y, Qi Y, Jin Z, *et al*, "Noninvasive imaging in cancer immunotherapy: The way to precision medicine," Cancer letters, 466, 13-22 (2019).
[6] Stack E C, Wang C, Roman K A, *et al*, "Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis," Methods, 70(1): 46-58 (2014).
[7] Bian C, Wang Y, Lu Z, *et al*, "ImmunoAIzer: A Deep Learning-Based Computational Framework to Characterize Cell Distribution and Gene Mutation in Tumor Microenvironment," Cancers, 13(7), 1659 (2021).
[8] Esteva A, Robicquet A, Ramsundar B, *et al*, "A guide to deep learning in healthcare," Nature medicine, 25(1),24-29 (2019).
[9] Wang S, Wang T, Yang L, *et al*, "ConvPath: A software tool for lung adenocarcinoma digital pathological image analysis aided by a convolutional neural network," EBioMedicine, 50, 103-110 (2019).
[10] E. A. Burlingame, A. A. Margolin, J. W. Gray, *et al*., "SHIFT: speedy histopathological-to-immunofluorescent translation of whole slide images using conditional generative adversarial networks," Medical Imaging 2018: Digital Pathology, 10581 (2018).
[11] Burlingame, Erik A., *et al*, "SHIFT: speedy histological-to-immunofluorescent translation of a tumor signature enabled by deep learning," Scientific reports, 10(1): 1-14 (2020).
[12] Mirza, Mehdi, and Simon Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv: 1411.1784 (2014).
[13] Goodfellow, Ian, *et al*, "Generative adversarial nets," Advances in neural information processing systems, 27 (2014).

[14] Salgado, Roberto, et al, “The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014,” Annals of oncology, 26(.2), 259-271 (2015).
[15] Bian C, Wang Y, An Y, et al. “A computational prediction method based on modified U-Net for cell distribution in tumor microenvironment,”Medical Imaging 2021: Digital Pathology. International Society for Optics and Photonics, 11603, 116030C (2021).
[16] Vahadane, Abhishek, Vahadane A, Peng T, Sethi A, *et al*, “Structure-preserving color normalization and sparse stain separation for histological images,” IEEE transactions on medical imaging, 35(8), 1962-1971(2016).
[17] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox, “U-Net: Convolutional networks for biomedical image segmentation,” International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 234-241 (2015).
[18] Szegedy, Christian, *et al*, “Inception-v4, inception-resnet and the impact of residual connections on learning,”Thirty-first AAAI conference on artificial intelligence (2017).
[19] Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” 2016 fourth international conference on 3D vision (3DV). IEEE, 565-571 (2016).
[20] Zhang, Yizhe, *et al*, “Deep adversarial networks for biomedical image segmentation utilizing unannotated images,” International conference on medical image computing and computer-assisted intervention. Springer, Cham, 408-416 (2017).