DOI: 10.19884/j.1672-5220.202307005

Systematic PD-L1 Slide Analysis Based on Multi-Objective Learning

CHEN Zhao[1], GUO Danqi[1], WANG Qian[1*], SHEN Yiting[1], WANG Qingguo[2]
1. School of Computer Science and Technology, Donghua University, Shanghai 201620, China
2. Radiology Department, Shanghai General Hospital, Shanghai 200080, China

Abstract: In treatment of cancers, especially non-small-cell lung cancers such as lung squamous cell carcinoma (LUSC), tumor proportion score (TPS) of a programmed death-ligand 1 (PD-L1) slide is essential for selecting tumor therapies. Many parameters of tumor cells (TCs) are vital to cancer diagnosis. Although the indexes can be estimated via the computational analysis, there is seldom a unified system that could acquire different nucleus information simultaneously. To address the issues, multi-objective learning pipeline (MOLP) is proposed to predict TPS, cell counts, nucleus contours and categories altogether from PD-L1 slides of LUSC. The main network comprises two branches, one estimating TPS via the cell analysis and the other directly regressing TPS. It minimizes the difference between these two approximated values of TPS to gain robustness. The cell-analysis branch increases confidence of the estimated TPS by nucleus segmentation, classification and counting. It also enables the system to estimate appearance parameters of TCs for LUSC diagnosis. Experiments on a large image set show that MOLP is feasible and effective. The TPS predicted by MOLP exhibits statistically significant correlation with pathologists' scores, with a mean absolute error (MAE) of 4.97 (95% confidence interval (CI): −0.56–10.49) and a Pearson correlation coefficient (PCC) of 0.97 ($p < 0.001$).

Key words: programmed death-ligand 1 (PD-L1) slide; tumor proportion score (TPS); multi-objective learning; classification; segmentation; counting
CLC number: TP391.1 **Document code**: A
Article ID: 1672-5220(2024)03-0221-10

Open Science Identity (OSID)

0 Introduction

Immunohistochemistry staining technique can highlight certain biomarkers in histopathology slides that are vital to cancer diagnosis and treatments. It is known that tumors can evade the immune system via exploitation of inhibitory checkpoint pathways. In the programmed death-1 (PD-1) and programmed death-ligand 1 (PD-L1) pathways, PD-L1 that is expressed by tumor cells (TCs) binds to PD-1 on T-cells, inhibiting T-cell receptor signaling and blocking antitumor immune response[1]. Thus, PD-L1 expression becomes a crucial biomarker of PD-1/PD-L1 therapy response, which can be indicated by the tumor proportion score (TPS), i.e., the ratio between the count of positive tumor cells (TC(+)) and the count of TCs in a PD-L1 stained whole slide image (WSI). While the TPS is essential for selecting tumor treatments, other numerical indexes such as cell counts and appearance parameters are also valuable. Traditionally, detecting and measuring the cells are time-consuming, and only allow the pathologists to provide an approximate range rather than an exact percentage for the TPS. For the appearance parameter estimation, the situations are more or less the same. Although the estimates can be used for giving diagnoses and prescribing medications, they may be not accurate enough to support further investigation into the cancers.

Given the prevalence of the digital WSI, these TC parameters can be easily computed by image analysis models, rather than human beings. Deep learning has greatly improved histopathology image analysis[2-3]. However, there lacks a framework that can obtain different aspects of pathology information systematically. The existing methods usually adopt a multi-stage strategy. With multiple complicated models being attached to each other, they segment, classify and count TCs and compute TPS successively with the independent modules[4-6], thus being prone to computational inefficiency and accumulated errors.

Therefore, this paper proposes a unified system, namely, multi-objective learning pipeline (MOLP). It contains two parts: a multi-objective network (MON) and a nucleus parameter estimator (NPE). It performs TPS estimation, TC counting, nucleus segmentation and nucleus classification to extract comprehensive information from PD-L1 slides of lung squamous cell carcinoma (LUSC). The main network, MON, comprises a regression branch (RB) and a cell-analysis branch (CB) for direct and indirect TPS estimation, respectively.

The contributions of this work are as follows.

Received date: 2023-07-22
Foundation items: National Natural Science Foundation of China (Nos. 61702094 and 62301142); "Chenguang Program" Supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission, China (No. 18CG38)
* Correspondence should be addressed to WANG Qian, email: wangqian@dhu.edu.cn
Citation: CHEN Z, GUO D Q, WANG Q, et al. Systematic PD-L1 slide analysis based on multi-objective learning[J]. *Journal of Donghua University (English Edition)*, 2024, 41(3): 221-230.

1) The proposed MOLP predicts TPS direct from the PD-L1 slides in an end-to-end fashion, which is much more convenient and accurate than the multi-stage methods that take steps to compute TPS[6].

2) Apart from directly estimating TPS by RB, the MON predicts TPS based on the cell analysis to improve accuracy and credibility and minimizes the discrepancy between the different TPS values to gain robustness.

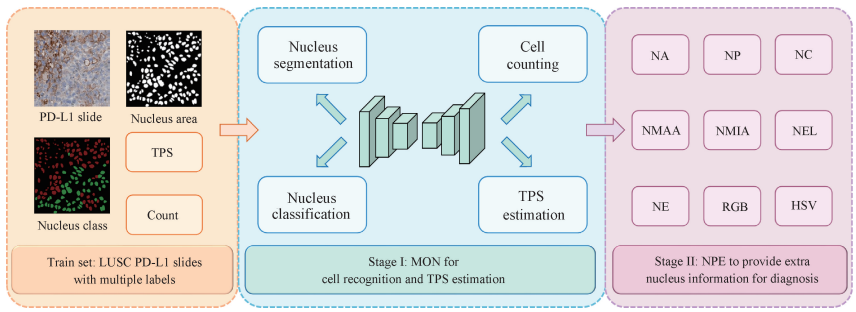
3) The proposed MOLP can provide comprehensive information for pathologists to consider from different perspectives to make accurate diagnoses as it automatically computes size, shape and color parameters of the TC based on segmentation and classification masks. Experiments on a large PD-L1 data set of LUSC can demonstrate the feasibility and effectiveness of MOLP.

1 Related Work

Most work of histopathology image analysis by deep learning focuses on nucleus/cell segmentation and classification[7-9]. Only a few methods deal with TPS estimation in PD-L1 slides, while most of these methods include multiple stages[5-6,10] which segment, classify and count TCs and finally compute TPS by separate models[4,11-12]. Kang et al.[6] proposed a multi-stage ensemble strategy for TPS and assessment of PD-L1 expression. Liu et al.[10] combined three modules to predict TPS automatically. Mi et al.[4] designed a quantitative analysis platform for PD-L1 immunohistochemistry. Since no model is guaranteed one hundred percent correct, the errors of each model would accumulate as the data are passed on from one stage to another. This could seriously decrease the accuracy of TPS and undermine the confidence of the results. Moreover, there may be discrepancy between any two decision models. Even though one stage is accurate, it may not yield the features which are exactly required by the next stage. Therefore, an end-to-end network with multiple objectives is proposed to perform robust TPS estimation.

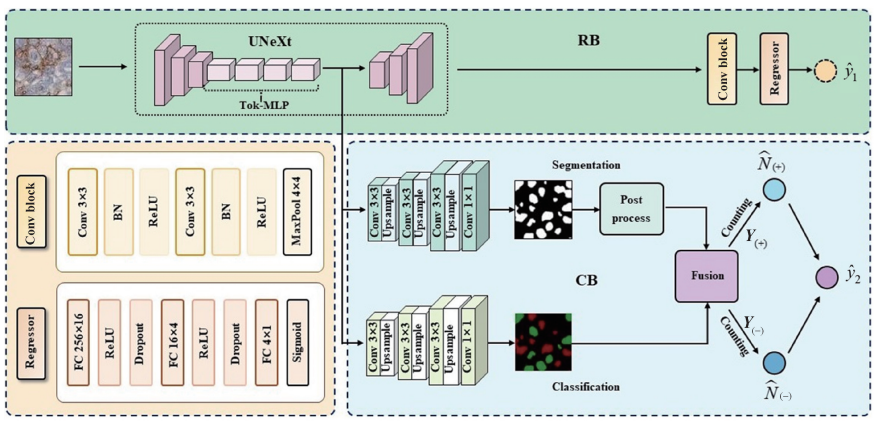
2 Methods

The proposed MOLP is illustrated in Fig. 1. Branches of MON, namely RB and CB, are shown in Fig. 2. The labels and the outputs are listed in Table 1. The former direct predicts TPS from the input PD-L1 slides and the latter performs multiple relevant tasks including nucleus segmentation, classification and counting to refine TPS values. Furthermore, the diversified outputs of CB are exploited to produce appearance parameters of TCs for LUSC diagnosis. Let $\boldsymbol{X} \in \mathbf{R}^{H\times W\times B}$ represent an input patch cropped from an original PD-L1 slide, where H, W and B are the height, the width and the band number, respectively. In this case, $H = W = 64$, and $B = 3$. The details of MOLP are described as follows.



NA—nucleus area; NP—nucleus perimeter; NC—nucleus circularity; NMAA—nucleus major axis; NMIA—nucleus minor axis; NEL—nucleus ellipticity; NE—nucleus eccentricity; RGB—red-green-blue value; HSV—hue-saturation value.

Fig. 1 MOLP



BN—batch normalization; ReLU—rectified linear unit; FC—full connection; Conv—convolutional; UNeXt—upgraded version of U-Net; Tok-MLP—tokenized multi-layer perceptron; MaxPool—maximum pooling.

Fig. 2 MON in MOLP

Table 1 Symbols of labels and outputs of MOLP

Label		Output	
Symbol	Explanation	Symbol	Explanation
y	Real TPS	$\hat{y}_1$	TPS predicted by RB
		$\hat{y}_2$	TPS predicted by CB
$\boldsymbol{Y}^{(\mathrm{seg})} = [y_{ij}^{(\mathrm{seg})}]$, $i = 1, 2, \cdots, H$, $j = 1, 2, \cdots, W$	Segmentation mask	$\hat{\boldsymbol{Y}}^{(\mathrm{seg})} = [\hat{y}_{ij}^{(\mathrm{seg})}]$, $i = 1, 2, \cdots, H$, $j = 1, 2, \cdots, W$	Segmentation map predicted by CB
$\boldsymbol{Y}^{(\mathrm{cla})} = [y_{ic}^{(\mathrm{cla})}]$, $i = 1, 2, \cdots, HW$, $c = 1, 2, \cdots, C$	Classification mask	$\hat{\boldsymbol{Y}}^{(\mathrm{cla})} = [\hat{y}_{ic}^{(\mathrm{cla})}]$, $i = 1, 2, \cdots, HW$, $c = 1, 2, \cdots, C$	Classification map predicted by CB
$N_{(+)}$	Count of TC(+)	$\hat{N}_{(+)}$	Count of TC(+) predicted by CB
$N_{(-)}$	Count of negative TCs(TC(−))	$\hat{N}_{(-)}$	Count of TC(−) predicted by CB

2.1 TPS

TPS is a percentage defined as

$$y = N_{(+)}/N \times 100\%, \tag{1}$$

where N is the total count of TCs in a PD-L1 WSI[13], $N = N_{(+)} + N_{(-)}$. It indicates levels of PD-L1 expression which is a crucial biomarker used in non-small cancer (e.g., LUSC) diagnosis and treatment. If TPS is lower than 1%, PD-L1 expression is considered negative; if TPS ranges from 1% to 49%, there is a positive but low PD-L1 expression; if TPS is equal to or above 50%, there is a high level of PD-L1 expression. Thus, with TPS, doctors can identify tumor patients who may respond to the treatment using PD-1/PD-L1 inhibitors. Traditionally, TPS is estimated manually as pathologists go through a slide, locate and examine tumor nests and roughly determine whether TPS is above the critical threshold (i.e., 50%) with bare eyes. Obviously, this manual approximation is laborious and prone to errors. It would be ideal if TPS could be predicted automatically and accurately, which, fortunately, can be realized by computational models such as deep learning networks. Under this circumstance, we propose a new method, MOLP for TPS estimation.

2.2 MON

MON is responsible for estimating TPS from PD-L1 stained slides. As shown in Fig. 2, MON involves RB constructed by UNeXt[14] to recognize cell patterns, a convolutional block to down-sample the features and a regressor made of several FC layers to estimate TPS. The feature maps $\boldsymbol{F} \in \mathbf{R}^{64\times64\times16}$ yielded by the fourth Tok-MLP of UNeXt[14] are fed to the convolutional block and the parallel CB. The loss of RB is the regression error $\ell^{(\mathrm{reg})}$ defined as

$$\ell^{(\mathrm{reg})} = \| y - \hat{y}_1 \|^2. \tag{2}$$

However, direct regressing TPS from large complicated PD-L1 slides could be inaccurate and slow. In pathology diagnosis, doctors are used to examining sizes, shapes and colors of typical TCs along with the qualitative indexes. TPS alone is far from enough for the experts to make any rational medical decisions. Thus, the network employs CB to provide diversified information on TCs and boost up the performance of TPS estimation. The main steps of CB are as follows.

1) Segment and classify nuclei by two convolutional blocks trained in parallel with the same structure as indicated in Fig. 2 and the same input $\boldsymbol{F} \in \mathbf{R}^{64\times64\times16}$ produced by the fourth Tok-MLP of UNeXt in RB. The segmentation loss $\ell^{(\mathrm{seg})}$ is

$$\ell^{(\mathrm{seg})} = 1 - \mathrm{Dice}(\boldsymbol{Y}^{(\mathrm{seg})}, \hat{\boldsymbol{Y}}^{(\mathrm{seg})}) = 1 - \frac{2\sum_{j=1}^{W}\sum_{i=1}^{H} y_{ij}^{(\mathrm{seg})}\hat{y}_{ij}^{(\mathrm{seg})}}{\sum_{j=1}^{W}\sum_{i=1}^{H} y_{ij}^{(\mathrm{seg})} + \sum_{j=1}^{W}\sum_{i=1}^{H}\hat{y}_{ij}^{(\mathrm{seg})}}, \tag{3}$$

where the measurement $\mathrm{Dice}(\cdot)$ evaluates of intersection-over union of two masks[15]. The classification loss $\ell^{(\mathrm{cla})}$ is

$$\ell^{(\mathrm{cla})} = -\frac{1}{HW}\cdot\sum_{i=1}^{HW}\sum_{c=1}^{C} y_{ic}^{(\mathrm{cla})}\log(\hat{y}_{ic}^{(\mathrm{cla})}). \tag{4}$$

2) Refine the predicted segmentation map for each patch by morphology operations (first erosion and then dilation, both using an 8 pixel×8 pixel kernel) to enlarge and smooth the areas of nuclei.

3) Fuse the refined segmentation and classification results to obtain separate masks for TC(+) and TC(−). For each instance (i.e., nucleus) in the segmentation map, compute coordinates of its central point, project it to the classification map, find out its predicted class and let the class be the label of the segmented instance. Then use one template to hold all the positive nuclei and another for the negative nuclei. In this way, the binary masks $\boldsymbol{Y}_{(+)} \in \mathbf{R}^{H\times W}$ and $\boldsymbol{Y}_{(-)} \in \mathbf{R}^{H\times W}$ for TC(+) and TC(−), respectively, are created. As segmentation and classification results are

combined, their inconsistent and misleading information is eliminated. Moreover, the misclassifications around the nuclei are cleaned out, leaving nice and clean maps for the successive counting procedure.

4) Count TC(+) and TC(−) and produce $\hat{N}_{(+)}$ and $\hat{N}_{(-)}$ by recording the enclosed areas standing for nuclei on the fused masks $\boldsymbol{Y}_{(+)}$ and $\boldsymbol{Y}_{(-)}$, respectively. The counting errors for TC(+) and TC(−) are denoted by $\ell_{(+)}$ and $\ell_{(-)}$, respectively:

$$\ell_{(+)} = \|N_{(+)} - \hat{N}_{(+)}\|^2, \tag{5}$$

$$\ell_{(-)} = \|N_{(-)} - \hat{N}_{(-)}\|^2. \tag{6}$$

5) Finally, compute TPS from the cell counts by Eq. (1).

The architecture details of CB are given in Fig. 2. Since UNeXt[14] is proven to be effective, segmentation and classification blocks adopt the last three convolutional layers of the decoder of UNeXt with different weights as shown in Fig. 2. After every convolutional and pooling layer, there is BN and nonlinear mapping by ReLU. It can be seen that in CB there are multiple objectives: minimization of TPS regression, nucleus segmentation, classification, positive cell counting and negative cell counting errors, which are all involved in the overall loss ℓ,

$$\ell = \ell^{(\mathrm{dis})} + \lambda_1 \ell^{(\mathrm{reg})} + \lambda_2 \ell^{(\mathrm{seg})} + \lambda_3 \ell^{(\mathrm{cla})} + \lambda_4 \ell_{(+)} + \lambda_5 \ell_{(-)}, \tag{7}$$

where $\lambda_i (i = 1, 2, \cdots, 5)$ represents the trade-off coefficient; $\ell^{(\mathrm{dis})}$ is the mean squared error loss. $\ell^{(\mathrm{dis})}$ is added to reduce the discrepancy between the TPS values predicted by CB and RB and stabilize the two branches. It makes the network robust to strong variations in the PD-L1 slides stained with different shades of colors and crowded with TCs crushing on top of each other. With each objective optimizing a unique yet inter-connected part of MON, task-aware features are extracted and different information is combined to gain generalizability. Since CB considers far more factors than RB, the former usually results in more accurate TPS than the latter. Therefore, the values of TPS predicted by RB are used as the final output of MON.

2.3 NPE

The fused masks $\boldsymbol{Y}_{(+)}$ and $\boldsymbol{Y}_{(-)}$ as well as the original image patch $\boldsymbol{X}$ are used to compute nucleus parameters and provide diversified information to facilitate LUSC diagnosis. The parameters are listed as follows[16].

1) NA A_{N}: the number of pixels inside a nucleus as shown in Fig. 3, computed on the fused masks.

2) NP P_{N}: the number of pixels on the contour of a nucleus computed on the fused masks.

3) NC C_{N}:

$$C_{\mathrm{N}} = 4\pi A_{\mathrm{N}} / P_{\mathrm{N}}^2. \tag{8}$$

4) NMAA D_{NMAA}: the longest distance between any two pixels on the contour of a nucleus.

5) NMIA D_{NMIA}: the sum of two longest distances between NMAA and any pixels on the contour of a nucleus as shown in Fig. 3.

6) NEL L_{N}:

$$L_{\mathrm{N}} = D_{\mathrm{NMAA}} / D_{\mathrm{NMIA}}. \tag{9}$$

7) NE E_{N}:

$$E_{\mathrm{N}} = \sqrt{(D_{\mathrm{NMAA}}/2)^2 - (D_{\mathrm{NMIA}}/2)^2} / (D_{\mathrm{NMAA}}/2). \tag{10}$$

8) RGB: the average value of all the pixels inside a nucleus computed on the original RGB image $\boldsymbol{X}$.

9) HSV: the average value of all the pixels inside a nucleus computed on the HSV image transformed from $\boldsymbol{X}$.

It can be seen that A_{N} and P_{N} are size parameters, C_{N}, D_{NMAA}, D_{NMIA}, L_{N} and E_{N} are shape parameters, and RGB and HSV are color parameters. NP, NA, NMAA and NMIA are illustrated in Fig. 3. The whole procedures of MOLP including cell recognition, TPS estimation and parameter computation are summarized in Algorithm 1. The network is optimized via gradient descent, implemented by the Adam optimizer[17].

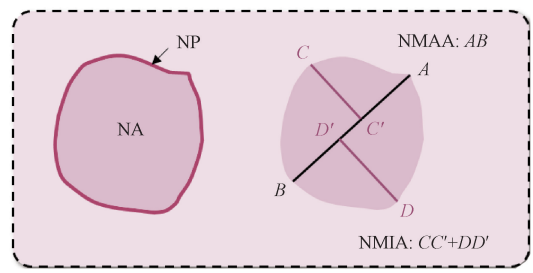


Fig. 3 Geometric diagram for computing nucleus parameters

Algorithm 1 MOLP

Input: each patch $\boldsymbol{X}$ from PD-L1 slides and multiple labels y, $\boldsymbol{Y}^{(\mathrm{seg})}$, $\boldsymbol{Y}^{(\mathrm{cla})}$, $N_{(+)}$, $N_{(-)}$.

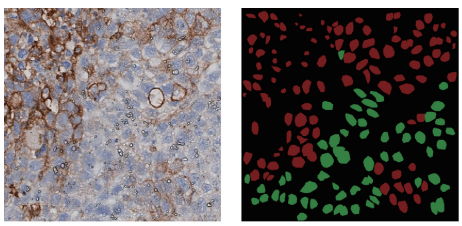
Step 1: feed $\{\boldsymbol{X}, y, N_{(+)}, N_{(-)}, \boldsymbol{Y}^{(\mathrm{seg})}, \boldsymbol{Y}^{(\mathrm{cla})}\}$ to MON;
Step 2: optimize MON with Eqs. (2)–(7);
Step 3: obtain $\{\hat{y}_1, \hat{N}_{(+)}, \hat{N}_{(-)}, \hat{\boldsymbol{Y}}^{(\mathrm{seg})}, \hat{\boldsymbol{Y}}^{(\mathrm{cla})}, \boldsymbol{Y}_{(+)}, \boldsymbol{Y}_{(-)}\}$;
Step 4: compute nucleus parameters in Subsection 3.2 with $\boldsymbol{X}$, $\boldsymbol{Y}_{(+)}$ and $\boldsymbol{Y}_{(-)}$ according to Eqs. (8)–(10).

Output: $\hat{y}_1$, $\hat{N}_{(+)}$, $\hat{N}_{(-)}$, $\hat{\boldsymbol{Y}}^{(\mathrm{seg})}$, and $\hat{\boldsymbol{Y}}^{(\mathrm{cla})}$; nucleus parameters (NA, NP, NC, NMAA, NMIA, NEL, NE, RGB and HSV).

3 Experiments

3.1 Dataset

The original PD-L1 stained slides of LUSC are provided by Shanghai Histo Pathology Diagnostic Center and Fudan University Shanghai Cancer Center, China. The dataset[18] used in this work contains 43 PD-L1 images with 1 000 pixel × 1 000 pixel cropped from 4 WSIs scanned with KF-PRO-120 (0.248 1 μm/pixel, 40 × magnification). Each image is with a fully annotated classification mask, as shown in Fig. 4. The labels used for multi-objective learning, nucleus segmentation mask, cell counts and TPS are easily generated from this classification mask. These images are randomly divided into two subsets: one with 35 images for training and the other with 9 images for inference. After padding, each 1 000 pixel × 1 000 pixel image is further cropped into 256 patches, each with 64 pixel × 64 pixel. The patches are direct fed to the proposed network. Thus, each input patch for Algorithm 1 can be denoted by $\boldsymbol{X}_i$ with its real TPS as y_i, where i is the index of patches.



(a) (b)

Fig. 4 Example from LUSC PD-L1 dataset: (a) 1 000 pixel × 1 000 pixel image with blue denoting tumor nuclei and brown for membrane (the darker the brown is, the more likely the tumor cell is to be positive); (b) classification mask (red denotes TC(+) and green denotes TC(−))

3.2 Setup

The proposed MOLP is compared with various pre-trained networks which are U-Net[19], U-Net + +[20], ResNet18, ResNet50 and ResNet101[21]. For TPS regression, several FC layers and a sigmoid mapping function are attached to the end of the competing models, since they are originally designed for segmentation or classification. As MOLP is optimized by the Adam optimizer, the learning rate is set as 0.001, which is multiplied by 0.4 after every five epochs, the batch size is set as 16, the weight decay is 8×10^{-5}, and the maximum number of training epochs is set as 50. In Eq. (7), $\lambda_1 = \lambda_3 = 1$, $\lambda_2 = 0.5$, and $\lambda_4 = \lambda_5 = 10$. The other networks are fine-tuned with their parameters adjusted with our best efforts. TPS estimation is evaluated by a mean absolute error (MAE)[22], 95% confidence interval (CI) indicating reliability of the results and a Pearson correlation coefficient (PCC) with $p < 0.001$ showing how consistent the results are with pathologists' annotations[23]. Nucleus parameter estimation is evaluated by a mean relative error (MRE)[24]. For evaluation of nucleus segmentation, a dice index (DI) is computed[25]. For cell classification, overall accuracy (OA), precision, recall and F1-score are computed and classification maps are visualized[8]. Meanwhile, with the nuclei and the background being regarded as two different classes, the classification indexes are also applied to the segmentation maps. The indexes for each 64 pixel×64 pixel patch in the test set are computed; the results for each 1 000 pixel × 1 000 pixel image obtained by titling its patch-wise segmentation/classification maps are also evaluated. All the algorithms are coded by PyTorch on a platform with 43 GB RAM and RTX 2080 Ti GPU.

4 Results and Discussion

4.1 Main analytic results

The major results of MOLP, the estimated TPS and nucleus parameters are presented in Tables 2 and 3, respectively. The proposed system MOLP is able to predict TPS with high confidence and compute nucleus parameters for diagnosis practice. High CI and PCC suggest that the approximated results are reliable. As shown in Fig. 5, while for most images, the TPS estimation errors are lower than 8%; for the sixth test image, the absolute error is as high as 26.35%. This is because there are too many TC(−) which are not properly identified. However, the error would not affect diagnosis results according to the medical thresholds in Subsection 3.1. In reality, the predictions of the system would be only used as reference for pathologists who normally leverage their a high-level expertise in addition to the computational results in the final decision.

Table 2 TPS estimation by MON

Parameter	64 pixel×64 pixel patches					1 000 pixel ×1 000 pixel images				
	y/%	$\hat{y}_1$/%	MAE	95% CI	PCC	y/%	$\hat{y}_1$/%	MAE	95% CI	PCC
Value	83.45	85.91	8.62	(7.48, 9.76)	0.72	84.43	87.12	4.97	(−0.56, 10.49)	0.97

Table 3 Average nucleus parameters given by NPE

Nucleus parameter		Real		Estimated		MRE/%	
		TC(+)	TC(−)	TC(+)	TC(−)	TC(+)	TC(−)
Size	NA	970.771	1 244.002	798.037	1 951.926	17.79	56.91
	NP	131.770	145.651	107.973	192.294	18.06	32.02
Shape	NC	0.664	0.713	0.751	0.663	13.10	7.01
	NMAA	47.514	52.995	37.743	63.154	20.56	19.17
	NMIA	30.947	33.825	26.726	42.782	13.64	26.48
	NEL	1.604	1.626	1.422	1.474	11.35	9.35
	NE	0.710	0.728	0.641	0.665	9.72	8.65
Color	R	133.900	151.929	138.655	152.905	3.55	0.64
	G	120.990	152.938	131.206	153.847	8.44	0.59
	B	124.464	170.424	140.772	170.497	13.10	0.04
	H	88.305	118.283	113.509	119.688	28.54	1.19
	S	53.158	33.400	33.465	31.331	37.05	6.19
	V	140.469	171.233	147.112	170.952	4.73	0.16

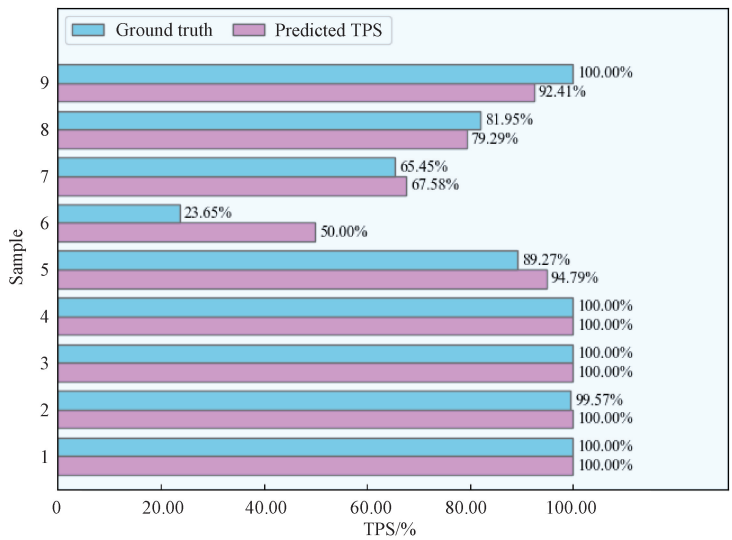


Fig. 5 Bar charts of TPS for each 1 000 pixel ×1 000 pixel image

Figure 6 visualizes distributions of shape parameters for TC(+) and TC(−) by violin plots, where the width of each "violin" denotes frequency and the box inside shows median, interquartile range and 95% CI[26]. It demonstrates that the distributions of the estimated parameters, especially NE, share some resemblance to those of the real values. L_N/E_N of TC(+) and TC(−) are very close, mostly ranging from 0.7 to 0.8, meaning that the shapes of TC(+) and TC(−) are not very different, which can be reflected from the original slides as shown in Fig. 4. Although the nucleus parameters seem not to be able to direct indicate positive and negative LUSC cells, they provide quantitative information for the pathology analysis and could be used for classification of different cells in other application scenarios.

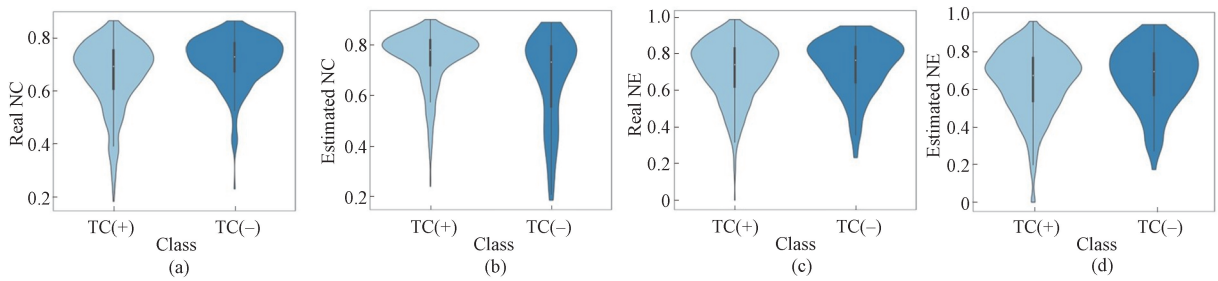


Fig. 6 Violin plots of distributions of nucleus parameters in 1 000 pixel×1 000 pixel images; (a) real NC; (b) estimated NC; (c) real NE; (d) estimated NE

4.2 Auxiliary task performance

Figure 7 provides bar charts of indexes of individual tasks for 1 000 pixel×1 000 pixel images in the test set. It validates the robustness of MOLP in multi-tasking while confirming that MOLP can predict TPS along with diversified cellular information to enrich users' understanding of the PD-L1 slides. Figure 8 visualizes segmentation, classification and counting results of MOLP for a 1 000 pixel ×1 000 pixel image in the test set. It can be seen that most nuclei are correctly identified by the predicted contour maps with low counting errors. Although segmentation and classification errors are inevitable, they do not affect TPS estimation or diagnosis results, as indicated by the insignificant difference between the real TPS and the estimated TPS in Fig. 5. For example, there is only an error of 2.13% for TPS estimation with the real TPS of 65.45% and the estimated value of 67.58%. Both are much higher than the critical threshold (50%), indicating a high level of PD-L1 expression, according to Subsection 2.1. In computer-aided LUSC diagnosis, these intermediate results are mostly displayed on application interfaces as references for the pathologists, showing the process of TPS estimation and confidence about the estimated TPS, and helping to make further decisions.

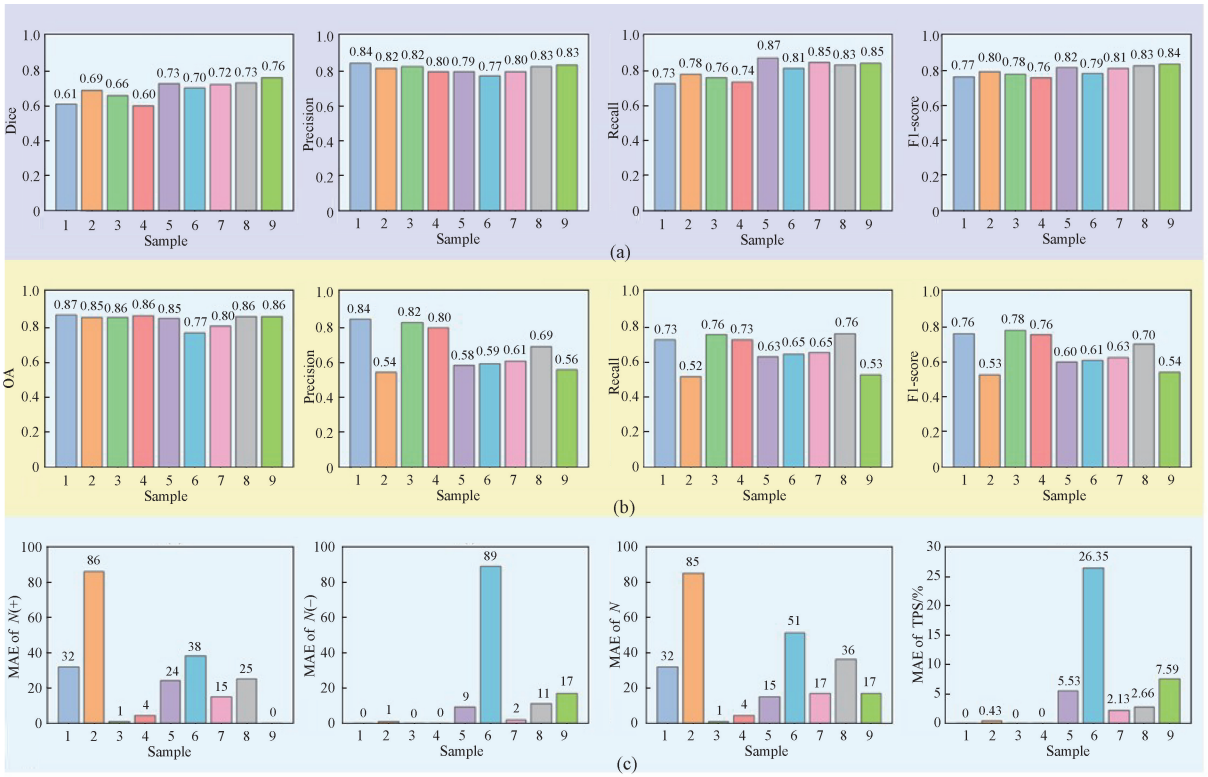


Fig. 7 Bar charts of nucleus segmentation, classification and counting results by MOLP: (a) segmentation value; (b) classification value; (c) MAE of counting and TPS estimation

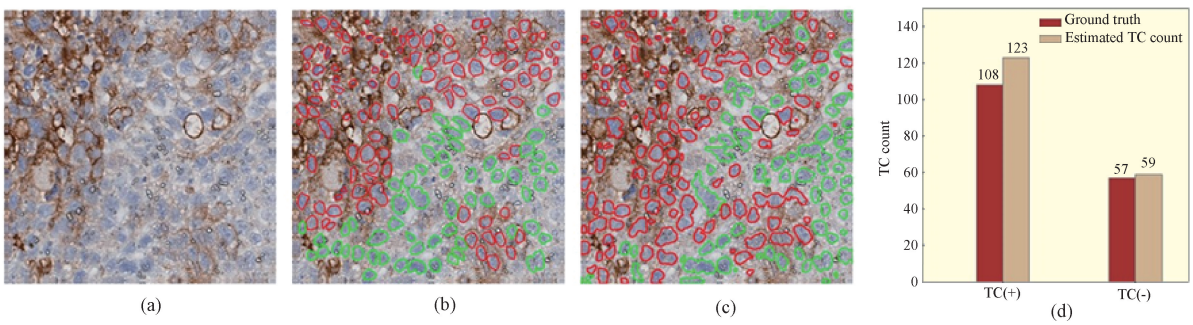


Fig. 8 Visualization of nucleus classification contour maps and TPS results: (a) an original 1 000 pixel×1 000 pixel image; (b) real contour maps overlaid on images; (c) contour maps generated from classification maps inferred by MOLP (red indicates TC(+) and green indicates TC(−)); (d) bar charts of real and estimated cell counts

4.3 Efficacy of MON

After the feasibility and the efficacy of the whole system are demonstrated, the performance of MON of MOLP is analyzed in this part.

4.3.1 Comparative study

Table 4 gives MAE of TPS estimation by different methods. MOLP outperforms the competing networks widely used in cell-related tasks.

Table 4 MAE of TPS estimation by different methods

Method	MAE of 64 pixel×64 pixel patches	MAE of 1 000 pixel×1 000 pixel images
U-Net[19]	14.17	15.57
U-Net++[20]	13.83	15.57
ResNet18[21]	14.15	7.07
ResNet50[21]	14.52	14.21
ResNet101[21]	16.69	16.42
MOLP	8.62	4.97

4.3.2 Ablation study

Table 5 presents the results of different ablated models of MOLP. The complete MON outperforms the ablated models at the 1 000 pixel ×1 000 pixel image level, validating the effectiveness of UNeXt as the backbone of RB and the segmentation block in CB.

Table 5 MAE of TPS estimation of ablated models of MOLP

Model	MAE of 1 000 pixel× 1 000 pixel images
With U-Net instead of UNeXt for RB	7.44
w/o RB	5.07
w/o RB + w/o the segmentation block in CB	6.07
MOLP	4.97

Note: w/o stands for without.

4.4 Convergence of MON

Figure 9 proves the convergence of the overall loss and each individual loss of MON, as defined in Eqs. (2)–(7). The smooth convergence curves clearly demonstrate stability and efficiency of MOLP, as each loss converges after proper training and different goals of multi-objective learning are accomplished. Specifically, through the convergence of the discrepancy loss defined in Eq. (7), the gap between RB and CB is gradually narrowed and the different features produced by the two branches are integrated into one another. Learning a variety of features, the network gains robustness to diversified images and generalizability on multiple tasks, including reducing TPS prediction errors.

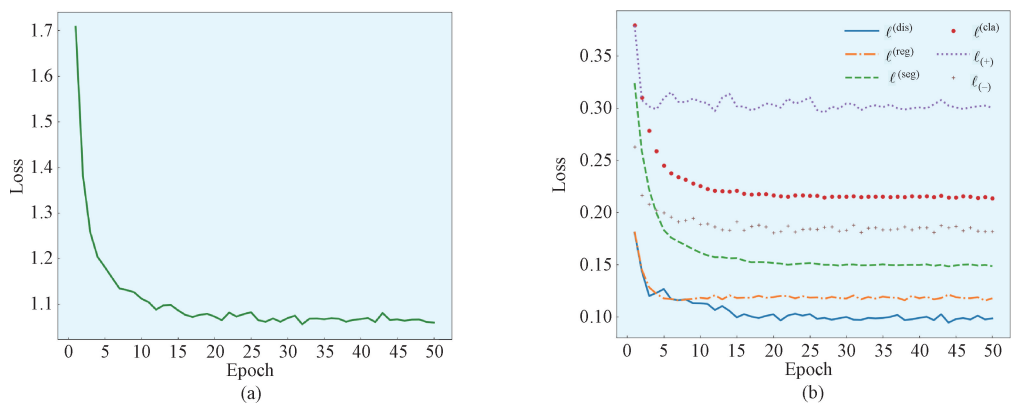


Fig. 9 Convergence curves: (a) overall loss; (b) individual loss of MON

4.5 Discussion

The results presented above show that the proposed MOLP can accurately predict TPS from the PD-L1 slides and provide nucleus parameters. It improves TPS estimation by multi-objective learning, lowering the TPS errors via nucleus segmentation, classification and counting. However, there are still some limitations to overcome. The cell-recognition module could be embedded with better multi-objective learning mechanisms to improve nucleus segmentation, classification, counting and eventually TPS estimation. The NPE could be replaced with some trainable networks and optimized with the cell-recognition module, forming a larger end-to-end network and producing nucleus parameters directly for pathologists.

5 Conclusions

To facilitate LUSC diagnosis and treatment, this work proposes a systematic method MOLP for predicting TPS, cell counts, nucleus contours and categories from the PD-L1 slides of LUSC. MOLP consists of two stages: multi-objective learning and nucleus parameter analysis. The main network MON comprises two branches, namely CB and RB, whereas the former estimates TPS via the cell analysis (including nucleus segmentation, classification and counting) and the latter direct regresses TPS. The difference between these two approximated values of TPS is minimized to gain robustness. Moreover, segmentation and classification results of CB are used to estimate the appearance parameters of TCs to provide extra information for LUSC diagnosis. Experiments on a large set of PD-L1 slides have demonstrated feasibility, efficacy and stability of MOLP.

References

[1] DOROSHOW D B, BHALLA S, BEASLEY M B, et al. PD-L1 as a biomarker of response to immune-checkpoint inhibitors [J]. *Nature Reviews Clinical Oncology*, 2021, 18: 345-362.

[2] WANG J S, MAO X T, WANG Y, et al. Automatic generation of pathological benchmark dataset from hyperspectral images of double stained tissues[J]. *Optics & Laser Technology*, 2023, 163: 109331.

[3] LI Q L, HE X F, WANG Y T, et al. Review of spectral imaging technology in biomedical engineering: achievements and challenges[J]. *Journal of Biomedical Optics*, 2013, 18(10): 100901.

[4] MI H B, XU K L, XIANG Y, et al. A quantitative analysis platform for PD-L1 immunohistochemistry based on point-level supervision model [C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. [S.l.]: International Joint Conferences on Artificial Intelligence Organization, 2019: 6554-6556.

[5] PAN B J, KANG Y X, JIN Y, et al. Automated tumor proportion scoring for PD-L1 expression based on multistage ensemble strategy in non-small cell lung cancer [J]. *Journal of Translational Medicine*, 2021, 19(1): 249.

[6] KANG Y X, LI H S, HAN X, et al. Automated tumor proportion scoring for assessment of PD-L1 expression based on multi-stage ensemble strategy [M]//Machine Learning in Medical Imaging. Cham: Springer Nature Switzerland AG, 2020: 70-79.

[7] YI J R, WU P X, JIANG M L, et al. Attentive neural cell instance segmentation[J]. *Medical Image Analysis*, 2019, 55: 228-240.

[8] GRAHAM S, VU Q D, RAZA S E A, et al. Hover-Net: simultaneous segmentation and classification of nuclei in multi-tissue histology images[J]. *Medical Image Analysis*, 2019, 58: 101563.

[9] DOGAR G M, SHAHZAD M, FRAZ M M. Attention augmented distance regression and classification network for nuclei instance segmentation and type classification in histology images[J]. *Biomedical Signal Processing and Control*, 2023, 79: 104199.

[10] LIU J X, ZHENG Q, MU X, et al. Automated tumor proportion score analysis for PD-L1 (22C3) expression in lung squamous cell carcinoma[J]. *Scientific Reports*, 2021, 11: 15907.

[11] WIDMAIER M, WIESTLER T, WALKER J, et al. Comparison of continuous measures across diagnostic PD-L1 assays in non-small cell lung cancer using automated image analysis [J]. *Modern Pathology*, 2020, 33(3): 380-390.

[12] CHEN J H, ZHU Y A, CHEN Z. Graph-embedded online learning for cell detection and tumour proportion score estimation [J]. *Electronics*, 2022, 11(10): 1642.

[13] WU J H, LIU C L, LIU X Q, et al. Artificial intelligence-assisted system for precision diagnosis of PD-L1 expression in non-small cell lung cancer [J]. *Modern Pathology*, 2022, 35(3): 403-411.

[14] VALANARASU J M J, PATEL V M. UNeXt: MLP-based rapid medical image segmentation network [M]//Lecture Notes in Computer Science. Cham: Springer Nature Switzerland AG, 2022: 23-33.

[15] MILLETARI F, NAVAB N, AHMADI S A. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). New York: IEEE, 2016: 565-571.

[16] ZHANG Q, WANG Y, QIU S, et al. 3D-PulCNN: pulmonary cancer classification from hyperspectral images using convolution combination unit based CNN [J]. *Journal of Biophotonics*, 2021, 14(12): e202100142.

[17] KINGMA D P, BA J. Adam: a method for stochastic optimization[EB/OL]. (2017-01-30) [2023-07-01]. https://arxiv.org/abs/1412.6980v9.

[18] CHEN Z, CHEN Z, LIU J X, et al. Weakly supervised histopathology image segmentation with sparse point annotations[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(5): 1673-1685.

[19] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical

image segmentation [M]//Lecture Notes in Computer Science. Cham: Springer International Publishing Switzerland, 2015: 234-241.
[20] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHSH N, et al. UNet++: redesigning skip connections to exploit multiscale features in image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1856-1867.
[21] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2016: 770-778.
[22] WILLMOTT C J, MATSUURA K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance [J]. *Climate Research*, 2005, 30: 79-82.
[23] LEE RODGERS J, NICEWANDER W A. Thirteen ways to look at the correlation coefficient [J]. *The American Statistician*, 1988, 42(1): 59-66.
[24] NASER M Z, ALAVI A H. Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences[J]. *Architecture, Structures and Construction*, 2023, 3(4): 499-517.
[25] CARASS A, ROY S, GHERMAN A, et al. Evaluating white matter lesion segmentations with refined Sørensen-dice analysis [J]. *Scientific Reports*, 2020, 10: 8242.
[26] HINTZE J L, NELSON R D. Violin plots: a box plot-density trace synergism[J]. *The American Statistician*, 1998, 52(2): 181-184.

多目标系统化学习的 PD-L1 切片分析方法

陈 昭[1], 郭丹琦[1], 王 倩[1*], 沈熠婷[1], 王庆国[2]
1. 东华大学 计算机科学与技术学院, 上海 201620
2. 上海市第一人民医院 放射科, 上海 200080

摘 要: 在肿瘤尤其是如肺鳞癌 (lung squamous cell carcinoma, LUSC) 的非小细胞肺癌的治疗中, 基于程序性死亡受体-配体 1 (programmed cell death-ligand 1, PD-L1) 染色切片的阳性肿瘤细胞比例评分 (tumor proportion score, TPS) 可为治疗方案的选择提供重要依据。肿瘤细胞 (tumor cell, TC) 的许多参数对癌症诊断至关重要。虽然可以通过计算分析来预测这些参数, 但很少有一个统一的框架可以同时获得细胞的不同病理信息。为此, 提出了一种多目标学习框架 (multi-objective learning pipeline, MOLP), 从 LUSC 的 PD-L1 切片中预测 TPS、细胞数目、细胞核轮廓和类别。主干网络包括两个分支: 一个分支通过细胞分析估算 TPS, 另一个分支直接通过回归分析估算 TPS。MOLP 通过最小化两个分支的 TPS 预测差值来提高其鲁棒性。细胞分析支路可实现细胞核分割、分类和计数, 不仅增强了 TPS 估计的可信度, 还使得 MOLP 能够估计肿瘤细胞的外观参数以用于 LUSC 诊断。在大规模图像集上的实验结果证明了 MOLP 的可行性和有效性。MOLP 预测的 TPS 与病理医师的评分呈现出统计学上的显著相关性: 平均绝对误差仅为 4.97 (95%置信区间: −0.56~10.49), 皮尔逊相关系数为 0.97 (p<0.001)。

关键词: 程序性死亡受体−配体 1 (PD-L1) 切片; 阳性肿瘤细胞比例评分 (TPS); 多目标学习; 分类; 分割; 计数