PLOS ONE

Tertiary lymphoid structures (TLS) identification and density assessment on H&E-stained digital slides of lung cancer

Panagiotis Barmpoutis[1,2]*, Matthew Di Capite[3], Hamzeh Kayhanian[1], William Waddingham[1], Daniel C. Alexander[2], Marnix Jansen[1], Francois Ng Kee Kwong[3]

1 Department of Pathology, UCL Cancer Institute, University College London, London, United Kingdom, 2 Centre for Medical Image Computing, University College London, London, United Kingdom, 3 Department of Histopathology, Norfolk and Norwich University Hospital, Norwich, United Kingdom

* p.barmpoutis@ucl.ac.uk

Abstract

Tertiary lymphoid structures (TLS) are ectopic aggregates of lymphoid cells in inflamed, infected, or tumoral tissues that are easily recognized on an H&E histology slide as discrete entities, distinct from lymphocytes. TLS are associated with improved cancer prognosis but there is no standardised method available to quantify their presence. Previous studies have used immunohistochemistry to determine the presence of specific cells as a marker of the TLS. This has now been proven to be an underestimate of the true number of TLS. Thus, we propose a methodology for the automated identification and quantification of TLS, based on H&E slides. We subsequently determined the mathematical criteria defining a TLS. TLS regions were identified through a deep convolutional neural network and segmentation of lymphocytes was performed through an ellipsoidal model. This methodology had a 92.87% specificity at 95% sensitivity, 88.79% specificity at 98% sensitivity and 84.32% specificity at 99% sensitivity level based on 144 TLS annotated H&E slides implying that the automated approach was able to reproduce the histopathologists' assessment with great accuracy. We showed that the minimum number of lymphocytes within TLS is 45 and the minimum TLS area is $6,245 \mu m^2$. Furthermore, we have shown that the density of the lymphocytes is more than 3 times those outside of the TLS. The mean density and standard deviation of lymphocytes within a TLS area are $0.0128/\mu m^2$ and $0.0026/\mu m^2$ respectively compared to $0.004/\mu m^2$ and $0.001/\mu m^2$ in non-TLS regions. The proposed methodology shows great potential for automated identification and quantification of the TLS density on digital H&E slides.

Introduction

Understanding the host immune response to cancer is a critical area of investigation. This has resulted in the recent introduction of various immunotherapeutic drugs (targeting checkpoint inhibition) in the treatment of lung, renal and skin cancers. In addition, the host immune response is also partly mediated by the Tertiary Lymphoid Structures (TLS) [1]. The latter are

discrete entities of lymphoid cells which are recognised on histological H&E-stained sections, as they share some histological features with lymph nodes [2]. In general, TLS are not present under normal conditions in some organs and have been observed in pathogen infection, auto-immune disorders, allograft rejection, and in several types of cancer [3–5]. However, in contrast to autoimmune disorders high densities of TLS in cancers including breast, colorectal and lung cancer, are usually associated with positive patient prognosis, outcomes and improved immunotherapy response [3,6,7].

The presence and importance of TLS in lung cancer were first reported by Dieu-Nosjean *et al.* in 2016 [6], who used immunohistochemistry, gene expression assays, and flow cytometry on large series of lung tumors. They demonstrated that TLS are the sites for the generation of the local and systemic T- and B-cell responses against tumours. Furthermore, in lung cancer, previous studies have identified three maturation stages of TLS culminating in germinal centre formation with significant relevance to patient survival [8]. The authors have described TLS development along the stages of secondary lymphoid organ formation and shown that the second (primary follicle-like stage) and third (secondary follicle-like maturation stages) are dependent on co-expression of CD21, CD23 and CXCL13, but that the first maturation stage (early stage, E-TLS), is characterized by dense lymphocytic aggregates without CD21 and CD23 expression.

TLS density can be assessed in diagnostic H&E sections and can, thus, be easily introduced in routine pathology to serve as a relevant prognostic parameter [9]. TLS are identifiable on H&E sections by histopathologists as discrete entities with curved and smooth outline and contain tightly packed mature lymphoid cells. However, there is no current consensus agreement on the definition of TLS, even though their presence has been evaluated in previous studies by morphology on H&E slides as early as 1990 [10]. For example, it is uncertain if there is a minimum number of mature lymphoid cells in the TLS. Although the lymphoid cells appear much more densely packed in a TLS compared to lymphocytes within normal or inflamed tissue, the minimum density of lymphocytes defining TLS remains unspecified. In addition, the minimum size of a TLS is not agreed. The assessment of TLS density over a large histological area is also very time consuming and subject to interpretation variation. Previous studies have assessed TLS based on representative areas of tumour rather than doing this on the whole tumour area [9,11]. However, only a limited number of studies are in progress for developing automated methods for TLS detection and analysis. Silina *et al.* describes a quantitative pathology approach for the identification and quantification of different TLS maturation stages using seven-color immunofluorescent staining and segmentation algorithms of Inform software [11]. As such, being able to evaluate TLS density across the overall area of the tumour would be more accurate. Furthermore, identification of TLS from routine H&E histological images would allow easier integration into clinical workflows.

Various techniques and methods, based on either hand-crafted or deep learning features, have been developed for digital pathology image segmentation tasks aiming to label regions of an image according to what is being shown and aid pathologists to make diagnostic and treatment processes more efficient. Hand-crafted developed methodologies, where a set of local or global features are extracted, mostly include thresholding methods [12], region growing methods based on seed points growing [13], exploitation of morphology features [14,15], watershed transformation [16,17], active contour models [18,19], Markov Random Fields [20] and dynamic image segmentation methods [21]. On the other hand, many segmentation methods have used deep-learning techniques aiming to address the problem by extracting knowledge directly from the data. There are numerous deep learning methods that have been developed for medical image segmentation. More specifically, these include autoencoders [22,23], deep convolutional neural networks (CNNs) [24,25], cascaded networks [26] and fully

convolutional networks [27]. However, the training of complex deep learning networks requires a large number of images and computational power as well as considerable effort and time for their annotation by experts [28].

To this end, in this study, we first propose an automated approach for the identification and quantification of TLS in H&E histological images by applying a method that combines a DeepLab v3+ network, an active contour model and a lymphocytes segmentation approach. Secondly, we aim to translate the visual recognition of TLS by histopathologists into a universally reproducible set of mathematical values for the standardisation of TLS recognition: area occupied by TLS, the minimum number of lymphocytes present and their density (number/ unit area). A heat map of lymphocytes is then built, thus allowing us to define TLS in lung tissue (cancer and normal). Based on the above data, we propose formal mathematical criteria for the definition of TLS.

Materials and methods

The framework of the proposed methodology for the detection of TLS regions and their lymphocytes is shown in Fig 1. Initially, an H&E image was fed into a modified DeepLab v3+ network for the detection of candidate TLS regions and an active contour model was then applied in order to refine the boundaries of the TLS regions. Then, segmentation of lymphocytes was performed for the identification of the following features: the number of lymphocytes, the size of TLS regions and the number of lymphocytes per unit area of TLS. The estimated features were used for post validation of candidate TLS regions aiming to filter out the false-positive detected TLS regions.

Identification of the candidate TLS regions

For the identification of the candidate TLS regions, a modified pre-trained DeepLab v3 + model [29] with Inception-ResNet-v2 as the main feature extractor, which employs dropout to avoid overfitting, was utilized. The DeepLab models have been extensively used in the task of semantic medical image segmentation and tested on large volumes of image datasets [30–34]. These models provide a capability in learning multi-scale contextual features through Atrous Spatial Pyramid Pooling (ASPP) and use a decoder module for the refinement of the segmentation results, especially along object boundaries. In this work, the ASPP is a module that employs multiple parallel atrous convolutional layers with different rates to learn multi-scale information of image aiming to identify different sizes of TLS regions and to retain the balance between context assimilation and fine localization. This network was selected due to the good balance it achieves between accuracy and computational complexity. Specifically, Inception-ResNet-v2 outperforms other common configurations with regards to accuracy and complexity [35]. The model was pre-trained on ImageNet and then fine-tuned with training data prepared for this work. Deep neural networks pre-training [36] can be seen as a case of transfer learning [37], in which a neural network further trained on a source dataset is subsequently fine-tuned to a target dataset. The pre-training on the source dataset enables the deep neural networks to learn useful low level features in their first layers, such that good results can be achieved with fewer examples in the fine-tuning stage, which mostly adapts the higher level features in the last layers, hence requiring less labeled training data. It is worth mentioning that color normalization was applied in all the dataset images.

Furthermore, an augmentation method was utilized to enlarge the image samples in order to better fine-tune the DeepLab v3+. Data augmentation artificially enlarges the size of the training dataset by applying spatial warps, which has been proven a very effective strategy in many image analysis tasks [38]. Even though the number of training cases might appear small,
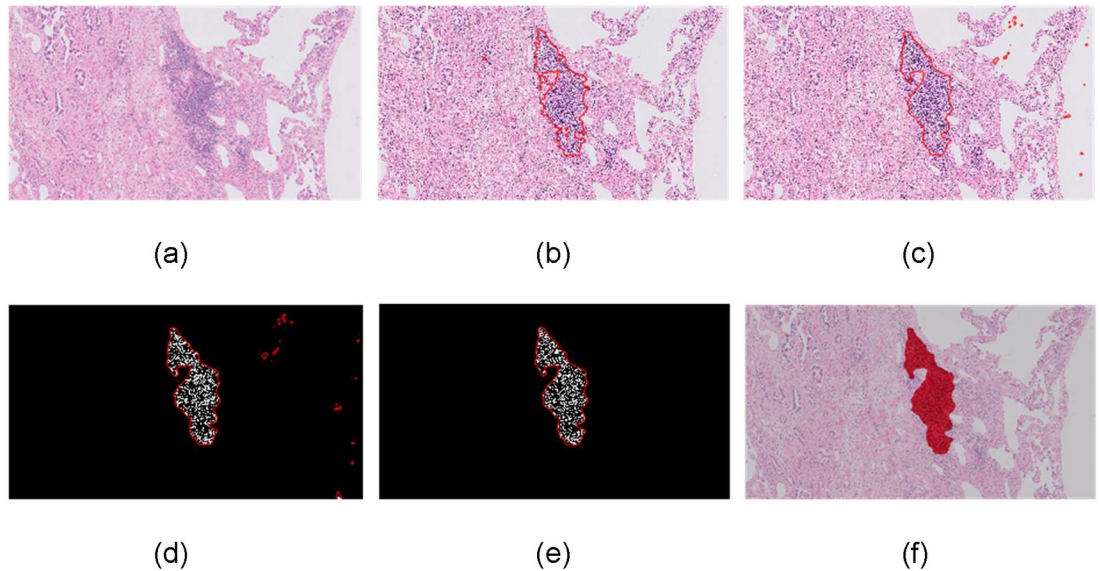
Fig 1. The proposed methodology for TLS detection. a) H&E images feed a DeepLab v3+ network, b) candidate TLS regions are identified, c) an active contour approach is applied for boundary refinement of the TLS regions, d) segmentation of lymphocytes is performed, e) post validation of TLS through rejection of candidate non-TLS regions, f) final detection of TLS regions.

https://doi.org/10.1371/journal.pone.0256907.g001

we combined the abundant pixel-level information with pre-training and with data augmentation in order to increase the variability of training images and to avoid overfitting of the network [39].

Additionally, a modified loss function was defined in order to adjust the model to better deal with the boundaries of TLS. Thus, introducing a weighting factor w, we force the model to be sensitive to the TLS boundaries and regions enclosed within the TLS. More specifically, the loss function is defined as follows:

$$Loss = -\sum_{p=1}^{N} w_p r_p \log(t_p) \tag{1}$$

where w_p, r_p and t_p denote the weighting factors, the reference values and the predicted values at pixel p respectively, and N is the total number of pixels. Regarding the weighting factors, we set the $w_p = 2$ when p is a TLS pixel and otherwise, we set $w_p = 1$.

TLS boundary refinement

Since the lymphoid cells appear in higher density in TLS regions in order to obtain precise TLS contours, we adopted an active contour approach [40] that utilizes local intensity distribution to drive the evolving curve. More specifically, the local intensities within its neighborhood are assumed to follow a Gaussian probability distribution:

$$p_{i,x}\big(I(y)|m_i(x), \sigma_i(x)^2\big) = \frac{1}{\sqrt{2\pi}\sigma_i(x)} \exp\left(-\frac{(I(y) - m_i(x))^2}{2\sigma_i(x)^2}\right) \tag{2}$$

where $m_i(x)$ and $\sigma_i(x)$ are mean and standard deviation of the intensities in each local region.

Thus, the local Gaussian distribution fitting energy is estimated as follows:

$$E = -\int_{\Omega}\int_{inside(C)} K_{\sigma}(x-y)logp_{1,x}(I(y)|m_1(x),\sigma_1(x)^2)dydx - \int_{\Omega}\int_{outside(C)} K_{\sigma}(x$$
$$- y)logp_{2,x}(I(y)|m_2(x),\sigma_2(x)^2)dydx \qquad (3)$$

where $\Omega \in \mathcal{R}^2$ represents the image domain, C a closed contour and the neighborhood of each pixel is defined by using a truncated Gaussian kernel K_{σ}. Thus, the applied model is able to differentiate between regions with intensity heterogeneity and also between regions with similar intensity means but different intensity variances [41]. The active contour model is initialized by the candidate TLS regions detected in the previous step.

Lymphocytes segmentation

In digital histopathology the cell segmentation is the task of the automated splitting of microscopic tissue images into segments, which represent individual cells. Many cell segmentation methods have been developed, utilizing both traditional techniques and deep learning methods in the field of medical image analysis. They achieve comparable accuracy rates and they identify single cells through watershed transformation [42], using active contours [43], modelling the cells with a set of circle or ellipses [14,44–46], while many other methods utilize deep neural networks [47–49].

In this work, we propose an improved methodology based on an ellipsoidal model [14] that iteratively identifies and counts the cells (Fig 2) aiming to keep good balance between the estimated cells' shape and overlapping parts of touching cells through a single validation criterion and at the same time to overcome the limitations of previously developed methods [14] that in many cases erroneously reject small touching cells. To this end, initially, input RGB images
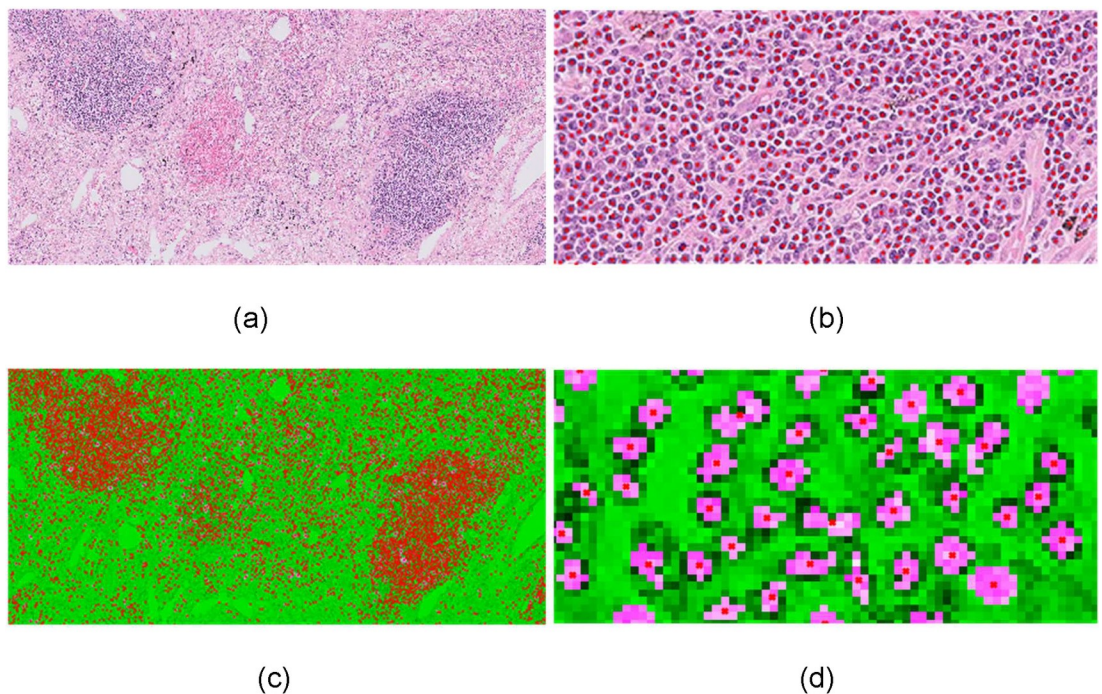

(a)

(b)

(c)

(d)

Fig 2. Lymphocyte segmentation. (a) input 10x H&E-stained image, (b) lymphocytes identification -red dots, (c) separation of merged or overlapped cells through the proposed ellipsoidal modelling, (d) detailed image of single-cell segmentation.

were converted to grayscale and filtered using a Gaussian filter with a 3×3 kernel in order to remove small artifacts. Furthermore, a histogram equalization filter was applied in order to enhance the differences between lymphocytes and other tissue components. Subsequently, an effective method using an adaptive threshold approach was applied. More specifically, this method set the threshold based on the local mean intensity in the neighbourhood of each pixel. Thus, the formula used for thresholding was defined as follows:

$$O(i,j) = \begin{cases} 255, I(i,j) \geq T \\ 0, I(i,j) < T \end{cases} \tag{4}$$

where $O(i,j)$ is the resulting pixel of output image at (i,j), $I(i,j)$ is the pixel of the input image and T is the selected local threshold value. In the binary image, in order to suppress small artifacts, morphological operations consisting of erosion, dilation and removal of small elements were applied.

For the separation of touching cells an improved ellipsoidal modelling approach is proposed. Initially, we estimated the distance transformation of the binary image M of p pixels that represents the connected cells and we estimated the regional maxima of this. Considering that the number and location of local maxima corresponds to these of nuclei, we rejected the touching maxima. The remaining maxima comprise the list of candidate seeds. Based on the hypothesis that cells can be spatially modeled as ellipsoids E_C, the pixels of cells were then modeled using a Gaussian distribution. More specifically, a Gaussian mixture model was applied with the number of clusters C being equal to that of candidate seeds and the mixture parameters, namely mean and variance, to be estimated using the expectation-maximization (EM) algorithm. For the initialization of the EM algorithm we used k-Nearest Neighbor classification using Euclidean distance as the distance metric in order to estimate the initial parameters. The EM is an iterative method consisting of two steps: (i) expectation, which computes the likelihood with respect to the current estimates and (ii) maximization (Eq 5), which maximizes the expected log likelihood (Eq 6) as follows:

$$Q(\theta|\theta^{(t)}) = E_{Z|X,\theta^{(t)}}[logL(\theta; X, Z)] \tag{5}$$

$$\theta^{(t+1)} = arg \ \max_{\theta} Q(\theta|\theta^{(t)}) \tag{6}$$

where Q is the expected values of the log likelihood function θ, X is the pixel coordinates, Z is the latent variables and $\theta^{(t)}$ is the current parameters.

Having estimated the ellipsoidal models of cells, we need to identify the optimal number of seeds rejecting or approving the candidate seeds from the previously estimated list. Thus, in this approach, we proposed a single fitness validation criterion estimating this for the total number of combinations of candidate seeds, aiming to accurately identify the total number of cells. The criterion takes into account the foreground, the background and the overlapping cell areas that are included by the estimated ellipses and the total area of the extracted ellipses. More specifically, the total area covered by the estimated ellipses is defined as follows:

$$E = \sum_{p \in M} E_C(p) \tag{7}$$

the foreground area of the binary image M that is included by the estimated ellipses is defined

as follows:

$$A_F = \sum_{p=1} M(p)E(p) \tag{8}$$

the area of the background area of the binary image *M* that is included by the estimated ellipses is defined as follows:

$$A_B = \sum_{p=1} [1 - M(p)]E(p) \tag{9}$$

and the overlapping parts of the ellipses of the touching cells for the total number of the identified ellipses is defined as follows:

$$A_T = \sum_{i=1} \sum_{p=1} E_{C_i}(p) \cap E_{C_j}(p), j = 1, j \neq i \tag{10}$$

Based on the calculation of these metrics, we estimated the fitness degree of the estimated ellipsoidal components against the 2D cell data and we selected the candidate seeds that maximize the following:

$$S = max\left(\frac{A_F - A_B - A_T}{E}\right) \tag{11}$$

The final segmentation of the clustered cells was performed by applying Bayesian classification which assigns each pixel *p* to cluster C_i with the maximum posterior probability. Finally, as lymphocytes typically have small ($7–10\mu m$), round, and dark nuclei with little cytoplasm, which is distinctive from malignant cells or stromal cells [50], we used morphological and textural features, namely size and shape cells, the average value and the skewness of the intensity histogram of the cell, in order to reject the non-lymphocytes.

Post validation of candidate TLS regions

Following the candidate TLS refinement and lymphocytes detection, the falsely detected candidate TLS regions were rejected, in order to validate the identified candidate TLS regions and decrease the false positive TLS identification rates. More specifically, hypothesizing that the lymphocytes density of TLS regions is much higher than lymphocytes within the rest of a tissue, 3 features were extracted and used for the rejection of candidate non-TLS regions. To this end, the number of lymphocytes, the size and the number of lymphocytes per unit area of each candidate TLS region were extracted. After the estimation of the features, an SVM classifier was deployed towards the aim of arriving at a final decision regarding whether an identified candidate TLS region is an actual TLS region or a false-positive candidate case.

Dataset description

Formalin-fixed paraffin-embedded tissue, surplus to diagnostic purposes, was obtained from patients undergoing lung cancer resection. Informed consent was obtained from the donor prior to surgery for use of surgically-excised tissues for research purposes. This study was approved by the local Ethical Committee of the University of East Anglia (Ref No. 2017/2018–119 HT). Histological cases were retrieved from the archive of the Norfolk and Norwich University Hospital histopathology department. Tumours were classified according to the 2015 WHO classification [51]. For each patient, TLS assessment was based on a representative tissue block, with adequate tumour material and interface between normal and tumor tissue well represented. Annotation was performed on tumour tissue slides from 18 patients with primary

lung cancer. There were 8 tumours, of which there were 14 adenocarcinoma (5 acinar predominant, 1, papillary, 2 solid predominant and 6 lepidic predominant), 3 squamous cell carcinomas and 2 sarcomatoid carcinomas. The age range of the patients was 54–82 years old, with an average of 69 ± 2.4 years. The tumour size was between 8 and 62 mm, with an average of 28 ± 6.4. TLS was defined as all dense lymphocytic aggregates and 144 TLS were annotated by 2 histopathologists (FK, MDC) with 100% concordance. To assess the generalizability of the model, two datasets (D1 and D2) were created consisting of 5 and 13 patients respectively for internal training-validation and further validation in independent populations.

Results

For the evaluation of the proposed method we conducted extensive tests using the two datasets. Initially, through the first dataset, we internally validated the efficiency for TLS identification of the proposed methodology by performing an ablation analysis and leave-one-out cross-validation and then through the second dataset we externally validated the generalizability of the proposed model in a different population. Furthermore, we compared the efficiency of identification of TLS regions, using state of the art approaches. Finally, for the identified TLS regions for both datasets, we used box plots to show the range of the number and density of lymphocytes as well as the TLS area.

Identification of TLS regions

The presented TLS identification and density assessment methodology comprised three main components, namely DeepLab model, its active contour approach and post-validation processing. The applied DeepLab v3+ model was used to identify the candidate TLS regions in a semantic image segmentation task, the active contour approach for boundary refinement of the candidate TLS regions, while the post-validation scheme was used to reject the candidate non-TLS regions.

Initially, for the evaluation of the components of the proposed methodology, in the first dataset we performed tissue slide-level leave-one-out analysis. Thus, we found that the DeepLab v3+ model achieves a performance that reaches an Area under the Receiver Operating Characteristic (AUROC) curve of 0.9584. More specifically, defining sensitivity levels, at 95%, 98% and 99%, the DeepLab v3+ model achieves 85.79%, 80.95% and 74.98% specificity rates respectively. In addition to the DeepLab v3+ model, the use of boundary refinement improves the overall performance reaching an AUROC of 0.96. This slight increase in AUROC is translated into higher specificity rates at the predefined sensitivity levels, reaching 86.97% specificity at 95% sensitivity, 80.97% specificity at 98% sensitivity and 74.99% specificity at 99% sensitivity level. The adoption of the post-validation scheme and rejection of the candidate non-TLS regions improves the performance further, reaching an AUROC of 0.9609. Thus, the corresponding specificity and sensitivity rates for the proposed model are 87.02% specificity at 95% sensitivity, 80.97% specificity at 98% sensitivity and 74.99% specificity at 99% sensitivity level (Fig 3).

Furthermore, in order to confirm that the performance of the proposed methodology remains robust, we carried out an external validation analysis using the second dataset. To this end, we used the first dataset as the training dataset, and the proposed methodology reaches an AUROC equal to 0.9589. More specifically, it achieves 92.87% specificity at 95% sensitivity, 88.79% specificity at 98% sensitivity and 84.32% specificity at 99% sensitivity level. The qualitative results of Fig 4 demonstrate that the total number of the components used in this methodology contribute to the overall accuracy of TLS identification. Based on the input 10x H&E-stained images (Fig 4A), the DeepLab v3+ accurately detects the candidate TLS regions
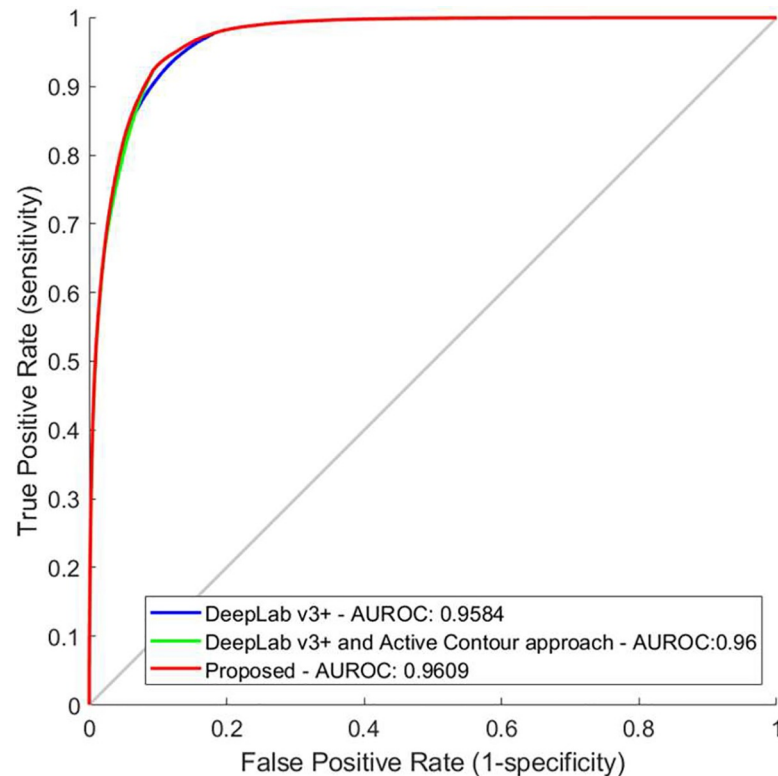
Fig 3. Ablation analysis and performance on tissue slide-based analysis of the a) DeepLab v3+ model, b) DeepLab v3 + and active contour approach and c) proposed methodology. AUROC, area under the receiver operating characteristic curve as shown in the image; FPR, false positive rate; TPR, true positive rate.

https://doi.org/10.1371/journal.pone.0256907.g003

(Fig 4B). Then the active contour approach obtains precise TLS boundaries (Fig 4C) while the post-validation processing step rejects candidate non-TLS regions (Fig 4D) that have partially similar characteristics to TLS regions.

A comparison of state-of-the-art methods

For the validation of the proposed methodology, three state of the art methods were deployed. Deploying a SegNet model [52], a U-Net model [53] and a classic lymphocytes-based thresholding method [54] the proposed method outperforms these (Fig 5) identifying more accurate the TLS regions (Fig 6).

More specifically, the SegNet is a CNN architecture for semantic segmentation proposed by researchers at University of Cambridge [52]. The main SegNet applications regarding segmentation tasks such as semantic segmentation of prostate cancer [55], gland segmentation from colon cancer histology images [56] and brain tumor segmentation from multi-modal magnetic resonance images [57]. For the evaluation of the SegNet in the task of TLS identification, we used the same protocol as the proposed methodology and we found that its performance in the external validation set reaches an AUROC of 0.9307. This is translated to 80.87% specificity at 95% sensitivity, 72.14% specificity at 98% sensitivity and 70.12% specificity at 99% sensitivity level.

Then we chose U-Net, a popular deep-learning-based method to solve microscopy image segmentation problems [53]. Deploying the U-Net we found that its performance reaches an AUROC of 0.7489 and 15.2% specificity at 95% sensitivity level. It is worth mentioning that in
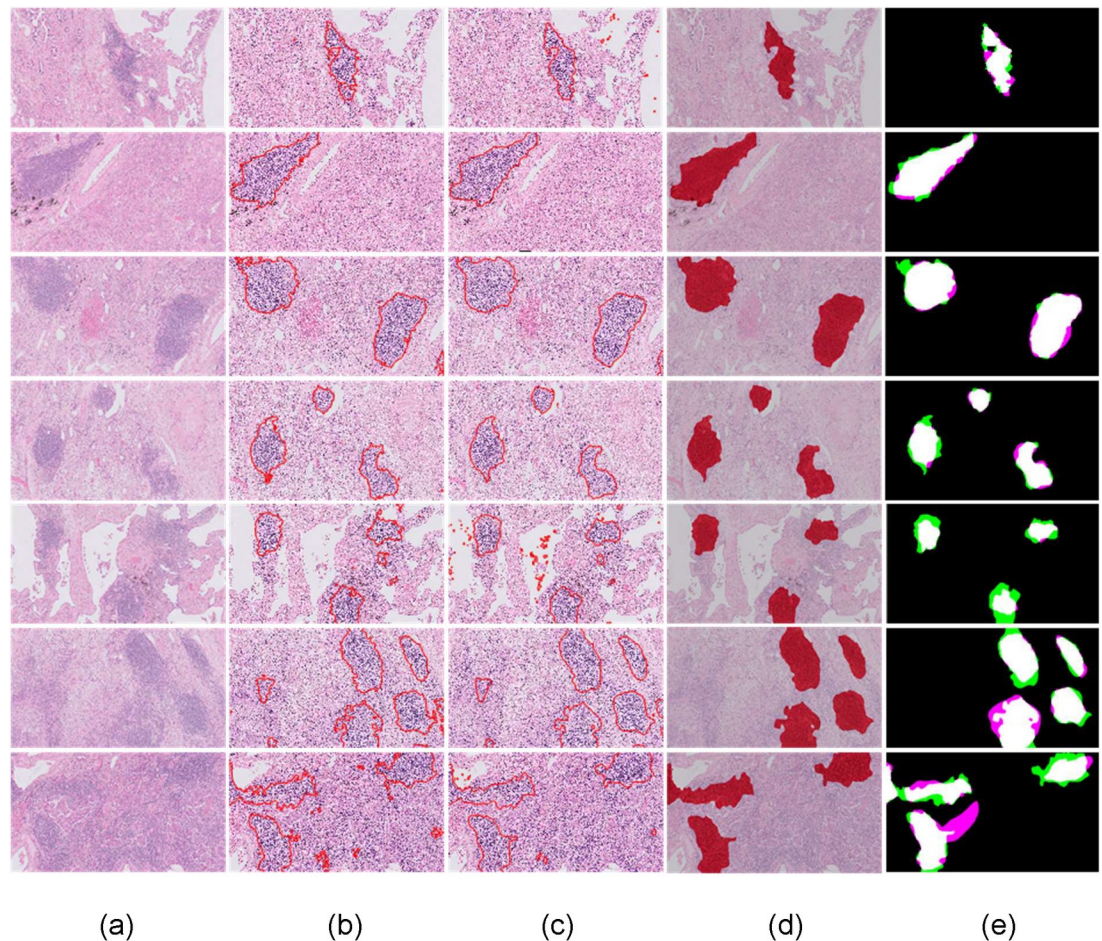
(a)   (b)   (c)   (d)   (e)

Fig 4. Identification of TLS regions and density of lymphocytes. a) input 10x H&E-stained image, b) DeepLabv3+ identified TLS regions, c) application of active contour model d) post validation of candidate TLS regions—detected TLS regions e) identified TLS regions (classification threshold at 0.5) in comparison with ground truth: White color (True Positive), green color (False Positive), magenta color (False Negative).

contrast to the proposed network the U-Net was not pre-trained. This might explain the latter's insufficient capture of the experimental variation. Finally, we applied a thresholding method [54] aiming to distinguish regions with contrasting density lymphocytes levels. The thresholding method reaches an AUROC of 0.9021 and 57.75% specificity at 95% sensitivity, 38.71% specificity at 98% sensitivity and 22.41% specificity at 99% sensitivity level.

Criteria for the definition of TLS regions

In this work, for the standardisation of TLS recognition, we aim to translate the visual recognition of TLS by histopathologists into a universally reproducible set of mathematical values (Fig 7). Thus, through the lymphocytes segmentation procedure, we show that the TLS regions include a minimum number of 45 lymphocytes and that the minimum area of TLS regions is $6,245 \mu m^2$. The mean number of lymphocytes in TLS regions is 620.8 while the mean area of TLS regions is $48,387 \mu m^2$. The largest number of lymphocytes in TLS regions is 2936 and the largest area of TLS regions is $230,604 \mu m^2$. The minimum number of lymphocytes per area of TLS regions is $0.0074/\mu m^2$ with a mean and standard deviation of $0.0128/\mu m^2$ and $0.0026/\mu m^2$ respectively and a maximum of $0.0189/\mu m^2$.
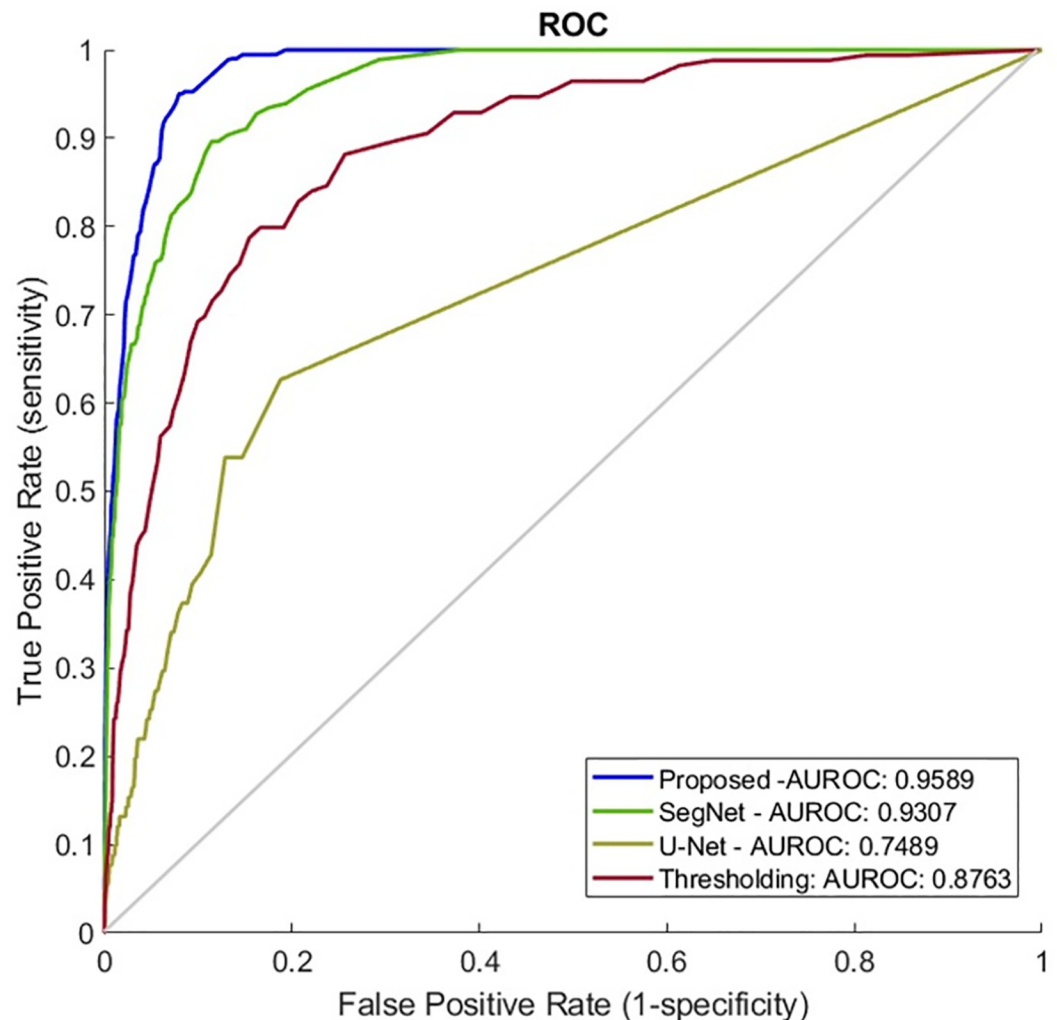
Fig 5. Performance comparison on external validation analysis. AUROC, area under the receiver operating curve as shown in the image; FPR, false positive rate; TPR, true positive rate.

It is worth mentioning that the aforementioned criteria were defined through the analysis of TLS regions in both datasets created. In additional analyses, we carried out tests for the comparison between the lymphocyte's density values of identified TLS and that outside of TLS regions. Thus, in the comparison of the lymphocytes density values of identified TLS regions with the values outside of the TLS regions we identified that the minimum number lymphocytes of TLS regions is $0.0019/\mu m^2$, with a mean and standard deviation of $0.0040/\mu m^2$ and $0.0010/\mu m^2$ respectively and a maximum of $0.0063/\mu m^2$ (Fig 8). Finally, it is worth mentioning that for the evaluation of the proposed lymphocytes segmentation method, we manually annotated four hundred cells. Thus, the method was compared with an ellipsoidal model [14] and the classical watershed algorithm outperforming these with a correctly segmented rate of 91.5% in contrast to 89.5% and 86% respectively. Thus, heat maps of lymphocytes were constructed from the input 10x H&E-stained images and their corresponding detected TLS regions (Fig 9). In the heat maps, the dark blue color represents the background and the other cells, while the lighter blue to red colors represent the lymphocytes from lower to higher
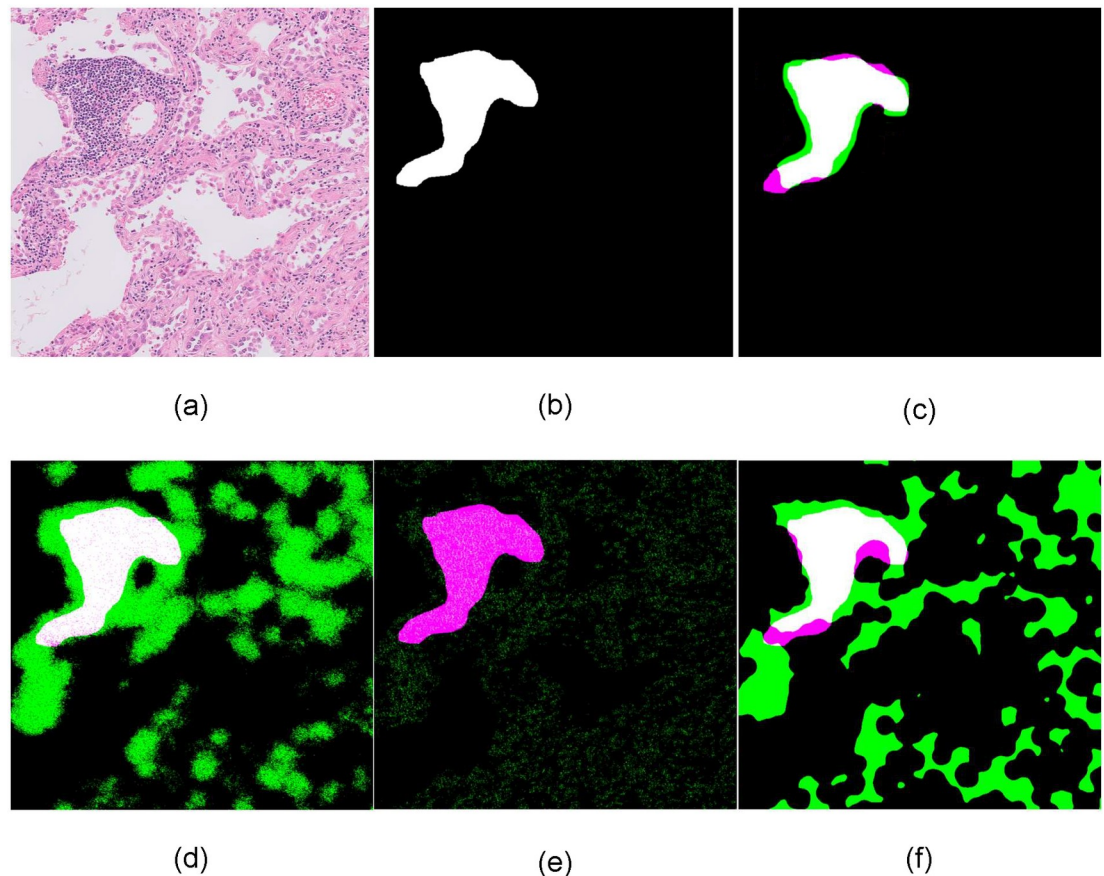
Fig 6. Qualitative comparison on tissue slide-based analysis. a) input 10x H&E-stained image, b) ground truth c) DeepLabv3 + identified TLS regions, d) SegNet identified TLS regions, e) U-Net identified TLS regions, f) lymphocytes density thresholding identified TLS regions, in comparison with ground truth: White color (True Positive), green color (False Positive), magenta color (False Negative).

https://doi.org/10.1371/journal.pone.0256907.g006

density. Thus, it is clearly shown that TLS regions were easily recognized within the lung cancer tissue from the lymphocytes heatmap.

Conclusion

Histologically, the TLS are recognised as mature and tightly packed lymphoid cells forming discrete entities with a smooth circular outline. Previous studies have determined the density of TLS based on the expression of high endothelial venules, DC Lamp expressing cells and other markers. However, based on our diagnostic experience, this is likely to represent an underestimate of the density of TLS. The TLS, detected by immunohistochemistry, represent only a subset of all the TLS which can be seen histologically on an H&E slide. In observations, Omer and Ng kee kwong have shown that there are TLS which do not show expression of CD21 and CD23 expressing cells, which are used as markers of TLS. Therefore, TLS, demonstrated with specific immunostain, represents an underestimate of the total number of TLS. In the study by Silina *et al* [11], the total TLS number can only be identified on histological interpretation on the H&E slides. Although histopathologists have recognized the presence of TLS as Crohn's-like lymphoid aggregates adjacent to colorectal adenocarcinomas since at least 1990 [10], current methods of detecting and quantifying TLS vary in the literature. Other
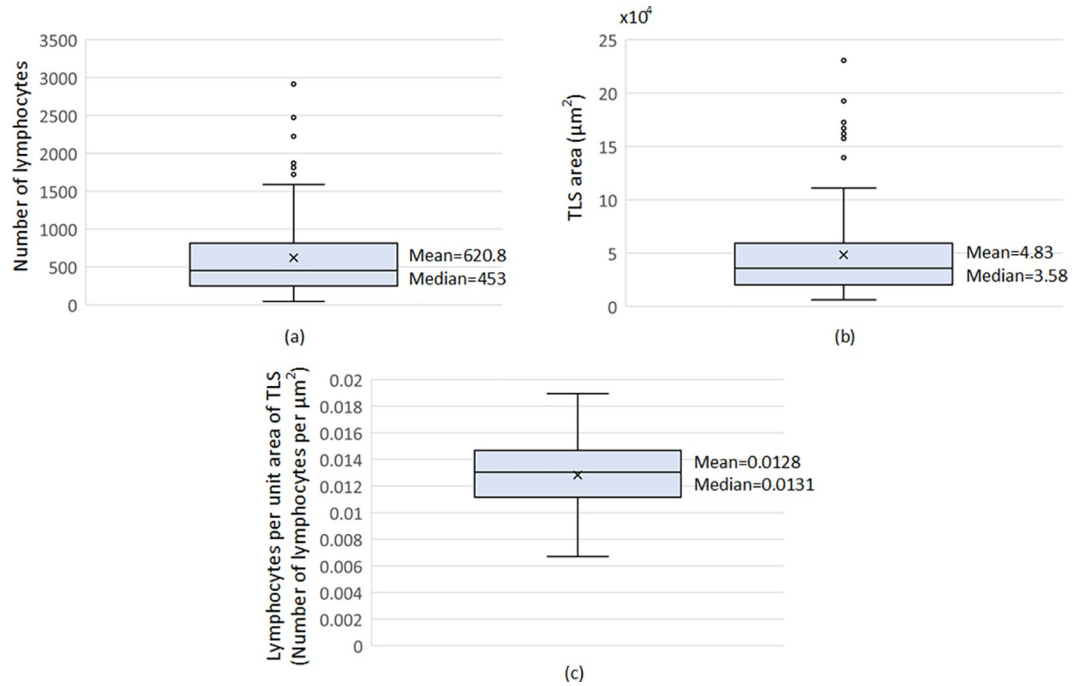
Fig 7. Box plots for (a) the number of lymphocytes, (b) size and (c) density (number of lymphocytes per unit area of the TLS regions) for the hematoxylin and eosin (H&E) stained histological slides included in both of D1 & D2.

factors, such as the TLS diameter has been shown to affect the prognosis of colorectal cancer [58]. Therefore, the counting of TLS over a large histological area should be facilitated such that other factors could be integrated in the prognostic determination of the TLS.

We first develop a machine learning automated approach to mimic the identification of the TLS by histopathologists. This had a 92.87% specificity at 95% sensitivity, 88.79% specificity at
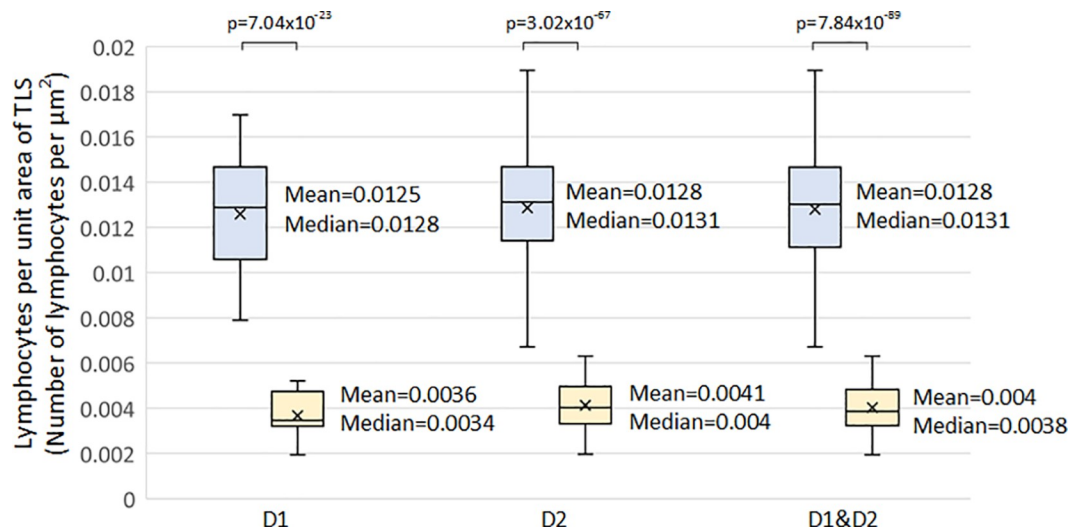

Fig 8. Box plots comparing lymphocyte density between the identified TLS regions and non-TLS regions in hematoxylin and eosin (H&E) stained histological slides included in dataset 1 (D1), in dataset 2 (D2) and in both of them (D1&D2).
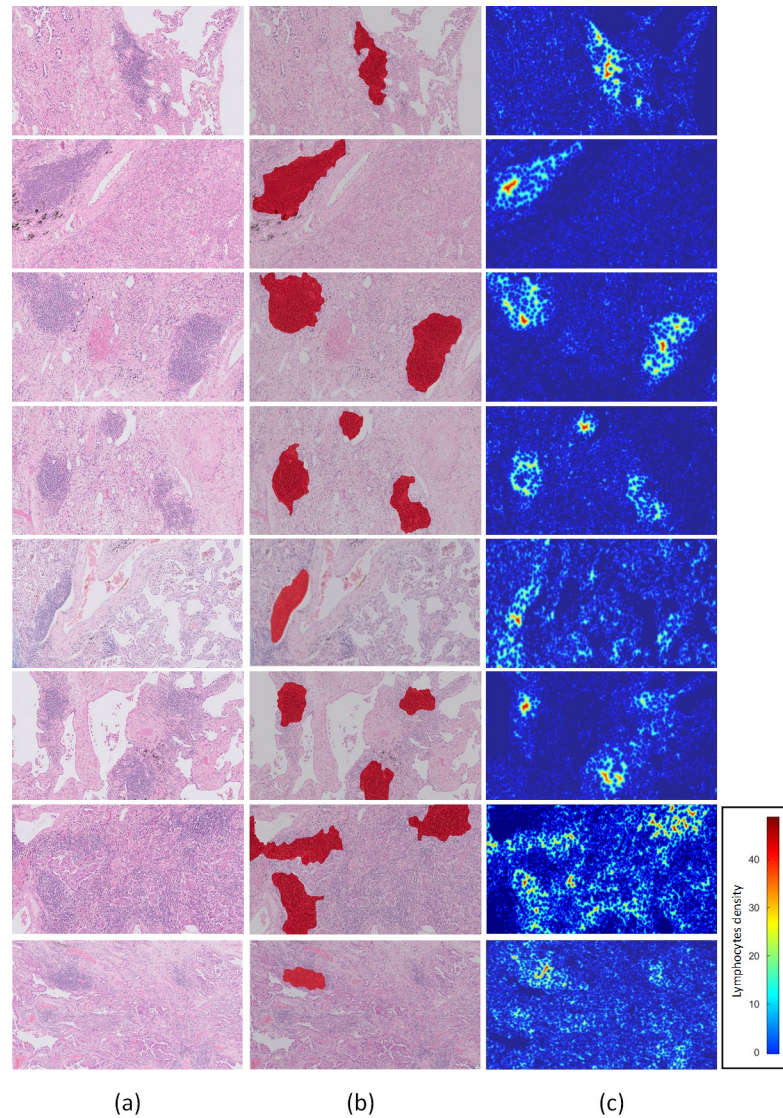
Fig 9. Identification of TLS regions and density of lymphocytes. a) input 10x H&E-stained image, b) detected TLS regions c) heat maps of lymphoyctes.

98% sensitivity and 84.32% specificity at 99% sensitivity level, implying that the automated approach was able to reproduce histopathologists assessment with great accuracy. This was compared to existing methods of vision recognition and our methodology was superior to those. After the lymphocytes have been segmented, we then used this approach to try to determine the mathematical criteria defining a TLS. Firstly, it appears apparent that there should be a minimum number of lymphocytes within the TLS. In this limited series of TLS, we have shown that the minimum number of lymphocytes is 45. It is also apparent that the TLS has a minimum area and we have shown this to be $6,245 \mu m^2$. Visually, we recognise that the lymphocytes are more tightly packed in TLS compared to those in the same tissue, but outside of the TLS. We have here shown that the density of lymphocytes is approximately 3 times those outside of the TLS. The mean density of lymphocytes within a TLS area is $0.0128 \mu m^2$ which is much higher compared to $0.004 \mu m^2$ in non-TLS regions.

We therefore argue that future studies should use the following criteria for definition of TLS: lymphocyte number within the TLS, the minimum area of the TLS, and density of lymphocytes within a given area. As the density of the lymphocytes within the TLS is much higher than those outside the TLS, the algorithm has demonstrated a distinction between TLS and the tumour infiltrating lymphocytes. The latter are especially relevant to the prognostic factor of breast cancer and therefore, this current trained deep learning model cannot be used to detect regions of tumour infiltrating lymphocytes. A limitation of this current study is that we were not able to correlate TLS density with the patient outcome, owing to the low number of patients in each dataset. Therefore, future studies should correlate the above 3 criteria with outcome of lung cancer patients in large datasets. Based on these findings, it would be possible to adjust the particular values used to define the TLS, to give an optimum definition based on clinical outcome data. Standardisation of TLS density assessment by machine learning will allow the comparison of the host immune response between different subtypes of the same tumour, and between the same tumour across different studies. It will also allow for better correlation between TLS density and outcome.

Another limitation is that this study is based on lung cancer histological images. Future studies are needed to prove that this methodology will be validated in other cancer types and in material from other centers. However, in our experience, the TLS are histologically identical between the different cancer types and we expect this algorithm to be validated in other cancer subtypes too. In our histological experience, lung biopsies are quite narrow and therefore only show part of the TLS, if they were present. Therefore, the smooth outline of the TLS would not be seen in the biopsy and therefore, the algorithm is unlikely to be effective. The model would need to be adapted to detect TLS in biopsies. This can be the subject of another study.

This algorithm provides for the detection of all TLS, but does not make a distinction between the different types of TLS previously described by Silina *et al.* [9–11]. However, we have made observations whereby TLS with germinal centres (secondary follicle-like TLS) constitute 53% of all TLS in a series of lung cancer cases (Omer and Ng kee kwong). We believe that the number of these TLS subtypes would be dependent on the total number of TLS. In this study, we used morphology to detect TLS. A future step would be to use machine learning to quantify TLS of different subtypes.

Our data shows that the performance of the proposed methodology for TLS identification remains high even when the methodology is trained with a limited number of data. The estimated accuracy rates mean that there is a possibility that this AI method can be used in larger scale studies and possibly future clinical diagnostic studies. Although the tumour inflammatory response is assessed on H&E slides using tumour infiltrating lymphocytes, the assessment of TLS is not used in diagnostic practice as there is no precise or standard definition [1]. We therefore believe that our methodology can be used in future studies, including those of non-pulmonary tumour sites. Furthermore, our study provides the foundations for the standardisation of the definition and quantification of the TLS on standard H&E histological images. In addition, we have shown that machine learning can provide a fast and reliable method of quantification. This would possibly lead to its widespread adoption in routine histopathological practice.

Acknowledgments

Author Contributions

References

1. Sun L, Zhang L, Yu J, Zhang Y, Pang X, Ma C, et al. Clinical efficacy and safety of anti-PD-1/PD-L1 inhibitors for the treatment of advanced or metastatic cancer: a systematic review and meta-analysis. Scientific reports. 2020 Feb 7; 10(1):1–3. https://doi.org/10.1038/s41598-019-56847-4 PMID: 31913322

2. Sautès-Fridman C, Lawand M, Giraldo NA, Kaplon H, Germain C, Fridman WH, et al. Tertiary lymphoid structures in cancers: prognostic value, regulation, and manipulation for therapeutic intervention. Frontiers in immunology. 2016 Oct 3; 7:407. https://doi.org/10.3389/fimmu.2016.00407 PMID: 27752258

3. Munoz-Erazo L, Rhodes JL, Marion VC, Kemp RA. Tertiary lymphoid structures in cancer–considerations for patient prognosis. Cellular & molecular immunology. 2020 Jun; 17(6):570–5.

4. Dieu-Nosjean MC, Goc J, Giraldo NA, Sautès-Fridman C, Fridman WH. Tertiary lymphoid structures in cancer and beyond. Trends in immunology. 2014 Nov 1; 35(11):571–80. https://doi.org/10.1016/j.it.2014.09.006 PMID: 25443495

5. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. Nature Reviews Cancer. 2019 Jun; 19(6):307–25. https://doi.org/10.1038/s41568-019-0144-6 PMID: 31092904

6. Dieu-Nosjean MC, Giraldo NA, Kaplon H, Germain C, Fridman WH, Sautès-Fridman C. Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. Immunological reviews. 2016 May; 271(1):260–75. https://doi.org/10.1111/imr.12405 PMID: 27088920

7. Cabrita R, Lauss M, Sanna A, Donia M, Larsen MS, Mitra S, et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. Nature. 2020 Jan; 577(7791):561–5. https://doi.org/10.1038/s41586-019-1914-8 PMID: 31942071

8. Posch F, Silina K, Leibl S, Mündlein A, Moch H, Siebenhüner A, et al. Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer. Oncoimmunology. 2018 Feb 1; 7(2): e1378844. https://doi.org/10.1080/2162402X.2017.1378844 PMID: 29416939

9. Silina K, Soltermann A, Attar FM, Casanova R, Uckeley ZM, Thut H, et al. Germinal centers determine the prognostic relevance of tertiary lymphoid structures and are impaired by corticosteroids in lung squamous cell carcinoma. Cancer research. 2018 Mar 1; 78(5):1308–20. https://doi.org/10.1158/0008-5472.CAN-17-1987 PMID: 29279354

10. Graham DM, Appelman HD. Crohn's-like lymphoid reaction and colorectal carcinoma: a potential histologic prognosticator. Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc. 1990 May 1; 3(3):332–5. PMID: 2362940

11. Siliņa K, Burkhardt C, Casanova R, Solterman A, van den Broek M. A quantitative pathology approach to analyze the development of human cancer-associated tertiary lymphoid structures. In Tertiary Lymphoid Structures 2018 (pp. 71–86). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-8709-2_5 PMID: 30141008

12. Win KY, Choomchuay S. Automated segmentation of cell nuclei in cytology pleural fluid images using OTSU thresholding. In2017 International Conference on Digital Arts, Media and Technology (ICDAMT) 2017 Mar 1 (pp. 14–18). IEEE.

13. Pan L, Gu L, Xu J. Implementation of medical image segmentation in CUDA. In2008 International Conference on Information Technology and Applications in Biomedicine 2008 May 30 (pp. 82–85). IEEE.

14. Dimitropoulos K., Barmpoutis P., Koletsa T., Kostopoulos I. and Grammalidis N., 2017. Automated detection and classification of nuclei in pax5 and H&E-stained tissue sections of follicular lymphoma. Signal, Image and Video Processing, 11(1), pp.145–153.

15. Verma VS. New morphological technique for medical image segmentation. In2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT) 2017 Feb 9 (pp. 1–5). IEEE.

16. Zhanpeng HU, Qi Z, Shizhong J, Guohua C. Medical image segmentation based on the watersheds and regions merging. In2016 3rd International Conference on Information Science and Control Engineering (ICISCE) 2016 Jul 8 (pp. 1011–1014). IEEE.

17. Liang Y, Fu J. Watershed algorithm for medical image segmentation based on morphology and total variation model. International Journal of Pattern Recognition and Artificial Intelligence. 2019 May 31; 33 (05):1954019.

18. Chen X, Williams BM, Vallabhaneni SR, Czanner G, Williams R, Zheng Y. Learning active contour models for medical image segmentation. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2019 (pp. 11632–11640).

19. Pratondo A, Chui CK, Ong SH. Integrating machine learning with region-based active contour models in medical image segmentation. Journal of Visual Communication and Image Representation. 2017 Feb 1; 43:1–9.

20. Lee HG, Lee SC. Nucleus segmentation using Gaussian mixture based shape models. IEEE journal of biomedical and health informatics. 2017 May 2; 22(1):235–43. https://doi.org/10.1109/JBHI.2017.2700518 PMID: 28475068

21. Na L, Yan J, Shu L. Application of PSO algorithm with dynamic inertia weight in medical image thresholding segmentation. In2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom) 2017 Oct 12 (pp. 1–4). IEEE.

22. Kucharski D, Kleczek P, Jaworek-Korjakowska J, Dyduch G, Gorgon M. Semi-supervised nests of melanocytes segmentation method using convolutional autoencoders. Sensors. 2020 Jan; 20(6):1546. https://doi.org/10.3390/s20061546 PMID: 32168748

23. Su H, Xing F, Kong X, Xie Y, Zhang S, Yang L. Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders. InInternational Conference on Medical Image Computing and Computer-Assisted Intervention 2015 Oct 5 (pp. 383–390). Springer, Cham.

24. Dabass M, Vig R, Vashisth S. Review of histopathological image segmentation via current deep learning approaches. In2018 4th International Conference on Computing Communication and Automation (ICCCA) 2018 Dec 14 (pp. 1–6). IEEE.

25. Krithiga R, Geetha P. Breast cancer detection, segmentation and classification on histopathology images analysis: a systematic review. Archives of Computational Methods in Engineering. 2020 Aug 10:1–3.

26. Sun S, Yuan H, Zheng Y, Zhang H, Jiang Z. Cancer Sensitive Cascaded Networks (CSC-Net) for Efficient Histopathology Whole Slide Image Segmentation. In2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) 2020 Apr 3 (pp. 476–480). IEEE.

27. BenTaieb A, Hamarneh G. Topology aware fully convolutional networks for histology gland segmentation. InInternational Conference on Medical Image Computing and Computer-Assisted Intervention 2016 Oct 17 (pp. 460–468). Springer, Cham.

28. Dimitropoulos K, Barmpoutis P, Zioga C, Kamas A, Patsiaoura K, Grammalidis N. Grading of invasive breast carcinoma through Grassmannian VLAD encoding. PloS one. 2017 Sep 21; 12(9):e0185110. https://doi.org/10.1371/journal.pone.0185110 PMID: 28934283

29. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. InProceedings of the European conference on computer vision (ECCV) 2018 (pp. 801–818).

30. Xiao WT, Chang LJ, Liu WM. Semantic segmentation of colorectal polyps with DeepLab and LSTM networks. In2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) 2018 May 19 (pp. 1–2). IEEE.

31. Wan T, Xu S, Sang C, Jin Y, Qin Z. Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. Neurocomputing. 2019 Nov 6; 365:157–70.

32. Altini N, Cascarano GD, Brunetti A, Marino F, Rocchetti MT, Matino S, et al. Semantic segmentation framework for glomeruli detection and classification in kidney histological sections. Electronics. 2020 Mar; 9(3):503.

33. Jiang YQ, Xiong JH, Li HY, Yang XH, Yu WT, Gao M, et al. Recognizing basal cell carcinoma on smartphone-captured digital histopathology images with a deep neural network. British Journal of Dermatology. 2020 Mar; 182(3):754–62. https://doi.org/10.1111/bjd.18026 PMID: 31017653

34. Tang W, Zou D, Yang S, Shi J, Dan J, Song G. A two-stage approach for automatic liver segmentation with Faster R-CNN and DeepLab. Neural Computing and Applications. 2020 Jan 16:1–0.

35. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception-resnet and the impact of residual connections on learning. InProceedings of the AAAI Conference on Artificial Intelligence 2017 Feb 12 (Vol. 31, No. 1).

36. Donahue J, Jia Y, Vinyals O, Hoffman J, Zhang N, Tzeng E, et al. Decaf: A deep convolutional activation feature for generic visual recognition. InInternational conference on machine learning 2014 Jan 27 (pp. 647–655). PMLR.

37. Pan SJ, Yang Q. A survey on transfer learning. IEEE Transactions on knowledge and data engineering. 2009 Oct 16; 22(10):1345–59.

38. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012; 25:1097–105.

39. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. InInternational conference on medical image computing and computer-assisted intervention 2016 Oct 17 (pp. 424–432). Springer, Cham.

40. Wang L, He L, Mishra A, Li C. Active contours driven by local Gaussian distribution fitting energy. Signal Processing. 2009 Dec 1; 89(12):2435–47.

41. Ji Z, Xia Y, Sun Q, Cao G, Chen Q. Active contours driven by local likelihood image fitting energy for image segmentation. Information Sciences. 2015 Apr 20; 301:285–304.

42. Tareef A, Song Y, Huang H, Feng D, Chen M, Wang Y, et al. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. IEEE transactions on medical imaging. 2018 Mar 12; 37 (9):2044–59. https://doi.org/10.1109/TMI.2018.2815013 PMID: 29993863

43. Molnar C, Jermyn IH, Kato Z, Rahkama V, Östling P, Mikkonen P, et al. Accurate morphology preserving segmentation of overlapping cells based on active contours. Scientific reports. 2016 Aug 26; 6(1):1–0. https://doi.org/10.1038/s41598-016-0001-8 PMID: 28442746

44. Liao M, Zhao YQ, Li XH, Dai PS, Xu XW, Zhang JK, et al. Automatic segmentation for cell images based on bottleneck detection and ellipse fitting. Neurocomputing. 2016 Jan 15; 173:615–22.

45. Barbier M, Jaensch S, Cornelissen F, Vidic S, Gjerde K, de Hoogt R, et al. Ellipsoid segmentation model for analyzing light-attenuated 3D confocal image stacks of fluorescent multi-cellular spheroids. PloS one. 2016 Jun 15; 11(6):e0156942. https://doi.org/10.1371/journal.pone.0156942 PMID: 27303813

46. Panagiotakis C, Argyros A. Region-based Fitting of Overlapping Ellipses and its application to cells segmentation. Image and Vision Computing. 2020 Jan 1; 93:103810.

47. Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nature methods. 2019 Jan; 16(1):67–70. https://doi.org/10.1038/s41592-018-0261-2 PMID: 30559429

48. Al-Kofahi Y, Zaltsman A, Graves R, Marshall W, Rusu M. A deep learning-based algorithm for 2-D cell segmentation in microscopy images. BMC bioinformatics. 2018 Dec; 19(1):1–1. https://doi.org/10.1186/s12859-017-2006-0 PMID: 29291722

49. Moshkov N, Mathe B, Kertesz-Farkas A, Hollandi R, Horvath P. Test-time augmentation for deep learning-based cell segmentation on microscopy images. Scientific reports. 2020 Mar 19; 10(1):1–7. https://doi.org/10.1038/s41598-019-56847-4 PMID: 31913322

50. Lu Z, Xu S, Shao W, Wu Y, Zhang J, Han Z, et al. Deep-learning–based characterization of tumor-infiltrating lymphocytes in breast cancers from histopathology images and multiomics data. JCO clinical cancer informatics. 2020 May; 4:480–90. https://doi.org/10.1200/CCI.19.00126 PMID: 32453636

51. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JH, Beasley MB, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. Journal of thoracic oncology. 2015 Sep 1; 10(9):1243–60. https://doi.org/10.1097/JTO.0000000000000630 PMID: 26291008

52. Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence. 2017 Jan 2; 39(12):2481–95. https://doi.org/10.1109/TPAMI.2016.2644615 PMID: 28060704

53. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. InInternational Conference on Medical image computing and computer-assisted intervention 2015 Oct 5 (pp. 234–241). Springer, Cham.

54. Withey DJ, Koles ZJ. Medical image segmentation: Methods and software. In2007 Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging 2007 Oct 12 (pp. 140–143). IEEE.

55. Ing N, Ma Z, Li J, Salemi H, Arnold C, Knudsen BS, et al. Semantic segmentation for prostate cancer grading by convolutional neural networks. InMedical Imaging 2018: Digital Pathology 2018 Mar 6 (Vol. 10581, p. 105811B). International Society for Optics and Photonics.

56. Tang J, Li J, Xu X. Segnet-based gland segmentation from colon cancer histology images. In2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC) 2018 May 18 (pp. 1078–1082). IEEE.

57. Alqazzaz S, Sun X, Yang X, Nokes L. Automated brain tumor segmentation on multi-modal MR image using SegNet. Computational Visual Media. 2019 Jun; 5(2):209–19.

58. Ueno H, Mochizuki H, Hashiguchi Y, Ishiguro M, Miyoshi M, Kajiwara Y, et al. Extramural cancer deposits without nodal structure in colorectal cancer: optimal categorization for prognostic staging. American journal of clinical pathology. 2007 Feb 1; 127(2):287–94. https://doi.org/10.1309/903UT10VQ3LC7B8L PMID: 17210518