

RESEARCH

Open Access



# A graph-learning based model for automatic diagnosis of Sjögren's syndrome on digital pathological images: a multicentre cohort study

Ruifan Wu<sup>1†</sup>, Zhipei Chen<sup>2,3†</sup>, Jiali Yu<sup>2,3†</sup>, Peng Lai<sup>1</sup>, Xuanyi Chen<sup>2,3</sup>, Anjia Han<sup>4</sup>, Meng Xu<sup>2,3</sup>, Zhaona Fan<sup>2,3</sup>, Bin Cheng<sup>2,3</sup>, Ying Jiang<sup>1\*</sup> and Juan Xia<sup>2,3\*</sup>

## Abstract

**Background** Sjögren's Syndrome (SS) is a rare chronic autoimmune disorder primarily affecting adult females, characterized by chronic inflammation and salivary and lacrimal gland dysfunction. It is often associated with systemic lupus erythematosus, rheumatoid arthritis and kidney disease, which can lead to increased mortality. Early diagnosis is critical, but traditional methods for diagnosing SS, mainly through histopathological evaluation of salivary gland tissue, have limitations.

**Methods** The study used 100 labial gland biopsy, creating whole-slide images (WSIs) for analysis. The proposed model, named Cell-tissue-graph-based pathological image analysis model (CTG-PAM) and based on graph theory, characterizes single-cell feature, cell-cell feature, and cell-tissue feature. Building upon these features, CTG-PAM achieves cellular-level classification, enabling lymphocyte recognition. Furthermore, it leverages connected component analysis techniques in the cell graph structure to perform SS diagnosis based on lymphocyte counts.

**Findings** CTG-PAM outperforms traditional deep learning methods in diagnosing SS. Its area under the receiver operating characteristic curve (AUC) is 1.0 for the internal validation dataset and 0.8035 for the external test dataset. This indicates high accuracy. The sensitivity of CTG-PAM for the external dataset is 98.21%, while the accuracy is 93.75%. In comparison, the sensitivity and accuracy for traditional deep learning methods (ResNet-50) are lower. The study also shows that CTG-PAM's diagnostic accuracy is closer to skilled pathologists compared to beginners.

<sup>†</sup>Ruifan Wu, Zhipei Chen and Jiali Yu contributed equally to this work.

<sup>†</sup>Zhipei Chen contributed equally as co-first author.

<sup>†</sup>Jiali Yu contributed as co-senior author.

\*Correspondence:

Ying Jiang  
jiangy32@mail.sysu.edu.cn

Juan Xia  
xiajuan@mail.sysu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Interpretation** Our findings indicate that CTG-PAM is a reliable method for diagnosing SS. Additionally, CTG-PAM shows promise in enhancing the prognosis of SS patients and holds significant potential for the differential diagnosis of both non-neoplastic and neoplastic diseases. The AI model potentially extends its application to diagnosing immune cells in tumor microenvironments.

**Keywords** Artificial intelligence, Graph learning, Sjögren's syndrome, Digital pathology, Single-cell feature

## Introduction

Sjögren's Syndrome (SS) is a rare chronic multisystem autoimmune disorder, affecting only 8% of the global population [1–3]. It primarily affects adult females and is characterized by persistent inflammatory features [4]. The hallmark feature of SS is impaired functionality of salivary and lacrimal glands, resulting in significant dryness of the oral mucosa [5]. The clinical manifestations of SS encompass oral and ocular dryness, resemble those observed in other conditions characterized by dry oral symptoms, including Sjögren's syndrome-like disease. These manifestations pose a diagnostic challenge due to their similarity and necessitate lip gland biopsy for accurate identification. A strong correlation exists between the onset of SS and systemic lupus erythematosus (15–36%) and rheumatoid arthritis (20–32%) [6]. SS may potentially involve systemic organ damage and the development of B-cell lymphomas, affecting the respiratory, digestive, urinary, and reproductive system [7–11]. For instance, it can lead to kidney damage and in severe cases progress to kidney failure and mortality. In comparison to isolated rheumatoid arthritis, the presence of Sjögren's Syndrome intensifies joint inflammation and joint damage, leading to heightened extra-articular symptoms such as pain, fatigue, anemia, and pulmonary function alterations. Additionally, it escalates the mortality rate among rheumatoid arthritis patients [12]. Therefore, the early diagnosis of SS holds paramount significance in the treatment of patients with rheumatoid arthritis or systemic lupus erythematosus.

Presently, SS is assessed through four distinct diagnostic criteria: the European Classification (PEC) criteria [13], the American European Consensus Group (AECG 2002) classification criteria [14], the American College of Rheumatology (ACR 2012) criteria [15], and the 2016 criteria jointly established by the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) [16]. Among these four sets of criteria, the histopathological evaluation of salivary gland tissue plays an irreplaceable role in the diagnosis of SS, particularly in the assessment of lymphocyte foci. However, lymphocyte counts in SS necessitate pathologists to engage in subjective assessment and manual enumeration within each high-power field of whole-slide imaging (WSI). Nonetheless, SS poses practical challenges due to its time-consuming nature, subjective judgment,

potential counting inaccuracies, difficulty in diagnosing, and susceptibility to diagnostic errors.

Artificial Intelligence (AI) has witnessed substantial application in pathology image diagnosis [17, 18]. Deep learning facilitates computers in learning and synthesizing tasks akin to human comprehension. Multiple research reports have attested to the high performance and precision of deep learning algorithms in pathology image diagnostics, with diagnostic outcomes paralleling or surpassing human performance, encompassing ailments afflicting organs, for example, the kidneys, liver, brain, and lungs [19–22]. The majority of the work is accomplished using convolutional neural networks, such as ResNet-50 [23]. However, due to the large scale of pathological images, pixel-level analysis using convolutional neural networks (CNNs) poses a computational challenge and training difficulties.

In recent years, graph-based approaches for pathological imaging have gained increasing attention due to their interpretability and ability to effectively extract valuable information from images. The most common method for constructing a graph structure is to consider cells as nodes in the graph [24–27]. In addition to cell graphs, another common approach is to construct graph models using superpixels extracted from pathological images to utilize tissues information [28, 29]. However, current graph-based methods typically perform image classification on the entire or local images for diagnostic purposes, making it challenging to achieve fine-grained analysis. Moreover, these methods often struggle to effectively leverage the valuable information from both individual cells and the surrounding tissue. To this end, we propose a cell-centric, tissue-assisted graph-based pathological image cell analysis model called Cell-tissue-graph-based pathological analysis model (CTG-PAM). CTG-PAM focuses on cells as the primary entities for feature extraction and classification, while incorporating the contextual information of tissues to complement cell features.

This study propounds a model based on the 2016 ACR-EULAR diagnostic standards for SS. Leveraging graph-based learning and WSI, this model enables expeditious and accurate scoring and diagnosis of SS, thereby furnishing clinical and pathologists with objective, precise, reproducible, and standardized diagnostic capabilities.

## Materials and methods

The baseline characteristics of the patients in the two datasets were summarized in Table S1 in supporting information.

### Training and internal validation data sets

Thirty-six labial gland biopsy patients were obtained from the Department of Pathology, Guanghua School of Stomatology, Hospital of Stomatology, Sun Yat-sen University (SYSU) in China from January 2009 to June 2023 and were selected chronologically. Patients with hematoxylin-eosin staining (H&E) biopsy slides of low quality, including slides with drying effects, significant air bubbles, or broken glass, were excluded. An automatic digital slide scanner (NanoZoomer S210, HAMAMATSU, Japan) at  $\times 40$  magnification was used to create WSIs. Two out of the 36 WSIs, one of which is diagnosed as SS and the other as non-SS, were employed to train the network for lymphocyte classification. The remaining 34 WSIs were utilized for validation. Four experienced academic pathologists graded and quantified the biopsy slides. Pathologist 1 was involved with annotating portions of lymphocytes and non-lymphocytes on the slides, where the two slides in the training set for model training and the remaining slides in internal validation data set for testing in the experiments. We categorized the doctors mentioned above into two groups based on their experience in pathology: the junior group (less than three years of pathology experience) and the senior group (more than three years of pathology experience). When the pathologist disagrees with the diagnosis, they will discuss and agree on their opinion. Finally, we compared their accuracy, sensitivity, and time against the automatic method.

### External test data set

Sixty-four labial gland biopsy patients with SS were randomly selected from the First Affiliated Hospital of SYSU from January 2012 to January 2023. Independent preparations by each hospital were used for hematoxylin and eosin staining and WSI scanning. In total, 64 WSIs were obtained by using the method already mentioned. Pathologist 1 was involved with annotating portions of lymphocytes and non-lymphocytes on the slides in the external data set for testing in the experiments.

### Immunohistochemical (IHC) staining

Following rehydration of paraffin-embedded tissue sections, antigen retrieval was achieved by exposing the sections to high heat and pressure in an EDTA buffer (pH 9.0) for 10 min. After blocking endogenous peroxidase activity and applying tissue blocking agents, the sections were incubated with primary antibodies against CD45 (1:100, clone: PD7/26+2B11, MXB, Fuzhou, China) at

room temperature for 4 h. Biotin-conjugated secondary reagents were then applied for 1 h. The immunohistochemistry procedure was completed with DAB (3,3'-diaminobenzidine) staining and hematoxylin counterstaining.

### Cell-tissue-graph-based pathological analysis model

We proposed a graph-based method to automated diagnosis for SS, referred to as Cell-tissue-graph-based pathological analysis model (CTG-PAM). Our method is primarily divided into five stages: preprocessing, Computation of single-cell features and single-tissue features, Peripheral group feature evolution for cell-cell features and cell-tissue features, Cell classification for identifying lymphocytes, and Lymphocyte group detection for SS diagnosis (Fig. 1). Through preprocessing, we detect cells and tissue structures in pathological images. Feature evolution based on a graph model is a key aspect of CTG-PAM. After computing single-cell features, two graph models are used to derive two types of group features, which describe the characteristics of surrounding cells and the surrounding tissue structures, respectively. By combining single-cell features with these two types of group features, we perform cell classification to identify lymphocytes. Based on the recognition of lymphocytes, the model detects lymphocyte groups in the images, thereby achieving automatic diagnosis of SS. Below, we will introduce these five stages one by one.

#### Preprocessing

The primary focus of the preprocessing stage is the detection of cells and tissue structures within the regions of interest in pathological images. We initially utilize two models from prior research, namely “Stardist” [30] and “Segment Anything” [31] to detect cells and tissue structures respectively. Specifically, the detected cells and tissue structures are represented as two sets:

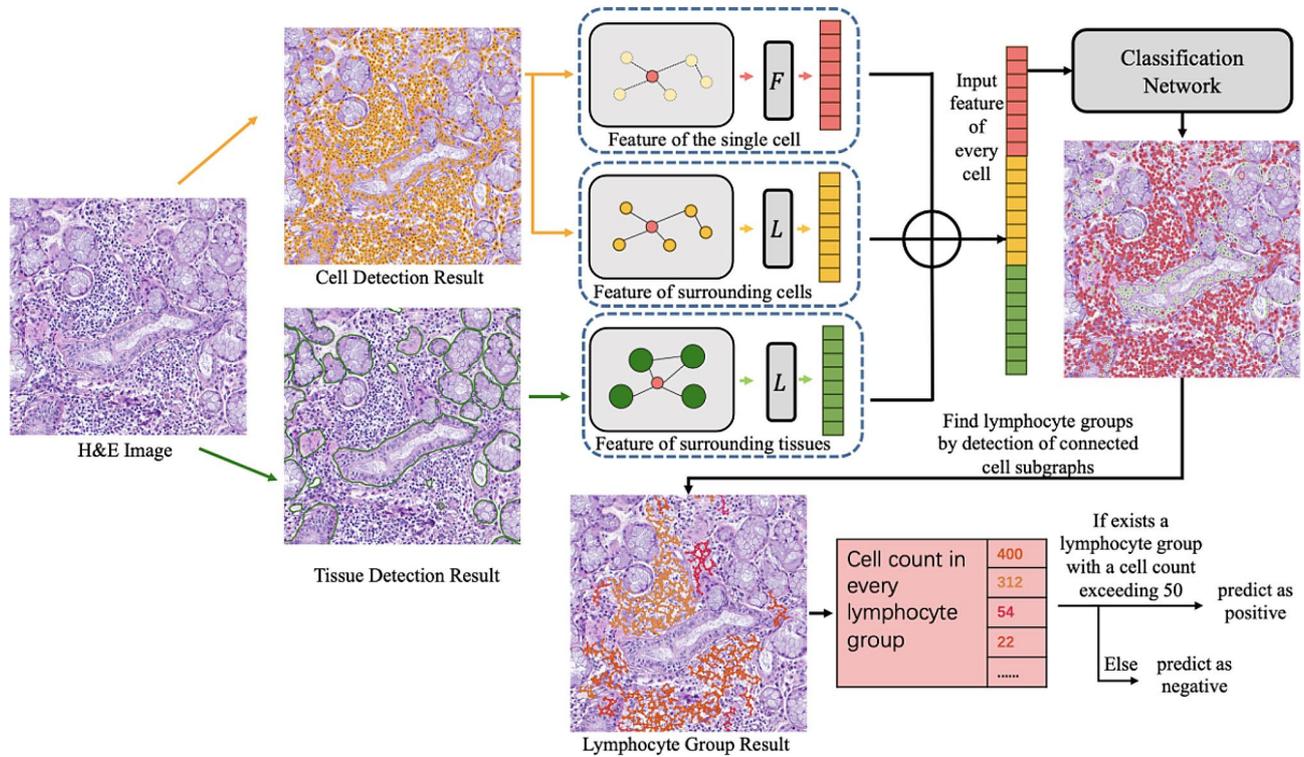
$$C = \{c_1, \dots, c_{n_c}\}, T = \{t_1, \dots, t_{n_t}\},$$

where  $n_c$  and  $n_t$  are the numbers of detected cells and tissue structures, respectively.

#### Computation of single-cell features and single-tissue features

Cell morphology is the foundation and gold standard for determining cell origin, type, and characteristics [32, 33]. Work based on handcrafted morphological features has achieved excellent performance in cell classification tasks [34–39]. Consequently, we use eight morphological features commonly employed in image processing to characterize the inherent geometric properties of cells and tissue structures. The following are the eight features:

Feature 1: mean of grayscale value,



**Fig. 1** The pipeline of the proposed CTG-PAM for Automatic diagnostic method for SS automatic diagnostic

$x_{\text{mean}}$  = The mean of grayscale value of pixels in the cell or tissue region.

$$x_{\text{eccentricity}} = \frac{\sqrt{\left(\frac{x_{\text{major}}}{2}\right)^2 + \left(\frac{x_{\text{minor}}}{2}\right)^2}}{\frac{x_{\text{major}}}{2}}$$

Feature 2: variance of grayscale value,

$x_{\text{variance}}$  = The variance of grayscale value of pixels in the cell or tissue region.

Feature 8: compactness,

$$x_{\text{compactness}} = \frac{(x_{\text{perimeter}})^2}{x_{\text{area}}}$$

Feature 3: area,

$x_{\text{area}}$  = Total number of pixels in the cell or tissue region.

Feature 4: perimeter,

$x_{\text{perimeter}}$  = Total number of pixels along the boundary of the cell or tissue region.

Feature 5: major axis length,

$x_{\text{major}}$  = The length of the long axis of ellipse fitting to the cell or tissue region.

Feature 6: minor axis length,

$x_{\text{minor}}$  = The length of the short axis of ellipse fitting to the cell or tissue region.

Feature 7: eccentricity,

In these features, area [34–39], perimeter [34–36], major axis length [35, 36, 38, 39], minor axis length [35, 36, 38, 39], eccentricity [35–37] and compactness [37, 39] are used for cell classification in existing researches. The features other than compactness are calculated using the “scikit-image” library [40], and compactness is calculated directly from the above formula. The feature computation process is denoted by the operator  $\mathcal{F}$ , which can be expressed as:

$$x_i^{(c)} = \mathcal{F}(c_i), x_j^{(t)} = \mathcal{F}(t_j), i = 1, \dots, n_c, j = 1, \dots, n_t,$$

where,

$$\begin{aligned}
x_i^{(c)} &= \left[ (x_{\text{mean}})_i^{(c)}, (x_{\text{variance}})_i^{(c)}, (x_{\text{area}})_i^{(c)}, (x_{\text{perimeter}})_i^{(c)}, (x_{\text{major}})_i^{(c)}, \right. \\
&\quad \left. (x_{\text{minor}})_i^{(c)}, (x_{\text{eccentricity}})_i^{(c)}, (x_{\text{compactness}})_i^{(c)} \right], x_j^{(t)} \\
&= \left[ (x_{\text{mean}})_j^{(t)}, (x_{\text{variance}})_j^{(t)}, (x_{\text{area}})_j^{(t)}, (x_{\text{perimeter}})_j^{(t)}, \right. \\
&\quad \left. (x_{\text{major}})_j^{(t)}, (x_{\text{minor}})_j^{(t)}, (x_{\text{eccentricity}})_j^{(t)}, (x_{\text{compactness}})_j^{(t)} \right] \in \mathbb{R}^8
\end{aligned}$$

These single-cell features and single-tissue features will serve as the foundation for the subsequent feature evolution that derives peripheral group features. It is worth noting that the selection of these eight fundamental features is not fixed. The number and types of fundamental features can be adjusted according to different application scenarios or diseases.

#### **Peripheral group feature evolution for cell-cell features and cell-tissue features**

When classifying each cell, doctors consider not only the characteristics of the cell itself but also its surrounding context. In existing studies [24–27], cell features are evolved to obtain more complex features for cell classification. We establish two graph models to derive two types of peripheral group features, including the features of surrounding cells, named cell-cell features, and surrounding tissue structures, named cell-tissue features, based on single-cell features and single-tissue features. In this process, each cell is sequentially assessed to determine if it is a lymphocyte, with the cell under evaluation referred to as the target cell. During this process, we establish two corresponding graph structures for each cell: the cell-cell graph (CCG) and the cell-tissue graph (CTG), where the CCG representing the relationships between the cell and its surrounding cells, and the CTG representing the relationships between the cell and its surrounding tissue structures. In both graphs, cells or tissue structures are nodes and edges are formed between those with relatively small distances, with predefined parameters governing the number of cells/tissue structures included and the distance threshold for edge creation. This approach of establishing edges based on the spatial distance between nodes is known as the k-nearest neighbor algorithm and is commonly used in cell graph models [26, 34, 41]. To approximate the actual distances between cells or between cells and tissue, we calculate the Euclidean distance between the center coordinates of two objects, subtracting the sum of their radii. The radii are estimated based on the area. Specifically, for any two cells or tissue structures  $u, v \in C \cup T$ , the distance is defined as:

$$D(u, v) = D_{\text{center}}(u, v) - \left( \sqrt{\frac{S_u}{\pi}} + \sqrt{\frac{S_v}{\pi}} \right),$$

where  $D_{\text{center}}(u, v)$  represents the Euclidean distance between the centers of  $u$  and  $v$ , with  $S_u$  and  $S_v$  representing the areas of  $u$  and  $v$ , respectively.

Below is the specific process for constructing CCG and CTG. Let's denote the node feature matrix, degree matrix and adjacency matrix of the graph model as  $X, D, A$ . Specifically, denote the corresponding matrices for CCG and CTG as  $X_{\text{CCG}}, D_{\text{CCG}}, A_{\text{CCG}}$  and  $X_{\text{CTG}}, D_{\text{CTG}}, A_{\text{CTG}}$ . The maximum numbers of nodes for CCG and CTG are given as  $n_{\text{CCG}}$  and  $n_{\text{CTG}}$ . And the distance thresholds are  $d_{\text{CCG}}$  and  $d_{\text{CTG}}$ . The node feature matrices represent the set of node features, i.e.,  $X_{\text{CCG}} \in \mathbb{R}^{8 \times n_{\text{CCG}}}$ ,  $X_{\text{CTG}} \in \mathbb{R}^{8 \times n_{\text{CTG}}}$ . First, from  $C$  select the nearest  $n_{\text{CCG}} - 1$  elements to the target cell  $c_*$  based on their distance, forming a node set  $N_{c_*} = \{u_1, \dots, u_{n_{\text{CCG}}}\}$ , which contain  $n_{\text{CCG}}$  cells. The edge connections are determined based on the computed  $X_{\text{CCG}}$  and the threshold  $d_{\text{CCG}}$ , i.e., for  $A_{\text{CCG}}(n_{\text{CCG}} \times n_{\text{CCG}})$ :

$$A_{\text{CCG},i,j} = \begin{cases} 1, & \text{if } D(u_i, u_j) < d_{\text{CCG}} \\ 0, & \text{other.} \end{cases}$$

$D_{\text{CCG}}$  can be obtained from  $A_{\text{CCG}}$  as a diagonal matrix:

$$D_{\text{CCG},i,j} = \begin{cases} \sum_{k=1}^{n_{\text{CCG}}} A_{\text{CCG},i,k}, & i = j \\ 0, & i \neq j. \end{cases}$$

Similarly, for CTG, based on the predefined parameters  $n_{\text{CTG}}$  and  $d_{\text{CTG}}$ ,  $A_{\text{CTG}}, X_{\text{CTG}}, D_{\text{CTG}}$  can be obtained. This completes the construction of CCG and CTG for the target cell  $c_*$ .

After constructing the two graphs, we utilize the LightGCN model [42] for unsupervised graph feature evolution based on the two graphs. LightGCN is a widely used model for graph embedding feature computation, achieving significant success in graph-based fields such as recommender systems [43, 44]. Specifically, given the adjacency matrix  $A(n \times n)$ , the degree matrix  $D(n \times n)$ , and the initial feature matrix  $X(n \times q)$  of the graph, where  $n$  represents the number of nodes in the graph and  $q$  represents the number of initial features. By performing calculations, we obtain the final extracted embedding features

$$E = \sum_{k=0}^K \alpha_k E^{(k)},$$

where

$$E^{(k+1)} = \left( D^{-\frac{1}{2}} A D^{\frac{1}{2}} \right) E^{(k)},$$

and  $E^{(0)} = X$ . We denote the LightGCN as  $\mathcal{L}$ . According to the results in the paper, we set  $\alpha_k$  to be uniformly distributed as  $\frac{1}{K+1}$ , and fix  $K = 3$ . Thus, we can simplify the LightGCN as

$$E = \mathcal{L}(A, D, X).$$

Specifically, applying LightGCN to both the cell-cell graph and the cell-tissue graph can be expressed as

$$\begin{cases} E_{\text{cell}} = \mathcal{L}(A_{\text{cell}}, D_{\text{cell}}, X_{\text{cell}}), \\ E_{\text{tissue}} = \mathcal{L}(A_{\text{tissue}}, D_{\text{tissue}}, X_{\text{tissue}}), \end{cases}$$

where  $A_{\text{cell}}$ ,  $D_{\text{cell}}$  and  $X_{\text{cell}}$  represent the adjacency matrix, the degree matrix and the feature matrix of the cell-cell graph,  $A_{\text{tissue}}$ ,  $D_{\text{tissue}}$  and  $X_{\text{tissue}}$  represent the adjacency matrix, the degree matrix and the feature matrix of the cell-tissue graph. The results of feature evolution  $E_{\text{cell}}$  and  $E_{\text{tissue}}$  encompass the evolved features of all nodes in two graphs. We extract the cell-cell features and cell-tissue features corresponding to the target cell from  $E_{\text{cell}}$  and  $E_{\text{tissue}}$ . These features represent the information of the surrounding cells and tissue structures of the target cell.

#### Cell classification for identifying lymphocytes

For any cell  $c_i \in C$ , after concatenating cell-cell feature vector  $x_i^{(cc)}$  from CCG and cell-tissue feature vector  $x_i^{(ct)}$  from CTG with the original single-cell feature vector  $x_i^{(c)}$ , we employ a pre-trained classifier for cell classification. Specifically, we utilize a double-hidden-layer fully connected neural network, i.e. multilayer perceptron (MLP), as the classifier, and the classification process can be expressed as

$$p_i = W_3 \sigma (W_2 \sigma (W_1 x_i + b_1) + b_2),$$

where  $x_i = x_i^{(c)} \oplus x_i^{(cc)} \oplus x_i^{(ct)}$ ,  $x_i \in \mathbb{R}^{24}$  is the feature vector of the cell  $c_i$ ,  $W_1$ ,  $W_2$  and  $W_3$  are weight matrices,  $b_1$  and  $b_2$  are bias vectors,  $\sigma$  is the activation function, and  $p_i$  represents the probability distribution over the final class labels for the cell  $c_i$ . The class label with the highest probability in  $p_i$  is chosen as the final prediction. We use cross-entropy as the loss function for network training and employ Adam optimizer to optimize the loss function. Additionally, we ensure consistency in the types of cells sampled in each training batch by using a random sampling method.

#### Lymphocyte group detection for SS diagnosis

In the lymphocyte group detection stage after identifying lymphocytes, we detect lymphocyte groups based on connected components in the lymphocyte graph structure, as a cluster containing more than 50 lymphocytes in

a pathological image is diagnosed as SS according to clinical practice. First, we construct a graph with all lymphocytes in the pathological image as nodes, and we establish edges between two lymphocytes if their distance is less than a given threshold. Using a connected component detection algorithm based on breadth-first search [45], we extract all connected subgraphs from this graph and count the number of nodes in each connected subgraph. If there exists a connected subgraph with a node count exceeding 50, which serves as a diagnostic criterion for SS in clinical practice, then the patient corresponding to that pathological image is diagnosed with SS.

#### Hyperparameters setting

Here, we set hyperparameters in CTG-PAM. The maximum numbers of nodes and distance thresholds for edges in CCG and CTG, i.e.,  $n_{\text{CCG}}$ ,  $n_{\text{CTG}}$ ,  $d_{\text{CCG}}$ ,  $d_{\text{CTG}}$ , are set to 20, 2, 50, 5, respectively. In the fully connected neural network for cell classification, the width of hidden-layers is set to 50, the activation function is ReLU, and the training learning rate is 0.001. The subgraph edge establishment threshold for lymphocyte groups detection is set to 30. The experimental results include a sensitivity analysis of the hyperparameters in both CCG and CTG.

#### Statistical analysis

Accuracy, Sensitivity, and Specificity are used to evaluate diagnostic performance, and the formula is shown below. Fisher exact test is used to compare the diagnostic Accuracy, Sensitivity, and Specificity between model and pathologists of different qualifications via GraphPad Prism 8.0 software. When values of  $p < 0.05$ , were considered statistically significant (\* $p < 0.05$ ; \*\*\* $p < 0.001$ ).

$$Accuracy = \frac{TP + TN}{TP + EP + TN + FN},$$

$$Sensitivity = \frac{TP}{TP + FN},$$

$$Specificity = \frac{TN}{TN + FP}.$$

## Result

#### Lymphocyte identification performance of CTG-PAM

Accurate identification of lymphocytes is crucial in the SS diagnosis process. Therefore, we evaluate the lymphocyte identification performance in CTG-PAM. The experiment is divided into three parts. First, we validate that whether the eight single-cell features effectively represent the image information of individual cells by comparing the classification performance based on single-cell features and images. Second, we conduct sensitivity

analysis for the hyperparameters in the CCG and the CTG, including the maximum number of nodes:  $n_{CCG}$  and  $n_{CTG}$ , and the distance threshold parameters for establishing edges between nodes:  $d_{CCG}$  and  $d_{CTG}$ . Third, we perform feature ablation experiments to show the effects of the CTG, which taking the cell-tissue information into the model for cell classification.

#### **Validation the reasonableness of single-cell features**

Eight fundamental features' effectiveness in representing cell information determines the effectiveness of subsequently evolved group features, thereby impacting the performance of cell classification. Therefore, we validated the reasonableness of the single-cell features by verifying whether the eight single-cell features can effectively represent individual cell information through the following experiment. It is generally believed that, compared to designing feature extraction methods for images followed by classification, automatic feature extraction based on deep learning, represented by CNNs, can fully utilize image information [46]. Therefore, we conducted cell classification on the internal validation dataset using an MLP based on the eight single-cell features and a CNN based on images of each cell. We compared the performance of these two methods to verify whether the eight single-cell features sufficiently extracted the information from the images. The MLP's hyperparameters were kept consistent with the previous setting. The CNN comprises two convolutional layers and two fully connected layers. The convolutional layers have 16 and 32  $3 \times 3$  kernels, respectively, each followed by a ReLU activation function and a  $2 \times 2$  max pooling layer. The two fully connected layers have the same widths and activation functions as those in the MLP.

The experimental results showed that both methods exhibited high sensitivity and low specificity for lymphocyte identification. Specifically, the accuracy, sensitivity, and specificity of the MLP based on the eight single-cell features were 0.9192, 0.9556, and 0.4902, respectively, while the accuracy, sensitivity, and specificity of the CNN based on images were 0.8894, 0.9225, and 0.5004, respectively.

On the one hand, the similar performance of the feature-based method to the image-based method suggests that the eight single-cell features adequately extract the image information of individual cells. On the other hand, the experimental results with unsatisfactory specificity also highlighted the need to utilize group features in addition to single-cell features for cell classification.

#### **Sensitivity analysis for the hyperparameters in CCG and CTG**

Among the three types of features (single-cell features, cell-cell features and cell-tissue features) used for cell classification, the calculation of single-cell features does

not involve hyperparameters that need adjustment. In contrast, the features of surrounding cells and surrounding tissues each include two hyperparameters: the maximum number of nodes ( $n_{CCG}$  and  $n_{CTG}$ ) and the distance threshold for establishing edges between nodes ( $d_{CCG}$  and  $d_{CTG}$ ). Through the adjustment and experimentation of these four parameters on the training dataset, we ultimately determined an optimal set of parameters. Specifically, the maximum number of nodes in the cell graph,  $n_{CCG}$ , is 20; the maximum number of nodes in the tissue graph,  $n_{CTG}$ , is 2; the distance threshold for edge connections in the cell graph,  $d_{CCG}$ , is 50; and the distance threshold for edge connections in the tissue graph,  $d_{CTG}$ , is 5. This set of parameters demonstrated the best cell classification performance in our experiments, with the accuracy of 0.9723, the sensitivity of 0.9741, and the specificity of 0.9518 in the internal validation dataset and the accuracy of 0.8024, the sensitivity of 0.7780, and the specificity of 0.9402 in the external dataset.

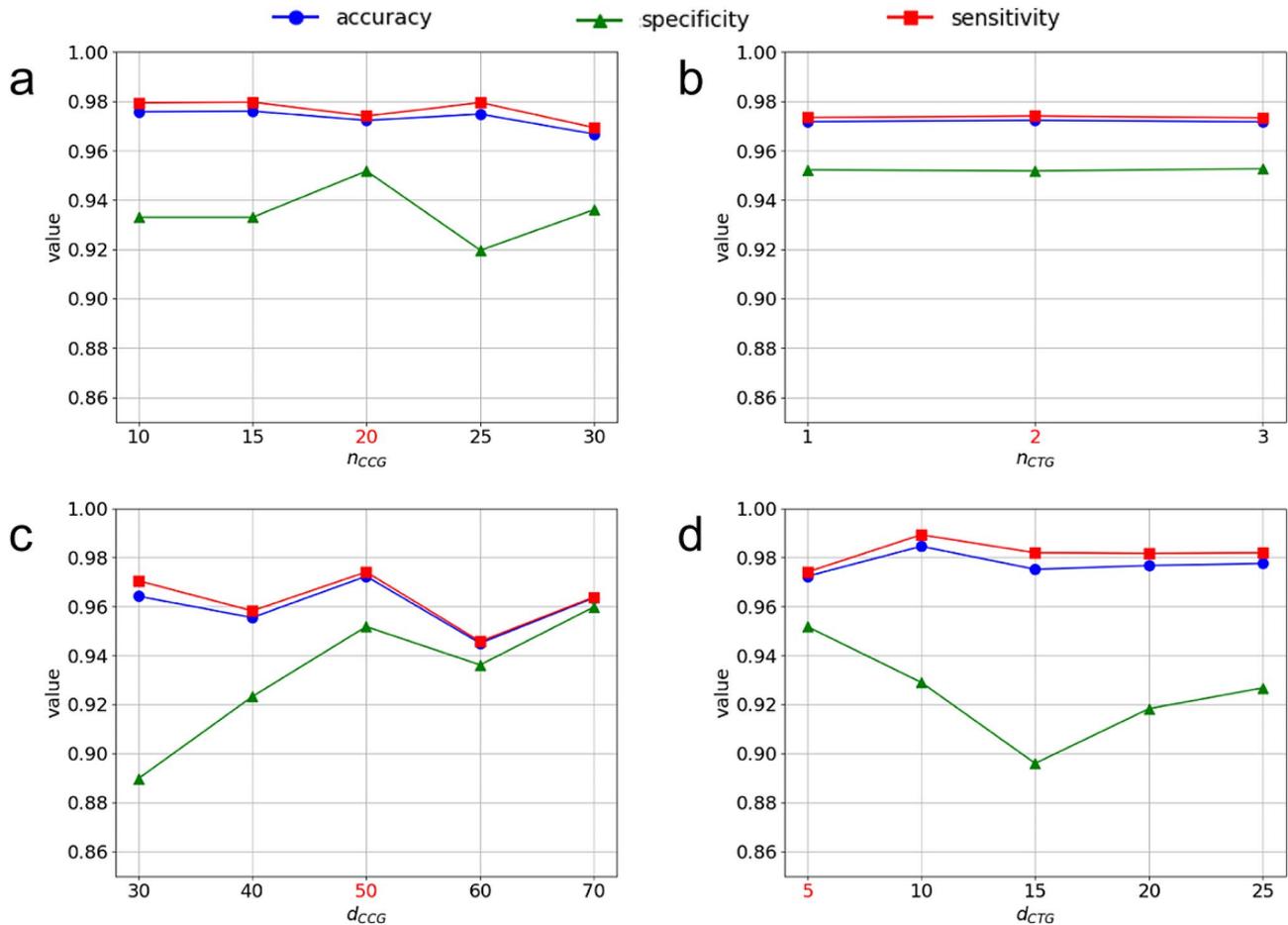
Using this parameter combination, we conducted a sensitivity analysis for each parameter, examining how changes in the parameters affected cell classification performance in Fig. 2. The results indicate that within a certain range of parameter variations, the cell classification performance does not undergo significant changes.

#### **Evaluating the impact of cell-tissue features**

We argue that cell-tissue information is important for accurate cell classification. To evaluate the impact of cell-tissue features provided by the CTG on classification performance, we designed the following experiment. We compared CTG-PAM with a model that does not use cell-tissue features. The model that uses only cell-cell features derived from CCG for cell classification is referred to as "CCG w/o CTG." The experimental results of cell classification performance for both CCG w/o CTG and CTG-PAM are shown in Table 1. Based on the experimental results, we found that adding cell-tissue features enhanced the model's cell classification performance. In the next section, we will further discuss the diagnostic effectiveness of these features combinations.

#### **Diagnostic performance of CTG-PAM and other models**

We employed a model, referred to cell group detection model (CGDM), focused solely on cell group detection without lymphocyte differentiation, specifically only utilizing the stage "Lymphocyte group detection for SS diagnosis" of CTG-PAM, for comparative purposes in subsequent experiments. We completed the qualitative classification of both the internal validation and external test data sets by using the ResNet-50 [23], CGDM, CCG w/o CTG, and CTG-PAM. The areas under the receiver operating characteristic curve from the internal validation and external test data sets of CTG-PAM



**Fig. 2** Sensitivity analysis of key parameters affecting cell classification performance. **(a)** maximum node number in cell graph ( $n_{CCG}$ ), **(b)** maximum node number in tissue graph ( $n_{CTG}$ ), **(c)** edge connection distance threshold in cell graph ( $d_{CCG}$ ), and **(d)** edge connection distance threshold in tissue graph ( $d_{CTG}$ ). The red numbers on the horizontal axes represent the optimal parameters we selected

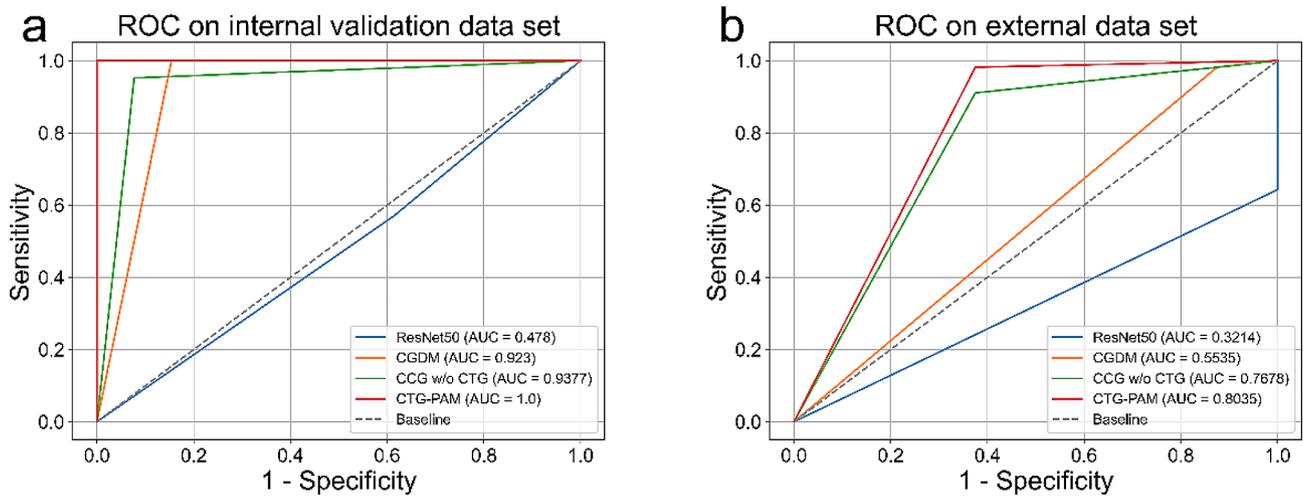
**Table 1** Performance of cell classification for identifying lymphocytes with and without cell-tissue features in the internal validation data set and the external data set

Method	Internal validation data set			External data set		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
CCG w/o CTG	0.9183	0.9214	0.9180	0.7169	<b>0.9957</b>	0.6674
CTG-PAM	<b>0.9723</b>	<b>0.9518</b>	<b>0.9741</b>	<b>0.8024</b>	0.9402	<b>0.7780</b>

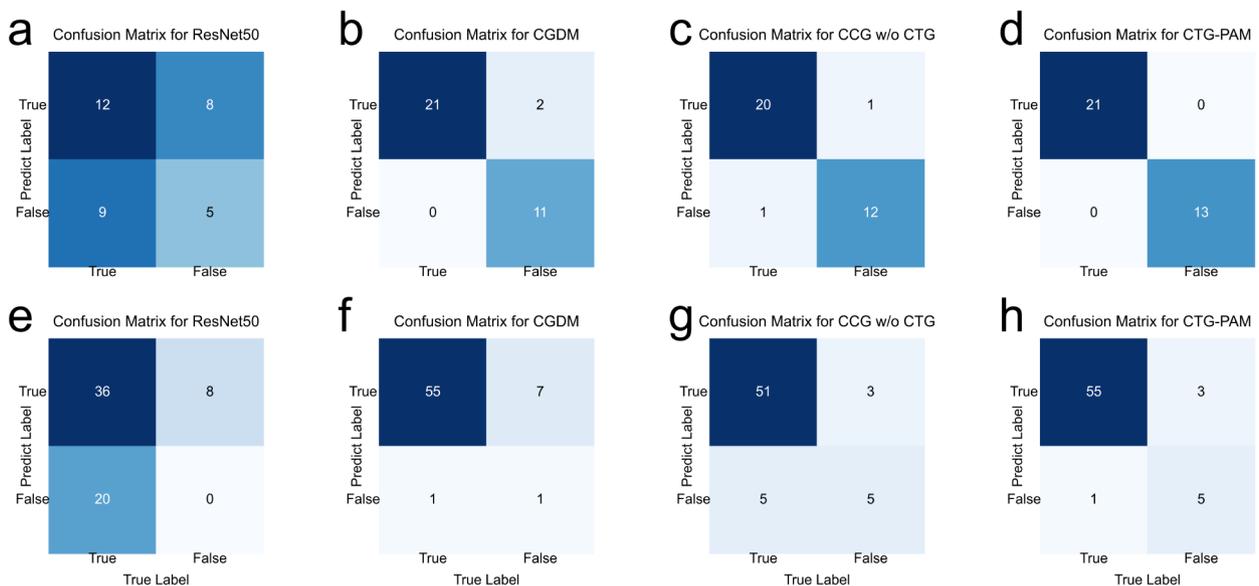
were 1.0 and 0.8035 respectively, which indicated that this was the best model (Fig. 3, A and B). The respective sensitivities for the internal and external data sets were 57.14% and 64.28% for ResNet-50, 100% and 98.21% for CGDM, 100% and 91.07% for CCG w/o CTG, and 100% and 98.21% for CTG-PAM, respectively. The corresponding accuracy were 50% and 56.25%, 94.11% and 87.5%, 97.06% and 87.5%, and 100% and 93.75%, individually. Our study showed that internal authentication was far superior to external authentication (Fig. 4; Tables 2 and 3). The CTG-PAM was visualized to identify the areas of tissue infiltration by lymphocytes, and both H&E and IHC confirmed that the model was able to learn from the characteristics of lymphocytes only (Fig. 5, A and B).

**Achieving specialist levels of SS detection performance: interpretation accuracy comparison among beginner pathologists, skilled pathologists and CTG-PAM**

CTG-PAM had exceeded skilled pathologists on prediction accuracy scores (Fig. 6, A and B). The average accuracy rates of the validation data set were 79.41%, 97.05% and 100% for beginner pathologists, skilled pathologists, and CTG-PAM, respectively. The average accuracy rates of the external data set were 84.37%, 90.62% and 93.75% for beginner pathologists, skilled pathologists, and CTG-PAM. On the other hand, beginner pathologists demonstrated markedly worse prediction accuracy scores. The AUC from the internal validation and external test data sets were 1 and 0.8035 for CTG-PAM, 0.8035 and



**Fig. 3** Performance of the CTG-PAM. **(a)** and **(b)**, The ROC and the AUC for diagnostic. **(a)**, Comparison of the AUC/ROC for ResNet-50, CGDM, CCG w/o CTG, and CTG-PAM by using the internal test data set. CTG-PAM had an AUC (1.0) significantly greater than those of the other two models. **(b)**, Comparison of the AUC/ROC for ResNet-50, CGDM, CCG w/o CTG, and CTG-PAM in the independent external test data set. CTG-PAM also provided the best AUC (0.8035) compared with those of the other ones



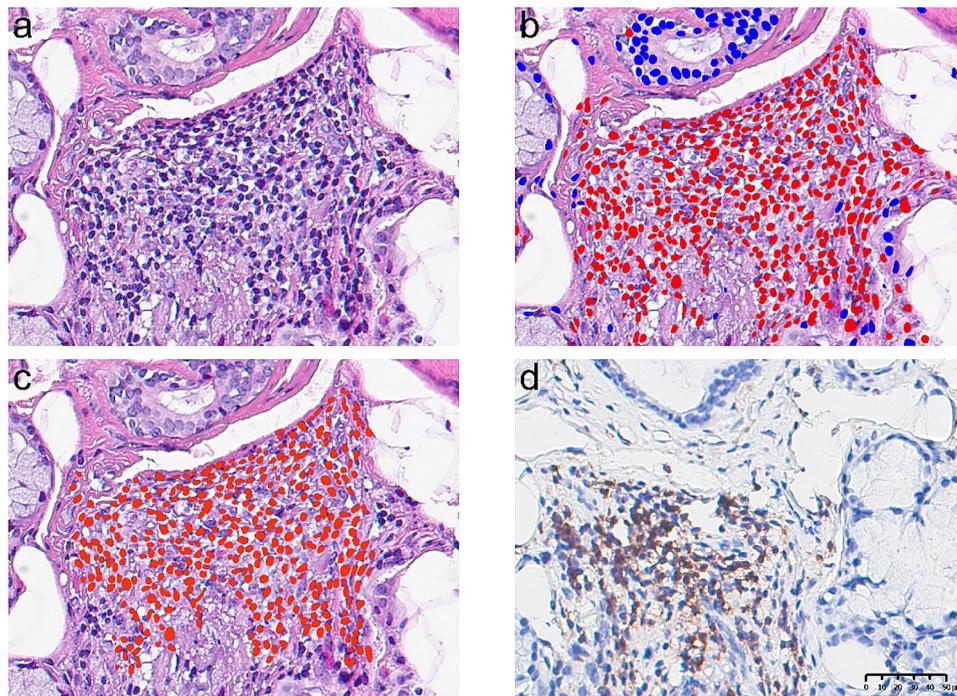
**Fig. 4** Confusion matrix of the models' diagnosis performance. **(a)-(d)**, Confusion matrices of the internal validation data set for models of ResNet-50, CGDM, CCG w/o CTG, and CTG-PAM, respectively. **(e)-(h)**, Confusion matrix of the independent external test data set for models of ResNet-50, CGDM, CCG w/o CTG, and CTG-PAM, respectively

**Table 2** Model and algorithm performance evaluation in internal validation data set

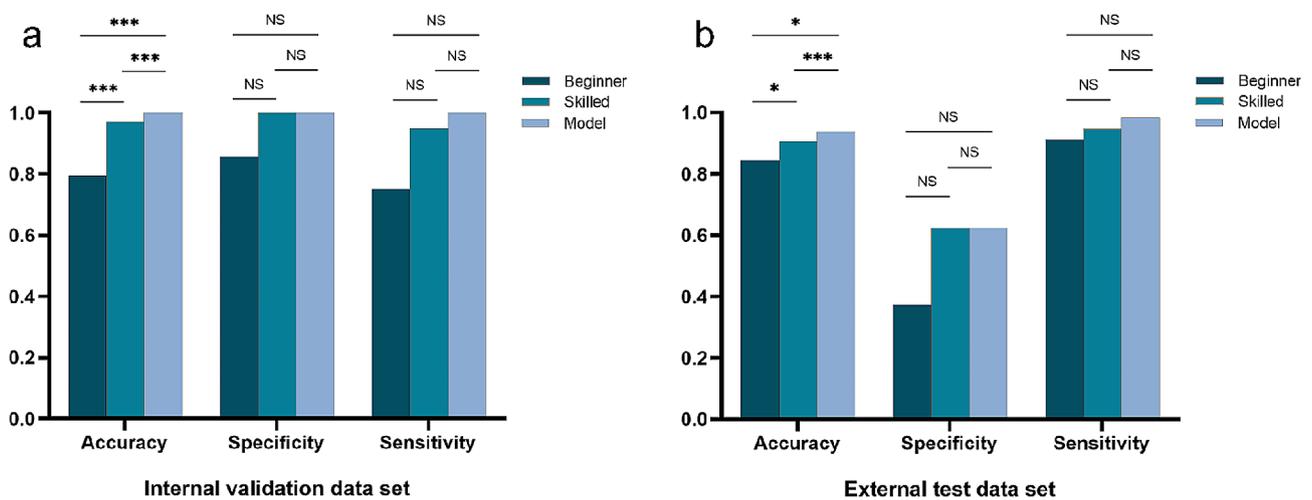
Method	TP	TN	FP	FN	Accuracy	Specificity	Sensitivity
ResNet-50	12	5	8	9	0.5000	0.3846	0.5714
CGDM	21	11	2	0	0.9411	0.8461	<b>1.0000</b>
CCG w/o CTG	21	12	1	0	0.9706	0.9231	<b>1.0000</b>
CTG-PAM	21	13	0	0	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>

**Table 3** Model and algorithm performance evaluation in the external data set

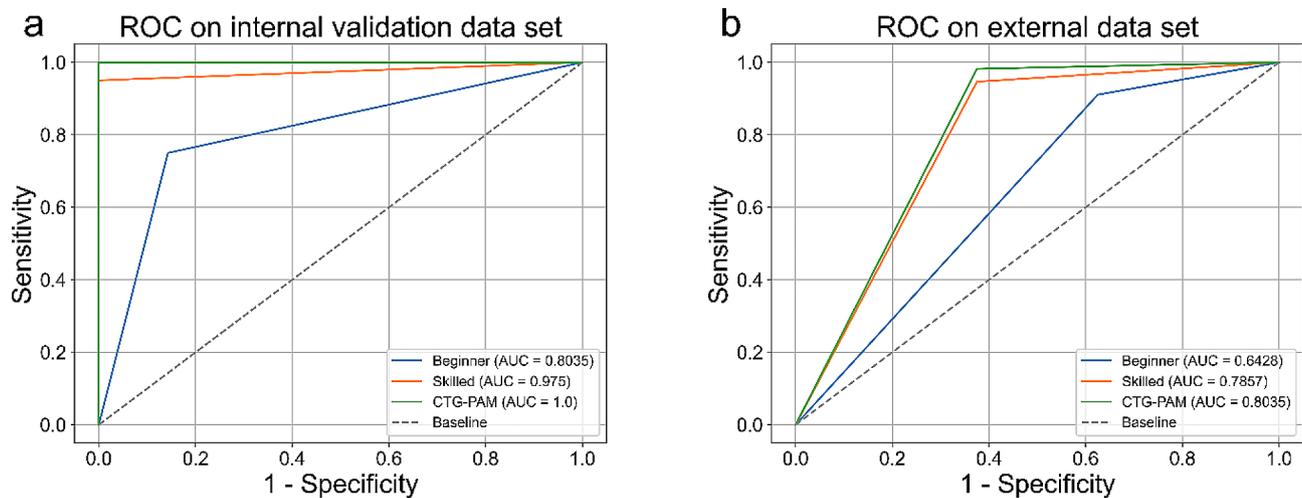
Method	TP	TN	FP	FN	Accuracy	Specificity	Sensitivity
ResNet-50	36	0	8	20	0.5625	0.0000	0.6428
CGDM	55	1	7	1	0.8750	0.1250	<b>0.9821</b>
CCG w/o CTG	51	5	3	5	0.8750	<b>0.6250</b>	0.9107
CTG-PAM	55	5	3	1	<b>0.9375</b>	<b>0.6250</b>	<b>0.9821</b>



**Fig. 5** The models' classification result of patch with a focus lymphocyte. (a), lymphocyte focus beside duct. (b), cell aggregates identify images, where distinct cell types are denoted by varying colors. The red color is employed to denote lymphocytes, whereas blue indicates duct cells. (c), a lymphocytes group, denoted in red color, detected by the connected subgraph detection technique on the constructed lymphocyte graphs. (d), CD45 confirmed that CTG-PAM identified the areas of lymphocyte focus



**Fig. 6** Comparisons between CTG-PAM and pathologists of different qualifications. (a). The diagnostic results between model and pathologists of different qualifications in the internal validation data set CTG-PAM, beginner pathologists, and skilled pathologists were significant differences among these three groups on the average accuracy rates. (b). The diagnostic results between model and pathologists of different qualifications in the external data set CTG-PAM, beginner pathologists, and skilled patho differed among these three groups on the average accuracy rates



**Fig. 7** Comparisons of ROC between CTG-PAM and two groups of human pathologist. **(a)**, Comparison of the AUC/ROC for CTG-PAM, beginner pathologists, and skilled pathologists by using the internal test data set. CTG-PAM had an AUC (1.0) significantly greater than pathologists. **(b)**, Comparison of the AUC/ROC for CTG-PAM, beginner pathologists, and skilled pathologists in the independent external test data set. CTG-PAM also provided the best AUC (0.8035) compared with those of the other ones

**Table 4** Diagnostic results between model and pathologists of different qualifications

Group	Internal validation data set			External data set		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
Beginner	0.7941	0.8571	0.7500	0.8437	0.3750	0.9107
Skilled	0.9705	1.0000	0.9500	0.9062	0.6250	0.9464
CTG-PAM	1.0000	1.0000	1.0000	0.9375	0.6250	0.9821

0.6428 for beginner pathologists, and 0.975 and 0.7857 for skilled pathologists. The average accuracy rates of the validation data set were 79.41%, 97.05%, and 100% for beginner pathologists, skilled pathologists, and CTG-PAM, respectively (Fig. 7, A and B). The average accuracy rates of the external data set were 84.37%, 90.62%, and 93.75% for beginner pathologists, skilled pathologists, and CTG-PAM. The different qualifications on the validation data set and external data set are shown in Table 4.

The CTG-PAM significantly reduces the diagnosis time compared to pathologists, as pathologists also need to invest time and effort in manually counting lymphocytes and lymphocytes foci under the microscope. We compared the time required for diagnosis between the CTG-PAM and pathologist judgment. The result showed that the CTG-PAM (22.22s, 95%CI, 21.49–22.96) took less time than beginner pathologist judgment (248.82s, 245.90–251.73) and skilled pathologist judgment (129.91s, 127.70–132.12) (Table S2, in supporting information).

## Discussion

Through single-cell features, cell cluster features, and cell-tissue features, our model can effectively accomplish cell classification, facilitating flexible subsequent research and analysis. To our knowledge, this is the first study to identify and evaluate a predictive

model for SS. The disease of SS has well-established and specific diagnostic criteria in pathological clinical settings, specifically, the presence of cell clusters with a lymphocyte count greater than 50, which confirms the diagnosis of SS [16]. Therefore, detecting cell clusters for automated diagnosis is a natural approach, and this can be achieved directly using connected component analysis in the cell graph. However, during the implementation of this method, it was found that its accuracy depends on the accurate detection of lymphocytes. Consequently, enabling the model to understand the definition of lymphocytes assumes paramount significance.

Therefore, a cell-level classification model that describes the characteristics of surrounding cells and the structure of surrounding tissues is necessary. Cell morphology is the foundation and gold standard for determining cell origin, type, and characteristics. Initially, our model utilized a relatively simple set of eight features, combined with the clinical expertise of pathologists in cell identification, to classify cells. However, the experimental results for lymphocyte identification based on single cells show that neither the extracted single-cell features nor the image-level classification of individual cells can provide satisfactory lymphocyte identification performance. This suggests that relying solely on single-cell features to

identify cell types is insufficient. In fact, in practical clinical applications, the differential diagnosis of non-neoplastic and neoplastic diseases often requires consideration of the characteristics of surrounding cells and tissues from multiple perspectives. Regarding SS, the disease's characteristic lies in the identification of lymphocytes surrounding the duct, which serves as SS's diagnostic criterion and foundation. This holds great significance for the condition. In contrast, for specific tumors such as breast cancer, esophageal cancer, and thyroid cancer, the diagnostic criteria primarily emphasize the overall morphology, size, type, and structural arrangement of tumor cells. Based on this, we further improved our model and developed CTG-PAM, a model that extracts cell features from these three perspectives. Training the model with a small amount of labeled data allows it to distinguish lymphocytes from other cell types. Classifying cells before detecting cell clusters effectively enhances detection accuracy and reduces false positives. CTG-PAM demonstrates high accuracy and sensitivity in cell classification and diagnosis, showing great potential in cases where the number of tumor cells is low and pathologists face diagnostic challenges.

Furthermore, the emergence of precision medicine has significantly transformed cancer management. Clinicians are increasingly interested in the relationship between tumors and immune cells in the surrounding microenvironment. Given CTG-PAM's precise ability to identify and count lymphocyte foci, this algorithm can extend its accuracy to the identification and counting of immune cells in the tumor microenvironment. This has the potential to significantly impact the treatment and prognosis of patients with breast cancer, gastric cancer, and squamous cell carcinoma. While the accuracy of cell classification may vary between different diseases, this does not affect diagnostic outcomes.

This study has several limitations. Despite the use of a total of 100 WSIs, the number of patients is small to draw a conclusive outcome regarding the application of deep learning in the pathology diagnosis of SS since three images from a single patient may have similarities. Further analysis with more patients and control subjects using H&E stained sections from multiple hospitals is needed. The control group consisted of patients who underwent lip gland biopsy but were not diagnosed with SS, without normal lip gland tissue or other parts such as the parotid gland. Distinguishing SS or sialadenitis of the parotid gland may be crucial, as the parotid gland is also one of the common sites for SS or sialadenitis. In this regard, the model may be effective in differentiating between sialadenitis and SS, as the lymphocytes in the sialadenitis of the parotid gland are scattered in the fibrous tissue, while SS is

characterized by lymphocytic infiltration around the ducts.

## Conclusion

Overall, the CTG-PAM shows great potential in improving the prognosis of patients with SS and has significant potential for diseases requiring multi-perspective differential diagnosis. It also demonstrates exceptional accuracy in cell classification and diagnosing SS, providing clinical pathologists with objective, precise, reproducible, and standardized diagnostic results.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05550-8>.

Supplementary Material 1

Supplementary Material 2

## Acknowledgements

We thank Dr. Xuecong Li from the Hospital of Stomatology of Sun Yat-sen University for aiding in the data acquisition.

## Author contributions

Z.P.C., R.F.W., Y.J., and J.X. conceived the study. Z.P.C., R.F.W., Y.J., and J.X. designed the experiments. Z.P.C., A.J.H., X.Y.C., M.X., Z.N.F., and X.C.L. provided data resources and collected data. R.F.W., Z.P.C., J.L.Y., and P.L. completed all codes of data processing, model training and testing. R.F.W., Z.P.C., and J.L.Y. performed the experiments analysis. R.F.W., Z.P.C., J.L.Y., and P.L. prepared the manuscript and revised the writing. Y.J., B.C., and J.X. supervised the research. All authors have read and agreed to publish the paper.

## Funding

Our study was funded by the National Natural Science Foundation of China (No.82170960, No.U23A20445 and No.81870769), Guangzhou Science and Technology Program Project (No.006038 and No.202201010895), the Key-Area Research and Development Program of Guangdong Province, China (No.2021B0101190003) and the Natural Science Foundation of Guangdong Province, China (No.2022A1515010831).

## Data availability

The pathological images used in this paper is not publicly available. However, they can be obtained by contacting the corresponding author ([xiajuan@mail.sysu.edu.cn](mailto:xiajuan@mail.sysu.edu.cn)) for scientific research purposes. The code of the CTG-PAM in this paper is available online at <https://github.com/Pan-Wu/CTG-PAM.git>.

## Declarations

### Ethics approval and consent to participate

This study protocol was reviewed and approved by the Hospital of Stomatology, Sun Yat-sen University (approval #KQEC-2023-84-01).

### Consent for publication

Not applicable.

### Competing interests

There are no competing interests in this work.

### Author details

<sup>1</sup>School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

<sup>2</sup>Department of Oral Medicine, Hospital of Stomatology, Sun Yat-sen University, Guangzhou, China

<sup>3</sup>Guangdong Provincial Key Laboratory of Stomatology, Guanghua School of Stomatology, Sun Yat-sen University, Guangzhou, China

<sup>4</sup>Department of Pathology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

Received: 5 February 2024 / Accepted: 29 July 2024

Published online: 08 August 2024

## References

- Goules AV, Tzioufas AG. Lymphomagenesis in Sjogren's syndrome: predictive biomarkers towards precision medicine. *Autoimmun Rev*. 2019;18:137–43.
- Bombardieri M, Argyropoulou OD, Ferro F, Coleby R, Pontarini E, Governato G, et al. One year in review 2020: pathogenesis of primary Sjogren's syndrome. *Clin Exp Rheumatol*. 2020;38(126):3–9.
- Travaglino A, Giordano C, Pace M, Varricchio S, Picardi M, Pane F, et al. Sjogren syndrome in primary salivary gland lymphoma. *Am J Clin Pathol*. 2020;153:719–24.
- Soret P, Le Dantec C, Desvaux E, Foulquier N, Chassagnol B, Hubert S, et al. A new molecular classification to drive precision treatment strategies in primary Sjogren's syndrome. *Nat Commun*. 2021;12:3523.
- Maciel G, Crowson CS, Matteson EL, Cornec D. Prevalence of primary Sjogren's syndrome in a US population-based cohort. *Arthritis Care Res*. 2017;69:1612–6.
- Stefanski AL, Tomiak C, Pleyer U, Dietrich T, Burmester GR, Dornier T. The diagnosis and treatment of Sjogren's Syndrome. *Dtsch Arztebl Int*. 2017;114(20):354–61.
- Park YS, Gauna AE, Cha S. Mouse models of primary Sjogren's syndrome. *Curr Pharm Des*. 2015;21:2350–64.
- Psianou K, Panagoulas I, Papanastasiou AD, de Lastic AL, Rodi M, Spantidea PI, et al. Clinical and immunological parameters of Sjogren's syndrome. *Autoimmun Rev*. 2018;17:1053–64.
- Fragkioudaki S, Mavragani CP, Moutsopoulos HM. Predicting the risk for lymphoma development in Sjogren syndrome: an easy tool for clinical use. *Medicine*. 2016;95:e3766.
- Pezoulas VC, Goules A, Kalatzis F, Chatzis L, Kourou KD, Venetsanopoulou A, et al. Addressing the clinical unmet needs in primary Sjogren's syndrome through the sharing, harmonization and federated analysis of 21 European cohorts. *Comput Struct Biotechnol J*. 2022;20:471–84.
- Brito-Zeron P, Kostov B, Fraile G, Caravia-Durán D, Maure B, Rascón FJ, et al. Characterization and risk estimate of cancer in patients with primary Sjogren syndrome. *J Hematol Oncol*. 2017;10:90.
- Hajjabbari A, Shenavar MI, Alizadeh Y, Banikarimi AS, Ghavidel PP. Secondary Sjogren's syndrome in 83 patients With Rheumatoid Arthritis. *Acta Med Iran*. 2016;54(7):448–53.
- Vitali C, Bombardieri S, Moutsopoulos HM, Balestrieri G, Bencivelli W, Bernstein RM, et al. Preliminary criteria for the classification of Sjogren's syndrome. Results of a prospective concerted action supported by the European community. *Arthritis Rheum*. 1993;36:340–7.
- Vitali C, Bombardieri S, Jonsson R, Moutsopoulos HM, Alexander EL, Carsons SE, et al. Classification criteria for Sjogren's syndrome: a revised version of the European criteria proposed by the American-European consensus group. *Ann Rheum Dis*. 2002;61:554–8.
- Shiboski SC, Shiboski CH, Criswell LA, Baer AN, Challacombe S, Lanfranchi H, et al. American College of Rheumatology classification criteria for Sjogren's syndrome: a data-driven, expert consensus approach in the Sjogren's international collaborative clinical alliance cohort. *Arthritis Care Res (Hoboken)*. 2012;64:475–87.
- Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjogren's syndrome: a consensus and data-driven methodology involving three international patient cohorts. *Arthritis Rheumatol*. 2017;69:35–45.
- Puladi B, Ooms M, Kintsler S, Houschyar KS, Steib F, Modabber A, et al. Automated PD-L1 scoring using artificial intelligence in head and neck squamous cell carcinoma. *Cancers*. 2021;13:4409.
- Valieris R, Amaro L, Osorio C, Bueno AP, Mitrowsky RAR, Carraro DM, et al. Deep learning predicts underlying features on pathology images with therapeutic relevance for breast and gastric cancer. *Cancers*. 2020;12:3687.
- Hermesen M, de Bel T, den Boer M, Steenbergen EJ, Kers J, Florquin S, et al. Deep learning-based histopathologic assessment of kidney tissue. *J Am Soc Nephrol*. 2019;30:1968–79.
- Cheng N, Ren Y, Zhou J, Zhang YW, Wang DY, Zhang XF, et al. Deep learning-based classification of hepatocellular nodular lesions on whole-slide histopathologic images. *Gastroenterology*. 2022;162:1948–61.
- Jin L, Shi F, Chun Q, Chen H, Ma YX, Wu S, et al. Artificial intelligence neuropathologist for glioma classification using deep learning on hematoxylin and eosin stained slide images and molecular markers. *Neuro Oncol*. 2021;23:44–52.
- Wang R, Dai W, Gong J, Huang MZ, Hu TD, Li H, et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. *J Hematol Oncol*. 2022;15:11.
- He K, Zhang X, Ren S et al. Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: IEEE; 2016. pp. 770–8.
- Jaume G, Pati P, Buzorgtabar B, Sun J et al. Quantifying explainers of graph neural networks in computational pathology. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: IEEE; 2021. pp. 8106–16.
- Studer L, Wallau J, Dawson H, Zlobec I, Fischer A. Classification of intestinal gland cell-graphs using graph neural networks. In: *International Conference on Pattern Recognition (ICPR)*. Milan: IEEE; 2021. pp. 3636–43.
- Wang J, Chen RJ, Lu MY, Baras A, Mahmood F. Weakly supervised prostate tma classification via graph convolutional networks. In: *IEEE International Symposium on Biomedical Imaging (ISBI)*. USA: IEEE; 2020. pp. 239–43.
- Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, Lindeman NI, et al. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans Med Imaging*. 2022;41:757–70.
- Lu W, Graham S, Bilal M, Rajpoot N, Minhas F. Capturing cellular topology in multi-gigapixel pathology images. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: IEEE; 2020. pp. 260–1.
- Raju A, Yao J, Haq MM, Jonnagaddala J, Huang J. Graph attention multi-instance learning for accurate colorectal cancer staging. In: Anne M, Purang A, Danail S, Diana M, Maria Z, Zhou K, et al. editors. *Medical Image Computing and Computer assisted intervention (MICCAI)*. Lima: Springer Cham; 2020. pp. 529–39.
- Schmidt U, Weigert M, Broaddus C, Broaddus C, Myers G. Cell detection with star-convex polygons. In: Frangi A, Schnabel J, Davatzikos C, Alberola-López C, Fichtinger G, editors. *Medical Image Computing and Computer assisted intervention (MICCAI)*. Cham: Springer International Publishing; 2018. pp. 265–73.
- [preprint] Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L et al. Segment anything. *arXiv preprint arXiv:2304.02643* (2023).
- Salto-Tellez M. More than a decade of molecular diagnostic cytopathology leading diagnostic and therapeutic decision-making. *Arch Pathol Lab Med*. 2018;142:443–5.
- Salek M, Li N, Chou HP, Saini K, Jovic A, Jacobs KB, et al. COSMOS: a platform for real-time morphology-based, label-free cell sorting using deep learning. *Commun Biol*. 2023;6:971.
- Pérez-Cano J, Valero IS, Anglada-Rotger D, Pina O, Salembier P, Marques F, et al. Combining graph neural networks and computer vision methods for cell nuclei classification in lung tissue. *Heliyon*. 2024;10(7):e28463.
- Zhou Y, Graham S, Koohbanani NA, Shaban M, Heng PA, Rajpoot N. CGC-Net: cell graph convolutional network for grading of colorectal cancer histology images. In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. Korea: IEEE; 2019. pp. 388–98.
- Pati P, Jaume G, Fernandes LA, Foncubierta-Rodríguez A, Feroce F, Anniciello AM, et al. HACT-Net: a hierarchical cell-to-tissue graph neural network for histopathological image classification. In: Sudre CH, Fehri H, Arbel T, Baumgartner CF, Dalca A, Tanno R, et al. editors. *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*. Cham: Springer International Publishing; 2020. pp. 208–19.
- Bora K, Chowdhury M, Mahanta LB, Kundu MK, Das AK. Automated classification of pap smear images to detect cervical dysplasia. *Comput Methods Programs Biomed*. 2017;138:31–47.
- Marinakos Y, Dounias G, Jantzen J. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Comput Biol Med*. 2009;39(1):69–78.

39. Chankong T, Theera-Umpon N, Auephanwiriyakul S. Automatic cervical cell segmentation and classification in pap smears. *Comput Methods Programs Biomed.* 2014;113(2):539–56.
40. van der Walt S, Schonberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, et al. Scikit-image: image processing in Python. *PeerJ.* 2014;2:e453.
41. Pati P, Jaume G, Foncubierta-Rodríguez A, Feroce F, Anniciello AM, Scognamiglio G, et al. Hierarchical graph representations in digital pathology. *Med Image Anal.* 2022;75:102264.
42. He X, Deng K, Wang X, Li Y, Zhang Y, Wang M. LightGCN: simplifying and powering graph convolution network for recommendation. In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'20)*. New York (NY): Association for Computing Machinery; 2020. P. 639–48.
43. Huang W, Hao F, Shang J, Yu W, Zeng S, Bisogni C, et al. Dual-LightGCN: dual light graph convolutional network for discriminative recommendation. *Comput Commun.* 2023;204:89–100.
44. Huang J, Xie R, Cao Q, Shen H, Zhang S, Xia F, et al. Negative can be positive: signed graph neural networks for recommendation. *Inf Process Manage.* 2023;60(4):103403.
45. Aric AH, Daniel AS, Pieter JS. Exploring network structure, dynamics, and function using NetworkX. In: Gael V, Travis V, Jarrod M, editors. *Proceedings of the 7th Python in Science*. 2008. pp. 11–5.
46. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–44.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.