

# Multi-level histomorphological features for mycosis fungoides diagnosis based on histological whole slide image of dermatology

Junjie Wang<sup>†</sup>, Congcong Xu<sup>†</sup>, Xiangxue Wang, Yiping Jiao, Cheng Lu, Yiqun Jiang\*, and Jun Xu\*

**Abstract—Background:** Mycosis fungoides (MF) is one kind of dermatology tumor in which lymphocytes become malignant and affect the skin. MF is easily confused with inflammatory diseases such as lichen planus in the early stage. Therefore, it is significant to develop automated diagnosis models to discriminate MF from benign or normal tissues. **Methods:** This paper presents a multi-level histomorphological features analysis method to differentiate MF from lichen planus automatically. It comprises three major parts: 1) Developing a fast and dense VGG (FD-VGG) for six tissues segmentation; 2) A multi-branch SENet (Multi-SENet) for cell classification; 3) An adaptive cell graph (ACG) algorithm to construct a cell graph for describing the spatial relationship of cells. First, a model ranking method is used to select the high score features of three levels, and then a correlation analysis is carried out to remove the redundant features. Finally, the preserved features of the three levels are fused and fed into the random forest classifier. **Findings:** An adaptive cell graph was developed to characterize the spatial relationship between lymphocyte and epithelial cells in epidermis tissue, which is able to describe the lymphatic infiltration of mycosis fungoides early-stage quantitatively. The diagnosis model with histomorphological features in three-level is 0.94 (AUC) in discriminating MF from lichen planus.

**Index Terms**—Computational pathology, Mycosis fungoides, Histomorphological features, FD-VGG, Multi-branch SENet, Adaptive cell graph.

## I. INTRODUCTION

Mycosis fungoides (MF) is a primary cutaneous T-cell lymphoma, which is originated from memory helper T cells [?]. It accounts for about 50% of all skin cancers, and the ratio of male to female is 4:5 [?]. Moreover, the average age of MF patients is from 55 to 60 years old, which is rare in children and adolescents. MF has a long development cycle and can be easily cured in the early stage [?]. However, MF will affect all internal organs once it is in the advanced stage [?]. Clinically, it is challenging to diagnose MF in the early stage because MF resembles inflammatory dermatoses, such as lichen planus, vitiligo and chronic cutaneous lupus erythematosus.

Histopathological image diagnosis is a gold standard for most diseases [?]. In the tissue-level, there are mainly five types of tissues in skin histological images: epidermis, dermis, hair follicle, sweat gland and fat [?]. For example, the histological images of mycosis fungoides show that most tumor cells are concentrated in the epidermis [?], and a few are hidden in the hair follicles [?]. The tumor can be located by identifying the epidermis and hair follicles tissues. Therefore, accurate segmentation of the epidermis and hair follicles tissues plays a vital role in the MF diagnosis. In the cellular level, as the tumor cells of MF are evolved from lymphocytes, the

J. Wang, X. Wang, Y. Jiao, and J. Xu are with the Institute for AI in Medicine, School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail:xujung@gmail.com).

C. Xu and Y. Jiang are with the Department of Dermatopathology, Hospital of Dermatology, Chinese Academy of Medical Science and Peking Union Medical College.

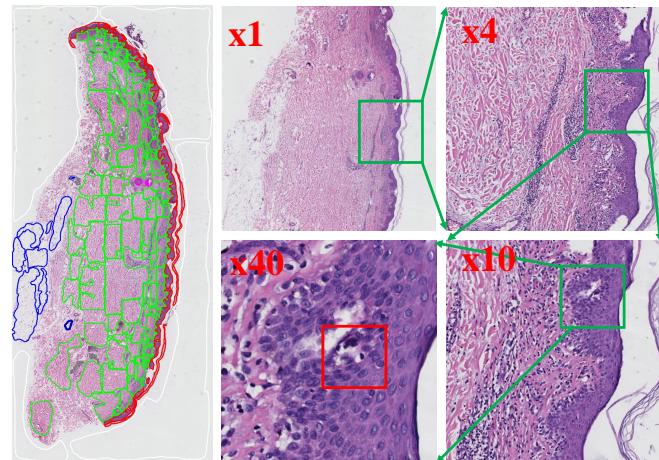
C. Lu is with the Dept. of Biomedical Engineering, Case Western Reserve University, OH44106, USA.

† are joint first authors.

\* are corresponding authors.

appearance of tumor cells is often accompanied by the abnormal proliferation of lymphocytes. Therefore, recognizing of lymphocytes is the key to MF diagnosis.

The development of mycosis fungoides includes the patch, plaque, and tumor stages [?]. In the patch stage (also known as the early stage), the cells' heteromorphism is not apparent [?], and tumor cells and lymphocytes can not be distinguished. Therefore, the morphological characteristics of cell level can not describe the abnormality of cells. Studies have shown that single or multiple aggregated lymphocytes often infiltrate into the epidermis (Fig.1) during this period, representing the change of cells' spatial position. This paper regards the spatial relationship between lymphocytes and epidermal cells as two types of points and attempts to construct a cell graph to describe the cell Spatial features.



**Fig. 1.** Histological images of mycosis fungoides at different magnification. The whole slide image of mycosis fungoides on the far left was annotated by pathologists where the annotated contours in different colors represent different types of regions. The image at  $\times 40$  shows lymphocytes infiltrating into the epidermis clearly in the red box.

There are some differences between MF images and inflammatory images. In clinical cases, the histological images of mycosis fungoides are often accompanied by mild spongy edema and sparse lymphocytes infiltration around the superficial dermis vessels. The features represent the global information of a histological image, which should be considered. However, this morphological change is similar to inflammatory skin diseases, such as lichen planus. They also show superficial zonal infiltration at the junction of the epidermis and dermis, which dramatically increases the doctors' burden for diagnosing mycosis fungoides. In summary, it is significant to leverage multiple-level features to achieve an early automated diagnosis of mycosis fungoides.

## II. PREVIOUS WORKS AND NOVEL CONTRIBUTIONS

The histological interpretation by pathologists is an effective method for mycosis fungoides diagnosis. However, this process is

tedious and time-consuming. In recent years, computational pathology has achieved great success in histological image analysis [?]. [?] trained a neural network for breast cancer image segmentation, which shows the successful application of deep learning in breast pathology. [?] applied VGG16 in brain image classification. [?] considered the feature learning among different tissues and got the segmentation result of tumor region and stroma. With the ongoing in-depth study of histological images, the large-scale problem of histological images was considered. [?] leveraged deep convolutional neural network (CNN) activation features to finish histopathology image segmentation. [?] used a U-Net to develop an automated diagnosis system for masses detected from entire mammograms of the breast.

At the same time, some scholars began to focus on cell classification studies. [?] performed nuclear segmentation and used AdaBoost classifier to classify each nucleus based on the strength, shape, texture and other nucleus features. [?] proposed a stacked sparse automatic encoder for cell detection in breast cancer dataset. However, this work does not distinguish the types of cells. With the development of multi-task network structure, [?] developed a cell classification network based on multi branches. The proposed horizontal and vertical maps innovatively solved cell adhesion and successfully applied it to the 7 cell classification problem in the colorectal dataset. Simultaneously, transfer learning began to emerge in this field. [?] successfully employed a multi-classification of OSCC by transfer learning. In recent years, many studies committed to understanding disease from different levels of pathology. [?] segmented tissue of cervical cancer slides, extracted global features and achieved automatic diagnosis of cervical cancer, which proves the effectiveness of tissue-level features for the diagnosis task.

In some studies, cell-level features, especially cell graph features, have been proved to have a good expression effect for slide level classification. For example, [?] predicted the relationship among low, medium and high levels of detection risk of ductal carcinoma by quantitative analysis of nuclear tissue morphology, including texture, topology, morphology and other nuclear level features. [?] presented a new CGCNet, which used the spatial relationship of single type cells to transform large tissue images into a graph for grading colorectal cancer. [?] proved the existence of "immune escape" in lung adenocarcinoma by using the change of cell location. In the same year, [?] used the spatial topological structure of cells to establish a spatial graph model to analyze skin melanoma histology. [?] applied the Delaunay triangulation to construct the cell graph, and then classified various tissues in colorectal data. [?] presented an innovative feature-driven method to construct cell graphs and successfully applied them to lung cancer and oropharyngeal cancer data sets. The above research revealed the effectiveness of cell graphs in histopathology, but they did not consider the spatial relationship between different types of cells. There is no research to use the spatial description between types of cells as the basis of cell graphs.

Inspired by these works, multi-level histomorphological features were presented for quantifying the differences between mycosis fungoïdes (MF) and lichen planus. The features include multi-level morphology and cell graph. A fast, dense VGG (FD-VGG) model was first presented for automated segmentation of 6 tissues, including epidermis, dermis, fat, hair follicle, sweat gland, and background, from whole slide image (WSI) of dermatology. Subsequently, a multi-branches SENet (Multi-SENet) model was developed to classify lymphocyte and epithelial cells from epidermis and dermis regions. An adaptive cell graph (ACG) algorithm constructs the cell graph features based on the positional relationship between lymphocyte and epithelial cells. Then, the tissue- and tile- features at the junction of epidermis and dermis from WSI of dermatology according to tissue

segmentation results by FD-VGG model. Finally, features in tissue-, tile- and cell- levels are fused and fed to a random forest classifier for discriminating MF from lichen planus.

The novel contributions of this study are

- 1) A fast and dense VGG (FD-VGG) network was developed for 6 types of tissue segmentation from the whole slide image of dermatology. The FD-VGG network used the convolution without padding so that the slide has no boundary effect in prediction. Also, a  $1 \times 1$  convolution instead of the full connection layer was used to ensure that the network is scalable to input images with any scales.
- 2) A Multi-SENet model was developed to discriminate lymphocyte and epithelial cells from epidermis and dermis tissues.
- 3) An adaptive cell graph (ACG) was constructed to characterize the spatial relationship between lymphocyte and epithelial cells in epidermis tissue, aiming to describe the lymphatic infiltration of mycosis fungoïdes early-stage quantitatively.
- 4) The mycosis fungoïdes diagnosis (MFD) system was developed based on the histomorphological features in tissue-, tile-, and cell- levels extracted from the whole slide image of dermatology.

### III. METHOD

In this paper, we use "model" to represent the segmentation or feature extraction algorithms while "module" is an independent unit used to construct the diagnosis system. The Mycosis Fungoïdes Diagnosis (MFD) system has two modules (see Fig.2):

- Module I: Computation image analysis module which comprises: I(a) Generation of training samples, I(b) Tissue segmentation model in training phase, I(c) Cell classification model in training phase, I(d) Histomorphological feature extraction models;
- Module II: Mycosis Fungoïdes diagnosis module which comprises: II(a) Image analysis on whole slide images, II(b) Feature extraction, II(c) Diagnosis.

#### A. Module I: Computational image analysis

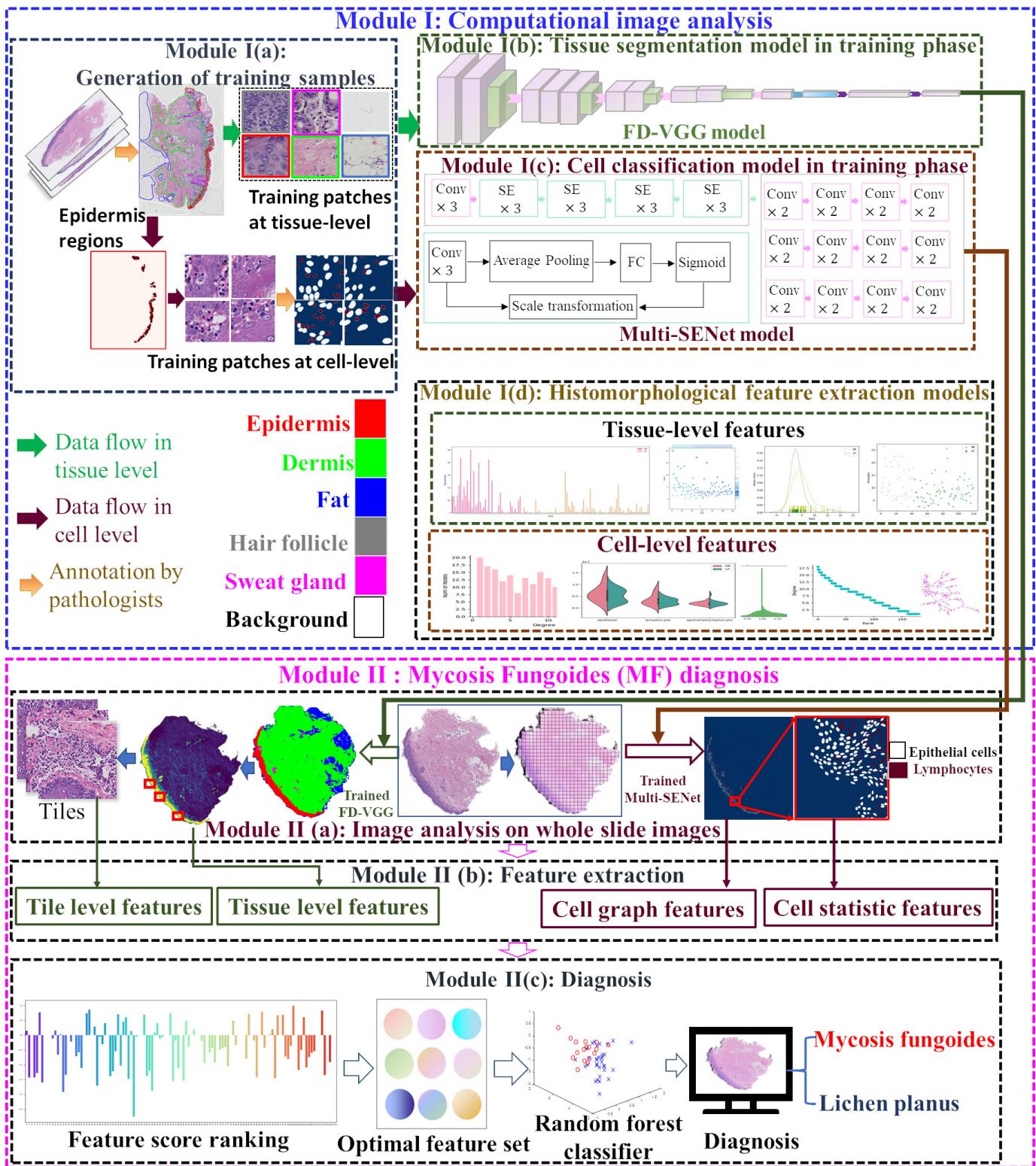
As Fig.2 shows, Module I comprises 4 components: I(a) Generation of training samples from WSI of dermatology in the training set, I(b) The training phase of FD-VGG model for tissue segmentation (see Fig.3), I(c) The training phase of Multi-SENet model for identifying the epithelial cells and lymphocytes based on the tissue segmentation results (see Fig.4), I(d) Histomorphological feature extraction models at the tissue- and cell-levels based on tissue and cell segmentation and classification results (see Fig.5).

#### B. Module I(a): Generation of training samples

As the WSIs are too large to be used for training, Module I(a) aims to generate training patches at the tissue- and cell levels from the WSIs in the training set. Firstly, 6 types of tissues were annotated by pathologists. Then the training patches at the tissue level were then generated according to the annotation results. Secondly, new patches were manually selected from epidermis regions. Pathologists subsequently annotated the lymphocytes and epithelial cells to generate training patches at the cell level.

**1) Module I(b): Tissue segmentation using FD-VGG:** The architecture of the FD-VGG model is shown in Fig.3. Compared with VGG [?], the FD-VGG model is optimized and adjusted as follows:

- 1) The last pooling layer uses average pooling instead of maximum pooling;
- 2) All convolution operations use convolution without padding;



**Fig. 2.** The flowchart of Mycosis Fungoides Diagnosis (MFD) system. It consists of module I (computational image analysis) and module II(diagnosis phase). Module I(a) is the dataset construction. Module I(b) is used for tissue segmentation, the module I(c) is applied to the cell classification. Module II(b) shows the tissue segmentation and cell classification testing phase. Module II(c) is the feature extraction phase which needs the extraction model in module I(c).

- 3) The  $1 \times 1$  convolution layer is used instead of the fully connected layer;
- 4) The batch normalization is used between each convolution layer and the active layer.

FD-VGG inherits the advantages of the VGG network and avoids some disadvantages, where items 1) and 3) ensure that the network can accept images of any size. Item 2) avoids the edge effect, which means these small pieces divided from large-scale histopathology will have an obvious fault at the edge when stitching. In addition, after each convolution layer, batch normalization in item 4) is used to normalize the convoluted feature maps, which can reduce the data jitter between different batches during training. The input image is processed by a convolution operator composed of  $3 \times 3$  convolution without padding, batch normalization, ReLu and max pooling. Then two convolutional operators are leveraged to get the final classification. In each convolutional operator, we gradually increase the number of feature maps from 64 to 512. Dense feature maps can retain more input images, increasing to 1024 after average pooling. The cross-entropy was chosen as the loss function of the network, which is defined as [?]:

$$\text{Loss} = - \sum_{i=1}^{n=6} c_i \cdot \log(p_i) \quad (1)$$

where  $c_i$  is the label category corresponding to the  $i - th$  output neuron and  $p_i$  is the probability value of the  $i - th$  output neuron after softmax computation.

**2) Module I(c): Cell classification using Multi-SENet:** The overall framework of the network is shown in Fig.4. Multi-SENet is divided into encoding (Fig.4(g)) and decoding stages(Fig.4(h)). In the encoder stage, the network shares an encoder. The patches input Multi-SENet to get high-level features through the encoder and then decode the features through the decoder. At the same time, there are three branches in the decoding phase: the main branch(Main), the auxiliary branch I(B(I)) and auxiliary branch II(B(II)). The main branch outputs the category information, so the label of cell classification constrains it (Fig.4(d)); the auxiliary branch I output the binary detection branch, which is constrained by the label of cell and background(Fig.4(c)); The auxiliary branch II outputs the horizontal and vertical map(Fig.4(e-f)) branch [?]. This paper combines it with the classification task of the main branch and achieves excellent results. This multi-task network learning mode avoids the waste of resources caused by the training of Multi separate tasks and enables the encoder in our model to learn more features. Moreover, the network can better promote the main task of learning. The network refers to the SENet module [?], and its attention mechanism can assign a weight to each feature map, which makes the image information retained by the valuable network for the task. The label of auxiliary task II is the horizontal-vertical map, which is based on the information of cell gradient, which helps the network recognize the glued cells with the change of gradient to identify the boundaries of cells. Since the network has three decoder branches, the setting of the loss function must be related to the training effect of each branch. In other words, the loss function should make each branch converge correctly. Therefore, the loss function is set as the weighted sum of the sub loss functions of each branch

$$L_{loss} = \lambda_1 L_1 + \lambda_2 L_2 + \lambda_3 L_3 \quad (2)$$

where  $L_1, L_2, L_3$  represent the loss function of auxiliary 1, auxiliary 2 and main task. As  $L_3$  is the main loss,  $\lambda_3$  is set 1. and  $\lambda_1, \lambda_2$  are set as 0.5 and 0.5.

$$L_1 = \text{Loss}_{\text{Dice}} + \text{Loss}_{\text{MSE}} \quad (3)$$

The choice of composite loss function prevents the Dicedice from not converging, which is defined as

$$\text{Loss}_{\text{Dice}} = 1 - \frac{2 |X \cap Y|}{|X| + |Y|} = 1 - \frac{2 \times \sum_{i=1}^N (Y_i(I) \times X_i(I)) + \epsilon}{\sum_{i=1}^N (Y_i(I)) + \sum_{i=1}^N (X_i(I)) + \epsilon} \quad (4)$$

where  $X, Y$ , and  $N$  represent the ground truth, the prediction, the pixel summation of the predicted image, respectively.  $\epsilon$  is a minimal value that prevents the denominator from 0. Mean square error(MSE) loss function is defined as

$$L_{MSE} = \text{loss}(\hat{y}, y) = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y)^2 \quad (5)$$

$L_1$  is the same to  $L_{MSE}$ . The main task is set as the cross entropycrossentro loss function with the superior training effect

$$L_3 = L_{CE} = - \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K X_{i,k}(I) \log Y_{i,k}(I) \quad (6)$$

where  $K$  is the number of categories, and  $I$  is the input image.

**3) Module I(c): Feature extraction models:** The feature extraction models are shown in Fig.5. Four types of features are extracted. Figs.5(a,b) are the tissue-level and tile- level features, respectively. Figs.5(c,d) are cell-level features. Fig.5(a) computes not only the tumor region features, including count, area, lengthetc. but also the traditional image features of probability maps. Tile level features have three parts: gray-level co-occurrence matrix(GLCM), local binary pattern(LBP), and color histogram, and it is based on tiles obtained from epidermis regions. Cell statistic features include the area of lymphocytes and epithelial cells. As pathologists pay close attention to lymphocytes and epithelial cells, the statistical features can also be regarded as the clinical features of mycosis fungicides diagnosis. Cell graph features are based on a graph, which shows all graph features.

### C. Module II: Diagnosis phase

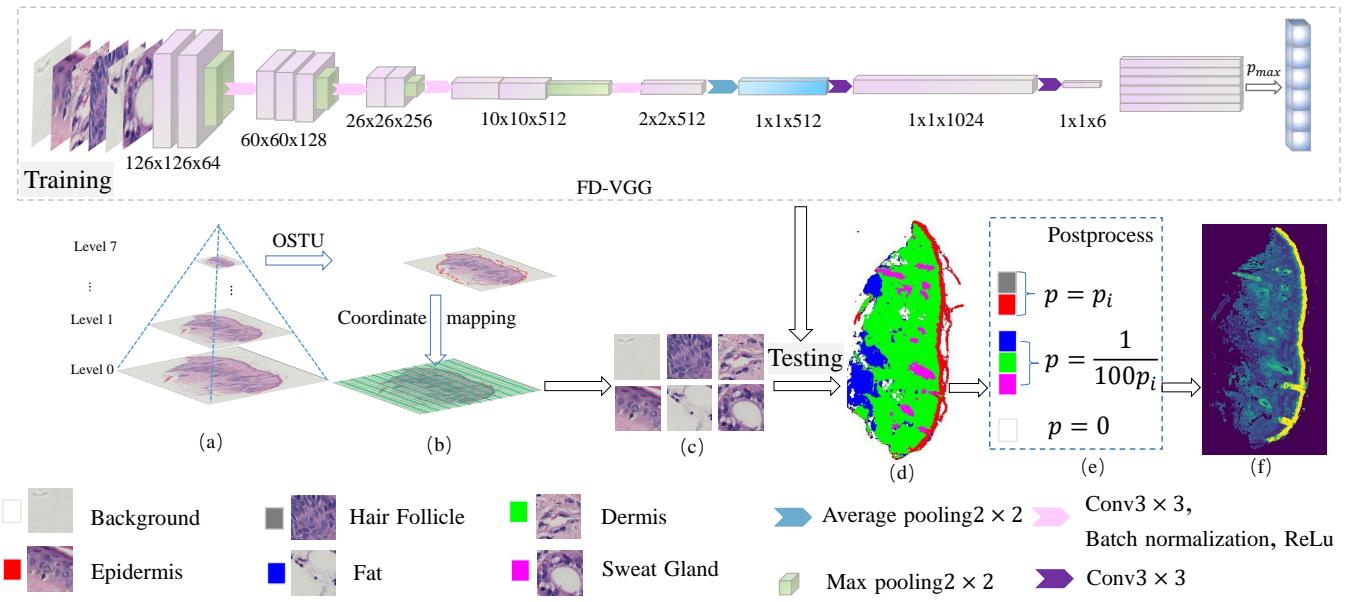
The whole diagnosis process is divided into three phases: testing phase, feature extraction phase and final diagnosis phase(Fig.2). The testing phase includes two parts: tissue segmentation testing and cell classification testing; in order to analyze MF whole slides from multiple levels, the feature extraction phase includes four parts(Fig.5): (a)Tissue level, (b)Tile level, (c)cell statistics (d)cell graph. Finally, the final diagnosis phase includes feature selection, feature correlation analysis, random forest classifier model [?].

**1) Module II(a):Image analysis on whole slide images:** The testing phase comprises 6 tissue segmentation (see Fig.3(a-f)) and cell classification (see Fig.4(h-k)). FD-VGG model was used to segment 6 tissues, and then postprocessing operation(e) to get the probability map of slides. This result will be used in the feature extraction phase.

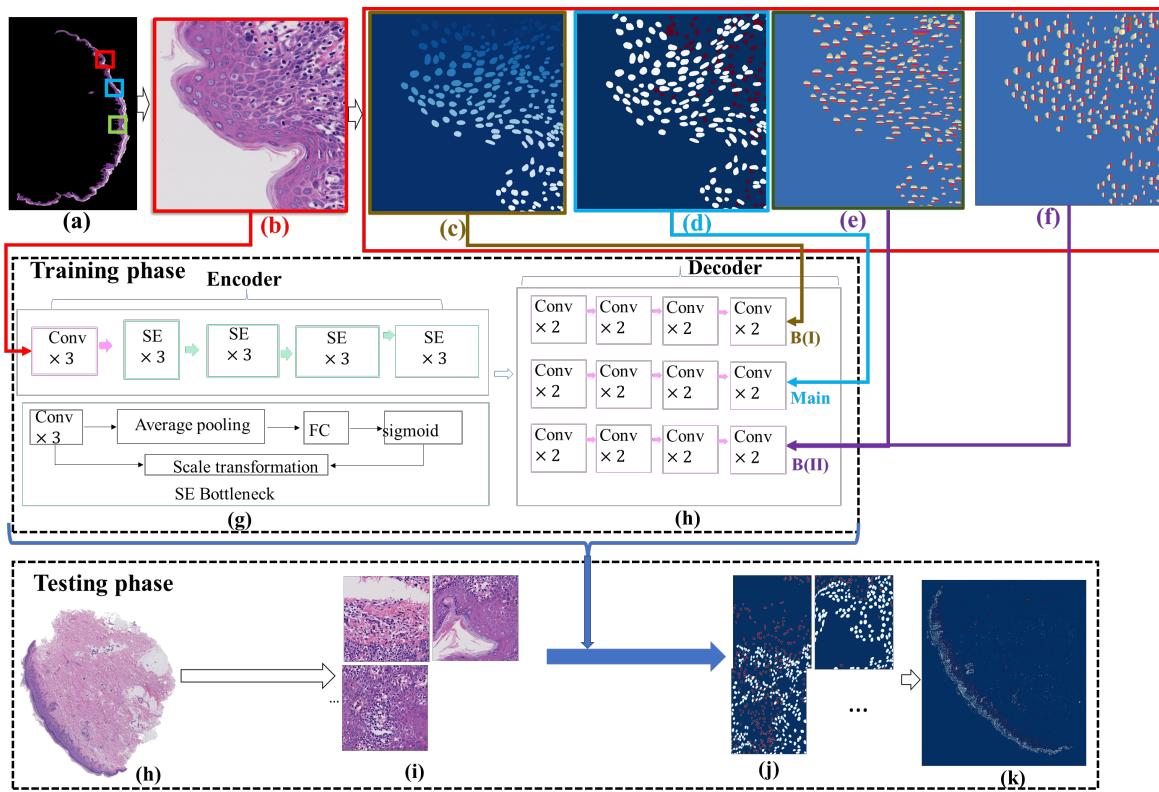
**2) Module II(b): Feature extraction-tissue level :** Tissue level features extraction is based on probability map(Fig.3(e)). In this paper, the tissue level features is shown in Fig.5(a).

**3) Module II(b):Feature extraction-tile level:** Tile level extraction extracts in Fig.5(b). Compared with tissue features, It pays more attention to the characteristics of local tissue. Therefore, it need to take epidermis tiles(Fig.2(Module II(a))).

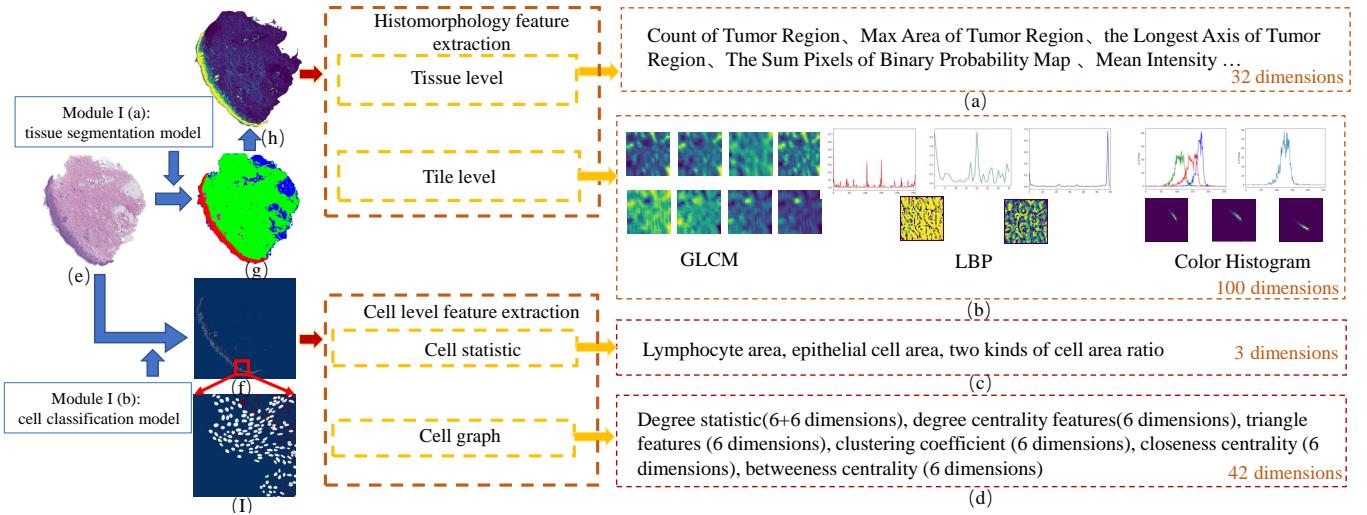
**4) Module II(b): Feature extraction-cell statistic:** The specific content of cell statistics is shown in Fig.5(c). The statistical characteristics of cells are the area of lymphocytes and epithelial cells and their ratio(Fig.1 in the appendix). Cell statistics are very significant because the abnormal growth of lymphocytes is one of the manifestations of early mycosis fungoides.



**Fig. 3.** Diagram describes the process of tissue segmentation for the WSI of dermatology with the FD-VGG model. The process comprises training and segmentation phases of 6 tissues segmentation (a-f).



**Fig. 4.** The flowchart of cell classification with Multi-SENet model. The model comprises a training phase and two types of cell classification from the WSI of dermatology.



**Fig. 5.** The diagram of feature extraction models. (a-d) are the four types of features. (g) is the segmentation result of (e) by Fig.2(module I(a)), (h) is the probability map of (g). (f) is the classification result of lymphocytes and epithelial cells by Fig.2(module I(b)).

**5) Module II(b): Feature extraction-cell graph:** In the clinic, the appearance of tumor cells in mycosis fungoides is often accompanied by the phenomenon of immersing into epidermis and follicle (Fig.6). However, tumor cells are differentiated from lymphocytes, so it is difficult to identify them with naked eyes. Also, it is difficult to describe this phenomenon with traditional texture, color, edge, and nuclear features because it mainly shows the changes in cells' spatial position rather than some changes in the cells themselves. In this paper, a cell graph construction method is proposed based on spatial relationship: adaptive cell graph(ACG) algorithm, which can calculate the adaptive distance to describe these two clinical phenomena (Fig.6) from the perspective of spatial relationship. For the extraction of graph features, the cell classification results are preprocessed to get the center point of each cell. Then, the ACG algorithm is used to build a graph based on the center point. The specific steps are

- 1) Taking lymphocytes as the center, the adaptive distance relationship between lymphocytes and epithelial cells is calculated by the formula (7).
- 2) Taking lymphocytes as the center, the adaptive distance between lymphocytes and adjacent lymphocytes is calculated. For the first step, the formula (7) is also used for construction, and the result can describe the spatial location between the lymphocyte and its adjacent lymphocyte.

$$L = \gamma \sqrt{(o_{m1} - p_{n1})^2 + (o_{m2} - p_{n2})^2} \quad (7)$$

As the whole slide is enormous, and the actual pixel distance between cells is sometimes long, we add the adaptive factor  $\gamma$  to convert the actual distance between cells into the relative distance, which calls the adaptive distance.

$$\gamma = \frac{1}{(||o_{m1} - p_{n1}|| + ||o_{m2} - p_{n2}||)} \quad (8)$$

where  $(o_{m1}, o_{m2})$ ,  $(p_{n1}, p_{n2})$  are the point of cell centers. The proposed adaptive factor has two advantages:

- 1) Increase the computational speed, as the value of  $L$  is controlled in a certain range.
- 2) Avoid the  $L$  is too large or too small, which affects the construction and expression accuracy of cell graph.

Since most of the epithelial cells are in the epidermis, ACG algorithm distance can ideally quantify the extent of lymphocyte infiltration into

the epidermis. Therefore, it is of great significance for the diagnosis of mycosis fungoides. Furthermore, in the clinic, there is a specific relationship between adjacent tumor cells [?]. Therefore, we also calculate the adaptive distance among lymphocytes. Moreover, the cell graph constructed by the ACG algorithm allows the constructed graph to reflect the clinical features. Therefore, the features of the cell graph can be well interpreted.

It is necessary to get edges and nodes when a graph is constructed. This paper regards each cell center as a node and the corresponding adaptive distance as an edge, and then we can get the cell graph. The flowchart is shown in Fig.7.

**6) Module II(b): Diagnosis:** The diagnosis phase(Fig.8) includes four parts: a) random forest model ranking, b) correlation analysis, c) removing redundant features, d) random forest classifier. The basic process are

- 1) Three kinds of features(Fig.8(①-③)) are sent to random forest model based ranking. It will give each feature a score and rank them(Fig.2(module 2(c))).
- 2) Choose the top 12 features in 1), then the correlation among them is analyzed in pairs using the pearson correlation coefficient [?].
- 3) For feature pairs with  $P > 0.05$ , only one of the two is selected.
- 4) A random forest model is trained by features retained from 3) and Fig.8(④)

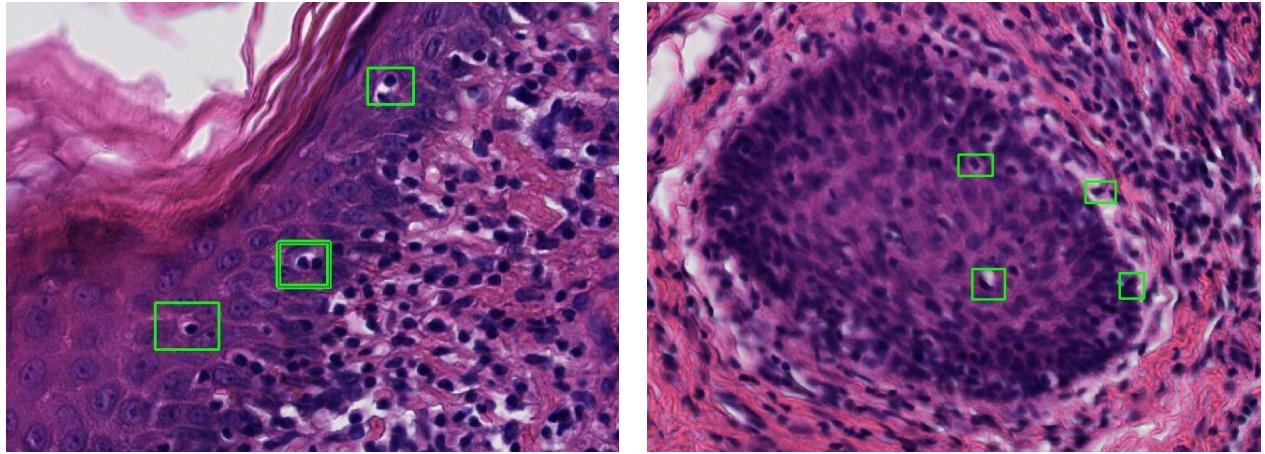
## IV. EXPERIMENT AND DISCUSSION

### A. Dataset description

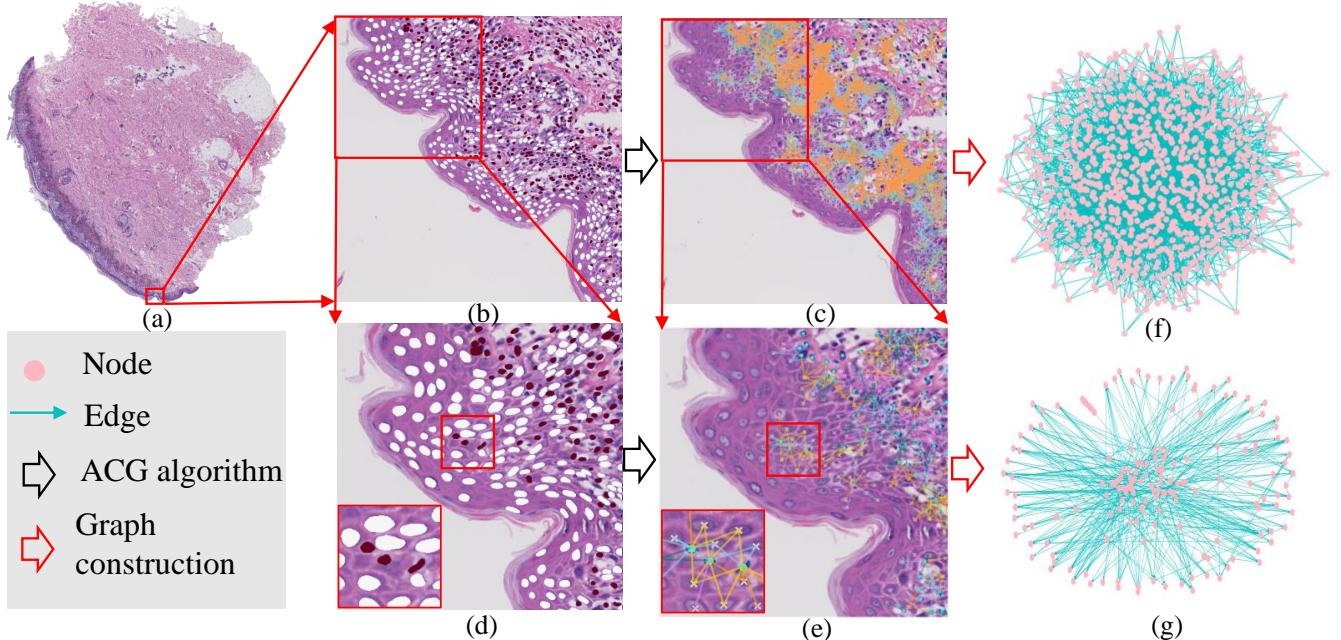
The data used in this paper were collected from the Department of Dermatopathology, Hospital of Dermatology, Chinese Academy of Medical Science and Peking Union Medical College. All slides were stained with Hematoxylin-Eosin (HE) staining protocol. Images were scanned with a NanoZoomer2.0-RS at 40 $\times$  magnification with a resolution of 50 $\mu\text{m}$  per pixel and annotated using ImageScope software. The compressed storage space was approximately 3.9 GB. A total of 119 slides were used in this study, which contained 42 mycosis fungoides and 77 lichen planus slides.

### B. Manual annotation

The 6 regions in WSIs were manually annotated by pathologists CX and YJ who have been working on mycosis fungoides diagnosis



**Fig. 6.** The example of pro epidermis and pro follicle of mycosis fungoides regions. The green boxes marked the tumor cells in epidermis and hair follicles regions.



**Fig. 7.** The flowchart of cell graph construction. The cell classification result is shown in (a), which got from the Multi-SENet model(Fig.4). (b), (d) are the zoomed local images. (c), (e) are the adaptive distance visualization, (f), (g) are the cell graph constructed by edges and nodes which are based on (c) and (e). ACG is the abbreviation of adaptive cell graph.

for many years. A total of 14 WSIs of mycosis fungoides were annotated. The 6 type of tissues are epidermis, dermis, fat, hair follicle, sweat gland and background and each tissue was annotated with a false color as shown in (Fig.3).

### C. Dataset for tissue segmentation and cell classification

For all 14 WSIs, we used 12 for training and 2 for validation. The number of patches taken from each tissue is shown in Fig.9.

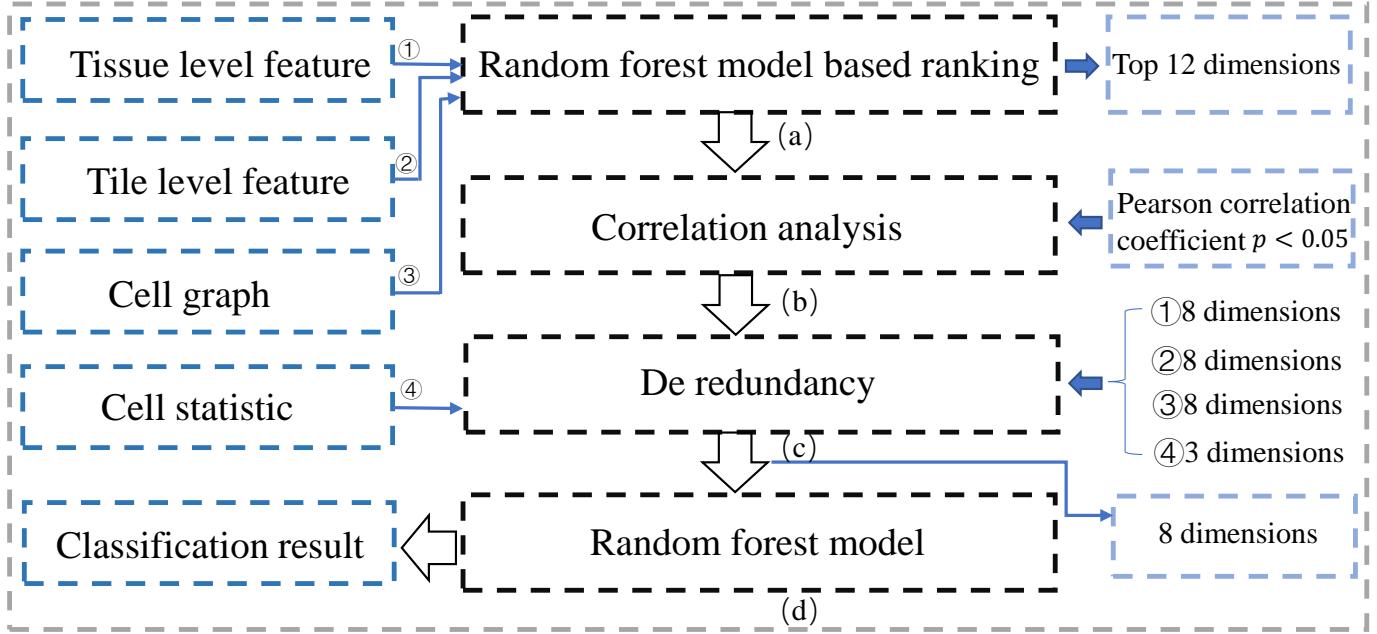
The 9 image tiles( $H = W = 2048$ ) were taken from the epidermis of mycosis fungoides and labeled by a professional pathologist. We divided these tiles into  $512 \times 512$ . There are 144 patches, 114 for training, 30 for validation.

### D. Experimental design and comparison strategy

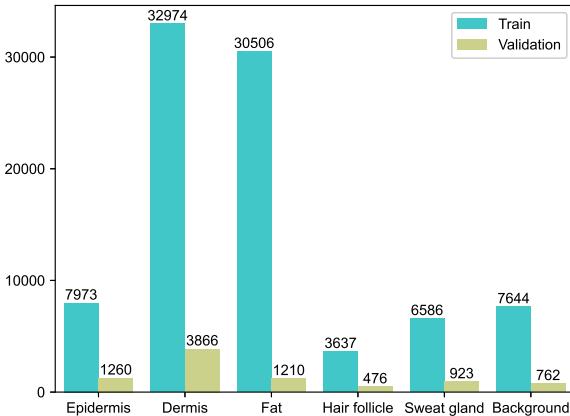
Three experiments were designed to show the effectiveness of multi-level histomorphology for the automated diagnosis of mycosis fungoides(MF).

**1) Experiment 1: multi-level histomorphology for tissue segmentation of mycosis fungoides(Module I(b)):** This experiment aims to show the effectiveness of tissue segmentation on mycosis fungoides. The FD-VGG model includes training and testing phases.

*a) Training phase:* In the training phase, as is shown in Fig.3, FD-VGG was used to segment the tissue region, and the segmentation result was post-processed to get the probability map of the tumor region. The training data set is shown in Fig.9. Each patch is  $144 \times 144 \times 3$ , and the validation set is the same shape. The data expansion used flip, rotation, cropping, color dithering. The learning rate was set to 0.01, momentum was set to 0.9, and the epoch was



**Fig. 8.** Diagnosis phase. ①, ②, ③, ④ are four types features. (a), (b), (c), (d) are the four diagnosis steps. Top 12 dimensions features are selected in (a) from tissue level, tile level, cell graph. (b) is the correlation analysis using pearson correlation coefficient. Features are retained if  $p > 0.05$ . Cell statistic features and the retained features are sent to random forest model(d).

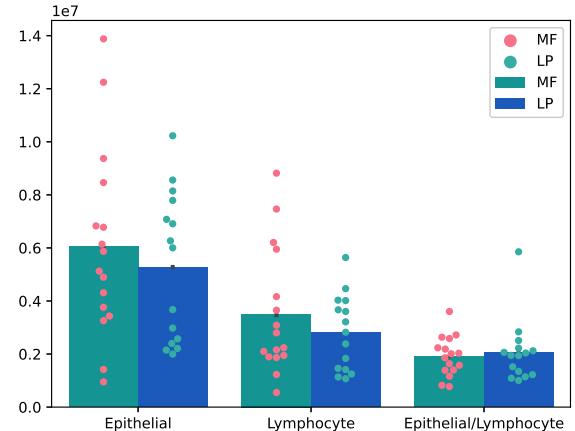


**Fig. 9.** The tissue segmentation dataset. It counts the number of patches in training phase and verification phase, and the size of patches is  $224 \times 224 \times 3$ .

100.

b) *Testing phase:* In testing phase(Figs.3(a-f)), the redundant background of whole slides was removed by OSTU [?] algorithm. Then the overlapping patches were selected according to a certain stride(we set stride to 36), which can weaken the obvious edge (edge effect) that appeared in the splicing process. The patches were sent into the trained model(FD-VGG) to get the probabilities that belong to each kind of tissue, and the probability map of mycosis fungoides tumor region was obtained by postprocessing(Fig.3(e)).

At the same time, we compared several classic network structures(Inception V3 [?], Resnet101 [?] and Res2net [?]) to prove the effectiveness of our model. Since the probability map(Fig.3(f)) is the basis of tissue-level features extraction and features drive the task of automatic diagnosis, our performance evaluation is quantified by the



**Fig. 10.** The diagram illustrates the number of Epithelial and Lymphocyte, and the ratio of Epithelial and Lymphocyte in MF and LP samples, respectively, where each dot represents a whole slide of tissue sample.

contribution in the diagnosis task. In other words, we do not take the segmentation accuracy of each tissue type as our evaluation standard. However, judge the effectiveness by the probability map of the tumor region. The result is shown in Fig.11(a).

2) *Experiment 2: multi-level histomorphology for cell classification(Module I(c)):* This experiment aims to show the effectiveness on the classification of lymphocyte and epithelial cells in mycosis fungoides. The Multi-SENet model is illustrated in Fig.4.

a) *Training phase:* As Fig.4(a-g) shows, Multi-SENet is used for cell classification task. The summary of datasets studied in this paper is shown in Table I.

The images in the training set are  $512 \times 512$  pixels, and they are randomly cut into  $352 \times 352$  pixels before being sent into Multi-SENet. Then data argument includes fuzzy, hue saturation and

**TABLE I**  
SUMMARY OF DATASETS STUDIED IN THIS PAPER.

2*Datasets	2*Purpose	2*Models	Training and testing		Annotations see Fig.2	2*Modules
			Training	Testing		
D1	Segmenting six types of skin tissues	FD-VGG	89293(Fig.9 for details)	8947	Module I(a))	Module I(b)
D2	Developing a cell classification model	Multi-SENet	114(512×512)	30	Module I(a)	Module I(c)

brightness change, brightness contrast change and gamma change. Therefore, Multi-SENet (Fig.4) can learn more complex information and avoid overfitting. Due to the data sample is not enough, the data expansion mode of each epoch in the training process is different, and the epoch is set to 1000. Finally, the encoder part is initialized using the SENet154, pre-trained on the ImageNet data. The learning rate is also set to 0.01, and the momentum is 0.9.

b) *Testing phase*: In the testing phase(Fig.4(h-k)), the whole slides of mycosis fungoides were also cut into  $512 \times 512$  patches after the background was removed by OSTU [?]. Then patches were sent to the trained Multi-SENet. The final cell classification result (k) is obtained by splicing (f). Due to the large size of the whole slides, the construction of the cell graph is a highly computational process. Therefore, it is inefficient to evaluate the model through comparative experiments. Therefore, this paper evaluates classification accuracy through four F1 scores, accuracy, recall and accuracy. Furthermore, as our model is a multi-task architecture, we designed three ablation experiments to illustrate the effectiveness of adding auxiliary branches inspired by [?].

3) *Experiment 3: multi-level histomorphology for diagnosis of dermatology whole slides(Module II(c))*: This experiment aims to evaluate the diagnosis effectiveness of mycosis fungoides (MF) and lichen planus(Fig.8)). We set six features ablation experiments to illustrate the effectiveness of multi-level fusion histomorphology features for diagnosis.

We first perform tissue segmentation and cell classification on mycosis fungoides whole slides(Fig.2module II (a)) and then extract the features (module II (b)) of tissue level, tile level, cell statistics and cell graph features based on these obtained results. The feature dimensions are 32, 100, 3 and 42(Fig.5). The top 12 features with the highest score were selected by the random forest model ranking method, and then correlation analysis and redundancy removal were carried out. The retained features are shown in Fig.10, Fig.12 and Fig.13. The final diagnosis accuracy is 0.94(AUC)(Fig.11(b)). The ablation experiments are shown in Fig.11(a) and Fig.11(c).

## V. EXPERIMENTAL RESULT AND DISCUSSION

### A. The segmentation accuracy of six types tissues of mycosis fungoides(Module I(b))

The results of three comparative models and FD-VGG were shown in the appendix(Figs.2). The prediction results of mycosis fungoides tumor area in three different networks is shown in b-c. (a) is the original whole slide labeled by pathologist. (b), (c), (d) illustrated the segmentation results of Inception V3, Resnet101, Res2net, and FD-VGG, respectively. The red and gray areas in (a) were tumor areas. In the prediction, the highlighted area is the predicted tumor area. From these results, the four networks can recognize the correct tumor area, but the three networks have a large misjudgment for the dermis area. In the green marked area, Inception V3, Resnet101 and Res2net all regard a large area as the tumor area, but FD-VGG correctly classifies it as normal tissue. The classification result is shown in Fig.11. The FD-VGG model proposed by us is the best, about 0.83. It shows the effectiveness of features based on our model.

### B. The classification accuracy of Multi-SENet(Module I(c))

The results of cell classification are significant for building cell graphs. They can directly affect the accuracy of feature extraction. Multi-SENet performs well in the classification task of epithelial cells and lymphocytes. As is shown in appendix(Figs.3), the yellow rectangular frame in (a) shows that the doctor has not labeled, and the network does not take the cells in a specific type. It indicates that the network learns the characteristics of distinguishing epithelial cells from lymphocytes and the characteristics of distinguishing two types of cells from other types of cells. However, our network can solve this problem well. At the same time, the third column (c) of the ablation experiment can not solve this problem when only has the auxiliary branch I, so it proves the effectiveness of adding the auxiliary branch II. The last row shows that for the already connected cells, the network also has a certain degree of stickiness when classifying (as is shown in f), so we can only regard it as an everyday center when selecting the center of the cells.

The prediction results of each group were evaluated accordingly by the accuracy, F1 score, IOU, performance, accuracy and recall rate(Table II). The first column is the result of removing auxiliary branches; the second column is the evaluation of removing auxiliary branch I; the third column is the result of removing auxiliary branch II; the last column is the evaluation result of the original model. The evaluation results were based on the same mycosis fungoides test dataset.

**TABLE II**  
EXPERIMENTAL EVALUATION RESULTS.

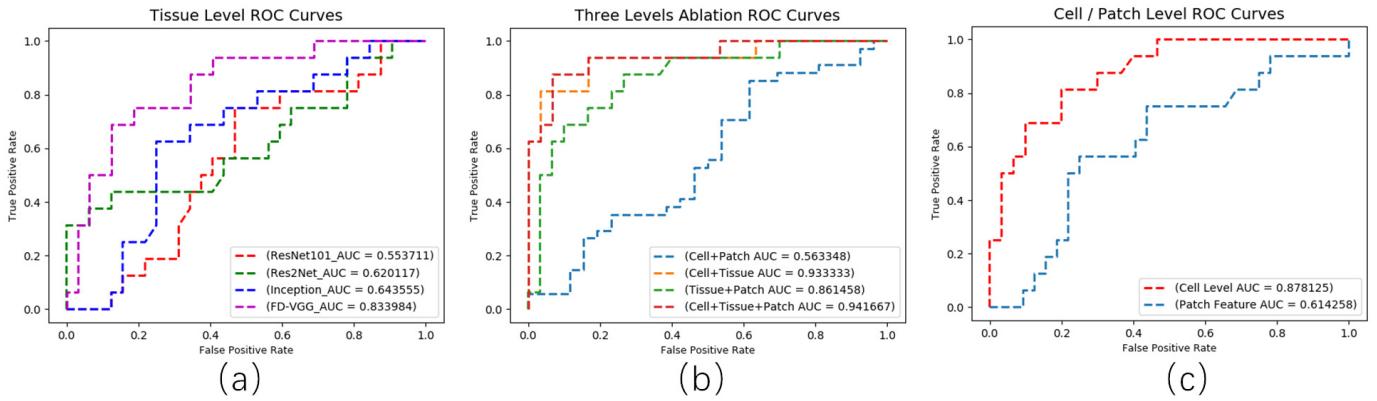
Evaluation	Main	Main+II	Main+I	Ours
Accuracy	0.940	0.931	0.940	<b>0.943</b>
F1	0.700	0.708	0.727	<b>0.738</b>
IoU	0.570	0.548	0.571	<b>0.580</b>
Precision	0.720	0.701	<b>0.727</b>	0.711
Recall	0.730	0.716	0.727	<b>0.760</b>

**TABLE III**  
ABLATION EXPERIMENTS

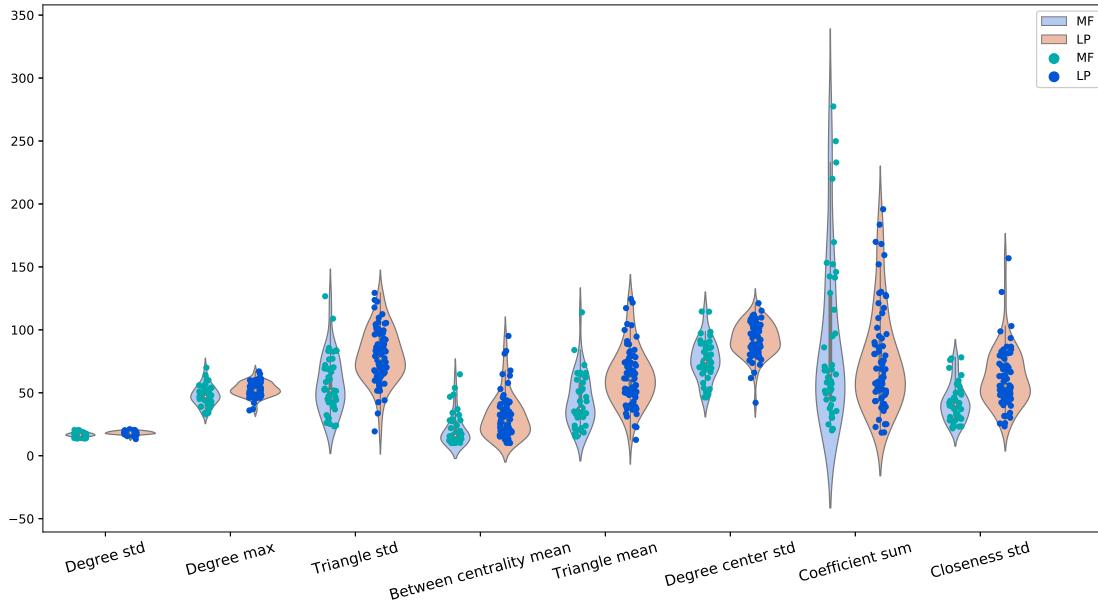
	Tissue Level	Tile Level	Cell Level	Ablation
1	✓			✓
2		✓		✓
3			✓	✓
4	✓	✓		✓
5		✓	✓	✓
6	✓		✓	✓

### C. Diagnosis accuracy assessment(Module II(c))

To demonstrate the effectiveness of multi-level morphology for automatic diagnosis, the diagnosis result is shown in Fig.11(a). The fusion features of cell level, tile level, tissue level get the best diagnosis result(AUC=0.942). The ablation experiments show the effect on single type features in diagnosis task(Fig.11(b), Fig.11(c)). Fig.11(a) shows the classification result based on tissue level features(FD-VGG) are 0.83. Cell level features(Fig.11(c)) are 0.88 and patch(tile) level are 0.61.



**Fig. 11.** ROC curves of (a) experiment 1 by Resnet101, Res2net, inception v3 and FD-VGG; (b) three levels ablation experiments and (c) cell and tissue levels ablation experiments.



**Fig. 12.** Top 8 cell graph level features. It mainly includes degree, triangle, betweenness centrality, coefficient and closeness centrality.

#### D. Features analysis

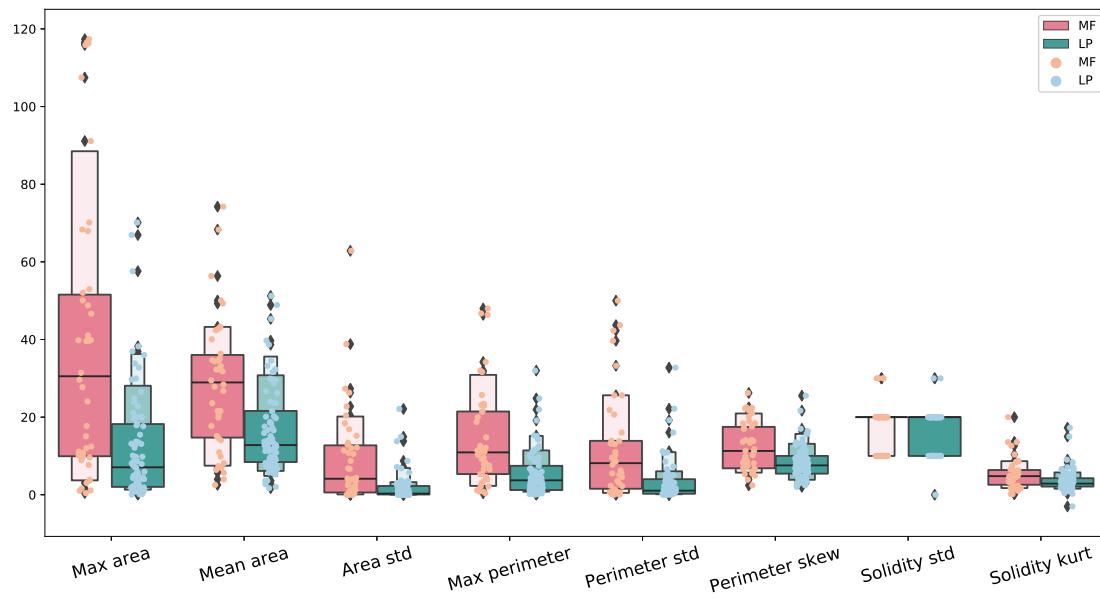
The features selected by tile-level have little contribution to the diagnosis results compared with the tissue and cell levels (AUC=0.61). It does not mean that tile features extracted at the junction of epidermis and dermis of mycosis fungoides are unnecessary, but our feature selection strategy should be improved. In addition, the texture features and color features(Fig.5(b)) we selected may not be able to describe lymphatic infiltration well because these traditional features pay too much attention to the edges, shape rather than spatial relationship. It is an important reason it cannot distinguish mycosis fungoides from lichen planus.

The best ablation result was obtained by cell-level features(AUC=0.878). On the one hand, the graph constructed by the adaptive cell graph(ACG) algorithm can describe the spatial relationship extensively. On the other hand, the features extracted by the cell graph prove to have a good expression on the spatial relationship of two types of cells. In the top 8 features(Fig.12), there are two features of node centrality: between centrality and closeness centrality. They are used to describe the contribution of nodes in the graph. Between

centrality(Fig.4 in the appendix) represents the effect of nodes. The bigger the node's circles are, the more contribution nodes make.

#### E. Computational environment

The automatic diagnosis of mycosis fungoides and lichen planus can be divided into three stages. The first stage is tumor region recognition, which FD-VGG realized. 12 labeled slides were used for training and 2 for verification. Each training slice extracts the image block in the marked area by sliding window, and the image block size is H=W=224. The data expansion uses flip, rotation, cropping, color dithering. The learning rate is set to 0.01, and the momentum is 0.9. The epoch is 100. The second stage is the cell classification stage. First, according to the doctor mark,  $2048 \times 2048 \times 3$  image blocks got the ground truth of auxiliary branch I and auxiliary branch II. The input image is randomly cropped to  $352 \times 352 \times 3$  pixels in the training process. Then the data is expanded by using fuzzy, hue saturation and brightness change, brightness contrast change and gamma change so that Multi-SENet can learn more complex information and avoid overfitting. Due to the lack of data, the data expansion mode of



**Fig. 13.** Top 8 tissue level features. The features focus on the statistics of tumor shape, area and solidity.

each epoch in the training process is different, and the epoch is set to 1000. The encoder part of the network is initialized using the SENet154, which has been pre-trained on the ImageNet data. The above models are implemented based on Pytorch, an open-source deep-learning library, and optimized by a random gradient descent method. The optimizer is Ranger, and all the experimental device is NVIDIA Geforce GTX2080ti GPU.

#### F. Concluding remarks

This paper proposes an automatic diagnosis method for cutaneous mycosis fungoides(MFD). First, to locate the tumor area accurately, we developed the FD-VGG network. Its performance in diagnosis tasks is better than other networks. Then we propose Multi-SENet to classify lymphocytes and epithelial cells in sections. Finally, we extract the tissue-, tile- and cell-level features based on the prediction results of FD-VGG and Multi-SENet. MFD based on multi-histomorphology features has achieved the SOTA performance in the diagnosis task of mycosis fungoides. We are even the first to fuse three-level features and get good results. However, due to the limitation of data, there is only one benign inflammation whole slide in our dataset, which is the most difficult to distinguish from mycosis fungoides. We do not have another dataset to verify and expand, such as subacute dermatitis, eczema, and psoriasis. We will further study for this work.

#### ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (Nos.62171230,U1809205,61771249,91959207,81871352); Natural Science Foundation of Jiangsu Province of China (No. BK20181411).

#### REFERENCES

- [1] Michael Girardi, Peter W Heald, and Lynn D Wilson. The pathogenesis of mycosis fungoides. *New England Journal of Medicine*, 350(19):1978–1988, 2004.
- [2] Sareeta RS Parker and Bridget Bradley. Treatment of cutaneous t-cell lymphoma/mycosis fungoides. *Dermatology nursing*, 18(6):566, 2006.
- [3] Youn H Kim, Howard L Liu, Serena Mraz-Gernhard, Anna Varghese, and Richard T Hoppe. Long-term outcome of 525 patients with mycosis fungoides and sezary syndrome: clinical prognostic factors and risk for disease progression. *Archives of dermatology*, 139(7):857–866, 2003.
- [4] Adele De Masson, John T O’Malley, Christopher P Elco, Sarah S Garcia, Sherrie J Divito, Elizabeth L Lowry, Marianne Tawa, David C Fisher, Phillip M Devlin, Jessica E Teague, et al. High-throughput sequencing of the t cell receptor  $\beta$  gene identifies aggressive early-stage mycosis fungoides. *Science translational medicine*, 10(440), 2018.
- [5] Manuel Ciba, Robert Bestel, Christoph Nick, Guilherme Ferraz de Arruada, Thomas Peron, Comin César Henrique, Luciano da Fontoura Costa, Francisco Aparecido Rodrigues, and Christiane Thielemann. Comparison of different spike train synchrony measures regarding their robustness to erroneous data from bicuculline-induced epileptiform activity. *Neural computation*, 32(5):887–911, 2020.
- [6] Ebrahim Nasr-Esfahani, Shima Rafiei, Mohammad H Jafari, Nader Karimi, James S Wrobel, Shadrokh Samavi, and SM Reza Soroushmehr. Dense pooling layers in fully convolutional network for skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 78:101658, 2019.
- [7] Brian J Nickoloff and CE Griffiths. Intraepidermal but not dermal t lymphocytes are positive for a cell-cycle-associated antigen (ki-67) in mycosis fungoides. *The American journal of pathology*, 136(2):261, 1990.
- [8] Jérémie Scheurer, Claudio Ferrari, Luis Berenguer Todo Bom, Michaela Beer, Werner Kempf, and Luis Haug. Semantic segmentation of histopathological slides for the classification of cutaneous lymphoma and eczema. In *Annual Conference on Medical Image Understanding and Analysis*, pages 26–42. Springer, 2020.
- [9] Rein Willemze, Elaine S Jaffe, Gunter Burg, Lorenzo Cerroni, Emilio Berti, Steven H Swerdlow, Elisabeth Ralfkiaer, Sergio Chimenti, José L Diaz-Perez, Lyn M Duncan, et al. Who-eort classification for cutaneous lymphomas. *Blood*, 105(10):3768–3785, 2005.
- [10] Gregory A Hosler and Kathleen M Murphy. Leukemia and lymphoma. part i. mycosis fungoides and sézary syndrome: Using molecular tools to aid in the diagnosis, staging, and therapy for mycosis fungoides and sézary syndrome. In *Molecular Diagnostics for Dermatology*, pages 133–165. Springer, 2014.
- [11] Chetan L Srinidhi, Ozan Ciga, and Anne L Martel. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*, page 101813, 2020.
- [12] Angel Cruz-Roa, Ajay Basavanhally, Fabio González, Hannah Gilmore,

- Michael Feldman, Shridar Ganesan, Natalie Shih, John Tomaszewski, and Anant Madabhushi. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *Medical Imaging 2014: Digital Pathology*, volume 9041, page 904103. International Society for Optics and Photonics, 2014.
- [13] B Srikanth and S Venkata Suryanarayana. Multi-class classification of brain tumor images using data augmentation with deep neural network. *Materials Today: Proceedings*, 2021.
- [14] Jun Xu, Xiaofei Luo, Guanhao Wang, Hannah Gilmore, and Anant Madabhushi. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing*, 191:214–223, 2016.
- [15] Yan Xu, Zhipeng Jia, Liang-Bo Wang, Yuqing Ai, Fang Zhang, Maode Lai, I Eric, and Chao Chang. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC bioinformatics*, 18(1):1–17, 2017.
- [16] Khaoula Belhaj Soulami, Naima Kaabouch, Mohamed Nabil Saidi, and Ahmed Tamtaoui. Breast cancer: One-stage automated detection, segmentation, and classification of digital mammograms using unet model based-semantic segmentation. *Biomedical Signal Processing and Control*, 66:102481, 2021.
- [17] Yinyin Yuan, Henrik Failmezger, Oscar M Rueda, H Raza Ali, Stefan Gräf, Suet-Feung Chin, Roland F Schwarz, Christina Curtis, Mark J Dunning, Helen Bardwell, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Science translational medicine*, 4(157):157ra143–157ra143, 2012.
- [18] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:101563, 2019.
- [19] Navarun Das, Elima Hussain, and Lipi B Mahanta. Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network. *Neural Networks*, 128:47–60, 2020.
- [20] Ye Tian, Li Yang, Wei Wang, Jing Zhang, Qing Tang, Mili Ji, Yang Yu, Yu Li, Hong Yang, and Airong Qian. Computer-aided detection of squamous carcinoma of the cervix in whole slide images. *arXiv preprint arXiv:1905.10959*, 2019.
- [21] Haojia Li, Jon Whitney, Kaustav Bera, Hannah Gilmore, Mangesh A Thorat, Sunil Badve, and Anant Madabhushi. Quantitative nuclear histomorphometric features are predictive of oncotype dx risk categories in ductal carcinoma in situ: preliminary findings. *Breast Cancer Research*, 21(1):1–16, 2019.
- [22] Yanning Zhou, Simon Graham, Navid Alemi Koohbanani, Muhammad Shaban, Pheng-Ann Heng, and Nasir Rajpoot. Cgc-net: Cell graph convolutional network for grading of colorectal cancer histology images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pages 0–0, 2019.
- [23] Khalid AbdulJabbar, Shan E Ahmed Raza, Rachel Rosenthal, Mariam Jamal-Hanjani, Selvaraju Veeriah, Ayse Akarcu, Tom Lund, David A Moore, Roberto Salgado, Maise Al Bakir, et al. Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. *Nature Medicine*, 26(7):1054–1062, 2020.
- [24] Henrik Failmezger, Sathya Muralidhar, Antonio Rullan, Carlos E de Andrea, Erik Sahai, and Yinyin Yuan. Topological tumor graphs: a graph-based spatial model to infer stromal recruitment for immunosuppression in melanoma histology. *Cancer research*, 80(5):1199–1209, 2020.
- [25] Sajid Javed, Arif Mahmood, Muhammad Moazam Fraz, Navid Alemi Koohbanani, Ksenija Benes, Yee-Wah Tsang, Katherine Hewitt, David Epstein, David Snead, and Nasir Rajpoot. Cellular community detection for tissue phenotyping in colorectal cancer histology images. *Medical image analysis*, 63:101696, 2020.
- [26] Cheng Lu, Can Koyuncu, German Corredor, Prateek Prasanna, Patrick Leo, XiangXue Wang, Andrew Janowczyk, Kaustav Bera, James Lewis Jr, Vamsidhar Velcheti, et al. Feature-driven local cell graph (flock): New computational pathology-based descriptors for prognosis of lung cancer and hpv status of oropharyngeal cancers. *Medical Image Analysis*, 68:101903, 2021.
- [27] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [28] Fatemeh Farahnak-Ghazani and Mahdieh Soleymani Baghshah. Multi-label classification with feature-aware implicit encoding and generalized cross-entropy loss. In *2016 24th Iranian Conference on Electrical Engineering (ICEE)*, pages 1574–1579. IEEE, 2016.
- [29] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.
- [30] Jaime Lynn Speiser, Michael E. Miller, Janet Tooze, and Edward Ip. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134:93–101, 2019.
- [31] Sarah I Jawed, Patricia L Myskowski, Steven Horwitz, Alison Moskowitz, and Christiane Querfeld. Primary cutaneous t-cell lymphoma (mycosis fungoïdes and sézary syndrome): part i. diagnosis: clinical and histopathologic features and new molecular and biologic markers. *Journal of the American Academy of Dermatology*, 70(2):205–e1, 2014.
- [32] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. Pearson correlation coefficient. In *Noise reduction in speech processing*, pages 1–4. Springer, 2009.
- [33] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.
- [34] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [35] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [36] Shanghua Gao, Ming-Ming Cheng, Kai Zhao, Xin-Yu Zhang, Ming-Hsuan Yang, and Philip HS Torr. Res2net: A new multi-scale backbone architecture. *IEEE transactions on pattern analysis and machine intelligence*, 2019.