Region of Interest Identification in the Cervical Digital Histology Images

Tetiana Biloborodova[1,2][0000-0001-7561-7484], Semen Lomakin[3], Inna Skarga-Bandurova[4][0000-0003-3458-8730], and Yana Krytska[3][0000-0003-4575-2559]

[1]University of Applied Science HTW Saar, Saarbrücken, Germany
beloborodova.t@gmail.com
[2] G.E. Pukhov Institute for Modelling in Energy Engineering, Kyiv, Ukraine
[3] Volodymyr Dahl East Ukrainian National University, Severodonetsk, Ukraine
[4] Oxford Brookes University, Oxford, UK

Abstract. The region of interest (RoI) identification has a significant potential for yielding information about relevant histological features and is imperative to improve the effectiveness of digital pathology in clinical practice. The typical RoI is the stratified squamous epithelium (SSE) that appears on relatively small image areas. Hence, taking the entire image for classification adds noise caused by irrelevant background, making classification networks biased towards the background fragments. This paper proposes a novel approach for epithelium RoI identification based on automatic bounding boxes (bb) construction and SSE extraction and compares it with state-of-the-art histology RoI localization and detection techniques. Further classification of the extracted epithelial fragments based on DenseNet made it possible to effectively identify the SSE RoI in cervical histology images (CHI). The design brings significant improvement to the identification of diagnostically significant regions. For this research, we created two CHI datasets, the CHI-I containing 171 color images of the cervical histology microscopy and CHI-II containing 1049 extracted fragments of microscopy, which are the most considerable publicly available SSE datasets.

Keywords: Region of Interest (RoI), Cervical Histology Image, Bounding Box.

1 Introduction

Adoption of digital technology into histological practice increases the productivity and accuracy of the interpretation of the pathology slides [1], provides an objective quantitative assessment of the microscopy images, reduces the variability of the diagnosis, and improves differential diagnosis [2, 3]. Meanwhile, accurate automatic interpretation of histopathological images comes up against the difficulty that we still do not have a principled approach for identifying areas of interest. The lack of a gold standard for automatic RoI identification is explained by the morphological diversity of histological images, a wide variety of shapes of epithelial regions, density and shape of cells in these regions, the presence of artefacts, e.g., blood, mucus, staining in the tissue samples, the

presence of columnar cellular regions and varying sizes of the RoI in images. For instance, cervical digital histology images are widely used for early diagnosis to prevent malignancy but only a relatively small part of the image containing SSE can be seen as a diagnostic area (see Fig. 1). Hence, considering the entire image for segmentation would add noise caused by irrelevant portions in the background and lead to the segmentation network being biased towards the background regions [4].

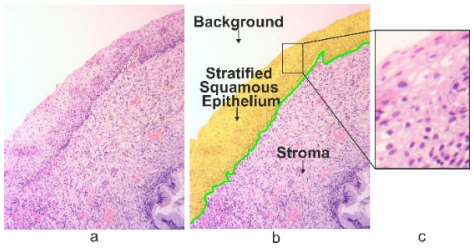

Fig. 1. An example of the CHI: (a) an initial image of 3288 × 4096 resolution at 40X magnification, (b) background, SSE and stroma (the yellow SSE region corresponds to the RoI in cervical histology microscopy and green line is picked to basement membrane, which separates the epithelium from the underlying stroma), (c) a zoomed-in view of the SSE

One of the possible solutions to increase detection performances is using bounding boxes (bb) around SSE as a region of interests (RoI) and then relying on region-based features for class identification [5]. This idea is both self-contained and attractive, but due to the highly diverse directions of SSE in histology images, it is often hard to come up with accurate RoI to pair with all the objects in a cervical image by using RoI with limited directions. The fine-grained, accurate RoI that follow directions and scales of actual boundaries of SSE suffer from high computational complexity during region identification and localization phases. For these reasons, most research has been focused on the analysis of manually selected RoI, making it difficult to replicate results on new data and ignoring the localization problem, which involves identifying RoI in new images. Early CHI image classification studies mainly focused on the discriminating models from the raw color input images to identify diagnostic elements for cell segmentation and differentiate them as RoI [6] or histological primitives [7, 8]. Most of them aim to image processing and identify areas of interest at the tissue level using histological primitives at the cellular level [9, 10] for cancer diagnosis. A set of studies on histology image analysis [11, 12, 13] utilize feature extraction and different attention strategies without relying on the RoI. However, as mentioned in [12], handcrafted color and texture features are rarely sufficiently robust, and any unlabeled object without a corresponding bounding box can be the source of an incorrect learning signal [14]. Superior technical solutions for SSE identification can be found in [15, 16]; however, to some extent, they also rely on the annotated SSE Finally, for many reasons, including concerns over privacy and confidentiality, a sufficient number of medical datasets of cervical histology microscopy could not be found in the public domain. We solve the problem of automatic localization and identification of the SSE on a new histological image dataset and propose a new multi-stage approach involving automatic

bb extraction and subsequent classification with a deep neural network to form the RoI representing the SSE in CHI. We demonstrate that this approach provides precise SSE identification even for the small size of the SSE in CHI with an overlapping clump of cells or masking by artifacts. Our specific contributions include:

- The CHDM-I dataset containing 171 full-size color images of the cervical histology microscopy, and CHI-II dataset containing 1049 fragments of microscopy. Compared with [13, 15, 16, 25], the CHDM-I proposes the largest number of samples among known public datasets.
- A method of automatic forming bb for RoI-based epithelium image segmentation.
- RoI construction pipeline based on bb, which guarantees efficiency and low complexity during region identification and localization phases.
- Assessment of the effectiveness of our approach through quantitative experiments with a validation of the findings by expert pathologists. Experiments also show that the proposed RoI identification approach can be used with other detector architectures providing significant improvements in the classification performance.

This paper is organised as follows. Section 2 discusses the related works. In Section 3, we present the pipeline for CHI processing and RoI identification. Section 4 describes experiments with presented datasets. Section 5 provides the conclusions.

2 Related work

Although RoI localization and identification is a well-known problem in analyzing histological images, there is still no general approach to all kinds of histology images. Much recent progress on an interpretation of the histology slide images has benefited a lot from the deep learning modes [10, 15, 17, 18, 19, 20, 21], including adversarial neural networks [17], trained on labeled data for basal membrane segmentation to detect cancer micro invasions; DenseNet [18] for tumor metastasis detection as the RoIs. The framework [18] comprises a patch-based classifier, an improved adaptive sampling method and a postprocessing filter on annotated data. A graph convolution network that admits a graph-based RoI representation [10] to incorporate local inter-patch context and, as in previous cases, RoI annotated data were used for RoI image classification. A comprehensive solution for identifying SSE as an area of interest is presented in [15]. RoI were used for differential diagnosis of cervical intraepithelial neoplasia (CIN). The procedure for epithelium localization was discussed in more detail in [19]. The classification of areas of interest in accordance with the CIN degree class using the DeepCIN network was proposed in [20, 21]. RoI identification was focused on differentially informative vertical segment regions [20]. Another group includes semi-supervised and weakly supervised techniques. A Multilayer Hidden Conditional Random Field framework for the CHI classification is an example weakly-supervised approach proposed in [13], where the CHI are mixed with the complex nucleus, interstitial and tissue fluid. Here, an immunohistochemical stained CHI dataset was used to test the effectiveness of the proposed model for the CHI classification. The classification was carried out

without pre-processing in order to detect the area of interest. Early research on the analysis of digital slides of cervical SSE [22] considers the threshold-based semi-automatic segmentation approach for obtaining the contours of cervical cancer RoI. Finally, a large group of unsupervised techniques [9, 23, 24, 25] demonstrate promising results in RoI identification and feature extraction. The detection framework using a bounding box to surround RoI for lung cancer images [9] offers fine-detailed, boundary-adherent super-pixel segmentation. This approach comprised super-pixel segmentation and super-pixel classification for RoI identification. Another approach [23] utilizes image-to-image translation for cancerous regions detection in histology images. It is based on training a model on tissues without pathological changes and then using it to detect visual abnormalities, which are cancerous tissue lesions. The paper [24] discusses RoI localization in whole slide images of breast biopsy slides based on understanding the visible patterns and predicting the diagnostically relevant areas. In [25], a region-based segmentation and grey-level co-occurrence matrix (GLCM) were applied for feature extraction and classification. It was noted that color and morphology of SSE show wide variability across different samples and even within the same image. Therefore, the simple segmentation approach demonstrates poor effectiveness for the SSE segmentation from cervical histological digital slides.

3 Method

In this section, we present the general pipeline of cervical digital histology image processing and RoI identification enriched with a new procedure of automatic bb construction embedded in this process. We start with data processing of cervical slide images to achieve out edge of SSE. Then we introduce the procedure of automatic bb construction. Third stage includes bb decomposition on patches and classification. Finally, we perform epithelium RoI identification and assessment. An overview of the methodology is shown in Fig. 2. Each stage consists of several procedures discussed below.

3.1 Cervical slide image processing

The process begins with traditional image processing to form a curve that defines the lateral border (outer edge) of the tissue. Tissue can be epithelium or not. To accomplish this, the following procedures are used: (a) converting the RGB color model to the grayscale model, (b) blurring the image according to Gaussian, (c) image binarization using thresholding function, (d) cervical contour detection. All procedures are applied to low-resolution 15% reduced images (initial image size is 4096 × 3288, reduced size is 614 × 493). Threshold, the length of the tangent, and the width and height of the bounding boxes are chosen empirically based on the low-resolution images. The grey-scale value is computed as the weighted average of the RGB values as Greyscale = 0.299R + 0.587G + 0.114B [26]. Then Gaussian blur is applied to reduce the high-frequency components of the image [27]. Image thresholding converts a grayscale images to a black and white ones, where white pixels correspond to the background (no tissue), and the black pixels correspond to the tissue. The search for the contour of black

pixels leads to the selection of the contours of the tissue surface and image boundaries. Further, the contours coinciding with the borders of the image are removed, and thus only the curve remains, indicating the outer surface of the tissue.

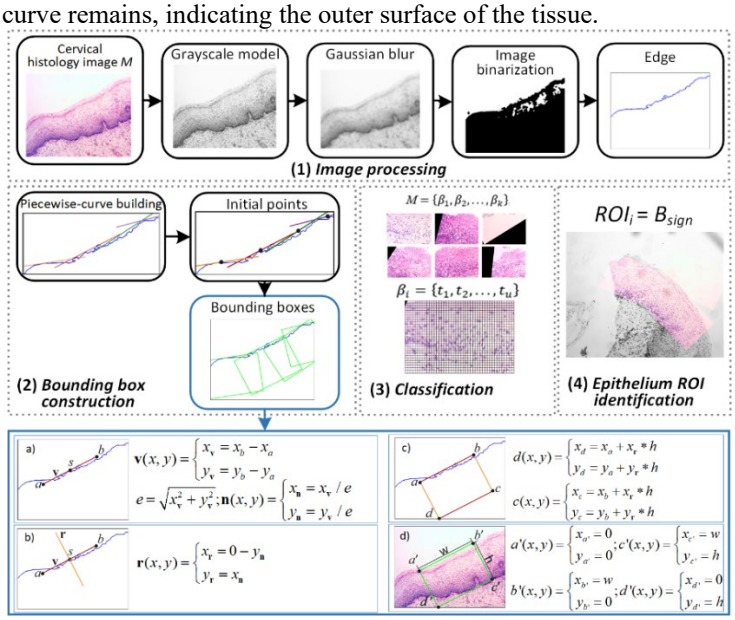

Fig. 2. An overview of proposed approach: (1) cervical slide image processing procedure, (2) bb construction presents piecewise curve and bb constraction (a, b, c, d shows bb constraction steps), (3) classification and (4) epithelium RoI identification.

3.2 Bounding box (bb) construction procedure

When plotting the contour of the tissue, the coordinates are stored in two arrays, the coordinates along the *x*-axis and the coordinates along the *y*-axis. Images have an origin in the upper left corner; thus, the *y*-axis is inverted. Since the number of points per tangent and the length / width of the bb values are indicated relative to the initial dimensions of the image, we reduce these values proportionally to the image reduction.

$$q' = q \cdot \left(\frac{z}{100}\right), \tag{1}$$

where q denotes the initial value, z denotes the reduction percentage of the original image, and q' is the new value. Shaping bb involves shaping piecewise curves and shaping rectangles, which are the basis of bb.

The contour processing. The entire length of the contour is computed iteratively until all points of the contour have been achieved. During each iteration, a separate fragment of the contour is processed, and its length depends on the number of points included in the fragment. In addition, each iteration is accompanied by a search for a tangent of indefinite length by fitting a polynomial using the least squares method [28]. For each fragment of the contour, the average value of the coordinates along the *x*-axis is calculated. Resulting value is compared with the coordinates of each point along the

x-axis. The point with the least difference is used to draw a tangent of fixed length.

Constructing a fixed-length tangent involves finding two points that define the tangent's boundaries. The input data are the length of the tangent, the slope of the tangent, and starting point from which two vectors are constructed in opposite directions. The ends of the constructed vectors define the tangent boundaries. If the slope of the tangent is zero, then the coordinates (*x*, *y*) along the *x*-axis and *y*-axis for two points *a* and *b* that define the boundaries of the tangent are calculated as follows

$$a(x,y) = \begin{cases} x_a = x_s - l/2, \\ y_a = y_s, \end{cases} \qquad b(x,y) = \begin{cases} x_b = x_s + l/2, \\ y_b = y_s, \end{cases} \tag{2}$$

where *x* and *y* define the coordinates of points along the x-axis and y-axis, respectively, *a* and *b* are the points that define the tangent line, *s* is the starting point, *l* is the length of the tangent.

If the slope of the tangent does not exist, that is, when the tangent line is parallel to the y-axis, and the slope is taken to be infinity, then the coordinates of the points are calculated using Eq. (3)

$$a(x,y) = \begin{cases} x_a = x_s, \\ y_a = y_s - l/2, \end{cases} \qquad b(x,y) = \begin{cases} x_b = x_s, \\ y_b = y_s + l/2. \end{cases} \tag{3}$$

If the slope does not meet the above conditions, then the tangent line coordinates are calculated as follows. First, the change in coordinates Δx along the *x*-axis and Δx along the *y*-axis is calculated:

$$\Delta x = \frac{\frac{l}{2}}{\sqrt{1+m^2}}, \qquad \Delta y = \Delta x \cdot m, \tag{4}$$

where *m* is the slope. Then, the coordinates of the points of the tangent line are calculated using Eq. (5)

$$a(x,y) = \begin{cases} x_a = x_s - \Delta x, \\ y_a = y_s - \Delta y, \end{cases} \qquad b(x,y) = \begin{cases} x_b = x_s + \Delta x, \\ y_b = y_s + \Delta y. \end{cases} \tag{5}$$

Owing this, the contiguous straight applied lengths are realized, making up a piece-wise-specified curve, most closely corresponding to one obtained at the previous stage. As a result, it generates *n*-dimensional continuous piecewise linear function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, where $\mathbb{R}$ is the set of points after processing stage, pointing to the tissue, and $\mathbb{R}^n$ is the set of points of the piecewise linear straight lines that best fits the original curve.

The bb building and extraction. The bbs are constructed through the forward and backward vectors. The coordinates of tangent line of a certain length are used to achieve the coordinates of the direction vector **v**(*x*, *y*). Then, **v** is turned to the normalized direction vector **n**(*x*, *y*). First, the scalar *e* of the space of **v** is calculated as it is presented in Algorithm 1, and used to achieve coordinates of **n**. Then **n** is rotated 90° to obtain a new normalized vector **r**(*x*, *y*) which is used to form the sides of the rectangle that is the basis of the bb. Next, *c*(*x*, *y*) and *d*(*x*, *y*) the coordinates of the two unknown vertices of the bb are calculated. Thus, having *a* and *b*, the known vertices of the bb rectangle, *h* is the height of the rectangle, and **r** is the normalized vector rotated by 90°, *c* and *d*, the unknown vertices of the rectangle are computed using Algorithm 1. The obtained coordinates of all four vertices are used for bb β_i construction and visualization.

Algorithm 1: Bounding box construction

Input: $f: \mathbb{R}^n \to \mathbb{R}$
Parameter: a, b, h
Result: $\beta_i(a,b,c,d) \epsilon M$
Get the direction vector for given points a,b

$$\mathbf{v}(x,y) = \begin{cases} x_{\mathbf{v}} = x_b - x_a \\ y_{\mathbf{v}} = y_b - y_a \end{cases}$$

Normalize the vector **v**

$$e = \sqrt{x_{\mathbf{v}}^2 + y_{\mathbf{v}}^2}$$

$$\mathbf{n}(x,y) = \begin{cases} x_{\mathbf{n}} = \dfrac{x_{\mathbf{v}}}{e} \\ y_{\mathbf{n}} = \dfrac{y_{\mathbf{v}}}{e} \end{cases}$$

Rotate the vector 90 degrees by swapping x and y, and inverting one of them

$$\mathbf{r}(x,y) = \begin{cases} x_{\mathbf{r}} = 0 - y_{\mathbf{n}} \\ y_{\mathbf{r}} = x_{\mathbf{n}} \end{cases}$$

Create a new line at b pointing in the direction of **v**

$$c\,(x,y) = \begin{cases} x_c = x_b + x_{\mathbf{r}} * h \\ y_c = y_b + y_{\mathbf{r}} * h \end{cases}$$

Create a new line at a pointing in the direction of v

$$d\,(x,y) = \begin{cases} x_d = x_a + x_{\mathbf{r}} * h \\ y_d = y_a + y_{\mathbf{r}} * h \end{cases}$$

The generated bbs provide necessary fragments from the original full-size image, and coordinates of the bb vertices correspond to the reduced version of the image. For this reason, additional calculations of the bb vertex coordinates are required for the full-size image:

$$a(x,y) = \begin{cases} x_a = 100 \cdot x_a/z \\ y_a = 100 \cdot y_a/z' \end{cases} \quad b(x,y) = \begin{cases} x_b = 100 \cdot x_b/z \\ y_b = 100 \cdot y_b/z' \end{cases}$$
$$c(x,y) = \begin{cases} x_c = 100 \cdot x_c/z \\ y_c = 100 \cdot y_c/z' \end{cases} \quad d(x,y) = \begin{cases} x_d = 100 \cdot x_d/z \\ y_d = 100 \cdot y_d/z' \end{cases} \tag{6}$$

where *a, b, c*, and *d* are the vertices of the rectangle and *z* is the reduction percentage of the original image.

Next, the coordinates of the supplementary rectangle are determined, which has identical dimensions, but does not have a slope and is located at one of the vertices at the point of origin. The coordinates of the outlined rectangle are defined as

$$a'^{(x,y)} = \begin{cases} x_{a'} = 0 \\ y_{a'} = 0 \end{cases}, \quad b'^{(x,y)} = \begin{cases} x_{b'} = w \\ y_{b'} = 0 \end{cases},$$
$$c'(x,y) = \begin{cases} x_{c'} = w \\ y_{c'} = h \end{cases}, \quad d'(x,y) = \begin{cases} x_{d'} = 0 \\ y_{d'} = h \end{cases}, \tag{7}$$

where *a', b', c'* and *d'* are the vertices of the construction rectangle, and *w* and *h* are the width and height of bb.

The coordinates of the vertices of the construction rectangle and the bb, give the perspective transformation matrix, which is used to extract a fragment of an image from a full-size image by deforming the rotated bb to obtain a straightened rectangle.

3.3 Extracted bb classification

After bb extraction, a new dataset of histological image fragments is generated. Depending upon the presence or absence of SSE, all extracted fragments are marked with a positive class label in there any SSE, and a negative class label in the absence of SSE in the bb. The resulting dataset is used to train the neural network model.

3.4 Epithelium RoI identification

The microscopic image M includes a certain amount of bbs β which depends on settings of contour processing stage and number of obtained tangents. We model the microscopy image M as a set of k extracted bb β: $\mathrm{M} = \{\beta_1, \beta_2, \ldots, \beta_k\}$, where each bb $\beta_i = \{t_1, t_2, \ldots, t_u\}$ is represented by a set of u patches t_i. Each patch is represented by the following expression $t_i = (P_i^O), 1 \le i \le u$. For the i-th patch t_i, P_i^O is a set of values of the probabilities of belonging of the patch to the labels of the classes of the set of labels O. The set of labels of the classes O in the proposed approach is represented as $O_k = \{o_k^{neg}, o_k^{pos}\}, 1 \le k \le u$, where o_k^{neg} is the class label of the k-th patch t_i indicating the absence of SSE in the patch, and o_k^{pos} is the class label of the k-th patch t_i indicating the presence of SSE in the patch. Based on this, we describe P_i^O as follows: $P_i = \{p_i^{neg}, p_i^{pos}\}, 0 \le p \le 1$. Since each selected bb may contain or may not contain SSE of cervical microscopy, they are defined as significant or insignificant, depending on the patch class labels obtained as a result of the classification and the probabilities of this class $\beta_i = \underset{i \in [o^{pos}, o^{neg}]}{arg\,max} \{t_i\}$. If a majority of the o^{pos} patch class positive labels are associated with presence of SSE, the β_i bb is defined as significant; otherwise, β_i bb is qualified as insignificant, i.e. $\beta_i' = \emptyset$. Hence, each image M can contain combined sets of two types of bb, B_{in} and B_{sign}, where B_{in} is a set of insignificant bb and B_{sign} denotes the dataset of significant bb containing SSE In these ways, the image defined as $M = B_{in} \cup B_{sign}, B_{in} \cap B_{sign} = \emptyset$, where insignificant set of bb B_{in} represents an empty set of objects $B_{in} = \emptyset$ subject to $\beta_i' = \emptyset$ and $1 \le i \le j$ and significant set of bb B_{sign} represents a set of patches labeled with class o^{pos} with a probability value $p = argmax_i\{P_i\}$. Thus, we get B_{sign} for β_i' subject to $\beta_i' \ne \emptyset$ and $1 \le i \le f$. As a result, RoI_i of the i-th microscopic image M_i with SSE is defined as $\mathrm{RoI}_i = B_{sign}$.

4 Experiments

The experiments were conducted on cervical histology images datasets CHI-I and CH-II developed with the assistance of medical professionals and expert pathologists and use of the proposed method for constructing bb and SSE RoI identification. The software implementation was carried out using the Python programming language [29] and the OpenCV library [30]. The consultations of medical experts was provided at the processing stage to determine the parameters of the SSE: (1) the threshold value that cuts off image fragments without a tissue site; (2) the size of the bounding box based on the height of the epithelium. Consultations with medical experts were carried out at the stage of training the model to annotate the obtained fragments and when testing the proposed approach for validating the obtained fragments for the presence of SSE.

4.1 Dataset

For this research, we created two CHI datasets captured during the years 2018, 2019 and 2020 and provided by the Department of Pathology at the Municipal Institution "Severodonetsk City Multidisciplinary Hospital", Severodonetsk, Ukraine. Histological slides were made by freezing and paraffin wire methods, stained by hematoxylin and eosin staining. The thickness of the sections and the degree of coloration vary. Data are fully anonymized. Written informed consent was obtained from all subjects (patients) in this study. Digital slides were acquired at 10× objective magnification using Carl Zeiss Primo Star, SIGETA UCMOS camera and archived in 24-bit color JPEG format. Acquisition software is NIS-Elements F 3.2. The images have a resolution of 4096 × 3288 pixels.

The CHI-I dataset contains 171 color images of the cervical histology microscopy, 135 samples of cervix connective tissue with SSE and 36 samples without SSE CHI-I has two identifiers: (1) Image ID, the number of the histological specimen. One patient and, therefore, one diagnosis may correspond to several specimens belonging to one diagnostic case (visit); (2) Visit ID, the unique identifier for one diagnostic case (visit).

The CHI-II dataset contains 1049 fragments of microscopy, 644 bbs of positive class (contain SSE) and 405 bbs of negative class (background and/or artifacts without SSE). All extracted bbs inherit the visit ID from CHI-I. The set descriptions also specify the hyperparameters of the bb that make up the CHDM-II dataset: threshold lower limit, bb width and bb height. RoIs from CHI-II were annotated by expert pathologists.

4.2 Procedure

The cervical image processing stage forms an outer edge curve of tissue. To process the image, the following parameters were used: Gaussian Kernel Size for Gaussian Blur function (25,25), threshold value = 230, number of points in one curve section = 1500. To avoid issues relating to edge detection, initial images must be set up horizontally so that the outer surface of SSE is directed to the right or upward, possibly at an angle.

Piecewise curves are formed iteratively until all points of the contour obtained at the previous stage have been processed. During each iteration, a separate contour fragment is processed, and its length depends on the set number of points included in the fragment. Each iteration is accompanied by the search and construction of a tangent of indefinite length. Further, for each fragment of the contour, the average value of the coordinates of the points along the x-axis is calculated. The obtained value is compared with the coordinates of each point of the fragment along the x-axis. The point with the least difference is determined by the original one and will be used to construct a tangent of fixed length. Finally, the coordinates of the starting points are used to calculate the angle of inclination of the straight line and form a rectangle. To obtain the rectangles (bb), the following parameters were used: bounding box width = 2250, bounding box height = 1500.

Bringing the reverse scaling of the resulting rectangles to the original image, we get bb - highlighted in green in the original image in Fig. 3a and in the form of separate highlighted fragments in Fig. 3b. The following image scaling options are used holds the percentage by which image has to be scaled = 0.15. Depending on the specimen under study, the extracted fragments may not contain stratified squamous epithelium.

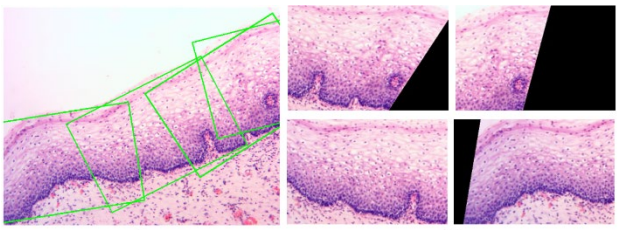

Fig. 3. Result of shaping and extracting bb.

The extracted fragments were annotated by experts and formed the CHDM-II dataset. The fragment class was annotated as positive only if the extracted patch contains an SSE fragment image in height, i.e., from the basement membrane to the outer layer. The negative class includes microscopic fragments that do not contain an SSE, a background without tissue. Patches containing SSE, but not meeting the requirements of the positive class, were not included in the set - Single segments which are given "0". The number of fragments labeled with the positive class was 644 samples, the negative class 405 samples.

The fragments presented in the CHDM-II dataset are used to train the model. As was mentioned in [15, 24], the main challenges in this domain are the limited availability of annotated data, and images of varying sizes. To overcome these issues at the fragment classification stage, each of the extracted fragments was divided into 1000 overlapping patches with a patch stride s and fed to the network input. The total of 1,049,000 labelled image patches were generated, with 90% of the data used for training and 10% of data used for testing. The fragment class label is assigned according to the class of most patches.

The DenseNet [31] is trained for 100 epochs with an Adam optimizer with a learning rate of 0.0001 under early stopping conditions. Training parameters: batch size = 128, patch size = 224. Weights are used to balance the simulation results to obtain a correct model. The model is run on PyTorch platform [32] using nVidia GeForce RTX 2060 Super with 8GB of memory. Model training completed under 5 h.

4.3 Assessment

To evaluate the model, we used a test set consisting of 54 images, 28 images of the negative class and 28 images of the positive class. We evaluate the DenseNet network for classification as s.s.e/non-SSE RoIs. The scoring metrics include accuracy (ACC), precision (PRC), recall (RR), F1-score ($F1$) [33]. The overall classification accuracy for both classes is 98%. We have used additional classification metrics, such as Fowlkes–Mallows index (FM), Matthews correlation coefficient (MCC). These metrics are more correctly reflected the quality of the prediction. With the test data we obtain FM=1 and MCC=0.959. Maximum possible value of these metrics is 1, which corresponds to the best classification possible, where all the elements have been perfectly classified. The proposed approach was validated using images that were not involved in training and testing. The following parameters were used for validation: kSize = (25,25), threshold value = 240, one curve points = 1500, bb width = 1600, bb

height = 1000, scale percent = 0.15. For model validation, we took several tough cases, e.g., SSE is highly compressed and skewed with respect to connective tissue, SSE boundaries and edges do not have a clear color delimitation with the fabric image. The selected fragments are classified using the trained model and analyzed by medical experts.

4.4 Comparison with State-of-the-arts

We compare our method to epithelium RoI identification with state-of-the-art histology slides RoI localization and detection. In the proposed approach the classifier is jointly trained on positive and negative image fragments. Table 1 summarizes the accuracy results of the current research. It shows the highest results achieved in the studies reviewed. The results show that the use of the proposed approach made it possible to achieve a higher classification accuracy compared to the breast cancer RoI allocation study by 18.4% and 19.44% than in studies [20] and [10], respectively.

Table 1. Quantitative comparisons with the state-of-the-art methods for RoI identification based on classification accuracy rates (%)

Histology RoI	ACC	PRC	RR	F1
breast cancer [24]	79.60	–	–	–
breast cancer[10]	78.56	–	–	–
lung cancer [9]	–	71.27	73.33	–
anomaly (cancerous) [23]	–	–	–	92
basal membrane [17]	–	61.2	63.6	62.4
cervical intraepitelial neoplasia [20]	88.5	–	–	88.0
cervical cancer [13]	–	93.75	**100**	91.43
cervical SSE [15]	97.8	–	–	95.6
cervical SSE, Ours	**98**	**96**	**100**	**98**

The results show that the proposed approach to identifying epithelium RoI allows achieving more accurate classification results than the classification results in current research aimed at histology slides RoI detection.

5 Conclusion

In this paper, we presented an approach to identifying epithelium RoI in cervical digital slides. To implement the proposed approach, we have created a cervical digital slides dataset. We used a processing step that included color model transformation, image blurring, image binarization and edge detection to highlight the curve indicating the outer edge of the cervical tissue. We proposed a new method for bb building and extraction, which allows isolating fragments of SSE from cervical digital slides, thus, creating a dataset of patches containing epithelium and not containing epithelium for further training of the classification model. The results demonstrate that the proposed approach made it possible to achieve classification accuracy up to 98% which is higher compared to the related works. Also, we have proposed a new method for identifying epithelium RoI in cervical digital slides. The proposed technique is also applicable in other histology images where the lateral border of epithelium should be seen as RoI. The CHI-I and CHI-II datasets, models, and code can be accessed on the CHI GitHub

page. We hope that the proposed approach for epithelium RoI identification in cervical digital slides will allow conducting more comprehensive studies to analyze the diagnostic patterns of cervical diseases in histopathology.

References

[1] Ho, J., Ahlers, S.M., Stratman, C., Aridor, O., Pantanowitz, L., Fine, J.L., Kuzmishin, J.A., Montalto, M.C., Parwani, A.V.: Can digital pathology result in cost savings? A financial projection for digital pathology implementation at a large integrated health care organization. *Journal of pathology informatics*, 5(1):33 (2004).

[2] Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B.: Histopathological image analysis: A review. In: *IEEE Reviews in Biomedical Engineering*, 2: 147–171 (2009).

[3] Jaume, G., et al.: Quantifying Explainers of Graph Neural Networks in Computational Pathology. In *CVPR*, pp. 8106-8116 (2021).

[4] Zamzmi, G., Sachdev, V., Antani S.: Trilateral Attention Network for Real-time medical image segmentation. *arXiv preprint arXiv:2106.09201* (2021).

[5] Ding, J., Xue, N., Long, Y., Xia, G.-S., Lu, Q.: Learning RoI Transformer for Oriented Object Detection in Aerial Images. In CVPR, pp. 2849–2858 (2019).

[6] Chen, K., Zhang, N., Powers, L., Roveda, J.: Cell nuclei detection and segmentation for computational pathology using deep learning. In *Spring Simulation Conference* (*SpringSim*), pp.1–5 (2019).

[7] Ginley, B., Jen, K-Y., Rosenberg, A., Yen, F., Jain, S., Fogo, A., Sarder, P.: Neural network segmentation of interstitial fibrosis, tubular atrophy, and glomerulosclerosis in renal biopsies. *arXiv preprint arXiv:2002.12868* (2020).

[8] Hermsen, M., de Bel, T., den Boer, M., Steenbergen, E.J., Kers, J., Elorquin, S., Roelofs, J.J.T.H., Stegall, M.D., Mariam P Alexander, Byron H Smith, Bart Smeets, Luuk B Hilbrands, Jeroen A W M van der Laak. Deep learning-based histopathologic assessment of kidney tissue. *Journal of the American Society of Nephrology*, 30:1968–1979 (2019).

[9] Junzhou, H., Li., R.: Fast regions-of-interest detection in whole slide histopathology images. *Histopathology and Liquid Biopsy*. IntechOpen, vol. 67 (2020).

[10] Aygüneş, B., Aksoy, S., Gökberk Cinbiş, R., Kösemehmetoğlu, K., Önder, S., Üner, A.: Graph convolutional networks for region of interest classification in breast histopathology. In *Medical Imaging 2020: Digital Pathology*, pages 113200K (2020).

[11] Gu, J., Fu, C.Y., Ng, B.K., Liu, L.B., Lim-Tan, S.K., Lee, C.G.L.: Enhancement of early cervical cancer diagnosis with epithelial layer analysis of fluorescence lifetime images. *PLoS One*, 10(5):e0125706 (2015).

[12] Xia Li, Zhenhao Xu, Xi Shen, Yongxia Zhou, Binggang Xiao, Tie-Qiang Li. Detection of Cervical Cancer Cells in Whole Slide Images Using Deformable and Global Context Aware Faster RCNN-FPN. *Current Oncology*, 28(5):3585-3601 (2021).

[13] Li, C. et al. Cervical Histopathology Image Classification Using Multilayer Hidden Conditional Random Fields and Weakly Supervised Learning. *IEEE Access*, 7:90378-90397 (2019).

[14] Yang, Y., Liang, K.J., Carin, L. Object Detection as a Positive-Unlabeled Problem. *arXiv preprint arXiv.* 20202002.04672 (2020).

[15] Sornapudi, S., et al. Cervical Whole Slide Histology Image Analysis Toolbox. *medRxiv 2020.07.22.20160366* (2020).

[16] Gallwas, J., Jalilova, A., Ladurner, R., Kolben, T.M., Kolben, T., Ditsch, N., Homann, C., Lankenau, E., Dannecker, C. Detection of cervical intraepithelial neoplasia by using optical coherence tomography in combination with microscopy. *Journal of Biomedical Optics*, 22(1):16013, 2017.

[17] Wang, D., Gu, C., Wu, K., Guan, X.: Adversarial neural networks for basal membrane segmentation of microinvasive cervix carcinoma in histopathology images. In *2017 International Conference on Machine Learning and Cybernetics (ICMLC)*, pp. 385-389, (2017).
[18] Ruan, J., Zhu, Z., Wu, C., Ye, G., Zhou, J., Yue, J.. A fast and effective detection framework for whole-slide histopathology image analysis. *PlosOne*, 16(5):e0251521, (2021).
[19] Sudhir Sornapudi, Jason Hagerty, R. Joe Stanley, William V. Stoecker, Rodney Long, Sameer Antani, George Thoma, Rosemary Zuna, Shellaine R. Frazier. EpithNet: Deep Regression for Epithelium Segmentation in Cervical Histology Images. *Journal of Pathology Informatics,* 11:10, (2020).
[20] Sudhir Sornapudi, R. Joe Stanley, William V. Stoecker, Rodney Long, Zhiyun Xue, Rosemary Zuna, Shellaine R. Frazier, Sameer Antani. DeepCIN: Attention-based cervical histology image classification with sequential feature modeling for pathologist-level accuracy. *Journal of Pathology Informatics*, 11:40, (2020).
[21] Sudhir Sornapudi, R. Joe Stanley, William V. Stoecker, Rodney Long, Zhiyun Xue, Rosemary Zuna, Shelliane R. Frazier, Sameer Antani Feature based Sequential Classifier with Attention Mechanism. *arXiv preprint arXiv:2007.11392*, (2020).
[22] B Weyn, W.A.A Tjalma, P Vermeylen, A van Daele, E Van Marck, W Jacob. Determination of tumour prognosis based on angiogenesis-related vascular patterns measured by fractal and syntactic structure analysis. *Clinical Oncology*, 16(4):307-316, (2004).
[23] Dejan Stepec, Danijel Skocaj. Unsupervised Detection of Cancerous Regions in Histology Imagery using Image-to-Image Translation. In *Proceedings of the IEEE / CVF Conference on Computer Vision and Pattern Recognition*, pages 3785-3792, (2021).
[24] Mercan, E., Aksoy, S., Shapiro, L.G., Weaver, D.L., Brunyé, T.T., Elmore, J.G. Localization of Diagnostically Relevant Regions of Interest in Whole Slide Images: a Comparative Study. *Journal of Digital Imaging*, 29(4):496-506 (2016).
[25] Wang, Y., Crookes, D., Osama Sharaf Eldin, Shilan Wang, Peter Hamilton, Jim Diamond. Assisted diagnosis of cervical intraepithelial neoplasia (CIN). *IEEE Journal of Selected Topics in Signal Processing*, 3(1):112-121, (2009).
[26] Bovik, A.C. (ed.). The essential guide to image processing. Academic Press, (2009).
[27] Nixon, M., Aguado A. Feature extraction and image processing for computer vision. *Academic press*, (2019).
[28] Simanca S.R., Sutherland, S. Mathematical problem solving with computers. *The University at Stony Brook*, (2002).
[29] Van Rossum, G.: Python Programming Language. In *USENIX annual technical conference*, 41:36 (2007).
[30] Bradski, G., Kaehler, A. OpenCV. *Dr. Dobb's journal of software tools*, 3 (2000).
[31] Huang G., Liu Z., van der Maaten L., Weinberger, K.Q. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708 (2017).
[32] Paszke A., et al. Automatic Differentiation in Pytorch (2017).
[33] Pal A., et al. Deep Metric Learning for Cervical Image Classification, in *IEEE Access*, vol. 9, pp. 53266-53275 (2021).
[34] CHI. https://github.com/beloborodova-t/CHI/tree/main/Data