Research Article

New Morphological Features for Grading Pancreatic Ductal Adenocarcinomas

Jae-Won Song and Ju-Hong Lee

Department of Computer & Information Engineering, Inha University, 253 Yonghyun-dong,
Nam-gu, Incheon 402-751, Republic of Korea

Correspondence should be addressed to Ju-Hong Lee; juhong@inha.ac.kr

Pathological diagnosis is influenced by subjective factors such as the individual experience and knowledge of doctors. Therefore, it may be interpreted in different ways for the same symptoms. The appearance of digital pathology has created good foundation for objective diagnoses based on quantitative feature analysis. Recently, numerous studies are being done to develop automated diagnosis based on the digital pathology. But there are as of yet no general automated methods for pathological diagnosis due to its specific nature. Therefore, specific methods according to a type of disease and a lesion could be designed. This study proposes quantitative features that are designed to diagnose pancreatic ductal adenocarcinomas. In the diagnosis of pancreatic ductal adenocarcinomas, the region of interest is a duct that consists of lumen and epithelium. Therefore, we first segment the lumen and epithelial nuclei from a tissue image. Then, we extract the specific features to diagnose the pancreatic ductal adenocarcinoma from the segmented objects. The experiment evaluated the classification performance of the SVM learned by the proposed features. The results showed an accuracy of 94.38% in the experiment distinguishing between pancreatic ductal adenocarcinomas and normal tissue and a classification accuracy of 77.03% distinguishing between the stages of pancreatic ductal adenocarcinomas.

1. Introduction

Pathological diagnosis is currently performed subjectively via the knowledge and experience of doctors after inspection of tissue slides through light microscopes. This subjective diagnosis has some problems. First, tumor screening at high magnification by light microscope requires a lot of time and effort [1, 2]. Also, the individual competence of doctors has a decisive effect on the final diagnosis. This means that the pathological diagnosis of the same tumor by two different doctors may vary, because it is not based on objective quantitative feature analysis [3–5].

The advent of digital pathology has led to a new type of pathological diagnosis. Digital pathology changes glass slides into digital images. It provides a convenient screening environment that also makes it possible to objectively diagnose tumors through quantitative feature analysis using a computer [6–8]. Nevertheless, a pathological diagnosis needs much time and workload because the diagnosis is manually processed. Therefore, there are still problems related to the

subjective diagnosis of tumors. To overcome these problems, studies on computer-aided diagnosis (CAD) techniques based on digital pathology are growing. Many studies of CAD-based digital pathology have covered not only detection of tumors but also grading the stages of tumors. In pathology, tumor stage grading is useful to identify the extent of the disease and determine the appropriate treatment for a patient [9]. In fact, it is known that pathologists receive intensive trainings on grading stages of tumors in order to prescribe the correct type of treatment [7]. Currently, many studies of pathological diagnosis with CAD techniques are being carried out on breast and prostate tumors.

First, most CAD studies related to prostate cancer are based on the Gleason grading system [10]. Tabesh et al. [11] proposed an automated system based on machine learning to diagnose prostate cancer and grade the stages (low and high) of cancer with the Gleason grading system. They extracted color, texture, and morphometric features at global and object levels of a given tissue image. Classifying algorithms such as Gaussian, k-nearest neighbor (k-NN), and Support Vector

Machine (SVM) learned the features of cancer diagnosis and Gleason grading. Naik et al. [12] proposed a diagnostic system for distinguishing between intermediate Gleason grades. They identified the candidate gland region using a Bayesian classifier with low-level information and eliminated false positive regions identified as glands using empirical domain information. After that, the morphologic features were extracted from the identified glands, and Gleason grades 3, 4, and benign were classified through the SVM that learned the features. Huang and Lee [13] classified prostate cancer images into 4 grades based on a Gleason grading system. They used Bayesian, k-NN, and SVM classifiers for classifying stages of the cancer. And, to teach the classifiers Gleason grading, features are proposed by using differential box-counting and entropy-based dimension estimation techniques. In addition to that, there are many studies for the grading and diagnosis of prostate tumors [14–17].

There are also CAD studies related to breast tumors. Anderson et al. [18] worked on a problem for distinguishing ductal hyperplasia (DH), which is benign, and ductal carcinoma in situ (DCIS), which is malignant. In this study, they automatically segmented breast ducts using knowledge-guided machine vision and proposed measuring duct cribriformity and architectural complexity to quantitatively analyze the duct patterns in proliferative lesions to distinguish between DH and DCIS. Bilgin et al. [19] proposed a method to diagnose breast cancer using graph theory techniques. They segmented given tissue images using a k-means algorithm and generated different cell graphs using the positional coordinates of cells for each segmented image. An SVM model that can classify given tissue images into benign, invasive, and non-invasive (ductal carcinoma *in situ*) was learned by quantitative metrics that are computed from the generated cell graphs. Basavanhally et al. [20] proposed a grading system that identifies and grades the extent of lymphocytic infiltration (LI), a known viable prognostic indicator. First, they detected lymphocytes using region growing and Markov random field algorithms. Then, the architectural features were extracted from the detected lymphocytes, and the extent of LI was classified into low, medium, and high grades by an SVM classifier learned by the features. In addition, there are a lot of studies for the diagnosis and grading of breast tumors [21–24].

Many pathological CAD studies make an effort to analyze the pathological characteristics of and design the methods for quantitatively measuring each disease, because there are many methods for pathological diagnosis according to the type of disease and lesion. Currently, in addition to the breast and prostate tumors mentioned above, some studies on colonic [25], bladder [26], neuroblastoma [27–29], and follicular lymphoma [30, 31] tumors have been performed. Pathological CAD studies are still in the early stages and focus on a few tumors. There are many methods for diagnosing different types of tumors. Therefore, more studies of pathological CADs must be performed.

The aim of this study is to detect pancreatic ductal adenocarcinoma (PDAC) and classify them by stages. To achieve this, we propose new morphological features for diagnosing and grading PDAC. The region that is inspected to diagnose PDAC is a duct that consists of lumen and epithelium. Therefore, this paper segments the given image into lumen, epithelial nuclei, and nonepithelial nuclei and extracts the morphological features for diagnosing PDAC from the segmented objects. After that, the diagnosis and grading stages of PDAC are performed using the SVM model learned by the extracted features. This paper has several sections. Section 2 describes the pathological characteristics and the morphological features needed in diagnosing PDAC. Section 3 shows the configuration of systems used to diagnose PDAC and the segmentation methods of objects. Section 4 discusses the proposed new morphological features for quantitatively measuring the pathological characteristics of PDAC described in Section 2. In Section 5, the performances between SVM classifiers learnt by the proposed and existing classical morphological features are compared with each other to show the suitability of the proposed features to detect and grade the stages of PDAC. Section 6 evaluates and statistically analyzes the results. Finally, Section 7 presents a conclusion of this study.

2. Pathological Characteristics of PDAC

Pancreatic cancer is the second most common gastrointestinal neoplasm that causes death, after colon cancer [33]. And of all pancreatic neoplasms, PDAC accounts for 85–95%. Approximately 80% of all PDAC patients are between 60 and 80 years of age, and cases in people below the age of 40 are rare. The incidence of PDAC is about 50% higher in men than in women. By race, those of African ancestry have the highest rate of PDAC [34]. The best way of treating PDAC is known as curative resection. However, because of the rare possibility of diagnosing PDAC in the early stage, only 5–22% of PDAC patients can take the curative resection at the time the cancer is discovered [35]. Therefore, an accurate determination of the degree of cancer development is crucial factor for the treatment.

PDAC progression is divided by histological and cytological features and mitotic activity into Grade 1, well-differentiated carcinomas; Grade 2, moderately differentiated carcinomas; and Grade 3, poorly differentiated ductal adenocarcinomas [36, 37]. Grade 1 consists of a duct-like structure combined with medium-sized neoplastic glands. Tubular or cribriform patterns are typical. There may also be small irregular papillary projections without a distinct fibrovascular stalk, particularly in large duct-like structures. Mitotic activity is low. The mucin-producing neoplastic cells tend to be columnar, have eosinophils, and occasionally exhibit pale or even clear cytoplasm. Some neoplastic cell nuclei show loss of polarity. Grade 2 is characterized by a mixture of medium-sized duct-like and tubular structures of variable shapes, embedded in desmoplastic stroma. The duct shape is commonly that of incompletely formed glands. Compared with Grade 1, Grade 2 shows a greater variation in nuclear size, chromatin structure, and prominence of the nucleoli. The cytoplasm is usually slightly eosinophilic, but clear cells are occasionally abundant. Mucin production appears to be decreased, and intraductal *in situ* components
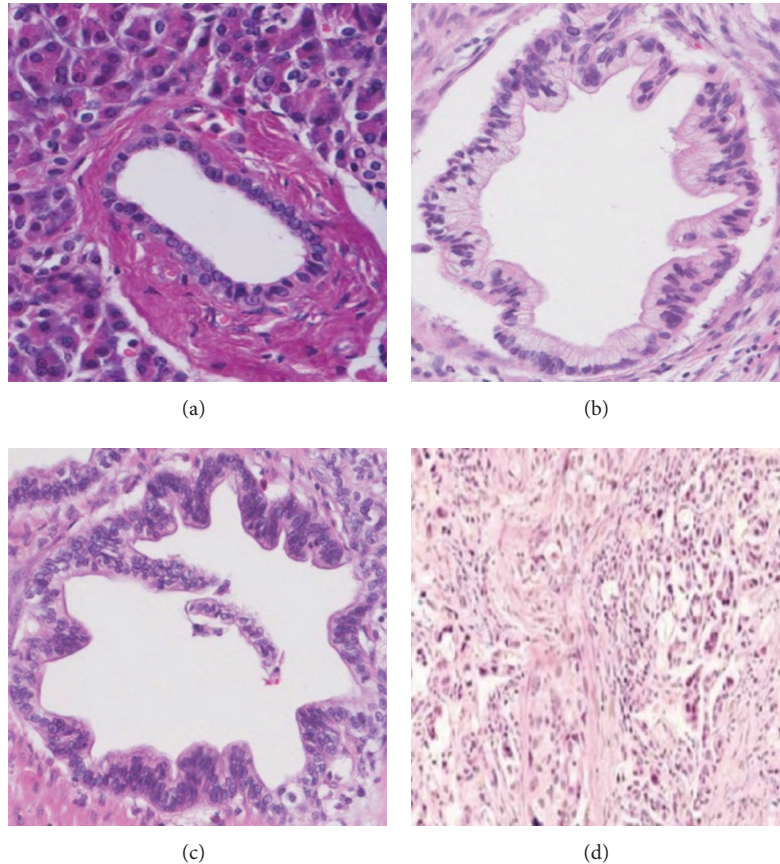
(a)

(b)

(c)

(d)

FIGURE 1: (a) Normal; (b) Grade 1, well differentiated; (c) Grade 2, moderately differentiated; (d) Grade 3, poorly differentiated.

are somewhat less frequent than in Grade 1. Grade 3 is infrequent. It is composed of a mixture of densely packed, small, and irregular glands as well as solid tumor cell sheets and nests that entirely replace the acinar tissue. While typical large, duct-like structures and intraductal tumor components are absent, there may be small squamoid features, spindle cells, or anaplastic foci. The neoplastic cells show marked pleomorphism, little or no mucin production, and brisk mitotic activity. Figure 1 shows Normal, Grade 1, Grade 2, and Grade 3 tissue images.

As described above, Grade 3 is not common. Also, Figure 1 shows that Grade 3 is certainly morphologically different from Grades 1 and 2. Therefore, this paper focuses on the detection of PDAC and the differentiation of Grades 1 and 2.

3. System Overview

In this paper, system configuration to diagnose PDAC consists of three phases as follows: segmentation and feature extraction, model learning and validation, and diagnosis. In the first phase, after preprocessing the given tissue image, the image is segmented into three parts. These three parts are the lumen region, epithelial nuclei, and nonepithelial nuclei. Then, according to the characteristics of each part, the features to be used for the classification model are extracted

and stored into a feature database. The second phase is the learning and validation of the SVM classification model using the features extracted in the previous step. The final phase carries out PDAC diagnosis for a tissue sample using the generated SVM classification model.

3.1. Segmentation for Major Interest Objects. In this section, we describe the method of segmenting three object types in a tissue image. Two of three object types are the lumen and epithelial nucleus constituting a duct. The last one is nonepithelial nucleus. Figure 2 shows the overall process of identifying three object types from a tissue image.

3.1.1. Lumen Segmentation. In this paper, the lumen of the tissue image is segmented by a seeded region growing (SRG) algorithm [38]. A beginning point should be designated for the use of the SRG. In the previous research [32], segmentation of the lumen region was automated by identifying candidate seed points within the lumen region. The proposed method is as follows. First of all, in order to facilitate the application of SRG, median filtering algorithm and background correction algorithm [39] are applied on a given image, and then maximum Entropy Threshold [40] is applied to produce a binary image $\mathbf{A}$. From the produced binary image $\mathbf{A}$, Direction Cumulative Map $H(\mathbf{A})$ is generated to
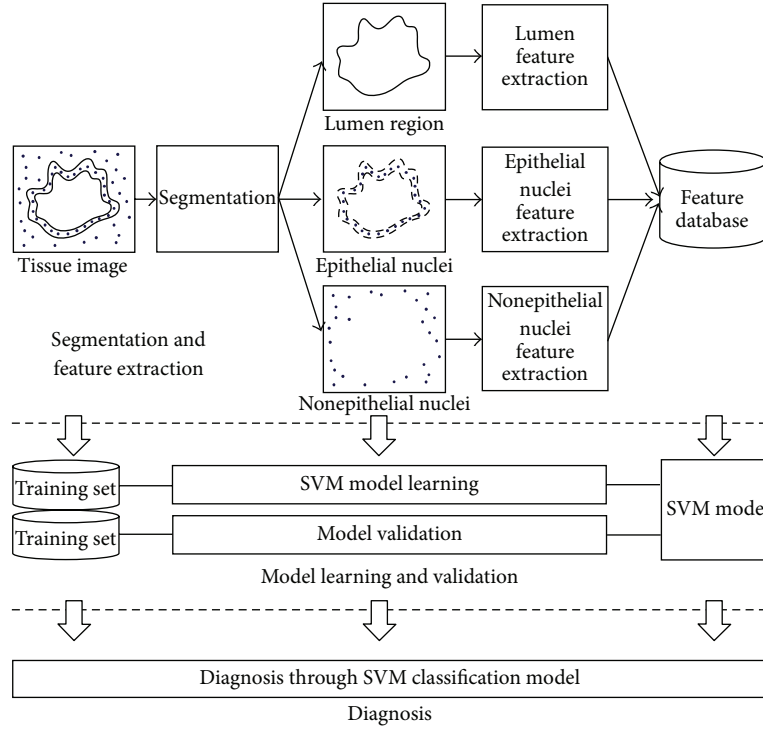
FIGURE 2: System overview for diagnosing PDAC.

find seed points. The $H(\mathbf{A})$ is generated by cumulating only white pixels of four directions (left, right, up, and down) of the binary image $\mathbf{A}$ and calculating the sum of the square root of the cumulated values. The $H(\mathbf{A})$ will have higher values around the central area of lumen region. Therefore, the local maximum points of $H(\mathbf{A})$ would be used as candidate seed points for the SRG algorithm. However, if candidate seed points are generated directly from $H(\mathbf{A})$, it might generate candidate seed points for unnecessarily narrow areas. As a solution for this problem, $H_T(\mathbf{A})$ with a threshold of lower value is, instead, employed to acquire candidate seed points. The *Otsu* method [41] is used to determine the threshold. With having the acquired candidate seed points set as a beginning point of parameter, the lumen region can be segmented by SRG algorithm. The boundary of segmented lumen region will be denotated as $\mathbf{B^O}$. Figure 3 shows the process of segmenting the lumen boundary that has been explained so far.

3.1.2. The Identification of Epithelial Nuclei and Nonepithelial Nuclei.
In this phase, we segmented the nuclei of tissue images and separate them into epithelial and nonepithelial nuclei. The process is as follows.

(1) Nuclei Segmentation. This step identifies all nuclei in a tissue image. First, the impurities shown on the tissue image are eliminated by median filter. Then, the color thresholding based on k-means [42] removes the parts such as cytoplasm and lumen that are unnecessary to identify nuclei. Next, the holes of the nuclei in the thresholded image are filled with a *hole filling* algorithm [43]. Finally, the nuclei are

separated using a *Watershed* algorithm [44]. Then, a set of the segmented nuclei are denoted by $\mathcal{N}$. Figure 4(a) shows the segmented nuclei.

(2) Division the Epithelial Nuclei and the Nonepithelial Nuclei. This step divides the segmented nuclei acquired in the previous step into the epithelial nuclei and the nonepithelial nuclei. Epithelial cells surround the lumen. Therefore, epithelial nuclei are identified by selecting the nearest nuclei to the lumen boundary, $\mathbf{B^O}$, from a set of nuclei $\mathcal{N}$. A set of epithelial nuclei is denoted by $\mathcal{N}^{\mathscr{E}}$ and defined as

$$\mathcal{N}^{\mathscr{E}} = \bigcup_{p \in \mathbf{B^O}} \left\{ n \,\middle|\, \underset{n \in \mathcal{N}}{\operatorname{argmin}} \left(\text{Distnace} \left(\text{Centroid}\,(n)\,, p \right) \right) \right\}, \tag{1}$$

where p is a point in $\mathbf{B^O}$ and the Centroid($\cdot$) is a function returning the center point of a given object. The Distance($\cdot,\cdot$) is a function returning the Euclidean distance between two given points. Segregation of epithelial nuclei from $\mathcal{N}$ is a procedure in which the nearest nucleus from each point $p \in \mathbf{B^O}$ is firstly selected, and the selected nucleus $\mathbf{n}$ is then included in a set of epithelial nuclei, $\mathcal{N}^{\mathscr{E}}$. Algorithm 1 presents this procedure. Figure 4(b) shows the selected epithelial nuclei (marked as red).

Nonepithelial nuclei are acquired by eliminating the identified epithelial nuclei $\mathcal{N}^{\mathscr{E}}$ from a set of nuclei $\mathcal{N}$ as follows:

$$\mathcal{N}^{\mathcal{N}\mathscr{E}} = \mathcal{N} - \mathcal{N}^{\mathscr{E}}. \tag{2}$$

(a)

(b)

(c)

(d)
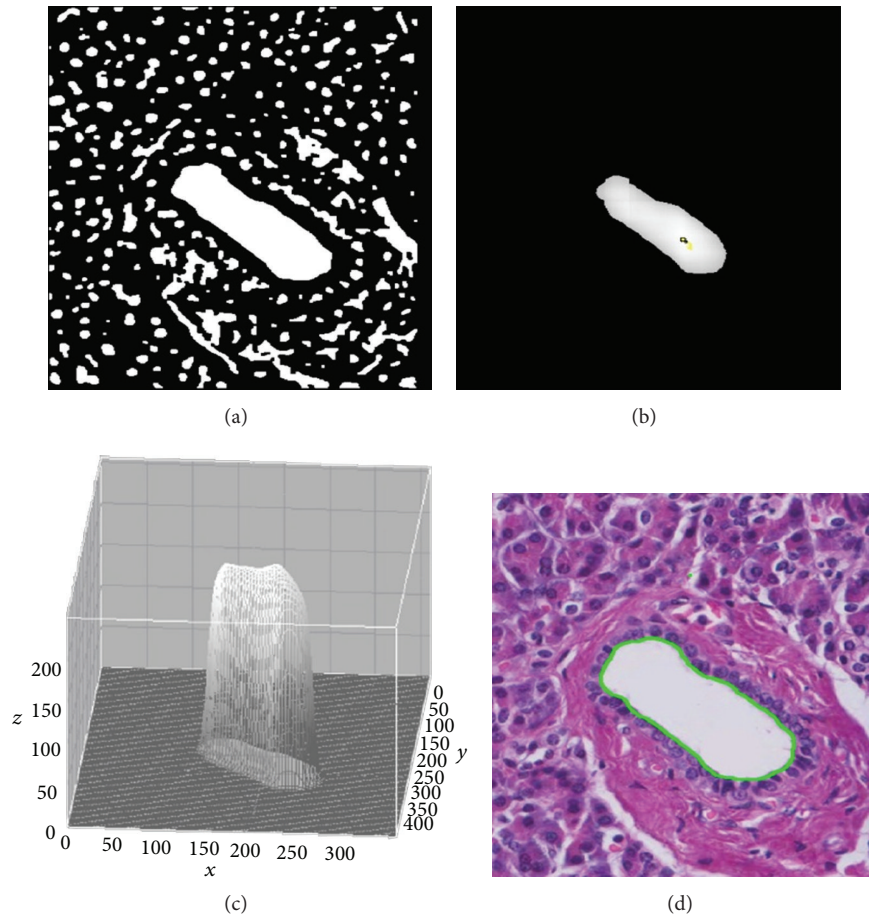
Figure 3: (a) The preprocessed binary image for identifying the lumen boundary; (b) $H_T(\mathbf{A})$ and candidate seed points (yellow points) for the image; (c) 3D plot for $H_T(\mathbf{A})$. It is scaled as a range from 0 to 255; (d) the boundary of lumen segmented by candidate seed (green line).

(a)

(b)

Figure 4: (a) A set of the segmented nuclei, $\mathcal{N}$; (b) a set of the epithelial nuclei selected from $\mathcal{N}$, $\mathcal{N}^g$ (marked as red).

Input: A set of nuclei $\mathcal{N}$, A original lumen boundary $\mathbf{B}^{\mathbf{O}}$
Output: A set of epithelial nuclei $\mathcal{N}^{\mathcal{E}}$
Method:
(1) $\mathcal{N}^{\mathcal{E}} \leftarrow \{\}$
(2) **for all** point p **in** $\mathbf{B}^{\mathbf{O}}$ **do**
(3) $n \leftarrow$ getNearestNuclus $(p, \mathcal{N})$
(4) $\mathcal{N}^{\mathcal{E}} \leftarrow \mathcal{N}^{\mathcal{E}} \cup \mathbf{n}$
(5) **end for**
(6) **return** $\mathcal{N}^{\mathcal{E}}$

ALGORITHM 1: Selection_Epithelial_Nuclei ($\mathcal{N}$).

TABLE 1: List of notations.

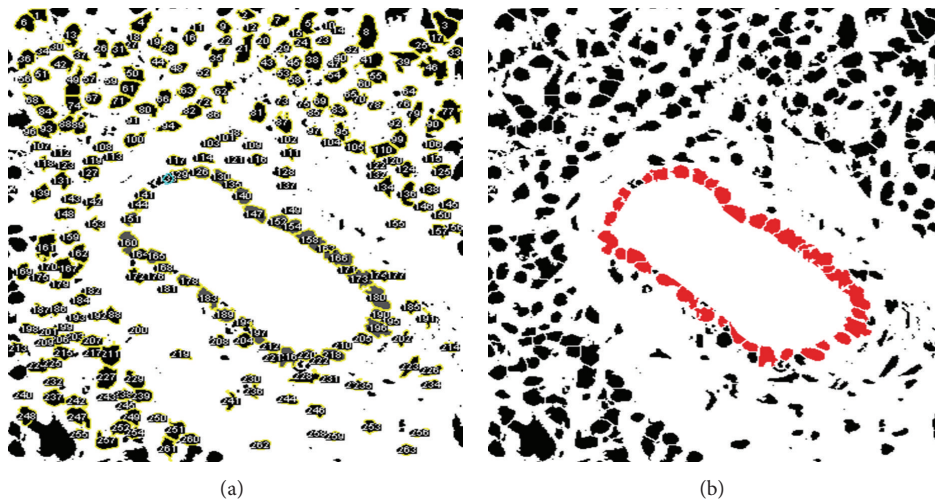| Symbol	Description
p	The p is a pair of numbers that are x and y coordinates $p = (x, y)$
$\mathbf{A}$	The matrix representation for preprocessed binary image
$H(\mathbf{A})$	The direction cumulative histogram for $\mathbf{A}$
$H_T(\mathbf{A})$	The thresholded $H(\mathbf{A})$
$\mathbf{B}^{\mathbf{O}}$	A sequence of points that consist of the original lumen boundary
$\mathbf{B}^{\mathbf{C}}$	A sequence of points for the convex hull of $\mathbf{B}^{\mathbf{O}}$
$\mathbf{B}^{\mathbf{I}}$	A sequence of points for the ideal lumen boundary that is estimated from $\mathbf{B}^{\mathbf{O}}$
$\mathcal{N}$	A set of nuclei $\mathcal{N} = \{n_1, n_2, \ldots, n_m\}$
$\mathcal{N}^{\mathcal{E}}$	A set of epithelial nuclei $\mathcal{N}^{\mathcal{E}} \subset \mathcal{N}$
$\mathcal{N}^{\mathcal{N}\mathcal{E}}$	A set of nonepithelial nuclei $\mathcal{N}^{\mathcal{N}\mathcal{E}} = \mathcal{N} - \mathcal{N}^{\mathcal{E}}$
Distance$(\cdot,\cdot)$	The Euclidean distance function
Centroid$(\cdot)$	The function returning center point of given object
$A(t)$	The atypia-amplitude function
$L(t)$	The perimeter from p_0 (start point) to p_t (tth point)
$\mathbf{AR}$	A set of atypia regions $\mathbf{AR} = \{ar_1, ar_2, \ldots, ar_m\}$
R^O	The region surrounded by $\mathbf{B}^{\mathbf{O}}$
R^C	The region surrounded by $\mathbf{B}^{\mathbf{C}}$
R^I	The region surrounded by $\mathbf{B}^{\mathbf{I}}$
Area $(\cdot)$	The function returning area of given region
$\mathbf{D}(\cdot)$	The data set for given feature set. $\mathbf{D}(\cdot) = \{(\mathbf{x}_1(\cdot), y_1), \ldots, (\mathbf{x}_m(\cdot), y_m)\}$, $\mathbf{x}(\cdot)$ is feature vector for given feature set and y_i is class label for $\mathbf{x}(\cdot)$

3.2. Notations. Table 1 summarizes the notations used in this paper.

4. Proposed Features

A major object examined to diagnose PDAC in a tissue sample is a duct. As described in Section 2, PDAC is classified as Grade 1, 2, or 3 by morphological changes of the duct which is composed of the lumen and the epithelial cells. Therefore, in this section, we propose methods to extract the specific morphological features of the segmented lumen and epithelial nuclei for PDAC diagnosis.

4.1. Lumen Features. In PDAC, a duct seems to have the shape of an incomplete gland with a papillary form and a loss of nucleus polarity. As the processing stage progresses, the shape

of the duct becomes more irregular with various atypia. In this subsection, the method representing atypia of a duct and the morphological features for measuring it are discussed.

4.1.1. Representing Atypia of Duct. Generally, a lumen of normal duct seems like a convex hull because atypia rarely appears. Unlike normal formation, as PDAC progresses, atypia of the lumen boundary in the duct becomes more and more irregular. From this standpoint, estimation of ideal lumen boundary of a given duct will be possible and portrayal of atypia of the original lumen will be feasible based on the boundary.

(1) Ideal Lumen Boundary. In this step, an original lumen boundary and an estimated ideal lumen boundary are represented as $\mathbf{B}^{\mathbf{O}}$ and $\mathbf{B}^{\mathbf{I}}$, respectively. $\mathbf{B}^{\mathbf{O}}$ and $\mathbf{B}^{\mathbf{I}}$ are, respectively,
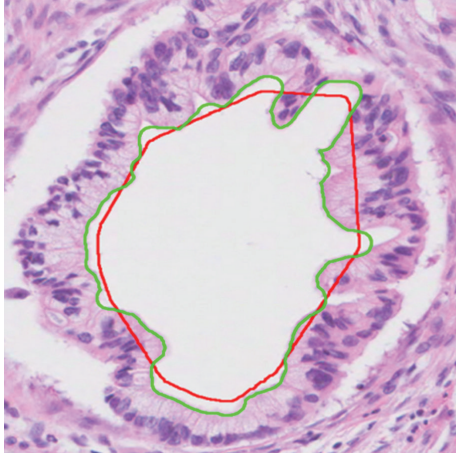
FIGURE 5: Original lumen boundary $\mathbf{B^O}$ (green line) and ideal lumen boundary $\mathbf{B^I}$ (red line).

a sequence of points constituting each lumen boundary. The procedure to estimate the ideal lumen boundary, $\mathbf{B^I}$, is as follows. First, the convex hull, $\mathbf{B^C}$, is obtained from the original lumen boundary $\mathbf{B^O}$. Then, because $\mathbf{B^C}$ is bigger than $\mathbf{B^O}$, the ideal lumen boundary, $\mathbf{B^I}$, is acquired through downsizing $\mathbf{B^C}$ to $\mathbf{B^O}$. The scaling factor, s, for scaling $\mathbf{B^C}$ to $\mathbf{B^O}$ is calculated as follows:

$$s = \sqrt{\frac{\text{Area}\left(R^O\right)}{\text{Area}\left(R^C\right)}}, \qquad (3)$$

where R^O represents an area bordering $\mathbf{B^O}$ while R^C does an area bordering $\mathbf{B^C}$. Area($\cdot$) is a function returning area of given region. Therefore, the ideal lumen boundary $\mathbf{B^I}$ is a sequence of points that consist of the boundary of the downsized region by the scaling factor s about the center of R^C. Figure 5 shows the original lumen boundary, $\mathbf{B^O}$, with a green line and the ideal lumen boundary, $\mathbf{B^I}$, with a red line.

(2) Atypia-Amplitude Signature. In this stage, 1D signature will be proposed as a means to depict atypia of a lumen employing an original lumen boundary and an ideal lumen boundary. The proposed 1D signature visualizes atypia of a lumen by measuring atypia-amplitude between an original lumen boundary and an ideal lumen boundary. The atypia-amplitude is an orthogonal distance with sign between an original lumen boundary and an ideal lumen boundary. It is measured by an atypia-amplitude function $A(t)$ as follows:

$$A(t) = \text{sgn}_t\left(q\right) \times \min_{p_t \in \mathbf{B^I}, \, q \in \mathbf{B^O}} \left(\text{Distance}\left(p_t, q\right)\right)$$
$$\text{such that } p_t \perp q, \qquad (4)$$

where t is an index variable, indicating an order of points within an ideal lumen boundary $\mathbf{B^I}$ and $\text{sgn}_t(q)$ is a function representing sign of a vertical distance between $p_t \in \mathbf{B^I}$ and $q \in \mathbf{B^O}$, which returns $+1$ or -1, respectively, when a point, q,

is located either outside or inside the boundary with a point p_t on $\mathbf{B^I}$. Figure 6(a) shows a process of measuring atypia-amplitude at $p_t \in \mathbf{B^I}$. In this example, $A(t)$ has positive value because q is located outside the boundary of p_t.

The 1D signature using the function, $A(t)$, is named as the atypia-amplitude signature. The atypia-amplitude signature is plotted with points of x and y. Then x-coordinate is the $L(t)$ meaning the perimeter from p_0 (starting point of $\mathbf{B^I}$) to $p_t \in \mathbf{B^I}$, and y-coordinate is the $A(t)$ implying atypia-amplitude at point p_t of $\mathbf{B^I}$. $L(t)$, the perimeter from p_0 (start point of $\mathbf{B^I}$) to p_t, is a sum of the Euclidean distances of each points within the given breadth. The $L(t)$ is as follows:

$$L(t) = \sum_{i=1}^{t} \text{Distance}\left(p_{i-1}, p_i\right), \qquad L(0) = 0. \qquad (5)$$

Figure 6 shows how to plot the atypia-amplitude signature by $A(t)$. Table 2 shows the atypia-amplitudes for Figures 1(a), 1(b), and 1(c).

4.1.2. Features for Measuring Atypia of Duct. This section introduces features that measure atypia of the lumen quantitatively using atypia-amplitude signature and the ideal lumen boundary developed in the previous phase. The proposed features are RMSAA (root-mean-squared atypia-amplitude), TSAV (total sum of atypia volatilities), AtypiaRatio, and #AtypiaRegions (it means the number of atypia regions).

(1) RMSAA (Root-Mean-Squared Atypia-Amplitude). RMSAA is measured by atypia-amplitude signature developed in the previous phase. It is the square root of the mean-squared atypia-amplitude (MSAA). $A(t)$, the vertical distance between $\mathbf{B^I}$ and $\mathbf{B^O}$, can be interpreted as the residual that represents the difference between the sample value of and the fitted value of the estimated regression model. Likewise, MSAA corresponds to a mean-squared error (MSE) [45] that measures the average squares error of the regression model. RMSAA is defined as follows:

$$\text{RMSAA} = \sqrt{\text{MSAA}} = \sqrt{\frac{1}{m} \sum_{t=1}^{m} A(t)^2}. \qquad (6)$$

Here, m is the number of points in $\mathbf{B^I}$.

(2) TSAV (Total Sum of Atypia Volatilities). Variation of the lumen boundary becomes more irregular since the lumen becomes more complex as PDAC develops. TSAV measures the degree of irregularity of the lumen boundary shown in the progress of PDAC. For the calculation of TSAV, major inflection points of atypia-amplitude signature are identified and the sum of diversion at those points is taken into account. In this paper, we use the Perceptual Important Point (PIP) method [46, 47] to find the major inflection points of the atypia-amplitude signature. The PIP method finds critical points that represent important trends of time series data. In this paper, the conventional PIP algorithm that detects a fixed number of PIPs has been modified to find all critical
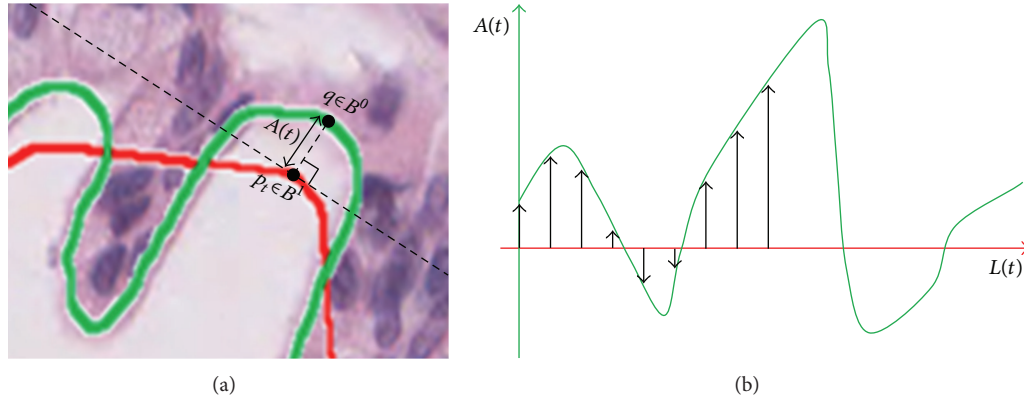
(a)

(b)

FIGURE 6: The atypia-amplitude signature with $A(t)$.

TABLE 2: The atypia-amplitude signature for Figures 1(a), 1(b), and 1(c).

points in the atypia-amplitude signature. The detail about the modified PIP algorithm is included in Appendices A and B. Figure 7 shows a part of atypia-amplitude signature of Grade 2 in Table 2 and PIPs observed by the modified PIP algorithm.

The TSAV is computed by (7) as the total sum of atypia volatilities (AVs) at PIPs. AV_i at a PIP, p_i, is defined by angle (θ_i) between two vectors, $\mathbf{a}_i = p_i - p_{i-1}$ and $\mathbf{b}_i = p_{i+1} - p_i$:

$$
\begin{aligned}
\text{TSAV} &= \sum_{i=2}^{m-1} \text{AV}_i \\
&= \sum_{i=2}^{m-1} \left(\arccos \left(\frac{\mathbf{a}_i \cdot \mathbf{b}_i}{\|\mathbf{a}_i\| \|\mathbf{b}_i\|} \right) \right) \\
&\quad (\mathbf{a}_i = p_i - p_{i-1}, \mathbf{b}_i = p_{i+1} - p_i).
\end{aligned}
\tag{7}
$$

Here, m is the number of PIPs detected from the atypia-amplitude signature.

(3) Atypia Ratio and the Number of Atypia Regions. The shape of duct becomes more complex and papillary becomes more vivid as PDAC develops. The original lumen region R^O of the developed PDAC does not fit into the ideal lumen region, R^I, extending beyond or contracting into the R^I. Thus, in this section, we measure AtypiaRatio and #AtypiaRegions (the number of atypia regions) to assess such characteristics. First, a set of atypia regions, $\mathbf{AR} = \{ar_1, ar_2, \ldots, ar_m\}$, is composed with regions which are generated by separating results of $(R^I \cup R^O) - (R^I \cap R^O)$ using $\mathbf{B^I}$. Figure 8 shows the identified atypia regions in Grade 2 tissue image.
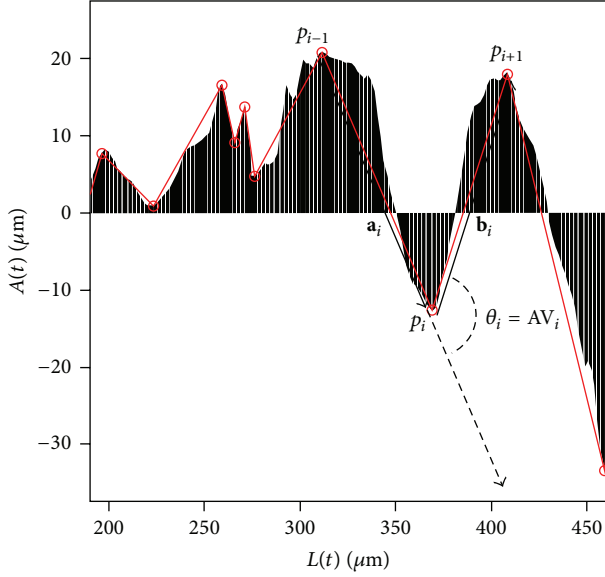
FIGURE 7: AV_i at a PIP, p_i, is measured as the angle (θ_i) between $\mathbf{a}_i$ and $\mathbf{b}_i$.

The obtained atypia regions are used to come up with AtypiaRatio and #AtypiaRegions:

$$\text{AtypiaRatio} = \frac{\sum_{i=1}^{m} \text{Area}(\text{ar}_i)}{\text{Area}(R^O)}$$

$$\text{\#AtypiaRegions} = |\{\text{ar} \,|\, \text{ar} \in \mathbf{AR}, \; \text{Area}(\text{ar}) > \text{Threshold}\}|,$$

(8)

where Area($\cdot$) is a function returning the size of a given region. The m means the cardinality of the $\mathbf{AR}$. The AtypiaRatio feature represents the overall degree of distortion within a duct, and the #AtypiaRegions feature quantitatively measures the papillary duct by counting the atypia regions. Then, small atypia regions representative of the papillary are excluded from counting by thresholding. The value for thresholding is set 300 μm^2 by consensus of pathologists at the Pathology Department of Yeongnam University.

4.2. Epithelial Cell Feature. Epithelium is another component composing a duct. In most cases, the epithelial cells of PDAC are the mucin-producing neoplastic cells that tend to be columnar, and their nuclei show loss of polarity [34, 48]. In this phase, we introduce the methods of extracting these features from the segmented epithelial nuclei.

(1) Cytoplasm Length. A duct of normal tissue is surrounded by cube-like epithelial cells. In the PDAC, a duct has columnar epithelium with abundant cytoplasm. The nuclei of columnar cells are oval-shaped. The cytoplasm length of columnar epithelium with abundant cytoplasm is longer than cuboidal epithelium. So, measuring the cytoplasm length of epithelial cells represents whether or not the epithelial cell trend is columnar. We proposed the feature, CytoplasmLength measuring the cytoplasm length of the epithelial cells

in [32]. The CytoplasmLength is the orthogonal distance between the epithelial nucleus and the original lumen boundary $\mathbf{B^O}$:

$$\text{CytoplasmLength}(n) = \text{Distance}(\text{Centroid}(n), q)$$
$$\text{such that, Centroid}(n) \perp q,$$

(9)

where q is a point in $\mathbf{B^O}$ that is orthogonal to Centroid(n). Figure 9 shows the measured CytoplasmLengths of Normal and Grade 1. In Figure 9, red regions are epithelial nuclei and green line is the identified lumen boundary. The blue lines between nuclei and lumen boundary are the measured CytoplasmLengths.

(2) The Standard Deviation of CytoplasmLength. This feature measures the loss of nuclear polarity that is one of the features of PDAC. The epithelial nuclei with the loss of nuclear polarity have a large deviation between cytoplasm lengths of them because epithelial nuclei are arranged irregularly along the lumen boundary. In contrast, the CytoplasmLengths of normal epithelium have a small deviation because the epithelial nuclei are arranged along the lumen boundary. Therefore, we measure the loss of nuclear polarity by calculating the standard deviation of the CytopalsmLength of epithelial nuclei:

$$\text{CytoplasmLength}_{\text{SD}}$$
$$= \sqrt{\frac{1}{m} \sum_{n \in \mathcal{N}^\mathcal{E}} \left(\overline{\text{CytoplasmLength}} - \text{CytoplasmLength}(n) \right)^2},$$

(10)

where m is the cardinality set $\mathcal{N}^\mathcal{E}$ and $\overline{\text{CytoplasmLength}}$ is the average of CytoplasmLength.

5. Experiments

5.1. Image Acquisition and Experimental Environment. We received 21 normal tissue slides and 26 PDAC tissue slides from the Pathology Department of Yeongnam University for our experiments. Those received tissue slides were stained via hematoxylin and eosin. Those tissue slides were scanned into digital slides using the *ScanScope CS System* [49] at 20x magnification. Each digital image of slides is variable depending on acquired tissue. Table 3 shows the information of digital slide.

In order to assess proposed features, we manually generated images for the experiments from these digital slides to make sure that each includes a duct. Each of the experiment images was formatted into a 24-bit tiff, and their size varied depending on the size of each duct. Important issue in diagnosis is inter- and intraobserver variability leading the diagnosis to be inconsistent, inaccurate, and biased [50]. A similar issue arises from ground truth data of experts for configuration of and performance assessment of a diagnosis system. A feasible way of reducing variability issue of ground truth is to construct ground truth with participation of several experts [51]. This article has three pathologists of the
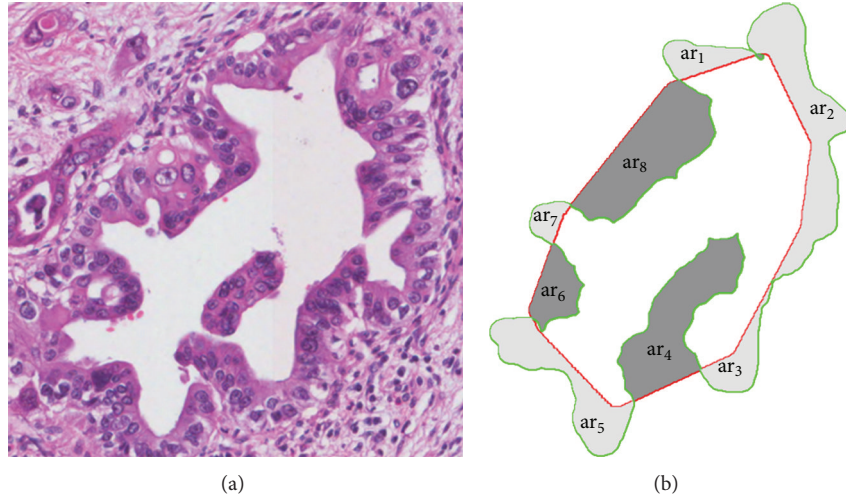
FIGURE 8: (a) Grade 2 tissue image, (b) atypia regions that are generated by region R^O and region R^I of (a).

$\overline{\text{CytoplasmLength}}$ = 3.64 m; CytoplasmLength$_{\text{SD}}$ = 0.86

(a)

$\overline{\text{CytoplasmLength}}$ = 8.64 μm; CytoplasmLength$_{\text{SD}}$ = 3.00
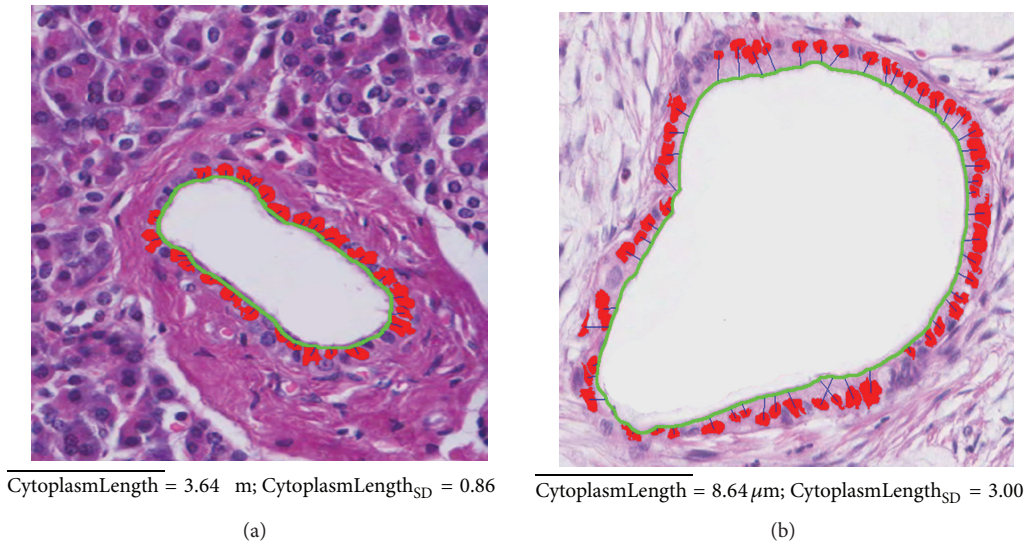
(b)

FIGURE 9: CytoplasmLengths (blue lines) of epithelial cells for Normal and Grade 1.

Pathology Department of Yeongnam University participated in an assessment of ground truth. Each image of duct generated from digital slides has been labeled into class with discreet consensus of those three pathologists. Table 4 presents the number of experiment images labeled by the experts for the assessment of ground truth.

We segmented the given tissue images into three parts (lumen, epithelial nuclei, and nonepithelial nuclei) and extracted existing classical morphological features with the proposed features from each part. Table 5 shows the features used in the experiments for diagnosing PDAC. The features extracted from each segment are asterisked. 1~12 rows in Table 5 are the existing classical features [5, 24, 39, 52]. 13~ 18 rows are the proposed features in this paper and our previous study [32]. Because a number of epithelial and nonepithelial cells were found in a captured tissue image, the features of each object were extracted and then averaged

TABLE 3: Information of the digital slide.

Digital slide size	Variable size
Image resolution	0.492 μm/pixel
Image type	SVS/JPEG2000
Image channels	3
Image bit depth	8 bits
Magnification	20x
Organization	Tiled
Tile width	240 pixels
Tile height	240 pixels

to represent features of the tissue image. The experiment environment for feature extraction was performed on a computer with an AMD Athelon II 3 GHZ CPU and 2 G

TABLE 4: The obtained experimental images.

Type	Number of images
Normal	80
Grade 1	80
Grade 2	80

RAM running Windows7 64 bit. The existing and proposed methods for extracting features were implemented by using ImageJ [53], an image processing package based on the JAVA programming language.

5.2. Experiment Design. We compared the performance of the classifiers learned by the classical and proposed features to demonstrate the quality of morphological features that are proposed to diagnose PDAC. In this paper, SVM, a well-founded learning technique based on statistical learning theory [54], was employed as the learning method of the classifier. The SVM shows good generalized performance, because it minimizes the combination of the empirical risk and the VC (Vapnik-Cheronenkis) dimension [55].

The experiment evaluated the classification performance for two cases: classification between Normal and PDAC tissues and classification between Grade 1 and Grade 2 of PDAC. To measure how the proposed features improve the accuracy of classification, the classifiers were learned by feature sets that are configured as existing classical, proposed, and combination features for the three segmented objects (lumen, epithelial nuclei, and nonepithelial nuclei). Table 6 shows the symbol and dimension for the configured feature sets used in classification experiments.

Experiment data were generated according to each feature set in Table 6 for the experiments in two cases (Normal versus PDAC and Grade 1 versus Grade 2). Thus, for the experiments to diagnose PDAC in the first case, 13 data sets for 13 feature sets were generated as follows: $\mathbf{D}(\text{CLF})$, $\mathbf{D}(\text{PLF})$, $\mathbf{D}(\text{ALF})$, $\mathbf{D}(\text{CEF})$, $\mathbf{D}(\text{PEF})$, $\mathbf{D}(\text{AEF})$, $\mathbf{D}(\text{CNF})$, $\mathbf{D}(\text{CDF})$, $\mathbf{D}(\text{PDF})$, $\mathbf{D}(\text{ADF})$, $\mathbf{D}(\text{CTF})$, $\mathbf{D}(\text{PTF})$, and $\mathbf{D}(\text{ATF})$. $\mathbf{D}(\cdot)$ is the data set that is configured by the given feature set as a parameter. It is denoted as $\mathbf{D}(\cdot) = \{(\mathbf{x}(\cdot)_1, y_1), \ldots, (\mathbf{x}(\cdot)_m, y_m)\}$, where $\mathbf{x}(\cdot)_i$ is the ith feature vector corresponding to a given feature set of parameters (symbols of Table 6) and y_i is its class label. The y_i is either -1, which is Normal, or 1, which is PDAC.

Similarly, for the experiments to grade stages of PDAC in the second case, 13 data sets were generated. In these data sets, $y_i = -1$ means Grade 1 whereas $y_i = 1$ means Grade 2. To evaluate the performance of the SVM classifier for each feature set, we configured a training set and test set from the generated data set for the feature set. The ratio of training set to test a set was 60 to 40. In the first experiment (Normal versus PDAC), the experiment data set of PDAC is configured by sampling 80 data from 160 PDAC data of either Grade 1 or Grade 2. Because the number of Normal sample is 80, we limit the number of PDAC for fair evaluation of classifiers. Table 7 shows the number of training and test data sets used in the experiments of two cases.

In this paper, SVM classifiers used the soft margin method and RBF kernel [54]. Therefore, model parameter

C and kernel parameter γ are required. The optimal classifier parameters (C^*, γ^*) in which the classification accuracy for 10 cross-validation [56] in the training set is maximized are selected from parameter pairs of $(C, \gamma) \in \{10^{-1}, 10^{-0.5}, \ldots, 10^4\} \times \{2^{-5}, 2^{-4.5}, \ldots, 2^0\}$ by *Grid Search* [57].

The number of experiment images used in this study is insufficient. Therefore, the classification accuracy of the generated model might be biased [58]. In statistics, to solve this problem, the bootstrap resampling technique [59] is used. We used the bootstrap resampling technique for the unbiased evaluation of classifiers for each feature set. First, we generated 10 training sets and 10 testing sets from data set $\mathbf{D}(\cdot)$ corresponding to a given feature set for bootstrap evaluation (refer to Table 6). Therefore, the classification performance for each feature set is measured by averaging evaluation results of individually optimized classifiers for 10 training sets and 10 testing data sets. The performance measures used in the experiment are as follows: true positive (TP), true negative (TN), false positive (FP), false negative (FN), sensitivity (SN), Specificity (SP), positive predictive value (PPV), negative predictive value (NPV), and accuracy (ACC). The descriptions for TP, TN, FP, and FN are explained in each experiment. The rest of performance measures are defined as follows:

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

$$SP = \frac{TN}{TN + FP},$$

$$PPV = \frac{TP}{TP + FP}, \tag{11}$$

$$NPV = \frac{TN}{TN + FN},$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}.$$

5.3. Experimental Results

Case 1 (Normal versus PDAC). Table 8 and Figure 10 show the bootstrap evaluation of classifier learned by each feature set, for distinguishing between the Normal and PDAC. The standard deviations of the evaluation results are displayed in parentheses. First, we will compare of classifiers that are learned by feature set extracted from lumen object. In this comparison, the accuracy of classifier learnt by PLF is 91.56% which is about 18% higher than the classification accuracy with CLF (73.44%). The classification accuracy with ALF feature set configured as combination of CLF and PLF has, contrarily, decreased to 87.35%. The results showed that the PLF feature set is more suitable for diagnosing PDAC than classifiers with CLF. However, no improvement in the performance of the classifier with the ALF that is a combination of PLF and CLF was revealed in the results.

Secondly, we will compare classifiers with feature set extracted from epithelial nuclei objects. PEF and AEF show the same accuracy of 87.50%. However, in case of a classifier with AEF, the standard deviation of accuracy was observed

TABLE 5: Morphological features used in the classification experiment.

No.	Feature	Description	NEN[1]	Lumen	EN[2]
1	Area	Area of selection in square pixels	*	*	*
2	Perimeter	The length of the outside boundary of the selection	*	*	*
3	Width	Width of the smallest rectangle enclosing the selection	*	*	*
4	Height	Height of the smallest rectangle enclosing the selection	*	*	*
5	MajorAxis	Major (primary) axis length of the best fitting ellipse	*	*	*
6	MinorAxis	Minor (secondary) axis length of the best fitting ellipse	*	*	*
7	Circularity	$4\pi \times (\text{Area}/\text{Perimeter}^2)$. A value of 1.0 indicates a perfect circle	*	*	*
8	Feret's diameter	The longest distance between any two points along the selection boundary	*	*	*
9	AspectRatio	MajorAxis/MinorAxis	*	*	*
10	Skewness	The third order moment about the mean	*	*	*
11	Roundness	$4 \times \text{Area}/(\pi \times \text{MajorAxis}^2)$. The measure of the sharpness of particle's edge and corners	*	*	*
12	Solidity	Area/ConvexArea	*	*	*
13	RMSAA	Root-mean-squared atypia-amplitude	*		
14	TSAV	Total sum of atypia volatilities for PIPs	*		
15	AtypiaRatio	The ratio of atypia region	*		
16	#AtypiaRegions	The number of atypia regions for identifying papillary	*		
17	CytoplasmLength	The cytoplasm length of epithelial nucleus [32]			*
18	CytoplasmLength$_{SD}$	The standard deviation of CytoplasmLength			*

NEN[1]: nonepithelial nuclei; EN[2]: epithelial nuclei.

TABLE 6: The symbol and the dimension of feature sets that are configured by the object features and combined features.

Object	Symbol	Feature set	Dimension
NEN	CNF	Classical nonepithelial nuclei features	12
Lumen	CLF	Classical lumen features	12
	PLF	Proposed lumen features	4
	ALF	CLF + PLF	16
EN	CEF	Classical epithelial nuclei features	12
	PEF	Proposed epithelial nuclei features	2
	AEF	CEF + PEF	14
Duct	CDF	Classical duct features (CLF + CEF)	24
	PDF	Proposed duct features (PLF + PEF)	6
	ADF	CDF + PDF	30
Tissue	CTF	Classical features extracted from three objects (CLF + CEF + CNF)	36
	PTF	Proposed features extracted from three objects (PLF + PEF + CNF)	18
	ATF	All features (CLF + CEF + PLF + PEF + CNF)	42

as 1.47 which is far stable than PEF of 3.21. Interestingly, AEF is a combination of the PEF and the CEF, its accuracy is dependent on the proposed PEF. Also, the notable point is that the feature dimension of PEF is only two. These results do not only show that the PEF is very suitable for identifying the PDAC but they also prove their effectiveness in diagnostic cost aspect.

In diagnosis of PDAC, a duct that is composed of lumen and epithelium is an important region. Therefore, we thought that the experiment with the combination of features extracted from lumen and epithelial nuclei is of very meaning. For this, we prepared three combination feature sets which include CDF(CLF + CEF), PDF(PLF + PEF), and ADF(CDF + PDF). With these sets, we performed classification. In these experiments, the classification accuracy is improved to 94.38% when the PDF feature set was used. The classification accuracy was measured about 3~7% higher than PLF and AEF (or PEF) that showed best classification performance in each object. It showed that combination of lumen and epithelial nuclei features helps diagnose PDAC.

Consequently, the experimental results of CTF, PTF, and ATF using combination of the feature sets extracted from three objects in a tissue image depended on experimental results conducted in duct object. Thus, it is implied that there are no improvements as a result of combining all features. Further, the experiment with PTF that consists of PDF and CNF showed 2% lower in its accuracy than PDF alone.

Subsequently, ROC (receiver-operating characteristic) analysis with regard to classifiers learned by each feature set was performed. ROC analysis is being widely used in

TABLE 7: The number of training and testing data sets for learning and evaluating the SVM classifier.

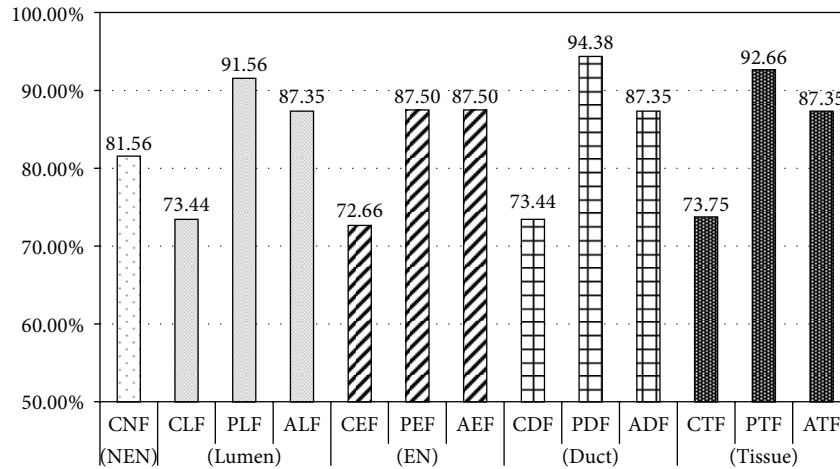| Experiment	Class	Number of training data	Number of testing data
Case 1: Normal versus PDAC	Normal	48	32
	PDAC	48	32
Case 2: Grade 1 versus Grade 2	Grade 1	48	32
	Grade 2	48	32

FIGURE 10: Comparison of classification accuracy for the Normal and PDAC for each feature set.

medical study as a benchmark of accuracy and comparison of diagnosis. ROC analysis examines ROC curves drawn by TP rate (Sensitivity) and FP rate (1-Sensitivity). The examination presents diagnosis accuracy with area under the ROC curve. Swets classified the degree of the accuracy, according to the value of AUC (area under the ROC curve), into noninformative (AUC = 0.5), less accurate ($0.5 < AUC \leq 0.7$), moderately accurate ($0.7 < AUC \leq 0.9$), highly accurate ($0.9 < AUC < 1$), and perfect tests (AUC = 1) [60, 61]. In other words, as ROC curve approaches to left hand corner, the accuracy is interpreted as higher. Figure 11 shows the average ROC graph and value of AUC of 13 classifiers learnt by features of each object.

In the ROC analysis, a classifier learned by PDF displays the highest value of AUC with 0.96. It proves the meaningfulness of combination of epithelial nuclei and lumen features likewise the performance evaluation of the classifier. Subsequent to PDF, a classifier with PEF shows a slightly higher AUC value of 0.94 than that of PLF with 0.93. It is interpreted that classifiers learned by features inclusive of the proposed PDF and PEF show fairly accurate diagnosis with AUC value of above 0.9.

Overall, experiments including proposed feature set show better performance than classifiers with classical feature set. As mentioned in Section 2, the experiments showed that the duct is an important region in diagnosing PDAC. PDF that is composed of PLF and PEF has led to improvement of classifier performance. Also, the classifiers with the proposed PLF and PEF extracted from lumen and epithelial nuclei, respectively, show higher performance than classical feature

sets, CLF, CEF, and CNF. It implies that simple morphological features such as Area and Perimeter, are inadequate for finding complicated characteristic of PDAC.

Case 2 (Grade 1 versus Grade 2). In this step, we distinguished between two stages, Grade 1 and Grade 2, of PDAC. As with the experiments in Case 1, we generated 10 training and testing sets for Grade 1 and Grade 2, and we evaluated classification accuracy. The results are shown in Table 9 and Figure 12. The classification results of each feature sets for distinguishing between Grade 1 and Grade 2 show lower classification accuracy than the experiments of distinguishing between Normal and PDAC. Firstly, in the experiments with the lumen object, the classification accuracy with PLF (77.03%) was about 19% higher than that with CLF (57.97%). In particular, specificity of PLF (79.69%) is measured 34% higher than CLF (45.94%), showing performance gain by 73%. In the experiment with ALF, combination of CLF and PLF, the classifier accuracy showed poorer classification performance than the classifier learnt by PLF.

Next, the experiments with the feature sets extracted from epithelial nuclei showed lower accuracy in classification than the experiments with epithelial cells in Case 1. As epithelium cells in all stages of Grade 1 and Grade 2 of PDAC showed columnar and loss of polarity characteristics, distinguishing of stages through them is difficult. Nevertheless, accuracy of the classifier with PEF (70.78%) using only two features increased by about 14% compared to CEF (56.41%). In an experiment using AEF, the accuracy was measured lower than PEF, indicating no performance improvement through the combination of CEF and PEF.
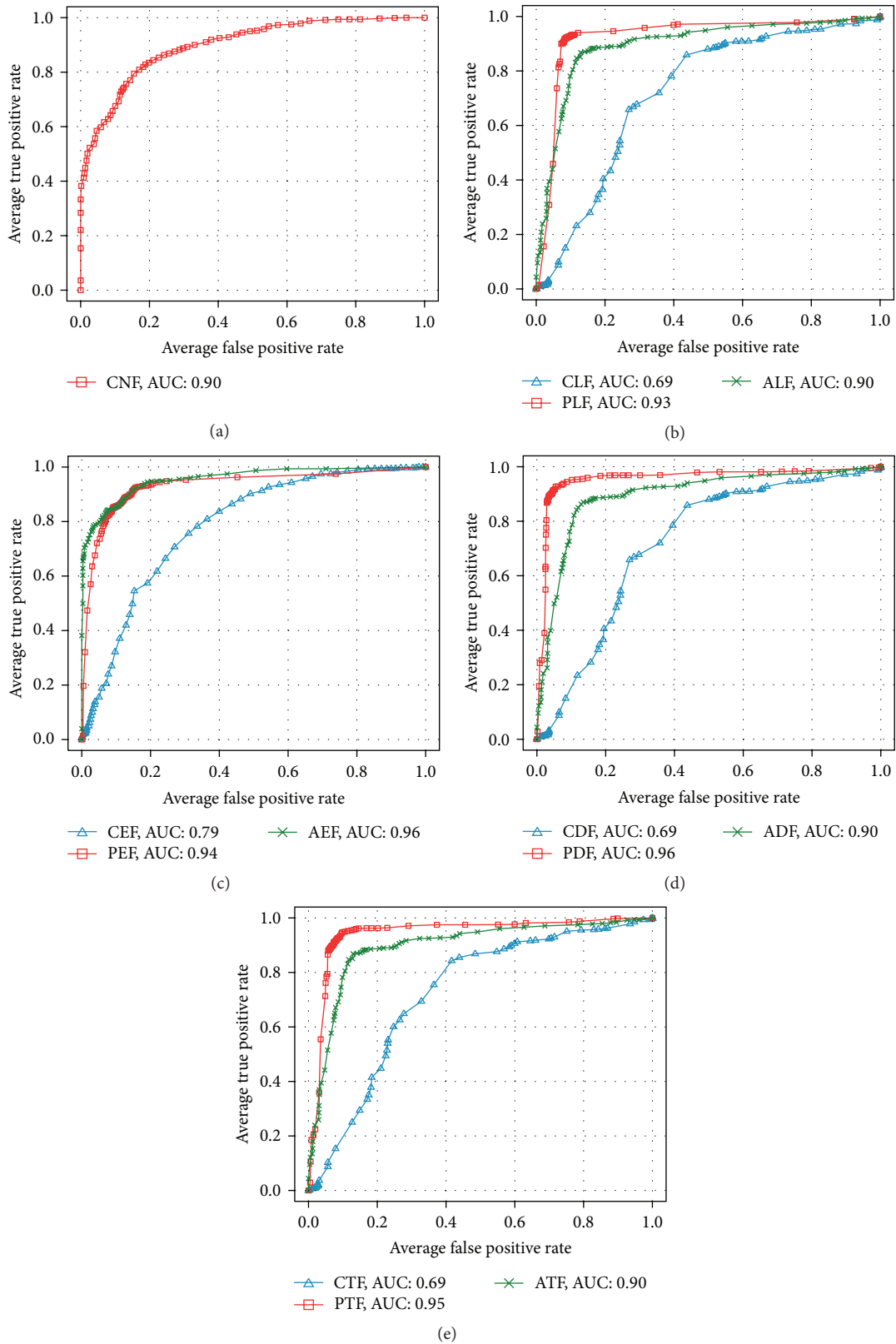
(a)

(b)

(c)

(d)

(e)

FIGURE 11: Comparison of ROC curves and AUC values for classifiers in Case 1.

TABLE 8: Evaluation results for distinguishing Normal and PDAC to each feature set.

Object	Feature set	TN	FP	FN	TP	SN (%)	SP (%)	PPV (%)	NPV (%)	ACC (%)
NEN	CNF	26.50	6.30	5.50	25.70	82.82	80.32	81.07	82.50	81.56
		(1.58)	(2.11)	(1.58)	(2.11)	(4.94)	(6.60)	(5.17)	(4.22)	(3.95)
Lumen	CLF	27.30	12.30	4.70	19.70	85.32	61.57	69.38	81.72	73.44
		(2.21)	(3.27)	(2.21)	(3.27)	(6.92)	(10.21)	(4.22)	(6.25)	(3.13)
	PLF	29.30	2.70	2.70	29.30	91.57	91.57	91.99	91.88	91.56
		(1.77)	(2.00)	(1.77)	(2.00)	(5.52)	(6.26)	(5.45)	(4.70)	(3.14)
	ALF	27.50	3.60	4.50	28.40	85.94	88.75	88.96	86.36	87.35
		(0.85)	(2.46)	(0.85)	(2.46)	(2.66)	(7.68)	(6.92)	(2.12)	(3.57)
EN	CEF	25.10	10.60	6.90	21.40	78.44	66.88	70.72	75.75	72.66
		(1.66)	(2.84)	(1.66)	(2.84)	(5.20)	(8.86)	(4.97)	(3.99)	(3.84)
	PEF	27.50	3.50	4.50	28.50	85.94	89.07	89.52	86.67	87.50
		(1.51)	(2.95)	(1.51)	(2.95)	(4.72)	(9.23)	(6.80)	(3.25)	(3.21)
	AEF	27.00	3.00	5.00	29.00	84.38	90.63	90.55	85.73	87.50
		(2.05)	(1.89)	(2.05)	(1.89)	(6.42)	(5.89)	(5.01)	(4.53)	(1.47)
Duct	CDF	27.30	12.30	4.70	19.70	85.32	61.57	69.38	81.72	73.44
		(2.21)	(3.27)	(2.21)	(3.27)	(6.92)	(10.21)	(4.22)	(6.25)	(3.13)
	PDF	29.80	1.40	2.20	30.60	93.13	95.63	95.78	93.50	94.38
		(1.40)	(1.51)	(1.40)	(1.51)	(4.37)	(4.71)	(4.34)	(3.80)	(2.35)
	ADF	27.50	3.60	4.50	28.40	85.94	88.75	88.96	86.36	87.35
		(0.85)	(2.46)	(0.85)	(2.46)	(2.66)	(7.68)	(6.92)	(2.12)	(3.57)
Tissue	CTF	26.90	11.70	5.10	20.30	84.06	63.44	70.34	80.92	73.75
		(2.28)	(3.74)	(2.28)	(3.74)	(7.13)	(11.70)	(5.32)	(6.11)	(3.44)
	PTF	29.80	2.50	2.20	29.50	93.13	92.19	92.74	93.26	92.66
		(1.32)	(2.22)	(1.32)	(2.22)	(4.11)	(6.95)	(6.15)	(3.76)	(3.21)
	ATF	27.50	3.60	4.50	28.40	85.94	88.75	88.96	86.36	87.35
		(0.85)	(2.46)	(0.85)	(2.46)	(2.66)	(7.68)	(6.92)	(2.12)	(3.57)

TP (true positive): the number of PDACs that are correctly classified as PDACs.
FP (false positive): the number of Normals that are incorrectly classified as PDACs.
FN (false negative): the number of PDACs that are incorrectly classified as Normals.
TN (true negative): the number of Normals that are correctly classified as Normals.
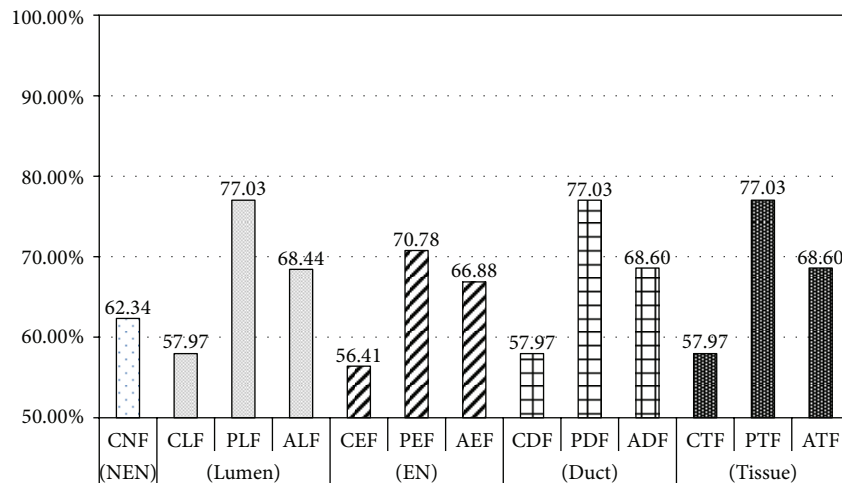

FIGURE 12: Comparison of classification accuracy for distinguishing between Grade 1 and Grade 2 for each feature set.

Table 9: Evaluation results for distinguishing between Grade 1 and Grade 2 to each feature set.

Object	Feature set	TN	FP	FN	TP	SN (%)	SP (%)	PPV (%)	NPV (%)	ACC (%)
NEN	CNF	19.60	11.70	12.40	20.30	61.25	63.44	64.35	63.29	62.34
		(6.26)	(4.97)	(6.26)	(4.97)	(19.56)	(15.52)	(8.93)	(9.49)	(7.78)
Lumen	CLF	22.40	17.30	9.60	14.70	70.00	45.94	58.39	66.13	57.97
		(5.99)	(6.95)	(5.99)	(6.95)	(18.70)	(21.70)	(10.63)	(16.21)	(5.23)
	PLF	23.80	6.50	8.20	25.50	74.38	79.69	79.65	75.40	77.03
		(1.32)	(3.95)	(1.32)	(3.95)	(4.11)	(12.35)	(10.15)	(4.96)	(6.83)
	ALF	25.20	13.40	6.80	18.60	78.75	58.13	67.05	73.46	68.44
		(2.90)	(6.31)	(2.90)	(6.31)	(9.06)	(19.72)	(9.88)	(5.85)	(7.39)
EN	CEF	18.60	14.50	13.40	17.50	58.13	54.69	56.45	57.03	56.41
		(4.30)	(4.14)	(4.30)	(4.14)	(13.44)	(12.95)	(7.17)	(9.10)	(7.53)
	PEF	22.20	8.90	9.80	23.10	69.38	72.19	71.89	70.04	70.78
		(1.14)	(2.92)	(1.14)	(2.92)	(3.55)	(9.14)	(6.94)	(4.08)	(5.26)
	AEF	22.40	11.60	9.60	20.40	70.00	63.75	65.95	68.16	66.88
		(2.32)	(2.27)	(2.32)	(2.27)	(7.25)	(7.10)	(5.42)	(5.84)	(5.50)
Duct	CDF	22.40	17.30	9.60	14.70	70.00	45.94	58.39	66.13	57.97
		(5.99)	(6.95)	(5.99)	(6.95)	(18.70)	(21.70)	(10.63)	(16.21)	(5.23)
	PDF	23.80	6.50	8.20	25.50	74.38	79.69	79.65	75.40	77.03
		(1.32)	(3.95)	(1.32)	(3.95)	(4.11)	(12.35)	(10.15)	(4.96)	(6.83)
	ADF	25.10	13.20	6.90	18.80	78.44	58.75	67.25	73.40	68.60
		(2.88)	(6.21)	(2.88)	(6.21)	(9.02)	(19.42)	(9.75)	(5.88)	(7.31)
Tissue	CTF	22.40	17.30	9.60	14.70	70.00	45.94	58.39	66.13	57.97
		(5.99)	(6.95)	(5.99)	(6.95)	(18.70)	(21.70)	(10.63)	(16.21)	(5.23)
	PTF	23.80	6.50	8.20	25.50	74.38	79.69	79.65	75.40	77.03
		(1.32)	(3.95)	(1.32)	(3.95)	(4.11)	(12.35)	(10.15)	(4.96)	(6.83)
	ATF	25.10	13.20	6.90	18.80	78.44	58.75	67.25	73.40	68.60
		(2.88)	(6.21)	(2.88)	(6.21)	(9.02)	(19.42)	(9.75)	(5.88)	(7.31)

TP (true positive): the number of Grade 2s that are correctly classified as Grade 2s.
FP (false positive): the number of Grade 1s that are incorrectly classified as Grade 2s.
FN (false negative): the number of Grade 2s that are incorrectly classified as Grade 1s.
TN (true negative): the number of Grade 1s that are correctly classified as Grade 1s.

Table 10: The statistics for nonepithelial nuclei features and the results of F-test.

Lumen features	Normal (df = 79) $\overline{N}$ (C.I[a])	Grade 1 (df = 79) $\overline{G1}$ (C.I)	Grade 2 (df = 79) $\overline{G2}$ (C.I)	F-tests ($\mu_N = \mu_{G1} = \mu_{G2}$) F-value
Area (μm^2)	$3.68E + 01$ ($\pm 1.57E + 00$)	$4.09E + 01$ ($\pm 1.67E + 00$)	$3.84E + 01$ ($\pm 1.57E + 00$)	$1.13E + 01^*$
Perimeter (μm)	$2.66E + 01$ ($\pm 6.35E - 01$)	$2.95E + 01$ ($\pm 7.09E - 01$)	$2.84E + 01$ ($\pm 6.67E - 01$)	$3.30E + 01^*$
Width (μm)	$7.50E + 00$ ($\pm 1.85E - 01$)	$8.48E + 00$ ($\pm 2.34E - 01$)	$7.97E + 00$ ($\pm 2.18E - 01$)	$3.64E + 01^*$
Height (μm)	$7.92E + 00$ ($\pm 2.20E - 01$)	$8.46E + 00$ ($\pm 2.57E - 01$)	$8.31E + 00$ ($\pm 2.41E - 01$)	$9.47E + 00^*$
MajorAxis (μm)	$8.58E + 00$ ($\pm 1.76E - 01$)	$9.68E + 00$ ($\pm 1.78E - 01$)	$9.28E + 00$ ($\pm 1.83E - 01$)	$6.78E + 01^*$
MinorAxis (μm)	$5.37E + 00$ ($\pm 1.20E - 01$)	$5.28E + 00$ ($\pm 1.35E - 01$)	$5.17E + 00$ ($\pm 1.28E - 01$)	$4.17E + 00$
Circularity	$6.67E - 01$ ($\pm 1.08E - 02$)	$6.01E - 01$ ($\pm 1.17E - 02$)	$6.10E - 01$ ($\pm 1.24E - 02$)	$6.63E + 01^*$
Feret's diameter (μm)	$9.63E + 00$ ($\pm 2.02E - 01$)	$1.09E + 01$ ($\pm 2.28E - 01$)	$1.04E + 01$ ($\pm 2.12E - 01$)	$5.80E + 01^*$
Skewness	$3.35E - 01$ ($\pm 1.03E - 01$)	$4.23E - 01$ ($\pm 2.02E - 01$)	$4.00E - 01$ ($\pm 1.52E - 01$)	$5.85E - 01$
AspectRatio	$1.66E + 00$ ($\pm 3.58E - 02$)	$1.98E + 00$ ($\pm 6.07E - 02$)	$1.91E + 00$ ($\pm 5.62E - 02$)	$7.55E + 01^*$
Roundness	$6.50E - 01$ ($\pm 1.14E - 02$)	$5.73E - 01$ ($\pm 1.25E - 02$)	$5.84E - 01$ ($\pm 1.21E - 02$)	$8.30E + 01^*$
Solidity	$8.38E - 01$ ($\pm 4.40E - 03$)	$8.19E - 01$ ($\pm 4.58E - 03$)	$8.19E - 01$ ($\pm 4.81E - 03$)	$3.96E + 01^*$

C.I[a]: confidence interval.
*It indicates features whose null hypothesis was rejected with F-value > $F_{0.01}(2,237)$.
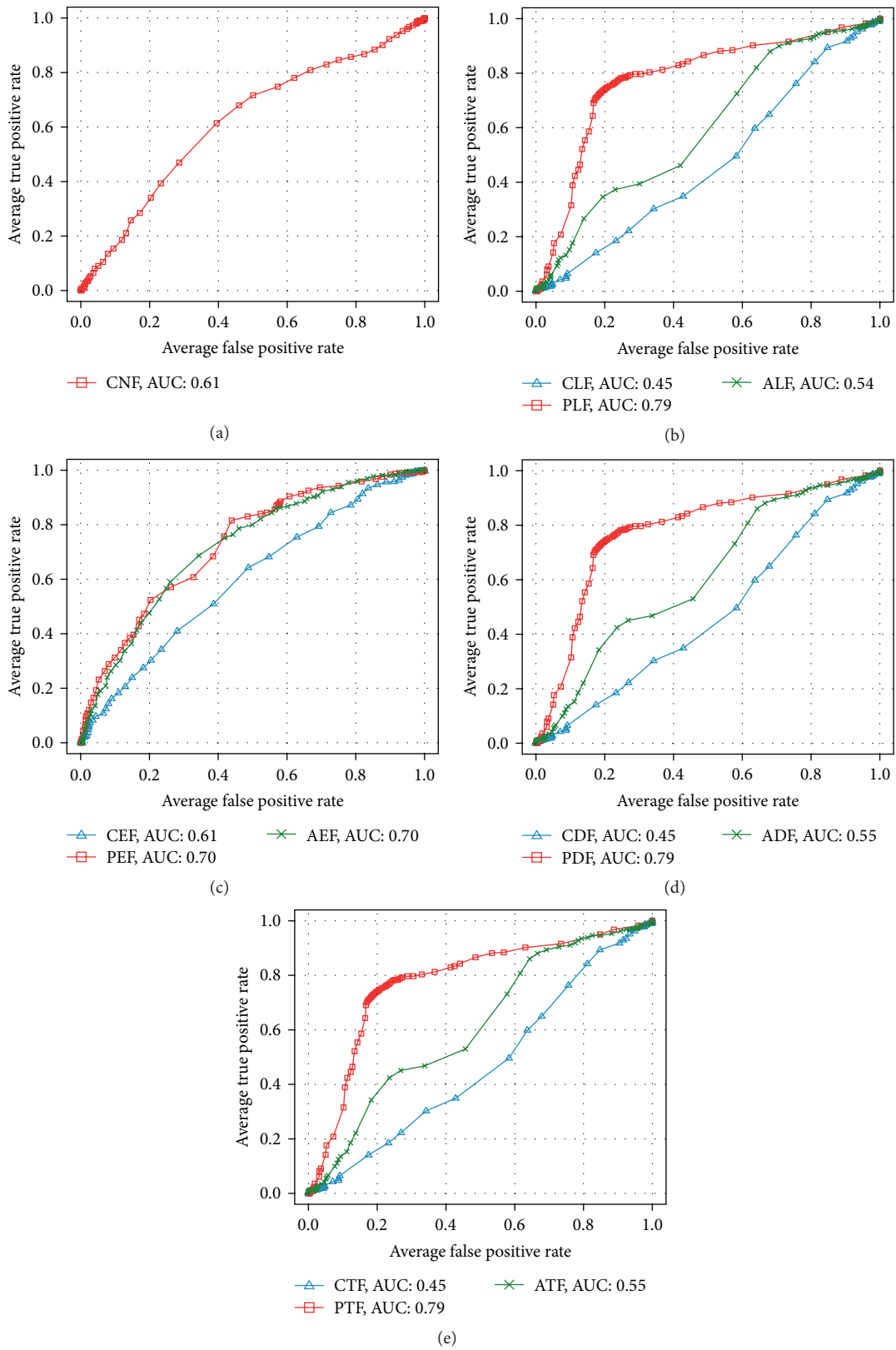
(a)

(b)

(c)

(d)

(e)

FIGURE 13: Comparison of ROC curves and AUC values for classifiers in Case 2.

TABLE 11: The statistics for lumen features and the results of F-test.

Lumen Features	Normal (df = 79) $\overline{N}$(C.I)	Grade 1 (df = 79) $\overline{G1}$(C.I)	Grade 2 (df = 79) $\overline{G2}$(C.I)	F-tests ($\mu_N = \mu_{G1} = \mu_{G2}$) F-value
Area (μm^2)	$3.78E + 04$ ($\pm 9.35E + 03$)	$6.77E + 04$ ($\pm 1.23E + 04$)	$1.19E + 05$ ($2.54E + 04$)	$4.00E + 01^*$
Perimeter (μm)	$1.49E + 03$ ($\pm 1.78E + 02$)	$2.09E + 03$ ($\pm 1.74E + 02$)	$2.73E + 03$ ($2.69E + 02$)	$6.05E + 01^*$
Width (μm)	$1.86E + 02$ ($\pm 2.64E + 01$)	$2.73E + 02$ ($\pm 2.93E + 01$)	$3.66E + 02$ ($4.70E + 01$)	$4.48E + 01^*$
Height (μm)	$1.79E + 02$ ($\pm 1.99E + 01$)	$2.39E + 02$ ($\pm 2.31E + 01$)	$3.05E + 02$ ($2.78E + 01$)	$4.88E + 01^*$
MajorAxis (μm)	$2.29E + 02$ ($\pm 2.94E + 01$)	$3.36E + 02$ ($\pm 3.23E + 01$)	$4.36E + 02$ ($5.01E + 01$)	$5.10E + 01^*$
MinorAxis (μm)	$1.83E + 02$ ($\pm 2.09E + 01$)	$2.42E + 02$ ($\pm 2.01E + 01$)	$3.20E + 02$ ($2.99E + 01$)	$5.70E + 01^*$
Circularity	$1.86E - 01$ ($\pm 1.10E - 03$)	$1.82E - 01$ ($\pm 2.70E - 03$)	$1.83E - 01$ ($2.59E - 03$)	$5.91E + 00^*$
Feret's diameter (μm)	$5.31E + 02$ ($\pm 6.42E + 01$)	$7.53E + 02$ ($\pm 6.43E + 01$)	$9.84E + 02$ ($1.00E + 02$)	$5.87E + 01^*$
Skewness	$2.15E - 01$ ($\pm 1.71E - 01$)	$-7.27E - 01$ ($\pm 1.91E - 01$)	$-9.88E - 01$ ($1.71E - 01$)	$8.79E + 01^*$
AspectRatio	$1.25E + 00$ ($\pm 5.58E - 02$)	$1.41E + 00$ ($\pm 1.12E - 01$)	$1.38E + 00$ ($1.10E - 01$)	$5.68E + 00^*$
Roundness	$8.15E - 01$ ($\pm 3.22E - 02$)	$7.47E - 01$ ($\pm 4.62E - 02$)	$7.60E - 01$ ($4.53E - 02$)	$5.21E + 00$
Solidity	$1.00E + 00$ ($\pm 0.00E + 00$)	$1.00E + 00$ ($\pm 0.00E + 00$)	$1.00E + 00$ ($0.00E + 00$)	$1.00E + 00$
AtypiaRatio	$5.11E - 02$ ($\pm 1.09E - 02$)	$2.72E - 01$ ($\pm 6.89E - 02$)	$4.69E - 01$ ($7.09E - 02$)	$9.22E + 01^*$
#AtypiaRegions	$7.50E - 02$ ($\pm 1.39E - 01$)	$2.48E + 00$ ($\pm 5.56E - 01$)	$4.84E + 00$ ($6.39E - 01$)	$1.61E + 02^*$
RMSAA	$1.03E + 00$ ($\pm 2.60E - 01$)	$8.85E + 00$ ($\pm 2.97E + 00$)	$1.88E + 01$ ($3.63E + 00$)	$7.51E + 01^*$
TSAV (rad)	$1.95E + 02$ ($\pm 1.88E + 02$)	$8.65E + 03$ ($\pm 3.45E + 03$)	$2.61E + 04$ ($6.45E + 03$)	$6.80E + 01^*$

* It indicates features whose null hypothesis was rejected with F-value > $F_{0.01}(2,237)$.

TABLE 12: The statistics for the epithelial nuclei features and the results of F-test.

Lumen features	Normal (df = 79) $\overline{N}$ (C.I)	Grade 1 (df = 79) $\overline{G1}$ (C.I)	Grade 2 (df = 79) $\overline{G2}$ (C.I)	F-tests ($\mu_N = \mu_{G1} = \mu_{G2}$) F-value
Area (μm^2)	$4.22E + 01$ ($\pm 2.67E + 00$)	$5.26E + 01$ ($\pm 2.82E + 00$)	$4.83E + 01$ ($\pm 2.94E + 00$)	$2.43E + 01^*$
Perimeter (μm)	$2.81E + 01$ ($\pm 1.09E + 00$)	$3.20E + 01$ ($\pm 9.63E - 01$)	$3.10E + 01$ ($\pm 9.88E - 01$)	$2.81E + 01^*$
Width (μm)	$8.04E + 00$ ($\pm 3.11E - 01$)	$9.11E + 00$ ($\pm 2.51E - 01$)	$8.65E + 00$ ($\pm 2.72E - 01$)	$2.58E + 01^*$
Height (μm)	$8.16E + 00$ ($\pm 3.38E - 01$)	$9.19E + 00$ ($\pm 3.17E - 01$)	$8.92E + 00$ ($\pm 2.59E - 01$)	$2.10E + 01^*$
MajorAxis (μm)	$8.87E + 00$ ($\pm 2.99E - 01$)	$1.01E + 01$ ($\pm 2.55E - 01$)	$9.66E + 00$ ($\pm 2.42E - 01$)	$3.73E + 01^*$
MinorAxis (μm)	$5.92E + 00$ ($\pm 2.01E - 01$)	$6.45E + 00$ ($\pm 2.10E - 01$)	$6.13E + 00$ ($\pm 2.25E - 01$)	$1.10E + 01^*$
Circularity	$6.87E - 01$ ($\pm 1.95E - 02$)	$6.47E - 01$ ($\pm 1.69E - 02$)	$6.33E - 01$ ($\pm 1.42E - 02$)	$1.86E + 01^*$
Feret's diameter (μm)	$9.94E + 00$ ($\pm 3.52E - 01$)	$1.13E + 01$ ($\pm 2.97E - 01$)	$1.09E + 01$ ($\pm 2.83E - 01$)	$3.46E + 01^*$
Skewness	$3.42E - 01$ ($\pm 1.19E - 01$)	$1.79E - 01$ ($\pm 8.28E - 02$)	$1.86E - 01$ ($\pm 9.83E - 02$)	$5.84E + 00^*$
AspectRatio	$1.51E + 00$ ($\pm 4.77E - 02$)	$1.60E + 00$ ($\pm 4.40E - 02$)	$1.62E + 00$ ($\pm 4.14E - 02$)	$1.23E + 01^*$
Roundness	$6.88E - 01$ ($\pm 1.73E - 02$)	$6.56E - 01$ ($\pm 1.41E - 02$)	$6.50E - 01$ ($\pm 1.29E - 02$)	$1.29E + 01^*$
Solidity	$8.48E - 01$ ($\pm 7.22E - 03$)	$8.41E - 01$ ($\pm 7.98E - 03$)	$8.32E - 01$ ($\pm 6.66E - 03$)	$8.87E + 00^*$
CytoplasmLength (μm)	$6.16E + 00$ ($\pm 8.20E - 01$)	$1.29E + 01$ ($\pm 1.16E + 00$)	$1.59E + 01$ ($\pm 1.01E + 00$)	$1.72E + 02^*$
CytoplasmLength$_{SD}$	$2.10E + 00$ ($\pm 4.06E - 01$)	$5.75E + 00$ ($\pm 7.69E - 01$)	$8.47E + 00$ ($\pm 7.17E - 01$)	$1.68E + 02^*$

* It indicates features whose null hypothesis was rejected with F-value > $F_{0.01}(2,237)$.

As opposed to the previous experiment that distinguished Normal from PDAC using PDF, an experiment of this case with PDF for classifying stages did not lead to improved performance of a classifier and the evaluation results were the same as experiments with PLF. It proves that the performance enhancement of classifier through PDF, a combination of PLE and PEF, is dependent entirely on PLF while PEF not showing any contribution.

The experiments with combination of feature sets (CTF, PTF, and ATF) extracted from three object types of tissue image showed the same results as those with combination of features sets (CDF, PDF, and ADF) extracted from a duct that consists two object types, lumen and epithelial nucleus.

Figure 13 presents ROC graphs and values of AUC averaged by classifiers that distinguishes PDAC stage based on sets of features of each object. In overall, lower performance was detected than Case 1. Classifiers with the classic morphologic features such as CNF, CLF, and CEF provide less accurate diagnosis with AUC values of 0.61, 0.45, and 0.61, respectively. On the contrary, the AUC values of classifiers learned by PLF and PEF are respectively 0.79 and 0.7, showing moderately accurate test results that is one-step higher than AUC values of classifiers learned by existing feature sets such as CEF, CLF, and CNF.

From these experiments of classifying PDAC, classifiers with PLF-contained feature sets show the best performances.

TABLE 13: LSDtest for *F*-test of nonepithelial nuclei features.

Nonepithelial nuclei features	LSD test			
	$|\overline{N} - \overline{G1}|$	$|\overline{N} - \overline{G2}|$	$|\overline{G1} - \overline{G2}|$	LSD-value
Area (μm^2)	$4.08E + 00^*$	$1.66E + 00$	$2.42E + 00^*$	$2.23E + 00$
Perimeter (μm)	$\mathbf{2.89E + 00^*}$	$\mathbf{1.81E + 00^*}$	$\mathbf{1.09E + 00^*}$	$\mathbf{9.34E - 01}$
Width (μm)	$\mathbf{9.75E - 01^*}$	$\mathbf{4.68E - 01^*}$	$\mathbf{5.07E - 01^*}$	$\mathbf{2.97E - 01}$
Height (μm)	$5.41E - 01^*$	$3.92E - 01^*$	$1.49E - 01$	$3.34E - 01$
MajorAxis (μm)	$\mathbf{1.10E + 00^*}$	$\mathbf{7.04E - 01^*}$	$\mathbf{3.99E - 01^*}$	$\mathbf{2.49E - 01}$
MinorAxis (μm)	—	—	—	—
Circularity	$6.66E - 02^*$	$5.68E - 02^*$	$9.75E - 03$	$1.62E - 02$
Feret's diameter (μm)	$\mathbf{1.22E + 00^*}$	$\mathbf{7.66E - 01^*}$	$\mathbf{4.57E - 01^*}$	$\mathbf{2.98E - 01}$
Skewness	—	—	—	—
AspectRatio	$3.24E - 01^*$	$2.60E - 01^*$	$6.40E - 02$	$7.24E - 02$
Roundness	$7.64E - 02^*$	$6.60E - 02^*$	$1.04E - 02$	$1.67E - 02$
Solidity	$1.92E - 02^*$	$1.88E - 02^*$	$3.46E - 04$	$6.40E - 03$

*It indicates that the absolute pairwise difference is greater than LSD value.

TABLE 14: LSD test for *F*-test of lumen features.

Lumen features	LSD test			
	$|\overline{N} - \overline{G1}|$	$|\overline{N} - \overline{G2}|$	$|\overline{G1} - \overline{G2}|$	LSD value
Area (μm^2)	$\mathbf{2.99E + 04^*}$	$\mathbf{8.14E + 04^*}$	$\mathbf{5.15E + 04^*}$	$\mathbf{2.39E + 04}$
Perimeter (μm)	$\mathbf{6.01E + 02^*}$	$\mathbf{1.25E + 03^*}$	$\mathbf{6.46E + 02^*}$	$\mathbf{2.95E + 02}$
Width (μm)	$\mathbf{8.69E + 01^*}$	$\mathbf{1.80E + 02^*}$	$\mathbf{9.27E + 01^*}$	$\mathbf{4.93E + 01}$
Height (μm)	$\mathbf{6.05E + 01^*}$	$\mathbf{1.26E + 02^*}$	$\mathbf{6.56E + 01^*}$	$\mathbf{3.32E + 01}$
MajorAxis (μm)	$\mathbf{1.07E + 02^*}$	$\mathbf{2.08E + 02^*}$	$\mathbf{1.01E + 02^*}$	$\mathbf{5.34E + 01}$
MinorAxis (μm)	$\mathbf{5.91E + 01^*}$	$\mathbf{1.37E + 02^*}$	$\mathbf{7.80E + 01^*}$	$\mathbf{3.34E + 01}$
Circularity	$3.90E - 03^*$	$3.18E - 03^*$	$7.25E - 04$	$3.13E - 03$
Feret's diameter (μm)	$\mathbf{2.22E + 02^*}$	$\mathbf{4.54E + 02^*}$	$\mathbf{2.32E + 02^*}$	$\mathbf{1.09E + 02}$
Skewness	$\mathbf{9.41E - 01^*}$	$\mathbf{1.20E + 00^*}$	$\mathbf{2.61E - 01^*}$	$\mathbf{2.48E - 01}$
AspectRatio	$1.63E - 01^*$	$1.34E - 01$	$2.90E - 02$	$1.34E - 01$
Roundness	—	—	—	—
Solidity	—	—	—	—
AtypiaRatio	$\mathbf{2.21E - 01^*}$	$\mathbf{4.17E - 01^*}$	$\mathbf{1.97E - 01^*}$	$\mathbf{7.99E - 02}$
#AtypiaRegions	$\mathbf{2.40E + 00^*}$	$\mathbf{4.76E + 00^*}$	$\mathbf{2.36E + 00^*}$	$\mathbf{6.90E - 01}$
RMSAA	$\mathbf{7.82E + 00^*}$	$\mathbf{1.77E + 01^*}$	$\mathbf{9.92E + 00^*}$	$\mathbf{3.77E + 00}$
TSAV (rad)	$\mathbf{8.45E + 03^*}$	$\mathbf{2.59E + 04^*}$	$\mathbf{1.74E + 04^*}$	$\mathbf{5.88E + 03}$

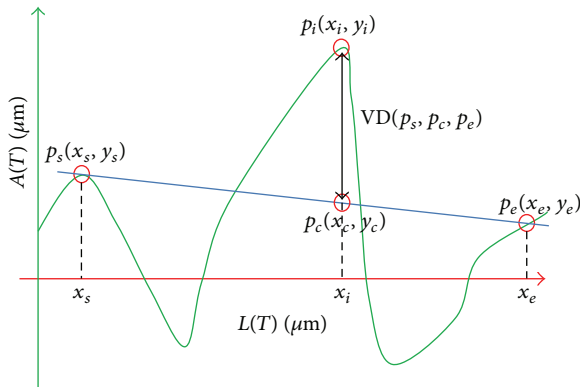*It indicates that the absolute pairwise difference is greater than LSD value.

FIGURE 14: VD(p_s, p_i, p_e) at p_i between PIPs, p_s and p_e.

As opposed to Case 1 distinguishing Normal from PDAC, no improvements were found with the combination of PEF and PLF. One particular aspect in these experiments is that experiment results of feature sets of lumen object are same as the results of experiments with feature sets of duct object and of tissue object. The feature sets extracted from duct and tissue are composed of the mix of feature sets from lumen object and other objects. It attests that features extracted from lumen are of positive influence to the classifier performance and contain the most information necessary for diagnosing stages of PDAC.

The experiments to classify Grade 1 and Grade 2 in Case 2 showed lower classification performance than the experiments to differentiate between Normal and PDAC in Case 2, both when the proposed features were used and when

TABLE 15: LSD-test for *F*-test of epithelial nuclei features.

Epithelial Nuclei Features	LSD-test			
	$|\overline{N} - \overline{G1}|$	$|\overline{N} - \overline{G2}|$	$|\overline{G1} - \overline{G2}|$	LSD-value
Area (μm^2)	**1.04E + 01***	**6.19E + 00***	**4.26E + 00***	**3.91E + 00**
Perimeter (μm)	$3.92E + 00$*	$2.92E + 00$*	$1.00E + 00$	$1.41E + 00$
Width (μm)	**1.07E + 00***	**6.07E − 01***	**4.66E − 01***	**3.89E − 01**
Height (μm)	$1.02E + 00$*	$7.60E − 01$*	$2.65E − 01$	$4.27E − 01$
MajorAxis (μm)	**1.21E + 00***	**7.96E − 01***	**4.19E − 01***	**3.71E − 01**
MinorAxis (μm)	$5.30E − 01$*	$2.19E − 01$	$3.12E − 01$*	$2.95E − 01$
Circularity	$3.96E − 02$*	$5.35E − 02$*	$1.38E − 02$	$2.36E − 02$
Feret's Diameter (μm)	$1.36E + 00$*	$9.36E − 01$*	$4.22E − 01$	$4.34E − 01$
Skewness	$1.64E − 01$*	$1.56E − 01$*	$7.53E − 03$	$1.40E − 01$
AspectRatio	$9.12E − 02$*	$1.11E − 01$*	$1.95E − 02$	$6.18E − 02$
Roundness	$3.19E − 02$*	$3.78E − 02$*	$5.90E − 03$	$2.07E − 02$
Solidity	$7.74E − 03$	$1.65E − 02$*	$8.74E − 03$	$1.02E − 02$
CytoplasmLength (μm)	**6.78E + 00***	**9.78E + 00***	**3.00E + 00***	**1.40E + 00**
CytoplasmLength$_{SD}$	**3.65E + 00***	**6.37E + 00***	**2.72E + 00***	**9.05E − 01**

*: It indicates that the absolute pairwise difference is greater than LSD-value.

the existing features were used. This can be explained by the fact that the characteristics of the PDAC commonly appear in Grade 1 and Grade 2 stages. Furthermore, the proposed features lack diagnosing stages of PDAC consisting of similar morphological characteristics since they are designed most of all to distinguish PDAC from Normal tissues. Even if it is so, the proposed feature sets perform better than the classical feature sets.

6. Discussion

In this section, we statistically analyzed features that were extracted from the three segmented parts (lumen, epithelial nuclei, and nonepithelial nuclei). Firstly, we assumed that if the extracted features are appropriate for diagnosing PDAC, then the value of features will be different among three populations (Normal, Grade 1, and Grade 2). To statistically show whether the extracted features are different among populations, we performed the ANOVA (ANalysis Of VAriance) for each feature. The null hypothesis for testing if the features are different among populations is as follows:

$$H_0 : \mu_N = \mu_{G1} = \mu_{G2}, \qquad (12)$$

where μ_N, μ_{G1}, and μ_{G2} are mean population means of Normal, Grade 1, and Grade 2, respectively. The significant test of ANOVA for the features is tested by *F-statistic*. Tables 10, 11, and 12 show *F-test* results of null hypothesis (12) for features of each of three object types at the 0.01 level of significance. The *F-test* results of features for three object types attest in statistics that there is difference among features in each group (Normal, Grade 1, or Grade 2) in most cases. In *F-test* of features extracted from lumen, the existence of disparities was confirmed between features of all groups (Normal, Grade 1, and Grade 2), except for Roundness and Solidity. In *F-test* for features of epithelial nuclei, all features show statistical difference between groups. As for the test

of nonepithelial nuclei, MinorAxis and Skewness were only features not showing statistically difference.

Next, for the post hoc analysis of features that reject the null hypothesis in *F-test*, we performed the multirange test to find whether there are significant differences between means of any population of two. In this paper, the commonly used Fisher's LSD (least significant difference) test was employed for the post hoc analysis of *F-test* [62]. In this paper, there are three populations (Normal, Grade 1, and Grade 2) examined. So, LSD-test for total $\binom{3}{2}$ pairs was performed. The null hypothesis for testing each pair is as follows:

$$H_{01} : \mu_N = \mu_{G1},$$
$$H_{02} : \mu_N = \mu_{G2}, \qquad (13)$$
$$H_{03} : \mu_{G1} = \mu_{G2}.$$

The LSD test results features for each of three object types are shown in Tables 13, 14, and 15. The bold values in Tables 13, 14, and 15 mean that the features of all the three hypotheses (13) are rejected in LSD test. Features marked with "-" are those not performed for LSD test since hypothesis test in (12) of *F-test* was rejected.

Firstly, in LSD test for features extracted from lumen, 12 features except for Circularity and AspectRatio have rejected the three null hypotheses of (13). Of which, 8 features are classical features and 4 features are proposed features. Although there is a number of classical features that have difference between groups, the experiment results using PLF showed more improved performance than experiments using CLF (refer to Tables 8 and 9) in Case 1 and Case 2.

In epithelial nuclei, 5 features including Cytoplasm-Length and CytoplasmLength$_{SD}$ rejected the null hypotheses (13). In the case of nonepithelial nuclei, only features of Perimeter, Width, MajorAxis, and Fereter's Diameter rejected the all null hypothesis of (13). The results of LSD test show that null hypothesis, H_{03}, of (13) to test significant difference
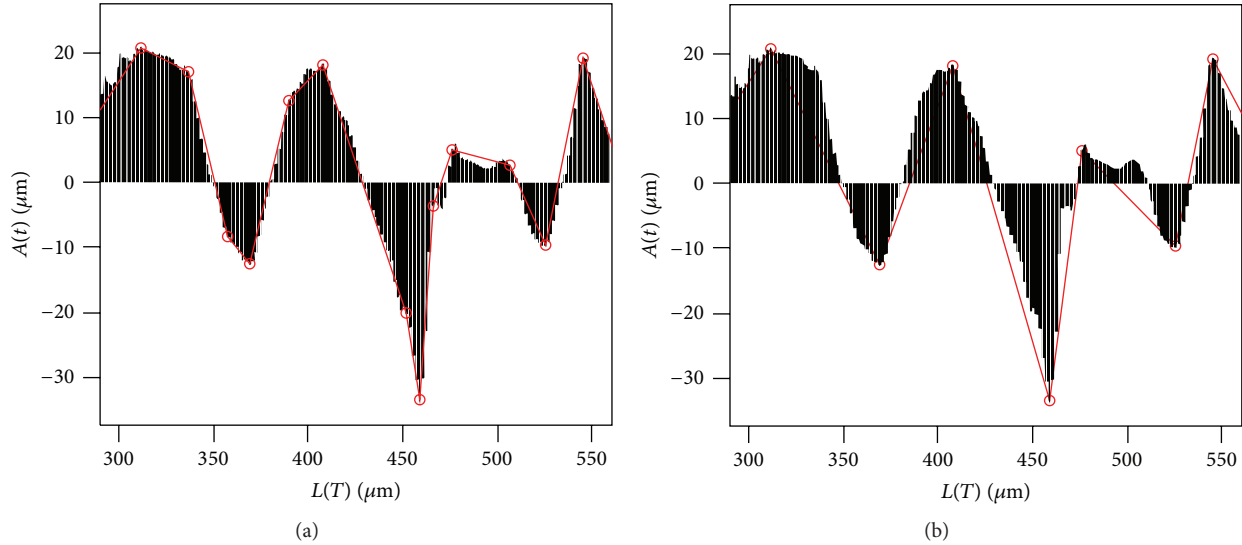
FIGURE 15: (a) PIPs identified by Algorithm 2; (b) PIPs after applying Algorithm 5 to (a) (red circles represent detected PIPs).

between Grade 1 and Grade 2 were not rejected in many features from two distinct object types, epithelial nucleus and nonepithelial nucleus.

In LSD test of lumen features, features that rejected null hypothesis (12) of *F-test* but not in H_{03} (13) are only two, Circularity and AspectRatio. However, in LSD tests of epithelial nucleus features, only 6 features of 14 features that rejected null hypothesis (12) of *F-test* reject H_{03}. In LSD-test of nonepithelial nuclei, only 4 features rejected H_{03}. Through LSD test, it has been confirmed that the lumen is the most important object in the diagnosis of PDAC and its stages. Furthermore, LSD test results describe the reason why classification performance of the experiments in Case 2 (Grade 1 versus Grade 2) is lower than that of them in Case 1 (Normal versus PDAC).

7. Conclusions

This paper proposed features to diagnose PDAC and to identify the stages of PDAC. PDAC is mainly diagnosed by investigating a duct that consists of lumen and epithelial cells. We segmented a tissue image into three parts: lumen, epithelial nuclei, and nonepithelial nuclei. Then, we proposed methods for extracting new morphological features from the epithelial cells and lumen parts that are segmented. In PDAC, the shape of the duct is more complex than Normal. Thus, this paper proposed the features for measuring atypia of the duct based on this perspective. We transformed the lumen into the atypia-amplitude signature with the atypia-amplitude function $A(t)$ to intuitively represent the variation of a duct and proposed RMSAA for measuring the deviation of the atypia-amplitudes and TSAV for measuring the volatility at PIP points of it. And, using the ideal lumen and original lumen regions, we measured AtypiaRatio, which represents the overall degree of distortion of a duct and #AtypiaRegions that quantify the papillary ducts. Also, we used features such as CytoplasmLength and CytoplasmLength$_{SD}$ to quantitatively

measure the morphological features from segmented epithelial nuclei. The experiments' results show that the proposed features are suitable to diagnose PDAC and to distinguish between the two stages, Grade 1 and Grade 2, of PDAC.

Appendices

In this paper, we use PIP method [46, 47] to find critical points of atypia-amplitude signature. However, the existing PIP method may not be able to find points that variation of the lumen boundary becomes more complex because it detects a fixed number of critical points. Therefore, we introduce the modified PIP detection method that finds all critical points of given atypia-amplitude signature. The modified PIP detection method is processed by the following two steps: (1) finding all PIPs of atypia-amplitude signature and (2) postprocessing for eliminating unnecessary PIPs.

A. Finding All PIPs of Atypia-Amplitude Signature

The PIP method finds critical points referred to as PIP (Perceptual Important Points) that represent important trends of time series data. In this paper, the existing PIP detection algorithm that detects a fixed number of PIPs has been modified to find all critical points in the atypia-amplitude signature. The modified PIP detection algorithm detects all PIPs with maximum vertical distances (VDs) [46, 47] above *Threshold* between adjacent PIPs. PIP_*Detection_For_Atypia_Amplitude_Signature* (Algorithm 2) shows the modified PIP detection algorithm in this paper.

As inputs, Algorithm 2 takes a sequence of points (denoted by *amp_list*) that forms the atypia-amplitude signature and the threshold (denoted by *T*) to detect PIPs. The algorithm output is a sequence of the detected PIPs (denoted by *pip_list*). The first step of Algorithm 2 initializes *pip_list* to

Input: Atypia-Amplitude Signature *amp_list*, Threshold for PIP detection *T*
Output: the sequence of the detected pip points *pip_list*
Method:
(1) $m \leftarrow$ *amp_list*.length
(2) *pip_list* $\leftarrow$ new List()   *//initialize pip_list*
(3) *pip_list*.add (*amp_list*.get(1))   *//the first point of amp_list is added in pip_list*
(4) *Sub_PIP_Detection* (1, *m*, *amp_list*, *T*, *pip_list*)   *//pip detection is performed at given range*
(5) **return** *pip_list*

ALGORITHM 2: *PIP_Detection_For_Atypia_ Amplitude_Signature* (*amp_list*, *T*).

Input: Start index *s*, End index *e*, Atypia-Amplitude Signature *amp_list*,
         Threshold for PIP detection *T*, Pip list *pip_list*
Method:
(1) max_vd_idx $\leftarrow$ *Max_ VD _Idx*(*s*, *e*, *amp_list*)
(2) max_vd $\leftarrow -\infty$
(3) **if** max_vd_idx **is not** 0 **then**
(4)   max_vd $\leftarrow$ *VD* (*s*, max_vd_idx, *e*, *amp_list*)
(5) **if** max_vd > *T* **then**
(6)   *Sub_PIP_Detection* (*s*, max_vd_idx, *amp_list*, *T*, *pip_list*)
(7)   *Sub_PIP_Detection* (max_vd_idx, *e*, *amp_list*, *T*, *pip_list*)
(8) **else**
(9)   *pip_list*.add (*amp_list*.get(*e*))
(10) **end if**

ALGORITHM 3: *Sub_PIP_Detection* (*s, e, amp_list, T*, *pip_list*).

Input: Start index *s*, End index *e*, Atypia-Amplitude Signature *amp_list*,
Output: Max VD Index *max_vd_idx*
Method:
(1) *max_vd_idx* $\leftarrow$ 0
(2) max_vd $\leftarrow -\infty$
(3) **if** $(e - s) > 1$ **then**
(4)   **for** $i \leftarrow s + 1$ **to** $e - 1$ **do**
(5)   vd $\leftarrow$ *VD* (*s, i, e, amp_list*)
(6)   **if** vd > max_vd **then**
(7)   *max_vd_idx* $\leftarrow i$
(8)   *max_vd* $\leftarrow$ vd
(9)   **end if**
(10)   **end for**
(11) **end if**
(12) **return** *max_vd_idx*

ALGORITHM 4: *Max_ VD _Idx* (*s, e, amp_list*).

the first point of *amp_list* (Algorithm 2, line 3). After that, all PIPs in a given *amp_list* are detected by *Sub_PIP_Detection* algorithm (Algorithm 3).

The *Sub_PIP_Detection* (Algorithm 3) finds a PIP within a given range of *amp_list*. The first and second inputs of the algorithm are start-index (denoted by *s*) and end-index (denoted by *e*) for the range of the *amp_list* to detect a PIP. Algorithm 3 first finds the location that has the maximum VD value in a given range (Algorithm 3, line 1). Then if the VD value of the found location is greater than a threshold *T*, the location is used as a pivot. The given range of the

algorithm is split into two ranges, and the *Sub_PIP_Detection* (Algorithm 3) is called with the new two ranges (Algorithm 3 lines 5–7). If the VD value is less than *T*, the partition of the range is stopped and the point at end-index *e* of the given range for the algorithm is added to *pip_list* (Algorithm 3, line 9). The *Sub_PIP_Detection* (Algorithm 3) is recursively called until all PIPs satisfying the threshold condition are detected in *amp_list* (Algorithm 4).

The distance metric VD used in the PIP detection is the vertical distance between the test point and the line connecting the two adjacent PIPs. That is, VD at a PIP

```
Input: Pip list pip_list,
Method:
(1) trend1 ← NIL
(2) trend2 ← NIL
(3) for i←1 to (pip_list.length − 2) do
(4)     pip1 ← pip_list.get (i)
(5)     pip2 ← pip_list.get (i + 1)
(6)     pip3 ← pip_list.get (i + 2)
(7)     if pip2.y > pip1.y then trend1 ← up
(8)     if pip2.y < pip1.y then trend1 ← down
(9)     if pip3.y > pip2.y then trend2 ← up
(10)    if pip3.y < pip2.y then trend2 ← down
(11)    if trend1 = trend2 then
(12)        pip_list.del (i + 1)
(13)        i ← i − 1
(14)    end if
(15) end for
```

ALGORITHM 5: *Post_Porocessing_PIPs* (**pip_list**).

$p_i(x_i, y_i)$ between two adjacent PIPs, $p_s(x_s, y_s)$ and $p_e(x_e, y_e)$, is as follows:

$$\begin{aligned} \mathrm{VD}(p_s, p_i, p_e) &= |y_c - y_i| \\ &= \left| \left(\frac{y_e - y_s}{x_e - x_s} \times (x_c - x_s) + y_s \right) - y_i \right|, \end{aligned} \quad \text{(A.1)}$$

where y_c is the value of a linear function determined by two points, p_s and p_e, when $x_c = x_i$ is given. Figure 14 shows the VD between the line connecting the two adjacent PIPs (p_s and p_e) and the test point p_i.

B. Postprocessing for Eliminating Unnecessary PIPs

Figure 15(a) shows the PIPs identified by the modified PIP algorithm (Algorithm 2). But some unnecessary PIPs can be found because the modified PIP algorithm identifies all PIPs satisfying the condition that the maximum VD is greater than a threshold. That is, several PIPs can be found where the change of the big trend has not occurred, as shown in Figure 15(a). These PIPs are points that have a trend in the same direction. To remove such unnecessary PIPs, postprocessing for the identified PIPs (*pip_list*) by the modified PIP algorithm (Algorithm 2) is performed. The postprocessing algorithm for PIPs is as Algorithm 5. The *Post_Processing_PIPs* (Algorithm 5) assesses whether each PIP of *pip_list* identified by Algorithm 2 is a maxima or minima when trend is reversed (Algorithm 5, lines 4–10). Then, if a PIP of the *pip_list* is not a maxima or minima, the PIP is removed in the *pip_list* (Algorithm 5, lines 11–14). Figure 15(b) shows the result of Algorithm 5 for Figure 15(a).

Acknowledgments

References

[1] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: a review," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147–171, 2009.

[2] A. E. Tutac, D. Racoceanu, T. Putti, W. Xiong, W. K. Leow, and V. Cretu, "Knowledge-guided semantic indexing of breast cancer histopathology images," in *Proceedings of the International Conference on BioMedical Engineering and Informatics (BMEI '08)*, vol. 2, pp. 107–112, May 2008.

[3] S. M. Ismail, A. B. Colclough, J. S. Dinnen et al., "Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia," *British Medical Journal*, vol. 298, no. 6675, pp. 707–710, 1989.

[4] A. Andrion, C. Magnani, P. G. Betta et al., "Malignant mesothelioma of the pleura: interobserver variability," *Journal of Clinical Pathology*, vol. 48, no. 9, pp. 856–860, 1995.

[5] C. Demir and B. Yener, "Automated cancer diagnosis based on histopathological systematic images: a systematic survey," *Rensselaer Polytechnic Institute*, 2005.

[6] S. Al-Janabi, A. Huisman, and P. J. Van Diest, "Digital pathology: current status and future perspectives," *Histopathology*, vol. 61, no. 1, pp. 1–9, 2012.

[7] A. Madabhushi, "Digital pathology image analysis: opportunities and challenges," *Imaging in Medicine*, vol. 1, no. 1, pp. 7–10, 2009.

[8] R. S. Weinstein, A. R. Graham, L. C. Richter et al., "Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future," *Human Pathology*, vol. 40, no. 8, pp. 1057–1069, 2009.

[9] J. I. Epstein and G. J. Netto, *Biopsy Interpretation of the Prostate*, Lippincott Williams & Wilkins, Philadelphia, Pa, USA, 4th edition, 2007.

[10] D. F. Gleason and G. T. Mellinger, "Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging," *Journal of Urology*, vol. 167, no. 2, pp. 953–959, 2002.

[11] A. Tabesh, M. Teverovskiy, H. Y. Pang et al., "Multifeature prostate cancer diagnosis and gleason grading of histological images," *IEEE Transactions on Medical Imaging*, vol. 26, no. 10, pp. 1366–1378, 2007.

[12] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi, "Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information," in *Proceedings of the Microscopic Image Analysis with Applications in Biology (MIAAB '07)*, 2007.

[13] P.-W. Huang and C. H. Lee, "Automatic classification for pathological prostate images based on fractal analysis," *IEEE Transactions on Medical Imaging*, vol. 28, no. 7, pp. 1037–1050, 2009.

[14] A. W. Wetzel, "Evaluation of prostate tumor grades by content-based image retrieval," in *Proceedings of the 27th AIPR Workshop on Advances in Computer-Assisted Recognition*, vol. 3584, pp. 244–252, 1999.

[15] S. Doyle, M. Hwang, K. Shah, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated grading of prostate cancer using architectural and textural image features," in *Proceedings of the 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI '07)*, pp. 1284–1287, April 2007.

[16] S.-K. Tai, Y.-C. Wu, C.-Y. Li, Y. J. Jan, and S. C. Lin, "Computer-assisted detection and grading of prostatic cancer in biopsy image," 2010, http://www.libsearch.com/view/1046167.

[17] Y. Peng, Y. Jiang, S. T. Chuang, and X. J. Yang, "Computer-aided detection of prostate cancer on tissue sections," *Applied Immunohistochemistry & Molecular Morphology*, vol. 17, no. 5, pp. 442–450, 2009.

[18] N. H. Anderson, P. W. Hamilton, P. H. Bartels, D. Thompson, R. Montironi, and J. M. Sloan, "Computerized scene segmentation for the discrimination of architectural features in ductal proliferative lesions of the breast," *The Journal of Pathology*, vol. 181, no. 4, pp. 374–380, 1997.

[19] C. Bilgin, C. Demir, C. Nagi, and B. Yener, "Cell-graph mining for breast tissue modeling and classification," in *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '07)*, pp. 5311–5314, 2007.

[20] A. N. Basavanhally, S. Ganesan, S. Agner et al., "Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, pp. 642–653, 2010.

[21] S. Doyle, S. Agner, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features," in *Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI '08)*, pp. 496–499, May 2008.

[22] W. H. Wolberg, W. N. Street, D. M. Heisey, and O. L. Mangarsian, "Computer-derived nuclear features distinguish malignant from benign breast cytology," *Human Pathology*, vol. 26, no. 7, pp. 792–796, 1995.

[23] R. Fernandez-Gonzalez, T. Deschamps, A. Idica, R. Malladi, and C. Ortiz de Solorzano, "Automatic segmentation of histological structures in mammary gland tissue sections," *Journal of Biomedical Optics*, vol. 9, no. 3, pp. 444–453, 2004.

[24] W. H. Wolberg, W. N. Street, and O. L. Mangasarian, "Breast cytology diagnosis with digital image analysis," *Analytical and Quantitative Cytology and Histology*, vol. 15, no. 6, pp. 396–404, 1993.

[25] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, and A. Murray, "Fractal analysis in the detection of colonic cancer images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, no. 1, pp. 54–58, 2002.

[26] H. K. Choi, T. Jarkrans, E. Bengtsson et al., "Image analysis based grading of bladder carcinoma. Comparison of object, texture and graph based methods and their reproducibility," *Analytical Cellular Pathology*, vol. 15, no. 1, pp. 1–18, 1997.

[27] M. N. Gurcan, J. Kong, O. Sertel, B. B. Cambazoglu, J. Saltz, and U. Catalyurek, "Computerized pathological image analysis for neuroblastoma prognosis," *Annual Symposium proceedings*, pp. 304–308, 2007.

[28] O. Sertel, J. Kong, H. Shimada, U. V. Catalyurek, J. H. Saltz, and M. N. Gurcan, "Computer-aided prognosis of neuroblastoma on whole-slide images: classification of stromal development," *Pattern Recognition*, vol. 42, no. 6, pp. 1093–1103, 2009.

[29] B. B. Cambazoglu, O. Sertel, J. Kong, J. Saltz, M. N. Gurcan, and U. V. Catalyurek, "Efficient processing of pathological images using the grid: computer-aided prognosis of neuroblastoma," in *Proceedings of the 16th International Symposium on High Performance Distributed Computing (HPDC '07)*, pp. 35–41, New York, NY, USA, June 2007.

[30] K. Belkacem-Boussaid, M. Pennell, G. Lozanski, A. Shana'ah, and M. N. Gurcan, "Effect of pathologist agreement on evaluating a computer-aided assisted system: recognizing centroblast cells in follicular lymphoma cases," in *Proceedings of the 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI '10)*, pp. 1411–1414, April 2010.

[31] O. Sertel, G. Lozanski, A. Shanáah, and M. N. Gurcan, "Computer-aided detection of centroblasts for follicular lymphoma grading using adaptive likelihood-based cell segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2613–2616, 2010.

[32] J. W. Song, J. H. Lee, T. S. Park, S. J. Chun, and J. H. Choi, "Mucinous cystadenoma classification system using automated epithelial tissue detection," in *Proceedings of the International Conference on Machine Vision*, 2010.

[33] G. Klöppel, R. H. Hruban, D. S. Longnecker, G. Adler, S. E. Kern, and T. J. Partanen, "Ductal adenocarcinoma of the pancreas," in *Word Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Digestive System*, International Agency for Research on Cancer, Lyon, France, 3rd edition, 2000.

[34] R. H. Hruban, P. Boffetta, C. Iacobuzio-Donahue et al., "Ductal adenocarcinoma of the pancreas," in *WHO Classification of Tumours of the Digestive System*, vol. 3, pp. 217–224, International Agency for Research on Cancer, Lyon, France, 4th edition, 2010.

[35] C. J. Yeo, J. L. Cameron, T. A. Sohn et al., "Six hundred fifty consecutive pancreaticoduodenectomies in the 1990s: pathology, complications, and outcomes," *Annals of Surgery*, vol. 226, no. 3, pp. 248–260, 1997.

[36] G. Klöppel, G. Lingenthal, M. Von Bulow, and H. F. Kern, "Histological and find structural features of pancreatic ductal adenocarcinoma in relation to growth and prognosis: studies in xenografted tumours and clinico-histopathological correlation in a series of 75 cases," *Histopathology*, vol. 9, no. 8, pp. 841–856, 1985.

[37] J. Lüttges, S. Schemm, I. Vogel, J. Hedderich, B. Kremer, and G. Klöppel, "The grade of pancreatic ductal carcinoma is an independent prognostic factor and is superior to the immunohistochemical assessment of proliferation," *Journal of Pathology*, vol. 191, no. 2, pp. 154–161, 2000.

[38] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[39] K. Rodenacker and E. Bengtsson, "A feature set for cytometry on digitized microscopic images," *Analytical Cellular Pathology*, vol. 25, no. 1, pp. 1–36, 2003.

[40] P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing*, vol. 41, no. 2, pp. 233–260, 1988.

[41] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans Syst Man Cybern*, vol. 9, no. 1, pp. 62–66, 1979.

[42] A. Z. Chitade and S. K. Katiyar, "Color based image segmentation using K-means clustering," *International Journal of Engineering Science and Technology*, vol. 2, no. 10, pp. 5319–5325, 2010.

[43] R. C. González and R. E. Woods, *Digital Image Processing*, Prentice Hall, New York, NY, USA, 2008.

[44] C. Pan, C.-X. Zheng, and H.-J. Wang, "Robust color image segmentation based on mean shift and marker-controlled watershed algorithm," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, vol. 5, pp. 2752–2756, November 2003.

[45] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[46] F.-L. Chung, T. C. Fu, V. Ng, and R. W. P. Luk, "An evolutionary approach to pattern-based time series segmentation," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 5, pp. 471–489, 2004.

[47] K.-P. Chan and A. W. Fu, "Efficient time series matching by wavelets," in *Proceedings of the 15th International Conference on Data Engineering (ICDE '99)*, pp. 126–133, 1999.

[48] M. Kosmahl, U. Pauser, M. Anlauf, and G. Klöppel, "Pancreatic ductal adenocarcinomas with cystic features: neither rare nor uniform," *Modern Pathology*, vol. 18, no. 9, pp. 1157–1164, 2005.

[49] "ScanScope CS-Asperio," 2012, http://www.aperio.com/lifescience/capture/cs.

[50] M. Kallergi, "Evaluation strategies for medical-image analysis and processing methodologies," in *Medical Image Analysis Methods*, L. Costaridou, Ed., vol. 18, CRC Press, New York, NY, USA, 2005.

[51] J. K. Udupa, V. R. LeBlanc, H. Schmidt et al., "A methodology for evaluating image segmentation algorithms," in *Proceedings of the Medical Imaging 2002: Image Processing*, pp. 266–277, February 2002.

[52] T. Allen, *Particle Size Measurement: Volume 1: Powder Sampling and Particle Size Measurement*, Springer, New York, NY, USA, 1996.

[53] W. Rasband, "ImageJ, 2012-1997," 2012, http://imagej.nih.gov/ij/.

[54] V. N. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, New York, NY, USA, 1st edition, 1998.

[55] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[56] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, Springer, New York, NY, USA, 2nd edition, 2009.

[57] C.-C. Chang and C.-J. Lin, "LIBSVM:a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, 2011.

[58] D. L. Verbyla and J. A. Litvaitis, "Resampling methods for evaluating classification accuracy of wildlife habitat models," *Environmental Management*, vol. 13, no. 6, pp. 783–787, 1989.

[59] H. J. Adér, G. J. Mellenbergh, and D. J. Hand, *Advising on Research Methods: a Consultant's Companion*, Johannes van Kessel, Huizen, Netherlands, 2008.

[60] J. A. Swets, "Measuring the accuracy of diagnostic systems," *Science*, vol. 240, no. 4857, pp. 1285–1293, 1988.

[61] M. Greiner, D. Pfeiffer, and R. D. Smith, "Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests," *Preventive Veterinary Medicine*, vol. 45, no. 1-2, pp. 23–41, 2000.

[62] L. J. Williams and H. Abdi, "Fisher's least significance difference (LSD) test," in *Encyclopedia of Research Design*, pp. 491–494, SAGE, Thousand Oaks, Calif, USA, 1st edition, 2010.