Age-based computer-aided diagnosis approach for pancreatic cancer on endoscopic ultrasound images

Murat Ozkan[1,2], Murat Cakiroglu[3], Orhan Kocaman[4], Mevlut Kurt[5], Bulent Yilmaz[5], Guray Can[5], Ugur Korkmaz[5], Emre Dandil[2,6], Ziya Eksi[7]

[1]Departments of Electricity and Energy, Bolu Vocational School and [5]Department of Gastroenterology, Abant Izzet Baysal University, Bolu, [2]Departments of Computer and Information Engineering, [3]Mechatronics Engineering and [7]Electronics and Computer Education, Sakarya University, Sakarya, [4]Department of Gastroenterology, Bezmiâlem Vakıf University, Istanbul, [6]Department of Computer Engineering, Bilecik Şeyh Edebali University, Bilecik, Turkey

ABSTRACT

Aim: The aim was to develop a high-performance computer-aided diagnosis (CAD) system with image processing and pattern recognition in diagnosing pancreatic cancer by using endosonography images. **Materials and Methods:** On the images, regions of interest (ROI) of three groups of patients (<40, 40-60 and >60) were extracted by experts; features were obtained from images using three different techniques and were trained separately for each age group with an Artificial Neural Network (ANN) to diagnose cancer. The study was conducted on endosonography images of 202 patients with pancreatic cancer and 130 noncancer patients. **Results:** 122 features were identified from the 332 endosonography images obtained in the study, and the 20 most appropriate features were selected by using the relief method. Images classified under three age groups (in years; <40, 40-60 and >60) were tested via 200 random tests and the following ratios were obtained in the classification: accuracy: 92%, 88.5%, and 91.7%, respectively; sensitivity: 87.5%, 85.7%, and 93.3%, respectively; and specificity: 94.1%, 91.7%, and 88.9%, respectively. When all the age groups were assessed together, the following values were obtained: accuracy: 87.5%, sensitivity: 83.3%, and specificity: 93.3%. **Conclusions:** It was observed that the CAD system developed in the study performed better in diagnosing pancreatic cancer images based on classification by patient age compared to diagnosis without classification. Therefore, it is imperative to take patient age into consideration to ensure higher performance.

Key words: Computer-aided diagnosis (CAD), endoscopic ultrasound (EUS) images, pancreatic cancer

INTRODUCTION

Pancreatic cancer is the fourth leading cause of cancer death among men and women.[1] As pancreatic malignant tumors are aggressive, the overall 5-year survival rate is below 5%.[2] Tumor removal is possible only in 10%-15% of the cases, and the 5-year survival rate of this group is about 10%.[3,4] Therefore, identification of tumors at early stages and assessment

Access this article online	
Quick Response Code:	**Website:** www.eusjournal.com
	DOI: 10.4103/2303-9027.180473

Address for correspondence

Dr. Mevlut Kurt, Department of Gastroenterology, Faculty of Medicine, Abant Izzet Baysal University, PK: 14280, Golkoy, Bolu, Turkey. E-mail: dr.mevlutkurt@gmail.com

of their spread is believed to result in a distinctive increase in survival rates.[5]

Today, imaging methods such as endoscopic ultrasound (EUS), magnetic resonance (MR) imaging, and computed tomography (CT) are used in the identification of pancreatic tumors. However, as is the case with other methods, the attention and experience of the physician are crucial here as well. Any lesion that is overlooked may bring the patient to a point of no return, and complications resulting from procedures may follow. Hence, there is an urgent need for new technologies that will supplement existing imaging techniques.[6]

In recent years, many studies in the biomedical field have focused on computer-aided diagnosis (CAD) systems to facilitate the identification of various diseases. The main goals in these studies are decreasing misdiagnosis rates, decreasing the rate of medical operations such as biopsies, saving time, and decreasing costs related to examination among others.[7]

Today, use of image processing and artificial intelligence techniques has considerably increased in many medical practice fields such as tumor diagnosis via mammogram, eye diseases, and cancer diagnosis. There are some studies in the literature related to CAD of some diseases such as pancreatic cancer and pancreatitis.[6,8,9] Literature review regarding the CAD systems developed for pancreatic cancer shows that the designed systems present common diagnostic performances for all age groups. However, it is known that the type of pancreatic problem changes according to the age of the patient and some other variables. In their study, Rajan *et al*.[10] observed age-based pancreatic changes on EUS images. In an experiment with over 120 patients, the researchers classified the patients in three groups based on their age (in years) as <40, 40-60 and >60 and observed age-dependent pancreatic changes. Based on the findings, it is evident that a CAD system that will be developed for pancreas will be more effective if it is based on age groups. The current study proposes a novel CAD system to differentiate between healthy and cancerous pancreas in three different age groups by using EUS images. The classifier in the designed system may receive EUS images for all age groups together as input for training and testing as well as receiving them separately. The proposed system is composed of four main phases: preprocessing, segmentation, feature extraction/selection, and classification.

MATERIALS AND METHODS

EUS image set
EUS images used in the current study were collected from 172 individuals who consulted the Bezmiâlem Vakıf University Faculty of Medicine between the dates January 1, 2013 and September 30, 2014. EUS-guided fine-needle aspiration biopsy (FNA) was used in the diagnosis of pancreatic cancers. The image data set was composed of 202 cancer and 130 noncancer EUS images obtained from 87 female and 85 male patients aged 17-84 years. All images were obtained by using a linear echoendoscope (Pentax EG-3870UTK, Pentax Europe GmbH, Hamburg, Germany, Hitachi EZU-MT28-S1, Hitachi Aloka Medical, Tokyo, Japan). Images in the data set were examined by two gastroenterology experts with at least 10 years of experience in the field. The identified pancreatic region was extracted with the help of a program to determine the regions of interest (ROI).

As can be seen in Table 1, EUS images were obtained from the patients on the basis of three different age groups. A total of 68 images were obtained from 40 patients in the group under 40 (11 cancer, 29 noncancer). Mean age in this group was 31.5 years. The group aged 40-60 years was composed of 36 cancer and 22 noncancer patients. A total of 75 EUS images were obtained from this group with the mean age of 50.3 years. The last group was composed of individuals over 60, and 189 EUS images in total were obtained from a total of 74 patients (46 cancer, 28 noncancer) in this group. Mean age for this group was calculated as 70.9 years, 95% confidence interval (CI) was used in data analysis, and the value of $P < 0.05$ was accepted as statistically significant.

The designed computer-aided diagnosis (CAD) system
As Figure 1 presents, the designed system is composed of four main phases. EUS images obtained for the

Table 1. The details of the image set used in the study

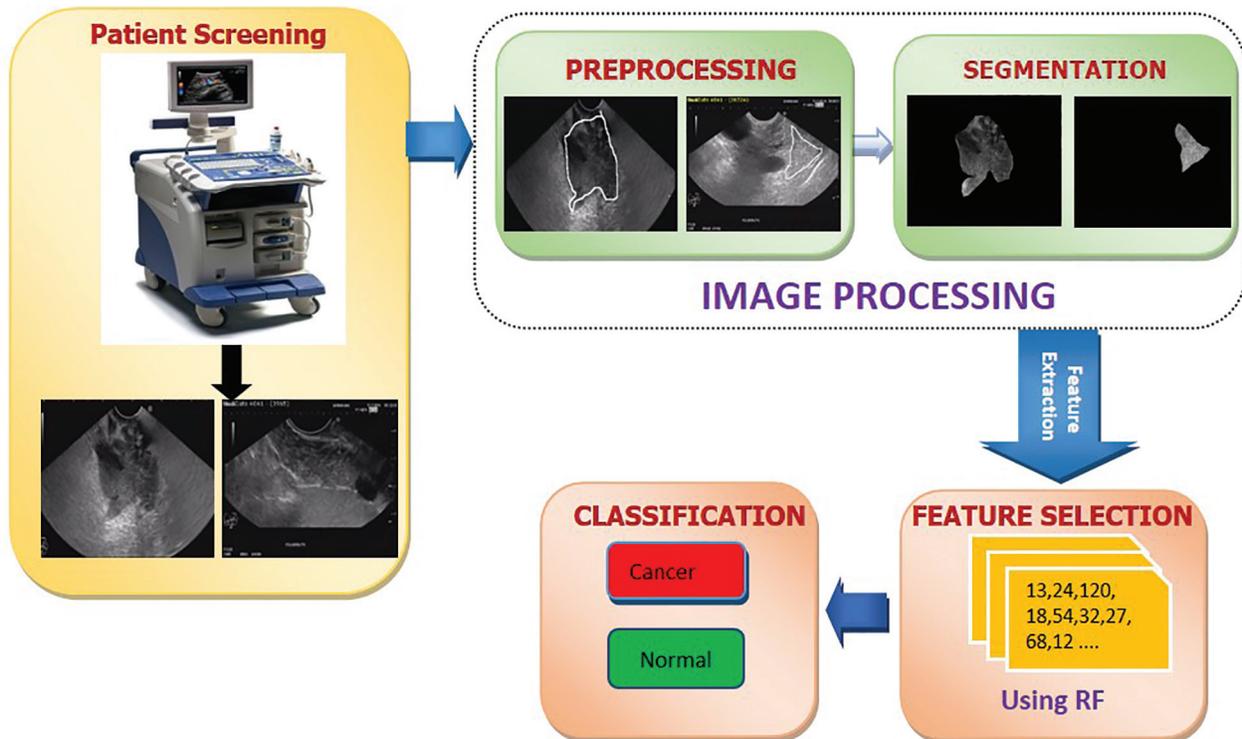| Age	<40		40-60		>60	
	N	**Image**	**N**	**Image**	**N**	**Image**
Pancreatic cancer	11	21	36	41	46	140
Normal pancreas	29	47	22	34	28	49
Average age		31.5		50.29		70.94
Sex						
Female		25		27		35
Male		15		31		39

Figure 1. Processing steps of the proposed CAD approach

study were preprocessed in the first phase to increase image quality. In the following phase, the pancreatic region was extracted from the whole image by the two experienced gastroenterology experts, and segmentation was completed. Feature extraction was undertaken in the next phase to identify the characteristics of the ROI and the most effective features were selected. The last phase included classification by using the selected features.

Image preprocessing

The purpose of image preprocessing is to ensure enhanced image quality, noise removal, and distinctive object edges in the images. A 5 × 5 median filter was applied to remove noise. Laplacian filters were used to ensure image sharpening. Histogram equalization was used to remove grains to complete the preprocessing phase. Following these operations, EUS images were reduced to the size 256 × 256. Figure 2 presents two examples of EUS images: before and after preprocessing.

Segmentation

Following the preprocessing phase of EUS images, the pancreatic region was identified and extracted by the two gastroenterology experts. This phase was completed with the help of experts as EUS image resolutions

were lower compared to MR and CT images, the whole pancreas was not observed in a single frame and the borders of the pancreatic region were not distinctive. Therefore, ROI was identified with the help of the experts. Figure 2 presents two examples of pancreatic segmentation following the extraction of ROI.

Feature extraction

Following ROI extraction, digital features of cancer and noncancer pancreatic regions were extracted. Gray level co-occurrence matrix (GLCM; 88 features),[11,12] standard statistical features (6 features),[13] wavelet decomposition energy features (10 features),[14] and boundary fractal features (18 features) methods were used to extract these features. In total, 122 features for each image were extracted.

Feature selection

One hundred twenty-two (122) features identified through feature extraction needed to be processed as input in the Artificial Neural Network (ANN). However, as the number of features extracted in this study was rather large in size, which would have taken a long time to be processed by ANN, identifying and using the most effective features was a more efficient approach. Hence, the Relief-F[15,16] feature reduction method was used on the 122 features identified in the
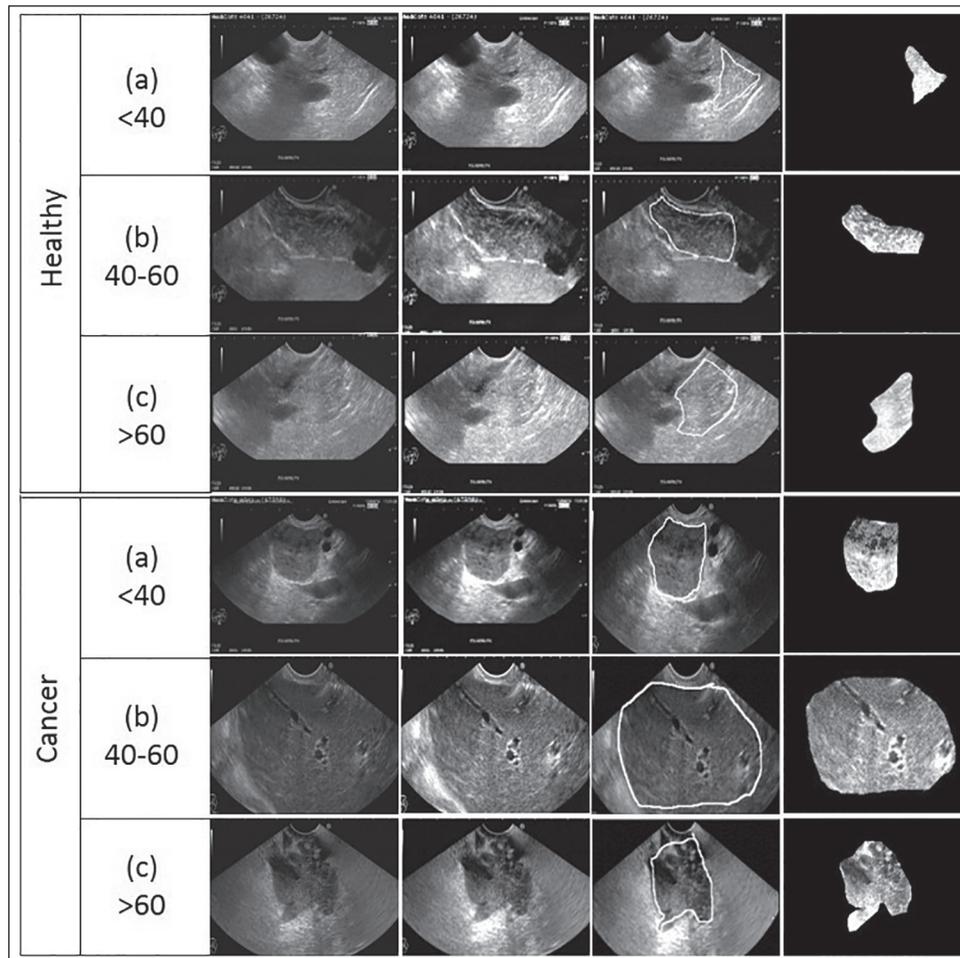
Figure 2. General operational steps of the designed system (Original Image, Enhanced Image, ROI Extraction, Segmentation)

previous phase to determine the most effective features that would be used by the system to make decisions. The test identified 20 features that would provide optimum results.

Classification

The 20 features that were selected in the study were used as input in ANN for classification. ANN is one of the most preferred artificial intelligence techniques that aim to create a system similar to the operations of the human brain by imitating it.[17] Multilayered, feed-forward perceptron was used in the system that classified cancer and noncancer pancreas.

The developed system has three layers: the input, hidden, and output layers. Input to ANN was undertaken via digital features obtained from the images. ANN involves a phase of training as the first step. In this phase, digital features related to noncancer and cancer images were fed into the system separately for each age group. Hence, the system learned the images for cancer and noncancer. The flowing phase includes testing in this phase, and the system was asked to make a decision after being fed the cancer and noncancer images. The system makes decisions based on two different options in its output. Obtaining "1" digitally as the output indicates that the image is cancerous, while "0" indicates a noncancer image.

RESULTS

A total of 332 EUS images of 202 cancer and 130 noncancer pancreas were used in assessing the ANN-based pancreatic cancer diagnosis system proposed in this study. MATLAB program (The MathWorks, Inc., MA, USA) image processing and artificial intelligence attachments were utilized in all the implementations of the study. All experimental studies were undertaken on a personal computer (PC) equipped with 3.4 GHz i7 processor, 8 GB memory, and Windows 7 operating system (Microsoft Corporation, WA, USA).

Accuracy, sensitivity, and specificity are the leading performance criteria preferred in medical diagnosis systems. Patients were first classified according to age groups, and classification results for each group were identified. Later, images for all patients were classified together to measure performance. All the obtained results are provided in Table 2. Examination of Table 2 shows that 21 cancer and 47 noncancer pancreatic images were used for classification in the age group under 40. 13 of the cancer data points were selected for training, whereas 8 were used for testing. 30 of the noncancer images were used for training and 17 were used for testing. Performance results for this age group were as follows: accuracy: 92%, sensitivity: 87.5%, specificity: 94.1%. 41 of a total of 75 pancreatic images for the 40-60 years age range were cancer data, whereas 34 were noncancer. 27 of these 41 cancer data were utilized for training and 14 were used for testing. 22 of the 34 noncancer data were used for training and 12 were used for testing. Performance results for this age group were found as follows: Accuracy: 88.5%, sensitivity: 85.7%, specificity: 91.7%. 140 cancer and 49 noncancer pancreatic images were used for classification for patients over 60. 110 of the cancer data were used for training, whereas 30 were selected for testing. 31 of the noncancer pancreatic data were used for training and 18 were used for testing. Performance results for this age group were as follows: Accuracy: 91.7%, sensitivity: 93.3%, specificity: 88.9%.

When all age ranges were evaluated together, 260 of 332 images were used for training and 72 were used for testing. 160 of the training data belonged to cancer images and 100 belonged to noncancer images. Data used for training included 42 cancer and 30 noncancer images. The following values were calculated for all age groups together: Accuracy: 87.5%, sensitivity: 83.3%, specificity: 93.3%. As seen from the results, classification of all age ranges together generates lower performance compared to separate classification of age ranges.

DISCUSSION

Pancreatic cancer has a high mortality and early diagnosis is the key factor in decreasing mortality. EUS imaging is the most common imaging method used for diagnosing pancreatic cancer. EUS can be used with FNA in differentiating benign and malignant tumors. However, FNA cannot be used in all health centers and the lack of specialized experts in FNA practices

Table 2: Performance results obtained from the experiments

Age		Predicted		Total	(%)
		Cancer (True)	Normal (False)		
<40	Cancer	7	1	8	Acc: 92±0.106
	Normal	1	16	17	Sn: 87.5±0.332
	Total	8	17	25	Sp: 94.11±0.136
40-60	Cancer	12	2	14	Acc: 88.46±0.117
	Normal	1	11	12	Sn: 85.71±0.374
	Total	13	13	26	Sp: 91.66±0.153
>60	Cancer	28	2	30	Acc: 91.66±0.057
	Normal	2	16	18	Sn: 93.33±0.147
	Total	30	18	48	Sp: 88.88±0.095
All ages	Cancer	35	7	42	Acc: 87.5±0.047
	Normal	2	28	30	Sn: 83.3±0.112
	Total	37	35	72	Sp: 93.33±0.075

may create serious problems. Therefore, diagnosis via EUS images will definitely offer convenience in clinical terms. However, the knowledge, experience, and skills of individual physicians may affect the results obtained from EUS use as well. Hence, using a computer-aided support system that will guide the physicians via EUS images will significantly contribute to more accurate and easier diagnosis.

The literature offers various studies related to CAD of cancer.[6,8,9,18] For instance, Das *et al.*[9] differentiated pancreatic cancer and chronic pancreatitis from normal pancreatic tissue by digital image analysis on EUS images. The researchers obtained cancerous pancreatic and normal pancreatic images from 12-22 cancer and normal patients. ROI were selected with the help of the experts to extract 228 features, out of which the 11 best features were selected. A 93% sensitivity rate was obtained using the ANN model.[9] In another study, Zhang *et al.*[6] differentiated between pancreatic cancer and normal cancer on EUS images. ROI were selected from 216 images obtained from 153 cancer and 63 non-cancer patients, and 97.98% sensitivity rate was obtained from the 29 features that were identified.[6] Zhu *et al.*[8] conducted a CAD by utilizing EUS images of pancreatic cancer and chronic pancreatitis patients in another study undertaken in 2013. In this study, images were obtained from 262 pancreatic cancer and 126 chronic pancreatitis patients, and 105 features were extracted. Sixteen (16) of these features were selected for use for classification by a support vector machine and a 94% sensitivity rate was obtained.[8]

The current study developed a new ANN-based approach that analyzed EUS images for the CAD of pancreatic cancer according to various age ranges. Cancer and noncancer EUS images in the current study were classified by age (in years) as under 40, between 40 and 60, and over 60 to be diagnosed. The study collected the EUS images from patients diagnosed with pancreatic cancer (93) and from noncancer patients (79), and 122 features were extracted. Relief-F method was used for feature reduction to obtain the 20 most appropriate features and the ANN was used to differentiate between cancerous and healthy tissues.

According to the results obtained from the experiments, performance results for all EUS images combined were "Accuracy: 87.5%, Sensitivity: 83.3%, Specificity: 93.3%," while better performances were obtained when age ranges were separately examined. For instance, values for the age range under 40 when evaluated separately were found to be Acc: 92%, Sn: 87.5%, Sp: 94.1%. For the age range 40-60, the values were Acc: 88.5%, Sn: 85.7%, Sp: 91.7%. Images for the age range over 60 yielded the following results: Acc: 91.7%, Sn: 93.3%, Sp: 88.9%. As these results suggest, classification of age groups for separate computer analyses increases system performance because the type and shape of pancreas may change according to age.[10] Therefore, training and testing the system on the basis of age groups increases system performance. It is believed that CAD of pancreatic diseases based on age ranges as proposed in the current study will contribute to increased performance of studies in the literature when applied to their study designs.

There are several limitations in the study. For instance, the number of EUS images that was used for the group under 40 years old is small. This is because of the small number of people afflicted by pancreatic cancer before the age of 40. The second important limitation of the study is the lack of comparison with other pancreatic diseases such as chronic pancreatitis, pancreatic pseudocysts, polyp, etc. It is rather difficult to accurately differentiate all these diseases on EUS images. We plan to carry out further studies using the available comparisons in this regard. Another limitation of the study is related to the use of a single health center and single equipment during the study. It is probable that images obtained via EUS equipment that offer lower-quality images have lower system performances. The last limitation is related to the lack of any real-time factor in the proposed system. Real-time operations of the system will be able to increase the probability of clinical use as they will allow the physician guidance during procedures. Hence, the purpose of future study in this field will be to add a real-time factor to the system in addition to postprocessing.

Conflicts of interest
There are no conflicts of interest.

REFERENCES

1. Siegel R, Ma J, Zou Z, *et al*. Cancer statistics, 2014. *CA Cancer J Clin* 2014;64:9-29.

2. Hidalgo M. Pancreatic cancer. *N Engl J Med* 2010;362:1605-17.

3. DiMagno EP, Malagelada JR, Taylor WF, *et al*. A prospective comparison of current diagnostic tests for pancreatic cancer. *N Engl J Med* 1977;297:737-42.

4. Gudjonsson B. Cancer of the pancreas. 50 years of surgery. *Cancer* 1987;60:2284-303.

5. Freeny PC, Marks WM, Ryan JA, *et al*. Pancreatic ductal adenocarcinoma: Diagnosis and staging with dynamic CT. *Radiology* 1988;166:125-33.

6. Zhang MM, Yang H, Jin ZD, *et al*. Differential diagnosis of pancreatic cancer from normal tissue with digital imaging processing and pattern recognition based on a support vector machine of EUS images. *Gastrointest Endosc* 2010;72:978-85.

7. Aslantas A, Dandil E, Saglam S, *et al*. CADBOSS: A computer-aided diagnosis system for whole-body bone scintigraphy. *J Cancer Res Ther* 2015. [In Press].

8. Zhu M, Xu C, Yu J, *et al*. Differentiation of pancreatic cancer and chronic pancreatitis using computer-aided diagnosis of endoscopic ultrasound (EUS) images: A diagnostic test. *PLoS One* 2013;8:e63820.

9. Das A, Nguyen CC, Li F, *et al*. Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. *Gastrointest Endosc* 2008;67:861-7.

10. Rajan E, Clain JE, Levy MJ, *et al*. Age-related changes in the pancreas identified by EUS: A prospective evaluation. *Gastrointest Endosc* 2005;61:401-6.

11. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;SMC-3:610-21.

12. Clausi DA. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Can J Remote Sensing* 2002;28:45-62.

13. Akilandeswari U, Nithya R, Santhi B. Review on feature extraction methods in pattern classification. *Eur J Sci Res* 2012;71:265-72.

14. Van de Wouwer G, Scheunders P, Van Dyck D. Statistical texture characterization from discrete wavelet representation. *IEEE Trans Image Process* 1999;8:592-8.

15. Tahir F, Fahiem MA. A statistical-textural-features based approach for classification of solid drugs using surface microscopic images. *Comput Math Methods Med* 2014;2014:791246.

16. Kira K, Rendell LA. The feature selection problem: Traditional methods and a new algorithm. *Amer Assoc Artificial* I 1992;129-34.

17. Dandil E, Çevik KK. Yapay sinir ağları için net platformunda görsel bir eğitim yazılımının geliştirilmesi. *International Journal of Informatics Technologies* 2012;5:19-28.

18. Fritscher-Ravens A, Brand L, Knöfel WT, *et al.* Comparison of endoscopic ultrasound-guided fine needle aspiration for focal pancreatic lesions in patients with normal parenchyma and chronic pancreatitis. *Am J Gastroenterol* 2002; 97:2768-75.