Ashour *et al. Health Inf Sci Syst (2018) 6:7*
https://doi.org/10.1007/s13755-018-0047-z

Health Information Science and Systems

RESEARCH

# Comparative study of multiclass classification methods on light microscopic images for hepatic schistosomiasis fibrosis diagnosis

Amira S. Ashour[1], Ahmed Refaat Hawas[1] and Yanhui Guo[2*]

**Abstract**

Hepatic schistosomiasis is a prolonged disease resulting mainly from the solvable egg antigen of schistosomiasis infection due to the host's granulomatous cell-mediated immune. Irreversible fibrosis results from the progress of the schistosomal hepatopathy. Sensitive diagnosis of this disease is based on the investigation of the microscopy images, liver tissues, and egg identification. Early diagnosis of schistosomiasis at its initial infection stage is vital to avoid egg-induced irreparable pathological reactions. Typically, there are several classification approaches that can be used for liver fibrosis staging. However, it is unclear which approaches can achieve high accuracy for analyzing and intelligently classifying the liver microscopic images. Consequently, this work aims to study the performance of the different machine learning classifiers for accurate fibrosis level staging of granuloma, namely cellular, fibrocellular and fibrotic granulomas as well as the normal samples. The classifiers include a multi-layer perceptron neural network, a decision tree, discriminant analysis, support vector machine (SVM), nearest neighbor, and the ensemble of classifiers. The statistical features of the microscopic images are extracted from the different fibrosis levels of granuloma, namely cellular, fibrocellular and fibrotic granulomas as well as the normal samples. The results established that the maximum achieved classification accuracies of value 90% were achieved using the subspace discriminant ensemble, the quadratic SVM, the linear SVM, or the linear discriminant classifiers. However, the linear discriminant classifier can be considered the superior classifier as it realized the best area under the curve of value 0.96 during the classification of the cellular granuloma as well as the fibro-cellular granuloma fibrosis levels.

**Keywords:** Hepatic schistosomiasis, Fibrosis, Statistical features, Ensemble classifier, Decision tree, Discriminant analysis, Support vector machine, Nearest neighbor

## Introduction

In developing countries, trematode schistosoma causes schistosomiasis, which is a prevalent disease that affects the liver tissues [1–3]. Liver fibrosis occurs due to the invariably of this schistosoma mansoni infection. Such fibrosis has several stages, namely cellular, fibrocellular and fibrotic granulomas that may be characterized by small focal areas of sever inflammation and excess extracellular matrix placed in periovular granulomas. In schistosomiasis, the interactions of the host-parasite assist the understanding of the liver fibrosis features, such as regulation, vascular changes, and the portal hypertension pathophysiology [4]. The progression of the fibrosis level is related to the liver function failure. Thus, monitoring the microscopic liver fibrosis images is essential for precise identification of the chronic liver diseases for further appropriate therapy. Quantitative assessment of the liver fibrosis level using image analysis provided superior and accurate results compared to the conventional assessment [5].

*Correspondence: yguo56@uis.edu
[2] Department of Computer Science, University of Illinois at Springfield, Springfield, IL, USA
Full list of author information is available at the end of the article

© Springer Nature Switzerland AG 2018.

Recently, for liver tissues classification, medical image processing and machine learning have been developed for computer-aided diagnosis systems. For liver images classification, Mahmoud-Ghoneim [6] applied texture analysis at different resolutions on the conventional grey scale images, and RGB (Red, Green, Blue) images as well as the Hue-Saturation-Intensity (HSI). At low resolution, significant characterizing features can be extracted from the green channel of the liver fibrosis images. However, at high resolution, the gray scale space provided superior results. Additionally, at all resolutions, the HSI space had high error percentage, thus, it is unsatisfactory for liver fibrosis classification. Stanciu et al. [7] used non-linear optical microscopy method, namely the two-photon excitation fluorescence (TPEF) for liver fibrosis assessment, and scoring by capturing images of a Thioacetamide-induced rat model. These images are then classified using a gradient based Bag-of-Features (BoF) approach. The results reported the assessed performance was influenced by the BoF parameters during the fibrosis scoring framework.

Early liver fibrosis diagnosis is a challenging issue, which inspired researchers to employ machine learning classifiers for the staging process. Based on the previous studies, until now, few automated fibrosis staging classifiers have been implemented. At the same time, there are several image classification approaches that can be employed for liver fibrosis staging, such as ensemble of classifiers, support vector machine (SVM), neural network, the decision tree, and k-nearest neighbor (KNN) [8]. Recently, for automated microscopic images classification of microscopic liver images, Cinque et al. [9] integrated the textural based segmentation technique with a support vector machine. This approach detected the existence of abnormal regions for further classification using the SVM. Consequently, it is essential to compare the performance and feasibility of the image classification procedures in diagnosing the liver fibrosis diseases. This can enable the follow-up studies on the automated liver fibrosis diagnosis.

Accordingly, for liver fibrosis staging, the present work conducted a comparative study of different classifiers, namely the discriminant analysis (linear/quadratic), support vector machine 'SVM' (linear/quadratic/cubic/fine Gaussian/medium Gaussian/coarse Gaussian), Nearest Neighbor 'KNN' (fine/medium/cosine/cubic/weighted/coarse), ensembles (subspace with discriminant/bagged with trees/subspace with KNN/ boosted with trees/RUS-Boosted with trees), neural network (multi-layer perceptron neural network 'MLP-NN'), and the decision tree (simple/medium/complex). This comparative study is applied to enumerate the different classifiers accuracies on liver fibrosis microscopic images of animal models (mice) liver samples.

The remaining sections are organized as follows. The methodology and a brief description of the involved classifiers are included in Sect. 2. The results are reported and discussed in Sect. 3. The conclusion is finally presented in Sect. 4.

## Methodology

The light microscopic samples of infected mice are captured by Schistosoma mansoni cercariae. The attained images included normal and different fibrosis three levels, which are cellular granuloma, fibrocellular granuloma, and fibrotic granuloma; respectively. Binarization, thresholding, and segmentation using the watershed method are used to identify the lesion regions. Then, the most significant features are selected, which are the area, Feret, minor, and the RawIntDen of the selected region in the image. These selected features can be defined as follows: the area represents the area in square pixels or in ($mm^2$, $\mu m^2$, etc.) of the region of interest (ROI) according to the calibration unit, the Feret represents the Feret's diameter, which is the longest distance between any two points along the ROI boundary, the minor is the secondary axis of the best fitting ellipse that fits the selected ROI, and Raw integrated density (RawIntDen) represents the sum of the pixel values in the selected ROI. These selected statistical features are employed to classify the cases in the dataset into one of the four cases using the different classifiers involved in the present comparative study.

### Discriminant analysis

In the present work, the linear discriminant analysis (LDA) and the quadratic discriminant analysis (QDA) are employed for the multiclass classification using a linear and a quadratic decision surface, respectively. These classifiers have easily computed closed-form solutions that do not have hyper-parameters to be tuned. Typically, the LDA can only learn linear boundaries, while QDA is flexible by using learn quadratic boundaries. For classification, the LDA calculates discriminant scores for each instance (sample) [10]. These scores are attained by finding the independent variables' linear combinations. For a single predictor variable $A = a$, the LDA classifier can be estimated using the following expression of the discriminant score $\left(\widehat{\eta}(a)\right)$ [11]:

$$\widehat{\eta}(a) = a \cdot \frac{\widehat{\delta}_w}{\widehat{\sigma}^2} - \frac{\widehat{\delta}_w^2}{2\widehat{\sigma}^2} + \log\left(\widehat{\gamma}_w\right) \tag{1}$$

where $\widehat{\eta}(a)$ is the expected discriminant score that used to classify the sample to its $w$th class within the response

variable according to the predictor variable $a$ value. For the $w$th class, $\widehat{\delta}_w$ represents the average of the training samples, $\widehat{\sigma}^2$ is the sample variances' weighted average and $\widehat{\gamma}_w$ represents the prior probability of the sample to which it belongs to specific class. Thus, each sample (instant) is assigned to the $w$th class, which has the largest $\widehat{\eta}(a)$, where the LDA computes the probability distribution to classify the sample to specific classifier. Typically, the LDA considers that the samples within each class are from a multivariate Gaussian distribution and the predictor covariance of the variables is common across all $w$ classes. These assumptions provide some enhancements over the logistic regression [12, 13].

Conversely, the QDA has different approach compared to the LDA as it considers that each class has its individual covariance matrix, where the predictor variables do not have mutual variance across the $w$ classes [14]. Consequently, the QDA can sense the divergent in the covariance of the variables and offer non-linear and more accurate classification decision boundaries compared to the LDA.

#### Support vector machines

The SVM is a supervised learning discriminative classifier that has separating hyperplane. It basically depends on determining the optimal hyperplane, which provides the largest minimum distance (margin) between the classes to separate all data samples of one class from those of the other classes [15]. The SVM classifier is fast and accurate. Nevertheless, it requires training and its model before using it. It can be used with multi-classes (as in the present work to classify the sample into one of the four classes), where the SVM model will generate a set of binary classification sub-problems using one SVM learner for each sub-problem. This binary classification can be performed i) the 'one-versus-one' strategy between every pair of the classes or using ii) the 'one-versus-all' strategy between one of the labels and the rest [16]. In the first strategy, the classification is performed using the 'max-wins voting' approach. The concept of this voting strategy is to assign the sample to one of the two classes, and then increase the vote for the assigned class. Afterwards, the class with the most votes is considered to determine the classification of the sample. On the contrary, the 'winner-takes-all' approach is applied to classify any new sample for the 'one-versus-all' strategy (second strategy). In this approach, the classifier with the highest output assigns the class.

To clarify the concept of the SVM, let $X$ and $Y$ represent the input and output sets; respectively, and $(x_1, y_1), \ldots, (x_a, y_a)$ is the training set. This training set is used to learn the classifier, which is expressed as follows:

$$y = f(x, \beta) \tag{2}$$

where $\beta$ are the function parameters and the decision function is given by:

$$f(x) = \sum_i \beta_i K(x_i, x) + \mathrm{s} \tag{3}$$

where $K(x_i, x)$ is the kernel function, which is used to implicit the nonlinear feature map. The present work applied several kernel functions, namely linear, quadratic, cubic, fine Gaussian, medium Gaussian, and coarse Gaussian.

#### Nearest neighbor based classifiers

In the present work, several varieties of the nearest neighbor (KNN) classifier are employed, namely the fine KNN, medium KNN, cosine KNN, cubic KNN, weighted KNN, and coarse KNN. Generally, in the training dataset, the KNN classified the samples (instances) according to their distance to other instances [17]. During the classification of any new instance, the kNN model search for the listed $k$ number of the nearest neighbors. However, this classifier may be misled by irrelevant features.

#### Decision tree

Accompanied by linear classifiers, the decision tree is considered one of the broadly used classification methods. It consists of internal (non-leaf) node that represents an attribute, each branch denotes the test output, and a class label is represented by each leaf node. The decision tree is extremely simple, intuitive and achieves interpretable estimates [18]. It has two main prediction steps, including the model training to build the tree, and then trained model can be used for predicting any new instances. In the present study, simple tree, medium tree, and complex tree are applied for multi-classification of the normal and the three fibrosis levels.

#### Other classifiers

Furthermore, the neural network, namely the multi-layer perceptron neural network 'MLP-NN' [19] as well as the ensembles of classifiers [20–22] is also included in the present work. The ensembles of classifiers that involved in the resent work are, namely the subspace with discriminant ensemble, the bagged with trees ensemble, the subspace with KNN ensemble, the boosted with trees ensemble, and the RUSBoosted with trees ensemble.

## Experimental results and discussions

In the current work, the used dataset is obtained from the Parasitology Department, Faculty of Medicine, Tanta University, Tanta, Egypt. It includes normal mice liver microscopic samples, cellular granuloma of level 1

**Table 1 Parameters' settings of the used classifiers**

| Main classifier type | Classifier specific type | Parameters' Settings |
|---|---|---|
| Discriminant analysis | Linear discriminant | Regularization is auto |
| | Quadratic discriminant | Regularization is diagonal covariance |
| Support vector machine | Linear SVM | Kernel function is linear<br>Kernel scale is automatic<br>Box constraint level is 1<br>Multiclass method is one-vs-one<br>Standardize data is true |
| | Quadratic SVM | Kernel function is quadratic<br>Kernel scale is automatic<br>Box constraint level is 1<br>Multiclass method is one-vs-one<br>Standardize data is true |
| | Cubic SVM | Kernel function is cubic<br>Kernel scale is Automatic<br>Box constraint level is 1<br>Multiclass method is one-vs-one<br>Standardize data is true |
| | Fine Gaussian SVM | Kernel function is Gaussian<br>Kernel scale is 0.5<br>Box constraint level is 1<br>Multiclass method is one-vs-one<br>Standardize data is true |
| | Medium Gaussian SVM | Kernel function is Gaussian<br>Kernel scale is 2<br>Box constraint level is 1<br>Multiclass method is one-vs-one<br>Standardize data is true |
| | Coarse Gaussian SVM | Kernel function is Gaussian<br>Kernel scale is 8<br>Box constraint level is 1<br>Multiclass method is one-vs-one<br>Standardize data is true |
| Decision tree | Complex tree | Maximum number of splits is 100<br>Split criterion is Gini's diversity index<br>Surrogate decision splits is off |
| | Medium tree | Maximum number of splits is 20<br>Split criterion is Gini's diversity index<br>Surrogate decision splits is off |
| | Simple tree | Maximum number of splits is 4<br>Split criterion is Gini's diversity index<br>Surrogate decision splits is off |
| Nearest neighbor | Fine KNN | Number of neighbors is 1<br>Distance metric is euclidean<br>Distance weight is equal<br>Standardize data is true |
| | Medium KNN | Number of neighbors is 10<br>Distance metric is euclidean<br>Distance weight is equal<br>Standardize data is true |
| | Coarse KNN | Number of neighbors is 100<br>Distance metric is euclidean<br>Distance weight is equal<br>Standardize data is true |
| | Cosine KNN | Number of neighbors is 10<br>Distance metric is cosine<br>Distance weight is equal<br>Standardize data is true |
| | Cubic KNN | Number of neighbors is 10<br>Distance metric is cubic (Minkowski)<br>Distance weight is equal<br>Standardize data is true |

**Table 1 continued**

| Main classifier type | Classifier specific type | Parameters' Settings |
|---|---|---|
| | Weighted KNN | Number of neighbors is 10<br>Distance metric is euclidean<br>Distance weight is squared inverse<br>Standardize data is true |

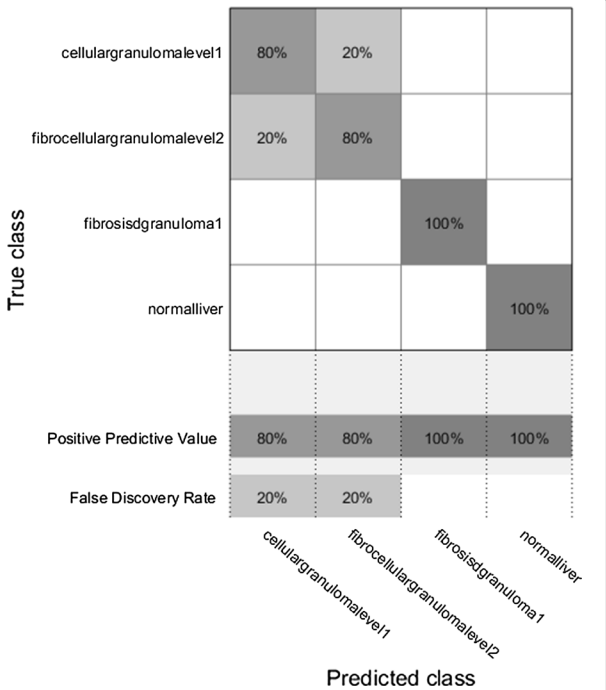


**Fig. 1** Confusion matrix of the linear discriminant

fibrosis, fibrocellular granuloma of level 2 fibrosis, and the fibrotic granuloma of level 3 fibrosis. The dataset includes 60 microscopic images of liver sections at different fibrosis levels and the normal liver case (15 images from each class). The comparative study between all the mentioned classifiers and the varieties of each is conducted in the present work in terms of the classification accuracy, where the setting parameters of the used classifiers are reported in Table 1.

To evaluate the performance of the different classifiers, the true positive rates/ false negative rates, and the positive predictive values/ false discovery rates are obtained for each classifier to measure the accuracies of the different classifiers. In addition, the Receiver Operating Characteristic (ROC) curves are included to represent the classifier results during the test phase. Generally, the ROC curve illustrates the FPR (false positive rate) representing the number of the incorrect positive classification regarding the negative instances, and the TPR (true positive rate) representing the number of correct positive results about all positive instances.

The present study included twenty-two classifiers; however, top four classifiers, namely the linear discriminant, linear SVM, quadratic SVM, and the subspace discriminant ensemble, provided the best accuracy. Thus, we highlighted in some details the results of those superior classifiers as follows. However, since the subspace discriminant ensemble employed the linear discriminant as its learner type with 30 learners using the ensemble subspace method, to avoid repetition, we do not include again the same using subspace discriminant ensemble.

## Discriminant analysis staging performance

### *Linear discriminant classification performance*

The confusion matrix of the linear discriminant classifier is illustrated in Fig. 1, showing the positive predictive values/false discovery rates. The ROC curves are demonstrated in Fig. 2a–d for the normal and fibrosis levels; respectively, during the staging process.

The confusion matrix in Fig. 1 reports 90% accuracy of the linear discriminant classifier. In addition, Fig. 2 presents the ROC curves including the area under the curve (AUC) to measure the classification accuracy. Figure 2 establishes that the linear discriminant classifier has AUC = 1, which indicates perfect classification of both the normal and the fibrosis granuloma due to the absence of the granulomas and the fibrosis regions in the normal cases and the very big area of the fibrosis granuloma regions, which provided significant differences in the extracted features of these classes compared to all the four classes in the present study. However, the AUC = 0.96 during the classification of the cellular granuloma, and the fibro-cellular granuloma fibrosis levels indicating good classification.

### *Quadratic discriminant classification performance*

The confusion matrix of the quadratic discriminant classifier using diagonal covariance regularization is illustrated in Fig. 3. The ROC curves for the different classes during the staging process are illustrated in Fig. 4.

The confusion matrix in Fig. 3 indicates 83.3% accuracy of the quadratic discriminant classifier. Furthermore, Fig. 4 illustrates the ROC curves with the AUC values, showing that the quadratic discriminant

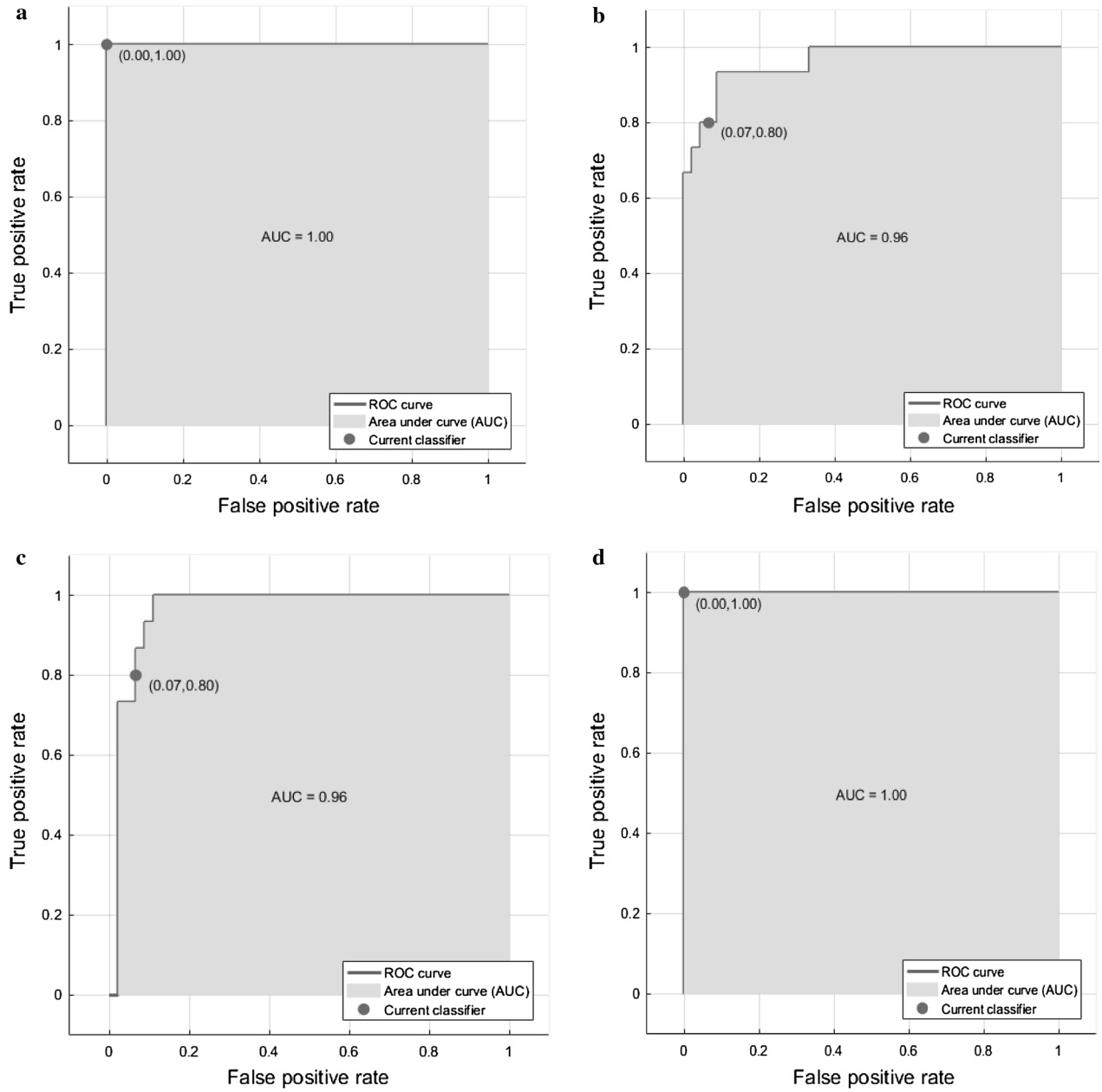


**Fig. 2** The ROC curves of the linear discriminant with the **a** normal liver case, **b** cellular granuloma, **c** fibro-cellular granuloma, and **d** fibrosis granuloma

classifier achieves the same perfect classification performance as the linear discriminant classifier with both the normal and the fibrosis granuloma. Nevertheless, the AUC = 0.90 during the cellular granuloma as well as the fibro-cellular granuloma fibrosis classification. This indicates the superiority of the linear discriminant classifier compared to the quadratic discriminant classifier performance during the staging process of the fibrosis levels 1 and 2.

#### Support vector machines staging performance

In the present work, several varieties of the SVM classifiers are used based on the used kernel. The performance evaluation in terms of the accuracy values using each type indicates that the linear SVM and the quadratic SVM achieve the superior accuracies of value 90% each compared to the other SVM types. The other SVM varieties realize the following accuracy values, 88.3, 85, 81.7, and 81.7% using the Cubic SVM, Fine Gaussian SVM,

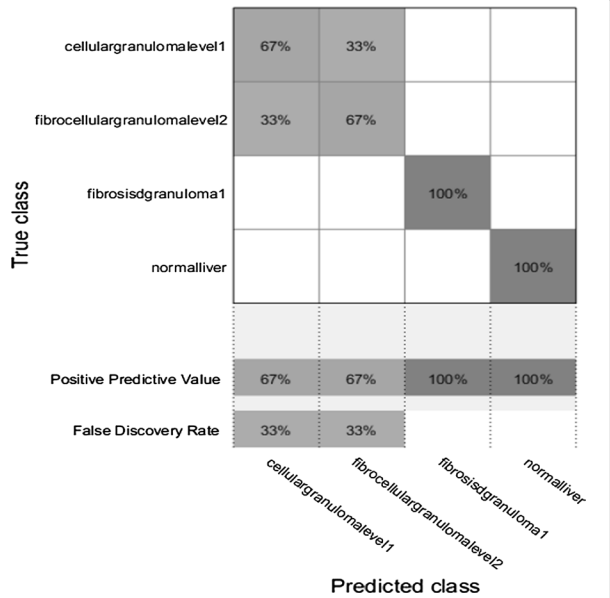


**Fig. 3** Confusion matrix of the quadratic discriminant showing the positive predictive values/false discovery rates

Medium Gaussian SVM, and the Coarse Gaussian SVM; respectively. Figure 5 illustrates the confusion matrices of both the linear SVM and the quadratic SVM using the 'one-versus-one' strategy; respectively.

Figure 5 indicates that each of the liner SVM, and quadratic SVM achieve 90% accuracy values. Figure 6 illustrates the ROC curve indicating the AUC values, where Fig. 6a1–d1 refers to the results of the normal liver case, cellular granuloma, fibro-cellular granuloma, and the fibrosis granuloma; respectively, when classified using the linear SVM. In addition, Fig. 6a2–d2 illustrate to the results of the normal liver case, cellular granuloma, fibro-cellular granuloma, and the fibrosis granuloma; respectively, when classified using the quadratic SVM.

Figure 6 illustrates the ROC curves with the AUC values, showing that the linear SVM and the quadratic SVM classifiers accomplish the same perfect classification of both the normal and the fibrosis granuloma with AUC = 1. Nevertheless, the linear SVM outperforms the quadratic SVM during the cellular granuloma as they achieve AUC = 0.96 and AUC = 0.91; respectively. Also, during the classification of the fibro-cellular granuloma fibrosis, the linear SVM has AUC = 0.95, AUC = 0.96, respectively, while the quadratic SVM has AUC = 0.94 in the classification of these two classes. This indicates the superiority of the linear SVM compared to the quadratic SVM classifier performance during the staging process.

**Performance evaluation comparative study**

Figure 7 illustrates radar graph showing the accuracies of the twenty-two employed classifiers to stage the fibrosis level compared to the normal liver cases in the present four-class classification problem.

Figure 7 reports that the maximum accuracy is 90% owing to the large size dataset as only 60 light microscopic samples are captured from the normal and the three levels of fibrosis (15 images from each class). Even though, the linear discriminant, linear SVM, quadratic SVM, and the subspace discriminant ensemble are considered the superior classifiers as they achieved the maximum accuracy of 90%. Additionally, the Cosine KNN, cubic SVM, medium KNN, and cubic KNN achieved 88.30% accuracy, while the Fine Gaussian SVM has 85% accuracy. Additionally, 81.70% accuracy values are obtained using the quadratic discriminant, bagged trees ensemble, medium Gaussian SVM, and the Coarse Gaussian SVM. However, all the decision tree classifiers (simple-, medium-, and complex-tree) along with the fine KNN and the weighted KNN achieved 78.30% accuracies, while the subspace KNN ensemble has 71.70% accuracy. Generally, three classifiers, namely Coarse KNN, Boosted trees ensemble, and RUSBoosted trees ensemble are failed in the fibrosis staging as they have 25%. In addition, Fig. 8 reports the computational processing time for the classifiers involved in the present study in terms of the prediction speed in observations/second.

Figure 8 depicts that the subspace KNN ensemble requires the least prediction speed of 44 observations/second, however, it achieves 71.7% accuracy, then the Subspace discriminant ensemble that requires 68 observations/second and achieves the superior accuracy of 90%. However, the maximum required prediction speed is 2700 observations/second in the case of the simple tree classifier, which also achieves 78.3% accuracy.

The preceding results reported the superiority of four classifiers, namely linear discriminant, linear SVM, quadratic SVM, and the subspace discriminant ensemble in terms of the classification accuracy, where each of the four classifiers achieved 90% classification accuracy. Consequently, another comparison is conducted in terms of the AUC and the prediction speed to determine the best classifier, which is related to the processing/computational time as reported in Table 2.

Table 2 establishes that in terms of the AUC for the classification/staging of the four classes, the linear discriminant classifier is considered the superior classifier as it achieves AUC = 1.0, 0.96, 0.96, 1.0 values, respectively, while the other classifiers achieved less AUC values in

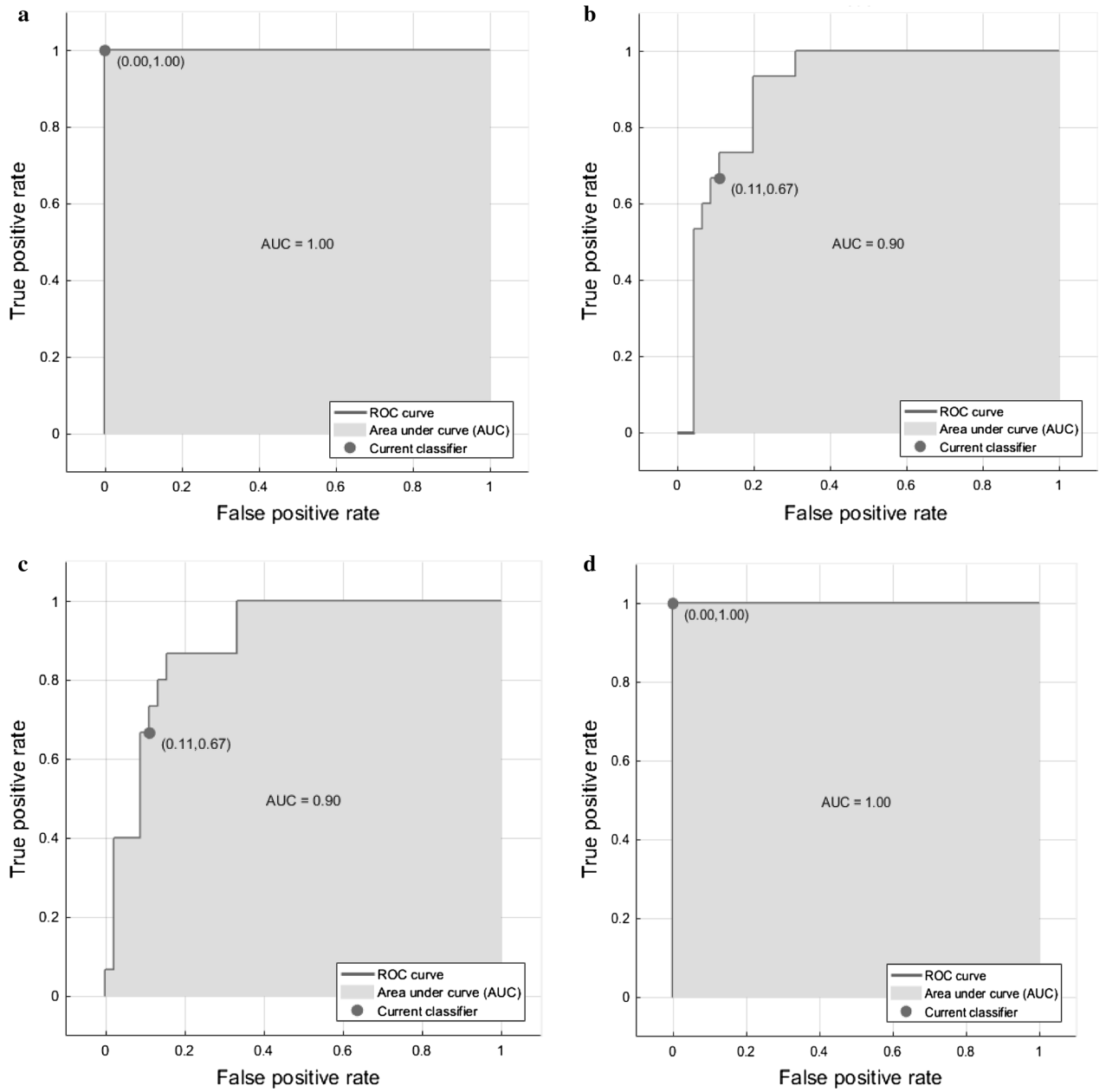


**Fig. 4** The ROC curves of the quadratic discriminant with the **a** normal liver case, **b** cellular granuloma, **c** fibro-cellular granuloma, and **d** fibrosis granuloma

the classification of the level 1 and level 2 in the fibrosis stages. However, in terms of the computational time, the subspace discriminant ensemble takes the least prediction speed of 68 observations per second, while the linear discriminant classifier takes the highest prediction speed of 860 observations per second.

Generally, the variation in the twenty-two classifiers' performance is owing to the ability of each classifier to handle the extracted statistical features, and the small size dataset in the current study. Consequently, only four classifiers had the top/ superior classification accuracy values of 90% compared to the remaining 18 classifiers, which involved in the present work, where the LDA, the linear SVM classifier, quadratic SVM classifier, and the subspace discriminant ensemble classifier achieved good performance on accuracy due to the characteristics of each classifier and their matching with the extracted features of the liver fibrosis images. Generally, the LDA

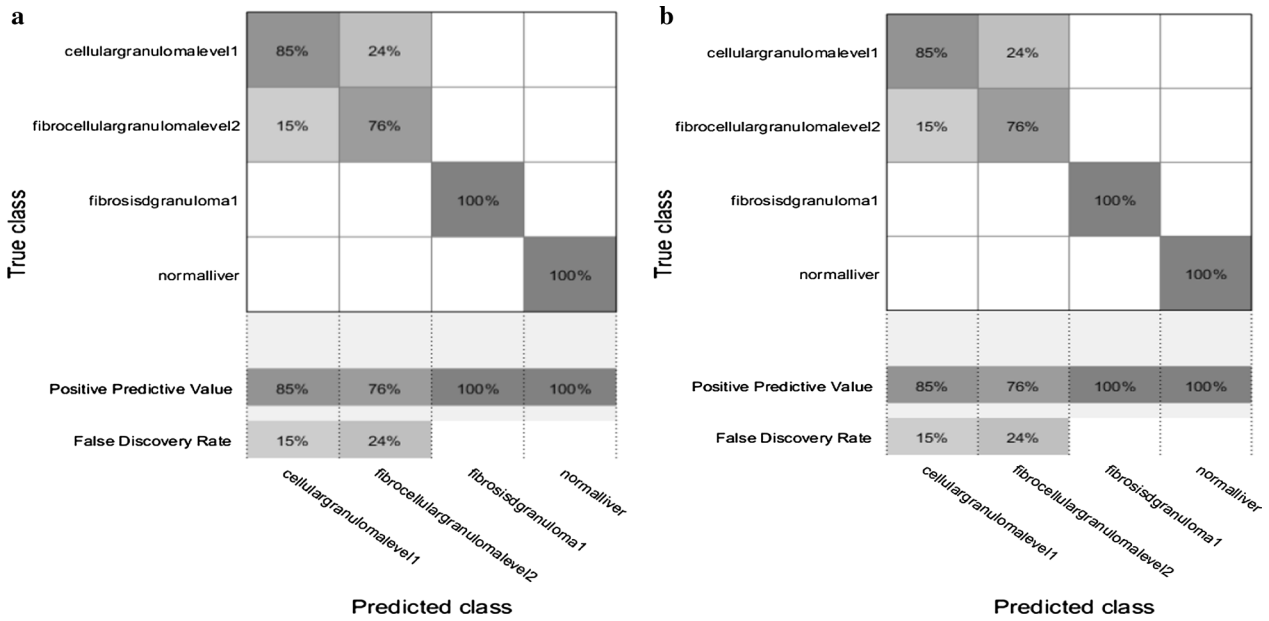


**Fig. 5** Confusion matrix showing the positive predictive values/false discovery rate of the **a** liner SVM, and **b** quadratic SVM

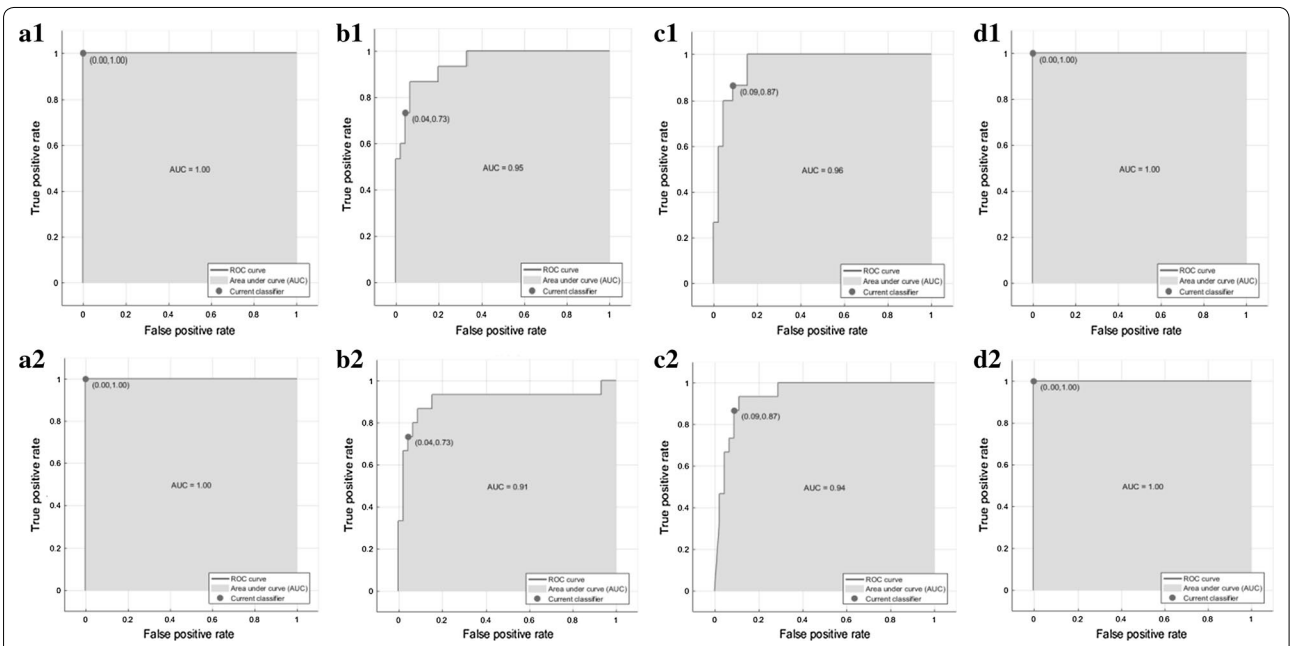


**Fig. 6** The ROC curves of the linear SVM (1st row) and quadratic SVM (2nd row) with the **a1**, **a2** normal liver case, **b1**, **b2** cellular granuloma, **c1**, **c2** fibro-cellular granuloma, and **d1**, **d2** fibrosis granuloma

provides more class separability as it draws between the classes a decision region. It finds the project axes to project the data samples of the dissimilar classes to be far from each other and to close the data samples of the same class. Hence, the LDA generates a linear combination of the data samples, which achieves the largest differences between the classes [23]. At the same time, the SVM determine decision boundaries according to the decision planes, which separate the set of features of the different classes. The SVM determines the hyper-plane that maximizes the separation margin between the different classes to divide the data space [24]. In addition, it is

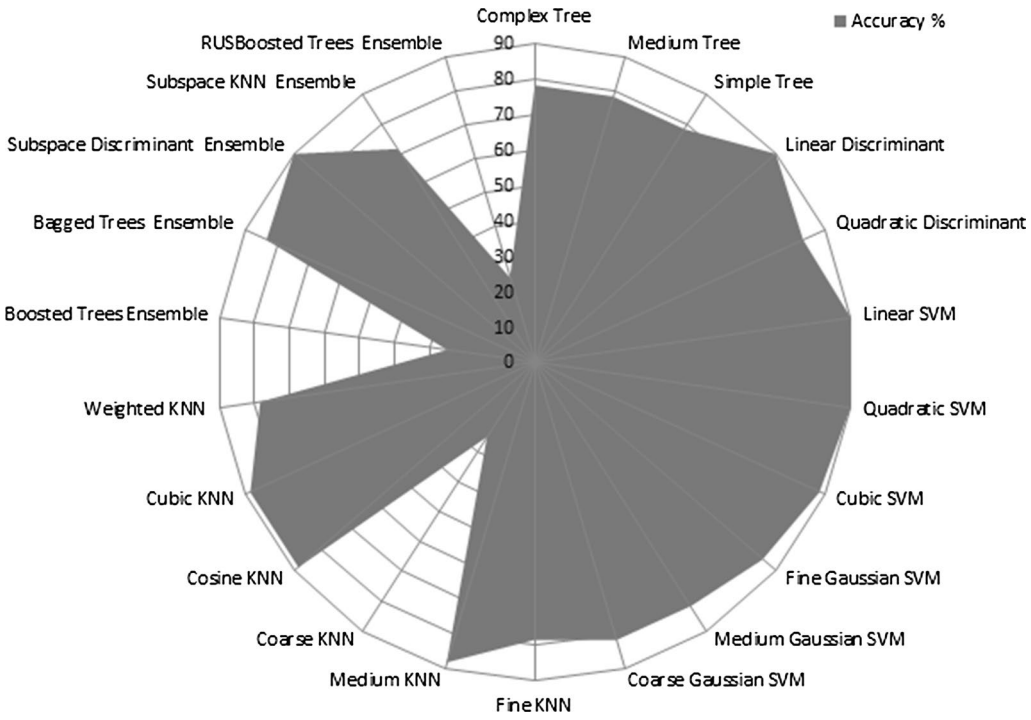


**Fig. 7** Radar graph of the different classifiers' accuracies

established that the classifiers which are based on linear functions, such as the linear SVM and the linear discriminant achieved the superior overall performance.

## Conclusion

Hepatic fibrosis is one of the serious diseases that has several stages and requires early detection and classification for accurate diagnosis and treatment. The current work was interested to determine the superior classifier in such medical problem for further implementation for a computer aided diagnosis system based on the automated classification. Consequently, mice animal models were used to capture microscopic images for liver fibrosis staging in schistosomiais. Twenty-two classifiers were employed in the current study after the analysis of the microscopic images to extract the statistical features.

The results demonstrate the superiority of the linear discriminant classifier, linear SVM classifier, quadratic SVM classifier, and the subspace discriminant ensemble classifier in terms of the classification accuracy, as they achieved 90% accuracy values. To be more specific, the linear discriminant classifier realized the superior performance in terms of the values of the AUC, while it took the highest computational time as it took prediction speed of 860 observations per second. However, the subspace discriminant ensemble took the least prediction speed of 68 observations per second.

Since the present study evaluated the classifiers using the significant statistical features only, namely the area, Feret, minor, and the RawIntDen of the selected region in the image, it is recommended to extract another features based on the morphology and the texture with integrating these features with the statistical ones. Furthermore, due to the superiority of the linear discriminant, linear SVM, quadratic SVM, and the subspace discriminant ensemble, it is recommended to test these classifiers at

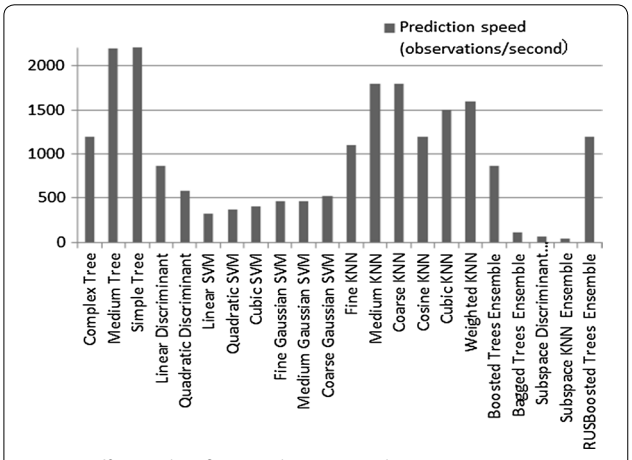


**Fig. 8** Different classifiers' prediction speed

**Table 2 AUC and the prediction speed of the classifiers**

| Classifier type | Class name | Area under curve (AUC) | Prediction speed (observations/second) |
|---|---|---|---|
| Linear discriminant | Class 1: normal liver | 1.0 | 860 |
| | Class 2: cellular granuloma (Level 1) | 0.96 | |
| | Class 3: fibro-cellular granuloma (Level 2) | 0.96 | |
| | Class 4: fibrosis granuloma (Level 3) | 1.0 | |
| Linear SVM | Class 1: normal liver | 1.0 | 320 |
| | Class 2: cellular granuloma (Level 1) | 0.95 | |
| | Class 3: fibro-cellular granuloma (Level 2) | 0.96 | |
| | Class 4: fibrosis granuloma (Level 3) | 1.0 | |
| Quadratic SVM | Class 1: normal liver | 1.0 | 370 |
| | Class 2: cellular granuloma (Level 1) | 0.91 | |
| | Class 3: fibro-cellular granuloma (Level 2) | 0.94 | |
| | Class 4: fibrosis granuloma (Level 3) | 1.0 | |
| Subspace discriminant ensemble | Class 1: normal liver | 1.0 | 68 |
| | Class 2: cellular granuloma (Level 1) | 0.94 | |
| | Class 3: fibro-cellular granuloma (Level 2) | 0.94 | |
| | Class 4: fibrosis granuloma (Level 3) | 1.0 | |

different parameters settings to improve their performance. In addition, other ensemble configurations can be used and compared to with results of this study. Using the segmented images before feature extraction is recommended instead of using the original images can lead to superior performance. Generally, large size dataset is recommended, which can inspire the use of the deep learning classifier for the fibrosis staging problem. Thus, other classifiers, other ensemble of classifiers or deep learning classifiers can be involved after extracting morphological features and their integration with those statistical features.

**Author details**
[1] Department of Electronics and Electrical Communication Engineering, Faculty of Engineering, Tanta University, Tanta, Egypt. [2] Department of Computer Science, University of Illinois at Springfield, Springfield, IL, USA.

**Acknowledgement**
The authors are thankful to Dr. Dalia Salah Ashour and Dr. Dina M. Abou Rayia, Department of Medical Parasitology, Faculty of Medicine, Tanta University, Egypt, for performing the parasitology part of the study and providing us with the used microscopic images dataset at the different fibrosis stages as well as the normal case.

Received: 29 May 2018 Accepted: 7 August 2018
Published online: 20 August 2018

**References**

1. Burke ML, Jones MK, Gobert GN, Li YS, Ellis MK, McManus DP. Immunopathogenesis of human schistosomiasis. Parasite Immunol. 2009;31(4):163–76.
2. Lambertucci JR, Cota GF, Pinto-Silva RA, Serufo JC, Gerspacher-Lara R, Drummond SC, et al. Hepatosplenic schistosomiasis in field-based studies: a combined clinical and sonographic definition. Mem Inst Oswaldo Cruz. 2001;96:147–50.
3. Hotez PJ, Savioli L, Fenwick A. Neglected tropical diseases of the Middle East and North Africa: review of their prevalence, distribution, and opportunities for control. PLoS Negl Trop Dis. 2012;6(2):e1475.
4. Andrade ZDA. Schistosomiasis and liver fibrosis. Parasite Immunol. 2009;31(11):656–63.
5. Gailhouste L, Le Grand Y, Odin C, Guyader D, Turlin B, Ezan F, et al. Fibrillar collagen scoring by second harmonic microscopy: a new tool in the assessment of liver fibrosis. J Hepatol. 2010;52(3):398–406.
6. Mahmoud-Ghoneim D. Optimizing automated characterization of liver fibrosis histological images by investigating color spaces at different resolutions. Theor Biol Med Model. 2011;8(1):25.
7. Stanciu SG, Xu S, Peng Q, Yan J, Stanciu GA, Welsch RE, et al. Experimenting liver fibrosis diagnostic by two photon excitation microscopy and bag-of-features image classification. Sci Rep. 2014;4:4636.
8. Ali S, Smith KA. On learning algorithm selection for classification. Appl Soft Comput. 2006;6(2):119–38.
9. Cinque L, De Santis A, Di Giamberardino P, Iacoviello D, Placidi G, Pompili S, et al. Design of a classification strategy for light microscopy images of the human liver. In: International conference on image analysis and processing. Cham: Springer; 2017. p. 626–636.
10. Mika S, Ratsch G, Weston J, Scholkopf B, Mullers K.R. Fisher discriminant analysis with kernels. In: Proceedings of the 1999 IEEE signal processing society workshop on neural networks for signal processing IX, 1999. p. 41–48.
11. Dudoit S, Fridlyand J, Speed TP. Comparison of discrimination methods for the classification of tumors using gene expression data. J Am Stat Assoc. 2002;97(457):77–87.
12. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning, vol. 112. New York: Springer; 2013.
13. Kuhn M, Johnson K. Applied predictive modeling, vol. 26. New York: Springer; 2013.
14. McLachlan G. Discriminant analysis and statistical pattern recognition, vol. 544. New York: Wiley; 2004.
15. Steinwart I, Christmann A. Support vector machines. New York: Springer; 2008.
16. Duan KB, Keerthi SS. Which is the best multiclass SVM method? An empirical study. In: International workshop on multiple classifier systems. Berlin: Springer; 2005. p. 278–285
17. Boiman O, Shechtman E, Irani M. In defense of nearest-neighbor based image classification. In: IEEE conference on computer vision and pattern recognition, 2008. CVPR 2008. IEEE; 2008. p. 1–8.

18. Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. IEEE Trans Geosci Remote Sens. 2004;42(8):1778–90.
19. Sarkaleh AK, Poorahangaryan F, Zanj B, Karami A. A neural network based system for Persian sign language recognition. In: 2009 IEEE international conference on signal and image processing applications (ICSIPA). IEEE; 2009. p. 145–149
20. Dietterich T.G. Ensemble methods in machine learning. In: International workshop on multiple classifier systems. Berlin: Springer; 2000. p. 1–15
21. Kuncheva LI, Rodríguez JJ, Plumpton CO, Linden DE, Johnston SJ. Random subspace ensembles for fMRI classification. IEEE Trans Med Imaging. 2010;29(2):531–42.
22. Fraz MM, Remagnino P, Hoppe A, Uyyanonvara B, Rudnicka AR, Owen CG, Barman SA. An ensemble classification-based approach applied to retinal blood vessel segmentation. IEEE Trans Biomed Eng. 2012;59(9):2538–48.
23. Panahi N, Shayesteh MG, Mihandoost S, Varghahan BZ. Recognition of different datasets using PCA, LDA, and various classifiers. In: 2011 5th international conference on application of information and communication technologies (AICT). IEEE; 2011. pp. 1–5.
24. El-Naqa I, Yang Y, Wernick MN, Galatsanos NP, Nishikawa RM. A support vector machine approach for detection of microcalcifications. IEEE Trans Med Imaging. 2002;21(12):1552–63.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.