RESEARCH

Ensemble of subspace discriminant classifiers for schistosomal liver fibrosis staging in mice microscopic images

Amira S. Ashour[1], Yanhui Guo[2*] , Ahmed Refaat Hawas[1] and Guan Xu[3]

Abstract

Schistosomiasis is one of the dangerous parasitic diseases that affect the liver tissues leading to liver fibrosis. Such disease has several levels, which indicate the degree of fibrosis severity. To assess the fibrosis level for diagnosis and treatment, the microscopic images of the liver tissues were examined at their different stages. In the present work, an automated staging method is proposed to classify the statistical extracted features from each fibrosis stage using an ensemble classifier, namely the subspace ensemble using linear discriminant learning scheme. The performance of the subspace/discriminant ensemble classifier was compared to other ensemble combinations, namely the boosted/trees ensemble, bagged/trees ensemble, subspace/KNN ensemble, and the RUSBoosted/trees ensemble. The simulation results established the superiority of the proposed subspace/discriminant ensemble with 90% accuracy compared to the other ensemble classifiers.

Keywords: Liver fibrosis, Schistosomiasis, Ensemble classifier, Statistical features

Introduction

Schistosomiasis is a serious disease triggered by parasitic flatworms called Schistosomes, which are widely spread in the developing countries due to the contaminated water. Early diagnosis saves the patient's life, which is identified by the existence parasite's eggs in the stool/urine of the individual and can be confirmed by discovering antibodies in the blood [1]. This disease causes liver fibrosis that can be assessed quantitatively and automatically using microscopic image analysis for detecting the liver fibrosis stage and minimizing the inter-observation variations [2]. For automated quantitative assessment of liver fibrosis, Sun et al. [3] used nonlinear optical microscopy. Mabey et al. [4] used the tissue and cellular information to identify the fibrosis progression based on the microscopic images.

Recently, for liver tissues classification, the artificial intelligence procedures were employed for image processing and computer-aided diagnosis. From the histological images, Mahmoud-Ghoneim [5] optimized the computerized features of the liver fibrosis by inspecting the three color spaces at different resolutions for texture classification, where classification is a supervised remarkable machine learning process. Several techniques can be used for classification, including the k-nearest neighbor (KNN), neural network, support vector machine (SVM), and the decision tree [6].

A standard practice for confirming the fibrosis level and screening is to examine the microscopic images of the liver tissue samples. From optical microscopy images, Saito et al. [7] implemented an automated approach for intestinal parasites based on a pattern classifier using active learning procedures. In order to achieve accurate diagnosis, the ensemble methodology that weighs and combines some individual classifiers can be applied to attain a classifier, which outperforms the individual classifiers included in the ensemble. Rathore et al. [8] implemented an ensemble classification procedure using the discriminatory abilities of information rich hybrid feature spaces in colon biopsy microscopic images. Based on majority voting, an ensemble classifier, including linear,

*Correspondence: guoyanhui@gmail.com
[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

Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 2 of 10

sigmoid SVM, and radial basis function, was applied to classify the microscopic images using the selected features. Early detection and diagnosis of liver fibrosis are still challenging tasks. Worldwide, several researchers are inspired to effectively determine the liver fibrosis stage. However, according to the previous studies, very few automated image-based classifiers have been reported. Furthermore, there is no such ensemble methodology has not been included for liver fibrosis staging.

Consequently, the current work applied an ensemble of subspace and discriminant classifiers on the microscopic images from mice as animal model liver samples of the different fibrosis stages for liver fibrosis staging. The proposed ensemble classifier used the extracted statistical features. Moreover, a comparative study of different ensembles, namely the boosted/trees ensemble, bagged/trees ensemble, subspace/KNN ensemble, and the RUS-Boosted/trees ensemble, was also included.

The structure of the remaining sections is as follows. Section 2 includes the methodology and the proposed method in the present work. Section 3 reports the obtained results with comparative studies. Finally, Sect. 4 concludes the proposed study.

Methodology

The proposed staging system consists of the following phases: (i) preprocess the acquired microscopic liver images for normal and different fibrosis levels, (ii) extract the statistical features, and (iii) apply the ensemble classifier to classify the liver image to any of the four cases, namely normal liver tissue, cellular granuloma, fibrocellular granuloma, or fibrotic granuloma.

Image preprocessing

The captured samples from the normal liver tissues as well as the three fibrosis levels are preprocessed. The preprocessing and segmentation steps were performed using ImageJ software tools. Initially, the colored microscopic images are converted to grayscale image. Then, the thresholding is used to identify the fibrosed regions, and then the watershed of the Euclidian distance map (EDM) segmentation method is applied to the microscopic images. During the segmentation process, the EDM is measured and the ultimate eroded points (UEPs) are located, and then dilates each UEPs. Afterwards, the statistical features are extracted from the segmented images.

Statistical features

In the present work, the statistical features of the different samples at the different fibrosis levels are extracted which are the area, perimeter, circularity, mean, median, mode, Feret, and the IntDen of the fibrosis regions in the microscopic images. The most prominent features are selected to distinguish the four classes for further classification process. These selected features are namely the (i) the 'minor', which is the secondary axis of the best fitting ellipse of the fibrosis region, (ii) the 'Feret', which is the Feret's diameter defined as the longest distance between any two points on the boundary of the selected fibrosis region, (iii) the 'area', which is the area of fibrosis/selected region in square pixels based on the calibration unit, and (iv) the 'RawIntDen', which is the integrated density defined as the sum of the pixel values within the fibrosis selected region. Subsequently, the ensemble of the subspace and discriminant classifiers is deployed to classify the normal liver case and the different fibrosis stages.

Ensemble classifier based liver fibrosis staging

A classification process based on the features similarity is used to classify the liver fibrosis stages. In the current work, an ensemble of classifiers is proposed for labeling each microscopic liver image as normal or one of the fibrosis levels according to the selected statistical features.

Typically, the multiple-classifier techniques or the ensemble-based techniques are more desirable compared to their single-classifier counterparts as they reduce the poor selection possibility [7]. The ensemble classifier combines a set of classifiers that might produce superior classification performance compared to each individual classifier. The ensemble of classifiers is categorized generally into (i) classifier selection, where only the output of the classifier with the preeminent performance is selected as the final output, or (ii) classifier fusion, where the outputs of the individual classifiers are combined to determine the final decision as the individual classifiers are trained in parallel [8]. To select the final class label from the individual ones, precise predefined rules are applied. The most combination rules include the weighted majority voting, majority voting, Borda count, and behavior knowledge space common [9]. The selection of the ensemble size (number of classifiers in the ensemble) involves a balance between the accuracy and speed of the classifier, where over-trained classification may occur with too large ensembles and larger ensembles take longer training time for prediction.

Ensemble learning combines several models for improving the prediction performance, which has several approaches, such as (i) random subspace, which randomizes the learning algorithm by selecting a subset of features randomly (chosen subspace) before performing the training algorithm, and then the models' outputs are combined by majority vote, (ii) bagging (Bootstrap Aggregation), which creates a set of models that trained on a random data, then the predictions are aggregated/combined for final prediction using averaging, and (iii)

Ashour *et al. Health Inf Sci Syst* (2018) 6:21

Page 3 of 10

boosting is based on averaging/voting of multiple models, where it weights the constructed models based on their performance. In the current work, the majority voting rule is used with the subspace ensemble through linear discriminant.

Subspace discriminant ensemble

Subspace learning techniques have a significant role; especially with the linear discriminant analysis (LDA) scheme that engaged to determine a specific discriminant subspace of low-dimension [10–12]. Several studies were conducted to study effect of the different subspacing, weighting, and resampling techniques on the classification performance in the ensemble learning [13–15]. Ho [16] used random subspaced feature arrangements using the random subspace method (RSM) using a random sample of features to construct each learner for decreasing the error rates [17]. Nevertheless, this random selection of the features in the subspaces is considered the main shortcoming of the RSM, where poor discrimination ability may occur due to the random selection of the subsets in some cases. In this case, the final ensemble decision becomes poor. To decrease this drawback of the RSM, a majority voting (MV) method is used. Generally, a single classifier in the ensemble might use only a small part of the features from the feature space. In addition, each classifier has the ability to classify any new/unknown instance. The MV method uses each classifier to separately predict the new/unknown instance's class. Afterwards a majority vote between the predictions is employed to adopt the final class of the instance (final classification result). In this work, a framework based on the discriminant learning is applied to classify the fibrosis levels and the normal case using subspaces, which are the main elements of the learning algorithm.

The RSM ensemble construction methods using a modified feature space is considered to build the ensembles of learners, unlike boosting and bagging ensemble methods [18]. Typically, the individual classifiers are constructed using the subset of features. In the present work, the steps of the used RSM technique are illustrated as follows.

The classifiers' outputs in the proposed procedure are combined with the MV method. In the MV, unlabeled (new/unknown) instance classification is performed based on the class that has the most frequent vote (the highest number of votes) from the classifiers in the ensemble. The description of the MV is as follows:

$$Class(a) = \underset{c_i \in dom(y)}{\arg \max} \left(\sum_v h\big(y_v(a), c_i\big) \right) \tag{1}$$

where $y_v(a)$ is the classification of the classifier 'v' and $h\big(y_v(a), c_i\big)$ represents an indicator function, which is given by:

$$h\big(y_v(a), c_i\big) = \begin{cases} 1 & y = c \\ 0 & y \neq c \end{cases} \tag{2}$$

Experimental results and discussion

In the present work, Schistosoma mansoni cercariae was used to infect the mice in the Parasitology Department, Faculty of Medicine, Tanta University, Egypt. Afterwards 60 microscopic images of liver sections at different fibrosis levels were captured (15 images from each class), namely (i) level 1 (cellular granuloma), (ii) level 2 (fibrocellular granuloma), and (iii) level 3 (fibrotic granuloma) along with normal samples. Figure 1 illustrates samples from each fibrosis level and the steps mentioned previously in order to extract the statistical features.

Performance evaluation of the proposed subspace discriminant

The subspace discriminant ensemble was designed using the majority voting rule, where the random subspace ensemble method was used with linear discriminant learner type of 30 learners and two subspace dimension. The confusion matrix is illustrated in Fig. 2. The ROC curves are demonstrated in Fig. 3a through d for the normal and three fibrosis levels; respectively.

Random Subspace Method Procedure

Start

    Select T random subsets with c classes

    Construct c^t classifiers
    Construct random subspace ensemble model
    Combine the classifiers and ***assign*** each instant to the corresponding class using the majority voting

End

Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 4 of 10

Fig. 1 **a1**–**a3** original image, **b1**–**b3** gray scale image, **c1**–**c3** segmented image using Watershed
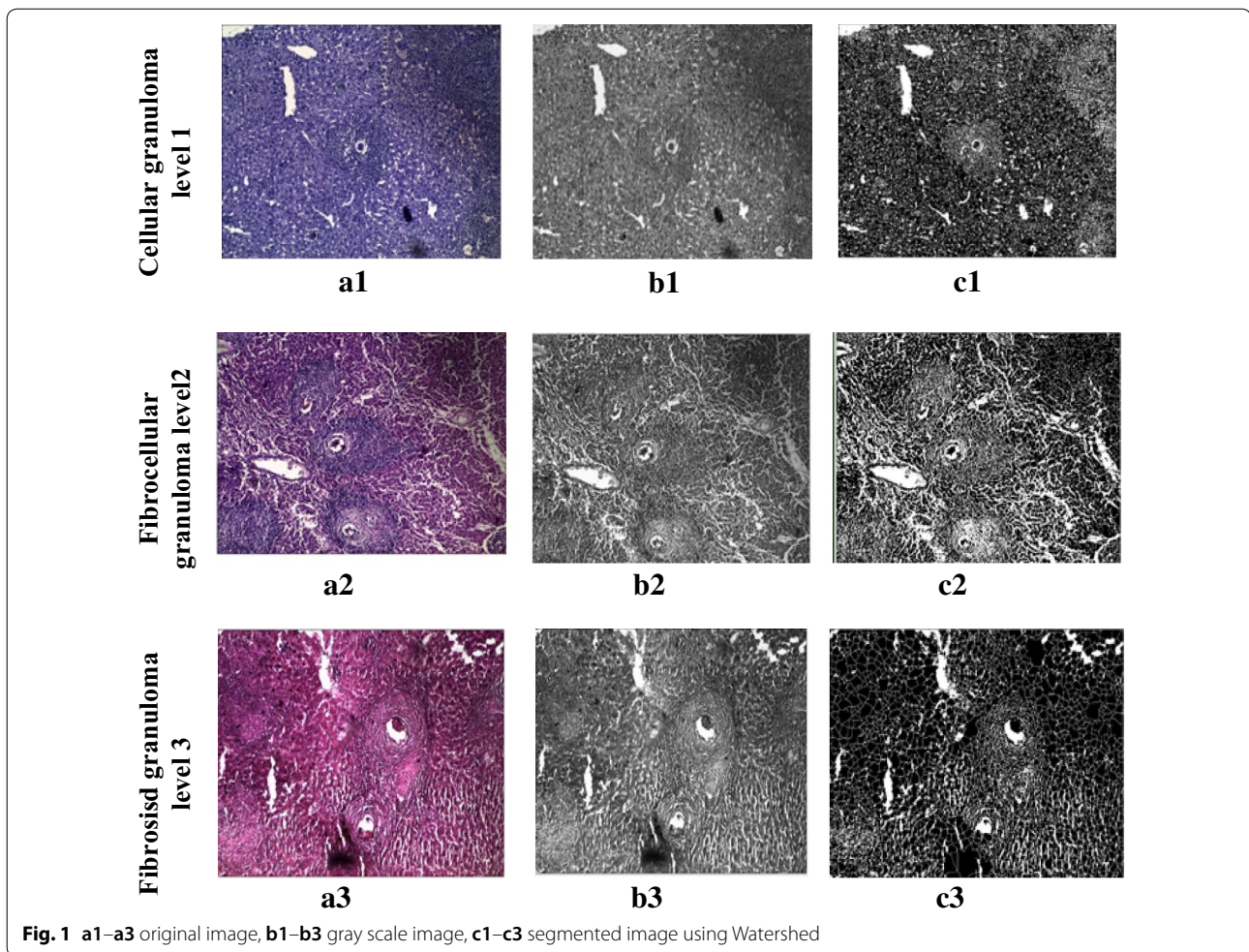
Figure 3 illustrates the ROC curve that represents (i) the false positive rate (FPR), which indicates the number of the incorrect positive results with respect to all the negative instances during the test and (ii) the true positive rate (TPR), which represents the number of correct positive results with respect to all positive instances. Typically, the classification accuracy is measured by AUC curve. Figure 3 reports that the proposed classifier achieved perfect classification with both the normal and fibrosis at level 3, while good classification with AUC = 0.94 during the classification of fibrosis cases at levels 1 and 2. These results are owing to the absence of the fibrosis and granulomas in the normal cases and the very big area of the fibrosis granuloma, while, in level 1

and 2 cellular- and fibrocellular- granuloma exist; respectively. The preceding results reported 90% accuracy, where the prediction speed was 68 observation/second.

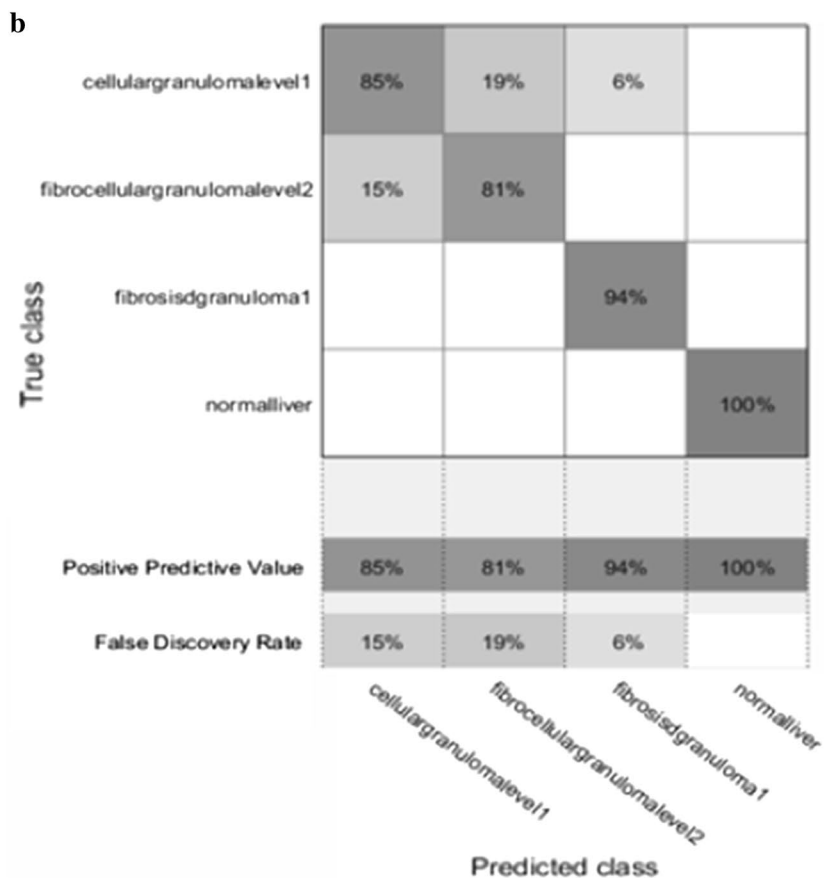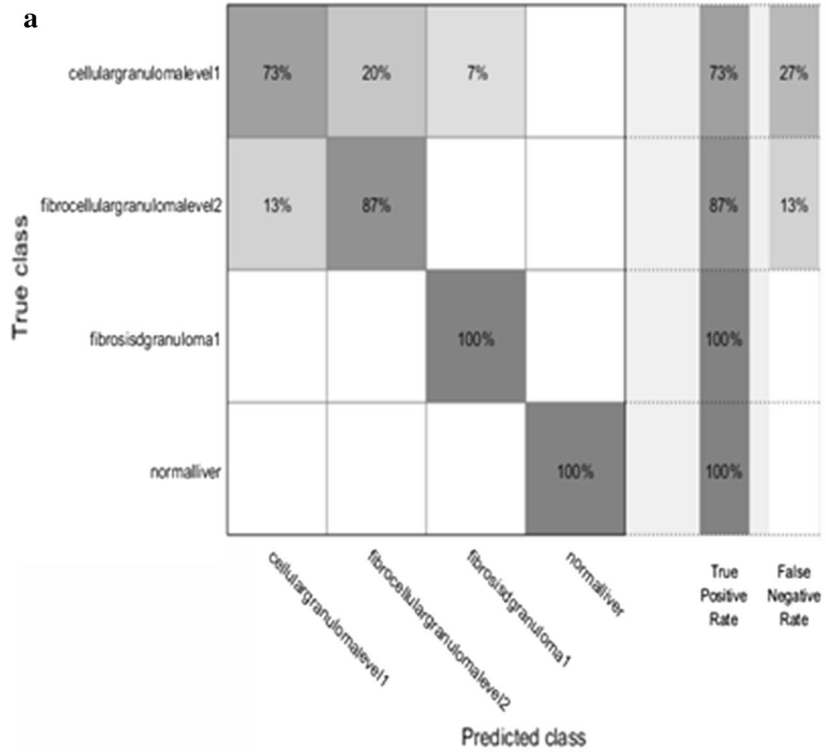## Comparative study with different classifiers of ensemble and neural network
A comparative study is conducted on different ensemble classifiers in terms of the classifiers' accuracies as follows.
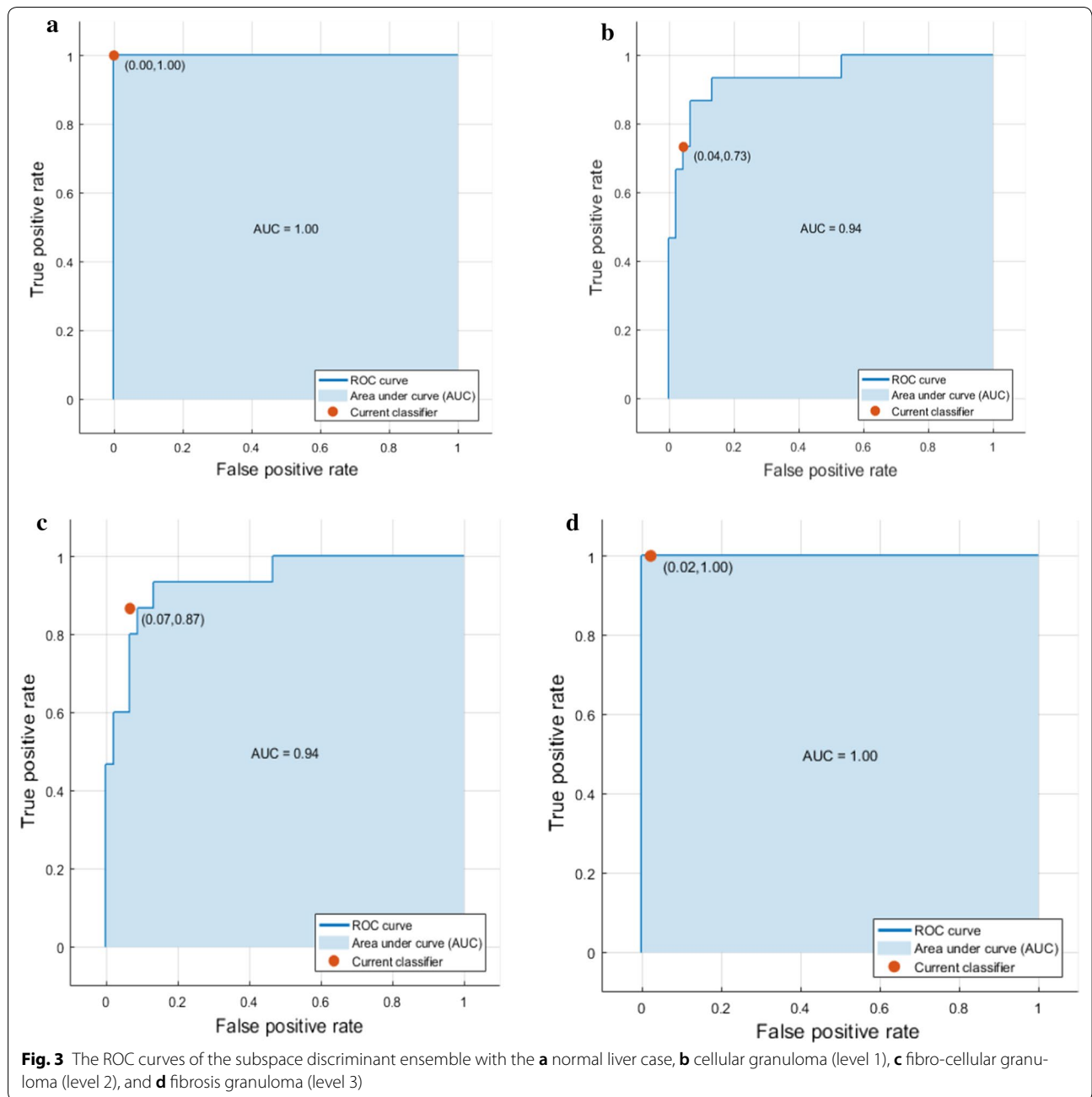
Bagged trees ensemble
The weight average rule uses the bag ensemble method with Decision tree learner type and 30 learners. The achieved results established 81.7% accuracy with prediction speed of 110 observation/second. The confusion

(See figure on next page.)
Fig. 2 Confusion matrix of the proposed subspace discriminant ensemble **a** true positive rates/false negative rates, and **b** positive predictive values/false discovery rates
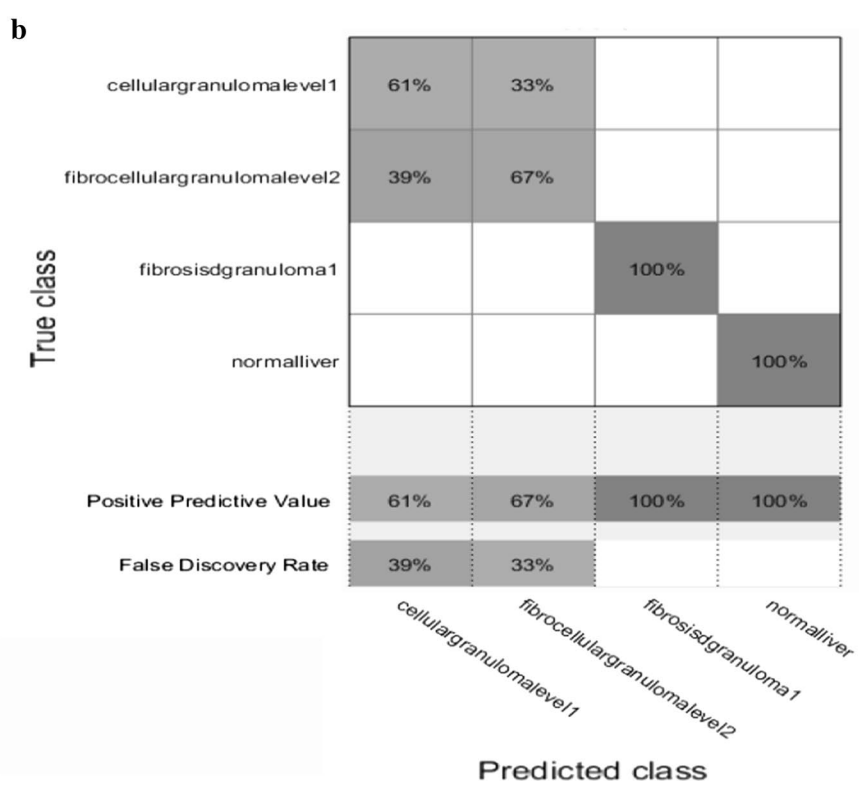
Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 5 of 10

a

b

Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 6 of 10

Fig. 3 The ROC curves of the subspace discriminant ensemble with the **a** normal liver case, **b** cellular granuloma (level 1), **c** fibro-cellular granuloma (level 2), and **d** fibrosis granuloma (level 3)
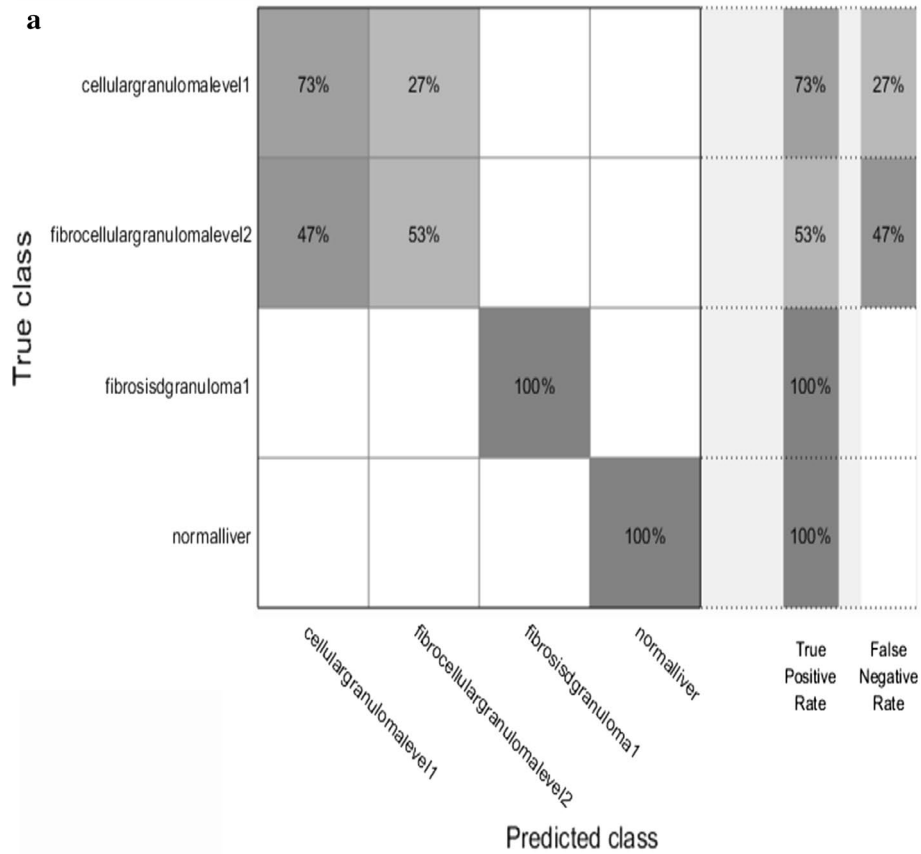
matrix results showing the true positive rates/false negative rates and the positive predictive values/false discovery rates are illustrated in Fig. 4. In addition, the ROC curves are demonstrated in Fig. 5a through d for the normal and three fibrosis levels; respectively.
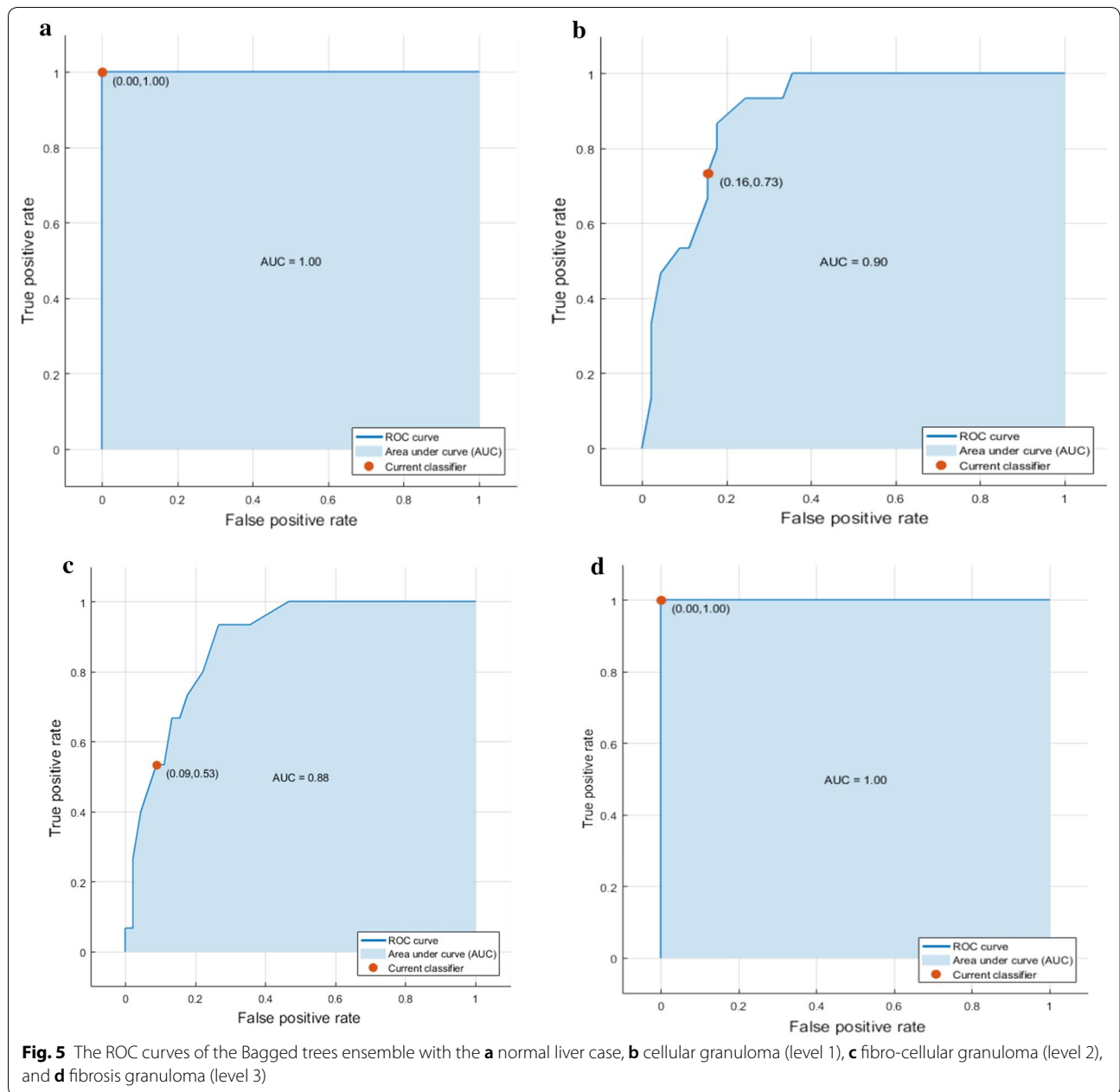
Subspace KNN ensemble
Subspace KNN, where the training parameters in this study are based on the simple Majority Vote rule with the Subspace ensemble method as in the proposed method. However, the learner type is Nearest Neighbor of 30 numbers of learners and 2 subspace dimensions. The performance of

(See figure on next page.)
Fig. 4 Confusion matrix of the bagged trees ensemble **a** true positive rates/false negative rates, and **b** positive predictive values/false discovery rates

Ashour *et al. Health Inf Sci Syst* (2018) 6:21

Page 7 of 10

Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 8 of 10

Fig. 5 The ROC curves of the Bagged trees ensemble with the **a** normal liver case, **b** cellular granuloma (level 1), **c** fibro-cellular granuloma (level 2), and **d** fibrosis granuloma (level 3)

this classifier is 73.3% accuracy with prediction speed of 44 observation/second.

Boosted trees ensemble

Boosted Trees, where the training parameters in this study are based on the Weighted Majority vote rule with the AdaBoost ensemble method. The learner type is Decision tree with maximum number of splits is 20, number of learners 30 and learning rate is 0.1. The performance

of this classifier is 25% accuracy with prediction speed of 870 observation/second.

RUSBoosted trees ensemble

RUSBoosted trees, where the training parameters in this study are Combined RUS and standard boosting procedure of AdaBoost with RUSBoost ensemble method. The learner type is the decision tree with maximum number of splits is 20 and number of learners 30 and learning rate

Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 9 of 10

is 0.1. The performance of this classifier is 25% accuracy with prediction speed of 1200 observation/second.

Multi-layer perceptron neural network

In addition, a comparison is conducted with the neural network of multi-layer perceptron neural network (MLP-NN) of one hundred hidden neurons. The NN realized accuracy of 88.3% to classify the different liver fibrosis levels as well as the normal case.

Comparative study evaluation

The reporting of the accuracy percentages of the preceding classifiers to discriminate between the normal case and the three liver levels staging is illustrated in Table 1.

Table 1 reports that both the boosted trees ensemble and the RUSBoosted trees ensemble classifiers failed to classify the fibrosis levels. However, the MLP-NN accomplished 83% accuracy, which is superior to the subspace KNN ensemble and the bagged trees ensemble. Generally, the proposed random subspace discriminant ensemble achieved the best accuracy of 90% value. These results illustrated that bagging provides better performance than boosting, and the RSM outperforms them both and the MLP-NN. Additionally, in terms of the computational time, the subspace KNN ensemble took the least computational time as it has prediction speed of 44 observation/second, while the RUSBoosted trees ensemble took the longest computational time as it has prediction speed of 1200 observation/sec. However, the proposed subspace discriminant ensemble took reasonable computational time as it has prediction speed was 68 observation/second. The superiority of the RSM classification is due to its ability to handle small dataset (samples) size due to its random subspaces process. However, bagging suffers from a shifting effect on the generalization error on small training sample sizes, also boosting failed to classify the small size dataset as it handles only large training

sample sizes [19]. Thus, it is recommended to conduct a comparative study on larger dataset with different classifier types.

Conclusions

This work offers significant contribution for liver fibrosis staging in schistosomiasis. The microscopic image analysis based on the statistical features was followed by using different ensemble of classifiers as well as the MLP-NN techniques and employed an ensemble of subspace discriminant classifiers for liver fibrosis staging. The results proved that the proposed random subspace discriminant ensemble realized the best accuracy of 90% compared to the other classifiers. In future, it is recommended to employ other ensemble rules and to increase the dataset size of the microscopic images. Furthermore, the morphological features can combined with the statistical features to realize better staging performance. In addition, the conventional neural network [20, 21] can be employed and compared with the proposed method.

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. [3] Department of Radiology, University of Michigan Medical School, Ann Arbor, USA.

References
1. Chaves NJ, Gibney KB, Leder K, O'brien DP, Marshall C, Biggs BA. Screening practices for infectious diseases among Burmese refugees in Australia. Emerging Infectious Dis. 2009;15(11):1769.
2. Xia JL, Dai C, Michalopoulos GK, Liu Y. Hepatocyte growth factor attenuates liver fibrosis induced by bile duct ligation. The American J Pathol. 2006;168(5):1500–12.
3. Sun W, Chang S, Tai DC, Tan N, Xiao G, Tang H, Yu H. Nonlinear optical microscopy: use of second harmonic generation and two-photon microscopy for automated quantitative liver fibrosis studies. J Biomed Opt. 2008;13(6):064010.
4. Mabey D, Peeling RW, Ustianowski A, Perkins MD. Tropical infectious diseases: diagnostics for the developing world. Nat Rev Microbiol. 2004;2(3):231.
5. Mahmoud-Ghoneim D. Optimizing automated characterization of liver fibrosis histological images by investigating color spaces at different resolutions. Theor Biol Med Modell. 2011;8(1):25.
6. Ali S, Smith KA. On learning algorithm selection for classification. Appl Soft Comput. 2006;6(2):119–38.
7. Kuncheva LI. Combining pattern classifiers: methods and algorithms. New York: Wiley; 2004.
8. Woods K, Kegelmeyer WP, Bowyer K. Combination of multiple classifiers using local accuracy estimates. IEEE Trans Pattern Anal Mach Intell. 1997;19(4):405–10.

Table 1 Accuracy percentage of the different classifiers compared to the proposed ensemble

General classifier type	Classifier	Accuracy %
Neural network	MLP-NN	83
Ensembles	Boosted/trees	25
	RUSBoosted/trees	25
	Subspace/KNN	73.3
	Bagged/trees	81.7
	Subspace/discriminant (proposed)	90

Ashour *et al. Health Inf Sci Syst (2018) 6:21*

Page 10 of 10

9. Polikar R. Ensemble based systems in decision making. IEEE Circuits Syst Mag. 2006;6(3):21–45.
10. Polikar R. Ensemble based systems in decision making. IEEE Circuits Syst Mag. 2006;6(3):21–45.
11. Zhang C, Ma Y, editors. Ensemble machine learning: methods and applications. New York: Springer Science & Business Media; 2012.
12. Rahman A, Verma B. Cluster-based ensemble of classifiers. Exp Syst. 2013;30(3):270–82.
13. Tao D, Tang X, Li X, Wu X. Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. IEEE Trans Pattern Anal Mach Intell. 2006;28(7):1088–99.
14. García-Pedrajas N, Ortiz-Boyer D. Boosting random subspace method. Neural Netw. 2008;21(9):1344–62.
15. Kotsiantis S. Combining bagging, boosting, rotation forest and random subspace methods. Artif Intell Rev. 2011;35(3):223–40.
16. Ho TK. The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell. 1998;20(8):832–44.
17. 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.
18. Panov P, Džeroski S. Combining bagging and random subspaces to create better ensembles. In: International Symposium on Intelligent Data Analysis. Springer, Berlin, Heidelberg; 2007. pp. 118-129.
19. Skurichina M, Duin RP. Bagging, boosting and the random subspace method for linear classifiers. Pattern Anal Appl. 2002;5(2):121–35.
20. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M,… & Sánchez CI. A survey on deep learning in medical image analysis. Medical Image Anal. 2017;42:60–88.
21. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng. 2017;19:221–48.