Appl Intell (2017) 47:397–408
DOI 10.1007/s10489-017-0897-0

Moth-flame swarm optimization with neutrosophic sets for automatic mitosis detection in breast cancer histology images

Gehad Ismail Sayed[1,2] · **Aboul Ella Hassanien**[1,2]

Published online: 23 March 2017

Abstract This paper presents an automatic mitosis detection approach of histopathology slide imaging based on using neutrosophic sets (NS) and moth-flame optimization (MFO). The proposed approach consists of two main phases, namely candidate's extraction and candidate's classification phase. At candidate's extraction phase, Gaussian filter was applied to the histopathological slide image and the enhanced image was mapped into the NS domain. Then, morphological operations have been implemented to the truth subset image for more enhancements and focus on mitosis cells. At candidate's classification phase, several features based on statistical, shape, texture and energy features were extracted from each candidate. Then, a principle of the meta-heuristic MFO algorithm was adopted to select the best discriminating features of mitosis cells. Finally, the selected features were used to feed the classification and regression tree (CART). A benchmark dataset consists of 50 histopathological images was adopted to evaluate the performance of the proposed approach. The adopted dataset consists of five distinct breast pathology slides. These slides were stained with H&E acquired by Aperio XT scanners with 40-x magnification. The total number of mitoses in 50 database images is 300, which were annotated by an expert pathologist. Experimental results reveal the capability of the MFO feature selection algorithm for finding the optimal feature subset which maximizing the classification performance compared to well-known and other meta-heuristic feature selection algorithms. Also, the high obtained value of accuracy, recall, precision and f-score for the adopted dataset prove the robustness of the proposed mitosis detection and classification approach. It achieved overall 65.42 % f-score, 66.03 % recall, 65.73 % precision and accuracy 92.99 %. The experimental results show that the proposed approach is fast, robust, efficient and coherent. Moreover, it could be used for further early diagnostic suspicion of breast cancer.

Keywords Computer aided diagnosis system · Cancer imaging · Neutrosophic sets · Moth-flame optimization

1 Introduction

Recently, breast cancer is considered as one of the leading causes of death. It is the most common among women in the most developing regions of the world with nearly one million new cases each year [26]. During the last decade, histological image grading has become widely used as an important prognosis indicator. The grading depends on the microscopic breast cancer cells similarity to normal breast tissue. It classifies cancer, which reflects progressively less normal appearing cells that have a worsening prognosis for low grade, intermediate grade, and high-grade [8]. There is a well-known system, namely Nottingham Grading System (NGS) (also called Elton-Ellis) for breast cancer grading [7]. NGS adds up scores of nuclear pleomorphic, tubule formation and mitotic count to grade breast carcinomas. For each of these criteria, the scores are added together to give a final overall score and its corresponding grade [6].

✉ Gehad Ismail Sayed
DarkSpot_1993@yahoo.com
http://www.egyptscience.net

[1] Faculty of Computers and Information, Cairo University, Giza, Egypt

[2] Scientific Research Group, Giza, Egypt

Early detection and accurate staging of cancer are considered an important issue in practical radiology. The number of mitotic cells in histology sections is an essential indicator for cancer diagnosis and assessment. Histologists through examining proliferated area usually perform mitosis counting process manually. Pathologists grade breast tissue samples under microscopes based on the cell structures deviation from normal tissues. A pathologist may have to examine and grade a high number of histopathological slides [29]. This process is a tedious and time-consuming. Automatic mitosis counting can reduce the time and the cost involved in this process. Also, it can minimize the errors and can improve the results comparability which obtained from different labs [24]. However, it is a challenging task due to the irregular shape of the mitosis cell, artifacts, and unwanted objects. In fact, different types of tissue, which exhibit highly variable appearance [18], mainly characterize different image areas. In addition to, in the most stages, mitosis cell looks very much like a non-mitosis one or like other dark-blue spots and ordinary cells, so that a human observer without extensive training cannot distinguish between them.

Several approaches for nuclei detection in H & E images were proposed in the literature. Khan et al. [24] proposed an approach for automatic mitosis detection of histopathology slide images. His approach depends on using statistical gamma-Gaussian mixture model (GGMM) to estimate the probability density function (pdf) of mitosis and non-mitosis cells. Support Vector Machine (SVM) classifier was adopted for mitosis detection based on the extracted statistical features from all candidates.

Sommer et al. [32] proposed another approach. His approach consists of two level classifications to detect mitosis. In level one, he used random forest classification to extract candidates. In level two, SVM was adopted to discriminate mitosis from non-mitosis cells. H. Irshad in [14] used blue ratio mapping to map the original histopathology image from RGB to blue ratio, color space to discriminate the nucleus region from the background. In the blue ratio color space, the pixels have high blue intensity relative to red and green channels. The nuclei based on this mapping appear as blue-purple areas, and thus it can be extracted by using only a simple threshold method and morphological operations. For candidates classification, features based on gray level co-occurrence matrix (GLCM), run-length matrices and scale invariant feature transform (SIFT) were extracted from each candidate. These features were used to feed into a decision tree (DT), linear SVM and non-linear SVM classifiers.

V. Roullier in [28] used graph-based multi-resolution approach for mitosis extraction in the breast cancer histological image. It is a top-down approach performs an unsupervised clustering by using domain specific knowledge at each resolution level. Another approach for automatic detection of mitosis cells of histopathology image was proposed at [18]. In this approach, statistical gamma-Gaussian mixture model (GGMM) was adopted to estimate the probability density function (pdf) of mitosis and non-mitosis cells. For mitosis detection, statistical based features and Gabor features were extracted from each candidate, and then these features were utilized to feed SVM. Other approaches were presented for mitosis detection purpose based on the exclusive independent component analysis (EICA) [11] and artificial neural networks [4]. Taskh [37] proposed an automatic mitosis detection system (AMDS) for breast cancer histopathological slide images based on using maximum likelihood estimation for candidate's segmentation and nonlinear SVM for candidate's classification. Local binary patterns based features were extracted from each candidate. His system obtains overall f-score 70.94 % for Aperio XT scanner images and 70.11 % for Hamamatsu images. Similar approaches were proposed in [35, 36].

In general, there are common challenges for automatic mitosis cell detection of histopathology slide images. The first challenge is that the mitosis cells have a very similar characteristic with non-cancerous cells and lymphocytes. Moreover, mitosis nuclei have significant variations in size, intensity pixel value and shape. The second challenge is the number of extracted mitoses is very high.

In this paper, an automatic mitosis detection approach was therefore proposed to overcome these challenges. Initially, Gaussian filter was applied to the original histopathology slide image to remove noise and enhance the image. Then, each pixel of the enhanced image was converted to the NS domain. Then, the truth subset image from each channel was enhanced by using morphology operations. Several features based on statistical, color, texture, shape and energy features were extracted from each connected component (candidate mitosis). Finally, the best discriminative features were selected based on using principles of MFO. These selected features were used to feed the classification and regression tree (CART) to classify each candidate to mitosis or non-mitosis.

The following is the organization of this paper. Section 2 explains the relevant basic concepts of NS and MFO. Section 3 shows the proposed mitosis detection approach. Section 4 explores and examines the experimental results and analysis with details of the data sets. Finally, conclusions are discussed in Section 5.

2 Preliminaries

2.1 Neutrosophic sets

Neutrosophic set (NS) theory was used to solve many problems related to indeterminacy. It is a generalization of an

intuitionistic set, fuzzy set, paraconsistent set, dialetheist set, paradoxist set, and a tautological set. The main difference between NS and fuzzy logic is that NS introduces a new membership called "indeterminate." This new membership component carries more information than fuzzy logic. NS and its properties were discussed briefly in [2]. In this paper, NS was used for mitosis candidates segmentation process.

In neutrosophy, each pixel of the image is transferred to the neutrosophic domain. The pixel in the neutrosophic domain represented as T, I, and F which mean the pixel is t % true, i % indeterminate, and f % false, where t varies in T, i varies in I, and f varies in F, respectively. In NS, $0 \leq t, i, f \leq 100$. However, in a classical set, $i = 0$, t and f are either 0 or 100 and in a fuzzy set, $i = 0$, $0 \leq t, f \leq 100$.

Definition (Neutrosophic image) Let U be a universe of discourse, and W be a set included in U, which is composed by bright pixels. The image P_{NS} in neutrosophic domain is characterized by three subsets T, I and F. Each pixel $P(i, j)$ in the image domain is transformed into neutrosophic domain $P_{NS}(i, j) = \{T(i, j), I(i, j), F(i, j)\}$, where $T(i, j)$, $I(i, j)$ and $F(i, j)$ are the probabilities belonging to white set, indeterminate set and non-white set, respectively [1]. They are defined as follows:

$$P_{NS}(i, j) = \{T(i, j), I(i, j), F(i, j)\} \tag{1}$$

$$\overline{g(i, j)} = \frac{1}{w \times w} \sum_{m=i-w/2}^{i+w/2} \sum_{n=j-w/2}^{j+w/2} g(m, n) \tag{2}$$

$$T(i, j) = \frac{\overline{g(i, j)} - \bar{g}_{min}}{\bar{g}_{max} - \bar{g}_{min}} g(m, n) \tag{3}$$

$$I(i, j) = 1 - \frac{Ho(i, j) - \bar{Ho}_{min}}{\bar{Ho}_{max} - \bar{Ho}_{min}} \tag{4}$$

$$F(i, j) = 1 - T(i, j) \tag{5}$$

$$Ho(i, j) = abs(g(i, j) - \overline{g(i, j)}) \tag{6}$$

where $\overline{g(i, j)}$ is the local mean value of the image, $\bar{g}_{max}$ is the first peak greater than mean of local maxima of histogram $\overline{g(i, j)}$, $\bar{g}_{min}$ is the last peak greater than local maxima of histogram $\overline{g(i, j)}$ and $Ho(i, j)$ is the homogeneity value of T at (i, j) , which is described by the absolute value of difference between intensity $g(i, j)$ and its local mean value $\overline{g(i, j)}$.

2.2 Moth flame optimization

S. Mirjalili developed the moth flame optimization (MFO) in 2015 [22]. Moths are fancy insects that are very similar to the butterfly family. In nature, there are greater than 160,000 various species of this insect. Larvae and adult are the two main milestones in their lifetime. The larva is converted to moth by cocoons. Special navigation methods in the night are the most interesting fact about moths. They used a mechanism called transverse orientation for their navigation. Moths fly using a fixed angle with respect to the moon, which is a very effective mechanism for long traveling distances in a straight line.

Let the candidate's solutions be moths and the problem's variables are the position of moths in the space. The P function is the main function where moths move around the search space. Each moth updates his position with respect to flame using the following equations:

$$M_i = P(M_i, F_j) \tag{7}$$

where M_i indicates the $i - th$ moth and F_j is $j_t h$ flame.

$$P(M_i, F_j) = D_i.e^{bt}.cos(2\pi t) + F_j \tag{8}$$

where D_i is the distance of the $i - th$ moth for the $j - th$ flame, t is is a random number in $[-1, 1]$ and b is a constant for defining the shape of the P. D is calculated using the following equation.

$$D_i = |F_j - M_i| \tag{9}$$

where M_i is the $i - th$ moth, F_j indicates the $j - th$ flame, and D_i indicates the distance of the $i - th$ moth to the $j - th$ flame.

Another concern, the moths update their position with respect to n different locations in the search space which can degrade the best promising solutions exploitation. Therefore, the number of flames adaptively decreases over the course of iterations using the following formula:

$$flame_{no} = round\left(FN - I * \frac{N - 1}{IN}\right) \tag{10}$$

where I is the current number of iterations, IN is the maximum number of iterations and FN is the maximum number of flames.

Global search ability and local search ability which knew as exploration and exploitation respectively, are the two conflicting milestones which can promote one result in degrading the other. As large search, space (exploration) can highly prevent the algorithm from finding an accurate approximation of the optimal solution (global optimum). On the other hand, large exploitation can be resulted in local optima stagnation and decrease the approximated optimum. However, the right balancing between them can guarantee a very accurate approximation of the global optimum. This can be avoided using population-based algorithms. A swarm-based algorithm such as moth-flame optimization is one of a population-based algorithm, which recently developed. Population-based algorithms balance between exploitation

and exploration through initially finding a rough approximation of the global optimum, and then through the algorithm running improve its accuracy. MFO uses the number of flame to balance between exploration and exploitation through adaptively decreasing the number of flames through the algorithm iterations.

The time complexity of MFO is mainly depending on the number of decision variables (dimensions), the number of the iteration, the number of moths (search agents) and sorting mechanism of the algorithm. Since MFO uses quick sort and the sort takes $O(nlogn)$ for the best case and $O(n^2)$ for the worst case, the time complexity is calculated as follows:

$$O(MFO) = O(MaxIter(nlogn) + O(position\ update)) \quad (11)$$

$$O(MFO) = O(MaxIter(n^2 + n * Dim)) \quad (12)$$

where n is the number of search agents, $MaxIter$ is the maximum number of iterations and Dim is defined as the number of decision variables

3 The proposed mitosis detection and classification approach

The proposed mitosis detection and classification approach consists of two main phases, namely candidate's extraction and candidate's classification. These phases are outlined in details in the following sections along with the steps involved with the characteristic features for each phase. The overall architecture of the proposed approach is described in Fig. 1.

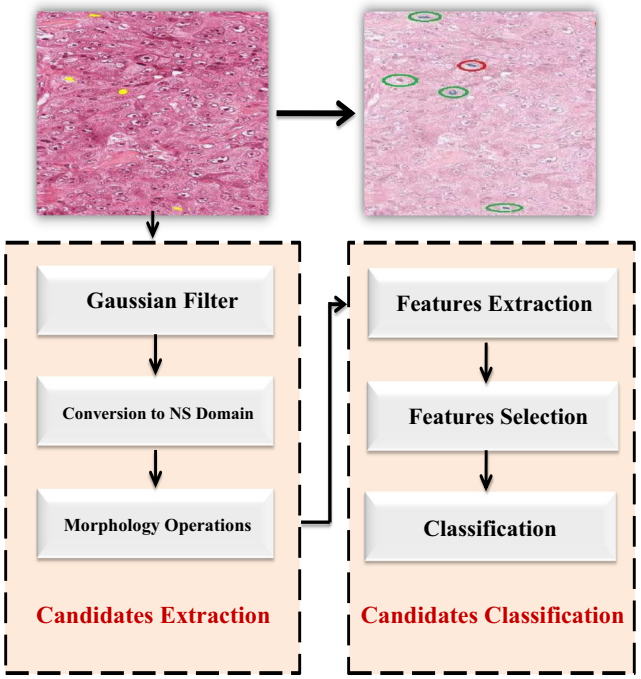

Fig. 1 The proposed mitosis detection and classification approach architecture

3.1 Candidate's extraction

3.1.1 Preprocessing

Noise is always undesirable, so removing this noise with preserving edges plays a vital role in image processing. In this paper, Gaussian filter was adopted to remove this noise [19]. Gaussian filter is a 2D convolution operator, which similar to the mean filter, but it uses a different kernel that represents the bell shape. It is defined in (13).

$$I(i, j) = \frac{1}{\sqrt{2\Pi\sigma^2}} exp\left(\frac{(i - \frac{w}{2})^2 + (j - \frac{w}{2})^2}{2\sigma^2}\right) \quad (13)$$

where w is the size of the window and σ standard deviation of the filter.

3.1.2 Conversion to neutrosophic domain

Conversion to neutrosophic domain In this paper, NS approach was used to extract mitosis candidates. Each pixel of histopathology image A was converted to NS domain $A\{(T^R, T^G, T^B), (F^R, F^G, F^B), (I^R, I^G, I^B)\}$ where T represents the Truth subset image, F represents False subset image and I represents Indeterminate subset image [16]. After converting the enhanced histopathology to NS domain, the truth subset image, T, from each channel was converted, by using an adaptive threshold method based on the local maximum, to a binary image. The proposed approach is described in Algorithm (1).

Algorithm 1 Neutrosophic approach

for $g := 1$ *to* 3 **do** *For each channel in image A*
  Convert each pixel of image to NS. ;
  Calculate local maximum of histogram $\overline{g(i,j)}$ and find the first peak $\bar{g}_{min}$ and the last peak $\bar{g}_{max}$ greater than local mean;
  Calculate Truth subset T(i,j) from $\bar{g}_{max}$, $\bar{g}_{min}$ and;
  Calculate False subset from T(i,j);
  Calculate Indeterminate subset from the absolute difference between $\overline{g(i,j)}$ and the original intensity image g(i,j);
  Enhance Truth subset of image T by applying an adaptive threshold method based on local mean intensity value to get T';
end
Get final segmented images (x, y) by using following equation:

$$s(x, y) = \begin{cases} True & \text{if } (T'_R = 1 \bigcap T'_B = 1 \bigcap T'_G = 1) \\ False & \text{otherwise} \end{cases} \quad (14)$$

3.1.3 Post-processing phase: morphological operators

In this phase, the segmented image $s(x, y)$ was enhanced by using morphology operations. The adopted operations are open and close morphology operations. The open

morphological operation consists of erosion followed by dilation. It used to eliminate small regions. Unlike open, close consists of a dilation followed by erosion. It used to fill holes and small gaps of the objects [30]. Then, connected component based approach with the number of connectivity equals to eight was applied to detect each candidate (8-connected component).

3.2 Candidates classification

3.2.1 Feature extraction

Several features were extracted from each candidate to discriminate mitosis from non-mitosis. The extracted features are based on 15 shape features (centroid, area, and perimeter) [33]. A 22 hair like texture features extracted from gray level co-occurrence matrix (GLCM) [10], 11 features obtained from Gray level run length matrix (CLRLM) [34], and 9 statistical features comprising of first, second, third and fourth moments, mean, standard deviation, skewness, kurtosis, and median. In addition to, 32 energy features derived from Gabor filtering in 8 directions and 4 frequencies [27] and 20 features obtained from Local Binary Pattern (LBP) with radius R = 2 [9]. Total of 115 features were extracted from each channel for each candidate, which resulted in the total number of extracted features 345.

3.2.2 Feature selection based on MFO

Conceptually, a large number of features are highly desirable to classify nuclei to mitosis or non-mitosis. However, some features are redundant, which fails to provide any additional class discriminatory information or irrelevant for classification. All these features can degrade the classification performance. The principle of moth-flame optimization was used for selecting the optimal subset of features.

In the beginning, many feature subsets were initialized randomly within a range from one to 345 'same as the total number of features', and then each subset was assigned to a moth position (a solution in the search space). Each subset has different size with various features combination. Each moth position was evaluated using a predefined fitness function. In our case, f-score obtained on average from 7-folds using K-nearest neighbor classifier (KNN) was used as the fitness function $f(\vec{M}_i)$. The fitness function was used to evaluate how the goodness of the moth position. MFO iterates to discover through all moths' positions (solutions) until reaching the optimal solution, which gives the highest f-score results. The overall proposed MFO feature selection algorithm is described in Algorithm 2 where $\vec{M}_i$ defined as moths positions at iteration i. Also, Fig. 2 shows the proposed MFO feature selection approach. The initial parameters setting of MFO is presented in Table 1.

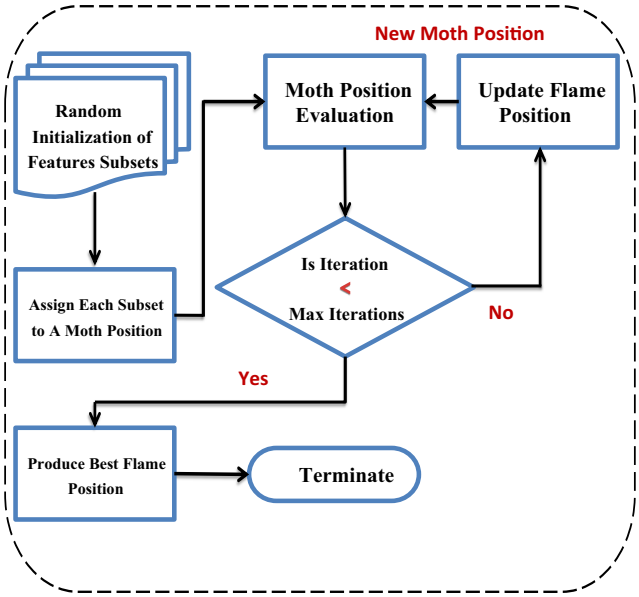

Fig. 2 MFO feature selection process

The MFO's parameters were used to fit only on our database. As it can be seen from Table 1, we set the number of dimensions to 345, which is similar to the total number of extracted features. This is due to the number of dimensions is mainly depending on the nature of the problem. The rest of parameters including the number of iterations and the number of search agents can be utilized to solve several benchmark problems. We set the total number of iterations and a total number of search agents to thirty, which is almost high relative to our number of dimensions. However, the used dataset is medical data; thus, the accuracy is highly required more than time. Besides, MFO was executed only once at training phase. In our paper, we compare the computational classification time before selecting salient features and using all features. The experimental results show that selecting salient features to reduce the time cost to almost half. Through setting the maximum number of iterations and population size, we can almost guarantee the algorithm nearly reach the global optima. All these parameters were set based on trial and error method, where we start with a few relatively agents and then increase the number. Through

Table 1 Parameters setting of MFO algorithm

Parameter	value(s)
Number of Search Agents	30
Number of features	345
Number of Iterations	30
Dimension	345
a	−1
b	1

this way, we set the ideal number of agents and the number of iterations to 30.

Algorithm 2 Moth flame optimization feature selection algorithm

Initialize the number of moth positions n and the maximum number of iterations Max_{itr};
for $(i = 1 : n)$ **do**
  Assign randomly a different feature combination subset to each moth position $\mathbf{M_i(t)}$;
  Evaluate the fitness function of each moth position $f(\mathbf{M_i})$;
end
for $(t < Max_{itr})$ **do**
  Rank the moths positions by sorting their fitness values and assign the values of the first value (highest f-score results) to flame position F using Equ. (10);
  for $(i = 1 : n)$ **do**
    for $(j = 1 : d)$ **do**
      Calculate D using Equ. (9) with respect to moth;
      Update $M(i, j)$ using Equ. (7) and (8) with respect to moth;
    end
  end
  Set $t = t + 1$;
end
Produce the best flame position F;

3.2.3 Classification

In this paper, classification and regression tree (CART) classifier was adopted in order to classify each candidate to mitosis and non-mitosis. CART classifies an instance by sorting them down the tree from root to leaf node, which provides a classification of that instance [20]. Three different splitting criteria were used, including Gini index, deviance and towing. In order to evaluate the performance of the proposed approach, the ground truth mitosis and non-mitosis were labeled automatically by separating mitosis using a CSV file and subtracting it from candidate mitosis to get non-mitosis.

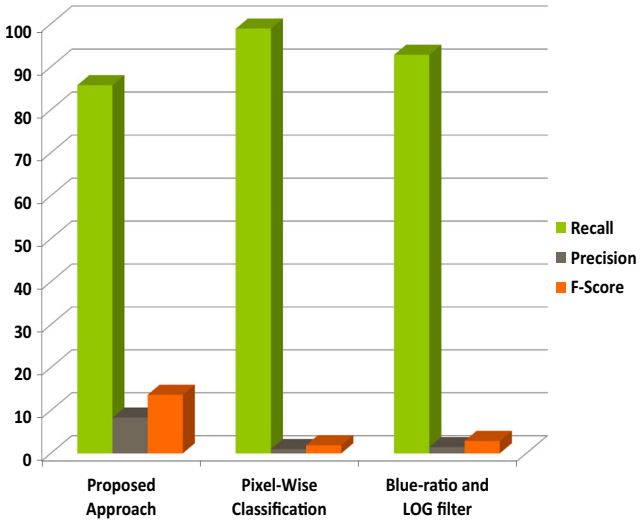

Fig. 4 Candidate's extraction results of the proposed approach with the other approaches

3.3 Description of MITOS dataset

The proposed approach was evaluated on a public data set related to the International Conference on Pattern Recognition (ICPR'12) [5]. It provides five distinct breast pathology slides. These slides were stained with H&E. In this experiment, 50 histopathology slide images, acquired by Aperio XT scanners with 40-x magnification were used. The resolution of histopathology image is 2084x2084 pixels. The total number of mitoses in 50 database images is 300, which were annotated by an expert pathologist. Figure 3 shows samples of mitosis and mitosis cells in the used data sets. The dataset is randomly divided using holdout method to two subsets, where 35 out of 50 histopathology slide image were used for training and feature selection and 15 for evaluating the

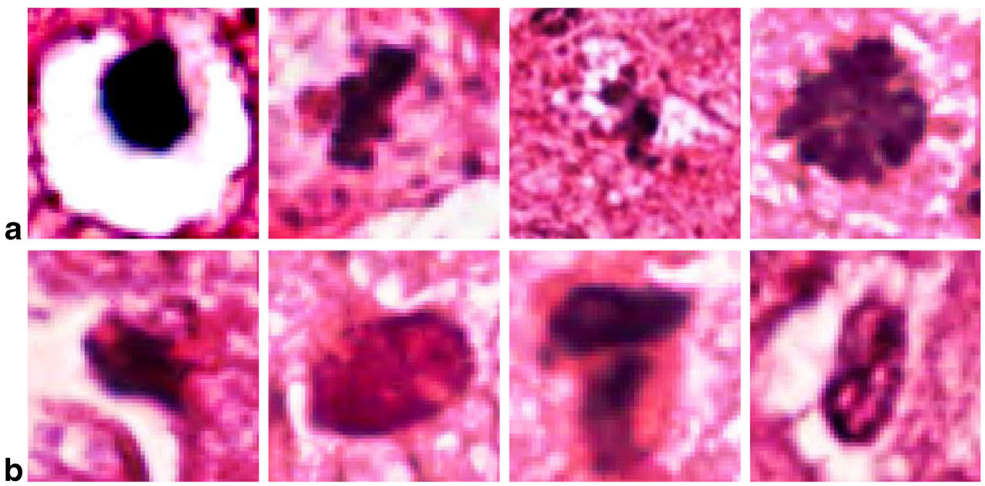

Fig. 3 The MITOS dataset; **a** mitosis **b** non-mitosis

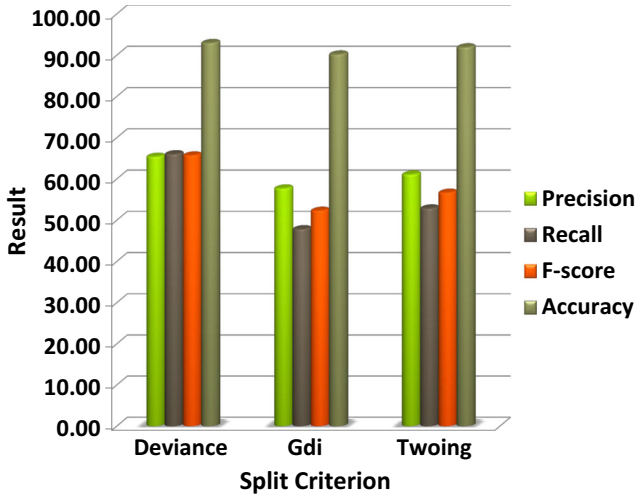

Fig. 5 Candidates classification results using MFO

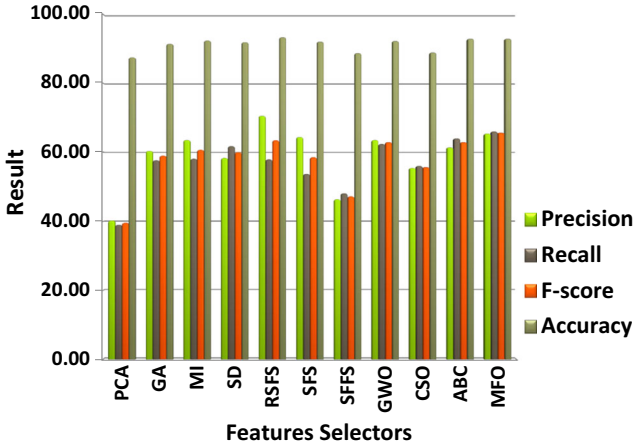

Fig. 7 Comparison between the proposed MFO feature selection approach and the other approaches regarding F-Score, Precision, and Recall

performance of the proposed approach using predefined measurements.

4 Experimental results and discussion

4.1 Evaluation metrics

The precision, recall, accuracy, and f-score were used to evaluate and measure the performance of the proposed approach. Equations from (15) to (18) define the precision, recall, accuracy, and f-score simultaneously. True Positive (TP) represents the number of truly detected mitoses. False Positive (FP) represents the number of non-mitosis that is misclassified as mitosis. There is a contradictory relation between FP and precision, where the less of FP, the higher of precision. Recall shows to what extent the algorithm success in detecting correct mitosis. False Negative (FN) states the number of last misclassified mitosis objects [31]. F-score represents the harmonic mean of precision and recalls that measure the weighted average between precision and recall. The best value of F-score is one, and the worst is zero.

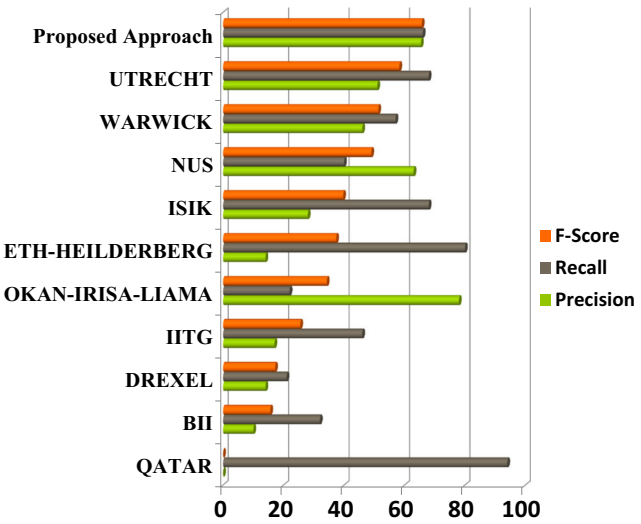

Fig. 6 Candidates classification results of the proposed approach with the other approaches

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

$$F - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{17}$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{18}$$

In addition, five measurements were used to evaluate stability, classification performance, average selected features (ASS) of GWO, MFO, ABC, and CSO feature selection algorithms. These measurements are mathematically defined as follows.

$$Std = \sqrt{\frac{\sum_{i=1}^{MaxIter} (Q_i - \mu)^2}{MaxIter}} \tag{19}$$

$$Best\ Fitness = max_{iter=1}^{MaxIter} Q_i \tag{20}$$

$$Worst\ Fitness = min_{iter=1}^{MaxIter} Q_i \tag{21}$$

Table 2 Comparison between using all original features and using only selected feature by MFO in terms of Precision, Recall and F-score

	Precision(%)	Recall(%)	F-Score(%)
All Features	53.63	57.34	55.42
MFO	65.42	66.03	65.73

Table 3 The processing time after using the proposed MFO feature selection approach and without using it

Parameter	Process Time in Seconds
Before MFO	12.87
After MFO	6.49

$$Mean\ Fitness = \frac{1}{MaxIter} \sum_{iter=1}^{MaxIter} Q_i \tag{22}$$

$$ASS = \frac{1}{MaxIter} \sum_{iter=1}^{MaxIter} \frac{length(Q_i)}{L} \tag{23}$$

where $MaxIter$ is the number of iterations and Q is the best score obtained so far at each time.

4.2 Results and analysis

In this subsection, the obtained result from each phase of the proposed approach including candidates' extraction and candidates' classification were evaluated. Figure 4 compares the obtained results of candidate's extraction phase with the obtained results from two other approaches [32] and [14] regarding f-Score, precision, and recall. As it can be observed from this figure, the proposed approach obtains better results compared with the others. Figure 5 compares the obtained results of candidates classification phase using different splitting criteria: GDI, Towing, and Deviance. As it can be seen, Deviance is the best splitting criteria as it obtains 65.42 % precision, 65.73 % recall, 65.73 % f-score and 92.99 % accuracy.

Figure 6 compares the obtained results of the proposed candidate's classification approach with the other competing approaches [29] regarding F-Score, Precision, and Recall. As it can be seen from this figure, the proposed approach obtains better results compared with the others.

Figure 7 compares the obtained candidate classification results of the proposed MFO approach using CART with deviance splitting criteria with other feature selection approaches regarding F-Score, Precision, and Recall. These approaches are Mutual Information (MI), Statistical Dependency (SD), Random Subset Feature Selection (RSFS), Principle Component Analysis (PCA), Sequential Floating Forward Selection (SFFS), Sequential Forward Selection (SFS), Genetic Algorithm (GA), Grey Wolf Optimizer (GWO) [23], Ant Bee Colony (ABC) [15] and Chicken Swarm Optimization (CSO) [15]. It can be observed that the proposed approach overtakes the other feature selection approaches.

Table 2 compares the mitosis detection results before using MFO feature selection (all features considered) and after using it. As it can be seen, using MFO feature selection approach enhances the obtained results with almost 12 % in terms of F-Score, Precision, and Recall. Table 3 compares the processing time, which the classification algorithm takes before using MFO feature selection approach and after using it. As it can be seen, after selecting the small subset of features the process time decreased, which is logic as feature dimensionality decreases. Figure 8 shows the visualization of the built tree using deviance splitting criteria where X; ($X_1 : X_{25}$, $X_{116} : X_{140}$ and $X_{231} : X_{255}$) are statistical features, ($X_{26} : X_{49}$, $X_{141} : X_{164}$ and $X_{256} : X_{279}$) are texture features, ($X_{50} : X_{61}$, $X_{165} : X_{176}$ and $X_{280} : X_{291}$) are GLRLM features, ($X_{62} : X_{94}$, $X_{177} : X_{209}$ and $X_{292} : X_{324}$) are Gabor features and ($X_{95} : X_{115}$, $X_{210} : X_{230}$ and $X_{325} : X_{345}$) are LBP features. As it can be seen that texture features give high priority than the other features.

Table 4 compares the obtained results of the proposed mitosis detection approach with other published approaches.

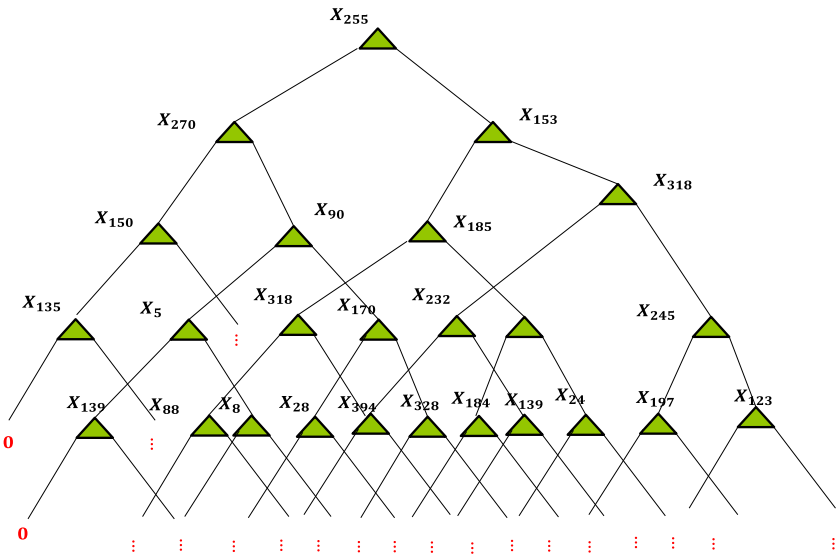

Fig. 8 Tree visualization of the used features

Table 4 The proposed approach vs. other approaches

	Approach	Data Size	Precision(%)	Recall(%)	F-Score(%)
[38]	Generic features and an ensemble of cascade adaboosts	35	54	62.7	58
[17]	Gamma-gaussian mixture model	35	70	72	70.98
[12]	Morphological and multichannel statistics features	50	56	71	63
[14]	Texture, SIFT features and HMAX biologically inspired approach	35	75	76	76
[25]	Teaching-Learning-Based Optimization	35	81	74.2	77.34
[13]	Texture, Morphology based Mitosis detection and Multispectral Intensity	50	63	66	65
The proposed approach	Neutrosphic and Moth Flame Optimization	50	65.42	66.03	66

Table 5 System detailed settings

Name	Detailed Settings
Hardware	
CPU	Core(TM) i3
Frequency	2.13 GHZ
RAM	2 GB
Software	
Operating System	Windows 7
Language	MATLAB R2012R

It can be seen that the proposed approach provides better results compared with those used 50 images. However, the other approaches, which use only 35 images, obtain the highest results. Therefore, we believe these provided results will be decreased as soon as the data size increases.

All the experimental results were performed on the same PC with the same configuration as shown in Table 5 to get an unbiased comparison of all swarms algorithms regarding CPU time. Besides, several statistical measurements were used to evaluate the performance of four swarm algorithms regarding best, worst, mean fitness, standard deviation and average selection size of selected features. Table 6 shows the obtained statistical results of MFO, GWO, ABC and CSO. It can be observed that MFO outperforms the other swarm algorithms. Also, it can be observed as the number of search agents m increases the computational time increases too, and the stability may be decreased. Additionally, we found the number of search equals to 30 is the best one,

Table 6 Statistical results of MFO, GWO, ABC and CSO

	m	Worst Fn.	Best Fn.	Mean Fn.	Std.	ASS	Time in Seconds
	10	0.55	0.56	0.56	0	0.55	33.14
	20	0.57	0.57	0.57	0	0.54	67.06
	30	0.6	0.61	0.6	0	0.53	122.86
	40	0.58	0.59	0.59	0.001	0.53	160.87
MFO	50	0.59	0.59	0.59	0	0.5	200.99
	10	0.48	0.49	0.49	0.007	0.64	327.88
	20	0.50	0.53	0.52	0.005	0.61	512.32
	30	0.49	0.53	0.52	0.007	0.62	646.32
	40	0.51	0.52	0.52	0.006	0.64	730.78
GWO	50	0.52	0.56	0.55	0.005	0.62	973.49
	10	0.5	0.51	0.5	0.007	0.65	114.95
	20	0.5	0.51	0.5	0.012	0.63	263.49
	30	0.50	0.55	0.53	0.02	0.56	364.74
	40	0.5	0.53	0.51	0.013	0.59	526.05
CSO	50	0.52	0.53	0.52	0.006	0.62	675.56
	10	0.54	0.54	0.54	0	0.55	797.97
	20	0.49	0.52	0.5	0.02	0.64	838.07
	30	0.50	0.53	0.52	0.02	0.63	962.62
	40	0.48	0.52	0.5	0.03	0.6	1026.35
ABC	0.51	0.51	0.51	0.5	0	0.64	1189.32

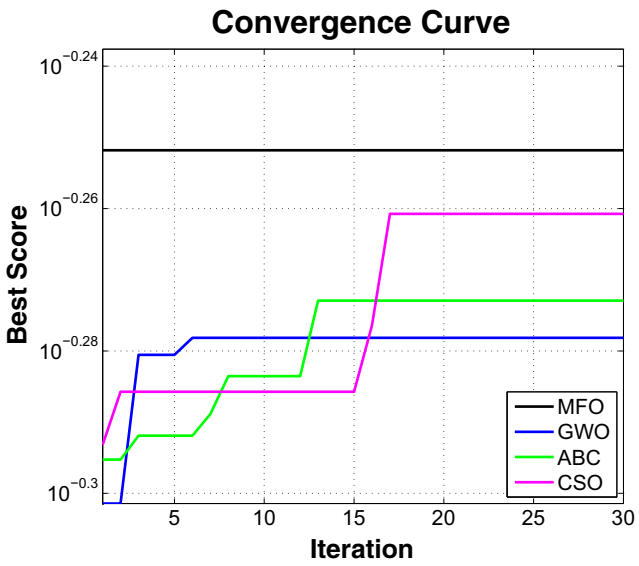

Fig. 9 Convergence curve of MFO, GWO, ABC and CSO

which obtains the highest f-score. These obtained results can prove that MFO is superior in selecting the minimum number of features with good classification performance. Figure 9 shows how the fitness value (obtained f-score) changes trough 30 iterations for MFO, GWO, ABC, and CSO. As it can be seen, as the number of iterations increases the fitness value increases too. Also, MFO converges faster the others which consistent with the obtained results in Table 6.

Figure 10 compares the obtained mitosis candidates' detection results in term of f-score, precision, accuracy and recall using different kernel functions: Linear, RBF, Quadratic, and Polynomial. As it can be observed in the obtained results, Linear is the best kernel function. It obtains overall 30.77 % precision, 39.51 % Recall, and 34.59 % F-Score. However, these results were very poor compared with CART. Also, it takes a long time compared with CART as it can be seen in Table 7. SVM takes one hour and 50 minutes, while CART takes a few seconds. As training data using SVM has complexity $O(n^2)$ and $O(n^3$ [21]. Also, it has been proven as the number of dataset samples increases, the performance of normal SVM decreases [3] as it can be seen in Fig. 11. The reason behind this, few numbers (mitosis) compared with the high number of non-mitosis. This kind of problem referred to as imbalanced dataset. Decision trees such as CART often perform well on the imbalanced datasets [39] due to the splitting rules used in the tree creation which can force the classes to be addressed. In this figure, various data set sizes were utilized to evaluate the performance of SVM, 500, 1000, 1500, 2000 and 2500 samples of mitosis and non-mitosis. The training and testing data were randomly chosen.

Table 7 The processing time using SVM (RBF Kernel Function) and CART (Deviance Splitting Criteria)

Parameter	Process Time in Seconds
SVM	5400
CART	6.49

Finally, Fig. 12 shows the obtained results produced from the proposed candidate detection approach with ground truth ones. The green circle indicates correct mitosis; the red circle indicates misclassified mitosis, and the black circle indicates undetected mitosis.

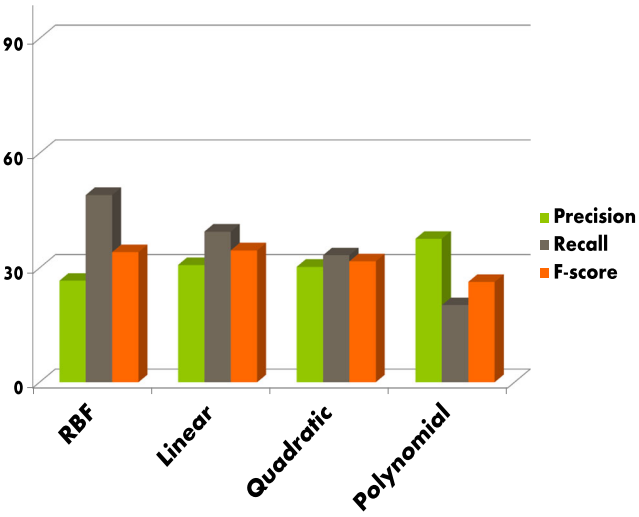

Fig. 10 Comparison between different kernel functions of SVM in terms of F-Score, Precision, Accuracy and Recall

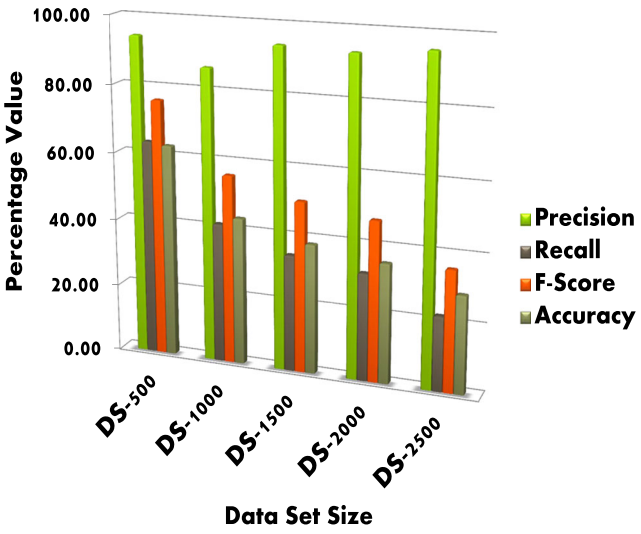

Fig. 11 Comparison between different sizes of the datasets regarding F-Score, Precision, Accuracy, and Recall

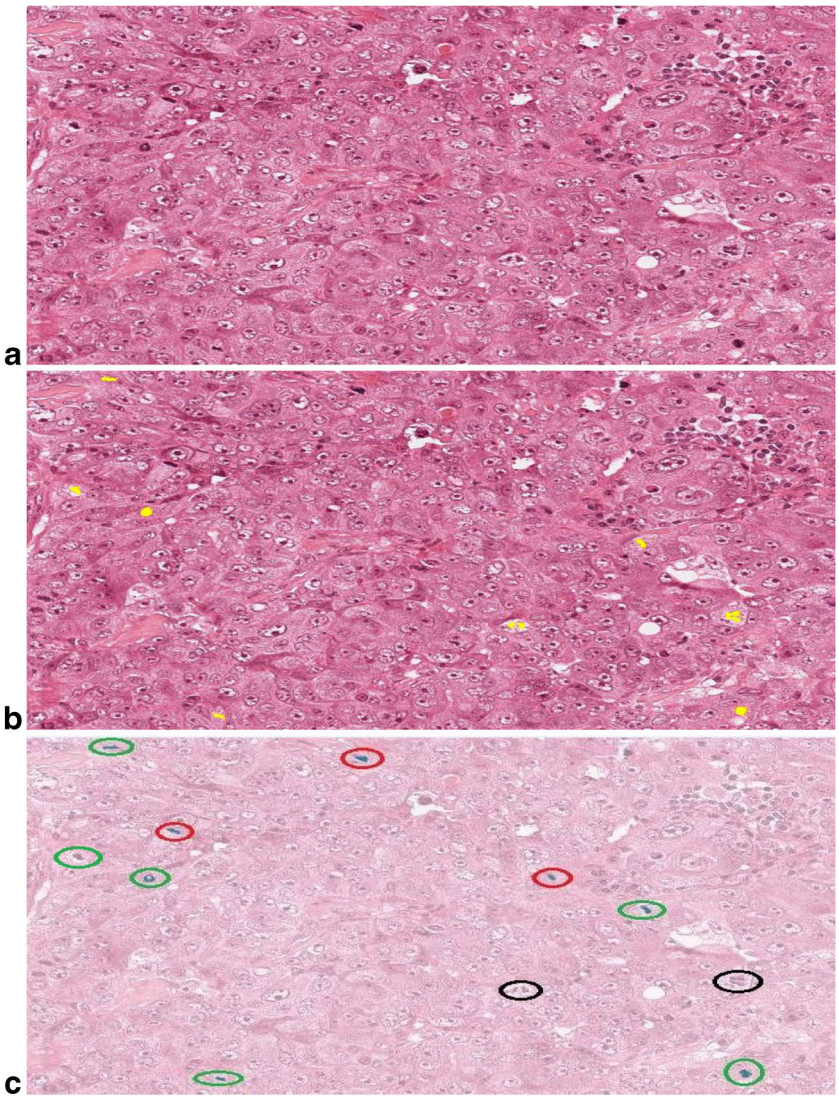

Fig. 12 Proposed mitosis detection approach results; **a** Original histopathology slide image, **b** Ground truth mitosis cell and **c** Automatic mitosis detection approach result

5 Conclusions

In this paper, an approach for mitosis detection and classification in histopathology slide images was presented. The proposed approach is based on NS and the principle of MFO based feature selection. A benchmark dataset consists of 50 histopathology slide image was used. The proposed MFO based feature selection was used to select the best discriminating feature subset which maximizing f-score. The experimental results show the efficiency and robustness of MFO based feature selection algorithm compared to other well-known and three other meta-heuristic algorithms, namely CSO, GWO and ABC feature selection algorithms. Two types of evaluation criteria were used to evaluate the proposed mitosis detection and classification approach. The first evaluation was used for evaluating the performance of the proposed MFO feature selection algorithm. Seven different measurements were adopted for this evaluation. These measurements were five statistical measurements, computational time, and convergence speed. The statistical measurements were the worst fitness value, the best fitness value, mean fitness value, standard deviation, and average features length. The second kind of evaluation was used to evaluate overall the proposed mitosis detection and classification approach regarding classification measures, namely accuracy, recall, precession, and f-score. The experimental results demonstrate that the proposed mitosis detection and classification approach obtains better results compared with the other published approaches. It obtains overall 65.42 % f-score, 66.03 % recall, 65.73 % precision, and accuracy 92.99 %. These high results could be used for further early diagnostic suspicion of breast cancer.

References

1. Şahin R (2014) Neutrosophic hierarchical clustering algoritms. Neutrosophic Sets and Systems. 2:18–25
2. Anter A, Hassenian A, AbuElSoud M (2014) Neutrosophic sets and fuzzy c-means clustering for improving ct liver image segmentation. In: 5th International Conference on Innovations in Bio-Inspired Computing and Applications, IBICA 2014. Springer, pp 193–203
3. Chang F, Yang C, Rong X, Lu C (2010) Tree decomposition for large-scale svm problems. J Mach Learn Res 11:2935–2972
4. Cires D, Giusti A, Gambardella L, Schmidhuber J (2013) Mitosis Detection in breast cancer histology images with deep neural networks. In: 16th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer, Nagoya, pp 411–418
5. Contest I (2012) Mitosis detection in breast cancer histological images. http://ipal.cnrs.fr/ICPR2012/
6. Damjanov I, Fan F (2007) Cancer grading manual
7. Elston C, Ellis I (1991) Pathological prognostic factors in breast cancer. i. the value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology. 19(5):403–410
8. Guican M, Boucheron L, Can A, Madabhushi A, Rajpoot N, Yener B (2009) Histopathological image analysis: a review. Biomed Eng EEE 2:147–171
9. Guo Z, Zhang D (2010) A completed modelling of local binary pattern operator for texture classification. IEEE Trans Image Process 19(6):1657–1663
10. Haralick RM, Shanmugam K, Dinstein IH (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3(6):610–621
11. Huang C, Lee H (2012) Automated mitosis detection based on exclusive independent component analysis. In: Proceedings of the IEEE international conference on pattern recognition. IEEE, Tsukuba, pp 1856–1859
12. Irshad H (2013) Automated mitosis detection in histopathology using morphological and multi-channel statistics features. J Pathol Informatics 4(1):4–10
13. Irshad H (2014) Automated mitosis detection in color and multispectral high-content images in histopathology: application to breast cancer grading in digital pathology. In: 12th European Congress of Digital Pathology, Paris, pp 18–22
14. Irshad H, Jalali S, Roux L, Racoceanu D, Hwee LJ, Le Naour G (2013) Automated mitosis detection using texture, sift features and hmax biologically inspired approach. J Pathol Inform 4(2):12–18
15. Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm. J Glob Optim 39(3):459–471
16. Khalid M, Abd Elfattaha M, Hassanien A, Schaefer G (2014) A binarization algorithm for historical arabic manuscript images using a neutrosophic approach. In: 9th International Conference on In Computer Engineering and Systems (ICCES). IEEE, pp 266–270
17. Khan A, ElDaly H, Rajpoot N (2013) A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. J Pathol Inf 4(1):1–11
18. Khan AM, El-Daly H, Rajpoot NM 2012. A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. In: 21st International conference on Pattern Recognition. IEEE, Tsukuba, pp 149–152
19. Kotecha JH (2003) Gaussian particle filtering. IEEE Trans Signal Process 51(10):2592–2601
20. Lawrence R, Wright A (2001) Rule-based classification systems using classification and regression tree (cart) analysis. Photogramm.Eng Remote Sens 67(10):1137–1142
21. Lopez A, Li X, Yu W (2013) Support vector machine classification for large datasets using decision tree and fisher linear discriminant, vol 36. Gener Comput Syst Elsevier, pp 57–65
22. Mirjalili S (2015) Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl-Based Syst Elsevier 89:228–249
23. Mirjalili S, Lewis A (2014) Grey wolf optimizer. Adv Eng Softw 69:46–61
24. Naik S, Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J (2008) Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology. In: Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Paris, pp 284–287
25. Nateghi R, Danyali H, Helfroush M, Tashk A (2014) Intelligent cad system for automatic detection of mitotic cells from breast cancer histology slide images based on teaching-learning-based optimization. Comput Biol J 2014:1–9
26. NCI (2012) National cancer institute. (Last Accessed: 14/02/2016); Available from: http://www.cancer.gov/cancertopics
27. Palm C, Lehmann TM (2002) Classification of color textures by gabor filtering. Machine Graphics and Vision 3(11):195–220
28. Roullier V, Lézoraya O, Elmoataz A (2011) Multi-resolution graph-based analysis of histopathological whole slide images: application to mitotic cell extraction and visualization. Comput Med Imaging Graph Elsevier. 35(7):603–615
29. Roux L, Tutac A, Lomenie N, Balensi D, Racoceanu D, Veillard A, Leow W, Klossa J, Putti T (2009) A cognitive virtual microscopic framework for knowlege-based exploration of large microscopic images in breast cancer histopathology. In: Annual international conference of the IEEE in engineering in medicine and biology society, Minneapolis, pp 3697–3702
30. Sharma R, sharma R (2014) Image segmentation using morphological operation for automatic region growing. International Journal of Innovative Research in Computer and Communication Engineering 2(9):5686–5693
31. Simon D, Member S, Simon D (2010) Analytic confusion matrix bounds for fault detection and isolation using a sum-of-squared-residuals approach. IEEE Trans Reliab 59(2):287–296
32. Sommer C, Fiaschi L, Hamprecht FA, Gerlich DW (2012) Learning-based mitotic cell detection in histopathological images. In: 21st International Conference on Pattern Recognition (ICPR, vol 2012. IEEE, Tsukuba, pp 2306–2309
33. Sutar GT, Shah AV (2014) Number plate recognition using improved segmentation. Int J Innov Res Sci Eng Technol 3(5):12360–12368
34. Tang X (1998) Texture information in run-length matrices. IEEE Trans Image Process 7(11):1057–7149
35. Tashk A, Helfroush M, Danyali H, Akbarzadehjahromi M (2013) An automatic mitosis detection method for breast cancer histopathology slide images based on objective and pixel-wise textural features classification. In: Proceedings of the 5th Conference on Information and Knowledge Technology, Tehran, pp 406–410
36. Tashk A, Helfroush M, Danyali H, Akbarzadehjahromi M (2014a) A cad mitosis detection system from breast cancer histology images based on fused features. In: Proceedings of the 22nd Iranian conference on electrical engineering, Tehran, pp 1925–1927
37. Tashk A, Helfroush MS, Danyali H, Akbarzadeh-Jahromi M (2014b) A novel cad system for mitosis detection using histopathology slide images. Journal of Medical Signals and Systems 4(2):139–149
38. Tek F (2013) Mitosis detection using generic features and an ensemble of cascade adaboosts. J Pathol Inform 4(1):4–12
39. V G (2012) An overview of classification algorithms for imbalanced. Int J Emerg Technol Adv Eng 2(4):42–47