Comput Syst Sci & Eng (2019) 2: 91–100

International Journal of
Computer Systems
Science & Engineering

A new enhanced learning approach to automatic image classification based on Salp Swarm Algorithm

Mohammad Behrouzian Nejad[1] **and Mohammad Ebrahim Shiri**[1,2]*

[1] *Department of computer Engineering, Borujerd Branch, Islamic Azad University, Borujerd, Iran*
[2] *Department of Mathematics and Computer Science, Amirkabir University, Tehran, Iran*
E-mail: Behrouzian@iaushadegan.ac.ir

In this paper we propose a new image classification technique. According to this note that most research focuses on extraction of features in the frequency domain, location, and reduction of feature dimensions, in this research we focused on learning step in image classification. The main aim is to use the heuristic methods to increase the function of the estimator of the learning algorithm and continue to achieve the desired state, as well as categorization without user interference and automatically performed by the model produced from the above steps. So, in this paper, a new learning approach based on the Salp Swarm Algorithm was proposed that was implemented and evaluated on learning algorithm Decision Tree, K-Nearest Neighbors and Naïve Bayes. The results demonstrate the improvement of the performance of learning algorithms in all the achieved criteria by using the SSA algorithm in comparison with traditional learning algorithms. In the accuracy, sensitivity, classification error and F1 criterion, the best performance of the proposed model is using the Decision Tree learning method with values of 99.17%, 100%, 0.83% and 95.65% respectively. In the specificity and precision criterion, the best performance of the proposed model is based on K-Nearest Neighbors learning method with values of 100%.

Keywords: Image Mining, Image Classification, Learning Algorithm, Salp Swarm Algorithm

1. INTRODUCTION

Classification is a machine learning problem that discusses how tags are assigned to new data based on a set of tagged data. The result is given by an input. In order to predict the result, the algorithm processes a training set containing a set of attributes and the result, usually called the target or the prediction of the attribute. The algorithm assigns the pixel classification in the image to favorite collections or classes. Several methods have been developed by researchers over the years. In order to classify a set of data into different classes or categories, the relationship between the data and the classes they are classified should be well understood [1]. In general, machine learning methods can be divided into two groups of supervised learning and unsupervised learning. In unsupervised learning does not use the set of training data with predefined categories. In supervised learning or learning from examples, the user is said to play the role of a teacher and must provide examples for each of the classes [2, 3]. Medical imaging is one of the most important modern medical instruments. Various types of imaging techniques used to diagnose different types of diseases. Important medical pictures can play an important role in helping to diagnose. However, in clinical applications, there usually is a lot of time to check the image manually. Therefore, an automatic diagnostic procedure will be helpful for doctors. With regard to the need for medical analysis, the classification of self-colliding images is one of the most important issues [4, 5]. Mostly used MRI images are used to discover the presence of a tumor in the brain. The information received from the image is very important in the treatment of the patient.

*Corresponding Author. E-mail: shiriahmadabadi@gmail.com

Image mining employs tacit knowledge of images. The image mining is an interdisciplinary effort requiring expertise in computer vision, pattern recognition, image processing, image retrieval, machine learning, databases, and artificial intelligence. One of the most important image mining techniques is the categorization of images that assign the existing images to several predefined categories of images that belong to them. To do this, the categories must first be identified and labeled by specialist. Then extract the main features of these images and use the machine learning algorithms to teach the system that the next time the images entered into the system are categorized in the correct group. The main steps in the process of categorizing images are: extracting features, reducing the dimensions of features, categorizing and evaluating [6-9].

So far, numerous researches have been done to categorize MRI images. Most research focuses on extraction of features in the frequency domain, location, and reduction of feature dimensions. Their goal was to produce a categorical model to increase the efficiency of the categorization of relevant images. It should be noted that the proposed methods have limitations in addition to the strengths they have. Considering the importance of the subject in this research, we try to take advantage of the strengths of the prior art approaches and to overcome the limitations and shortcomings of the previous methods proposed, and ultimately, to find the correct accuracy in the proposed method. Considering that in the classification of images the learning step is one of the most important steps and, on the other hand, learning algorithms cannot be accurate and accurate in traditional and manual methods. In this research, the aim is to use the heuristic methods can increase the function of the estimator of the learning algorithm and continue to achieve the desired state, as well as categorization without user interference and automatically performed by the model produced from the above steps. In this research we use from salp swarm algorithm to enhance performance of learning algorithms. Finally, the model is expected to perform well on the test data that has not yet been seen in the production model. In order to implement and apply the proposed model, algorithms and machine learning methods such as KNN, Naïve Bayes and Decision Tree will be used. Given that exist learning methods are traditional and that the results generated by these methods may have a high degree of reliability and reliability, this study focuses on improving the performance of the learning phase, so that an efficient and reliable model has achieved the correct classification of data. The results demonstrate the improvement of the performance of learning algorithms in all the achieved criteria by using the SSA algorithm in comparison with traditional learning algorithms. The overall structure of this paper is organized as follows: related works is described section 2. The proposed method and its evaluation are presented in section 3 and its details are defined. Finally the conclusion is given.

2. RELATED WORKS

Image classification tagging a pixel or a group of pixels is based on its gray value. Classification is one of the most common ways of extracting information. In classification, usually multiple features are used for a set of pixels, for example, many images of a particular object are required [4, 10].The important features of images are extracted from data set and used by machine learning algorithms that the system trains to categorize the images imported into the system in the correct group in the next time. The basic steps of the image classification process are: system training and system testing. The system has initially extracted the basic features from the input data, and after reducing the dimensions of feature set to accelerate the system's performance, in the next step by using a machine learning algorithm, it produces a model to categorize test data.

Most studies have focused on extracting and reducing the characteristics of images in the field of learning. In [11] authors develop Principal Component Analysis in feature selection step and kernel based Support Vector Machine in learning step. Authors in proposed method using PCA to reduced dimension of features to 1024 features. Then used from DWT, PCA, and KSVM with GRB kernel to achieving better performance in image classification.

The authors [12] used unstructured learning based on neural networks to classify MRI images and performed in three steps, the first stage of image processing consisting of histogram alignment, edge detection, noise filtering. In the second stage, the extraction of features was performed using Independent Component Analysis (ICA) and used for self-organizing mapping (SOM) classification. This technique is able to categorize images with accuracy of 98.6.

The authors in [13] described their method as spectral cluster analysis of independent components for the classification of tissue from the brain's MRI. The authors used a spectral-angled spectral-scattering feature clustering method. Proposed method provides the same prioritization for general and local features, thus striving for the efficacy of common methods for extraction of abnormal textures. Initially, multi-spectral MRI is divided into different categories by a spectral distance-based clustering. Then independent component analysis is applied to the cluster data. Further, SVM is used to analyze the texture. The results show that brain tissue classification with this high performance method, especially for small lesions and tumors, provides clinical analysis.

The authors in [14], because of the high dependence of the accuracy of the neural networks on the distribution of classes, have used multi-random data prior to classification in order to obtain an appropriate classification model that has at least the performance criteria optimal. In this work, for the purpose of extracting the property, we used a combination of multiple sampling of smart vectors quotient (LVQ), multilevel perceptron network (MLP), non-conductive network (SOM), and radial base function neural network (RBF) to categorize MRI images, this article shows a significant improvement in the classification stability.

The methodology of the authors in [15] is the classification of MRI to normal or one of seven different diseases. Initially, two levels of discrete two-dimensional wavelet transform (DWT 2D) of the input image are calculated. Returning to its conditional heterogeneous generalization, which is a statistical method, the parameters of this model form the main property vectors. After normalizing these features, LDA and PCA diminish the dimensions and then classify the images using KNN algorithms and support vector algorithms.

In the method [16], for the diagnosis of brain cancer using computer for initial prediction, authors used the features of the

tissue and logic of neuronal post-propagation classification and artificial neural network. This involves extracting the properties of the tissue from the MRI brain sample that was used by using a discrete wavelet transform and morphological characteristics, and for categorizing the bipolar neural network (BPNN). In this method, using image processing techniques, the classification of the surface has been remarkable.

The authors in [17] propose a tumor identification system using wavelet packet aspects and artificial neural networks in MRI images. Characteristics are extracted using a wavelet transform and classified by the neural network of the EBPTA algorithm as natural or abnormal images. By using the wavelet transform, they are better than the other clusters.

The authors proposed in [18] a method called Improved Performance, artificial neural networks for categorizing MRI images of two important parameters in neural networks, which resulted in increased accuracy and convergence rates. The major problem in neural networks are that high-accuracy networks require a high convergence time, and networks that have a low convergence time are also less accurate, hence the importance of concurrency and precision. In the method of authors in the Neural Network and Cohen Network, they have made improvements in the method of teaching these networks using free repetition and ultimately improve convergence speed and accurate results. The results of this method show improved productivity compared to conventional networks. The results show that the speed is approximately 5 to 10 times faster, and the computational complexity of these networks has decreased. Of the eight features based on the first histogram and the gray level of the matrix, the same thing happened in this work.

In [19], the authors proposed a hybrid system for tumor detection in MRI images and their classification using artificial neural networks (ANN) and K-nearest neighbor (KNN). In the proposed method, the authors extracted the features using a discrete wavelet transform (DWT) and then the component analysis method (PCA) was used to select the best features. Selected attributes are given as inputs to the KNN and ANN classifiers. The nearest neighbor and artificial neural networks include two stages, training and experimentation. KNN and ANN categorize MRI images as natural and abnormal images.

3. PROPOSE METHOD

Considering that most researches are focused on the selection and reduction of feature dimensions and, on the other hand, the most important step in generating a proper learning model for the learning phase, as well as the fact that existing learning methods are traditional and may be The results produced by these methods have a high error rate and low reliability. Therefore, this study focuses on improving the performance of the learning step so that an efficient and reliable model can be used to categorize the data correctly. The proposed method categorizes images automatically in related categories. We assume that the heuristic methods associated with machine learning algorithms can increase accuracy. Therefore, in this research, Salp Swarm Algorithm (SSA) methods are used to search the space of the parameters to select the optimal values for the learning method evaluator function.

Given that existing learning methods are traditional and that the results generated by these methods may have a high degree of faults, the focus in this study is on improving the performance of the learning phase, so that to achieve an efficient and reliable model of the correct classification of data, the aspect of innovation and novelty can be expressed as follows:

- Automatic categorization of images into related categories.
- Improved of machine learning algorithms by heuristic method.
- Using heuristic method to search for parameter space to select the optimal values ??for the learning method evaluator function.
- Automate the calculation and entry of optimal parameters related to the learning method.
- Automatic test output of the generated model and, if the results of calculation and retrieval are not appropriate, to select the appropriate values for the parameters.
- Producing an efficient learning model for automated image data collection using machine learning methods based on heuristic methods.

In the proposed categorization of images, the appropriate dataset will be selected. Then we need to specify the category of each image, i.e. the tagging action will be performed. Then, to apply the images to the next steps, we transform the images into a feature vector, and using a feature selection method; non-useful features are identified and try to find a limited number of features that are more important to be selected. This can increase the accuracy of the final classification. After determining the appropriate features, the classification operation will be done. There are different methods to categorizing with different performance. In this research, the aim is to use the SSA method to increase the functionality of the function of the estimator of the learning algorithm and to continue to achieve the desired state, and also to categorize without user intervention and automatically to be performed by the model produced from the above steps. . This is done using a model-tested with different criteria, which ultimately leads to the production of a reliable method. In this research, in the learning phase of the algorithms: Naive bayes, K-NN and decision tree are used which are improved using the SSA algorithm. Details of the proposed method are described below.

In this research we used from 120 T2-type MRI images of 256 $\times$ 256 pixels [20]. All these images are labeled by expert persons. Each image has one state: normal or abnormal. In other hand is labeled a healthy brain or a patient and a type of illness. In this dataset there are 12 normal images and 108 abnormal images that have different brain conditions. To feature extraction, discrete wavelet transform coefficients are used as a feature vector. Wavelet transform is a strong mathematical tool for extraction of features and can be used to extract wavelet coefficients from MRI images. Wavelets are the basic local functions. The main advantage of the wavelet is that provides local abundance information about the performance of a signal, especially for useful classifications. The sample of data shows in Fig. 1.

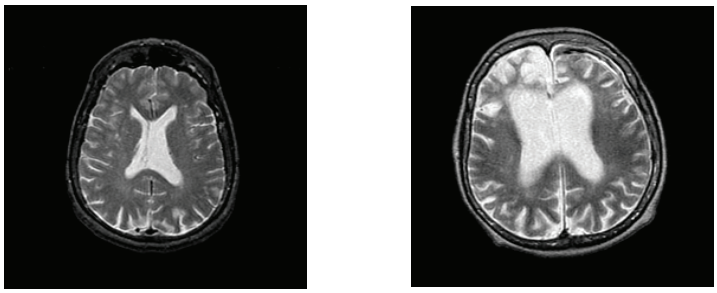


Figure 1 Sample of data.

3.1 Feature Extraction and Reduction

One of the discrete wavelets is the har wavelet. A har wavelet is a period of square wave. Because of its simplicity, most of this wavelet is selected. Using the wavelet, the given function can be analyzed at different levels of resolution. Wavelet transforms the ability to analyze signals in spatial and temporal domains, as well as greater flexibility, better compression and high performance. A discrete wavelet is a linear transformation that transforms a data vector whose size is a power of 2 to another numeric vector of the same size. This tool divides the data with different frequency components. Then each component is studied with the same scale and clarity. A discrete transform can be expressed using Eq. 1.

$$DWT_{x(n)} = \left\{ \begin{array}{l} d_{j,k} = \sum x(n) h_j^*(n - 2jk), \\ d_{j,k} = \sum x(n) g_j^*(n - 2jk), \end{array} \right\} \quad (1)$$

The coefficients $d_{j,k}$ express the components of the signal $x(n)$ and are equal to the wavelet function, while $a_{j,k}$ expresses the approximation coefficients in the signal $x(n)$.The functions $h(n)$ and $g(n)$ in the equation are coefficients of high pass and low pass filters. Parameters of j and k are the wavelet scale and the transmitter. In the previous step, a number of features are extracted from the images that directly affect the classification function. If the dimensions of the data are to be reduced, the final performance of the classification can be improved.

In the used method, a three-level decomposition operation will be performed on each 256 × 256 image in accordance with the application of the random wavelet. In this study, since the dimensions of the image are 256 × 256, in the first step the matrix is ??estimated at 128 × 128, in the second stage 64 × 64, and in the third stage, the estimated matrix is ??32 × 32. As a result, at this stage, we create an estimated linear matrix of 1024 × 1 for an image, which we will select as the image property matrix, and we will ignore the image detail matrix. One of the reasons regardless of the detail matrix, is the difference in the skeletal or structure of MRI images of healthy brain with MRI images of non-essential brain. As a result, we do not need to know the vertical, horizontal and diagonal edges. Fig. 2 shows the stages of the transformation of a wavelet waveform in three levels.

After extracting the features of the images, it consists of learning and testing data of all feature vectors in a matrix in the following dimensions, called the matrix of features.

Matrix of features = [Number of learning & test images * number of features]

Given the number of 120 images and 1024 attributes, the extracted feature matrix will have a matrix of 1024 × 120 Regarding the production of 1024 attributes for classification and attention to the fact that we have a lot of features and certainly this number of features in the final performance will be negatively affected, therefore, we must reduce the dimensions of the features. One of the most prominent feature selection methods is PCA. This tool is used to convert the existing input features into a new space of lower-dimensional features. In the PCA, it uses a correlation matrix to transform the space of the inputs into the space of features with lower dimensions than the largest vectors of the matrix of correlation. Projection techniques to under-space are widely used in diminution of dimensions, which reduces computational cost and computational complexity. According to the dataset, the PCA shows a linear representation of the lower dimension of the data that can maintain variance for reconstructing data. Using a PCA-based attribute-reduction feature, the bounded property vector chosen by this system can lead to an efficient classification algorithm. Therefore, the main idea behind the use of PCA in our approach is to reduce the dimensions of the extracted wavelet coefficients in the feature extraction stage, which results in the accuracy and efficiency of the results in a category. Fig. 3 shows this note.

One of the reasons that we used the principal components analysis method as a method of diminishing the dimension is that this method shows the correlation between the data variables and since MRI images have a high correlation, this method is used to reduce the data space to the feature space Convert. Before we discuss how to use this method, we need to provide a proper definition of the variance non-consistency. Two non-dominant variables are used when their covariance is zero. According to (2):

$$E\{y_1 y_2\} - E\{y_1\}E\{y_2\} = 0 \quad (2)$$

The first basic component is the combination of variables that express the maximum change value. The basic component of the second represents the largest amount of subsequent changes, and the basic component of the first is perpendicular. The basic component can be defined in terms of the variables. In fact, the principal component analysis method finds a super-page in the data space that minimizes the mean square of the image error.

First, the PCA finds the area where the image data sets of the data points on it have the greatest distance. In other words, the variance of these images is as much as possible. The second axis is obtained in the same way, provided it is perpendicular to the first axis. These two axes together put the pages in the best match. This method continues to find all the components of the

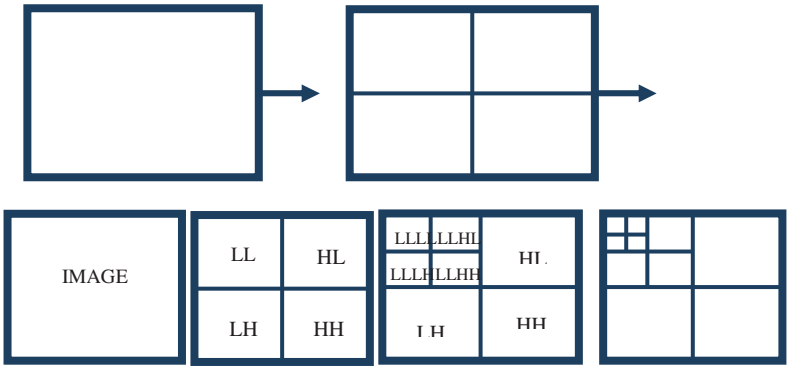


Figure 2 Steps to make the vortex wavelet transform on three levels.

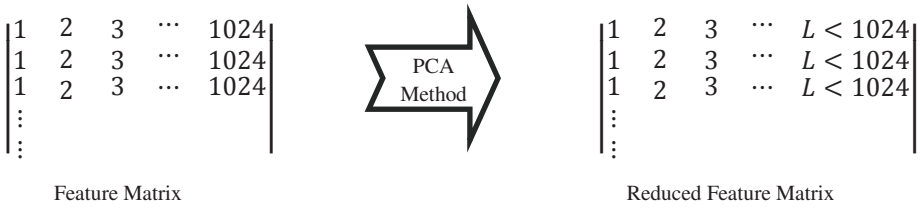


Figure 3 Feature reduction.

main component. The function of reducing the dimensions by using the main components analysis is as follows:

Consider the dataset X with N points, each of which has p dimension:

$$X = \begin{matrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_n \end{matrix} = \begin{matrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ x_{3,1} & x_{3,2} & \cdots & x_{3,p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,p} \end{matrix} \tag{3}$$

A) Calculate the mean of the matrix according to (3) (average of each dimension)

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{4}$$

B) Calculate the deviation from the mean of data so that we reduce each column from the mean of that column. And we save it in a matrix. According to (4)

$$XM = X - \bar{X} \tag{5}$$

C) Covariance of the XM matrix, which calculates a $p \times p$ matrix named C_XM. We store in a matrix. According to (5(

$$C_{XM} = \frac{\sum_{i=1}^{n}((X_i - \bar{X})^T \cdot (X_i - \bar{X}))}{N-1} \tag{6}$$

Where C_XM(i, j) represents the correlation between the two dimensions i and j. According to (6)

$$C_{XM}(i,j) = C_{XM}(j,i) = \frac{\sum_{k=1}^{n}(XM(k,i) \cdot XM(k,j))}{N-1} \tag{7}$$

If C_XM$(i, j) > 0$, then I, j similarly change if C_XM$(i, j) < 0$ then i, j opposite, and if C_XM$(i, j) = 0$ it can be concluded that I, j are independent.

D) Calculate the values and vectors of the covariance matrix C_XM. According to the rules of linear algebra, a symmetric matrix n × n has n special vector and n is a special value.

E) Special vectors and special values are arranged in descending order. Main components of importance are sorted out to low importance. To reduce the dimensions, we can eliminate minor components that are associated with the loss of a small amount of information. According to (7)

$$FeatureVector = (egi_1\ egi_2\ egi_3\ \cdots\ egi_n) \tag{8}$$

In the proposed method, we chose 13 values of special values and special vectors that were more important. The reason we chose this number is that with 13 special values, data projections can represent about 70.52 percent of the data.

3.2 Data Normalization

Data Normalization or normalizing variables is a way to put data that is not in a domain in the same domain. Simply put, a data miner may encounter situations where data properties include values that are within a different range. These large-scale features may have a much greater effect on the cost function

than the features with low values. This problem is resolved by normalizing the features so that their values fall into similar domains.

In this research, the normalization of Z-score was used according to (8) where μ is the mean and σ is the standard deviation of the data.

$$Z = \frac{x - \mu}{\sigma} \tag{9}$$

3.3 Salp Swarm Algorithm Explanation

Salp is a member of the Salp idea family with transparent barrel-shaped body same the jelly fishes. Also, they have moved same the jelly fish; whereas the water has been pumped via body as thrust to move forward [21].The salp shape is given in Fig. 4(a). The biological scholar shaves no exact information about this animal because they are living in far access environments, and they are hardly kept in laboratory environments [22]. Swarming behavior of salps is their prominent interesting feature. Indepth of oceans, salp shave commonly formed the swarm based salp chain. This swarm behavior has been presented in Fig. 4(b). Their living behavior is not exactly reported yet, but some scholars are convinced that this is performed to achieve best locomotion by fast coordinated foraging and changes [23].

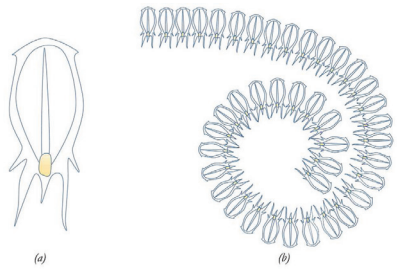


Figure 4 salps swarm behavior.

The position of salps has been updated by the equation (9):

$$x_j^1 = \begin{cases} F_j + C_1((ub_j + lb_j)C_2 + lb_j)C_3 > 0 \\ F_j - C_1((ub_j + lb_j)C_2 + lb_j)C_3 < 0 \end{cases} \tag{10}$$

Where, x_j^1 presents the leader salp position in the j_{th} dimension, F_j is the food source position in the j_{th} dimension, ub_j presents the upper bound of j_{th} dimension, lb_j presents the lower bound of j_{th} dimension, c_1, c_2, and c_3 are random numbers. Eq. 9 indicates that the leader just tune its position according to the food source. The c_1 is the key coefficient of SSA for balancing the exploration and exploitation which is defined as follows (10):

$$C_1 = 2e^{-(4l/L)^2} \tag{11}$$

Where, the current and maximum iterations are defined by and l and L, respectively. The c_2 and c_3 are random numbers uniformly created between [0, 1]. The updating followers 'position is performed by the following equation (11):

$$x_j^i = \frac{1}{2at^2} + v_0 t \tag{12}$$

Since, the optimization duration means the iteration, the difference between iterations is 1, and considering $v_0 = 0$, this equation is presented as follows (12):

$$x_j^i = \frac{1}{2(x_j^i + x_j^{i-1})} \tag{13}$$

Where, $i \geq 2$ and x_j^i presents the position of j_{th} follower salp in j_{th} dimension. Considering Eq. 9) and Eq. 12, the salp chains can be simulated.

3.4 Naïve Bayes

Naïve Bayes (NB) is an algorithm of Bayesian learning. Naive Bayes uses from Bayes Theorem, which aims to learn the conditional relationships between feature values and class tags. When an image is being classified to correct class, the classier learned by Naïve Bayes aims to determine the correct class to which the highest conditional probability has given the feature vector of this image. The feature vectors include a one or many values for the features. In this context, the class, which has the highest conditional probability, would be assigned to the image. The Naïve Bayes method can show how each feature relates to a category in the context of probability theory [24, 25].This algorithm is one of the fastest algorithms in categorization [26]. For simplicity, we consider two categories that are ranked $i = 0, 1$. The goal is to use an initial membership set with known membership members to create a score, so that higher scores are associated with Category 1 members and lower scores with category 0 members. So the ranking is achieved by comparing these privileges with one threshold. If $P(i|x)$ is the probability that a member with the vector of space $x = ()$ belongs to the category i, each uniform function $P(i|x)$ will create a suitable score. In particular, the $P(1|x)P(0|x)$ ratio will be appropriate. The initial probability is to divide $P(i|x)$ into $f(x|i)/P(i)$, where $f(x|i)$ is the conditional distribution x for the samples of the category i, and $P(i)$. The probability of belonging to a category is i (the initial probability of category i). That is, the ratio is obtained through relation (13).

$$\frac{P(1|x)}{P(0|x)} = \frac{f(x|1)P(1)}{f(x|0)P(0)} \tag{14}$$

To use this ratio in the categories, we need to estimate $f(x|i)$ and $P(i)$. If the instruction set is a random sample of the total population, then $P(i)$ can be calculated directly from the population of samples of the category i. To calculate $f(x|i)$, the simple binary method acts in such a way that the members of x are independent, Then each of the uniform distributions $f(x\_j|i)$, $j = 1, ..., P$ is measured separately. Therefore, the P-dimensional problem is reduced to problems of a form. One-dimensional estimate is simple and it requires smaller sets of learning sizes to get the right size. The most important features of naïve bayes are simplicity estimation and very fast and does not require complex repeating measurements.

3.5 Decision Tree

The decision tree classifier is a tree in which the internal nodes represent the features; the edges left out of the nodes are criteria for selecting the attributes, and the leaves representing the categories. The decision tree consists of two stages of growth and pruning. At the growth stage, a decision tree is created from educational data. At the pruning stage, the part of the tree is pruned in such a way that the test is not done on that branch. The decomposition tree group is divided into two categories CART and C4.5 based on the feature selection criteria in the decision nodes. CART is the classification and regression algorithm.C4.5 method makes decision tree by following top-down greedy model and descendant of ID3 decision tree. The decision tree can be pre-pruned or post-pruned to make it easily interpretable and simple [26].

3.6 K-Nearest Neighbors (KNN)

The K-Nearest Neighbors (KNN) algorithm is a model-based method. This algorithm first learns from training data and then predicts test data with the learned model, the model-free of KNN does not have training step and conducts classification tasks by first calculating the distance between the test data and training data to obtain its nearest neighbors and then conducting KNN classification. This algorithm has easily implementation and good classification performance. The KNN algorithm is a very popular method in machine learning and data mining, statistics and thus was voted as one of top ten data mining algorithms that produced. The KNN algorithm assigning an optimal K value with a fixed expert-predefined value for all tests and assigning different optimal K values for different test data. The K value is very important and affect to performance of classification [27].

3.7 Evaluation and Results

For evaluation of image classification models there are some criteria. This can be calculated depending on the value of TP, FP, TN, and FN [9]. Different scenarios for categories and images are shown in Table (1), according to the input data set for the classification of TP, FP, TN, and FN values for both positive and negative groups. The performance parameters can be calculated as Eq. (14) to (19).

Table 1 Different scenarios for categories and images.

	Predicted	
	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{15}$$

$$\text{Sensitivity (Recall)} = \frac{TP}{TP + FN} \tag{16}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{17}$$

$$F1_i = 2 \cdot \left[\frac{\text{Precision} * \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}\right] \tag{18}$$

$$ER = \frac{FN + FP}{TP + FP + FN + TN} = 1 - \text{Accuracy} \tag{19}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{20}$$

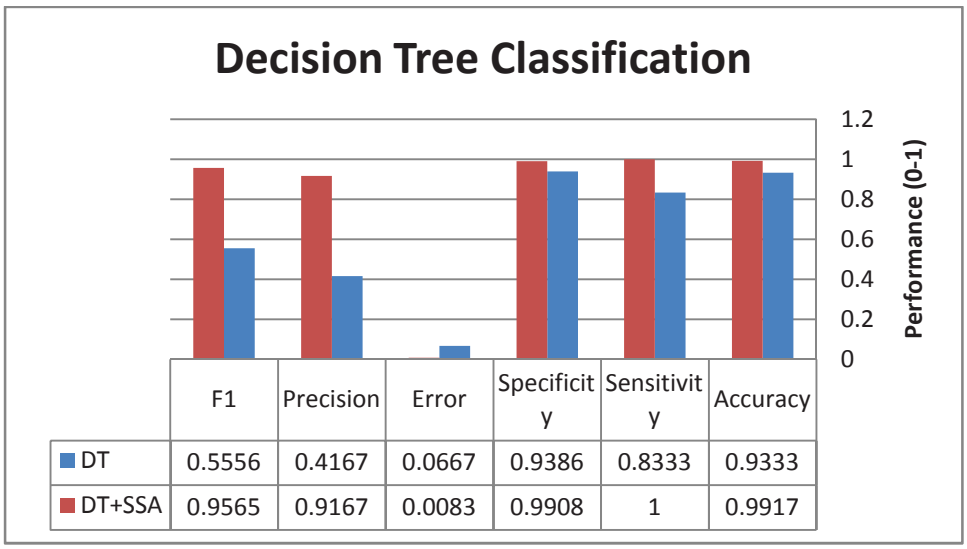
Where TP is True Positive, FP is False Positive, TN is True Negative, and FN is False Negative. At this stage, the proposed models are implemented and evaluated using the above criteria in matlab software 2017 version. Criteria include Accuracy, Sensitivity, Specificity, Error, Precision, and F1. All results are in percent. In this step, we used the DT classifier to evaluate the proposed method in the learning phase. Figure 3 shows the results of the evaluation of the proposed method using the DT classifier. The horizontal axis represents the evaluation criteria and the vertical axis represents the performance of the proposed method using the DT with SSA algorithm and DT algorithms in the base state. The results show the high efficiency of the proposed method using decision tree with SSA algorithm in Accuracy criterion of 99.17%, sensitivity criterion of 100%, Specificity criterion of 99.08%, Error criterion of 0.0083%, Precision criterion with 91.67% and the F1 criterion is 95.65%. As we can see, the proposed method is more efficient in comparison to DT learning methods in all of the proposed criteria.

Figure 6 shows the results of the evaluation of the proposed method using the KNN classifier. The horizontal axis represents the evaluation criteria and the vertical axis represents the performance of the proposed method using the KNN+SSA algorithm and KNN algorithms in the base state. The results show the high efficiency of the proposed method by using the KNN classifier with the SSA algorithm in the Accuracy criterion of 96.67%, the sensitivity criterion is 75%, the specificity criterion is 100%, the error criterion is 0.0333%, the Precision criterion is 100% and The F1 criterion is 85.71%.

As we can see, the proposed method is more effective in all of the criteria than KNN's methods of learning in basic state.

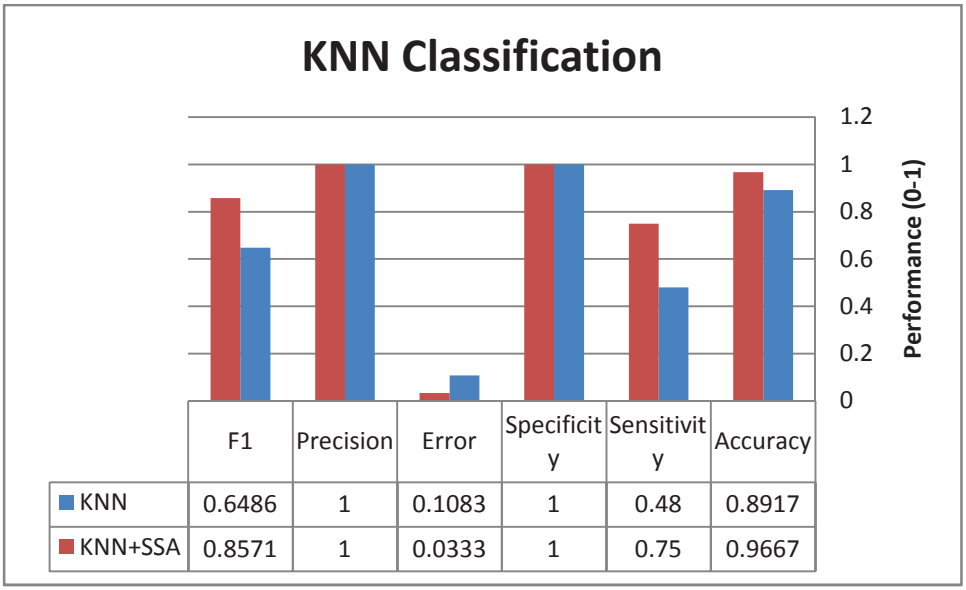
Figure 7 shows the results of the evaluation of the proposed method using the NB classifier. The horizontal axis represents the evaluation criteria and the vertical axis represents the performance of the proposed method using the NB+SSA algorithm and NB algorithms in the base state. The results show the high efficiency of the proposed method by using the NB classifier with the SSA algorithm in the Accuracy criterion of 97.5%, the sensitivity criterion is 100%, the specificity criterion is 97.3%, the error criterion is 2.5%, the Precision criterion is 75% and The F1 criterion is 85.71%.

As we can see, the proposed method is more efficient in comparison with NB learning methods in all of the proposed criteria.



	F1	Precision	Error	Specificity	Sensitivity	Accuracy
DT	0.5556	0.4167	0.0667	0.9386	0.8333	0.9333
DT+SSA	0.9565	0.9167	0.0083	0.9908	1	0.9917

Figure 5 The results of the evaluation of the proposed method using the DT Classifier.



	F1	Precision	Error	Specificity	Sensitivity	Accuracy
KNN	0.6486	1	0.1083	1	0.48	0.8917
KNN+SSA	0.8571	1	0.0333	1	0.75	0.9667

Figure 6 The results of the evaluation of the proposed method using the KNN classifier.

4. CONCLUSIONS

In this paper, a new learning approach based on the SSA algorithm was proposed that was implemented and evaluated on learning algorithm DT, KNN and NB. The results demonstrate the improvement of the performance of learning algorithms in all the achieved criteria by using the SSA algorithm in comparison with traditional learning algorithms. In the accuracy criterion, the best performance of the proposed model is using the DT learning method with a value of 99.17% compared to the proposed method using the KNN learning method with the value of 96.67% and the proposed method using the NB learning method with the value of 97.5%. In the Sensitivity benchmark, the best performances of the proposed model is using the DT and NB learning method with value of 100% compared to the proposed method using the KNN learning method with a value of 75%. In the Specificity criterion, the best performance of the proposed model is based on KNN learning method with a value of 100% compared with the proposed method using the DT learning method with the value of 99.08% and the proposed method using the NB learning method with the value of 97.03%. In the Error criterion, the best performance of the proposed model using the DT learning method with the value of 0.83% is compared with the proposed method using the KNN learning method with the value of 3.33% and the proposed method using the NB learning method with the value of 2.5%. In the Precision criterion, the best performance is related to the proposed model using the KNN learning method with value of 100% compared with the proposed method using the DT learning method with value of 91.67% and the proposed method using the NB learning method with the value of 75%. In the F1 criterion, the best performance of the proposed model based on the DT learning method with value of 95.65%, compared with the proposed method using the KNN learning method value of 85.71% and the proposed method using the NB learning method value of 85.71%.

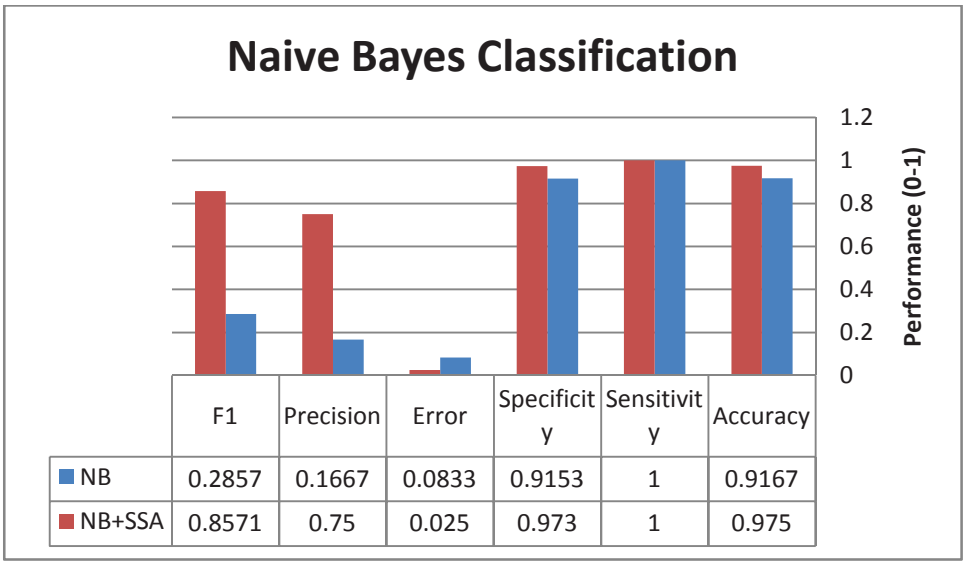


Figure 7 The results of the evaluation of the proposed method using the NB classifier.

REFERENCES

1. T. Das, "Machine Learning algorithms for Image Classification of hand digits and face recognition dataset", *International Research Journal of Engineering and Technology*, Vol. 4, No. 12, 2017, pp. 640-649.
2. I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, "Practical Machine Learning Tools and Techniques", *Morgan Kaufmann*, Fourth Edition, 2016.
3. N. M. Nasrabadi, "Pattern Recognition and Machine Learning", *Journal of Electronic Imaging*, Vol. 16, No. 4, 2007.
4. M. Kalekar, B. Sonawane, "A Survey on Medical Image Classification Techniques", *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 5, No. 7, 2017, pp. 13510-13516.
5. N. Orlov, L. Shamir, T. Macura, J. Johnston, "DMEckley, IG Goldberg, WND-CHARM: multi-purpose image classification using compound image transforms". *Pattern Recognit.Lett.*, Vol. 29, No. 11, 2008, pp. 1684–1693,
6. S. Sivasundari, R. S. Kumar, M. Karnan, "Review of MRI Image Classification Techniques", *International Journal of Research Studies in Computer Science and Engineering*, Vol. 1, No. 1, 2014, pp. 21-28.
7. H. Mohsen., E.-S. El-Dahshan, A. Salem , "A machine learning technique for MRI brain images", *In Proceedings of the 8th International Conference on Informatics and Systems*, 2012, pp. 161-165.
8. S. S. Shirke, J. A. Kendule, , S. G. Vawhare,, "An Approach for PCA and GLCM Based MRI Image Classification", *In Proceedings of the International Conference on Advanced Technologies*, 2016, pp. 265-274.
9. R. Ahmmed, A. Swakshar, M. F. Hossain, Md. A. Rafiq, "Classification of Tumors and It Stages in Brain MRI Using Support Vector Machine and Artificial Neural Network", *In Proceedings of the International Conference on Electrical, Computer and Communication Engineering*, 2017.
10. M. L. Antonie, O. R. Za¨ıane, A. Coman, "Application of Data Mining Techniques for Medical Image Classification", *In Proceedings of the Second International Conference on Multimedia Data Mining*, 2001, pp. 94-101.
11. Y. Zhang, L. Wu, "An MR brain images classifier via Principal Component Analysis and kernel support Vector Machine", *Progress In Electromagnetic Research*, Vol. 130, 2012, pp. 369-388.
12. S. Goswami, L. K. P. Bhaiya, "Brain Tumour Detection Using Unsupervised Learning Based Neural Network". *In Proceedings of the International Conference on Communication Systems and Network Technologies*, 2013, pp. 573-577
13. S. Sindhumol, A. Kumar, K. Balakrishnan, "Spectral clustering independent component analysis for tissue classification from brain MRI", *Biomedical Signal Processing and Control*, Vol. 8, 2013, pp. 667-674
14. B. H. Nayef, S. Sahran, R. I. Hussain, S. N. H. S. Abdullah, "Brain imaging classification based On Learning Vector Quantization", *In Proceedings of the 1st International Conference on Communications, Signal Processing and their Applications*, 2013, pp. 1-6
15. H. Kalbkhani, M. G. Shayesteh, B. Zali-Vargahan, "Robust algorithm for brain magnetic resonance image (MRI) classification based on GARCH variances series", *Biomedical Signal Processing and Control*, Vol. 8, 2013, pp. 909-919.
16. V. Gupta, K.S. Sagale, "Implementation of classification system for Brain Cancer using back propagation network and MRI", *International Conference on Engineering*, 2012, pp: 1-4
17. R. Mishra, "MRI based brain tumor detection using wavelet packet feature and artificial neural networks", *In Proceedings of the International Conference and Workshop on Emerging Trends in Technology*, 2010, pp. 656-659
18. J. D. Hemanth, C. Vijila, A. I. Selvakumar, J. Anitha, "Performance Improved Iteration-Free Artificial Neural Networks for Abnormal Magnetic Resonance Brain Image Classification", *Neuro computing*, Vol. 13, 2013, pp. 98-107.
19. E. S. A. El-Dahshan, T. Hosny, A. B. M. Salem, "Hybrid intelligent techniques for MRI brain images classification", *Digital Signal Processing*, Vol. 20, 2010, pp. 433-441
20. http://www.med.harvard.edu/AANLIB
21. L. Madin, "Aspects of jet propulsion in salps", *Canadian Journal of Zoology*, Vol. 68, 1990, pp. 765–777.
22. S. A. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Farism S. M. Mirjalili, "Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems", *Advances in Engineering Software*, Vol. 114, 2017, pp. 163-191.
23. Q. Bone, P. A. V. Anderson, A. Pulsford, "Communication between individuals in salp chains I morphology of system. *Proceedings of*

the Royal Society of London, Series B, Biological Sciences, Vol. 210, No. 1181, 1980, pp. 549-558.

24. D. Barber, "Bayesian Reasoning and Machine Learning", *Cambridge University Press*, 2012.
25. H. LIU, M. COCEA, W. DING, "DECISION TREE LEARNING BASED FEATURE EVALUATION AND SELECTION FOR IMAGE CLASSIFICATION", *In Proceedings of the International Conference on Machine Learning and Cybernetics,* 2017.
26. R. G. Ramani, K. Sivaselvi, "Classification of Pathological Magnetic Resonance Images of Brain using Data Mining Techniques", *In Proceedings of the 2nd International Conference on Recent Trends and Challenges in Computational Models*, 2017.
27. S. Zhang, X. Li, M. Zong, X. Zhu, R. Wang, "Efficient kNN Classification With Different Numbers of Nearest Neighbors", *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, 2017, pp. 1-12.