IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

VOLUME : 7 ISSUE : 05 Print / Issue Publication Date: 06-Nov-2022

ISSN : 2455-2143

DOI : 10.33564/IJEAST.2022.v07i05.004

Indexed In

Google SCHOLAR OPEN ACCESS doi Crossref

WWW.IJEAST.COM

editor@ijeast.com

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

OPTIMIZATION OF SUPPORT VECTOR MACHINES BY META-HEURISTIC APPROACH FOR DIFFERENTIAL DIAGNOSIS AND CLASSIFICATION OF ORAL DISEASES

R. Soltani, A. Jaouadi
ECE – Paris School of Engineering
37 quai de Grenelle, Paris - France

Abstract—**Classifying cancers as different groups is the key to a more accurate diagnosis. In every case, the goal is to reach an in-time diagnosis of the disease. The diagnosis, prediction, and control of oral cancers have been based on clinical and histopathological symptoms so far. Given the failure to confirm an in-time diagnosis in all the conventional methods, multiple biopsies are necessary for reaching a differential diagnosis, although patients are not satisfied with such biopsies. Machine learning algorithms can be employed to prevent such complications and reach faster and more accurate diagnosis. In this paper, Support Vector Machine algorithms were used for a four-class classification of a dataset containing 210 real samples. The research goal was to classify patients' cancers as four classes: normal, lichen planus, oral squamous cell carcinoma, and leukoplakia. To improve the accuracy of the classifier, the Support Vector Machine was combined with meta-heuristic algorithms such as Artificial Bee Colony algorithm and Ant Colony Optimization to propose a new approach. The proposed diagnostic approach showed a classification accuracy of 94.28%.**

Keywords— **Classification, Support Vector Machine (SVM), Artificial Bee Colony algorithm (ABC), Ant Colony Optimization algorithm (ACO), Neural Network, Oral Cancers.**

I. INTRODUCTION

Oral cancer is the sixth common cancer in the world [1], [2]. It is more prevalent among men. However, the oral cancer incidence rate has been increasing among women recently. This type of cancer is more common in developing countries in the East Asia and other countries such as India and Sri Lanka. It affects most parts of the mouth such as tongue, buccal, and hypopharynx as well as palate and lips. The first symptoms of many mucosal skin diseases such as lichen planus appear first in the oral mucosa. Leukoplakia is one of the pre-malignant lesions. In the absence of a proper diagnosis and treatment, it can turn into a malignant disease such as oral squamous cell carcinoma (SCC). SCC accounts for about 80% of the oral cavity malignancy. Factors such as smoking and alcohol consumption, personal hygiene, nutrition, age, genetic differences, human papillomavirus infection and race can be the major causes of these diseases. An initial assessment of oral pre-cancerous lesions has a significant impact on the mortality rate of oral cancer. Therefore, early diagnosis of the disease can greatly help the treatment. Since multiple biopsies often cause discomfort to the patients, preclinical analytical tools such as machine learning methods can be used practically. A wide variety of machine learning methods of data mining and artificial intelligence have been used for feature selection and classification in the diagnosis of diseases. Researchers have found that the use of modern techniques along with the main diagnosis and treatment templates can be far more practical and accurate[3], [4], [5], [6], [7], [8], [9], [10], [11], [12], and [13].

To address this public health issue, this study examines the performance of Support Vector Machines (SVM) machine learning algorithm for a real oral cancer images classification. The first purpose of this paper is to discuss the accuracy of SVM algorithm in oral cancer classification, which can be used as a diagnostic aid by the medical community. Comparison with Multilayer Perceptron (MLP) is made. The second one is how to improve SVM by combining them with Artificial Bee Colony algorithm (ABC) and Ant Colony Optimization algorithm (ACO). Although diagnosis of precancerous symptoms can be a great help to treatment, it is worth noting that the results obtained through the models used here must be approved by medical doctor's assessment.

The support vector machine (SVM) was introduced by Vladimir Vapnik in 1963 [14]. This algorithm was generalized by Vapnik et al. in 1995 for nonlinear models. Regarding classification and regression, SVM techniques are powerful methods for learning high-dispersion models. The benefits of SVM include relatively simple training and not being trapped in local optimums in addition to a relatively good response to high-dimensional data. There are two important issues in the use of SVM: how to select the input dataset and how to set and

initialize the best parameters of the kernel function. These two issues are closely related because the input set affects the selection of kernel parameters [15].

This paper contains 6 sections. After Introduction in Section 1, related works are discussed in Section 2. Section 3 defines oral diseases in detail. In Section 4, the dataset is presented. Section 5 describes the model and its implementation. Conclusion and discussions are presented in Section 6.

II. RELATED WORKS

Different methods have been proposed by related studies for the diagnosis of oral diseases. Some methods are hybrid, and some are pure. Less attention has been paid to the classification of the three types of oral diseases. There were also several methods for feature extraction.

A swathy et al. (2021) [16] proposed a breast cancer classification system that uses support vector machine (SVM) classifier based on integrated features (texture, geometrical, and color). Results showed that among different classifiers, SVM gave better results with a test accuracy of approximately 90%. Jeyaraj et al. (2019) [17] developed a deep Convolution Neural Network (CNN) algorithm for oral cancer detecting by investigating patient hyperspectral images. They have obtained classification accuracy of 91.4% with sensitivity 0.94 and a specificity of 0.91 for 100 image data sets. Chodorowski et al. (2002) [18] compared SVM with other traditional categories such as Fisher Linear Discriminant and Neural Networks with 5-fold cross validation for the classification of oral ulcers. In the 2-fold classification of cancerous and noncancerous types with different kernel functions, the linear SVM had a better performance with the accuracy of 90%. Regarding the 4-fold classification, SVM-RBF had a better performance than other techniques with the accuracy of 78%. With RBF kernel function, SVM has been more successful than the classic RBF and Fisher Linear Discriminant. Mutha Rama Krishnan et al. (2009) [19] auto-classified oral submucosa fibrosis used SVM. In this paper, the main theory of analysis was on the SECT cells. They transmitted images from the cell filter, and they used the target feature method to extract features. Finally, the proposed model had the accuracy of 88.69%, a sensitivity of 90.46% and a specificity of 87.54%. In another study (2012) [20], they also auto-diagnosed cancer and classified a dataset containing 42 cases as two groups of oral submucosa fibrosis with dysplasia and without dysplasia using a pair of classifiers such as the K-Nearest Neighbor, Gaussian Mixture Model, Fuzzy Model, Decision Tree, and Radial Basis Probabilistic Neural Network. They achieved the accuracy of 95.7% through the fuzzy classifier. Anoradeh et al. (2013) [21] presented a model based on the gray level co-occurrence matrix extraction technique and SVM for classifying oral cancers as benign and malignant groups. Input images achieved a accuracy of 92.5%, a sensitivity of 92.85% and a specificity of 92.3% by using the pre-processed contrast linear stretching method and the proposed method. In another study [22] , they also employed learning algorithms with different controllers such as k-nearest neighbor, back-propagation neural networks, naïve Bayesian method and SVM for the diagnosis of oral cancer on x-ray dental images by using different extracting methods such as the gray level co-occurrence matrix and gray level run length and intensity histogram. They achieved the accuracy of 94% for SVM, 83% for k-nearest neighbor, 85% for the back-propagation neural network and b77% for naïve Bayesian. Anirban Mukherjee et al. (2006) [23] analyzed four wavelet functions of GABOR, HAAR, DB2, DB4 to extract features in detecting oral pre-cancerous states and classifying them as three advanced, non-advanced and normal categories by using a feed forward 3-layer neural network . The efficiency of each one was measured by by the help of the Properly Classified Block Index (PCBI). The highest error rate was associated with non-advanced category; however, the rate of PCBI was always above 50%. They also used 17 images to evaluate the capability of the model, acting effectively and successfully against the advanced cancer. The model was based on the mentioned wavelets and HAAR wavelet. Venkatakrishnan et al. (2013) [24] used a histogram-based technique for feature extraction. They also employed SVM with different kernel functions and 4-fold cross validation method to classify images as two groups of normal and oral submucosa fibrosis. According to the conducted tests, SVM with Gaussian kernel function had the highest classification accuracy (94%). Pablo Rrosado et al. (2013) [25] presented an SVM-based model, clinical parameters, and molecular markers with two parallel methods (non-concave penalty coefficient and Newton's method) to determine the variables and the survival rate of patients with oral squamous cell carcinoma (OSCC). The tests conducted on both methods using cross validation indicated that the degree of classification accuracy for living patients was equal to 97.56%. As to the deceased patients, it was 100% at an average accuracy of 98.55%. Yi Li et al. (2010) [26] used the Raman spectral character and SVM with the RBF kernel function to classify a dataset containing 186 records including patients with leucoplakia, OSCC, and normal cells. The wavelet transformation was also used to select the feature. The cross-validation method was used for validation. Finally, they achieved accuracy of 98.75%, sensitivity of 97.66% and specificity of 100%. Siow-Wee Chang et al. (2013) [27] examined four methods including neuro-fuzzy adaptive inference system, artificial neural network, SVM with linear kernel function, logical regression, and the 5-fold validation on a dataset containing 31 samples with 15 features extracted by Relief-F, Pearson Correlation Coefficient, Genetic Algorithm, Pearson-Genetic Correlation Coefficient and Genetic-Relief-F for classifying oral diseases. Finally, the Relief-F-Genetic hybrid model and the Neuro-Fuzzy Adaptive System achieved accuracy of 91.83%. Rouhollah Maghsoudi et al. (2013) [28] used MultiLayer Perceptron (MLP) to detect lichen planus, leukoplakia, and OSCC on a set of 150 patients

with 4 features by testing and lowering the error rate with low successive Iteration that achieved the 0.20 error at the 100th Iteration. Kalaiaras et al. (2015) [29] used several methods of classification such as the random forest, Bayes classifier, C4.5, Apriori algorithm and SVM in addition to the Gabor wavelet for extraction of features in the category of oral cancer. Ultimately, the accuracy of 76% was recorded for the Apriori algorithm, 81% for C4.5, 79% for Bayes, 75% for SVM, and 83% for the random forest algorithm. Li Yeh Chuang et al. (2011) [30] compared the SVM with the RBF kernel function, the Bayes classifier, k-nearest neighbor, and the random forest for the diagnosis and prediction of oral cancer. They also used 10-fold cross validation and a holdout validation to evaluate a dataset with 238 records. They achieved the accuracy of 64.2% by the SVM with 10-fold cross validation. Neha Sharma et al. (2014) [31] used the three-layer perceptron neural network and SVM with RBF function to diagnose and predict the survival rate of patients with oral cancer on a dataset containing 33 variables and 1024 records. They also used the 10-fold cross validation. Finally, the accuracy of 73.56% and a sensitivity of 73.53% were recorded for the SVM, whereas an accuracy of 70.05% and a sensitivity of 36.65% were recorded for the multi-layer perceptron.

III. ORAL DISEASES

Common oral diseases include pre-cancerous and cancerous conditions. Lichen planus and leukoplakia are pre-cancerous conditions, whereas Oral Squamous Cell Carcinoma (OSCC) is a cancerous condition. Each of these conditions is described in brief.

3.1 Lichen planus

Lichen planus is an inflammatory autoimmune type of cutaneous-mucosal disease, which can affect squamous epithelium, skin, oral mucosa, and genital areas. Typically, lichen planus lesions involve buccal mucus bilaterally. Lichen planus may contain two red and white components with the feature of clinical classification cause the disorder. For clinical diagnosis of lichen planus, network surface or popular features should be present. If there are plaque-like, blouse, and injured areas besides them, the lichen planus lesions are diagnosed accordingly [32], [33], [34]

3.2 Leukoplakia

Oral leukoplakia is known as a white lesion of the oral mucosa, which cannot be defined as any other definable lesions. The incidence of leukoplakia as a pre-malignant lesion involves various genetic changes. This belief is supported by observing markers of genetic disorders that are presented differently in different leukoplakia conditions. The activation of oncogenes and elimination of damages to repressive genes and DNA repairing genes are all involved in genomic dysfunction that controls cell division [35]. Etiologic factors for leukoplakia include tobacco, alcohol, and human papillomavirus. Nearly 70% of oral leukoplakia is found on the vermilion of lips, buccal mucosa, and gums. But lesions on the tongue, lips and the floor of the mouth make up 90% of the lesions, indicating the development of dysplasia or carcinoma [36].

3.3 Oral Squamous Cell Carcinoma (OSCC)

Oral cancers often affect the tongue, oral throat, and the floor of the mouth. The incidence of cancer on the lips, gums, back surface of the tongue, and the palatal are less common. Oral cancer is an age-related phenomenon which reflects the time needed for the accumulation of genetic changes, continuous exposure to initiators, and promoters of cancer (including physical and chemical stimuli, viruses and hormonal effects), cellular age, and reduced immune function due to age increase. There are various diseases and lesions that can be potentially cancerous. These lesions include leukoplakia, sublingual leukoplakia, submucosa fibrosis, lichenoid lesions, a history of previous malignancy in the mouth, etc. Oral squamous cell carcinoma risk factors include tobacco use, alcohol consumption and betel leaf consumption, nutrition factors (vitamin A and carotenoids deficiency) and other factors (the use of dentures, irregular teeth or restorations, and chronic chewing habits). OSCC has clinical manifestations such as exophytic, endophytic, leukoplakia, and erythropoietic symptoms [36], [35].

IV. DATA SET

In this paper, we applied the dataset of research that has been done at Gilan University of Medical Sciences. It included 210 clinical images of oral mucosa, 154 images of which were related to the three studied lesions, i.e., 50 cases of erosive lichen planus and ulcerative, 53 cases of leukoplakia, 51 cases of oral SCC, and 56 images were related to healthy mucosa. The dataset contained 30 features. These 30 features are in fact key points obtained by a feature extraction algorithm called Surf [37]. Simple statistical information for these features is presented in Table 1.

Table -1 Simple Statistics of Each Feature

Feature Number	Mean	Standard Deviation	Minimum	Maximum
1	**6.024**	**11.986**	**0**	**124**
2	**10.448**	**24.622**	**0**	**162**
3	**5.538**	**15.298**	**0**	**136**

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

4	**3.51**	**5.731**	**0**	**36**
5	**8.181**	**15.56**	**0**	**122**
6	**7.186**	**16.366**	**0**	**134**
7	**6.324**	**10.091**	**0**	**71**
8	**8.386**	**18.447**	**0**	**146**
9	**4.671**	**6.933**	**0**	**43**
10	**6.476**	**11.952**	**0**	**118**
11	**5.752**	**17.832**	**0**	**206**
12	**5.338**	**9.106**	**0**	**57**
13	**4.738**	**5.713**	**0**	**37**
14	**4.69**	**6.511**	**0**	**39**
15	**6.005**	**9.017**	**0**	**71**
16	**8.629**	**20.94**	**0**	**130**
17	**9.576**	**21.978**	**0**	**165**
18	**5.79**	**14.201**	**0**	**126**
19	**4.448**	**6.567**	**0**	**43**
20	**4.01**	**6.676**	**0**	**41**
21	**6.824**	**7.656**	**0**	**59**
22	**3.814**	**6.601**	**0**	**50**
23	**8.248**	**16.578**	**0**	**120**
24	**6.395**	**10.069**	**0**	**85**
25	**9.21**	**23.468**	**0**	**282**
26	**5.576**	**9.772**	**0**	**96**
27	**5.967**	**10.645**	**0**	**92**
28	**5.481**	**8.781**	**0**	**65**
29	**3.914**	**5.187**	**0**	**33**
30	**3.729**	**6.6**	**0**	**46**

Diagnostic factors for clinical images are ready to enter the software to begin the simulation process, also the number of primary records of different classes are shown in Table 2.

Table -2 Number of Dataset Records

Class	F1	F2	F3	F4	F5	F6	F7	F8	F9	….	F30
Leukoplakia	**4**	**3**	**1**	**1**	**0**	**5**	**4**	**0**	**2**	….	**0**
Leukoplakia	**1**	**1**	**3**	**2**	**9**	**1**	**1**	**0**	**5**		**2**
Leukoplakia	**0**	**4**	**1**	**0**	**1**	**1**	**5**	**2**	**10**		**0**
Leukoplakia	**1**	**1**	**0**	**2**	**3**	**3**	**6**	**8**	**6**		**1**
Lichen Plan	**11**	**1**	**1**	**14**	**0**	**11**	**0**	**0**	**9**		**5**
Lichen Plan	**2**	**4**	**2**	**5**	**6**	**8**	**5**	**6**	**3**		**4**
Lichen Plan	**1**	**0**	**0**	**7**	**8**	**10**	**15**	**0**	**1**		**11**
SCC	**4**	**6**	**0**	**4**	**3**	**0**	**1**	**8**	**6**		**7**
SCC	**14**	**27**	**2**	**9**	**5**	**0**	**4**	**19**	**2**		**6**
SCC	**1**	**15**	**0**	**0**	**3**	**1**	**1**	**2**	**4**		**9**
SCC	**0**	**15**	**0**	**0**	**1**	**2**	**0**	**8**	**4**		**6**
Normal Mucosa	**18**	**4**	**10**	**2**	**5**	**4**	**7**	**6**	**5**		**0**
Normal Mucosa	**3**	**8**	**22**	**20**	**5**	**38**	**11**	**0**	**8**		**4**
Normal Mucosa	**0**	**0**	**1**	**0**	**17**	**0**	**11**	**0**	**2**	….	**2**

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

V. MODEL

Since convergence to lower levels of error and increased accuracy allow analysts to be more precise in diagnosis and treatment, this paper tried to improve the accuracy with the use of meta-heuristic algorithms. A slight improvement in accuracy can also produce significant results. There are several learning parameters that can be used to create SVM. Parameter C can change to control trade-off between the complexity of the decision rules and the frequency of the errors. The value of C is determined by the user. If the values of C are small, this parameter can increase the error rate, and if it is large, it ill lead to a behavior similar to a hard margin SVM [38]. The gamma parameter in the RBF kernel function can be optimized based on Fisher's discrimination. An ideal gamma along with its ideal range is obtained by a dynamic evaluation [39]. The goal is to find the ideal value of the parameters to produce the maximum of model performance. Furthermore, artificial bee colony algorithm (ABC) and ant colony optimization algorithm (ACO) were used to optimize C and gamma parameters in multi-class SVM.

The proposed model consists of two main steps:

i. The first step is to use a dataset with 210 records: All records are normalized between 0 and 1, then 10-fold cross validation method is used to divide the train and test data. Finally, a training-testing set is obtained. Then classification and diagnosing by using machine learning methods: First, the multi-class SVM (one-against-all) are applied along with five different kernel functions to the dataset in order to classify it as 4 categories of normal, lichen planus, leukoplakia, OSCC, in a way that each class is trained against other classes. The data of the class itself are attributed to the label +1, and the data of the other classes are attributed to the label -1. If the number of classes is N, then N number of SVMs are trained. Each of them corresponds to one of the classes. After training the classes at the testing stage, each test sample is applied to all SVMs. The winner class is the one whose SVM has the highest output. If the training data and the number of classes are high, SVM training time will be long. Finally, ABC and ACO were used to optimize C and gamma parameters in this study. To describe the function of the proposed model, SVM was applied to the fitness evaluation function of ABC and ACO. Therefore, the optimal values of the parameters were obtained after a certain number of iterations until convergence was achieved. The objective function that must be minimized was the mean squared error (MSE) so that the lower mean square error led to an increase in accuracy of the classification.

ii. The second step is to evaluate the efficiency of model using the values of the confusion matrix and important criteria. At this stage, multi-class SVM, as well as SVM-ABC and SVM-ACO, with different kernel functions, are executed. Each algorithm is evaluated using the confusion matrix, accuracy, recall, precision, and F-measure. The results are presented in the form of tables. Finally, to extract the best model, all the models will be compared. The proposed model diagram for the classification of oral diseases is presented in the Fig. 1.

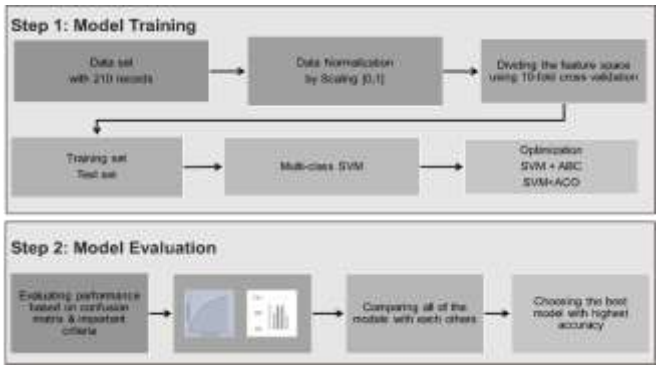

Fig. 1. Diagram of the proposed model

5.1 Support vectors machines

Linear SVM is used to categorize data that are linearly separable. Linear SVM tries to maximize the margins between the separator hyper-planes. The available template for the maximum margin is called the support vector. Similarly, the hyper-plane with the maximum margin is called maximized hyper-plane [19]. The separator hyper-plane g(x) is defined by the weight vector w and the bias b:

$$g(x) = w^T x + b \quad (1)$$

In this equation, $g(x) < 0$ for $y_i = -1$ and $g(x) \geq 0$ for $y_i = +1$. In other words, the training samples of the two categories are separated by the hyper-plane $g(x) = w^T x + b = 0$. SVM finds those hyper-planes that contain the largest amount of separation between the decision function of the two categories. Thus, the total distance between the two margins is equal to $(2/w^T w)$ which has been maximized. Mathematically, this

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

hyper-plane is obtained by minimizing the cost function.

$$j(w) = \frac{1}{2} w^T w \quad (2)$$

Subject to separability constraints

$$w^T x_i + b \geq 0 \text{ for } Y_i = +1 \quad (3)$$

Or

$$w^T x_i + b < 0 \text{ for } Y_i = -1 \; ; \; i = 1,2,\ldots,n \quad (4)$$

Equivalently, these equations can be written as the following:

$$y_i (w^T x_i + b) \geq 1 \; ; \; i=1,2,\ldots,n \quad (5)$$

To solve the problem of the quadratic exponential function, the saddle point of the Lagrange function (L_P) must first be obtained:

$$L_P (w,b,\alpha) = \frac{1}{2} ||w||^2 - \sum_{i=1}^{N} \alpha_i \left[y_i (w^T x_i + b) - 1 \right] \quad (6)$$

In equation (6), α_i refers to the diffusion coefficient of the Lagrange function; thus, finding critical points is very necessary because (L_P) must be minimized to w and b, and maximized to α_i. By Karush-Kuhn-Tucker (kkT) conditions of equation (6), the Lagrange function changes to optimize the (L_P) function. Therefore, the Lagrange dual function (L_D) could be maximized:

$$\text{MAX } L_D(\alpha) = \left[\sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j^T \right] \quad (7)$$

$$\text{MAX } L_D(\alpha) = \left[\sum_{i=1}^{n} \alpha_i - \frac{1}{2} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}^T H \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right] \quad (8)$$

In this equation, the value of ith and jth rows of Column H are equal to $H_{ij} = y_i y_i x_i^T x_i$

As a result, the Lagrange function is maximized to find the optimal hyper-plane. The α_0 solution for the quadratic optimization problem determines the parameters b* and w* for the optimal hyper-plane[19]. Therefore, the hyper-plane function (g(x)) is as follows:

$$g(x) = \text{sign} \left[\sum_{i=1}^{n} y_i \alpha_i^* x_i^T x_j + b^* \right] \quad (9)$$

Basically, one of the SVM strengths is the non-linear mapping of the input vector onto a high-dimensional feature space which is hidden from input and output views. This is done by the kernel functions [40] . Equation (9) can be shown as:

$$g(x) = \text{sign} \left[\sum_{i=1}^{n} y_i \alpha_i \, k(x_i x_j) + b \right] \quad (10)$$

Here, k is a kernel function measuring the distance and similarity of the input and the training vectors. This paper focused on RBF, polynomial, linear, quadratic, and MLP kernel functions.

Table – 3 Different Types of Kernel Functions

kernel	Equation*
RBF	$k(x_i,x_j)=\exp\left(-\frac{1}{2\sigma^2}\|x_i-x_j\|^2\right)$
MLP	$k(x_i,x_j)=\tanh(\beta_0 x_i^T x_j+\beta_1)$
Polynomial	$k(x_i,x_j)=(1+x_i^T x_j)^P$
Linear	$k(x_i,x_j)=x_i^T x_j$
Quadratic	$k(x_i,x_j)=(x_i^T x_j)^2$

The asterisk (*) "In the formulas (equations) of the kernel functions, x_i is the training input, x_j is the unlabeled input, σ^2 is the degree of variance, P is the polynomial degree, and β_0 and β_1 are constant values."

SVM is inherently binary classifiers. Sometimes, it is necessary to classify data as multi-class categories. The one-against-one and one-against-all methods are the techniques used in multi-class classifications [41] .

5.1.1 Initilisation and results

In this section, a multi-class SVM was investigated with different kernel functions and 20 executions as described in Table 4. The C parameters were randomly considered to be 0.8 for all kernel functions, and gamma parameters were 0.9 for the RBF kernel function.

Table - 4 Parameters Used in the Multi-Class SVM Algorithm

Parameter	**Value**
Kernel	**Polynomial, Quadratic, Linear, MLP, RBF**
Number of classes	**4**
Cross validation	**10-fold**
Gama	**0.9**
C	**0.8**

Accordingly, multi-class SVM with a quadratic kernel function achieved an accuracy of 91.9% (see Table 5), indicating its better performance compared to other classifiers. The mean of 20 iterations of each multi-class SVM indicated that this classifier was not associated with a significant deviation in terms of the performance correctness. In fact, the accuracy at each execution had a relatively small difference

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

from an average of 20 iterations The exact results of the comparison are shown in Table 5.

Table-5 The Performance of a Multi-Class SVM with Different Kernel Functions

kernel	Precision%	Recall%	F-measure%	Accuracy%
Quadratic	**92.41**	**92.10**	**91.96**	**91.90**
RBF	**86.45**	**78.82**	**80.24**	**79.05**
Linear	**89.68**	**89.59**	**89.51**	**89.52**
Polynomial	**88.33**	**87.75**	**87.59**	**87.62**
MLP	**80.56**	**77.66**	**77.98**	**77.62**

Table 5 also shows the value of the precision, the recall, the F-measure, and the accuracy for different kernels. One could note that the Quadratic kernel is the best kernel function that gives the best values for all metrics listed above.

In comparaison with Neural Network algorithm, we performed a Multilayer Perceptron (MLP) with same 10-fold cross-validation and 17 hidden layers. We found the following results: 76.4%, 76.2%, 76.2%, and 76.1% precision, recall, F-measure, and accuracy respectively.
We conclude then for this dataset, SVM performed much better than the artificial neural network (ANN) for all the metrics and specifically for the accuracy which is of 91.9% for SVM and only 76.1% for ANN.
In the following section, we aim at improving the value of the accuracy as we chose it as an optimal metric by combining the SVM algorithm with meta-heuristic approaches such as ABC and ACO.

5.2 Optimization

5.2.1 SVM+ABC

Forager bees are divided into three categories: employed bees, onlooker bees, and scouts. In the Karaboga Bee Algorithm, any source of food is a possible answer to the problem, and the amount of nectar available in each food source indicates the quality of the response, represented by the amount of suitability. This algorithm starts with the allocation of food sources (response) to the employed bees. On each iteration, each employed bee finds a source of food that is close to its source and evaluates the amount of nectar (fitness). The position of the ith source of food is represented by Xi = (xi1, xi2, ..., xid). F (xi) refers to the amount of nectar in the food source located in Xi. After seeing the employed bees dancing, an onlooker bee goes to the food source area located in Xi, with Pi probability[42] [43].

$$p_i = \frac{f(x_i)}{\sum_{k=1}^{s} f(x_k)} \qquad (11)$$

Here, s refers to the number of food sources or the number of employed bees. If the fitness of the new source is better than the best fitness obtained so far, the bee moves toward that new source of food and leaves the previous one; otherwise, it will remain in its previous source of food. When all the employed bees have completed this process, they share their fitness information with onlooker bees. Each onlooker bee chooses a food source based on the probability equation mentioned above. By doing so, good sources of food attract more onlooker bees than bad ones. After all onlooker bees have found a source of food, each one will search for a certain number of cycles in its neighborhood (limit) and find a source of food and measure its suitability (fitness). The best source of food among all the sources of food in the neighborhood of ith source and the ith source itself will be considered the new location of the ith source. And if the fitness does not improve, it becomes a scout and randomly looks for a new source of food. After the new location of each source of food is determined, another iteration of artificial bee colony algorithm begins. The whole process is iterated and iterated to satisfy the termination conditions[42] [43].

5.2.1.1 Initialization and results

The initialization of ABC algorithm is the first step to be performed to improve the accuracy of the modelling. We present in table 6 the different values of the parameters used in ABC algorithm and their explanations.

Table – 6 Parameters Used in the ABC Algorithm

Parameter	Value	Explication
NB	**20**	**The size of colony (employed bees + onlooker bees)**
FN	**NB/2**	**The number of food sources, which is half the size of the colony**
Limit	**30**	**A food source that cannot be optimized through the number of attempts identified by a parameter called 'limit' will be left by its employed bee.**
N_Iter	**20**	**The number of repetitive loops which is a sequestration criterion**
Domain	**[0.01, 1000]**	**The lower bound and upper bound for optimizing C value in SVM**

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

In this model, 20 iterations were considered for to the convergence of the model. Accuracy increased considerably. The results showed that the proposed SVM-ABC hybrid model increased the accuracy of the quadratic kernel function up to 93.33% versus 91.9% for SVM algorithm without no optimization. It is important to optimize parameters C and Gama by ABC to increase the rate of accuracy. Here we found for SVM-ABC hybrid model, C = 0.01 and Gamma = 51.28. The performances of SVM-ABC hybrid model for different kernel functions are presented in Table 7.

Table - 7 The Performance of a Multi-Class SVM + ABC with Different Kernel Functions

kernel	Precision%	Recall%	F-measure%	Accuracy%
Quadratic	**88.81**	**88.04**	**88.15**	**93.33**
RBF	**85.55**	**78.77**	**80.02**	**83.81**
Linear	**83.36**	**83.41**	**83.23**	**92.58**
Polynomial	**87.41**	**86.72**	**86.73**	**90.47**
MLP	**79.56**	**76.31**	**76.38**	**81.9**

In Fig. 2 we present the SVM-ABC optimization graph with quadratic kernel function as we found above that this kernel was the best one for SVM to obtain the maximum of the accuracy.
One could note that at the 20th iteration, the accuracy of SVM-ABC hybrid model is equal to 93.33% (orange dashed line), however for SVM alone the accuracy is less than 91.90% (green dotted-dashed line). One should precise that the value of the accuracy of SVM-ABC hybrid model is always greater than the one obtained from SVM alone for each iteration. We remark that this accuracy increases as function of the number of iteration and reaches a constant value for both SVM and SVM-ABC. The best value of accuracy 93.33% for ABC is reached at 11th iteration, however the SVM starts to get it maximum value at the 18th iteration. This tells us that SVM-ABC hybrid model converges much faster than the SVM algorithm without any optimization.

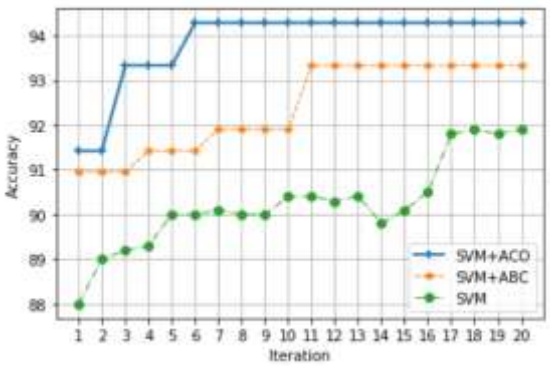

Fig. 2. Accuracy of SVM (green dotted-dashed line), SVM-ABC (orange dashed line) and SVM-ACO (blue solid line) with quadratic kernel.

5.2.2 SVM+ACO

Ant Colony Optimization (ACO) algorithm was initially proposed by Dorigo et al. in the 1990s [45], [46], [47] for the search of optimal paths in a graph. The algorithm was inspired by the behavior of ants looking for a path between their colony and a source of food. The foraging problem is an optimization problem where ants seek to collect a maximum of food with minimal energy. The energy is saved by finding the shortest path between the ants' nest and the source of food. Since many problems and several algorithms have emerged in different research areas.

The algorithm follows these different steps:

i. An ant chooses a path and traces a pheromone trail.
ii. All the ants travel a certain number of routes, each ant depositing a quantity of pheromone proportional to the quality of the route.
iii. Each edge of the best path is more reinforced than the others.
iv. Evaporation makes bad solutions disappear.

In this section we combine ACO with SVM algorithm, as we did previously with ABC, to optimize C and Gamma parameters. In the proposed novel ACO-SVM model, the parameters' values are dynamically optimized by implementing meta-heuristic process of artificial ants. And then the SVM model performs the classification task with the obtained optimal parameters' values. We found for SVM-ACO hybrid model, a C = 0.01 and Gamma = 30.73.

5.2.2.1 Initialisation and result

In table 8, we present the initial values of the parameters used in ACO algorithm.

TABLE - 8 Parameters Used in the ACO Algorithm

PARAMETER	VALUE	EXPLICATION
NL	2	THE NUMBER OF LEVEL
N_ANTS	20	THE NUMBER OF ANTS
N_ITER	20	THE NUMBER OF ITERATIONS
ALPHA	2	THE PHEROMONE IMPORTANCE
BETA	1	THE LOCAL IMPORTANCE HEURISTIC
RHO	0.02	THE EVAPORATION FACTOR
Q	100	THE PHEROMONE AMPLIFICATION FACTOR
TAU_0	1	THE INITIAL PHEROMONE LEVEL

International Journal of Engineering Applied Sciences and Technology, 2022
Vol. 7, Issue 5, ISSN No. 2455-2143, Pages 17-27
Published Online September 2022 in IJEAST (http://www.ijeast.com)

Subsequently, in the SVM-ACO hybrid model, 20 iterations were considered for to the convergence as we did with ABC. The results showed that the proposed SVM-ACO hybrid model increases the accuracy of the quadratic kernel function up to 94.28%. It is important to optimize parameters C and Gamma by ACO to increase the rate of accuracy as we fix the accuracy as an objective for our modelling. The results are presented in Table 9.

Table - 9 The Performance of a Multi-Class SVM + ACO with Different Kernel Functions

kernel	Precision%	Recall%	F-measure%	Accuracy%
Quadratic	**91.24**	**91.06**	**90.87**	**94.28**
RBF	**85.89**	**79.06**	**79.54**	**81.90**
Linear	**90.82**	**90.01**	**89.96**	**92.38**
Polynomial	**87.22**	**86.07**	**86.22**	**90.95**
MLP	**75.93**	**74.42**	**74.38**	**79.52**

We can see that quadratic kernel still gives the best results of all the metrics. In addition, SVM-ACO optimization graph with Quadratic Kernel Function are shown in Fig. 2. One should mention that the best accuracy for SVM-ACO is reached at the 6th iteration and its value is equal to 94.28% compared to 93.33% for SVM-ABC (obtained at the 11th iteration). Hence SVM-ACO is much faster than SVM-ABC to reach the maximum of the accuracy.

VI. CONCLUSION AND DISCUSSION

In this paper, the improvement and development of learning machines such as SVM were addressed by using ABC and ACO meta-heuristic algorithms for the diagnosis and classification of oral cancer diseases. Diagnosis of pre-cancerous symptoms can be a great help to treatment. Because of the high rates of oral cancer in recent years, various methods have been proposed for the diagnosis and classification of oral diseases; however, meta-heuristic and optimization algorithms have been used less often for diagnosis and treatment. In this study, after normalizing a dataset containing 210 records and 30 features, 10-fold cross validation methods were used to train and test the model. Then a multi-class SVM was compared with five different kernel functions. The multi-class SVM was extracted from quadratic optimized kernel function as the best intelligent diagnostic model with an accuracy of 91.9%. Subsequently, the multi-cross SVM was integrated with the ABC and ACO algorithms to obtain higher accuracies by optimizing the parameters C and Gamma. The use of ABC algorithm yielded interesting and promising results: an increase in the accuracy up to 93.33% with a convergence faster than SVM alone (the convergence reached by SVM was at the 18th iteration, however, accuracy reached by SVM-ABC hybrid model was reached at the 11th iteration). In the second part of the optimization, ACO gives more interesting results in terms of accuracy and rapidity. The accuracy was improved (94.28%) and reached only at the 6th iteration.

Since the increased rate of accuracy and decreased rate of errors are of greatly important in the classification issues, a new classification algorithm, consisting of two parts, called SVM-ABC and SVM-ACO, was introduced in the present study. These algorithms were combination of the support vector machines and the artificial bee colony algorithm or ant colony optimization. They were evaluated and compared through the criteria of precision, recall, and F-measure in addition to the accuracy and time convergence. The results showed that the proposed hybrid algorithms could achieve a higher degree of reliability in terms of performance accuracy than other known learning machines such as multi-class SVM only and MLP. This method emphasized that the combination of SVM and ABC or ACO in improving accuracy in the diagnosis and classification of oral diseases could be very efficient in expediting the treatment by specialists. The results also showed that the success of the proposed model depended on the high ability of SVM in the field of classification and the apparent capacity of ABC and ACO in the field of optimization as well as the ability to avoid early convergence and ultimately the ability to generalize a good performance. The feature of the provided learning machine enables us to use optimized and developed SVM as a certified and effective diagnostic system for oral diseases.

VII. REFERENCES

[1]. Dhanuthai, K., et al. (2018) . Oral cancer: A multicenter study. Medicina oral, patologia oral y cirugia bucal. **23**(1): p. e23.

[2]. Warnakulasuriya, S. (2009) . Causes of oral cancer–an appraisal of controversies. British dental journal. **207**(10): p. 471-475.

[3]. Alabi, R.O., et al. (2021) . Deep machine learning for oral cancer: From precise diagnosis to precision medicine. Frontiers in Oral Health: p. 97.

[4]. Al-Rawi, N., et al. (2022) . The Effectiveness of Artificial Intelligence in Detection of Oral Cancer. International Dental Journal.

[5]. Ilhan, B., P. Guneri, and P. Wilder-Smith. (2021) . The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. Oral oncology. **116**: p. 105254.

[6]. Krishna, A.B., et al. (2020) . Role of artificial intelligence in diagnostic oral pathology-A modern

approach. Journal of Oral and Maxillofacial Pathology: JOMFP. **24**(1): p. 152.

[7]. Welikala, R.A., et al. (2020). Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. IEEE Access. **8**: p. 132677-132693.

[8]. Majumder, S., et al. (2006). Multi-class classification algorithm for optical diagnosis of oral cancer. Journal of Photochemistry and Photobiology B: Biology. **85**(2): p. 109-117.

[9]. Chan, C.-H., et al. (2019). Texture-map-based branch-collaborative network for oral cancer detection. IEEE transactions on biomedical circuits and systems. **13**(4): p. 766-780.

[10]. Singh, R.K., et al. (2008). Hybrid SVM-Random Forest classication system for oral cancer screening using LIF spectra. in 2008 19th International Conference on Pattern Recognition.. IEEE.

[11]. Ray, M.T., et al. (2008). Detection of constituent layers of histological oral sub-mucous fibrosis: Images using the hybrid segmentation algorithm. Oral oncology. **44**(12): p. 1167-1171.

[12]. Lokesh, K., J. Kannabiran, and M.D. Rao.(2016). Salivary lactate dehydrogenase (LDH)-A novel technique in oral cancer detection and diagnosis. Journal of clinical and diagnostic research: JCDR. **10**(2): p. ZC34.

[13]. Ilhan, B., et al. (2020). Improving oral cancer outcomes with imaging and artificial intelligence. Journal of dental research. **99**(3): p. 241-248.

[14]. Vapnik, V.N. and V. Vapnik.(1998). Statistical learning theory. Vol. 1. Wiley New York.

[15]. Diehl, C.P. and G. Cauwenberghs.(2003). SVM incremental learning, adaptation and optimization. in Neural Networks. Proceedings of the International Joint Conference on. IEEE.

[16]. Aswathy, M. and M. Jagannath.(2021). An SVM approach towards breast cancer classification from H&E-stained histopathology images based on integrated features. Medical & biological engineering & computing. **59**(9): p. 1773-1783.

[17]. Jeyaraj, P.R. and E.R. Samuel Nadar.(2019) Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. Journal of cancer research and clinical oncology. **145**(4): p. 829-837.

[18]. Chodorowski, A., T. Gustavsson, and U. Mattsson.(2002). Support vector machines for oral lesion classification. in Biomedical Imaging. Proceedings.IEEE International Symposium onIEEE.

[19]. Krishnan, M.M.R., et al.(2009) Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis—An SVM based approach. Computers in biology and medicine. **39**(12): p. 1096-1104.

[20]. Krishnan, M.M.R., et al.(2012). Automated oral cancer identification using histopathological images: a hybrid feature extraction paradigm. Micron. **43**(2-3): p. 352-364.

[21]. Anuradha, K.(2013). Statistical feature extraction to classify oral cancers. Journal of Global Research in Computer Science. **4**(2): p. 8-12.

[22]. Anuradha, K. and K. Sankaranarayanan.(2013). Automated Diagnosis of Oral Cancer using Supervised Learning Algorithms: a Comparative study.

[23]. Mukherjee, A., et al.(2006). Performance analysis of different wavelet feature vectors in quantification of oral precancerous condition. Oral oncology. **42**(9): p. 914-928.

[24]. Venkatakrishnan, S., V. Ramalingam, and S. Palanivel. (2013). Classification of Oral Submucous Fibrosis using SVM. International Journal of Computer Applications. **78**(3).

[25]. Rosado, P., et al.(2013). Survival model in oral squamous cell carcinoma based on clinicopathological parameters, molecular markers and support vector machines. Expert systems with applications.. **40**(12): p. 4770-4776.

[26]. Li, Y., et al.(2010) . Research on the Raman spectral character and diagnostic value of squamous cell carcinoma of oral mucosa. Journal of Raman Spectroscopy. **41**(2): p. 142-147.

[27]. Chang, S.-W., et al.(2013). Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. BMC bioinformatics. **14**(1): p. 1-15.

[28]. Maghsoudi, R., A. Bagheri, and M.T. Maghsoudi .(2013). Diagnosis Prediction of Lichen Planus, Leukoplakia and Oral Squamous Cell Carcinoma by using an Intelligent System Based on Artificial Neural Networks. Journal of Dentomaxillofacial. **2**(2): p. 1-8.

[29]. KALAIARASAI, A. and M.K.M. AMANULLA.(2010). Unconscious Oral Cancer Detection using Data Mining Classification Approaches.

[30]. Chuang, L.-Y., et al.(2011). Support vector machine-based prediction for oral cancer using four snps in DNA repair genes. in Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong-Kong. Citeseer.

[31]. Naha Sharma , H.O.,(2014). Early detection and prevention of oral cancer : asociation rule mining on intevestigations. Wseas Transactions on Computers. **13**.

[32]. Donald, P.M. and R. George.(2015).Oral Atrophic Lichen Planus: Diagnosis and Management. .

[33]. Perdigão, P.F., et al. (2007) . Serotonin transporter

gene polymorphism (5-HTTLPR) in patients with oral lichen planus. archives of oral biology. **52**(9): p. 889-893.

[34]. Al-Hashimi, I., et al.(2007) Oral lichen planus and oral lichenoid lesions: diagnostic and therapeutic considerations. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology.. **103**: p. S25. e1-S25. e12.

[35]. Burket, L.W., et al..(2008). Burket's oral medicine: PMPH-USA.

[36]. Scully, C. and J. Bagan.(2009). Oral squamous cell carcinoma overview. Oral oncology. **45**(4): p. 301-308.

[37]. Maghsoudi, M.T. (2014). intelligent evaluation in diagnosis of white and red lesions in the mouth, in dental. Gilan university of medical sciences.

[38]. Cortes, C. and V. Vapnik, (1995). Support-vector networks. Machine learning. **20**(3): p. 273-297.

[39]. Wang, W., et al. (2003) Determination of the spread parameter in the Gaussian kernel for classification and regression. Vol. 55. 643-663.

[40]. Zhang, X., X. Chen, and Z. He, (2010).An ACO-based algorithm for parameter optimization of support vector machines. Expert Systems with Applications. **37**(9): p. 6618-6628.

[41]. Weston, J. and C. Watkins.(1998) Multi-class support vector machines, Citeseer.

[42]. Karaboga, D. and B. Akay.(2009) A comparative study of Artificial Bee Colony algorithm. Appl. Math. Comput.. **214**(1): p. 108-132.

[43]. Akay, B. and D. Karaboga.(2012). A modified Artificial Bee Colony algorithm for real-parameter optimization. Inf. Sci.**192**: p. 120-142.

[44]. Akay, B. and D. Karaboga.(2012) A modified artificial bee colony algorithm for real-parameter optimization. Information sciences. **192**: p. 120-142.

[45]. Dorigo, M. and C. Blum,(2005). Ant colony optimization theory). A survey. Theoretical computer science. **344**(2-3): p. 243-278.

[46]. Dorigo, M., G. Di Caro, and L.M. Gambardella.(1999). Ant algorithms for discrete optimization. Artificial life. **5**(2): p. 137-172.

[47]. Dorigo, M., V. Maniezzo, and A. Colorni.(1996). Ant system: optimization by a colony of cooperating agents. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). **26**(1): p. 29-41.

IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

ABOUT IJEAST

International Journal of Engineering Applied Science and Technology (IJEAST) is a peer-reviewed, open access journal that publishes high-quality research papers in the field of Engineering, Applied Science and Technology.

IJEAST aims to provide a platform for researchers, academicians, and professionals to share their innovative ideas, research findings, and practical experiences with the global scientific community.

FOCUS AREAS

- Engineering
- Applied Science
- Technology
- Innovation & Development
- Interdisciplinary Studies

PEER REVIEWED
All submissions are rigorously peer reviewed to ensure quality.

OPEN ACCESS
Free and unrestricted access to research for all.

GLOBAL REACH
Connecting researchers and professionals worldwide.

TIMELY PUBLICATION
We ensure a swift and efficient publication process.

For more information, visit our website
www.ijeast.com

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

editor@ijeast.com
www.ijeast.com
India

2455-2143