DEEPORCD: Detection of Oral Cancer using Deep Learning

R. Dharani¹ and S. Revathy²

¹Research Scholar, Faculty of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India
²Assoc. Prof., School of Computing, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Email : dharani.rit@gmail.com

Abstract. Oral cancer is a widespread and complex cancer with a high severity. Oral cancer is the eighth most common cancer in the world in India, with 130,000 deaths in each year. The tumor occurs in the salivary glands, tonsils, as well as in the neck, face and mouth. There are various diagnostic methods for oral cancer, such as a biopsy, in which a small tissue sample is taken from a part of the body and tested under a microscope also some screening methods. But the downside is that cannot clearly identify cancer cells and cannot classify the number of cells affected by cancer, so in this work cancer cells will find and classify that affected in the oral area through digital processing technology. The use of advanced technologies and an in-depth learning algorithm are possible for early detection and classification. This work uses three characteristics-extracting techniques such as the bag histogram of oriented gradients, wavelet features and the Zernike Moment. Once retrieving the texture characteristics, the fuzzy particle swarm optimization algorithm (FPSO) is applied to choose the best characteristic. Finally, these characteristics were classified using the Convolution Neural Network (CNN) classifier. For comparison of the efficiency of the proposed method, Recall Rate, Classification Accuracy, Precision Rate, and Error Rate. Evaluation outcomes demonstrated that the combination of ABC, FPSO and CNN performs better in the detection of oral cancer.

1. Introduction

Oral cancer has been reported to be uncontrollable in the enlargement of cells that provide with injured adjacent tissues [1]. Oral cancer detects less lifeless cells in the oral tissue early in the development of oral cancer known as ulcers. When metabolism means that dead cells are present in remote areas of the pretentious region or inside of the body. There are many different categories of cancer, with 90% of crab cells being medically referred to as OSCC (Oral Squamous Cell Carcinomas) [2]. Biological models along with clinical forms of related and lesion-free tumor models be able to identified in different location of the body through appearance models as well as stereotypes without staining. Machine learning approaches were used to forecast dissimilar biological models for OSCC, which would categorize non-cancerous along with malignant samples that were afterwards evaluated for the oral cancer stage [3]. The predictor will determine the accuracy of the interaction using three justification test kits as well as different stages of cancer. Sampling with justification of samples can forecast different tumor volume as well as the appearance of ulcers in the tissues, helping to predict dissimilar phases of oral cancer [4]. The major goal of the current procedure is to develop new tools for predicting the stage of growth of oral cancer tumors. Oral cancer occurs in regions such as the front of the tongue, the upper as well as lower parts of the mouth, the inside of the cheeks along with...
lips, the gums, as well as the region at the back of the wisdom teeth. Oral cancer symptoms are: The majority signs of cancer are inflammation or ulcers that do not cure as well as can reason for pain otherwise bleeding [5]. Risk aspects for oral cancer include a number of activities, for example smoking as well as alcohol. These two habits are measured most important risk aspects for oral cancer. Eating worms is so common in India that it also affects the inside of the gums [6]. The remaining risk aspects for instance Human Papillomavirus (HPV), sex and age. The major contributions of this manuscript are shortened below:

(i) CNN classifier is used to detect oral cancer disease. CNN classifier is trained through UCI Machine Learning Repository Data Set.

(ii) For extracting features, the combination texture, color and shape features are applied.

(iii) For segmenting oral cancer region Artificial Bee Colony (ABC) is used.

(iv) Fuzzy Partial Swam Optimization (FPSO) approach is applied for feature selection.

The manuscript of this document is prepared as follows: Section 2 examines a few of the related up to date literature. Section 3 illustrates a exhaustive explanation of the proposed method. Section 4 presents the experimental results, which include the general results of the CNN classification results comparison. Finally, a conclusion is given in Section 5.

2. Related Work

Fatihah Mohd et al. The goal of the researchers was to forecast the early stages of oral cancer through less correct output by Naïve Bayes, Multilayer Perceptron, KNearest Neighbors and Support Vector Machine methods leading to the stage of oral cancer. Oral cavities and analysis increase the accuracy of the classification. [7]. Ahmad LG et al. The aim of the researcher is to create a model for clinicians. Using tree-based decision-making methods, artificial neural network vector maintenance methods, and high-accuracy analysis of DATA, NN, and HDM. With the highest accuracy and the lowest error level, the ADMS rating representation is most excellent suited for discover of breast cancer reappearance. Evaluated to ANN as well as DT, the output show that SVM are the superior approach for forecasting [8].

The goal of Harikumar Rajaguru and Sunil Kumar Prabhakar [9] is to evaluate the detection accuracy of TNM phase systems employing the Multi-Layer Perceptron (MLP) as well as Gaussian Mi Mixed Models. The comparison of the two ranking groups here gave better results as the average accuracy for the stage. Ultraviolet scanning machines (LMMs) are used as post-data storage devices for the analysis of oral cancer, as well as the efficacy of LMS classifiers is evaluated with the efficiency of SMMs and MLPs. Amy F. Z. B. Al. SVM classification method was used to look for OSCC tumors by examining patient expression and RNA extraction, as well as microscopic analysis, showing normal prognosis of OSCC tissue [10]. Marc Aubreville et al. The aim is to evaluate new automation approaches for OSCC diagnostics using in-depth training and CNN methods on clear imaging. Here CNN's approach is to search for quotes, images, training, data and classification[11].

Shreyansh A et al. Data sets containing 251 equator X-rays were used, which were later subdivided into tests and training of experimental data such as in-depth study of ANA, transfer studies, along with CNN. Thus, they attained 88.46% accuracy [12]. Martin Halicek et al. Researchers tested the OSCC model through CNN to discover cancer. C. N. Xi achieved 80% accuracy in discovery of cancer [13]. Ramzi Ben Ali et al. The goal of the researchers was to categorize dental X-rays of damaged and normal teeth as well as to create innovative models for X-ray problems employing deep neurological technology [14]. Konstantina Kuru et al. The scholars compared the machine learning programs to dissimilar category of cancer forecast as well as prognosis.

The learning employed hereditary information to model through machine learning approaches as well as decision tree through methods for the selection and classification of traits [15]. The goal of researchers Wafaia K. Shams and Zaw Z. Hitike was to forecast the progression of oral cancer in OPL patients in which machine learning approaches using genetic representation were employed. The researchers used SVM, MLP, minimally invasive routine (RLS) and deep neural network to study the progression of oral cancer in patients with OPL records [16]. The author [17] discusses the significance of cancer discovery process. They presented the different ways discovery process on dissimilar cancer topics as well as described the way of artificial
techniques help improve cancer detection. They executed the categorization of brain tumors using standardized sample images of UCI study data sets. From the result, we have recognized cancer problems in the socio-economic population. They were 87.2% accurate in identifying brain cancer through deep neural networks. In Phillips et al. A study of the different practices of in-depth training to the application of different teaching methods used to classify medical images. In addition, computer-based computer techniques that make classification effective are discussed. The use of hypothetical imaging for cancer discovery compared to traditional techniques is talked about in the suggestion by Duke Kumama Al. [18] They evaluated several methods for cancer detection as well as make clear the benefits of symptomatic simulations. Therefore, HSI can be used for classification professionals. They perform the categorization employing a vector support machine using a self-mapping structure. In the work of Yuan et al. [19] and Dul Al. [20] Shows IT. P. IT. Over other imaging methods such as MRI and CT. They also use the detection of cerebrovascular disease through algorithms for in-depth training in their practice. Semi-automated methods for classifying cancer using in-depth learning algorithms are discussed in Wang L. [21]. They used ordered power machines for categorization in the effort of Kalantari et al. [22] amplification the use of hyperspectral to detect lung cancer. They employed CNN to divide into pictures. Against this background, the lack of a fully autonomous design of the cancer detection system through in-depth study techniques has been established. Most techniques require advanced system configuration, which leads to high system operating costs. This work conquer all the obstacles explained in Yuan et al. [22] with given that a regression formation of the neural network of computerized cancer discovery in existing medical imaging employing this new in-depth learning method

2.1 Observation

From the related work, it can be seen that the previously used methods and materials are mainly included for the detection of the presence of cancer, the categorization of cancer as well as the assessment of machine learning techniques. Therefore, oral cancer placement is an important mission in diagnosing oral cancer. This effort is not performed by some scholar, which is the most necessary mission in the study of prediction along with the curing of cancer patients for practitioners. Therefore, the current study focuses on the application of a variety of machine learning techniques that focus on effective stage analysis in the development of oral cancer

3. System Methodology

The overall flow of this work is shown in Figure 1. The upper side of the figure shows the training procedure. The sample database contains all training oral finally cancer images along with corresponding cancer types. After that, extract features from the oral cancer parts. The classifier model is trained by using the extracted features and saved for use in the future. The lower side of the Figure 1 shows the testing procedure.

During online process, the sample is undergone pre-processing techniques such as ABF to de-noise the sample and HE to improve the clarity of the sample. Pre-processing is followed by segmentation employing ABC approach. Features such as texture and intensity are retrieved by combining all the two strategies to calculate the unique features. After retrieving the features, an appropriate feature is selected to reduce the execution time. Finally, the disease and its impact are detected from the selected features. This research implements three feature extraction strategies such as the features based on wavelet transform, histogram of oriented gradients, Zernike Moment. Fuzzy Particle Swarm Optimization (FPSO) is employed to choose the most excellent among the features. Lastly these features are categorized by CNN method. In the rest of the section the proposed technique are described in detailed manner Sections, subsections and subsections.
3.1 Pre-processing

The first step before all process to be the contrast of the incoming CT scan was optimized using the histogram (HE) alignment technique. His Excellency was applied to adjust the intensity of the image to improve the contrast as shown in Equation (1). Let the scanned image represent $I_x$ from the $I_y$ matrix with pixel intensities in the range 0 to 256. $N$ specify the normal histogram of Figure I for the existing intensity.

$$I_y = \frac{\text{Number of pixels with available intensity } n}{\text{Total number of pixels}}$$  \hspace{1cm} (1)

Where $n = 0, 1, ..., 256$.

The improved image is revealed in Figures. 2.

**Figure 2.** Histogram Resultant Image  
2(a). Input Oral Image 2(b). Result of Histogram Equalized

Subsequently Adaptive Bilateral Filter (ABF) [23] is employed which is mentioned in Eq.(2). In this work ABF are applied on the oral image to increase the noise.
If $\nabla_r$ and $\delta$ are fixed, the ABF will degrade to a normal two-way filter. Low frequency filter, constant frequency is adopted for ABF. In addition, the ABF sharpens the image by increasing the edge gradient. $\Delta$ is calculated in ABF using Equation (3).

$$ABF_{x_0,y_0} = \sum_{a=-N}^{N} \sum_{b=-N}^{N} \exp \left( -\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2} \right) \times \exp(-\frac{(G[m,n] - G[x_0,y_0] - \delta[x_0,y_0])^2}{2\sigma^2})$$  \hspace{1cm} (2)

The size of the input picture window is stand for $(2W + 1) \times (2W + 1)$. All pixel is mentioned as $\beta_{(x_0, y_0)}$ through a center point $[x_0, y_0]$. Let MAXIMUM and MINIMUM represent each data value retrieval operation. Efficiency of FFAs with fixed douchean filter and effective range filter. Here $\nabla_{(d)} = 1$ is fixed and $\nabla_{(r)}$ modifys the value as exposed in the Figure. (3).

$$\delta[x_0,y_0] = \begin{cases} \text{MAXIMUM} (\beta_{x_0,y_0}) - G[x_0,y_0], & \text{if } \Omega_{x_0,y_0} > 0 \\ \text{MINIMUM} (\beta_{x_0,y_0}) - G[x_0,y_0], & \text{if } \Omega_{x_0,y_0} < 0 \\ 0, & \text{if } \Omega_{x_0,y_0} = 0 \end{cases}$$ \hspace{1cm} (3)

Figure 3. (a) Result of Histogram Equalized (b). Noise Removed Image

3.2 Segmentation

In pre-processed images, the segment detects objects or boundaries that help to obtain required portions of the images and divide them into areas to pinpoint the necessary data. During the classification, distinguishing the affected cells from a preoperative CT image is necessary and the above processes were first classified with Artificial Bee Colony (ABC) partition method as exposed in Figure 4. Methods for instance, Ant Colony, K-Means along with FCM have been employed to partition the oral cancer. The FCM algorithm has a longer time to compute. It is responsive to low speed and local noise. The K- means algorithm means that it is difficult to predict the quantity of groups [24]. In the Ant Colony method, the likelihood allocation alters through iteration along with is self-governing of the previous result to discover the most excellent answer. It obtains a long time for meeting to be unsure. These shortcomings were defeated through the ABC. ABC is easy, stretchy as well as strong. Its functioning is simple. It has less control parameters to investigate local results as well as deal with target values.

Figure 4(a). Noise Filtered Image 4(b). Result of ABC
3.3 Feature Extraction

Texture and intensity are the features resultant from the segmentation, LBP and wavelet techniques are employed to retrieve texture. This work excerpts 140 features which includes 96 Zernike moment, 26 wavelet and 18 CVH features.

Wavelet features
Wavelets are significant as well as are generally employed for describing features for texture retrieval. Wavelet's performance is effective in capturing fundamental frequency information and multi-resolution features. The wavelets signal continuously passes through a pair of low and high pass filters, an analytical filter that produces the conversion coefficient. Here H-level decomposition occurs, which results in different 3H + 1 bands. The LH frequency subgroup was used to compose the HL image decomposition was used to generate the HH horizontal image details used to compose the diagonal image details.

CVH features
CVH means histogram of CT value. A histogram of CT value is calculated for each return on investment. The number of containers in the histogram was determined experimentally. In fact, different CVHs with different tank numbers are obtained. Each CVH was tested for classification according to the K-NN classifier and the corresponding card was calculated. The largest number of bins that carry luggage is then taken.

Zernike moment features
The Zernike moments a descriptor of large forms in retrieving features. At this point, the pre-processed image input is placed according to the first histogram, which illustrates the large size of the preview. Zernike's time depends on the transformation along with extending of the tables in the ROI. Namely, the moments of the Zernike on two related images that are not extended as well as interpreted are dissimilar. Two procedures are used to solve Zernike dependency troubles. The centroid of every image mass is converted to the center of the related ROI. This procedure gets rid of the dependence of Zernike time on object transformation.

3.4 Feature Selection

To attain high-quality categorization outputs, various kinds of features are used at the similar time. Because various kinds of features may have additional data, this may lead to better classification efficiency by selecting discriminant features from different feature spaces. The benefit of feature assortment is to define the meaning of the unique feature set. Fuzzy Particle Swam Optimization (FPSO) is employed for choosing features.

Fuzzy Particle Swarm Optimization (FPSO)
To attain better outputs, various types of features are adopted simultaneously and FPSO is applied to choose features. FPSO has a knowledge base that combines information provided by experts in the form of ambiguous rules of language management, fuzzification interfaces, which effectively transform pure data into vague sets of input systems that use them. It goes along with the knowledge base. Reflecting through the filtering method and the conversion interface, which translates the fuzzy controls, thus obtains the actual control force from illusion technique.
Among various algorithms such as, DE, PSO,GA, PSO is highly appreciable as it has got prominent features such as fewer cost and guaranteed optimal solution. The influence employed by the particle over others around the locality is overwhelmed through a fuzzy variable.

3.5 Classification
3.5.1 Bag Classifier
The Bag classifier is largely employed for information retrieval (IR). Lately, it is also employed for computer application. In computer application software, it is to be validated to classify images in relation to datasets. In this classification, every figure is conducted as a file as well as is represented symbolically with a histogram created by the training image. Image classes are chosen and labelled first. At this time the classifier decides the class name of the training class that should be the label of the test picture class.

3.5.2 Naive Bayes Classifier
The Naive Bayes classifier relies on the likelihood form and distributes exact classes with the most probabilities next to the characteristic vectors. The subsequent likelihood $P(C_r/FV)$ of a particular group $C_r$ is measured through the FV characteristic vector. FV is calculated with the Basis Theorem. It is shown Equation (4)

$$P(C_r/FV) = \frac{P(FV|C_r)P(C_r)}{P(FV)}$$ (4)

This [24] method is depends on the Bayesian theory as well as is mainly suitable at what time the size of the input is large. Even though its effortlessness, it be able to frequently do better than the classification method.

3.5.3 K-NN Classifier
The K-Nearest Neighbour is an asymmetric scheme that is employed for classification as well as regression. The entrances have, for example, the nearest k of training in the characteristic space. In the K-NN rankings, the result is a group member. This object is recognised with the widely held of its neighbours, the most commonly classified object surrounded by its closest neighbours. If $K = 1$, the object is allocated to the group of the closest neighbour.

3.5.4 Adaboost Classifier
Adaboost classifier is being educated on additional learning instances. It gives high-quality adaptation to training instances, generating little training errors. It's pretty easy. It enhances classification performance through mixing poor prognosis with each other.

3.5.5 SVM Classifier
SVM is a dual categorization process that accepts two labeled classes and creates sample files to classify new unmarked/labeled data in one of two groups. It grouped stuffs with related feature records into clusters. SVM create a twin zone that extends the boundary among the negative as well as positive patterns. At last, categorization is done with deciding on the rate of the linear mixture of characteristics.

SVM is educated by transmitting recognized information through a formerly recognized conclusion value through producing a limited set of training. It is from this set of training that SVM derives its cleverness to categorize fuzzy information. In the SVM for the two categorization problems, the input information is recorded to a larger extent through the RBF kernel. Here, a linear classification of the high plane is performed in this changed space, using those model vectors close to the solution boundary.

3.5.6 ELM Classifier
ELM is a forward neural network for classification and functional training of single-layer or multi-layer hidden nodes where the constraints of the hidden node do not require to be defined.
In most cases, the initial weight of the hidden node is learned step by step, which is important to study the linear pattern as shown in Figure 5. The ELM classification tool can be shown below.

\[ q = W^2\alpha(W^1x) \]

where \( W_2 \) is the weight value among the hidden and output layer. \( \beta \) offers weight among the hidden layer along with the outcome. This is multiplied by \( q(x) \) to get the realization function.

ELM has limited capacity for optimization. Unrelated functions related to the sequence that fulfill the ELM approximation theorem can be used as output functions in hidden layers. The core matrix \( (K) \) is shown below.

\[ K = k(x_m, x_n) = q(x_m)q(x_n) = QQ^T \]

Where \( Q \) is the resultant matrix of the hidden layer and \( k(x_m, x_n) \) is the function of the nucleus of the hidden neuron. The ELM result is defined as Equation (5).

\[ F(x) = [k(x,x_1) ... (x,x_M)]\left[ \frac{1}{R} + K \right]^{-1}E \quad (5) \]

Where \( M \) is the sample number, \( R \) is the control factor, and \( E \) is the expected output format. Compared to other aristocratic classes, ELM is easy to use and fast learning process because it has a predefined network architecture. No need to set parameters manually. He does not spend much time setting up and training machine learning. The efficiency of the LMS collection is better than other classifications. It also applies to all non-linear activities, differential disruption, non-differential functions. It is also applicable for completely complex functions.

**CNN Classifier**

The structure of Convolutional Neural Networks is similar to the structure of conventional neural networks. Each neural network is composed of neurons with training weights, biased values that are of studying weight along with bias. Generally a neuron obtains input, computes its point product moreover the behaviour follows a non-linear characteristic. Each Convolutional Neuron Network is composed of single or additional convolutional layers as well as pooling layers or sub sampling layers. Generally convolutional neuron network is employed to sort the types of cancer cells and its severity. Down sampling is accomplished by the employment of pooling layer of CNN. The time consumption is lessened by diminishing the features extraction in convolution layers and this is depicted in Figure 6.
9

Figure 6. Oral Cancer Detection using ICNN

The two types of pooling layers of CNN are Maximum and Average pooling in CNN and when united the biggest pixel values are calculated in max whereas in average pooling, the mean values are in use to explanation. The response of this layer is served as input to the subsequent layer of convolutional neural network. Procedure for CNN based Classification is exposed in below algorithm.

1. Apply convolution filter in first layer
2. Filter sensitivity is reduced by smoothing the convolution filter
3. Signal transfer from one layer to subsequent layer is examined by the process layer
4. Restrict training time by Rectified Linear Unit (RELU)
5. The neurons in the flow layer are linked to every neuron in the subsequent layer
6. During offline processing, at the end add a layer of loss to suggest neural networks

4. RESULTS & DISCUSSIONS

4.1 Data set used

4.1.1 UCI Machine Learning Repository Data Set

UCI data determined for oral CT scans were considered to evaluate the efficacy of CNN. And various classes for classifying oral cancer. The UCI dataset is the largest publicly available library for cancer screening, with a 1018 CT scan of the oral cavity. Here the lung nodes appear in many parts of the CT scan. Figure 7 shows an example of an oral CT scan.
4.2 Performance Metrics used

*Sensitivity* ($Sn$) is termed as the division of malignant nodules forecasted accurately as exposed in Equation (6).

$$Sn = \frac{T_P}{T_P + F_{nN}} \quad (6)$$

*Specificity* ($Sp$) is termed as the division of benign nodules forecasted accurately as exposed in Equation (7),

$$Sp = \frac{T_{nN}}{T_{nN} + F_{nP}} \quad (7)$$

*Classification Accuracy (CA)*

$$CA = \frac{T_P + T_{nN}}{T_P + T_{nN} + F_{nP} + F_{nN}} \quad (8)$$

In Equation (8), where $T_P$ denotes the quantity of malignant nodules accurately forecasted. $F_{nN}$ denotes the quantity of malignant nodules incorrectly forecasted. $T_{nN}$ denotes the quantity of benign nodules accurately forecasted. $F_{nP}$ denotes the quantity of benign nodules incorrectly forecasted.

*Error Rate*

The error rate is a measurement system that measures the number of diseases that are misdiagnosed in a given image.

$$Error \ Rate = \frac{Number \ of \ images \ categorized \ incorrectly}{Total \ number \ of \ images}$$

4.3 Experimental Analysis

4.3.1 Trial No 1: Examination of Feature Derivative Techniques

To analyze the performance of this feature derivative scheme, it is compared with different techniques using the operating indicators such as Sensitivity, Error Rate, Accuracy and Specificity. The outputs of these indicators are tabulated in Table 1.

**Table 1. Examination of Feature Derivative Techniques**

| Technique | Sensitivity | Specificity | Accuracy | Error Rate |
|-----------|-------------|-------------|----------|------------|
| Technique A | 0.85 | 0.92 | 0.89 | 0.15 |
| Technique B | 0.88 | 0.89 | 0.87 | 0.17 |
| Technique C | 0.90 | 0.91 | 0.90 | 0.10 |

*Figure 7. Examples of UCI Machine Learning Repository Data Set*
The influences of feature derivative schemes that are employed in this test are successfully assessed. Table 1 illustrates that the highest SEN is 98.154 is found for All feature and it is more powerful as it is the highest value when compared with other techniques.

### 4.3.2 Trial No 2: Examination of Oral Cancer Partition Approaches

To analyze the performance of this feature derivative scheme, it is compared with different techniques using the operating indicators such as Sensitivity, Error Rate, Accuracy and Specificity. The outputs of these indicators are tabulated in Table 2.

| Examination Parameters | ACC  | SEN  | SPEC | ERR  |
|------------------------|------|------|------|------|
| Wavelet                | 82.58| 91.4 | 93.91| 17.42|
| CVH                    | 81.12| 90.51| 92.85| 18.88|
| Zernike                | 82.71| 89.82| 93.8 | 17.29|
| All                    | 91.651| 92.451| 92.321| 8.349|

The influences of feature derivative schemes that are employed in this test are successfully assessed. Table 2 illustrates that the highest SEN is 98.752 is found for ABC and it is more powerful as it is the highest value when compared with other techniques.

### 4.3.3 Trial No 3: Examination of Oral Cancer Classification Approaches

To analyze the performance of this cancer classification scheme, it is compared with different techniques using the operating indicators such as Sensitivity, Error Rate, Accuracy and Specificity. The outputs of these indicators are tabulated in Table 3.
### Table 3. Examination of Oral Cancer Classification Techniques through UCI

| Examination Parameters | ACC  | SEN  | SPEC | ERR  |
|------------------------|------|------|------|------|
| SVM                    | 94.23| 95.61| 94.91| 5.77 |
| Bagging                | 89.26| 90.75| 89.95| 10.74|
| Naive Bayes            | 85.71| 86.63| 86.42| 14.29|
| KNN                    | 84.23| 85.21| 84.96| 15.77|
| AdaBoost               | 91.76| 92.82| 92.43| 8.24 |
| IELM                   | 97.14| 98.39| 97.87| 2.86 |
| CNN                    | 98.2 | 98.45| 98.78| 1.8  |

The influences of feature derivative schemes that are employed in this test are successfully assessed. Table 3 illustrates that the highest SEN is 98.154 is found in CNN and it is more powerful as it is the highest value when compared with other techniques.

### 5. Conclusion

In this paper the oral cancer detection techniques of CT images are taken for the comparison. The comparison is done on three steps of the proposed work such as segmentation, feature extraction and classification. This work utilized six supervised machine learning methods, namely, SVM, Bagging, Naive Bayes, KNN, AdaBoost, CNN and IELM for comparison process. CNN classifiers can obtain an overall accuracy 97.21%. So the results indicate that CNN approach performs well for oral cancer detection of CT images. The sensitivity, specificity and error rate are commonly used to measure how well the method can identify the oral cancer. From these results, it is well know that CNN performs best than the other approach. As well as the similar experiments are also conducted for segmentation approaches. The result indicated that the ABC performs best in the segmentation step because of their high detection accuracy value than the other approaches. So these three approaches such as ABC, FPSO and CNN are used for development of new automated technique of oral cancer detection in CT images.

### References

[1] Oral cancer, https://en.wikipedia.org/wiki/Oral_cancer
[2] Oral Leukplakia LM328”https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM65276
[3] Oral cancer, https://www.webmd.com/oral-health/guide/oral-cancer#1
[4] Head-and-Neck-squamous-cell carcinoma, https://en.wikipedia.org/wiki/Head_and_neck_squamous_cell_carcinoma, Wikipedia (2018).

[5] Squamous-cell carcinoma, cancer, https://en.wikipedia.org/wiki/Head_and_neck_squamous-cell carcinoma

[6] Staging of Lip and Oral Cavity Cancer, https://www.ahns.info/patient-information/

[7] Fatihah Mohd, Noor Maizura Mohamad Noor, Zainab Abu Bakar, Zainul Ahmad Rajion, Analysis of Oral Cancer Prediction using Features Selection with Machine Learning, ICIT 2015 The 7th International Conference on Information Technology.

[8] Ahmad LG*, Eshlaghy AT, Poorebrahimi A, Ebrahimi M and Razavi AR, Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence, Health & Medical Informatics (2013), 2157-7420.

[9] Harikumar Rajaguru and Sunil Kumar Prabhakar, Performance Comparison of Oral Cancer Classification with Gaussian Mixture Measures and Multi Layer Perceptron, The 16th International Conference on Biomedical Engineering (2017) 123-129.

[10] Amy F. Ziober, Kirtesh R. Patel, Faizan Alawi, Phyllis Gimotty, A. Randall S. Weber, Michael M. Feldman, Ara A. Chalian, Gregory S. Weinstein, Jennifer Hunt, and Barry L. Ziober, Identification of a Gene Signature for Rapid Screening of Oral Squamous Cell Carcinoma, American Association for Cancer Research (2018).

[11] Marc Aubreville, Christian Knipfer, Nicolai Oetter, Christian Jaremenko, Erik Rodner, Joachim Denzler, Christopher Bohr, Helmut Neumann, Florian Stelzle, & Andreas Maier, Automatic Classification of Cancerous Tissue in Laserendomicroscopy Images of the Oral Cavity using Deep Learning, SCIENTIFIC Reports, (2017), 7: 11979.

[12] Symptoms of oral cancer, https://www.ahns.info/resources/education/patient_education/oralcavity

[13] Martin Halicek, Guolan Lu, James V. Little, Xu Wang, Mihir Patel, Christopher C. Griffith, Mark W. El-Deiry, Amy Y. Chen, Baowei Fei, Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging; J. Biomed. Opt. 22(6), 060503 (2017), doi: 10.1117/1.JBO.22.6.060503.

[14] Randomforest, https://towardsdatascience.com/therandom-forest-algorithm-d457d499ffcd

[15] Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, Dimitrios I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Computational and structural biotechnology journal, (2015), 18-17.

[16] Wafaa K. Shams and Zaw Z. Htike, Oral Cancer Prediction Using Gene Expression Profiling and Machine Learning, International Journal of Applied Engineering Research, (2017), 0973-4562.

[17] Heba M, El-Dahshan ESA, El-Horbaty ESM, Abdel-Badeeh M (2018) Classification using deep learning neural networks for brain tumors. Future Comput Inf J 3(1):68–71.

[18] Deepak Kumar J, Surendra Bilouhan D, Kumar RC (2018) An approach for hyperspectral image classification by optimizing SVM using self-organizing map. J Comput Sci 25(1):252–259.

[19] Yuan Y, Lin J, Wang Q (2015) Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization. IEEE Trans Cybern 46(12):2966–2977.

[20] Dou Q, Chen H, Yu L et al (2016) Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 35(5):1182–1195.

[21] Wang C, Gong L, Yu Q et al (2017) DLAS: a scalable deep learning accelerator unit on FPGA. IEEE Trans Computer Aided Des Integr Circuits Syst 36(3):513–517.

[22] Kalantari N, Ramamoorthi R (2017) Deep high dynamic range imaging of dynamic scenes. ACM Trans Graph 36(4):1–12.

[23] Buyue Zhang, and Jan P. Allebach, 2008, ‘Adaptive Bilateral Filter for Sharpness Enhancement and Noise Removal’, IEEE Transactions on Image processing vol. 17, no. 5.

[24] Dong-ping Tian and Nai-qian Li, 2009, ‘Fuzzy Particle Swarm Optimization Algorithm’, IEEE, International Joint Conference on Artificial Intelligence, pp. 263 – 267.
[25] A. Meenakshi and S. Revathy, "An Efficient Model for Predicting Brain Tumor using Deep Learning Techniques," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 1000-1007, doi:10.1109/ICCES48766.2020.9138029.

[26] G. Shekar, S. Revathy and E. K. Goud, "Malaria Detection using Deep Learning," 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, 2020, pp. 746-750, doi:10.1109/ICOE I48184.2020.9143023.

[27] Jeyaramya V, Navarajan J, Prathipa R, Ashok kumar L, ‘Detection of melanoma skin cancer using digital camera images’ ARPN Journal of Engineering and Applied Sciences, 10(7), pp. 3082-3085.