Diagnosis Of Oral Squamous Cell Carcinoma Using Machine Learning

Muhammad AsadIqbal, Khalid Masood, AnasRiaz, AneelaMehmood, Amina Atta

ABSTRACT: Malignancy (Cancer) is a disease wherein a wild development of cells happens that can likewise spread into the encompassing tissues. Oral malignancy is characterized as the advancement of destructive cells in any part of the mouth which include lips, tongue, cheeks, floor of the mouth, hard and delicate palates, sinuses, and oropharynx (the part at the rear of the mouth). In some cases, it is mentioned as Oral malignant growth and at different occasions may likewise as Oral cavity disease, however it is something very similar. There is a wide range of types of Oral Cancers and patients with Oral tumors must go through an extensive diagnostic system to identify the type of cancer along with its stage to develop a personalized and customized treatment plan according to the patient's need. The focus of this research work is squamous cell carcinoma by using biopsy images. Through this work, the oral cavity cancer detection can be easier and more accurate especially in far flung areas where the enough medical staff is not available.. The dataset contains two classes (normal and malignant). The classification model employs a number of machine learning algorithms such as decision tree and support vector machine and an overall accuracy of 98% is achieved.

Keywords: Oral Cancer, Mouth cancer, Machine learning, Classification, Medical science

1. INTRODUCTION

Since the evolution of mankind, human being is always curious about why and how things happen. Why is something the way it is, the question still arises. Also, humans have been trying to invent and discover new things that are better than others in every domain. With the passage of time, humans have revolutionized the medical procedures for diagnostic purposes of various diseases and in finding the different cures to those diseases. As there is a tremendous advancement in information technology, so the medical evolution is also happening at a greater and fastest pace. Scientists are facing the challenges of emerging new diseases and infections every day and to tackle this problem they are inventing new medical procedures for the diagnostics purposes along with the cure. In the multi-disciplinary medicinal domain, the disease diagnostical procedure varies a lot from physical examination to typical instruments to digital equipment's to computerized machines to online diagnostics and all built for same purpose and based on the same foundation but varies in procedure, working and construction.

Mouth Cancer:

Mouth cancer is the 6th most affected cancer in the world, almost every year 400,000 patients attacked by this disease [1]. Mouth cancer is one of the most common cancers globally, being closely affected due to smoking, alcohol drinking, chewing tobacco, and betel quid [2]. This disease has the main effect twenty-year-old young population but its most extensive effect is found at the age of forty all over the world. The ratio of mouth cancer is higher in men than women because the use of betel quid, tobacco and smoking is more common amongst men as compared to women. There are mainly four stages of mouth cancer. Most commonly, patients are diagnosed with this problem in the third and the fourth stages and until that time the disease is mostly not curable. In all types of mouth cancer, 89% of squamous cell carcinoma is involved which is originated in the tissues of mouth [3]. Treatment options that are available after the detection of mouth cancer include surgery, radiation, and chemotherapy, the mortality rate associated with mouth cancer has no improvement in the last 40 years [4]. The point of great concern is that there is no single computational software or any other method to detect the mouth cancer in early stages. According to the Global cancer observatory if we can detect any cancer in early stages, we can get rid of it. In this research, we are going to propose an automated algorithm to detect the cancer in earlier stages.

Malignant Oral Cancer Types:

1. Squamous cell carcinoma: It is stated that more than 90% of Oral cavity cancers are squamous cell carcinomas. Mouth and throat are lined with squamous cells and abnormality in these cells lead to this type of cancer.
2. Verrucous carcinoma: 5% of oral cancers fall into this category. Basically, it is a slow-growing cancer of squamous cells.
3. Minor salivary gland carcinomas: As name indicates it is the cancer of minor salivary glands which are present throughout themouth and throat lining. There are further types which include adenoid cystic carcinoma, mucoepidermoid carcinoma and polymorphous low-grade adenocarcinoma.
4. Lymphoma: This type involves oral cavity related lymph tissues such as tonsils and base of the tongue.

CAUSE OF ORAL CAVITY CANCERS:

1. Cancer develops because of mutations in cell's DNA in any part of the body. Similarly, Oral cancer develops when cells of lips or mouth (anywhere in oral cavity) undergoes mutation because of which cells continue to grow and divide and thus leads to abnormal accumulation of these cells in the form of a tumor inside the mouth. It is not clear yet that what leads to this mutation of the mouth cells, but doctors and sci entist have discovered through extensive research over a period of years that there are certain risk factors that could increase the chances of getting oral cancer.
2. These risk factors may include but are not limited to:
3. Use of tobacco in any form, that is, snuffing, chewing, having cigarettes, pipes, cigars etc.
4. Heavy consumption of alcohol or such drinks
5. Sometimes exposure to sunrays
6. When a person develops low immunity due to any reason

7. Could be due to a sexually transmitted viral infection that is human papilloma virus (HPV)
8. Age could also be one of the factors as it is observed that it occurs mostly in people over age of 40.
9. Family history of cancer

SYMPTOMS:

1. If it is just lip cancers the symptoms could be:
2. Continuous lip pain
3. Lip Numbness
4. Lip could be thickened
5. Lip sore that would not heal easily
6. Any kind of patch on lip
7. If it is in mouth the symptoms could be:
8. Persistent pain in jaw or tongue
9. Any lump or growth in the mouth
10. Difficulty or pain in swallowing
11. Difficulty or pain in chewing
12. Difficulty or pain in moving tongue
13. Difficulty or pain in moving jaw
14. Any colored patch formation inside the mouth that Is on tongue, gums, lining of mouth or tonsils.
15. Hoarseness or change in voice
16. Chronic sore throat
17. Sudden weight loss

TREATMENT:

Treatment of Oral cavity cancer depends upon the stage of the cancer. Following are different type of treatments used, the doctor could decide depending on the stage and severity that if only one of the treatments is enough or to go for the combination of two or more:

1. Surgery
2. Radiation therapy
3. Chemotherapy
4. Targeted drug therapy
5. Immunotherapy

PREVENTION OR PRECAUTION:

There is no proven method or way to prevent cancer but the risk of getting mouth cancer could be reduced by taking following measures:

1. Do not use tobacco in any form and if somebody is already doing, should stop.
2. Do not drink alcohol.
3. Avoid long sun exposures as much as possible.
4. Develop healthy habits and routine to have a strong immunity.
5. Visit Dentists regularly.

Problem Statement:

To solve the problem of detection of mouth cancer in earlier stages, multiple techniques and algorithms through computational method are used by many scientists in the past but there are still some technical gaps present which create room for other IT experts to work on in order to solve this problem. Patient as well as squamous cell carcinoma medical experts need to know about the suitable model for its diagnosis based on the related symptoms. The design and formulation of such an automated diagnostic model is proposed in my research.

Literature Review

In Jiang, C. F., C. Y. Wang, and C. P. Chiang. "Oral cancer detection in a fluorescent image by color image fusion." [6] The methodology used in this paper is ALA Fluorescent image acquisition method and Collar Image processing to reveal the cancer area which is used to identify oral malignancy for specialists to detect cancer.

Anuradha.K.,&Sankaranarayanan, K [7]. Describes in this paper the Detection of oral cancer based on the marker-controlled watershed algorithm. The paper talks about oral cancer detection using a cancer-based marker using the watershed algorithm. The technique used in this paper is Orthopantomogram.

M.A. Aswathy, M. Jagannath(Informatics in Medicine Unlocked) [8]. The modern medical science has several advanced methodologies and techniques for the identification of breast cancer. Digital pathology is one of the emerging trends in modern medicine.LingyanFengaYongChenb(graphene functionalized electrochemical aptasensor for selective label-free detection of cancer cells 2017) [9].He used an electrochemical sensor that can sense label-free cancer cell detection by using the first clinical trial II used aptamer AS1411 and functionalized graphene. By the use of AS1411 high binding affinity and specificity to the over expressed nucleolin on the cancer cell area, our developed electrochemical aptasensor can differentiate between affected cells and normal ones and detect as low as one thousand cells. By using DNA hybridization technique, this E-DNA sensor can be regenerated and is reusable for cancer cell detection.WeihongTan*§LingwenZeng(Aptamer−Nanoparticle Strip Biosensor for SensitiveDetection of Cancer Cells) [10] Used an aptamer−nanoparticle strip biosensor (ANSB) for the faster, specific, sensitive, and cheaper detection of present cancer cells. Their work was acknowledged for their high specificity and affinity, aptamers were first selected fromcells by using cell-SELEX (systematic evolution of ligands by exponential enrichment) process. So, when next time combined with the unique optical properties of gold nanoparticles (Au-NPs), ANSBs were constructed on a lateral flow machine. Ramos cells were used as a model target cell for proof of principle. The traditional cancer detection techniques are conducted by pathologists with manual methods, the procedure is followed by many steps. Health history and physical test, Complete blood count (CBC), Blood chemistry tests, Bleeding and clotting factors, Cytochemistry, Immunophenotyping, Cytogenetic, and molecular studies, Bone marrow aspiration and biopsy, Lumbar puncture, Lymph node biopsy, Blood chemistry tests, Chest x-ray, CT scan, MRI, Ultrasound [19]. Enhanced and exaggerated segmentation algorithm and feature extraction [20]. Tissue counter examination (TCA), depends on apportioning the entire picture into square components of equivalent size and after that highlights are determined from these square components of the picture. The highlights, in view of GLCM (Gray level co-occurrence matrix) and dark dimension histogram, permit the separation of homogeneous and high complexity or radiant tissue territories [21] [22]. Fuzzy membership capacities while understanding a similar grouping results for the information [23]. Mammography the stage contains a lot of steps which as a gathering effectively eliminate all data which is insignificant for the classification. After trimming and improving the pictures, which speaks to the information cleaning stage, highlights significant to the order are

separated from the cleaned pictures[24] . For a progressively exact of the highlights and for a further examination of the localization [25]. Every one of these division techniques delivered great outcomes on areas showing almost no cell swarming; nonetheless, they regularly neglected to isolate contacting cells precisely [26]. The watershed group of calculations has turned out to be a standout amongst the most regularly utilized division techniques inscribing the test of contacting cells. The essential confinement of the watershed draws that they regularly result in over division [27]. Convolution is an area task in which each yield pixel is the weighted whole of neighboring info pixels. The FFT spatial recurrence goal is identified with the picture measurement and to the equal pixel measurement. So, by changing the truncation recurrence rate esteem the protested cells are stamped [28] [29].

RESEARCH METHODOLOGY

AI based applications make revolution in the field of medical science, but the main risk factor in developing them is about the accuracy of input and output data. The datasets provide the foundation to AI based applications. Every application is designed on certain parameters and need specific environment to implement. The second most important thing in these kinds of projects are mathematical applications which are applied on datasets during the training of the model. This part presents the proposed model. The model further consists of several modules integrated together towards the efficient working of the proposed model. All the modules work together to achieve the goal of reliability.

The main objective of this research is to propose a model that will manage the services provided by the data center in a more efficient way.

Capability of Proposed Model:

The proposed model is capable enough to diagnose that the patient has OSCC or not. The system has:

a) The ability to learn from its experience to make decisions in a more efficient way.
b) The ability to differentiate between normal patient images and OSCC patients image data.
c) The system has the ability to convert 2D image data into 1D vector.

Proposed Model for OSCC:

The proposed machine learning classification model for OSCC consists of the following two major modules:

a. Statistical moments
b. Decision tree classification

Statistical Moments:

The dataset of proposed system is high dimensional image data and in order to apply machine learning on it the system requires to convert it into low dimensional dataset. For this purpose, central moments are used.

a. Moments:Refers to the statistical procedure to measure the graph function about the specific points. The details of first four moments are as following:

1. 1st moment = mean

$\overline{H} = \sum H_v / n$where n is the number of elements in vector.

2. 2nd moment = variance

$H^2 = \frac{\sum(H_v - \overline{H})^2}{n-1}$.

3. 3rd moment = skewness

$$\text{skewness} = \frac{(\text{mean} - \text{median})}{\text{std} - \text{dev}}$$

4. 4th moment= kurtosis

b. Moments of images:Images moments is the technical term which is used for digital images and computer vision is used for properties of the images. Images moments contain complete information about area, orientation and centroid of images. Image moments are standard statistical formulas which are mostly apply after segmentation of images. The image moments have different types, for example spatial, raw, hnhn and central etc. For this research work central moments are used. Let suppose B is the original image and after conversion shows in figure 1. Figure 1 shows that the picture now converts into 4x4 metrics and map on x, y axis.

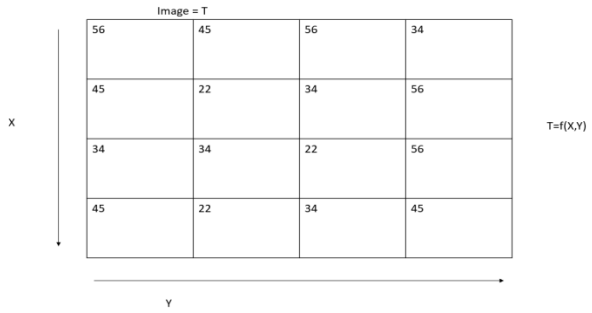

Figure 1: image representation in DIP

if T represent the image then the equation of central moment is given below.

$$\mu_{\varphi\omega} = \sum_X \sum_Y (X - \overline{X})^{\varphi} (Y - \overline{Y})^{\omega} f(X, Y)$$

Decision tree classification:

This phase of proposed method is the most important part of the proposed system. After the reduction of high dimensional dataset into low dimensional dataset this layer provides the foundation to the whole work. On this phase the model is trained for diagnosis of system on the base of previous knowledge. For this purpose, ML classification is used. This layer works as the brain of the proposed framework. It consists of three neural decision tree classifiers and is supervised machine learning technique which can be applied on classification and regression problems. Let suppose the table 1 represent the dataset which classifier need to train the model by using the following steps:

- Step-1: Tree starts with the node called root node.
- Step-2: then find the BA(best attribute)in the dataset using ASM(Attribute Selection Measure).
- Step-3: Divide the root node into subsets that contain possible values for the BA.
- Step-4: Generate the decision tree node, which contains the best attribute.
- Step-5: now repeat the process from step 1 to 4 to create the subtree of datasets

  To find out the best ASM the statistical equations are used which are called as:
  - Information gain
  - Gini index

Information gain
Information gain is a statistical process which is used to segment the dataset columns on basis of features. Information gain is providing the information attribute provide the more information about the system. Then try to minimize the value of information gain so the system creates the node on first column have high value of information gain. The general equation of information gain is given below:

$$\mathrm{IN} = \mathrm{en}(S) - \left(\bar{\partial}\right) * \mathrm{en}(\Delta)$$

Where
IN = information gain
en= entropy
S= Total number of samples
$\bar{\partial}$ =weight average
Δ= each feature
The general equation to find the entropy is as following:

$$\mathrm{en}(s) = \mathrm{p}(1)\log_2\mathrm{p}(1) - \mathrm{p}(2)\log_2\mathrm{p}(2)$$

Where:
P (1) and p (2) shows the probability of accuracy of class 1 and 2 as shown in table 1.
GINI Index
Gini index is used to check the purity and impurity of decision tree. The value with low gini index is preferred. Gini index is calculated by using following equation:

$$\mathrm{GI} = 1 - \sum_{\mathrm{j}} \mathrm{p}_{\mathrm{j}}^2$$

Table 1: sample table

m00	m10	m01	m20	m11	Class
38562375	4.25E+10	3.24E+10	6.16E+13	3.43E+13	1
43089390	5.44E+10	3.2E+10	8.24E+13	4.28E+13	2
39102720	2.74E+10	5.07E+10	2.69E+13	3.72E+13	2
35009205	2.42E+10	4.77E+10	2.33E+13	3.41E+13	1
38732460	2.69E+10	5.15E+10	2.62E+13	3.72E+13	2

PROPOSED ALGORITHM
Managing the reliability of a data center is a complex task that needs to be addressed. As management includes the low outage and downtime rates. Downtime and outage of the services are the major reasons of un-reliability of the systems.
Algorithm:
The proposed algorithm contains 6 steps and 2 phases as shown in figure 1. The algorithm will take the input, preprocess convert into low dimensional dataset, classify it and then gives the result. The main 2 phases of algorithm are:

1. Training phase
2. Testing Phase

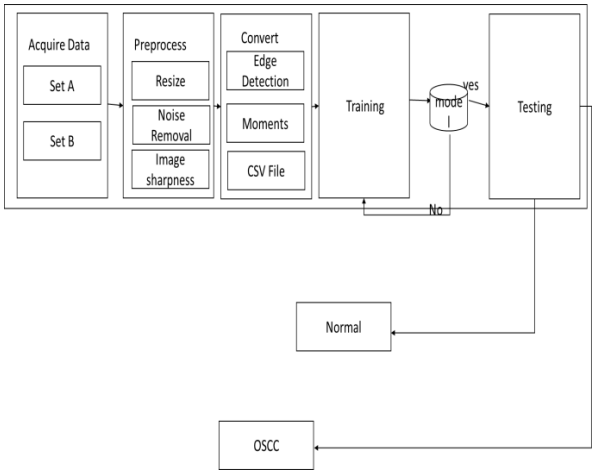

Figure 2: *Proposed Algorithm for OSCC*

4.2.1. Algorithm:

Input: OSCC parameters
Output: Detection
Class 1[] = Normal
Class 2[] =OSCC

Begins

1. Acquire Data
2. Preprocessing
3. Edge detection
4. Statistical Moments
5. Classification
6. Testing Results

SIMULATIONS/ EXPERIMENT AND RESULTS

In the proposed infrastructure the ability to make decisions based upon the knowledge obtained from the past will enhance the reliability and performance of the cancer detection. For development python 3.6 idle tool is used.

Dataset:
The dataset contains two sets ‘A’ and ‘B’. Set A represent normal images and set B contain OSCC images. The total number of images are 1225 in which 290 images are healthy images and reset of them are OSCC patient images. These images are microscopic images of patient biopsy. Figure 3 shows the normal patient biopsy image and figure 4 shows the results of OSCC patient biopsy.

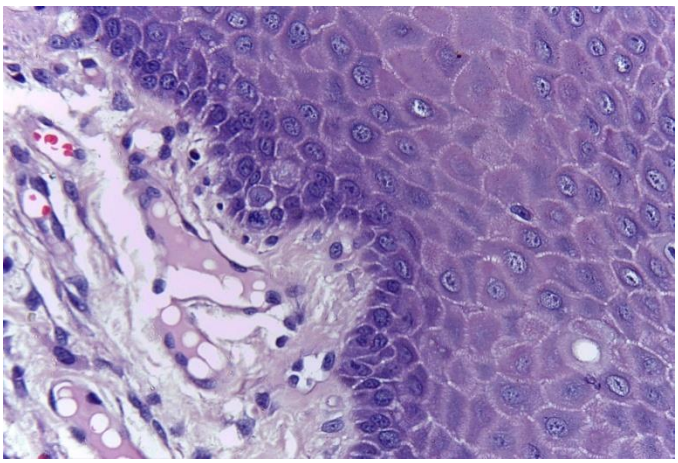

Figure 3: normal patient biopsy image

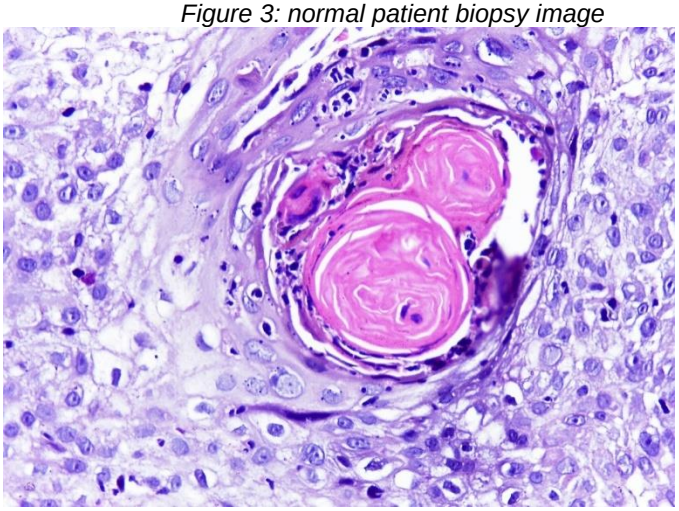

Figure 4: OSCC patient biopsy image

Preprocessing:

For OSCC detection system needs the previous data to learn. For this purpose, the proposed method is required to rationalize the dataset in specific order. So, the first step is to preprocess the image dataset to convert it into numeric form before performing training of machine learning model. For this purpose, the proposed system performs the following steps on the image:

- Resize the image
- Sharp the image
- Convert into binary
- Enhance the image

The final image after preprocessing is shown in figure 5.

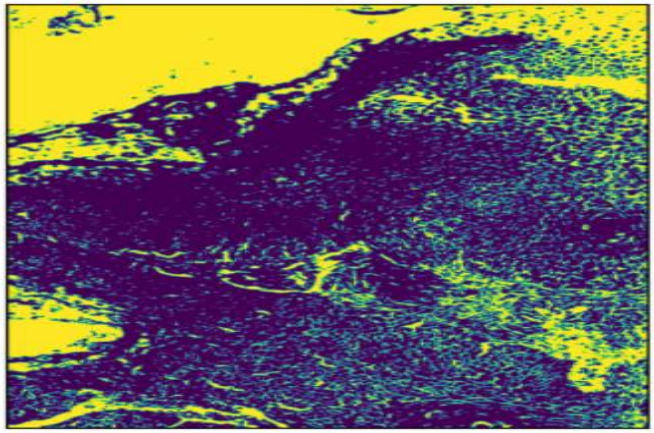

Figure 5: After preprocessing

Edge Detection:

After preprocessing the next step is to find the sharp edges of the images. For this purpose, canny edge detection algorithm is used. The resultant image of canny edge detection is shown in figure 6.

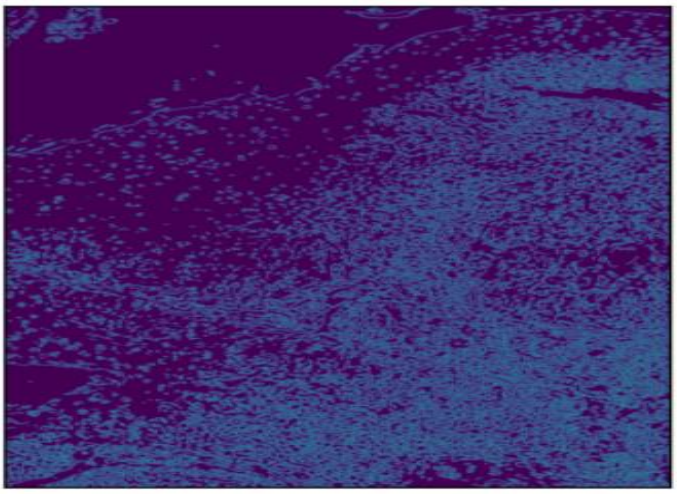

Figure 6: Edge detection

Moments:

To convert the image dataset into low dimensional dataset the proposed system uses the position variant statistical moment. The proposed system finds the statistical moment value of each image and store it into csv file and label the data as 1 and 2. In this 1 represents the normal patients and 2 represents the OSCC patients' data. After the system get the 1225 x 25 matrix in which 24 are the input values and column represents the output values. The screenshot of final dataset is shown in figure 7. The details of Position variant statistical moment is discussed in chapter 3

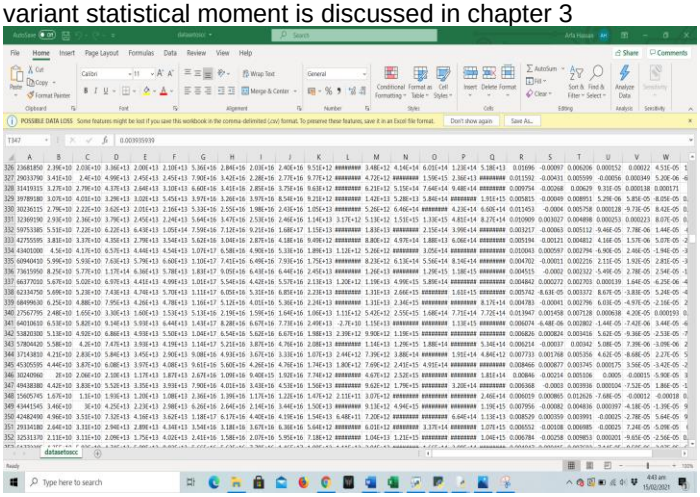

Figure 7: Screen shot of final dataset

Simulation
Input:
Statistical position variantMoments
Target Values:
Class Result
Output:
Accuracy
Division of Data Set for ML:
The table 2 shows the percentage of data that is divided for training, testing and validation purpose.

Table 2: *Storage Data Set Division for Decision tree*

Data Set Division	
Total Number of Samples	1225
Training Data	70%
Validation Data	15%

Test Data	15%

Result of Storage with 70 – 30 Division of Dataset:
The graph shows the accuracy results of data set for storage. The DD gives up-to 98% accurate results for the chosen data set.

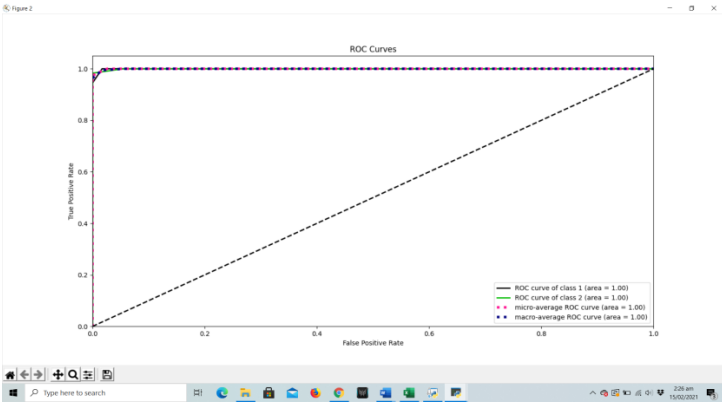

Figure 8: *Validation Curve with Dataset for Storage*

The data set is then divided with a different percentage and again is trained and validated on DD. The Confusion matrix is shown in figure 9.

```
[[290   0]
 [ 16 918]]
Score:  0.9869281045751634
```

Figure 9: *Confusion Matrix*

Limitations:
This research work is only for doctors and lab technicians who have deep knowledge of diseases. Human health issues are very critical and need a lot of intension and accuracy. The main limitation of this proposed method at this stage is that the system predicts about cancer cell using biopsy images which need professional help to take samples.

Future Work:
In future this research work can be extended toward normal camera images of oral cavity instead of biopsy images. Then the patients also able to examine their oral cavity problem.

Conclusion

Cancer is a disease in which an uncontrollable growth of cells occur that can also spread into the surrounding tissues. Oral cancer is defined as the development of cancerous cells in any part of the mouth which includes lips, tongue, cheeks, floor of the mouth, hard and soft palates, sinuses, and oropharynx (the part at the back of the mouth). Sometimes it refers to as Oral cancer and other times may also as Oral cavity cancer, but it is the same thing. There are many different types of Oral cancers, and patients with Oral cancers go through an extensive diagnostic and staging procedure to find out the type of cancer along with its stage in order to develop a tailored treatment plan for each patient according to the diagnosed type. The main focus of this research work is OSCC (oral squamous cell carcinoma). The data set contains 290 normal images and 934 OSCC effected pictures. The dataset is trained on decision tree machine learning algorithm that achieves 98% of accuracy.In future, more data will be collected and a few other stages of cancer would be added.

REFERENCES

[1] Vigneswaran N, Williams MD: Epidemiologic trends in head and neck cancer and aids in diagnosis. Oral & Maxillofacial Surgery Clinics of North America 26: 123-141, 2014.

[2] Reddy, S.S.; Sharma, S.; Mysorekar, V. Expression of Epstein–barr virus among oral potentially malignant disorders and oral squamous cell carcinomas in the South Indian tobacco-chewing population. J. Oral Pathol. Med. 2017, 46, 454–459. [CrossRef] [PubMed]

[3] Tabassum Yesmin Rahman1 (Image-Based Methods Available for Early Detection of Oral Cancer: A Review) International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 3, Issue 8, August 2015

[4] C.R.MuzakkirAhmed1(TO DETECT AND CLASSIFY ORAL CANCER IN MRI IMAGE USING FIREFLY ALGORITHM AND EXPECTATION MAXIMIZATION ALGORITHM) International Journal of Pure and Applied Mathematics Volume 116 No. 21 2017, 149-154

[5] K. Anuradha1*(A Review on Computer-Aided Detection Techniques of Oral Cancer) International Journal of Computer Science and Engine and Engineering. E-ISSN: 2347-2693.

[6] James J. Sciubba, 1999. Improving the detection of precancerous and cancerous oral lesions, American Dental Association, 130, 1445 – 1457. [2] Radha Sharma, —Oral Cancer goes viralǁ, Times of India, 27th November 2012, Http:// articles

[7] Jiang, C. F., C. Y. Wang, and C. P. Chiang. "Oral cancer detection in the fluorescent image by color image fusion." Engineering in Medicine and Biology Society, IEMBS'04. 26th Annual International Conference of the IEEE. Vol. 1. IEEE.

[8] Sharma, Neha, and HariOm. "Extracting Significant patterns for oral cancer detection using apriori algorithm." Intelligent Information Management 2014 (2014).

[9] LingyanFengaYongChenb(graphene functionalized electrochemical aptasensor for selective label-free detection of cancer cells 2017) Changchun Institute of Applied Chemistry, Graduate School of the Chinese Academy of Sciences, Chinese Academy of Sciences, Changchun, Jilin 130022, China.

[10] Weihong Tan*§LingwenZeng (Aptamer−Nanoparticle Strip Biosensor for Sensitive Detection of Cancer Cells Laboratory of Regenerative Biology, Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, Guangdong, China 510663. timesofindia.indiatimes.com/keyword/oral-cancer.

[11] RouhollahMaghsoudi, AbolfazlBagheri, Mohammad TaghiMaghsoudi, 2013.Diagnosis Prediction of Lichen Planus, Leukoplakia and Oral Squamous Cell Carcinoma by using an Intelligent System based on Artificial Neural Network, Journal of Dentomaxillofacial Radiology, Pathology, and Surgery. 2(2), 1 – 8.

[12] Jadhav. A.S, Banerjee, P.K.Dutta, R.R. Paul, M. Pal, P. Banerjee, K. Chaudhuri, J. Chatterjee, 2006. Quantitative analysis of histopathological features of precancerous lesion and condition using Image Processing Techniques, Proc.of the IEEE Symposium on Computer-Based Medical Systems.

[13] HariKumar.R, Vasanthi.N.S, Balasubramani.M, 2012. Performance Analysis of Artificial Neural Networks and Statistical Methods in Classification of Oral and Breast cancer stages. Int. J. of Soft Computing and Engineering. 2(3): 263 – 269.

[14] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018:GLOBOCAN estimates of incidence and mortality worldwide for 36 cancersin 185 countries [J]. CA Cancer J Clin. 2018;68:394–424.

[15]. Warnakulasuriya S. Global epidemiology of oral and oropharyngealcancer.OralOncol. 2009;45:309–16..

[16]. Jin-Ye F, Chun-Xiao W, Chen-Ping Z, et al. Oral cancer incidence inShanghai ---- a temporal trend analysis from 2003 to 2012[J]. Bmc Cancer.2018;18(1):686.

[17]. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015[J]. CA Cancer J Clin. 2016;66(2):115–32.

[18]. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL. Trotti Cancer Staging Handbook – AJCC Cancer Staging Manual (7th ed.). New York: Springer; 2010.

[19]. Barnes LL, Eveson JW, Reichart PADS. Pathology and genetics of head and neck tumours. Lyon: IARC Press; 2005.

[20]. Ming-hui D, Han-jiang WU. Clinical analysis of 5443 malignant oral and maxillofacial tumors in Hunan area. China J Oral Maxillofacial Surg. 2010;8(2):144–8.

[21]. Fu JY, Gao J, Zheng JW, et al. Descriptive analysis of oral squamous cell carcinoma incidence in south and East China. China J Oral and Maxillofac Surg. 2014;12(3):261–5.

[22]. Dai XM, Liu H, Wen YM, et al. Analysis of oral squamous cell carcinomaincidence of 3436 cases. Chin J Clin Oncol. 2002;29(9):674–5.[23].

[23]. Benzian H, Williams D. The challenge of oral disease: a call for global action[J].The oral health atlas. 2nd ed. Geneva: FDI World Dental Federation; 2015.

[24]. Rikardsen OG, Bjerkli IH, Uhlin-Hansen L, et al. Clinicopathological characteristics of oral squamous cell carcinoma in northern Norway: a retrospective study. Bmc Oral Health. 2014;14(1):103.

[25]. Suba Z. Gender-related hormonal risk factors for oral cancer [J]. PatholOncol Res. 2007;13(3):195–202.

[26]. Suba Z. GyörgyiMaksa, SzilviaMihályi, et al. role of hormonal risk factors in oral cancer development [J]. Orv Hetil. 2009;150(17):791–9.

[27]. Farhood Z, Simpson M, Ward GM, et al. Does anatomic subsite influence oral cavity cancer mortality? A SEER database analysis [J]. Laryngoscope. 2019;129(6):1400–6.

[28]. Huang CH, Chu ST, Ger LP, et al. Clinicopathologic evaluation of prognostic factors for squamous cell carcinoma of the buccal mucosa. J Chin Med Assoc. 2007;70(4):164–70.

[29]. Bobdey S, Sathwara J, Jain A, et al. Squamous cell carcinoma of buccal mucosa: an analysis of prognostic factors [J]. South Asian J Cancer. 2018;7(1)