

Survey Paper on Oral Cancer Detection using Machine Learning

Madhura V¹, Meghana Nagaraju², Namana J³, Varshini S P⁴, Rakshitha R⁵

^{1,2,3,4}Students, Dept of Computer Science and Engineering, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India

⁵Assistant Professor, Dept of Computer Science and Engineering, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India

Abstract - Oral cancer is the most common type of cancer. It was an irrecoverable disease but now the progress in technology has made it curable if it is diagnosed in early stages. Oral cancer is increase in the number of cells which has the capability to affect its neighbor cells or tissues. It happens when cells divide out of control and form a growth, or tumor. In spite of having various advancements in fields like radiation therapy and chemotherapy the mortality rate is persistent. Therefore early detection of cancer is important. In this paper we are using Machine Learning as domain for detection of oral cancer considering the datasets of a victim. Then it is classified using apriori algorithm. We are developing a health sector application which also uses Data Mining and data extraction for prediction techniques, classification rules for oral cancer prediction and uses association rules to perceive the relationship between the oral cancer attributes.

Key Words: Oral cancer detection, machine learning, data mining technique, association rule mining, apriori algorithm.

1. INTRODUCTION

Cancer has been characterized as a heterogeneous disease consisting of various subtypes. Oral cancer is the most dangerous type of cancer. It is caused in various parts such as lips, tongue, hard and soft palate and floor of mouth. Oral cancer is caused due to tobacco use of any kind, including cigars, pipes, chewing tobacco, snuffs, or alcohol consumption. Its detection and diagnosis is very important or else it can be fatal. Therefore early diagnosis of a cancer type have become a necessity in cancer research. The ability of ML tools to detect key features from composite datasets reveals their significance. The predictive models discussed here are based on various supervised ML techniques as well as on different input features and data samples. Given the rapidly growing trend on the application of ML methods in cancer research, we present here the most recent publications that employ these techniques as an aim to model cancer risk or patient outcomes.

2. LITERATURE REVIEW

Arushi Tatarbe, Tanupriya Choudhury, Teoh Teik Toe, Seema Rawat [1] proposed a Oral Cancer Detection Using Data Mining Technique which used different algorithms of

data mining to detect oral cancer. Data mining is referred to as a prominent technique employed by various health institutions for classification of life threatening diseases, e.g. cancer, dengue and tuberculosis. In this proposed approach WEKA (Waikato Environment for Knowledge Analysis) is applied with ten cross validation to calculate and collect output. WEKA consists of a large variety of data mining machine learning algorithms. First the system classifies the oral cancer dataset and then analyses various data mining methods in WEKA through Explorer and Experiment interfaces. The major aim is to classify the dataset and help to collect important and useful material from the data and choose an appropriate algorithm for accurate prognostic model from it. This paper uses two two interfaces of Weka i.e., Explorer and Experimenter.

Explorer: This interface is responsible for preprocessing the dataset and further filtering it. Later it analyses the accuracy of the classification using the following algorithms:

1. Naive Bayes
2. J48
3. SMO
4. REP
5. Random tree

Experimenter: This interface is responsible for analyzing the data in the dataset by using algorithms to classify the dataset using the test and train sets. The following algorithms are used

1. Naive Bayes
2. J48
4. REP
5. Random tree

Woonggyu Jung, Jun Zhang, Jungrae Chung, Petra Wilder Smith, Matt Brenner, J. Stuart, Nelson, Zhongping Che [2] proposed Optical Coherence Tomography (OCT) which is a new classification capable of cross sectional imaging of biological tissue. Due to its many technical advantages such as high image resolution, fast acquisition time, and noninvasive capabilities, OCT is very useful in various medical applications. Because OCT systems can function with a fiber optic probe, they are applicable to almost any anatomic structures attainable either directly, or by endoscopy. OCT has the capacity to provide a fast and noninterfering means for early clinical detection, diagnosis, screening, and monitoring of precancer and cancer.

The goal of this study was to evaluate the feasibility of OCT for the diagnosis of multiple stages of oral cancer

progression. In this paper, conventional 2-D OCT images, and also 3-D volume images of normal and precancerous lesions are presented. The results demonstrate that OCT is a potential tool for cancer detection with complete diagnostic images.

D.Padmini Pragna, Sahithi Dandu, Meenakzshi M, C. Jyotsna, Amudha J [3] proposed Health Alert System to Detect Oral Cancer. The proposed health alert system enables the patients in identifying the disease in the initial stage itself. It accepts the Computerized Tomography (CT) scanned images of the cancer affected region and can detect the presence of malignancy. The obtained CT image is preprocessed using Adaptive Median Filter and the features such as Texture, Shape, Water Content, Linear Binary Pattern (LBP), Histogram of Oriented Gradients (HOG) and Gray Level Co-occurrence Matrix (GLCM) are extracted from preprocessed images. The unnecessary features are excluded using features election process and Support Vector Machine (SVM) classification algorithm is used to classify it as benign or malignant. Proposed Health Alert system has an accuracy of 97%.

After feature extraction, the features are selected for classification purpose. Here SVM classifier is used and the results are analyzed with K-nearest neighbor (KNN) algorithm to find the best classification technique.

Both the algorithms classifies whether the given image is cancerous or not and In addition to that, SVM classifier predicts the level of cancer.

Accuracy obtained for SVM classifier is 97%, precision is 95%, sensitivity is 95% and specificity is 96%.

Harikumar Rajaguru [4] Department of ECE Bannari Amman Institute of Technology Sathyamangalam, Innd Sunil Kumar Prabhakar Department of ECE Bannari Amman Institute of Technology Sathyamangalam, India proposed Oral Cancer Classification from Hybrid ABC-PSO and Bayesian LDA model. In this paper, Hybrid Artificial Bee Colony – Particle Swarm Optimization (ABC-PSO) algorithm and Bayesian Linear Discriminant Analysis (BLDA) is used to classify the risk level of oral cancer. The results show that when Hybrid ABC-PSO classifier is used, a classification accuracy of 100% is obtained while for BLDA classifier, a classification accuracy of about 83.16% is obtained.

Gaussian Mixture Models (GMM) and Multilayer Perceptrons were used for the classification of oral cancer. The normal Premalignant and malignant pathological conditions were classified using Principal Component Analysis(PCA).With the help of laser Induced Fluorescence a novel feature selection was done.

In this paper, the risk of oral cancer is being classified by the aid of Hybrid ABCPSO classifier and BLDA classifier. The results show that for all the stages of oral cancer, the Hybrid

ABC-PSO classifier gives a classification accuracy of 100% while for BLDA classifier, in T1 stage it gives an accuracy of 91.66%, for T2 stage it gives an accuracy of 74.49%, for T3 stage it gives an accuracy of 81.23% and for T4 stage it gives an accuracy of 85.25%. The average classification accuracy for BLDA classifier is about 83.16% while the average classification accuracy for Hybrid ABC-PSO classifier is 100%. The results are also compared to the classification accuracy percentage done by clinical procedure.

Julia D. Warnke-Sommer Department of Pathology and Microbiology University of Nebraska Medical Center Omaha, NE, USA and Hesham H. Ali Department of Computer Science University of Nebraska Omaha Omaha, NE, USA [5] proposed Evaluation of the Oral Microbiome as a Biomarker for Early Detection of Human Oral Carcinomas. In this paper metagenomics analysis pipeline and SVM model training techniques are used by using machine learning metagenomics analysis pipeline is used for processing and extracting features from metagenomics read data sets and this allows for the consistent extraction of biological features and minimizes variation that would result from an ill-defined approach. By using PCA plots show a clear separation between the cancer associated microbiome samples and healthy samples. This is evidence that the bacterial abundances and KO category abundances can be used to successfully separate healthy microbiomes from cancer associated microbiomes. Finally obtained features from the metagenomics extraction pipeline to train and test three SVM models.

The goal of this research is to evaluate oral microbiome based biomarkers for early oral carcinoma detection. The outcome of this research is a machine-learning based framework for microbiome-based early cancer detection. The ability to identify at risk patients using minimally invasive biomarkers will allow for more rapid treatment plan development and improved outcome.

Kazuhiro Tominaga Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan, Mana Hayakawa Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan, Shinobu Sato Department of Applied Chemistry Kyushu Institute of Technology Kitakyushu, Japan, Masaaki Kodama Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan, Manabu Habu Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan and Shigeori Takenaka Department of Applied Chemistry Kyushu Institute of Technology Kitakyushu, Japan[6] proposed Electrochemical telomerase assay for oral cancer screening . In this paper Exfoliated cells from whole oral cavity (EO) that is the EO samples were collected by scratching the whole oral cavity with a sponge-type brush. Collection method could be used as part of a self-screening system for patients concerned about oral cancer and Exfoliated cells from local lesions (EL) are used by using electrochemical telomerase assay (ECTA).

Telomerase has long been known as a cancer marker here they established a new electrochemical telomerase assay (ECTA) method that is superior to a telomerase repeat amplification protocol assay, a popular method for detection of telomerase activity. In the present study, we employed 3 types of clinical samples to examine the sensitivity and specificity of ECTA. Three types of samples were obtained from 44 oral squamous cell carcinoma patients and 16 healthy volunteers; exfoliated cells from the whole oral cavity, exfoliated cells from local lesions, and tissue samples from lesions. Telomerase activity was determined in the samples using ECTA. The increase in current in the oral cancer group was significantly higher than that in the healthy volunteer group for each sample type, while there were no significant differences among the 3 sample types in the oral cancer group. The sensitivity and specificity of ECTA was 94.6% and 88.6%, respectively. The ECTA method showed excellent detection of telomerase activity and is considered applicable as an oral cancer screening system.

Youmin Wang, Milan Raj, H. Stan McGuff, Ting Shen and Xiaojing Zhang [7] Department of Biomedical Engineering, University of Texas at Austin, Austin, USA proposed Portable oral cancer detection using miniature confocal imaging probe with large field of view. In this paper micro electromechanical system (MEMS) micrometer, MEMS scanning mirror, handled imaging probes, confocal imaging System is done by using digital image processing. Demonstrate MEMS micro mirror enabled hand held confocal imaging probe for portable oral cancer detection, where large field of view (FOV) is achieved through the lissajous scanning operation of MEMS mirror. This designed fabricated and characterized handled confocal imaging probe utilizing the MEMS micro mirror as core scanning device with versatile imaging pattern control Submicron lateral resolution higher than $3\mu\text{m}$ with FOV up to $100\mu\text{m} \times 100\mu\text{m}$ was achieved. To ensure Java GUI and image processing and rendering algorithms for lissajous scanning was developed. Additional functions such as mosaic imaging developed for FOV enhancement at suitable axial resolution. The compact handheld confocal imaging system shows promising experimental results in clinical test for early oral cancer diagnosis and treatments.

3. CONCLUSIONS

This paper deals with the different algorithms and techniques used for detecting oral cancer. It illustrates the use of some of the techniques like Metagenomic analysis, SVM, 3D and 2D scanning methodology, ABC-PSO, MEMS micrometer to detect oral cancer in different stages.

REFERENCES

[1] Arushi Tatarbe, Tanupriya Choudhury, Teoh Teik Toe, Seema Rawat "Oral Cancer Detection Using Data Mining tool".

[2] Woongyu Jung, Jun Zhang, Junrae Chung, Petra WilderSmith, Matt Brenner, J. Stuart, Nelson, Zhongping Che "Optical Coherence Tomography (OCT)"

[3] D.Padmini Pragna, Sahithi Dandu, Meenakzshi M, C. Jyotsna, Amudha J "Health Alert System to Detect Oral Cancer".

[4] Harikumar Rajaguru Department of ECE Bannari Amman Institute of Technology Sathyamangalam, Innd Sunil Kumar Prabhakar Department of ECE Bannari Amman Institute of Technology Sathyamangalam, India. "Oral Cancer Classification from Hybrid ABC-PSO and Bayesian LDA model".

[5] Julia D. Warnke-Sommer Department of Pathology and Microbiology University of Nebraska Medical Center Omaha, NE, USA and Hesham H. Ali Department of Computer Science University of Nebraska Omaha Omaha, NE, USA "Evaluation of the Oral Microbiome as a Biomarker for Early Detection of Human Oral Carcinomas"

[6] Kazuhiro Tominaga Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan, Mana Hayakawa Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan, Shinobu Sato Department of Applied Chemistry Kyushu Institute of Technology Kitakyushu, Japan, Masaaki Kodama Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan, Manabu Habu Department of Oral and Maxillofacial Surgery Kyushu Dental University Kitakyushu, Japan and Shigeori Takenaka Department of Applied Chemistry Kyushu Institute of Technology Kitakyushu, Japan "Electrochemical telomerase assay for oral cancer screening".

[7] Youmin Wang, Milan Raj, H. Stan McGuff, Ting Shen and Xiaojing Zhang Department of Biomedical Engineering, University of Texas at Austin, Austin, USA "Portable oral cancer detection using miniature confocal imaging probe with large field of view".