EARLY PHASE ORAL CANCER DIAGNOSIS USING CLAHE AND ADAPTIVE REGION GROWING

M. Lydia Packiam Mettilda[1*], Dr. P. Raajan[2] , Dr. S Antelin Vijila[3]

Abstract –
Cancers are a general term used to describe all diseases that cause uncontrolled and abnormal cell division. Oral cancer has been identified as one of the most common cancer kinds in the world. Oral cancer is caused by a number of modifiable risk factors, such as sugar consumption, cigarette use, alcohol use, poor hygiene, and their underlying societal and economic drivers. There are several non-communicable diseases (NCDs) that share similarities with these risk factors. The human papilloma virus, 16, or HPV 16, is a unique factor for producing mouth cancer that is unrelated to cigarette use. Oral squamous cell carcinoma (OSCC), often known as oral cancer, is an ulceroproliferative oral mucosa disease that can affect any area of the mouth, including the lips and tongue. Clinical paradigms, technological advancements, and variations in patient makeup present both opportunities and obstacles for the treatment of OSCC. Imaging will continue to play a bigger role in the staging, planning, and monitoring of patients with OSCC. Imaging techniques can now be used to non-invasively identify molecular and cellular changes in cells. Medical imaging applications employ pre-processing and segmentation techniques. Images of oral cancer (such as Histopathology images) must be preprocessed in order to raise their quality, make important features more visible, and make accurate analysis and diagnosis possible. After pre-processing, segmentation is essential for defining the limits of oral tumours or lesions within histopathology images. By taking into account image quality, this study seeks to increase the accuracy of oral cancer diagnosis. In order to enhance the picture quality and clarity for the task of feature extraction, it has concentrated on oral histopathology image pre-processing and segmentation procedures.

Keywords: Squamous cell carcinoma, medical imaging, pre-processing, segmentation

[1*]Research Scholar, (Reg. No: 21123092292009) PG and Research Department of Computer Science, Muslim Arts college, Thiruvithancode, Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli -627012, Tamil Nadu, India. E-mail: lydiaalwyn@gmail.com.
[2]Associate Professor, PG and Research Department of Computer Science, Muslim Arts college, Thiruvithancode, Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli -627012, Tamil Nadu, India. E-mail: raajanp99@gmail.com.
[3]Assistant Professor of Computer Science & Engineering, Manonmaniam Sundaranar University, Tirunelveli – 627012, Tamilnadu, India. E-mail: antelinvijila@gmail.com.

***Corresponding Author:** -M. Lydia Packiam Mettilda
[*]Research Scholar, (Reg. No: 21123092292009) PG and Research Department of Computer Science, Muslim Arts college, Thiruvithancode, Affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli -627012, Tamil Nadu, India. E-mail: lydiaalwyn@gmail.com.

DOI: 10.53555/ecb/2022.11.11.81

I. INTRODUCTION

A lot of study has been done to make machines intelligent in the Fourth Industrial Revolution (4IR) era that we are currently living in. Computers can now learn from their mistakes and comprehend the world in terms of a hierarchy of concepts thanks to this technology. Pattern recognition provides an essential tool for activities involving healthcare analytics [1][2]. Oral cancer may display a number of behavioral traits. For the treatment of oral cancer to be appropriate and effective, early detection and accurate prognostic prognosis are essential. The models of machine learning and deep learning have received recognition for their potential to transform the treatment of cancer by improving diagnostic precision and prognostication. When a tumors appears in a portion of the mouth, it is called oral cancer. It could be on the tongue's surface, the inside of the cheeks, the palate, the lips, or the gums. Additionally, tumours can form in the salivary glands, tonsils at the back of the mouth, and pharynx, the portion of the throat that connects the mouth to the windpipe. Understanding the type of oral cancer will aid in performing an accurate diagnosis and therapy. Oral cancer is characterized by unchecked cell development that over time invades and affects the neighbouring organs. Smoking, chewing tobacco, drinking alcohol excessively, being infected with HPV, and exposure to the sun are all risk factors for oral cancer.

Head and neck cancer can also have the following forms: [3]laryngeal cancer, which affects the voice box. Nasopharyngeal cancer, which affects the region behind the nose that makes up the pharynx. Oropharyngeal cancer, which affects the area of the throat that is immediately behind the mouth. Cancer of the thyroid gland, a butterfly-shaped gland on either side of the windpipe. cancer of the hypopharynx, the area of the throat directly behind the larynx. Cancer of the oesophagus, the food pipe, and the nose and sinusesSquamous cell carcinomas, commonly known as squamous cell cancers, make up the majority of oral cancers. Squamous cells, which are thin, flat cells that make up the lining of the mouth and throat, are where these malignancies begin[4]. It's crucial to get an early diagnosis[5] to prevent oral cancer. People who seem healthy and are not suspected of having oral cancer are routinely given screening exams.

The incidence of Oral Squamous Cell Carcinoma is 3% in United States of America whereas it is 30% in India and other Asian countries [6]. According to the American Cancer Society, approximately 48,000 Americans develop oral cancer every year and 8500 people die of the disease annually. India records more than 1, 00,000 cases of oral cavity cancers every year. India has the highest prevalence of oral cancer in the world (19/100, 000 population). It is the most common cancer in men and the third most common cancer in women, and constitutes 13%–16% of all cancers. Of all the oral cancers, 95% are related to the use of tobacco products.By finding oral cancer before any symptoms appear, screening lowers mortality.

The Process of an automated pattern recognition system can be divided into two basic tasks: the **description** task constructs an attributes of an object using ***feature extraction*** techniques, and the ***classification task*** impute a group label to the object based on those attributes with a classifier. The description and classification tasks work together to decide the most accurate level for each unlabeled object examined by the pattern recognition system. [7].

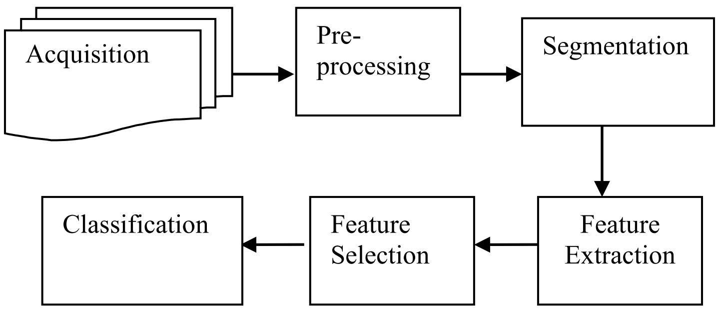


Figure1.1 Pattern recognition systems

The general processing activities of pattern recognition are [8].

- **Image acquisition**: This step involves capturing oral cancer images using histopathology imaging modalities.
- **Pre-processing**: techniques such as smoothing filters can help reduce noise, resulting in a cleaner image with improved clarity and diagnostic value.

- **Segmentation:** algorithms such as thresholding, region growing, or active contour models to identify and extract specific regions of interest in the oral cancer images, such as tumors or lesions.
- **Feature extraction:** Extract relevant features from the segmented regions, such as shape, texture, or intensity-based features, to characterize the oral cancer characteristics for further analysis or classification.
- **Feature selection**: Select a subset of the extracted features that contribute the most to the discrimination of different patterns. Feature selection helps reduce dimensionality and focus on the most informative features.
- **Classification:** Use the preprocessed data to train a pattern recognition model. This step involves selecting an appropriate algorithm or model, initializing its parameters, and optimizing the model using the training dataset. The model learns the patterns and their associated labels from the training data.

The proposed study focused on two main sections

i) to reduce noise and artifacts in the histopathology images

ii) to accurately segment specific regions of interest relevant to oral cancer diagnosis, treatment planning, or monitoring

II. RELATED WORKS

Many efforts towards oral cancer classification have been tried and most of the give results for wide ranges. Table 2.1 shows the various pre-processing and segmentation methods which have been used in image analysis of oral cancer.

Table 2.1 Various techniques for pre-processing and segmentation

Ref	Pre-processing	Segmentation
[9]	Histogram Specification	Collagen segm entation using neu ral network
[10]	Histogram Normalization	Blob Detection
[11]	Histogram Normalization	-
[12]	Savitzky-Golay filter	-
[13]	Gray level normalization	-
[14]	Gaussian Blur	Otsu’s thresholding
[15]	Gabor filter	
[16]	Anisotropic diffusion filtering	Fuzzy C-Means and neutrosophic algorithm
[35]	Histogram Normalization	Otsu's segmentation,maximally stable extremal regions(MSER)
[27]	Anisotropicfilter	Morphological segmentation
[25]	Adaptive Bilateral Filter	Artificial Bee Colony Partition
[22]	-	region-based active contoursegmentation
[20]	fuzzy-based (CLAHE)	-
[43]	Median filter	-
[38]	Gabbor Filter	
[21]	Gabor Filter	-
[41]	Gaussian Filter	
[44]	Adaptive Histogram Equalization	

III PROPOSED METHOD

In the proposed method, a new pre-processing filter has been proposed to improve image quality by keeping all the edges clearly visible.

3.1 Contrast Limited Adaptive Histogram Equalization (CLAHE)

The CLAHE (Contrast Limited Adaptive Histogram Equalization) filter with clip limits follows the following algorithm steps:

Input: Grayscale image

Parameters: Clip limit, Tile size, Tile overlap

Output: Filtered oral cancer image

1. Convert the input image to grayscale if it is in color.
2. Divide the image into small, overlapping tiles.
3. Calculate the histogram of each tile.
4. Clip the histogram of each tile based on a specified clip limit. This means that histogram values above the clip limit are truncated.

$$\beta = \frac{M}{N}(1 + \frac{\alpha}{100}(S_{max} - 1))$$

5. Redistribute the clipped histogram values uniformly.
6. Calculate the cumulative distribution function (CDF) for each tile's histogram.
7. Map the intensity values of each tile to a new range using the CDF.
8. Combine the processed tiles to form the final CLAHE filtered image.

The clip limit parameter determines the amount of contrast enhancement in the filtered image. Higher values of the clip limit result in more contrast enhancement, while lower values restrict the amplification of histogram peaks, leading to less contrast enhancement.

3.2 Adaptive Region Growing Segmentation

The algorithm starts with an image, a set of seed points (initial points of interest), and a similarity threshold that determines the criterion for region growing. It iteratively grows regions by expanding from the seed points and adding neighboring pixels that meet the similarity criterion.

Input: Filtered Oral cancer image

Initialization:

1. Select seed points or regions of interest as the starting points for region growing.
2. Create an empty label matrix with the same dimensions as the input image.

Seed Region:

1. Choose a seed point or region from the set of initial seed points.
2. Assign a unique label to the seed point or region in the label matrix.

Region Growing:

1. Iterate over the neighboring pixels of the seed point or region.
2. Check the similarity criteria between the neighboring pixel and the growing region.
3. If the neighboring pixel satisfies the similarity criteria:
4. Assign the same label to the neighboring pixel in the label matrix.
5. Add the neighboring pixel to the growing region.
6. Expand the growing region by including more neighboring pixels.
7.

$$T_{upper} = mgv(n) + [ud(n).w + c(n)]$$
$$T_{lower} = mgv(n) - [ld(n).w + c(n)]$$

Iteration:

1. Repeat the region growing process until no more neighboring pixels satisfy the similarity criteria or until a stopping condition is met.
2. The stopping condition can be a maximum region size, a threshold on the similarity measure, or a predefined number of iterations.

Result:

1. The final label matrix represents the segmented regions of the input image.
2. Each connected region in the label matrix corresponds to a separate segment or object.

IV EXPERIMENTAL RESULTS AND DISCUSSION

The Kaggle Dataset, which included 1224 oral histopathology pictures (290 non-cancerous and 934 malignant) from 230 patients, provided the Histopathologic Image. Instances of both classes from the dataset are shown in Figure 4.1. The images were obtained using a Leica ICC50 HD microscope at two different magnifications (100x and 400x) from tissue slides that had been stained with hematoxylin and eosin (H&E). 439 photos of malignant epithelium and 89 images of normal epithelium were magnified by 100 times each, while 201 normal images and 495 OSCC images were magnified by 400 times each (Figure 4.2).

Table 4.1 Details of dataset

S.NO	DATASET NAME	DATASET LINK
1	Histopathologic OralCancer Detection	https://www.kaggle.com/ashenafifasilkebede/dataset

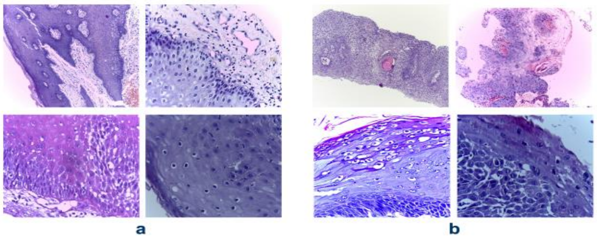


Figure 4.1: Vignettes of H&E stained oral histopathology images from the Oral Cancer dataset capturing normal epithelium (a) and cancerous epithelium(b)

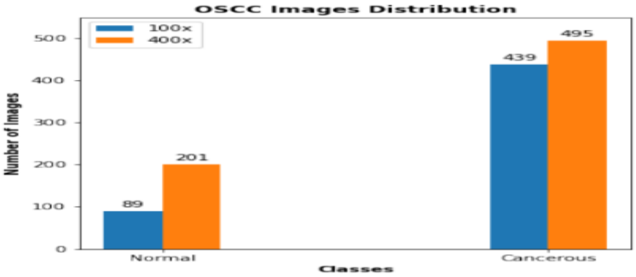


Figure 4.2. Distribution of images in the dataset.

The Normal and Cancerous class contains 290 and 934 images, respectively. The images are in two different magnifications — 100x (89 Normal, 439 Cancerous) and 400x (201 Normal, 495 Cancerous)Figure 4.3 shows the example of an image and the result of applying clahe filter.

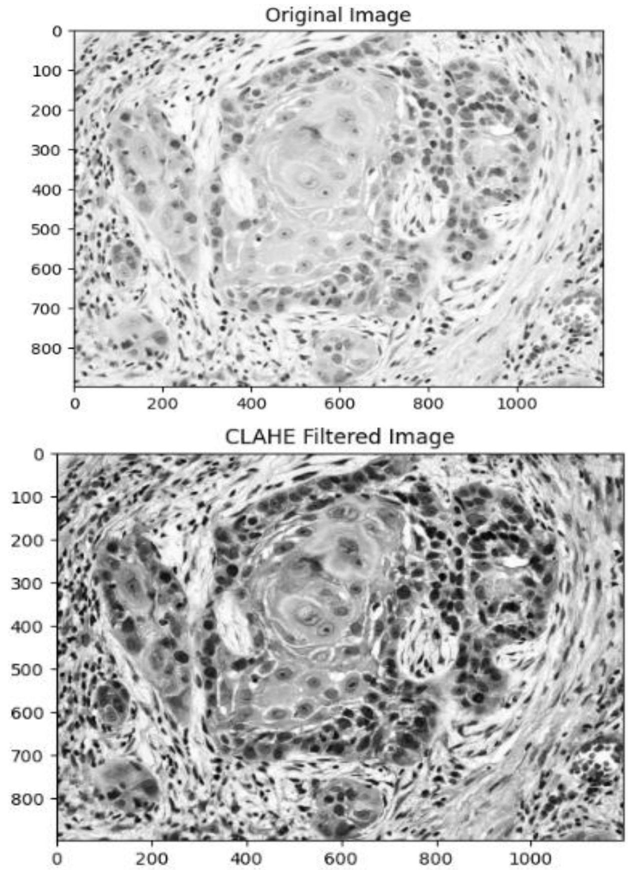
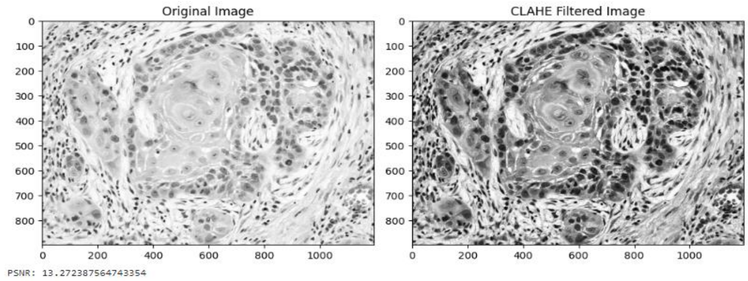


Figure 4.3 Applied CLAHE filter

CLAHE improves the visibility of structures in different regions of an image and enhance the visualization of important features.

4.2 System Performance Measures

For performance measurement, the code calculates the Intersection over Union (IoU) metric as an example. Peak Signal-to-Noise Ratio (PSNR)[27]: PSNR measures the ratio between the maximum possible signal power and the power of the noise present in an image. Higher PSNR values indicate better quality images. It produced as 13.2723.



V CONCLUSION

Oral cancer prevalence is rising. Although histopathology is one of the greatest ways to find oral cancer, pathologists may have trouble locating the lesions. The medical staff may benefit from the procedures described in this research and their accuracy of detection. Images are taken and a number of operations are carried out in this study to categorise them as normal or abnormal. Filtering techniques such as clahe filter and segmentation method adaptive region growing were used to remove speckle noise and segment the region of interest. After applying these techniques the image quality obtained and assessed with metrics. These metrics include Peak signal to Noise Ratio (PSNR). Following are some of the research's potential future improvements:

- Future research on the features extracted from oral cancer data is required to produce better and more accurate outcomes.
- Using optimization approaches, adding a fuzzy hierarchical classifier to random forest.

REFERENCES

1. Lu, J., Sun, J., and Wang, S., “Pattern recognition: An overview”,International journal of Computer Science and Network Security, Vol.6, No.6, pp.57-61, 2001.
2. Ghorpade, S., Ghorpade, J., and Mantri, S., “Pattern Recognition using Neural Networks”, International Journal of Computer Science and Information Technology”, Vol 2, No.6, pp.92-97, 2010.
3. https://www.cancer.gov/types/head-and-neck/head-neck-fact-sheet
4. Mehrotra R, Gupta DK. Exciting new advances in oral cancer diagnosis: avenues to early detection. Head Neck Oncol. 2011 Jul 28; 3:33. doi: 10.1186/1758-3284-3-33. PMID: 217980 30; PMCID: PMC3170277.
5. Vivek Borse, Aditya Narayan Konwar, Pronamika Buragohain, “Oral cancer diagnosis and perspectives in India”, Sensors International, Volume 1, 2020,100046, ISSN 2666-3511.
6. https://www.indiancancersociety.org/oral-cancer/
7. Olszewki, R.T., “Generalized Feature Extraction for Structural Pattern Recognition in Time-Series Data”, pp.1, 2001.
8. Lu, J., Sun, J., and Wang, S., “Pattern recognition: An overview”,International journal of Computer Science and Network Security, Vol.6, No.6, pp.57-61, 2001.
9. Muthu Rama Krishnan M, Shah P, Chakraborty C, Ray AK. Statistical analysis of textural features for improved classification of oral histopathological images. J Med Syst. 2012 Apr;36(2):865-81. doi: 10.1007/s10916-010-9550-8. Epub 2010 Jul 16. PMID: 20703647.
10. Galib, S., Islam, F., Abir, M., & Lee, H. K. (2015). Computer aided detection of oral lesions on CT images. *Journal of Instrumentation*, *10*(12), C12030.
11. Kiruthikaa K V1 And Dr Vijay Franklin. Detection Of Oral Cancer In Hyperspectral Images Using Restricted Boltzmann Machines, J2Volume 12, Issue 3 - Serial Number 3ICMMNT-2021 International Virtual Conference on Materials, Manufacturing and Nanotechnology, 30th June, 2021. Pages 1760-1766.
12. Yu M, Yan H, Xia J, Zhu L, Zhang T, Zhu Z, et al. Deep convolutional neural networks for tongue squamous cell carcinoma classification using Raman spectroscopy. Photodiagnosis Photodyn Ther 2019; 26:430–5. https://doi.org/ 10.1016/j.pdpdt.2019.05.008.
13. Jiliang Ren J, Qi M, Yuan Y, Duan S, Tao X. Machine Learning-Based MRI Texture Analysis to Predict the Histologic Grade of Oral Squamous Cell Carcinoma. AJR Am JRoentgenol. 2020 Nov;215(5):1184-1190. doi: 10.2214/AJR.19.22593. Epub 2020 Sep 15. PMID: 32930606.

14. Santisudha Panigrahi, Jayshankar Das, Tripti Swarnkar, "Capsule network based analysis of histopathological images of oral squamous cell carcinoma", Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 7, 2022, Pages 4546-4553, ISSN 1319-1578
15. Chan C-H, Huang T-T, Chen C-Y, Lee C-C, Chan M-Y, Chung P-C. Texture-map based branch-collaborative network for oral cancer detection. IEEE Trans Biomed Circuits Syst 2019;13:766–80. https://doi.org/10.1109/TBCS.2019.2918244
16. Anooja Ali, Pooja G, Prajeela MP, Riddhi Rakesh, Tabassum Taj, "Machine Learning Approach for Cancer Detection," *International Journal of Computer Sciences and Engineering*, Vol.07, Issue.14, pp.224-228, 2019.
17. Rahman TY, Mahanta LB, Choudhury H, Das AK, Sarma JD. Study of morphological and textural features for classification of oral squamous cell carcinoma by traditional machine learning techniques. Cancer Rep (Hoboken). 2020 Dec;3(6):e1293. doi: 10.1002/cnr2.1293. Epub 2020 Oct 7. PMID: 33026718; PMCID: PMC7941561.
18. Prabhakaran, R. and Mohana, Dr. J., Detection of Oral Cancer Using Machine Learning Classification Methods (June 30, 2020). International Journal of Electrical Engineering and Technology, 11(3), 2020, pp. 384-393, Available at SSRN: https://ssrn.com/abstract=3638829
19. Dharani, R. & Revathy, S.. (2021). DEEPORCD: Detection of Oral Cancer using Deep Learning. Journal of Physics: Conference Series. 1911. 012006. 10.1088/1742-6596/1911/1/012006.
20. B R, Nanditha & Annegowda, Geetha. (2022). Oral Cancer Detection using Machine Learning and Deep Learning Techniques. International Journal of Current Research and Review. 14. 64-70. 10.31782/IJCRR.2021.14104.
21. R. Marzouk, E. Alabdulkreem, S. Dhahbi, M. K. Nour, M. Al Duhayyim *et al.*, "Deep transfer learning driven oral cancer detection and classification model," *Computers, Materials & Continua*, vol. 73, no.2, pp. 3905–3920, 2022.
22. Panigrahi S, Nanda BS, Bhuyan R, Kumar K, Ghosh S, Swarnkar T. Classifying histopathological images of oral squamous cell carcinoma using deep transfer learning. Heliyon. 2023 Feb 6;9(3): e13444. doi: 10.1016/j.heliyon.2023.e13444. PMID: 37101475; PMCID: PMC10123069.
23. Rahman AU, Alqahtani A, Aldhafferi N, Nasir MU, Khan MF, Khan MA, Mosavi A. Histopathologic Oral Cancer Prediction Using Oral Squamous Cell Carcinoma Biopsy Empowered with Transfer Learning. Sensors (Basel). 2022 May 18;22(10):3833. doi: 10.3390/s22103833. PMID: 35632242; PMCID: PMC9146317.
24. Alanazi AA, Khayyat MM, Khayyat MM, Elamin Elnaim BM, Abdel-Khalek S. Intelligent Deep Learning Enabled Oral Squmous Cell Carcinoma Detection and Classification Using Biomedical Images. Comput Intell Neurosci. 2022 Jun 30; 2022: 7643967. doi: 10.1155/2022/7643967. PMID: 35814555; PMCID: PMC9262470.
25. Deif MA, Attar H, Amer A, Elhaty IA, Khosravi MR, Solyman AAA. Diagnosis of Oral Squamous Cell Carcinoma Using Deep Neural Networks and Binary Particle Swarm Optimization on Histopathological Images: An AIoMT Approach. Comput Intell Neurosci. 2022 Sep 30;2022: 6364102. doi: 10.1155/2022/6364102. PMID: 36210968; PMCID: PMC9546660.
26. Ananthakrishnan B, Shaik A, Kumar S, Narendran SO, Mattu K, Kavitha MS. Automated Detection and Classification of Oral Squamous Cell Carcinoma Using Deep Neural Networks. Diagnostics (Basel). 2023 Feb 28;13(5):918. doi: 10.3390/diagnostics13050918. PMID: 36900062; PMCID: PMC10001077.
27. https://chat.openai.com/share/a07b7ef5-ddfb-45d7-b94a-4d36e26385f0