Oral Cancer Detection by CNN

C. ANITHA[1], S. SASIKUMAR[2]

[1] M. Tech., Assistant Professor, Department of Electronics and Communication Engineering, Jeppiaar Engineering College, Chennai.

[2] Student, Department of Electronics and Communication Engineering, Jeppiaar Engineering College, Chennai.

Abstract*— Oral cancer is a major global health issue accounting for 177,384 deaths in 2018 and it is most prevalent in low- and middle-income countries. Enabling automation in the identification of potentially malignant and malignant lesions in the oral cavity would potentially lead to low-cost and early diagnosis of the disease. Building a large library of well-annotated oral lesions is key. In this paper we developed neural networks based oral cancer detection using different image processing techniques. For that we are using Convolutional Neural Network for the automated detection of oral lesions for the early detection of oral cancer. Oral cancer is a quite common global health issue. Early diagnosis of cancerous and potentially malignant disorders in the oral cavity would significantly increase the survival rate of oral cancer. Previously reported smartphone-based images detection methods for oral cancer mainly focus on demonstrating the effectiveness of their methodology, yet it still lacks systematic study on how to improve the diagnosis accuracy on oral disease using hand-held smartphone photographic images. We conducted a retrospective study. First, a simple yet effective centered rule image-capturing approach was proposed for collecting oral cavity images. Then, based on this method, a medium-sized oral dataset with five categories of diseases was created, and a resampling method was presented to alleviate the effect of image variability from hand-held smartphone cameras. Finally, a recent deep learning network (HRNet) was introduced to evaluate the performance of our method for oral cancer detection.

Indexed Terms*—Automated object tracking drone, Object detection using OpenCV, Object detection, Face recognition.

I. INTRODUCTION

Oral cancer is one of the most common cancers worldwide and is characterized by late diagnosis, high mortality rates and morbidity. GLOBOCAN estimated 354,864 new cases and 177,384 deaths in 2018. Two-thirds of the global incidence of oral cancer occurs in low- and middle-income countries (LMICs), half of those cases are in South Asia. Tobacco use, in any form, and excessive alcohol use are the major risk factors for oral cancer. A factor most prominent in South and Southeast Asia is the chewing of betel quid which generally is comprised of areca nut, slaked lime, betel leaf and may contain tobacco. Nowadays, these quids are available commercially in sachets and are popular in public due to vigorous marketing strategies. Oral cancer is typically associated with late presentation, particularly in LMICs, where more than two-thirds present at late stages and as a result survival rates are poor. Management of cancers, especially at the late stages, is very costly. The lack of public awareness and the lack of knowledge of health professionals concerning oral cancer is an important reason for late detection. Late diagnosis does not need to be a defining attribute as oral cancer is often preceded by visible oral lesions termed as oral potentially malignant disorders (OPMDs) which can be detected during routine screening by a clinical oral examination (COE) performed by a general dentist.

II. LITERATURE SURVEY

TOPIC : A novel lightweight deep convolutional neural network
AUTHOR : Fahed Jubair , Yazan Hassona
YEAR : 2021
DESCRIPTION:
To develop a lightweight deep convolutional neural network (CNN) for binary classification of oral lesions into benign and malignant or potentially malignant using standard real-time clinical images. A small deep CNN, that uses a pretrained EfficientNet-B0 as a lightweight transfer learning model, was proposed. A data set of 716 clinical images was used to train and test the proposed model. Accuracy, specificity, sensitivity, receiver operating characteristics (ROC) and area under curve (AUC) were used to evaluate

performance. Bootstrapping with 120 repetitions was used to calculate arithmetic means and 95% confidence intervals (CIs). Deep CNNs can be an effective method to build low-budget embedded vision devices with limited computation power and memory capacity for diagnosis of oral cancer. Artificial intelligence (AI) can improve the quality and reach of oral cancer screening and early detection.

TOPIC : Automatic classification and detection of oral cancer in photographic images using deep learning algorithms
AUTHOR : *KritsasithWarin , WasitLimprasert , SiriwanSuebnukarn , Suthin Jinapornthan , Patcharapon Jantana*
YEAR : 2021
DESCRIPTION:
Oral cancer is a deadly disease among the most common malignant tumors worldwide, and it has become an increasingly important public health problem in developing and low-to-middle income countries. This study aims to use the convolutional neural network (CNN) deep learning algorithms to develop an automated classification and detection model for oral cancer screening. The study included 700 clinical oral photographs, collected retrospectively from the oral and maxillofacial center, which were divided into 350 images of oral squamous cell carcinoma and 350 images of normal oral mucosa. The classification and detection models were created by using DenseNet121 and faster R-CNN, respectively. Four hundred and ninety images were randomly selected as training data. In addition, 70 and 140 images were assigned as validating and testing data, respectively. The DenseNet121 and faster R-CNN algorithm were proved to offer the acceptable potential for classification and detection of cancerous lesions in oral photographic images.

III. PROPOSED SYSTEM

In this paper deep learning-based computer vision approaches were assessed for the automated detection and classification of oral lesions for the early detection of oral cancer, these were image classification with ResNet-101 and object detection with the Faster R-CNN.

3.1.1 DISADVANTAGES OF EXISTING SYSTEM

- Deep learning can take more times to run.
- Deep learning needs some advanced hardware requirements for processing means high cost.

3.2 PROPOSED SYSTEM

In this paper computer vision based approaches were assessed for the automated detection of oral lesions for the early detection of oral cancer, these were image classification with Convolutional Neural Network is processed.

3.2.1 SCOPE OF THE PROJECT

The goal of oral cancer screening is to detect mouth cancer or precancerous lesions that may lead to mouth cancer at an early stage — when cancer or lesions are easiest to remove and most likely to be cured.

3.2.2 ADVANTAGES OF PROPOSED SYSTEM

- This process can take less time to run when compared to existing system.
- This technique needs normal hardware requirements followed by low cost and user friendly when compared to existing technique.

3.3 BLOCK DIAGRAM:

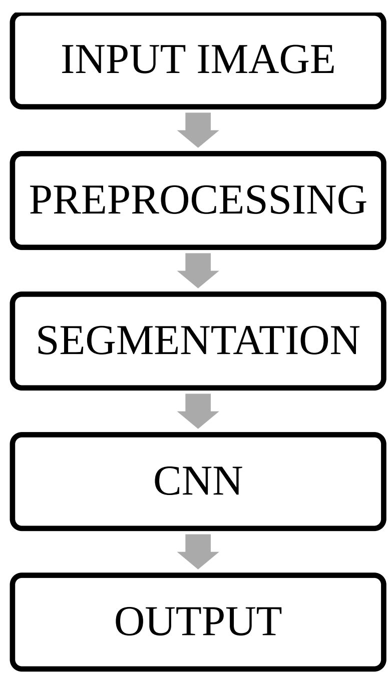


Figure 3.1: Block Diagram Of Proposed System

3.4 RGB COLOR IMAGE:

The RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue.

The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors as shown in fig 3.4.

Figure 3.4 : RGB Color Image

3.5 GRAYSCALE:

In photography and computing,
a grayscale or greyscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest.

Grayscale images are distinct from one-bit bi-tonal black-and-white images, which in the context of computer imaging are images with only the two colors, black, and white (also called bilevel or binary images). Grayscale images have many shades of gray in between. Grayscale images are also called monochromatic, denoting the presence of only one (mono) color (chrome) as shown in fig 3.5.

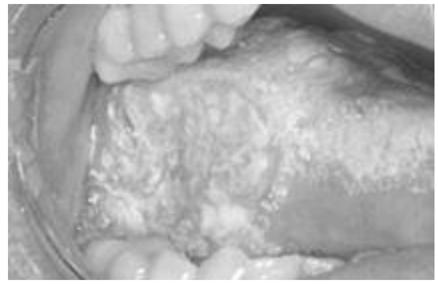


Figure 3.5: Gray Scale

3.6 MORPHOLOGICAL OPERATIONS:

To find the exact features we have to segment the lung region from the chest CT scan image for easy computation. For segmenting the lung region from the chest CT scan image morphological operation is carried out.

3.6.1 SMOOTHING:

This algorithm is based on the observation that a gray-level opening smoothes a gray-value image from above the brightness surface given by the function a[m,n] and the gray-level closing smoothes from below.

$$\begin{aligned} MorphSmooth(A,B) &= C_G(O_G(A,B),B) \\ &= \min(\max(\max(\min(A)))) \end{aligned}$$

Note that we have suppressed the notation for the structuring element B under the max and min operations to keep the notation simple. Its use, however, is understood.

3.6.2 GRADIENT:

For linear filters the gradient filter yields a vector representation with a magnitude and direction. The version presented here generates a morphological estimate of the gradient magnitude:

$$\begin{aligned} Gradient(A,B) &= \frac{1}{2}\left(D_G(A,B) - E_G(A,B)\right) \\ &= \frac{1}{2}\left(\max(A) - \min(A)\right) \end{aligned}$$

3.6.3 LAPLACIAN:

The morphologically-based Laplacian filter is defined by:

$$\begin{aligned} Laplacian(A,B) &= \frac{1}{2}\left(\left(D_G(A,B)-A\right)-\left(A-E_G(A,B)\right)\right) \\ &= \frac{1}{2}\left(D_G(A,B)+E_G(A,B)-2A\right) \\ &= \frac{1}{2}\left(\max(A)+\min(A)-2A\right) \end{aligned}$$

3.7 SEGMENTATION:

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like marching cubes.

3.8 CNN:

CNN stands for Convolutional Neural Network which is a specialized neural network for processing data that has an input shape like a 2D matrix like images.

CNN's are typically used for image detection and classification. Images are 2D matrix of pixels on which we run CNN to either recognize the image or to classify the image. Identify if an image is of a human being, or car or just digits on an address.

in that case applying the same formula, we get

(W−F+2P)/S+1 => (5–3 +2)/1 + 1=5,

now the dimension of output will be 5 by 5 with 3 color channels(RGB).

Let's see all this in action

If we have one feature detector or filter of 3 by 3, one bias unit then we first apply linear transformation as shown below

output= input*weight + bias

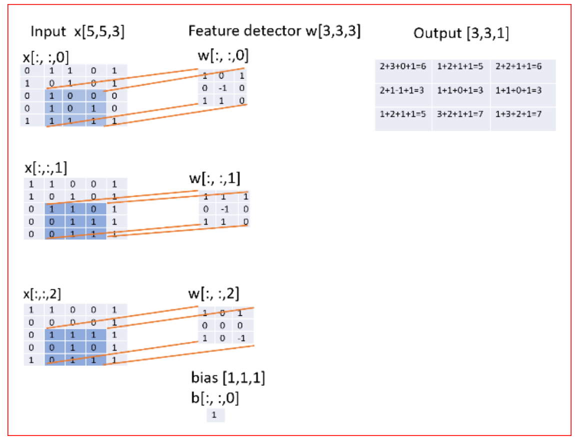


Fig 3.8 (a) : RGB Color Parameters

Input image of 5 by 5 with the three color channels and a feature detector or filter of 3 by 3 with a bias unit and stride is 1

No. of parameters = (3 * 3 * 3)+1 = 28

For 100 feature detectors or filters, number of parameters will 2800.

After every convolution operation which is a linear function, we apply ReLU activation function. ReLU activation function introduces non linearity in convolutional layer.It replaces all negative pixel values with zero values in the feature map.

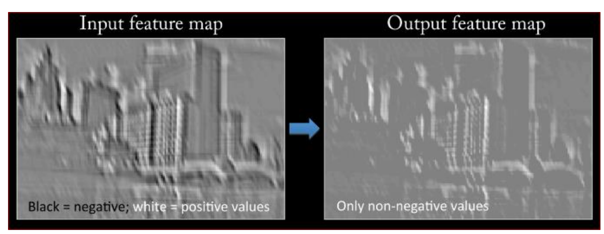


Fig 3.8(b) : ReLU Activation Function

Now that we have completed the feature detection from local areas we will combine all such feature detection from spatial neighborhood to build the picture.

Remember you are a detective scanning an image in dark, you have now scanned the image from left to right and top to bottom. now we need to combine all the feature to recognize the image.

3.8.1 POOLING:

we now apply pooling to have translational invariance.

Invariance to translation means that when we change the input by a small amount the pooled outputs do not change. This helps with detecting features that are common in the input like edges in an image or colors in an image

We apply the max pooling function which provides a better performance compared to min or average pooling.

when we use max pooling it summarizes the output over a whole neighborhood. we now have fewer units compared to the feature map.

In our example, we scan over all the feature maps using a 2 by 2 box and find the maximum value.

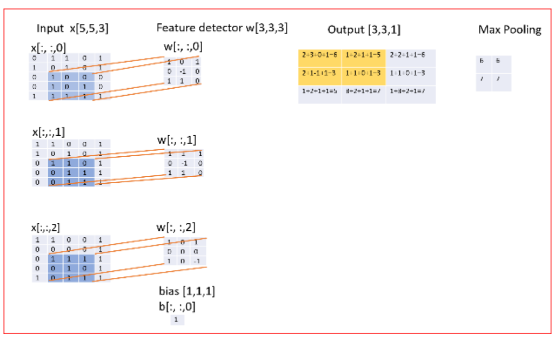


Fig 3.8.1 (a) : Feature Maps

Applying max pooling to the output using a 2 by 2 box.
Highlighted region in yellow has a max value of 6
so now we know that a convolutional network consists of

- Multiple convolutions performed in parallel — output is linear activation function
- Applying nonlinear function ReLU to the convolutional layers
- Use a pooling function like max pooling to summarize the statistics of nearby locations. This helps with “Translational Invariance”
- we flatten the max pooled output which are then inputs to a fully connected neural network

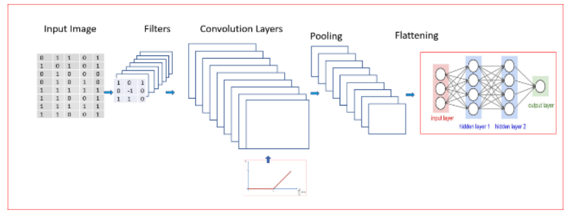


Fig 3.8.1(b): Full Convolutional Neural Network

3.9 APPLICATIONS:

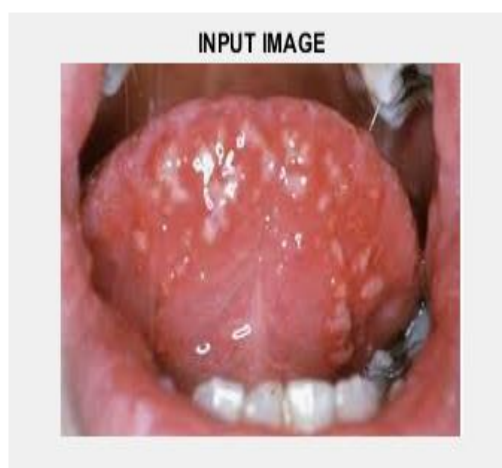
- To support early detection, diagnosis and optimal treatment of oral cancer.
- Image segmentation plays an essential role in many medical applications.
- To achieve robust and accurate segmentation.
- It is easy for doctors and lab technicians for detecting process.
- The detection process has less amount of cost.

IV. RESULTS AND DISCUSSION

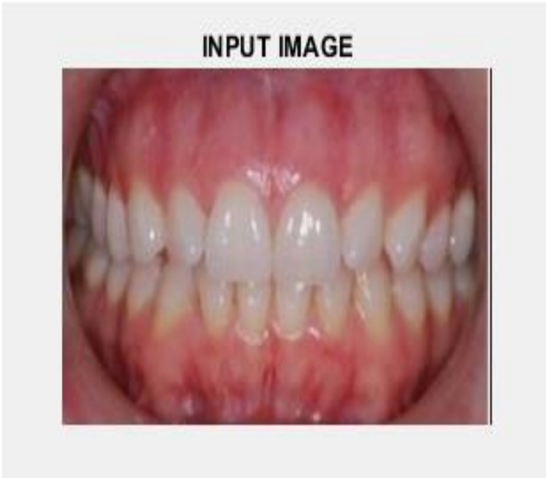
4.1 SNAPSHOT OF CANCER AND NON-CANCER IMAGES:

4.1.1 INPUT IMAGE:

Read and Display an input Image. Read an image into the workspace, using the imread command or camera. In image processing, it is defined as the action of retrieving an image from some source, usually a hardware-based source for processing. It is the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed as shown in fig 4.1.1.



Cancer Image



Normal Image
Figure 4.1.1: Input Image

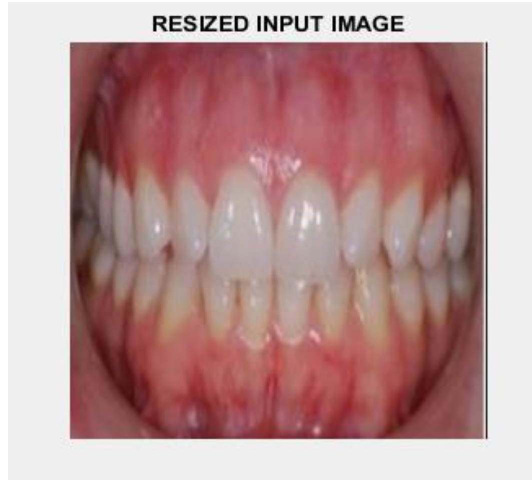


Figure 4.1.1(a) : Resized Image

4.1.2. PRE-PROCESSING IMAGE:

Pre-processing is a common name for operations with images at the lowest level of abstraction both input and output are intensity images. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing. Image pre-processing methods use the considerable redundancy in images. Neighboring pixels corresponding to one object in real images have essentially the same or similar brightness value. Thus distorted pixel can often be restored as an average value of neighboring pixels.

1. RESIZED IMAGE

All the input images are resized into same dimensions. If the specified size does not produce the same aspect ratio as the input image, the output image will be distorted as shown in fig 4.1.1 (a).

2. ENHANCED IMAGE:

Gamma correction, or often simply gamma, is a nonlinear operation used to encode and decode luminance or tristimulus values in video or still image systems for improving the image visibility as shown in fig 4.1.1(b).

Figure 4.1.1 (b) : Enhanced Image

3. GAUSSIAN FILTERED IMAGE:

A Gaussian filter is a linear filter. It's usually used to blur the image or to reduce noise. If you use two of them and subtract, you can use them for "unsharp masking" (edge detection). The Gaussian filter alone will blur edges and reduce contrast as shown in fig 4.1.1(c).

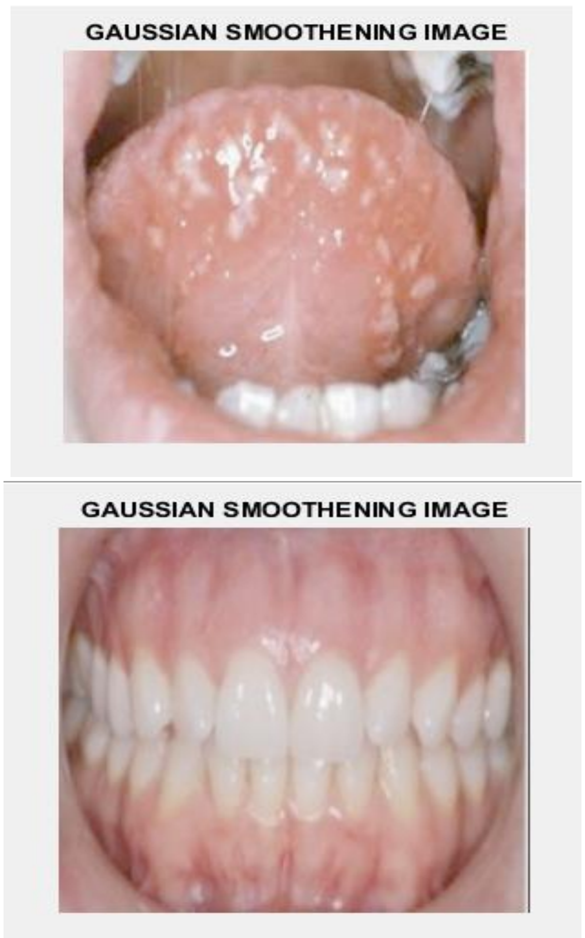


Figure 4.1.1 (c) : Gaussian Filtered Image

4.1.3. SEGMENTATION :

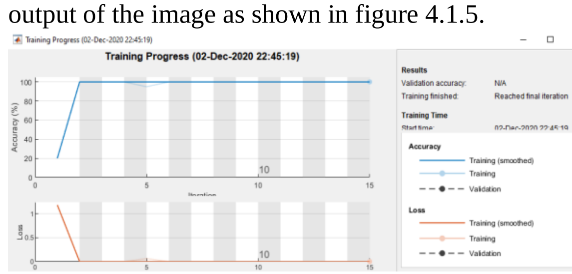
Image segmentation is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions, often based on the characteristics of the pixels in the image. In computer vision, Image Segmentation is the process of subdividing a digital image into multiple segments (sets of pixels, also known as super pixels. Segmentation is a process of grouping together pixels that have similar attributes. Image Segmentation is the process of partitioning an image into non-intersecting regions such that each region is homogeneous and the union of no two adjacent regions is homogeneous Pixels in a region are similar according to some homogeneity criteria such as color, intensity or texture so as to locate and identify objects and boundaries (lines,curves,etc) in an image. Segmentation accuracy determines the eventual success or failure of computerized analysis procedure.

4.1.4 CNN (CONVOLUTION NEURAL NETWORK):

Image segmentation is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions, often based on the characteristics of the pixels in the image. In computer vision, Image Segmentation is the process of subdividing a digital image into multiple segments (sets of pixels, also known as super pixels. Segmentation is a process of grouping together pixels that have similar attributes. Image Segmentation is the process of partitioning an image into non-intersecting regions such that each region is homogeneous and the union of no two adjacent regions is homogeneous Pixels in a region are similar according to some homogeneity criteria such as color, intensity or texture so as to locate and identify objects and boundaries (lines,curves,etc) in an image. Segmentation accuracy determines the eventual success or failure of computerized analysis procedure

4.1.5 OUTPUT:

Using the Matlab, we get the output is cancer or not cancer detected. We use for loop, if class coding that M-script code. First of all ,we train the code and after that ,selecting image from dataset image. We get output of the image as shown in figure 4.1.5.



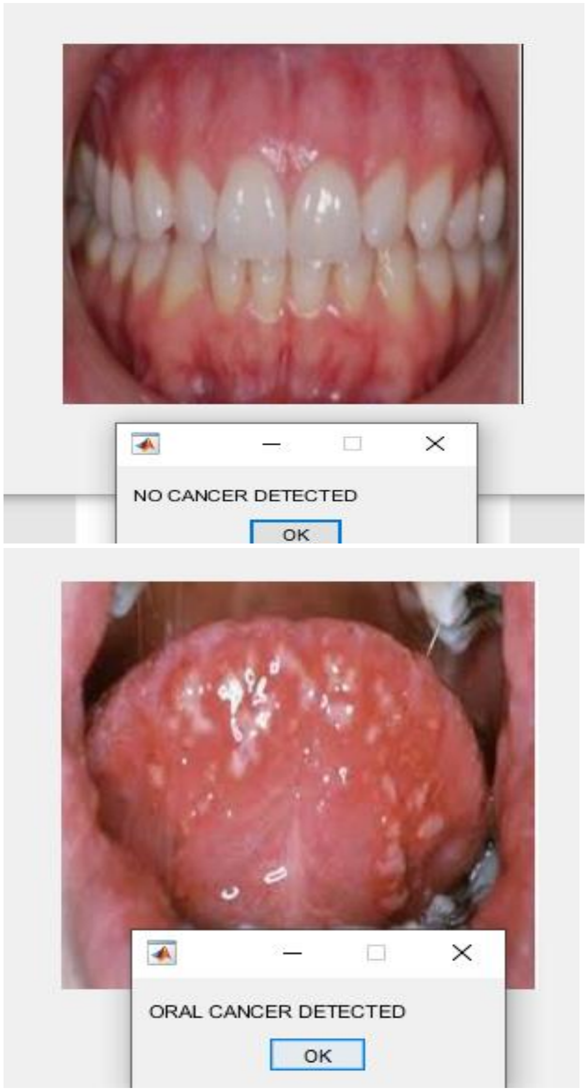


Figure 4.1.5: Output Image

V. CONCLUSION AND FUTURE SCOPE

In this paper, we propose an automatic method for the detection of oral cancer from the mouth images using image processing techniques. .In the experimental results, background and bones parts are distinguished very efficiently. The experimental results identified in this paper will positively helpful to the doctors for automated oral cancer identification system. In future, with more time and with more comprehensive research the proposed system can be made more accurate. Also new oral cancer detection algorithms can be added so as to give the doctor a wider variety of options to choose from.

Dental radiography analysis plays an important role in clinical diagnosis, treatment and surgery as radiographs can be used to find hidden dental structures, malignant or benign masses, dental cancer, bone loss and cavities. The goal of oral cancer screening is to detect mouth cancer or precancerous lesions that may lead to mouth cancer at an early stage — when cancer or lesions are easiest to remove and most likely to be cured.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal,(2018) ‘Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries’ ,CA: A Cancer J. Clinicians, vol. 68, no. 6, pp. 394-424.

[2] J.Rimal, A.Shrestha, I.K.Maharjan, S.Shrestha, and P.Shah, (2019)’Risk assessment of smokeless tobacco among oral precancer and cancer patients in eastern developmental region of nepal’, Asian Pacic J. Cancer Preven- tion, vol. 20, no. 2, pp. 411-415.

[3] J. G. Doss,W. M. Thomson, B. K. Drummond, and R. J. R. Latifah,(2011) ‘Validity of the FACT-H&N (v 4.0) among Malaysian oral cancer patients’,Oral Oncol., vol. 47, no. 7, pp. 648-652.

[4] G M. Aubreville, C. Knipfer, N. Oetter, C. Jaremenko, E. Rodner, J. Denzler, C. Bohr, H. Neumann, F. Stelzle, and A. Maier,(2017), ‘Automatic classication of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning’, Sci. Rep., vol. 7, no. 1, p. 119-279.

[5] H. Amarasinghe, R. D. Jayasinghe, D. Dharmagunawardene, M. Attygalla, P. A. Scuffham, N. Johnson, and S. Kularatna, (2019),’Economic burden of managing oral cancer patients in sri lanka: A cross-sectional hospital - based costing study,’BMJ Open, vol. 9, no. 7, Art. no. e027661.

[6] R. D. Jayasinghe, L. P. G. Sherminie, H. Amarasinghe, and M. A. Sitheeque, (2016),’Level of awareness of oral cancer and oral potentially malignant disorders among

medical and dental undergraduates', Ceylon Med. J., vol. 61, no. 2, p. 77.

[7] M. M. R. Krishnan, V. Venkatraghavan, U. R. Acharya, M. Pal, R. R. Paul, L. C. Min, A. K. Ray, J. Chatterjee, and C. Chakraborty, (2012), ``Automated oral cancer identication using histopathological images: A hybrid feature extraction paradigm',Micron, vol. 43, nos. 23, pp. 352364.

[8] *N. Haron, R. B. Z*ain, W. M. Nabillah, A. Saleh, T. G. Kallarakkal, A. Ramanathan, S. H. M. Sinon, I. A. Razak, and S. C. Cheong, (2017), 'Mobile phone imaging in low resource settings for early detection of oral cancer and concordance with clinical oral examination', Telemed. e-Health, vol. 23, no. 3, pp. 192-199.

[9] O. Kujan, A.-M. Glenny, R. Oliver, N. Thakker, and P. Sloan, (2009),'Screening programmes for the early detection and prevention of oral cancer', Austral. Dental J., vol. 54, no. 2, pp. 170-172.

[10] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J.Winn, and A. Zisserman,(2015), 'The Pascal visual object classes challenge: A retrospective', Int. J. Comput. Vis., vol. 111, no. 1, pp. 98-136.