A Review on Imaging Modalities and Techniques for Oral Malignancy Detection

Nanditha B R, Geetha Kiran A
Malnad College of Engineering, Hassan, Karnataka, India.

ABSTRACT

One of the most prevalent forms of cancers across the world is oral cancer which has a high rate of mortality mainly due to late diagnosis. Early diagnosis and proper medication provide the best chance of cure for oral cancer patients. Different imaging modalities have been developed by researchers for oral lesion screening and various image analysis techniques have been proposed for analyzing the oral images for malignancy detection. This paper aims to review various works in the literature related to different techniques for oral lesion screening and analysis of oral images for the detection of oral malignancies.

Key Words: Malignant, Benign, Oral cancer, Lesions, Texture features

INTRODUCTION

Cancer is a condition of the human body where an uncontrolled division of abnormal cells takes place. Among all types of cancers worldwide, oral cancer takes the sixth position. It accounts for approximately 4% of all cancers and 2% of all cancer deaths worldwide. In India it is the common malignant neoplasm, accounting for 20-30% of all cancers. There are many risk factors for oral cancer that include extensive use of tobacco, alcohol consumption, smoking and others. Normally, oral cancer patients visit doctors at a later stage because most of the lesions are asymptomatic and they do not affect the normal day-to-day activities of the patients. Late diagnosis of oral cancer has led to increasing in mortality rate which can be reduced with timely diagnosis and proper treatment.

Oral cancer normally occurs in the oral cavity or throat. Different kinds of lesions occur in the oral cavity. These can be benign, premalignant or malignant lesions. Figure 1 elucidates different kinds of oral lesions. A premalignant lesion has the potential to develop into a cancerous lesion. The most common oral precancerous lesions are oral leukoplakia, oral submucous fibrosis, oral erythroplakia, and Lichen Planus. Benign lesions include lichenoid reactions, Candidiasis, pemphigus Vulgaris, aphthous ulcers, mucocele and others. These lesions have different textures, colours, shapes and sizes. Normally doctors suggest a biopsy and histopathological examinations for confirming the specific kind of lesion. The biopsy is an invasive, time-consuming and expensive procedure. Many times patients are unwilling to undergo a biopsy since it is painful. In low socio-economic places where there is no proper access to medical facilities and resources, it may not be possible for each individual to undergo a biopsy examination. Thus, it may not be possible to identify premalignant lesions at an early stage and hence might lead to a high fatality rate. In such an environment, it would be desirable to develop imaging models and techniques which are less expensive and easily accessible to all sections of people. Many researchers have been working towards oral malignancy identification by developing various oral screening models for capturing oral images and image analysis techniques for analyzing the oral images for identifying malignancies. The following sections present a detailed description of these methods.

IMAGING MODALITIES FOR ORAL CANCER SCREENING

Normally, doctors use intraoral cameras for visualizing the oral cavity of patients during the routine clinical examination which is shown in Figure 2. A novel smartphone-based
imaging modality which uses an intraoral camera for the detection of oral cancer has been proposed by Bofan Song et al.\textsuperscript{7} in which an integration of two modes i.e autofluorescence light and white light has been used for capturing images of the oral cavity. This modality is useful in low-resource settings where mass screening can be carried out by social workers to screen high-risk populations. The dual-mode images are analyzed by using cloud-based image processing software and it can be communicated to remote specialists to decide the existence of malignancy.

A multimodal optical imaging system that combines white-light images, autofluorescence images and microendoscopic images for evaluating oral premalignant lesions in a non-invasive manner has been proposed by Eric C.Yang et al.\textsuperscript{8}. This imaging system captures white light and autofluorescence images and generates a heat map that delineates suspicious regions based on the red-green ratio of the images. A high-resolution microendoscopic device that has higher specificity than fluorescence imaging is used to explore the suspicious regions for identifying the biopsy sites. This system is very useful in identifying the site of biopsy for suspicious lesions and avoiding biopsy for benign lesions.

A surgical headlight system has been modified to act as a visualization and imaging device that can capture images of the oral cavity by integrating multiple modalities like fluorescence imaging, white-light imaging and orthogonal polarization. The device is battery powered and includes LED lights for visualization of the oral cavity and acquisition of images both in white light and fluorescent light. It is a low-cost and portable device and hence can serve as a promising device for oral cancer screening in low-resource settings.

A simple hand-held device that uses fluorescent reflectance for oral cavity visualization directly by the human observer or by camera recording for detecting oral malignancies has been proposed by Lane et al.\textsuperscript{9}. This device makes use of blue light illumination for detecting tissue changes between normal and malignant lesions. This device can act as an adjunct to white-light reflectance for oral cancer screening and also it serves as a guiding device for identifying biopsy sites.

Optical instruments serve as useful tools for screening oral cancer because of their non-invasive nature and also due to the ease of use\textsuperscript{11}. With the intent of facilitating minimum invasiveness in oral screening, the evaluation of optical instrument oralook has been considered by Morikawa et al.\textsuperscript{12} to demonstrate its usefulness as a convenient real-time device for oral cancer visualization. The subjective and objective evaluations are performed with fluorescent visualization images captured using this device which showed that optical instruments serve as better adjunct devices for oral cancer screening. One more useful screening device for differentiating cancerous and normal lesions through the use of fluorescence images is illumiscan\textsuperscript{13}. This optical instrument has been used for differentiating squamous cell carcinoma from oral leukoplakia by analyzing the images captured using illumiscan. Results indicate that there is a significant difference in the luminance values of leukoplakia and squamous cell carcinoma. This device can be integrated with conventional screening techniques to serve as a better tool for oral cancer screening.

Apart from the visualization and image acquisition devices described so far, other imaging modalities are useful for staging and grading oral cancer. These include computer tomography (CT)\textsuperscript{14}, magnetic resonance imaging (MRI)\textsuperscript{15}, positron emission tomography (PET) and others.

Another popular imaging technique used for the diagnosis of squamous cell carcinoma is confocal laser endomicroscopy\textsuperscript{16} which has high rates of magnification and enables penetrating deep into the tissues for better diagnosis of oral malignancies.

TRADITIONAL TECHNIQUES TO ANALYSE ORAL IMAGES FOR MALIGNANCY DETECTION

Literature reveals the use of traditional image processing and machine learning techniques for oral image analysis to identify oral malignancies. Researchers have tried to develop different methods for analyzing oral images captured using various imaging modalities and acquisition devices. Table I lists the different types of images used by various researchers for oral malignancy detection using various techniques. These images include standard white light images, fluorescent images, hyperspectral images, microscopic images and others.

Images need to be analyzed for concluding whether there are malignancies present in the oral images or not. Different image analysis techniques have been developed that make use of image features for identifying whether they are normal or abnormal images.

Features serve as discriminating factors for identifying oral lesions as benign or malignant. Various features like colour, texture and shape have been identified by researchers as good differentiating features for analyzing oral images. A brief description of different features used for oral image analysis is provided in the following sections.

Gray Level Co-Occurrence Matrix (GLCM)

GLCM features are used to measure the variation in intensity at a particular pixel of interest. The GLCM specifies how frequently different combinations of grey levels co-occur in an image.\textsuperscript{17} The pairwise spatial co-occurrences of pixels separated by a particular angle and distance are computed by using GLCM. The GLCM texture features are contrast, correlation, dissimilarity, homogeneity, energy, entropy, mean,
variance and standard deviation. Table 2 lists all GLCM features and their characteristics that are useful in analyzing oral images for detecting malignancies.

**Gray Level Run Length Matrix (GLRLM)**
A run is constituted by adjacent pixels with the same intensity in a particular direction. The GLRL matrix is a two-dimensional tabulation of such runs against quantized pixel intensities. A coarse texture would normally result in relatively longer runs while comparatively shorter runs would correspond to fine textures. The GLRL features are short-run emphasis (SRE), long-run emphasis (LRE), low grey-level run emphasis (LGRE), high grey-level run emphasis (HGRE), grey-level non-uniformity (GLN) and run-length non-uniformity (RLN). Researchers have used GLCM and GLRL features for the identification and classification of benign and malignant lesions.

Microscopic images of oral lesions have been classified into benign and malignant lesions by using GLCM and histogram techniques. They have made use of 134 images of normal tissues and 135 images of malignant tissues. Six first-order texture features namely mean-variance, skewness, kurtosis, energy and entropy have been used. T-test and principal component analysis (PCA) methods have been used for selecting significant features. They achieved 100% accuracy with a linear support vector machine (SVM) classifier.

Intensity-based first-order statistical features, GLCM and GLRL features have been used for multiclass classification of colour oral images by using backpropagation based artificial neural network classifier. The boxplot analysis method has been used to select 11 useful features for classification. Classification accuracy of 97.92% has been achieved.

Texture features, shape and morphology features, histogram oriented gradient (HOG) features, wavelet colour features, Tamura’s features and Law’s texture energy (LTE) features have been used for automated detection and classification of cancer from microscopic biopsy images. KNN classifier was used and an accuracy of 92.19% was achieved.

J.V Raja et al. made use of fractal features and GLCM features like angular second moment (ASM), contrast, inverse difference moment (IDM) and entropy for identifying oral cancers from CT images. They found that the lesion group recorded higher mean fractal dimension, ASM, contrast and IDM than in the normal group.

A combination of first-order statistical features, GLCM features and GLRL features has been used for classifying cyst and tumour lesions from dental panoramic images. Support vector machine has been used for classification. It is found that the combination of first-order statistical features and GLRLM achieved the highest accuracy of 94.44%.

**Fractal Features**
Fractal features are found to be useful for measuring the texture of images. Fractals are found to be useful in describing the geometric structure of objects. They give information about the roughness or smoothness of a region and describe the heterogeneity of a region in an image.

The fractal dimension of an image correlates with its roughness and when used for image analysis, results indicate that the fractal dimension of an image increases with an increase in noise in the image.

Unsupervised classification of textured images is possible by calculating the fractal dimension (FD) using the differential box-counting method. Different fractal features namely, FD of the original image, high grey valued image, low grey valued image, horizontally smoothed image and vertically smoothed image perform the texture-based classification.

**Gabor Features**
Gabor features are very much suitable for the analysis of images since they contain a lot of texture information. Gabor filters have been extensively used for texture description in various image processing and analysis approaches. They decompose an image into multiple scales and orientations analyzing texture patterns more straightforward.

Gabor features find their use in many image analysis applications which convey local information in an image at different frequencies and orientations. Different texture features have been deployed for classifying oral histopathological images into normal and oral submucous fibrosis. Gabor features and fractal features are also among these features that achieve good results in performing the classification.

A computer-aided diagnosis (CAD) system for differentiating normal and abnormal thyroid nodules in biopsy images using wavelet features, grey level features and Gabor features has been proposed. The background is segmented from the foreground objects in the cytology images by the application of morphological transformation and watershed segmentation techniques. Then, statistical features are used for training different classifiers and an accuracy of 93.33% was achieved by the use of a neural network classifier trained with Gabor features.

**Colour Features**
Colour is an important feature that is found to be useful in analyzing images. Colour features are the basic characteristic of the content of images.

Different colour features have been used to discriminate oral malignancies such as oral lichenoid reactions from oral leu-koplakia. The oral images are represented in five different colour representations namely, Red-Green-Blue RGB, Irg, Hue-Saturation-Index (HSI), I11213 and La*b* have
been used for analysing oral cavity images. Results show that the HSI colour system provided the best classification accuracy. 70 out of 74 lichenoid reactions and 14 out of 20 leukoplakia were correctly classified.

A semi-automatic method for the detection of oral lesion boundaries has been proposed where colour images are analyzed by converting them into single-band images. Several single bands were derived from original three-band images: R, G and B bands from RGB colour space, H, S and I bands from Hue-Saturation-Index (HIS) colour space, Rn, Gn and Bn bands from normalized RGB colour space, I1, I2 and I3 bands from II1I2I3 colour space. An active contour model is applied on each of these bands for detecting the boundaries of lesions.

A three-stage Color-Based Feature Extraction (CBFE) system has been proposed which includes colour normalization, automatic feature extraction and PCA learning algorithm and classification. Two transformed images are computed based on the hue component in the normalized image namely, normalized red component image and normalized blue component image. They have used a dataset of 79 microscopic images and achieved more than 90% classification accuracy.

RGB and HSV colour representation modalities have been used for discriminating normal tissues from two potentially premalignant lesions, oral lichen planus and oral submucous fibrosis. Classification accuracy of 78% was achieved by employing the HSV colour space for image analysis.

**Local Binary Patterns**

Local binary pattern (LBP) is one of the useful methods for feature extraction. The original LBP operator uses a 3X3 square neighbourhood centred at the given pixel. The algorithm assigns either 0 or 1 to the 8 neighbouring pixels according to the equation

\[ N = \begin{cases} 
0 & \text{if } g_N < g_C \\
1 & \text{if } g_N \geq g_C 
\end{cases} \]

where \( N \) is the binary value assigned to the neighbouring pixel, \( g_N \) denotes the grey-level value of the neighbouring pixel and \( g_C \) is the grey-level value of the centre pixel. The resulting values are then concatenated into an 8 bit binary number. Its decimal representation is used for further computation. The global shape and the local texture of the images are obtained as a result of LBP operation.

The use of higher order spectra features (HoS) and local binary pattern (LBP) features has been evaluated for classifying normal, oral submucous fibrosis with dysplasia and oral submucous fibrosis without dysplasia. Twenty three HoS features and nine LBP features have been extracted and fed to a support vector machine (SVM) for automated diagnosis. Results show that LBP features provide a good sensitivity (82.85%) and specificity (87.84%), and the HoS features provide higher values of sensitivity (94.07%) and specificity (93.33%).

Seven features have been extracted from lesion images and the lesions have been classified as benign or malignant lesions. These features are perimeter, area, diameter, fractal dimension, lacunarity, histogram of oriented gradients, and local binary patterns. A classification system that was able to diagnose with an accuracy of 85% has been developed.

Different variants of LBP features, namely average LBP (ALBP), block-based LBP (BLBP), ELBP (Elliptical Local Binary Pattern), Uniform ELBP, LDP (Local Directional Pattern) and M-ELBP (Mean-ELBP) can also be used for classification tasks which can capture intrinsic and detailed micro-pattern features from images.

All these features have been deployed for classifying oral lesions into normal or abnormal lesions by using traditional machine learning algorithms like artificial neural networks, support vector machines, bayesian classifiers, decision tree classifiers, principal component analysis, linear discriminant analysis.

**DEEP LEARNING TECHNIQUES FOR ORAL MALIGNANCY DETECTION**

Deep learning is gaining popularity as an adjunctive tool for disease diagnosis and treatment to aid doctors in diagnosing various kinds of disorders. Many deep learning techniques for oral cancer diagnosis have been developed by researchers. Jeyaraj et al. have developed a deep learning method for detecting oral cancer from hyperspectral oral images. Deep convolution neural network (CNN) architecture with two partitioned layers, one for labelling and the other for classifying the labelled region has been proposed which achieved 91.4% classification accuracy with 94% sensitivity and 91% specificity.

Confocal laser endomicroscopic images have been classified into cancerous and non-cancerous by using deep convolution neural networks. LeNet-5 network has been used for building the CNN where the first layer is a convolution layer that consists of 64 filters sized 5x5 pixels, the next layer is the max-pooling layer of 3x3 pixels and one more is a convolution layer containing 32 filters sized 5x5 pixels, followed by a max-pooling layer of size 3x3 pixels and a fully connected layer with drop-out and an output layer. This method shows better performance compared to conventional classification techniques like texture-based methods.

Two deep learning techniques have been assessed by Roshan Alex et al for the early detection of oral cancerous lesions.
One deep learning method used is for classifying images using ResNet-101 and the second technique is the deployment of R-CNN for detecting lesions. ResNet-101 is a deep neural network with 101 layers and is used widely for various applications. Transfer learning has been used where ResNet has been pre-trained on the ImageNet database. The faster R-CNN has also been trained with the COCO database and data augmentation is also performed to increase the size of the dataset. Deep learning techniques used in this work for oral lesion detection and classification have shown promising results.

An open-source convolutional neural network toolbox for analysis of autofluorescence and white light images has been used for the classification of oral lesions into cancerous and normal lesions. To increase the size of a dataset, transfer learning has been used with the ImageNet database and also data augmentation techniques like flipping and rotation have been used. Results indicate that deep learning techniques are effective in classifying oral lesions from multiple modalities.

Hyperspectral images have been classified into squamous cell carcinoma (SCC), thyroid cancer and normal tissues using deep learning architecture. A convolutional neural network has been implemented using TensorFlow for the classification of cancerous and non-cancerous tissues. This architecture includes 6 convolutional layers, 3 fully connected layers and a softmax layer that generates the probability with which every pixel belongs to a class. The system classified cancerous and non-cancerous tissues with an accuracy of 80%.

Identification of oral precancerous lesions to classify them as benign or malignant using six different deep learning techniques has been developed. Of these techniques, a deep learning architecture based on vgg19 differentiated benign and malignant lesions with an accuracy of 98%. Another convolution neural network model based on resnet50 architecture could make a multiclass classification of oral lesions with an accuracy of 97%.

**DISCUSSION**

An attempt has been made in this paper to review the existing research work carried out by different researchers in analysing oral images for the identification of malignancies. Different kinds of images like white light images, fluorescent images, hyperspectral images, microscopic images have been used in the literature for analysing the abnormalities in the images. Different features have been used for the image analysis to classify them as benign or malignant based on the differences in their texture, colour and fractal features. These are found to be good discriminators of different kinds of lesions. Also, the use of different imaging modalities is very effective in capturing oral images for facilitating their analysis by the use of various image processing and deep learning techniques.

**CONCLUSION**

One of the serious health issues being faced by many developing countries is that of the high incidence of oral cancer. Early diagnosis and treatment of oral cancer can increase the survival rate of patients. An effort has been made in this paper to review different imaging modalities that have been devised for oral cancer screening and also different image processing and machine learning approaches that are available for detecting oral cancer. Various texture analysis methods such as GLCM, GLRL, Gabor features, fractal features and local binary patterns have been explored for image analysis. Also, the use of different colour features for oral lesion analysis has been reviewed. These features are very useful in identifying malignancies and also in classifying images into different types of lesions. Also, various deep learning techniques have been developed which are found to be very useful in identifying oral malignancies. In addition to all the existing screening devices and techniques for oral cancer detection, novel imaging modalities and image analysis methods are still required for the early detection of oral malignancies.

**ACKNOWLEDGEMENT**

The authors acknowledge the immense help received from the scholars whose articles are cited and included in references of this manuscript. The authors are also grateful to authors/editors/publishers of all those articles, journals and books from where the literature for this article has been reviewed and discussed.

**Financial Support**

No financial support has been obtained for this work.

**Conflict of interest**

Authors have no conflict of interest.

**Authors contribution**

All authors have substantially contributed to the conception and design of the manuscript and interpreting the relevant literature. The authors have drafted the manuscript and revised it carefully for important intellectual content.

**REFERENCES**


Figure 1: Different types of oral lesions: a) Benign lesions b) Premalignant lesions c) Malignant lesions.

Figure 2: Intra-oral camera used for oral cavity visualization.

Table 1: Different type of images used for analysis of oral malignancy

<table>
<thead>
<tr>
<th>Name of the system</th>
<th>Type of images used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histopathological analysis of oral cancer</td>
<td>Histopathological images</td>
</tr>
<tr>
<td>True color RGB image analysis using neural networks</td>
<td>Standard white light images</td>
</tr>
<tr>
<td>Classification of laser endo-microscopic images of oral cavity using deep learning techniques</td>
<td>Endomicroscopic images</td>
</tr>
<tr>
<td>Head and neck cancer detection using deep learning techniques</td>
<td>Hyperspectral images</td>
</tr>
<tr>
<td>Classification of orofacial diseases using deep learning techniques</td>
<td>Standard white light images</td>
</tr>
<tr>
<td>Deep learning techniques for oral cancer tissue classification</td>
<td>Histopathological images</td>
</tr>
<tr>
<td>Smartphone-based imaging modality for classification of oral malignancies</td>
<td>Autoflourescence and white light images</td>
</tr>
<tr>
<td>A framework for early detection of oral dysplasia</td>
<td>Histopathological images</td>
</tr>
<tr>
<td>CAD for classification of oral cancer employing deep learning algorithm</td>
<td>Hyperspectral images</td>
</tr>
<tr>
<td>3D-CNN for early diagnosis of oral cancer</td>
<td>CT images</td>
</tr>
<tr>
<td>Feature</td>
<td>Use in image analysis</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Contrast</td>
<td>Gives a measure of local variations in the GLCM</td>
</tr>
<tr>
<td>Correlation</td>
<td>Gives a measure of probability of occurrence of the given pair of pixels</td>
</tr>
<tr>
<td>Angular Second Moment (ASM)</td>
<td>Gives a measure of how similar the pixels are in a GLCM</td>
</tr>
<tr>
<td>Homogeneity</td>
<td>Gives a measure of how closely distributed the elements are in a GLCM</td>
</tr>
<tr>
<td>Entropy</td>
<td>Gives a measure of the amount of information and randomness of intensity distribution</td>
</tr>
<tr>
<td>Inverse Difference Moment (IDM)</td>
<td>Gives a measure of how closely distributed the GLCM elements are to the GLCM diagonal</td>
</tr>
<tr>
<td>Sum Average (SA)</td>
<td>Gives a measure of the average of the gray level sum distribution in an image</td>
</tr>
<tr>
<td>Sum Variance (SV)</td>
<td>Gives a measure of the dispersion of the gray level sum distribution in an image</td>
</tr>
<tr>
<td>Sum Entropy (SE)</td>
<td>Gives a measure of the disorders of the gray level sum distribution in an image</td>
</tr>
<tr>
<td>Difference Entropy</td>
<td>Gives a measure of the disorders of the gray level difference distribution in an image</td>
</tr>
</tbody>
</table>