Artificial Intelligence-based Oral Cancer Screening System using Smartphones

Parnasree Chakraborty
Electronics & Communication Engineering, BSA Crescent Institute of Science & Technology, India
prernasree@crescent.education (corresponding author)

Tharini Chandrapragasam
Electronics & Communication Engineering, BSA Crescent Institute of Science & Technology, India
tharini@crescent.education

Ambika Arunachalam
Electronics & Communication Engineering, BSA Crescent Institute of Science & Technology, India
ambika@crescent.education

Syed Rafiammal
Electronics & Communication Engineering, BSA Crescent Institute of Science & Technology, India
rafiammal@ crescent.education

Received: 2 September 2023 | Revised: 18 September 2023 | Accepted: 26 September 2023

 DOI: https://doi.org/10.48084/etasr.6364

ABSTRACT

About one-fifth of all oral cancer cases reported globally are from India. The low-income groups in India are affected most due to the wide exposure to risk factors such as tobacco chewing and insufficient access to early diagnostic tools. Visual examination and histological study are the standard for oral cancer detection. This paper proposes the idea of using Autofluorescence-based imaging techniques to detect and classify oral cancer using AI algorithms. Various features of the images along with medical history, age, gender, and tobacco usage are considered as inputs to the proposed Mobilenet classification architecture.

Keywords-oral cancer; AI; autofluorescent images; mobilenet architecture

I. INTRODUCTION

Cancer detection is an important research topic in medical science. Globally, oral is the sixth most common cancer, whereas there is not an effective tool for early diagnosis and treatment. In this paper, we propose an approach for cancer detection based on image analysis using Artificial Intelligence (AI). The main objective of the proposed idea is to collect a dataset from Indian patients and to analyze it using smart phone assisted AI algorithms. The goal is to collect and generate big data of autofluorescence images of the oral cavity and analyze the data using the Lightweight Mobilenet architecture for accurate classification of images for oral cancer detection while developing an android-based application to capture the image and visualize the output.

A portable laser device for oral cancer detection has been developed by the Mazumdar Shaw Cancer Centre (MSCC) [1] in Bangalore, India. The work was funded by the National Institute of Health's National Institute of Biomedical Imaging and Bioengineering, under the Indo-US collaborative program. These screening devices are costly and hence there is a need for low cost accurate devices. Several other commercial instruments have been developed for routine examination worldwide [2-5].

II. METHODOLOGY

The advent of information technology and its modern derivative AI can improve oral cancer screening, but to date, only a few efforts have been made to apply these techniques and relatively little research has been conducted to retrieve meaningful information from AI data. The suggested system consists of a handheld device (smartphone) which will illuminate the fluorescent light of wavelength 400 – 460 nm on mucosa. Normal mucosa emits a green autofluorescence when exposed to this fluorescent light due to the presence of naturally occurring fluorophores. Abnormal mucosa appears dark due to the reduction or change in the quantity and quality of fluorophores. Autofluorescence digital images will then be captured and transmitted through Bluetooth to hardware in which the data will be classified for oral cancer detection. This work will be carried out in two phases. The first phase involves data collection and the second phase involves implementation of AI algorithms in smart phones.

The detection network will consider an oral photograph as input and will generate one bounding box that will be located in the suspected lesion. The lesion area will be cropped as a candidate patch according to the detection results returned by the first step. The candidate patch is then fed to a classification network. The backbone networks of detection and classification work on a pre-trained model. The overall block diagram and process flow is shown in Figure 1. Epidemiology history will be collected for statistical analysis. A scoring system will be used to assign weightage for oral cancer detection [8]. The scoring system uses weighted inputs such as image, dataset of age, gender, smoking habits etc.

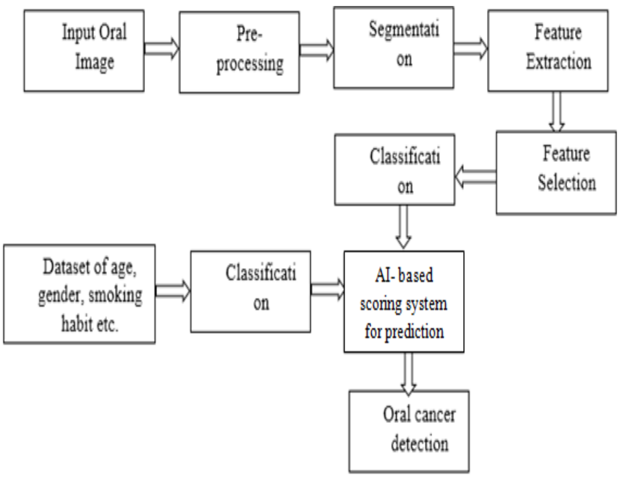

Fig. 1. The proposed system.

The development of an AI-based oral cancer screening system using a smartphone shown in Figure 1 involved several steps. First, data collection of a diverse dataset of oral cavity images, correctly labeled for cancerous and non-cancerous cases. Next, data preprocessing and augmentation techniques were applied to standardize and enhance image quality. Transfer learning was employed with pre-trained AI models like MobileNet, followed by fine-tuning for oral cancer classification. The model was then trained on the dataset, with hyperparameter tuning to optimize performance. Evaluation metrics such as accuracy, sensitivity, and specificity were used to assess the model's accuracy. Deployment on smartphones involves converting the model to mobile-friendly formats, while user-friendly interfaces are developed for image capture and results display. Continuous improvement through user feedback and updates ensures the system's effectiveness. Compliance with privacy and regulatory standards was paramount throughout the development process.

III. THE ORAL CANCER DATASET

The dataset used for this work is taken from [7]. It contains images of lips and tongue which are classified as cancerous and non-cancerous (Figure 2). The captured images originate from several distinct ENT (Ear, Nose, and Throat) hospitals, forming a diverse collection. The process of categorizing these images into their respective groups was undertaken in close collaboration with skilled ENT doctors who lent their expertise. Their specialized knowledge and insights were instrumental in ensuring the accurate classification of these images, which is paramount for any medical-related application, particularly in the context of ENT healthcare. This collaborative effort between the medical professionals and the image processing team underscores the importance of interdisciplinary cooperation in delivering high-quality healthcare solutions. The dataset contains a total of 175 images, from which 131 were classified as cancerous and 44 as non-cancerous images. A 75% of the total data was used for training and the remaining 25% for testing.

(a) (b)

Fig. 2. (a) Cancerous and (b) non-canserous sample images.

IV. MOBILENET BASED CLASSIFICATION OF ORAL CANCER IMAGES

MobileNetV2 is a Convolutional Neural Network (CNN) architecture. It is specifically designed for efficient image classification and feature extraction tasks. It utilizes a series of depth-wise separable convolutions, which are a key component of many modern CNN architectures designed for mobile and embedded devices. MobileNet is a popular architecture for deep learning and is well-suited for image classification tasks, especially when computational resources are limited, such as on mobile devices. To use MobileNet [10] for the classification of oral cancer images, these steps were followed:

- Data Collection and Preparation: A dataset of oral cancer images, labeled with the corresponding classes (e.g. cancerous, non-cancerous) was collected and split into training, validation, and test sets.
- Data Augmentation: Data augmentation techniques like rotation, scaling, and flipping were applied to increase the diversity of the training dataset.
- Preprocessing: Preprocessing was carried out for the images, typically by resizing to a fixed input size (e.g. 224 × 224 pixels). Normalization was done for the pixel values to have zero mean and unit variance.
- MobileNet Architecture: MobileNetV2 was chosen for its balance between performance and efficiency.
- Transfer learning: Transfer learning was carried out by starting with a pre-trained MobileNet model.
- Fine-Tuning: Modifications were carried to the MobileNet architecture by adding a custom output layer with the

number of units corresponding to the classification classes. This output layer uses a suitable activation function (e.g. softmax for multi-class classification) and connects to the last layers of the MobileNet architecture.

- Training: Training of MobileNet model was conducted in 75% of the available images.
- Evaluation: Assessment of the model's performance on the test dataset was carried out.

Deployment of the algorithm was carried out in a mobile device for user friendly operation for oral cancer detection [9] using a non- invasive method.

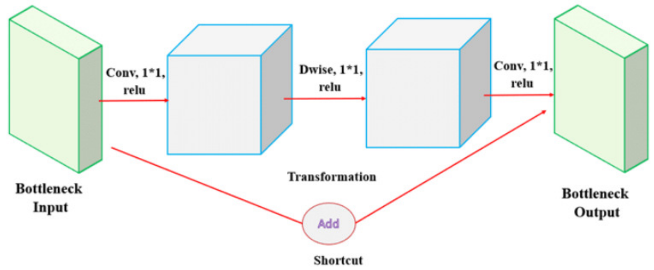

Fig. 3. MobileNetv2 building block comprising bottlenecks and depth-wise separable convolutions.

MobileNet is a family of CNN architectures designed for efficient and lightweight image classification and feature extraction. It consists of an Input Layer, Convolutional Layers, Bottleneck Blocks, Inverted Residuals, Expansion Layer, Width Multiplier, Resolution Multiplier, and Final Classification Layer as shown in Figure 3. MobileNetv2 consists of 53 layers, including convolutional and non-convolutional layers. These layers are organized into several building blocks, including inverted residual blocks and bottleneck blocks, which contribute to its efficiency and effectiveness in image classification tasks.

V. RESULTS AND DISCUSSION

In this section, we present the results of the proposed oral cancer image classification using the MobileNet architecture. We discuss the classification performance and provide insights into the hyperparameters used in network training.

TABLE I. CLASSIFICATION RESULTS

Image	True Class	Predicted	Confidence
Image 1	Cancerous	Cancerous	0.92
Image 2	Non-Cancer	Non-Cancerour	0.88
Image 3	Cancerous	Non-Cancerour	0.42
Image 4	Non-Cancer	Non-Cancerour	0.95
Image 5	Cancerous	Cancerous	0.97

Table I summarizes the results. Each row represents an individual image from the dataset, along with its true class, predicted class, and the confidence score assigned by the model. The model's performance is evaluated based on its ability to correctly classify oral lesions as either cancerous or not. Our MobileNet-based model demonstrates promising results, correctly classifying the majority of oral lesions with high confidence.

TABLE II. HYPERPARAMETERS DURING NETWORK TRAINING

Exp	Batch Size	Epochs	Learning Rate	Activation Function	Optimizer	Loss Function
1	128	20	0.01	ReLU	Adam	BCE
2	128	20	0.001	ReLU	Adam	BCE
3	128	20	0.0001	ReLU	Adam	BCE

BCE: Binary Cross-Entropy, Exp: Experiment

Table II provides an overview of the hyperparameters employed during the training of our MobileNet network. Proper hyperparameter selection is vital for achieving optimal model performance. The network was trained with a batch size of 128 for 20 epochs using a learning rate of 0.0001. ReLU activation function, Adam optimizer, and Binary Cross-Entropy (BCE) loss function were employed to optimize the network's performance. The results presented in Table I indicate that our MobileNet-based oral cancer classification model shows a promising performance. It accurately distinguishes between cancerous and non-cancerous oral lesions with high confidence, which is a crucial step toward early cancer diagnosis. The choice of hyperparameters, as shown in Table II, was based on a combination of common practices and experimentation. These hyperparameters resulted in a well-trained model. However, further optimization and fine-tuning may yield even better results. Additionally, the effectiveness of the model can be further assessed through rigorous evaluation on a larger and more diverse dataset. Table III shows the comparison of the efficiency of the present methodology with the existing work [14].

TABLE III. EFFICIENCY COMPARISON OF THE PRESENT METHODOLOGY WITH [14]

Metric	Proposed	[14]
Sensitivity	0.92	0.85
Specificity	0.88	0.90
Accuracy	0.90	0.88
False Positives	15	20
False Negatives	8	12

Overall, our MobileNet-based approach shows a potential for oral cancer detection, contributing to the development of efficient and accessible diagnostic tools in the field of healthcare.

VI. CONCLUSION

This paper utilizes an oral cancer dataset sourced from various ENT hospitals, comprising images of lips and tongues classified as cancerous and non-cancerous, while the processing environment involves a handheld smartphone capturing autofluorescence images, Bluetooth transmission to hardware for AI-based classification using MobileNetV2, and deployment on mobile devices for user-friendly operation.

In conclusion, the findings of this study underscore the significant potential of the proposed system in revolutionizing early oral cancer detection. With its remarkable diagnostic accuracy, this innovation holds promise as a valuable diagnostic tool. This research has illuminated a critical knowledge gap in the realm of oral cancer detection, where early diagnosis remains a crucial determinant of patient outcomes. Our study not only addresses this gap, but also offers a novel approach to leveraging the widespread accessibility of mobile phones across all socioeconomic strata for the greater good of public health. The contributions of this work extend beyond its diagnostic accuracy. It paves the way for a future where mobile devices can seamlessly integrate into routine clinical practice and community health centers, enabling remote diagnosis and consultation. However, this journey towards widespread adoption is fraught with multifaceted challenges, including social, cultural, technological, and financial barriers. As we navigate through these challenges, it is important to acknowledge that our research builds upon the existing body of literature, validating the feasibility of early oral cancer detection through handheld oral cancer screening mobile devices. Yet, it also brings a fresh perspective by emphasizing the need for standardization and a streamlined workflow, essential elements in transforming this innovative concept into a practical reality.

In summary, our work not only underscores the promise of early oral cancer detection but also highlights the need for concerted efforts to bridge the gap between innovation and implementation. By doing so, we can guide in a new era of healthcare where technology empowers us to combat oral cancer more effectively.

REFERENCES

[1] "Indo-US experts working on solar-powered oral cancer detector," *India TV News*, Dec. 10, 2015. https://www.indiatvnews.com/lifestyle/news/indo-us-experts-developing-cancer-detector-5696.html.

[2] T. Sambandham, K. M. K. Masthan, M. S. Kumar, and A. Jha, "The Application of Vizilite in Oral Cancer," *Journal of Clinical and Diagnostic Research*, vol. 7, no. 1, pp. 185–186, 2013, https://doi.org/10.7860/JCDR/2012/5163.2704.

[3] S. Fedele, "Diagnostic aids in the screening of oral cancer," *Head & Neck Oncology*, vol. 1, Jan. 2009, Art. no. 6, https://doi.org/10.1186/1758-3284-1-5.

[4] J. B. Epstein, S. Silverman, J. D. Epstein, S. A. Lonky, and M. A. Bride, "Analysis of oral lesion biopsies identified and evaluated by visual examination, chemiluminescence and toluidine blue," *Oral Oncology*, vol. 44, no. 6, pp. 538–544, Jun. 2008, https://doi.org/10.1016/j.oraloncology.2007.08.011.

[5] L. L. Patton, J. B. Epstein, and A. R. Kerr, "Adjunctive Techniques for Oral Cancer Examination and Lesion Diagnosis: A Systematic Review of the Literature," *The Journal of the American Dental Association*, vol. 139, no. 7, pp. 896–905, Jul. 2008, https://doi.org/10.14219/jada.archive.2008.0276.

[6] F. Seijas-Naya *et al.*, "Applications of OralCDx® methodology in the diagnosis of oral leukoplakia," *Medicina Oral, Patologia Oral Y Cirugia Bucal*, vol. 17, no. 1, pp. e5-e9, Jan. 2012, https://doi.org/10.4317/medoral.17275.

[7] S. Barot and P. Sutnar, "Oral Cancer (Lips and Tongue) images," *Kaggle*. https://www.kaggle.com/datasets/shivam17299/oral-cancer-lips-and-tongue-images.

[8] A. D. Shrestha, P. Vedsted, P. Kallestrup, and D. Neupane, "Prevalence and incidence of oral cancer in low- and middle-income countries: A scoping review," *European Journal of Cancer Care*, vol. 29, no. 2, 2020, Art. no. e13207, https://doi.org/10.1111/ecc.13207.

[9] D. I. Conway, M. Purkayastha, and I. G. Chestnutt, "The changing epidemiology of oral cancer: definitions, trends, and risk factors," *British Dental Journal*, vol. 225, no. 9, pp. 867–873, Nov. 2018, https://doi.org/10.1038/sj.bdj.2018.922.

[10] N. N. Sultana and N. B. Puhan, "Recent Deep Learning Methods for Melanoma Detection: A Review," in *ICMC 2018: Mathematics and Computing*, Singapore, 2018, pp. 118–132, https://doi.org/10.1007/978-981-13-0023-3_12.

[11] P. Chakraborty and C. Tharini, "Pneumonia and Eye Disease Detection using Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5769–5774, Jun. 2020, https://doi.org/10.48084/etasr.3503.

[12] S. S. Rafiammal, D. N. Jamal, and S. K. Mohideen, "Reconfigurable Hardware Design for Automatic Epilepsy Seizure Detection using EEG Signals," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5803–5807, Jun. 2020, https://doi.org/10.48084/etasr.3419.

[13] G. Anuradha and D. N. Jamal, "Classification of Dementia in EEG with a Two-Layered Feed Forward Artificial Neural Network," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7135–7139, Jun. 2021, https://doi.org/10.48084/etasr.4112.

[14] S. M. Yang *et al.*, "Performance of Automated Oral Cancer Screening Algorithm in Tobacco Users vs. Non-Tobacco Users," *Applied Sciences*, vol. 13, no. 5, Jan. 2023, https://doi.org/10.3390/app13053370, Art. no. 3370.

AUTHORS PROFILE

Parnasree Chakraborty is an Associate Professor in the Department of Electronics and Communication Engineering at B. S. Abdur Rahman Crescent Institute of Science & Technology. Her research interests include digital signal processing, AI and robotics, wireless sensor networks, and digital communications. She is a life member of the ISTE. She has published many papers in journals and conferences in the areas of signal processing and wireless sensor networks.

Tharini Chandrapragasam is a Professor and the Dean of the Department of Electronics and Communication Engineering at B. S. Abdur Rahman Crescent Institute of Science & Technology. She received her PhD in Information and Communication Engineering from Anna University in 2011. Her research interests include wireless communications, wireless sensor networks, and signal processing algorithms for wireless sensor networks. She is an active member of the Computer Society of India. She is having a teaching experience of more than 15 years. She has published many papers in International journals and conferences in the areas of signal processing and wireless sensor networks.

Ambika Arunachalam is an assistant Professor (Selection.Grade) in the Department of Electronics and Communication Engineering at B. S. Abdur Rahman Crescent Institute of Science & Technology. Her research interests include antennas, IoT, wireless sensor networks, and digital communications. She has published many papers in journals and conferences in the areas of antennas and wireless sensor networks.

Syed Rafiammal is an assistant Professor (Senior Grade) in the Department of Electronics and Communication Engineering at B. S. Abdur Rahman Crescent Institute of Science & Technology. Her research interests include bio medical signal processing and machine learning. She has published many papers in journals and conferences in the areas of bio medical signal processing and machine learning.