International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Feb-24)

http://dx.doi.org/10.12785/ijcds/150162

Convolutional Neural Network For Detection Of Oral Cavity Leading To Oral Cancer From Photographic Images

Shruthi K[1], A S Poornima[2], Musaddiq Shariff[3], Pradeep Singh S M[4], Subramanyam D P[5] and Varun M H[6]

[1-6]*Computer Science and Engineering, Siddaganga Institue of Technology, Tumakuru,Visvesvaraya Technological University, Belagavi, India*

E-mail address: shruthik@sit.ac.in, aspoornima@sit.ac.in, musaddiqshariff11@gmail.com, pradeepsinghtmk1999@gmail.com, subramanyamdp578@gmail.com, varun.m.h.2000@gmail.com

Received 27 Aug. 2023, Revised 30 Jan. 2024, Accepted 31 Jan. 2024, Published 5 Feb. 2024

Abstract: Oral cancer poses a substantial global health threat, as it continues to witness escalating incidence rates and consequential mortality on a widespread scale. To enhance patient outcomes, the crucial role of early detection cannot be overlooked. This research introduces an innovative real-time approach to detect various oral cavity conditions, focusing specifically on the prediction of oral cancer using the framework of deep learning. The methodology adopted integrates the user answering questions and oral cavity images analysis, amalgamating them to improve the accuracy and reliability of our predictive model. The comprehensive questionnaires gather extensive data on dietary habits, lifestyle factors, and potential risk factors associated with oral cancer. Leveraging deep learning models such as ResNet101, ResNet50, ResNet152, and VGG19, we classify oral cavity images as either cancerous or non-cancerous. By considering the relative weightage of the questionnaire responses and image analysis predictions, we compute a final probability of oral cancer. A diverse dataset is utilized to evaluate the performance of our proposed model, assessing its accuracy, sensitivity, specificity, and overall predictive capability. The resulting system aims to provide healthcare professionals with a real-time prediction tool featuring a user-friendly interface, thereby facilitating early detection and intervention. The outcomes of this study significantly contribute to the advancement of oral cancer detection methods, offering the potential to enhance patient outcomes through timely intervention.

Keywords: Convolutional Neural Networks, Deep Learning, Oral Cavity, Oral Potentially Malignant Disorder, Questionnaires.

1. Introduction

Global health faces a significant challenge from the escalating incidence and associated mortality of oral cancer. The prompt identification and timely intervention offer substantial potential to improve patient outcomes in the context of this ailment. Nevertheless, conventional diagnostic approaches frequently depend on subjective visual evaluations conducted by healthcare practitioners, resulting in possible setbacks in identifying and diagnosing conditions. From past few years, there has been a groundbreaking transformation in the field of medical image analysis, driven by the introduction of deep learning models usage. Using convolutional neural networks (CNNs), researchers have achieved significant advancements in applying these models to diverse medical imaging responsibilities, such as the timely identification of cancerous conditions in various anatomical regions.

Motivated by the complex neural connections found in the brain of a human, a complex structure emerges to form neural network. The networks contains number of layers with synthetic nodes or units that systematically analyze and convey information. From the extensive training, they develop the capacity to identify some patterns and reveal intricate connections within data, enabling accurate predictions. The revolutionary capabilities of neural networks have been applied in many domains, including but not limited to image classification, natural language processing, and object recognitions.

Convolutional Neural Networks (CNNs) have proven to be a crucial and efficient asset in the field of medical image analysis, especially when it comes to assisting in the detection of oral cancer. By training a CNN on a varied and extensive dataset containing images related to oral cancer, the network attains a high level of proficiency in precisely recognizing areas that suggest the presence of cancer. This groundbreaking advancement significantly empowers healthcare professionals by endowing them with enhanced early detection capabilities. Furthermore, it plays a pivotal

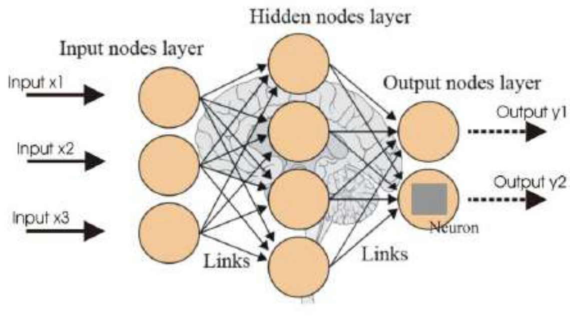

Figure 1. Neural Network Architecture

role in facilitating treatment planning and establishing an effective means for closely monitoring patients with oral cancer.

In our continuous research efforts, we introduce a novel method tailored for the instantaneous detection of diverse oral cavity conditions. This method prioritizes the proactive anticipation of oral cancer onset, harnessing the potential of an advanced deep learning framework. Our approach is marked by a harmonious amalgamation of patient surveys and images of the oral cavity, culminating in the creation of a predictive model.

This model not only enhances the overall accuracy of predictions but also elevates the reliability of the diagnostic process, marking a significant stride in the field of oral cancer detection and prognosis.

A. *Organization of Paper*

The rest of the paper is arranged as follows: related work is discussed in Section Two. In Section Three, the proposed methodology along with the proposed CNN model are highlighted. The methods and techniques used are discussed in Section Four. Results and Experimental studies are discussed in Section Five. UI Snapshots are completed in Section Six. Section Seven sheds on the limitations and considerations of the work. Finally, in the Section Eight, the conclusion is elaborated.

2. Related Work

Literature	N. Tenali, V. S. Desu, C. Boppa, V. Chowdary Chintala, and B. Guntupalli, "Oral Cancer Detection using Deep Learning Techniques". [1]
Methodology	Applied advanced deep learning to swiftly and accurately detect oral cancer from medical imaging, enabling cost-effective early interventions.
Key Findings	The integration of advanced deep learning methods is crucial for swiftly identifying and proactively applying measures against oral cancer. Employing these cutting-edge techniques enables healthcare practitioners to analyze essential medical imaging data with precision, ultimately enhancing the accuracy of oral cancer detection. This inventive strategy holds significant promise for the creation of economical and clinically relevant interventions, ensuring the early, noninvasive identification of oral cancer—a facet that highlights its utmost importance in the realm of healthcare.

Literature	M. Goswami, M. Maheshwari, P. D. Baruah, A. Singh, and R. Gupta, "Automated Detection of Oral Cancer and Dental Caries Using Convolutional Neural Network". [2]
Methodology	Implement system for early, precise oral health diagnosis without physical examination.
Key Findings	It is noteworthy to mention that the system assumes a pivotal role in facilitating the early diagnosis of oral health concerns. This innovative approach not only ensures a streamlined and efficient process but also furnishes dentists with a tool for precise and accurate identification of dental health issues, eliminating the necessity for direct physical examination.
Literature	C. C. Ukwuoma et al., "Detection of Oral Cavity Squamous Cell Carcinoma from Normal Epithelium of the Oral Cavity using Microscopic Images". [3]
Methodology	Authors introduce a novel method using pretrained deep learning models, synchronized with shared layers for early OCSCC detection. Transfer learning and Augmentor enhance dataset quality, validated through comparative analysis, showcasing the potential for early OCSCC detection and classification.
Key Findings	In this, authors present a revolutionary technique that utilizes a combination of pretrained deep learning models to identify early-stage oral cavity squamous cell carcinoma (OCSCC) in microscopic images. This method incorporates the integration of shared layers to synchronize the ensemble heads. Furthermore, we employ transfer learning by leveraging various pre-trained models, supported by the Augmentor library, to create datasets of exceptional quality. To validate the effectiveness of our approach, we perform a comparative analysis of our findings with those of another study that used the same dataset. This comparison highlights the potential and feasibility of our proposed methodology for the early detection and classification of OCSCC.

Literature	G. Sharma and R. Chadha, "Skin Cancer and Oral Cancer Detection using Deep Learning Technique". [4]
Methodology	Deep learning technique for prompt skin and oral cancer detection. It comprehensively reviews prior research, offering a visually elucidating flowchart for enhanced reader comprehension of the proposed methodology.
Key Findings	The focal point of the engaging study is centered on presenting a groundbreaking approach that harnesses the potential of deep learning methods to promptly detect skin and oral cancer. The paper thoroughly examines a wide range of prior research, providing an all-encompassing analysis of scientific papers dedicated to identifying these particular types of cancers. Furthermore, a visually elucidating flowchart is incorporated within the document, strategically designed to illustrate the proposed methodology, thereby augmenting reader comprehension.
Literature	R. Chinnaiyan, M. Shashwat, S. Shashank, and P. Hemanth, "Convolutional Neural Network Model based Analysis and Prediction of Oral Cancer". [5]
Methodology	Research predicts oral cancer stages using diverse, effective machine learning classifiers.
Key Findings	Within the context of the research paper, authors central focus centers on a thorough investigation and prediction of oral cancer using a varied range of machine learning classifiers. These classifiers are systematically applied to carefully curated datasets, and their performance undergoes rigorous examination through comprehensive evaluations. The results undeniably demonstrate the effectiveness of our methodology in accurately classifying various stages of carcinoma. By strategically employing machine learning techniques, we enable swift analysis and detection of carcinoma, and the integration of diverse methodologies highlights promising potential for the early diagnosis of this condition.

Literature	Y. C. Srivastava, "Oral cancer detection". [6]
Methodology	Introduces the 2017 WHO classification for oral potentially malignant disorders, exploring risk factors comprehensively.
Key Findings	It is crucial to underscore that experienced healthcare practitioners should face no difficulties in identifying precursor lesions that typically precede oral cancer. This is http:// journals.uob.edu.bh Int. J. Com. Dig. Sys. , No. (Mon-20..)) 191 because such lesions often present with distinct clinical features. Moreover, this article introduces the 2017 World Health Organization (WHO) classification system for 'oral potentially malignant disorders.' This classification encompasses a spectrum of conditions associated with various factors, including habits, immune-mediated or inflammatory influences, solar radiation-induced conditions like actinic cheilitis, and genetic disorders such as dyskeratosis congenita. Delving into the content of this paper, authors embark on a thorough exploration of common risk factors and oral potentially malignant disorders, with the aim of providing a comprehensive understanding of these crucial aspects.

Literature	H. Lin, H. Chen, L. Weng, J. Shao, and J. Lin, "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis". [7]
Methodology	The approach focuses on centered image capture, mitigates smartphone camera variability with a curated dataset, and evaluates efficacy using the HRNet deep learning network.
Key Findings	Addressing the urgent need for automated identification of oral disorders, an innovative solution emerged, integrating the use of smartphone-derived imaging and advanced deep learning algorithms [7]. This approach employed a simple yet impactful method for capturing images of the oral cavity, with a focus on adhering to a centered rule during the image capture process. To mitigate the variability introduced by smartphone cameras, a carefully compiled medium-sized dataset covering five distinct disease categories was meticulously curated, along with the introduction of a resampling technique. The efficacy of this methodology http://journals.uob.edu.bh Int. J. Com. Dig. Sys. , No. (Mon-20..)) 191 in detecting oral cancer was evaluated by employing the state-of-the-art deep learning network, HRNet, representing the current pinnacle of technological advancement in this domain.
Literature	R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer". [8]
Methodology	Study adopts ResNet-101 for image classification and Faster R-CNN for object detection, advancing automated oral lesion detection.
Key Findings	The study, as detailed in the aforementioned reference, showcases encouraging results. This is primarily attributed to the adoption of two distinct deep learning-based computer vision approaches. The first involves the utilization of ResNet-101 for image classification, while the second employs Faster R-CNN for object detection. These sophisticated techniques contribute to the overall success of the MeMoSA project in advancing the field of automated oral lesion detection.

Literature	N. B R and G. K. A. Sanathkumar M P, "Oral Cancer Detection using Machine Learning and Deep Learning Techniques". [9]
Methodology	Authors employ diverse models, including a 43-layer CNN, for automatic oral cancer image classification, aiming early identification.
Key Findings	In this, the authors present a groundbreaking approach to the development of automatic software for classifying images related to oral cancer. Leveraging the accumulated data, various machine learning models, including Naive Bayes, KNN, SVM, ANN, and a newly conceptualized 43-layer CNN network inspired by VGG-16, are employed. These models work collectively to automate the detection and classification of oral malignancies. The central objective of this research is to develop a system capable of automatically identifying potential malignant oral lesions at an early stage. A crucial aspect of this undertaking involves the creation of an extensive dataset containing meticulously annotated oral lesions.
Literature	[10]-[28]
Methodology	These researches typically involves literature review, data collection, analysis, and interpretation to advance understanding and inform prevention and treatment strategies.
Key Findings	To optimize treatment results and mitigate the risk of malignant progression, the timely identification and precise staging of oral cancer assume crucial significance. This document delivers a thorough examination of established procedures for the acquisition and characterization of oral samples, as well as the implementation of molecular risk evaluations. Through furnishing recommendations on appropriate sampling techniques and customized subsequent analyses tailored to specific research goals.

3. Proposed System

Figure 2 illustrates the operational sequence of the proposed system. The mobile application, crafted to be compatible with both Android and iOS platforms, empowers users with a range of functionalities, thereby augmenting their capabilities. The extensive suite of features includes, but is not limited to:

- A registration and login experience characterized by its smooth and uninterrupted flow, ensuring users encounter a seamless journey from initial sign-up to accessing their accounts effortlessly.
- Obtain entry to an extensive examination designed to assess oral health comprehensively. This evaluation includes the following components:
  1) Engaging with a thoughtfully crafted set of questions, providing comprehensive responses to each item, and actively participating in a meticulously designed questionnaire tailored to elicit detailed insights and thoughtful reflections.
  2) Capturing the images of the oral cavity is one of the key functionalities provided by the mobile application.
- The Django REST server receives the data which was gathered, and advanced deep learning algorithms come into play. Then, the server employs a trained model to which analyzes the image of the oral cavity and then it predicts the class.The model is designed to take three images of the same area to enhance the accuracy. Average probability of each class is calculated. Additionally, the user's questionnaire responses are evaluated by aligning them with predetermined options for each question. This provides further insights and information about the health of the user related to oral cancer.

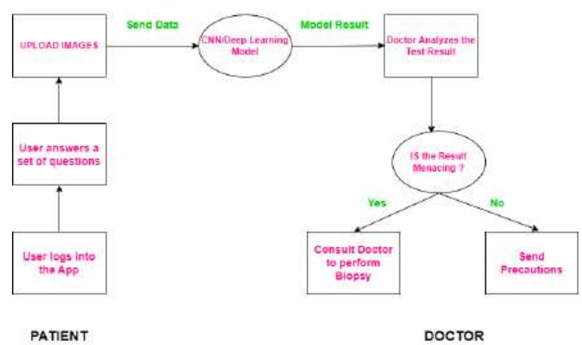

Figure 2. System Workflow

A. *System Scope*

- The functionality of this application empowers individuals, regardless of their medical background, to identify early alterations in the oral cavity. This is achieved through the utilization of their device's camera and the provision of responses to a set of inquiries regarding their daily habits. By leveraging these capabilities, even those without medical expertise can actively participate in the early detection of changes within their oral health.
- The purpose of the application is to assist oncologists and dentists by providing them with valuable insights into the severity of lesions. Additionally, it aids in precisely pinpointing the area that requires further examination through biopsy, enhancing the overall diagnostic process. The application serves as a valuable tool for healthcare professionals, facilitating a comprehensive understanding of

lesion characteristics and enabling targeted biopsy localization.

B. Objectives

- To enable a standard phone user to detect early changes in the oral cavity by answering a standard set of questions regarding their eating habits and using images of their oral cavity captured through a phone camera.
- Accuracy and performance of various models are trained and compared using the standard and real time dataset.
- Promoting awareness and adoption of the mobile application among the public, healthcare providers, and oral health organizations to increase early detection rates and improve oral health outcomes.
- Collaborating with healthcare professionals to validate the effectiveness and reliability of the mobile application for early detection of oral cavity abnormalities.

4. Material And Methods

The portrayed image in Figure 3 supplies an offers a thorough perspective on the suggested initiative.

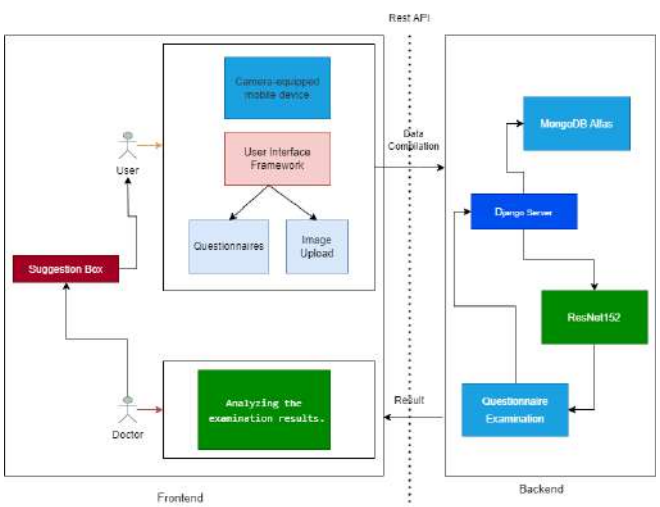

Figure 3. High Level Design

Through a smartphone application available on both the Google's Android and Apple's iOS platforms, people are able to participate in a comprehensive assessment of the health of their mouth. This engaging procedure comprises answering questions, taking pictures of the mouth cavity for evaluation, and assisting with user registration and login. A deep neural network model anticipates the category of the images and computes average probability after data is delivered to a Django REST service. The findings are weighted toward the questionnaire after The user's survey answers are converted into probability mappings. The doctor evaluates the final result, which is then shown to the user on the mobile application.

A. Oral Dataset

The website's data and live datasets sourced from a medical university make up the two primary parts of the training dataset utilized to train the ResNet152 model. This combination guarantees a wide and varied set of data for the model's training.
The dynamic dataset from the medical college provides precise and current data, the data from the website represents a wide range of situations. The model may learn from a variety of scenarios by combining the two sources, which improves the model's accuracy and resilience.
Incorporating live medical data into the learning phase not only improves the quality of the training but also guarantees that the model undergoes training on up-to-date and pertinent information, bringing it into close alignment with real-world applications in the fields of cancer diagnosis and oral health.

Figure 4. Cancer Image

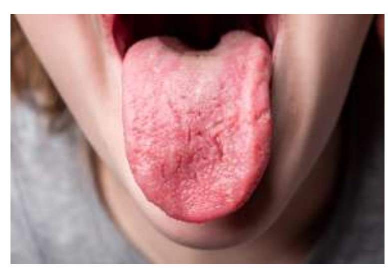

Figure 5. Non-Cancer Image

The malignancy pictures show various abnormalities that are irregular proliferations, ulcerations, or cancerous formations inside the mouth. As seen in Figure 4, these lesions are important markers of possible cancer.
However, as Figure 5 illustrates, the majority of the non-malignant images in this study depict ulcers, characterized by open sores or lesions brought on by a variety of circumstances but do not have malignant features.

B. Training and Testing Methodology

Two sets of images related to oral cancer are separated using a traditional 80/20 split. The dataset comprises a total of 150 images, encompassing both cancerous and non-cancerous instances. Within the learning phase of the Convolutional Neural Network (CNN) model, there is a pivotal role played wherein the images undergo processing through the network.

During this process, the parameters of the model are meticulously adjusted to reduce forecasting inaccuracies. Subsequent to the conclusion of the training process regimen, a distinct testing set is deployed to evaluate the efficacy and performance of the now-trained model. In this evaluation, a range of metrics, such as accuracy and precision, are systematically measured to provide a comprehensive assessment of the model's capabilities.

C. Network Architecture

1) VGG19

As seen in Figure 6, the structure of the VGG19 model is made up comprising a total of 19 layers, including 3 densely connected layers and 16 convolutional layers. The max-pooling layers utilize 2x2 filters with a stride of 2, and the convolutional layers efficiently use 3x3 filters with a stride of 1. The intricacy of this intricate network is additionally increased by a stepwise increase in the quantity of filters, which starts at 64 and doubles following each max-pooling layer.

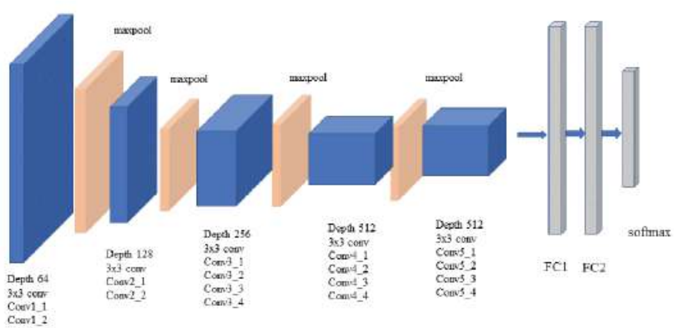

Figure 6. VGG19 Architecture

2) GoogLeNet

GoogLeNet makes use of a 22-layer deep neural network architecture. The inception module is a crucial part that, as Figure 7 illustrates, integrates parallel 1x1, 3x3, and 5x5 convolutional filters. Additionally, it uses max-pooling and average pooling strata along with extra auxiliary classifiers to promote intermediate feature acquisition.

3) ResNet101

Using residual connections to facilitate the learning of residual mappings, ResNet101 is an extensive neural network architecture with 101 layers. As shown in Figure 8, the residual block is the basic part and it uses two 3x3 convolutional layers with residual connections. After each downsampling step, the convolutional strata double the quantity of filters, starting at 64.

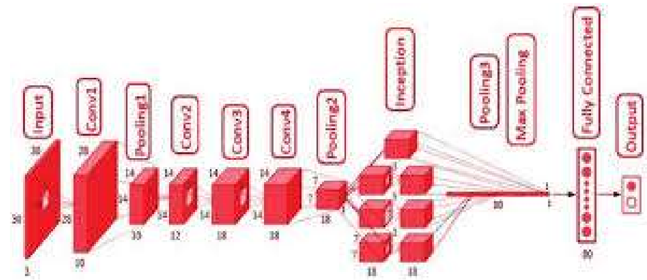

Figure 7. GoogLeNet Architecture

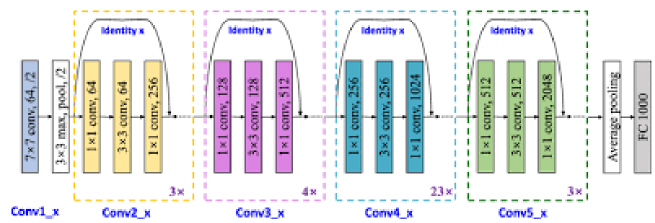

Figure 8. ResNet101 Architecture

4) ResNet152

ResNet152, which has 152 layers, is an extension of ResNet101, as shown in Figure 9. It has more depth but still adheres to the same residual block structure. Additionally, it begins with 64 filters and adds two more filters for every layer of downsampling. ResNet152's extra layers allow it to better represent the network and capture more complex features.

ResNet152 is utilized in two distinct applications.:

- Image Classification: ResNet152 has been widely employed in image classification tasks, demonstrating exceptional performance in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Its ability to accurately categorize images into multiple classes is attributed to its deep architecture and skip connections, allowing it to capture intricate features and enhance classification accuracy.
- Object Detection: ResNet152 has found application in tasks related to the identification and localization of multiple objects within an image. Leveraging the network's capacity for representation, object detection models can benefit from the learned deep features of ResNet152, resulting in enhanced accuracy and precision in detecting objects. This application

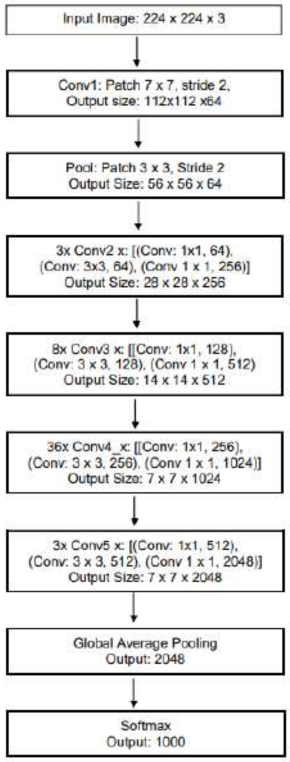

Figure 9. ResNet152 Architecture

contributes to improved overall performance in object detection tasks.

VGG19, GoogLeNet, ResNet101, and ResNet152 are the four well-known convolutional neural network (CNN) models that are compared in Figure 10. Every model has its own distinct features and architecture. Details for each model are listed in the table, including the quantity of parameters, number of layers, size of the filter, size of the input, and size of the output. A brief synopsis of these CNN architectures' salient features is made possible by this information.

Model	Filter Size	Layers	Accuracy	Number of Parameters	Input Size	Output Size
VGG19	3x3	19	69.44 %	143 million	224x224x3	7x7x512
GoogLeNet	Varies	22	63.89%	6.8 million	224x224x3	7x7x1024
ResNet101	7x7	101	72.22%	44.6 million	224x224x3	7x7x2048
ResNet152	7x7	152	80.56%	60.3 million	224x224x3	7x7x2048

Figure 10. Comparison of CNN Models

D. Processing and Classifying Images

- The entire system's workflow, as depicted in Figure 11. Standardize oral cavity images as input for further analysis and ensure consistent pixel value ranges and size, common image enhancement techniques such as image resizing and normalization are employed.

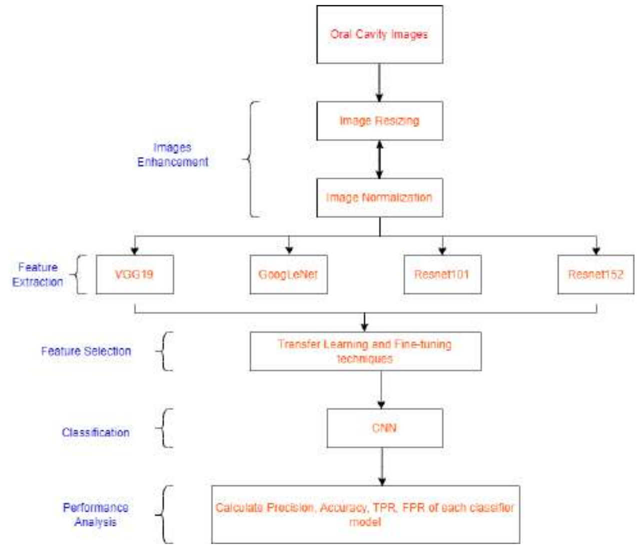

Figure 11. Work Flow

- In transfer learning, on of the pre-trained deep learning models like VGG19, GoogLeNet, ResNet101, and ResNet152 are employed to extract features. Leveraging the knowledge acquired from extensive datasets, these models extract the information from the oral cavity images.
- Through the adjustment and fine-tuning of parameters in pre-trained models on a dataset of oral cavity images, enhances the efficiency and accuracy through transfer learning and fine-tuning. This facilitates effective feature selection, ultimately enhancing the classification performance of CNN models. Performance metrics such as accuracy, precision, false positive rate (FPR), and true positive rate (TPR) are regularly evaluated to gauge the effectiveness of these models in terms of classification.

5. Results And Discussion

A. Experimenral Evaluation

Within this segment, we assessed three crucial metrics—namely, the confusion matrix, accuracy, and the Receiver Operating Characteristic (ROC) curve—to choose the most suitable model for the detection of oral cancer.

1) Confusion Matrix

The visual representation presented in Figure 12, identified as the confusion matrix, plays a pivotal role in elucidating the intricacies of the model's performance. Contained within its matrix structure, it offers a comprehensive breakdown of true positive predictions, true negative predictions, false positive predictions, and false negative predictions. This detailed classification of outcomes offers a nuanced insight into the model's efficacy in correctly identifying instances of oral cancer and differentiating non-cancerous cases. Going beyond the quantification of predictions, the confusion matrix emerges as a crucial tool for assessing accuracy and precision.

In essence, the confusion matrix acts as a powerful lens, enabling a meticulous examination of each model's predictive prowess. By offering a comprehensive view of both correct and incorrect classifications, it provides a nuanced perspective on the strengths and weaknesses inherent in the models' decision-making processes. This depth of insight is invaluable for not only evaluating the overall performance but also for identifying specific areas where enhancements or adjustments may be warranted. Consequently, the confusion matrix becomes more than a static representation; it transforms into a dynamic tool, empowering meaningful comparisons of performance across different models and shedding light on the subtle nuances of their predictive behaviors.

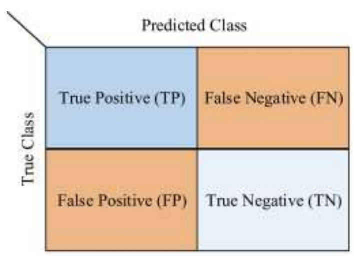

Figure 12. Confusion Matrix

Analyzing Figure 13 reveals a bar chart that adeptly juxtaposes the occurrences of false positives (FP) and false negatives (FN) among four distinct models: VGG19, GoogLeNet, ResNet101, and ResNet152. The x-axis corresponds to the model names, while the y-axis quantifies the instances of misclassifications. The height of each bar represents the specific count of false positives and false negatives attributed to each model.

After meticulous examination, a significant pattern becomes apparent. Specifically, ResNet152 stands out with the lowest occurrences of false positives and false negatives in comparison to the other models. This finding indicates that ResNet152 attains superior accuracy in accurately categorizing instances of oral cancer, leading to a reduced frequency of misclassifications. Conversely, VGG19, GoogLeNet, and ResNet101 demonstrate differing levels of misclassifications, evident in their elevated counts of false positives and false negatives.

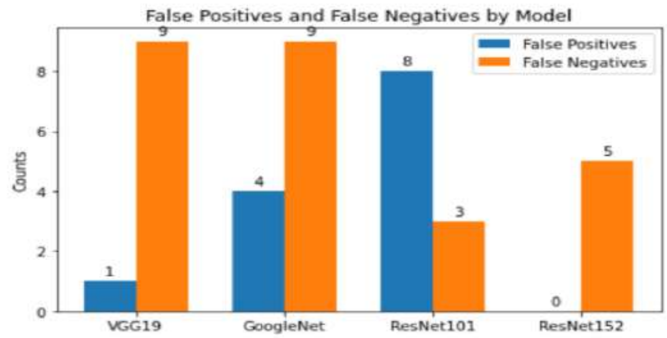

Figure 13. Bar graph compares models

2) Accuracy

Accuracy assessment is computed by using below formula.

$$Accuracy = \frac{\sum_{i=1}^{l} (TP_i + TN_i)}{TP_i + TN_i + FP_i + FN_i} \quad (1)$$

3) ROC Curve

In Figure 14, the ROC curve illustrates the balance between sensitivity and specificity across thresholds. The AUC succinctly captures the model's discriminatory ability. Analyzing ROC curves helps identify the optimal model with the highest AUC, facilitating nuanced comparisons.

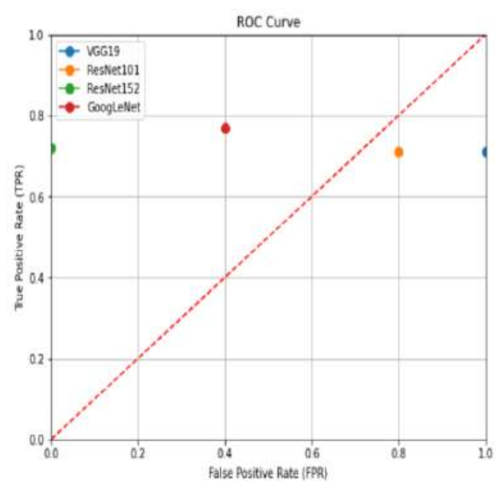

Figure 14. ROC curve

The presented findings bring attention to the ROC curve, depicted in Figure 14, which elucidates the distinction between false positive rates along the x-axis and true positive rates along the y-axis. Each data point on this visual representation corresponds to a distinct classification threshold, providing a visual insight into the effectiveness of the models in distinguishing between positive and negative instances. A greater elevation of the curve signifies superior performance. Notably, upon careful examination, the apparent superiority of ResNet152 over other models becomes evident, as it demonstrates an elevated true positive rate at lower false positive rates. This implies that ResNet152 effectively achieves a commendable balance between sensitivity and specificity in the identification of oral cancer.

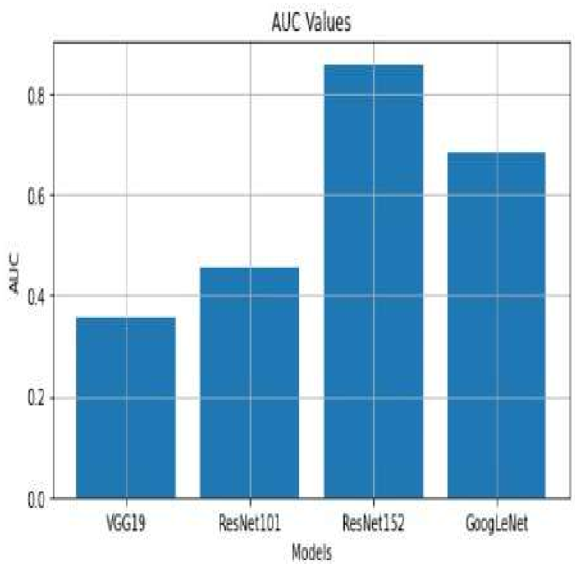

Figure 15. AUC Values for each Model

Presented in Figure 15, the associated bar graph provides a graphical portrayal of the Area Under the Curve (AUC) values for each model. This measurable parameter serves as a valuable evaluation of their respective performances, where higher AUC values indicate enhanced classification accuracy.

4) The choosen model - ResNet152

After an extensive evaluation of various deep learning models for our oral cancer detection project, ResNet152 emerged as the most robust and accurate choice. Its superior performance was evident in terms of accuracy and other key parameters. The intricate architecture of ResNet152, with its deep layer stacking and skip connections, allowed the model to capture complex features and patterns within the oral images. Not only did ResNet152 outperform other models in terms of overall accuracy, but it also demonstrated a remarkable ability to generalize well to new, unseen data. The decision to adopt ResNet152 as our primary model was further validated through rigorous testing, yielding a best test accuracy of 80.56 as shown in Figure 16. This exceptional level of accuracy positions ResNet152 as the optimal choice for our oral cancer detection framework, demonstrating its capacity to provide reliable and precise predictions on our dataset.

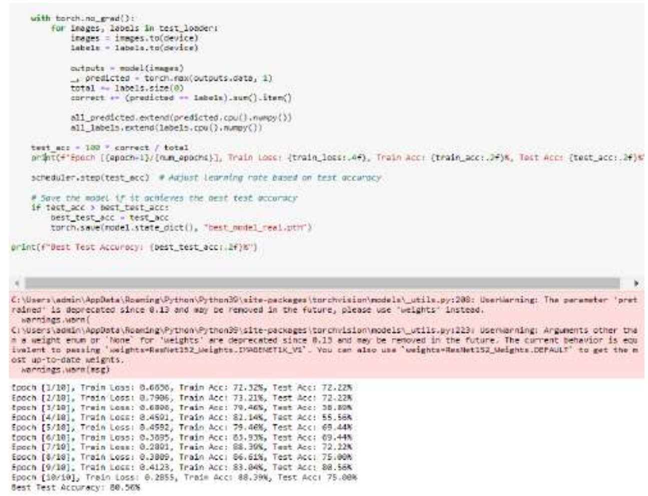

Figure 16. Test Accuracy of Resnet152

5) Diagnostic Methodology

Our project's algorithm operates on a dual-axis strategy, synergizing two fundamental components: an interactive user questionnaire and a sophisticated image classification model, ResNet152. The set of questions posed to the user is meticulously curated with inputs from an expert dental doctor, ensuring a comprehensive exploration of symptoms and relevant information. These questions are designed in a multiple-choice format, each assigned a specific weightage based on its criticality in the diagnostic process. Moreover, the options within each question carry individual weightages, reflecting their significance in the diagnostic decision-making process. Simultaneously, the ResNet152 model engages in image classification, leveraging its deep learning architecture to scrutinize intricate patterns within oral images. The amalgamation of these two dimensions is orchestrated through a meticulous aggregation process.

Probability calculations from both the questionnaire responses and the model's predictions are harmoniously combined to yield a comprehensive final result. This unique fusion of subjective user input and objective image analysis not only enhances the diagnostic accuracy but also provides a more holistic approach to oral cancer detection. The Figure 17 shows the probability calculation algorithm implemented.

6. User Interface Snapshots

Below are the snapshots of the user interface.

The register page for the user and doctor which consists of Username, Email, Create password is as shown in Figure 18.

In Figure 19, the user is asked about the basic information about the eating habits and the data collected

```python
def getProbabilityFromImage(self, b64image):
    # Convert the base64 image into a binary file
    # b64image = b64image.split("base64,")[1]
    file_content = base64.b64decode(b64image)
    file_path = f'{self.IMG_DIR_PATH}{calendar.timegm(time.gmtime())}.jpg'
    with open(file_path, "wb") as f:
        f.write(file_content)

    # Do the prediction
    img = image.load_img(file_path, target_size=(224, 224))
    img = np.array(img)
    img = img.reshape((1, 224, 224, 3))

    result_array = self.loaded_model.predict(img)
    print(result_array)
    result_array = result_array.flatten()

    # Delete the file before returning
    os.remove(file_path)

    # Return the probabilities for each class
    return list(result_array)

def getProbabilityFromQuestionnaire(self, responses):
    num_cancer = 0
    num_non_cancer = 0
    num_pre_cancer = 0
    p_of_cancer = 0
    p_of_non_cancer = 0
    p_of_pre_cancer = 0

    for i in range(len(responses)):
        if responses[i] == "" or responses[i] == None:
            continue
        num_cancer += self.options[i][responses[i]]["num_cancer"]
        num_non_cancer += self.options[i][responses[i]]["num_non_cancer"]
        num_pre_cancer += self.options[i][responses[i]]["num_pre_cancer"]

    total = num_cancer + num_non_cancer + num_non_cancer
    if total == 0:
        total = 1
    p_of_cancer = num_cancer / total
    p_of_non_cancer = num_non_cancer / total
    p_of_pre_cancer = num_pre_cancer / total

    return [p_of_cancer, p_of_non_cancer, p_of_pre_cancer]
```

Figure 17. Probability Calculation and Aggregation

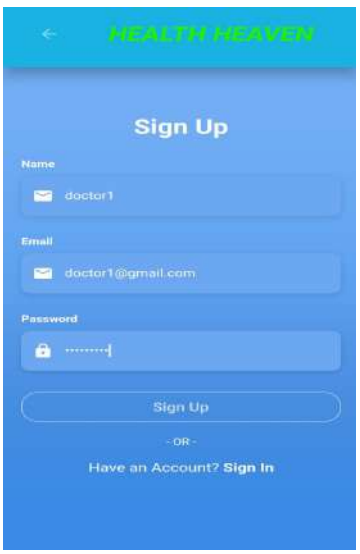

Figure 18. Doctor SignUp

Figure 19. Questionnaire Page

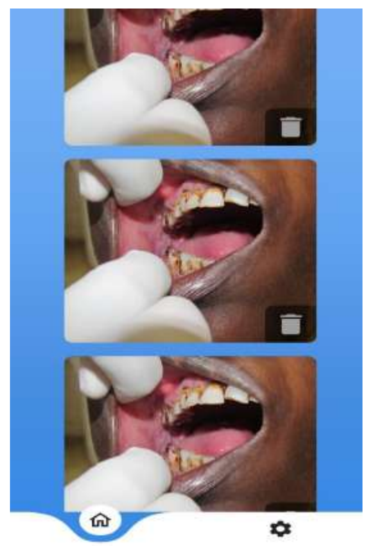

Figure 20. Image Upload Page

is sent to the AI/ML model for analysis.

The below steps indicate the flow of taking a test on the application:

- The user provides a response to the questionnaire as shown in Figure 19.
- The user clicks 3 images of the affected region and waits for the result as shown in Figure 20.
- The test details are passed onto the backend which classifies the image and evaluates the questionnaire.
- The test results and test details are stored in the database. Individual tests are maintained separately, unique test ID is provided to each of the tests and the same is displayed as shown in Figure 21.

The data provided by the user in the form of answers to the questions and images are processed and analyzed by the Deep Learning Model and the result is generated by classification into cancer or non-cancer. This is done by the probability obtained from the model. All the details of the test result including the probablilty is displayed at the doctor's side as shown in Figure 22.

The doctor is now able to access all the questions answered by the user during the test process. In addition to that the images of the oral cavity uploaded by the user is also displayed at the doctor's side. Now the Doctor is able

Figure 21. Test Details Page

Figure 22. Test Result Page

to decide the next procedure to be performed. Generally, if the lesions seems to be menacing and the model result predicted is cancer then Doctor suggests for Biopsy to be performed by the patient. If the lesions does not seems to be menacing then a precautionary measures must be followed by the user and the same is suggested by the Doctor. To enable this communication between Doctor and the user a suggestion box is provided at the doctor's side as shown in Figure 23.

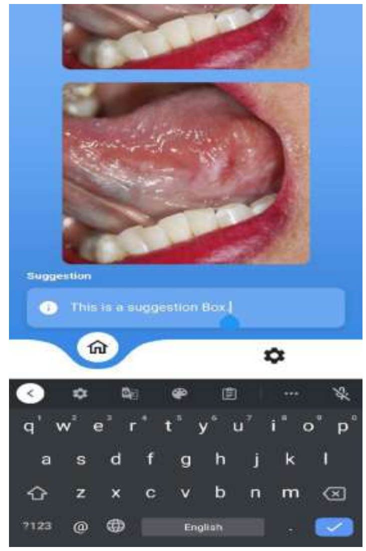

Figure 23. Suggestion Box Page

7. Limitations And Considerations

Despite the effectiveness of ResNet152 in oral cancer detection, our project contends with several inherent limitations that warrant careful consideration. The foremost challenge lies in the interpretability of deep learning models. The complex architecture of ResNet152, while contributing to its high accuracy, creates a black-box scenario, making it arduous to discern the decision-making rationale. This lack of interpretability could pose challenges in clinical settings, where understanding the basis of individual predictions is crucial for establishing trust and confidence in the model's diagnostic capabilities.

Another limitation revolves around the static nature of the image-based analysis. While ResNet152 excels in capturing patterns in static images, it may fall short in comprehensively addressing the dynamic nature of oral health. Oral conditions can evolve over time, and the model's reliance on fixed snapshots might limit its efficacy in capturing gradual changes or monitoring progressive abnormalities. Furthermore, the model's performance can be sensitive to external variables such as lighting conditions and image quality. This sensitivity introduces a potential source of variability in real-world scenarios, where these factors are often unpredictable. Acknowledging these limitations is pivotal as we strive for the continuous improvement of our oral cancer detection system, with ongoing efforts directed toward addressing these challenges and refining the model's adaptability and transparency.

8. Conclusion And Future Scope

A. Conclusion

In the intricate landscape of oral cancer detection, ResNet152 emerges not just as a choice but as a transformative force, dominating across a spectrum of critical metrics. Its ascendancy in accuracy, precision, and overall efficacy underscores a pivotal shift in our approach to diagnosing oral malignancies.

The accelerating prevalence of Oral Potentially Malignant Disorders in India represents more than a statistical trend; it embodies a growing public health challenge. Despite the commendable strides in technology by researchers, oncologists, and dentists, the diagnostic trajectory for oral cancer often unfolds belatedly, leading to heightened healthcare costs and discouraging prognoses. The pivotal nexus of change lies in early detection and intervention. This imperative is echoed by innovative devices like Vizilite, VELscope, and Cell Scope, each a beacon in the quest for timely identification. Particularly pressing is the need to curb the progression of Oral Squamous Cell Carcinoma (OSCC). In this context, our paper strives not only to highlight the significance of early detection but also to contribute to the evolving landscape of diagnostic methodologies. Harnessing the ubiquity of standard phone cameras, our aim is to empower a broader spectrum of healthcare providers with a tool that facilitates the swift identification of malignant lesions in the oral cavity. Through this pursuit, we envision not just a refinement of diagnostic processes but a paradigm shift in how we confront and combat oral cancer, with the ultimate goal of improving patient outcomes and steering towards a future where early detection is synonymous with prevention.

B. Future Scope

- To maximize the impact of our idea, future enhancements should focus on expanding collaborative networks with healthcare professionals, research institutions, and oral health organizations. By fostering collaborations and partnerships, we can leverage collective expertise, access larger datasets for training and validation, and gain valuable insights for refining the solution. This collaborative approach will ultimately contribute to the widespread adoption and continuous improvement of our system for early detection and prevention of oral cancer.

- A key area for future enhancement lies in continuously refining the mobile application's user interface (UI) and user experience (UX). Incorporating user feedback, conducting usability studies, and implementing iterative design processes can help optimize the app's navigation, functionality, and overall usability. This will ensure a seamless and intuitive experience for users, encouraging greater engagement and adoption of the early detection features.

References

[1] N. Tenali, V. S. Desu, C. Boppa, V. Chowdary Chintala, and B. Guntupalli, "Oral Cancer Detection using Deep Learning Techniques," Int. Conf. Innov. Data Commun. Technol. Appl. ICIDCA 2023 - Proc., pp. 168–175, 2023, doi: 10.1109/ICIDCA56705.2023.10100045.

[2] M. Goswami, M. Maheshwari, P. D. Baruah, A. Singh, and R. Gupta, "Automated Detection of Oral Cancer and Dental Caries Using Convolutional Neural Network," 2021 9th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir. ICRITO 2021, pp. 1–5, 2021, doi: 10.1109/ICRITO51393.2021.9596537.

[3] C. C. Ukwuoma et al., "Detection of Oral Cavity Squamous Cell Carcinoma from Normal Epithelium of the Oral Cavity using Microscopic Images," 2022 Int. Conf. Decis. Aid Sci. Appl. DASA 2022, pp. 29–36, 2022, doi: 10.1109/DASA54658.2022.9765023.

[4] G. Sharma and R. Chadha, "Skin Cancer and Oral Cancer Detection using Deep Learning Technique," IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron. ICDCECE 2022, pp. 1–5, 2022, doi: 10.1109/ICDCECE53908.2022.9792688.

[5] R. Chinnaiyan, M. Shashwat, S. Shashank, and P. Hemanth, "Convolutional Neural Network Model based Analysis and Prediction of Oral Cancer," 2021 Int. Conf. Adv. Electr. Electron. Commun. Comput. Autom. ICAECA 2021, pp. 1–4, 2021, doi: 10.1109/ICAECA52838.2021.9675533.

[6] Y. C. Srivastava, "Oral cancer detection.," J. Indian Dent. Assoc., vol. 40, no. 3, p. 75, 1968.

[7] H. Lin, H. Chen, L. Weng, J. Shao, and J. Lin, "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis," J. Biomed. Opt., vol. 26, no. 08, pp. 1–16, 2021, doi: 10.1117/1.jbo.26.8.086007.

[8] R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," IEEE Access, vol. 8, pp. 132677–132693, 2020, doi: 10.1109/ACCESS.2020.3010180.

[9] N. B R and G. K. A. Sanathkumar M P, "Oral Cancer Detection using Machine Learning and Deep Learning Techniques," Int. J. Curr. Res. Rev., vol. 14, no. 01, pp. 64–70, 2022, doi: 10.31782/ijcrr.2021.14104.

[10] G. Yang et al., "A Systematic Review of Oral Biopsies, Sample Types, and Detection Techniques Applied in Relation to Oral Cancer Detection," BioTech, vol. 11, no. 1, pp. 1–20, 2022, doi: 10.3390/biotech11010005.

[11] A. Mota-Ramírez, F. J. Silvestre, and J. M. Simó, "Oral biopsy in dental practice.," Med. Oral Patol. Oral Cir. Bucal, vol. 12, no. 7, 2007.

[12] Rengaswamy Sankaranarayanan, Kunnambath Ramadas, Hemantha Amarasinghe, Sujha Subramanian, and Newell Johnson "Oral Cancer: Prevention, Early Detection and Treatment", The International Bank for Reconstruction and Development / The World Bank, November 2015.

[13] N. Haj-Hosseini, J. Lindblad, B. Hasséus, V. V. Kumar, N. Subramaniam, and J. M. Hirsch, "Early Detection of Oral Potentially Malignant Disorders: A Review on Prospective Screening Methods with Regard to Global Challenges," J. Maxillofac. Oral Surg., 2022, doi: 10.1007/s12663-022-01710-9.

[14] C. Anitha, M. Tech., Assistant Professor, Department of Electronics and Communication Engineering, Jeppiaar Engineering College, Chennai. S. Sasikumar, Student, Department of Electronics and Communication Engineering, Jeppiaar Engineering College, Chennai "Oral Cancer Detection By CNN", June 2022.

[15] Gonzalez-Moles MA, Aguilar-Ruiz M, Ramos-Garcia P, "Early Diagnosis of Oral Cancer, Evidence Gaps and Strategies for Im-

provement", A Scoping Review of Systematic Reviews. Cancers (Basel), October 2022.

[16] Santosh Hunasgi, Anila Koneru, Vardendra Manvikar, Rudraraju Amrutha, K. M. Parveen Reddy "Oral biopsy in general dental practice" in Journal of Advanced Clinical Research Insights, 2017.

[17] P. Birur N et al., "Field validation of deep learning based Point-of-Care device for early detection of oral malignant and potentially malignant disorders," Sci. Rep., vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-18249-x.

[18] B. Ananthakrishnan, A. Shaik, S. Kumar, S. O. Narendran, K. Mattu, and M. S. Kavitha, "Automated Detection and Classification of Oral Squamous Cell Carcinoma Using Deep Neural Networks," Diagnostics, vol. 13, no. 5, pp. 1–19, 2023, doi: 10.3390/diagnostics13050918.

[19] K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, P. Jantana, and S. Vicharueang, "AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer," PLoS One, vol. 17, no. 8 August, pp. 1–14, 2022, doi: 10.1371/journal.pone.0273508.

[20] Q. Fu et al., "A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study," EClinicalMedicine, vol. 27, p. 100558, 2020, doi: 10.1016/j.eclinm.2020.100558.

[21] M. Das, R. Dash, and S. K. Mishra, "Automatic Detection of Oral Squamous Cell Carcinoma from Histopathological Images of Oral Mucosa Using Deep Convolutional Neural Network," Int. J. Environ. Res. Public Health, vol. 20, no. 3, 2023, doi: 10.3390/ijerph20032131.

[22] Q. Huang, H. Ding, and N. Razmjooy, "Oral cancer detection using convolutional neural network optimized by combined seagull optimization algorithm," Biomed. Signal Process. Control, vol. 87, no. PB, p. 105546, 2024, doi: 10.1016/j.bspc.2023.105546.

[23] N. C. S. K. Chaitanya, S. Chavva, E. Surekha, V. Priyanka, M. Akhila, and H. K. Ponnuru, "Head and Neck Cancers A Meta-analysis on efficacy of auto fluorescence in detecting the early dysplastic changes of oral cavity," pp. 233–236, 2020, doi: 10.4103/sajc.sajc.

[24] M. J. Jeng et al., "Multiclass classification of autofluorescence images of oral cavity lesions based on quantitative analysis," PLoS One, vol. 15, no. 2, pp. 1–18, 2020, doi: 10.1371/journal.pone.0228132.

[25] T. Y. Rahman, L. B. Mahanta, A. K. Das, and J. D. Sarma, "Histopathological imaging database for oral cancer analysis," Data Br., vol. 29, p. 105114, 2020, doi: 10.1016/j.dib.2020.105114.

[26] S. M. Fati, E. M. Senan, and Y. Javed, "Early Diagnosis of Oral Squamous Cell Carcinoma Based on Histopathological Images Using Deep and Hybrid Learning Approaches," Diagnostics, vol. 12, no. 8, 2022, doi: 10.3390/diagnostics12081899.

[27] S. Abati, C. Bramati, S. Bondi, A. Lissoni, and M. Trimarchi, "Oral cancer and precancer: A narrative review on the relevance of early diagnosis," Int. J. Environ. Res. Public Health, vol. 17, no. 24, pp. 1–14, 2020, doi: 10.3390/ijerph17249160.

[28] V. Borse, A. N. Konwar, and P. Buragohain, "Oral cancer diagnosis and perspectives in India," Sensors Int., vol. 1, no. September, p. 100046, 2020, doi: 10.1016/j.sintl.2020.100046.

Shruthi K Shruthi K is a Research scholar at the Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur, Karnataka, India. She received an M.Tech. degree in Computer Science and Engineering from VTU Belgaum in 2014. She has teaching experience of more than 8 years. Her research interests include Machine Learning, Blockchain Technology, Databases, and Cloud Computing.

Poornima A S Dr. Poornima A S is a Professor and head of the Department of Computer Science and Engineering at Siddaganga Institute of Technology, Tumkur, Karnataka, India. She was awarded a Ph.D. degree in Computer Science and Engineering in 2012. She has teaching experience of more than 20 years. Her research interests include wireless sensor networks, Artificial Intelligence, Machine Learning, and Data Science.

Musaddiq Shariff Musaddiq Shariff is an undergraduate student at the Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur, Karnataka, India. His research interests include web development, Artificial Intelligence, and Blockchain. He has worked on various projects like Digital Assets Management using Blockchain, and web development projects using HTML, JavaScript, and MySQL.

Pradeep Singh S M Pradeep Singh S M is an undergraduate student at the Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur, Karnataka, India. His research interests include Machine Learning, Deep Learning, Databases, and Cloud Computing. He has worked on projects like SMS Spam Detection using Machine Learning and Implementation of Proxy Chains in Networking.

Subramanayam D P Subramanayam D P is an undergraduate student at the Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur, Karnataka, India. His research interests include Machine Learning, Web Development, and Cloud Computing. He has worked on various projects like Agricultural Loan Approval Prediction using Machine Learning and Electricity Bill System Database.

Varun M H Varun M H is an undergraduate student at the Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumkur, Karnataka, India. His research interests include Machine Learning, Deep Learning, Databases. He has worked on projects like Movie Recommendation System using Machine Learning and Crime Rate Prediction using K-means algorithm.