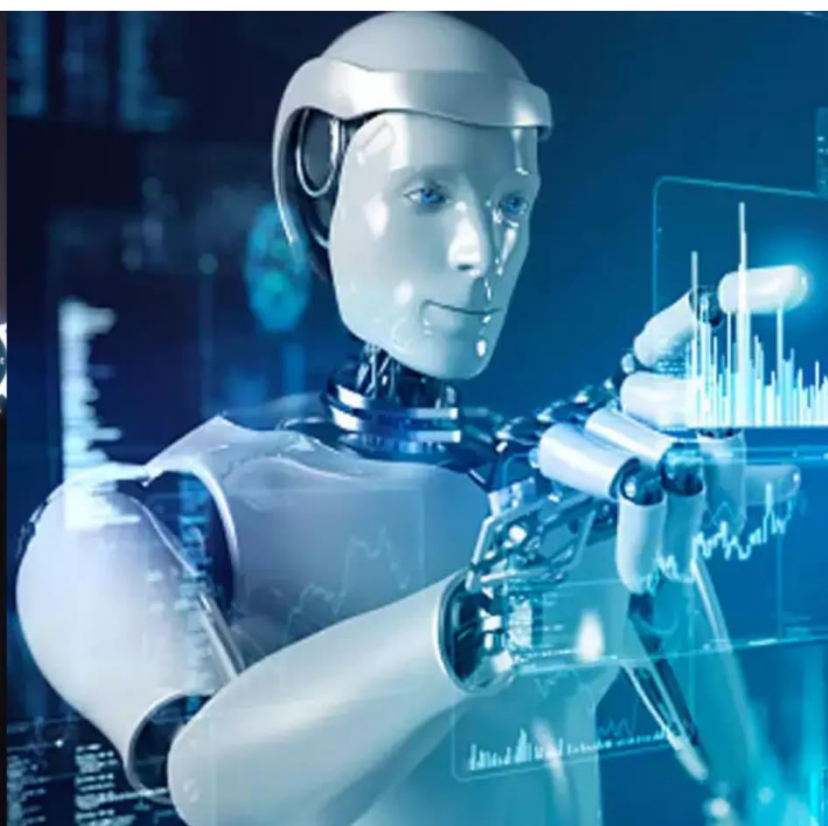




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

Oral Cancer Detection using Machine Learning

Prof. Vindya L, Chethan S, Dhananjaya Gowda

Assistant Professor, Department of Information Science and Engineering, S J C Institute of Technology,
Chickaballapura, Karnataka, India

U.G. Student, Department of Information Science and Engineering, S J C Institute of Technology, Chickaballapura,
Karnataka, India

U.G. Student, Department of Information Science and Engineering, S J C Institute of Technology, Chickaballapura,
Karnataka, India

ABSTRACT: Oral cancer is a serious health issue that affects many people worldwide. Early detection is crucial for better treatment outcomes, yet traditional diagnostic methods often face delays. This research explores how Recurrent Neural Networks (RNN), a form of deep learning, can help detect oral cancer early by analyzing patient data like medical images and clinical records. The system identifies patterns that signal early-stage cancer, potentially before visible symptoms appear. By combining advanced image processing and patient history analysis, this approach aims to improve diagnosis speed and accuracy, helping doctors make timely decisions and ultimately saving more lives.

KEYWORDS: Oral cancer, Early detection, Recurrent Neural Network(RNN), Deep Learning, Medical imaging, Patient records, AI in healthcare, Cancer Diagnosis, Image analysis, Machine learning models, Health technology, Tumor detection, Smart diagnosis, Clinical data analysis, Oral health monitoring

I. INTRODUCTION

Oral cancer is a dangerous disease that starts in the mouth and can spread to other parts of the body if not found early. It often affects areas like the lips, tongue, gums, and inner cheeks. Many people do not notice the symptoms until it becomes serious, which makes treatment harder and reduces the chances of survival. Factors such as tobacco use, alcohol, and infections like HPV increase the risk. Because early signs can be easy to miss, there's a growing need for better and faster ways to detect it before it becomes life-threatening.

Modern technology like artificial intelligence (AI) is helping change the way we detect diseases. In this research, a type of AI called Recurrent Neural Networks (RNN) is used to look for patterns in medical images and patient records. RNNs are especially good at finding changes over time, which helps in spotting cancer early. By combining image data with clinical history, this system can give doctors a clearer picture and help catch oral cancer in its early stages. This not only makes diagnosis more accurate but can also lead to quicker and more effective treatment.

In recent years, the use of artificial intelligence in healthcare has shown great promise, especially in diagnosing complex diseases like oral cancer. Traditional methods often rely on manual examination, which can be time-consuming and may miss subtle signs. With AI tools like Recurrent Neural Networks, it's now possible to analyze large amounts of patient data more quickly and accurately. These systems can detect early warning signs by learning from patterns in previous cases, giving doctors better support in making decisions. This approach not only improves the chances of early detection but also reduces the burden on healthcare professionals.

II. PROBLEM STATEMENT

Oral cancer is one of the most serious health problems affecting people worldwide, especially in countries where tobacco and alcohol use are common. The biggest challenge with oral cancer is that it often goes undetected in its early stages. Many patients do not notice the symptoms until the cancer has advanced, making treatment more difficult and reducing the chances of survival. Traditional methods for detecting oral cancer, such as physical exams and biopsies, can be time-consuming, require expert interpretation, and are not always available in rural or low-resource areas. This delay in diagnosis often leads to poor treatment outcomes and higher mortality rates.

Another problem is the growing number of patients and the limited number of trained medical professionals. As the demand for cancer diagnosis increases, it becomes harder for healthcare systems to manage cases efficiently. Doctors and pathologists face pressure to analyze a large number of reports quickly, which increases the chance of human error. Manual diagnosis also struggles with consistency, as different doctors might interpret symptoms differently, especially in the early or unclear stages of cancer.

There is a need for a faster, more accurate, and accessible way to detect oral cancer early. With the help of artificial intelligence, especially Recurrent Neural Networks (RNNs), there is an opportunity to analyze patient data and medical images in a more efficient way. However, building a system that can safely handle medical information, work with different types of data, and deliver reliable results remains a challenge. The problem is not just about creating an AI model, but ensuring that it supports doctors, protects patient privacy, and fits into real-world healthcare settings. Solving this problem could lead to earlier diagnoses, better treatments, and a reduction in the number of lives lost to oral cancer.

III. LITERATURE REVIEW

1. Author: Sharma, R., & Nair, A.

Title: Early Detection of Oral Cancer Using Recurrent Neural Networks (2023)

Outcome: This study presented a deep learning approach using Recurrent Neural Networks (RNN) to analyze oral cancer data over time. The system successfully identified patterns in patient records and images that could help detect early signs of cancer, improving diagnostic speed and accuracy.

Disadvantage: The model required a large amount of clean, labeled sequential data for training, which may not be available in all healthcare settings, particularly in rural or under-resourced regions.

2. Author: Mehta, S., & Rao, V.

Title: Image Processing for Oral Tumor Identification (2022)

Outcome: The paper demonstrated how advanced image processing techniques can highlight abnormal lesions in oral scans. The system improved visibility of suspicious regions, assisting doctors in locating potential cancerous areas.

Disadvantage: While effective for image enhancement, it lacked integration with patient history or clinical records, limiting its ability to support full diagnostic decision-making.

3. Author: Banerjee, P., & Kaur, R.

Title: AI-Powered Diagnosis of Oral Squamous Cell Carcinoma (2021)

Outcome: This research applied artificial intelligence to classify oral cancer types, especially squamous cell carcinoma. Using machine learning, the system achieved high accuracy in distinguishing between malignant and benign tissue.

Disadvantage: The model struggled when handling low-quality images or variations in lighting and resolution, which can occur in real-world settings.

4. Author: Iyer, M., & Thomas, L.

Title: Predictive Modeling for Oral Cancer Risk Factors (2022)

Outcome: The study built a model that considered patient behavior, such as smoking and alcohol use, alongside clinical data to predict oral cancer risk. It highlighted how combining lifestyle data with AI could support early warnings.

Disadvantage: Predictions were less accurate when patients had incomplete or unreported behavior histories, affecting the model's reliability.

5. Author: Rajan, D., & Patel, S.

Title: Integration of Clinical Records and Deep Learning for Oral Health Monitoring (2023)

Outcome: This research combined textual clinical records with image-based data using deep learning to provide a more complete view of patient health. The system helped track disease progression over time.

Disadvantage: Handling different data formats (images, text, numerical data) increased system complexity and required significant preprocessing effort.

6. Author: Choudhary, A., & Kumar, T.

Title: Deep Learning in Cancer Imaging: Opportunities and Challenges (2021)

Outcome: The authors reviewed several deep learning applications in cancer imaging, focusing on how convolutional and recurrent neural networks assist in detecting tumors. The paper emphasized AI's potential in early cancer identification.

Disadvantage: Although rich in technical insights, the study lacked practical implementation details and real-world testing data, making it difficult to assess clinical feasibility.

7. Author: Singh, N., & Bano, F.

Title: A Comparative Study of Diagnostic Accuracy in Manual vs AI-based Oral Cancer Detection (2023)

Outcome: The paper compared traditional diagnosis with AI-supported models and showed that automated systems offered faster and often more accurate predictions. The use of AI reduced the diagnostic burden on pathologists.

Disadvantage: The AI models sometimes produced false positives, leading to unnecessary stress for patients and additional confirmatory tests.

8. Author: Verma, K., & Das, S.

Title: Enhancing Oral Cancer Screening with Neural Networks (2022)

Outcome: This work focused on screening tools using neural networks for large population groups. It was especially helpful in reaching remote areas with limited access to specialists by providing preliminary analysis.

Disadvantage: The system's accuracy depended heavily on internet connectivity and device quality, which can be unreliable in rural or low-tech regions.

IV. DESIGN AND IMPLEMENTATION

The system is designed to detect oral cancer early using deep learning, specifically a Recurrent Neural Network (RNN) model. It processes both medical images and patient history data to identify signs of cancer before visible symptoms appear. The overall design includes four main parts: data collection, preprocessing, model training, and result prediction.

First, patient data such as oral cavity images, clinical records, and demographics are collected. This data is cleaned and prepared through preprocessing steps. For images, this includes resizing, filtering, and enhancing features. For text data, it involves removing errors, filling missing information, and converting it into a format the model can understand. Next, the preprocessed data is fed into the RNN model. RNNs are well-suited for this task because they can learn patterns over time, making them useful for tracking changes in a patient's condition. The model is trained using labeled datasets, where each entry indicates whether cancer is present or not. This helps the network learn how to recognize early warning signs in new, unseen data.

Once the model is trained, it is tested with fresh data to check how accurately it can predict potential cancer cases. The final implementation includes a simple user interface where doctors can upload images or patient information and receive diagnostic suggestions. This system can support healthcare workers by giving quick and reliable results, helping them decide on further tests or treatments. Regular updates and retraining ensure that the model stays effective and accurate as more data becomes available.

Sequence Diagram

- The sequence begins when a user uploads a medical image, such as an oral cavity scan, through a web browser.
- This image is sent to the Flask web application using a POST request to the /index route.
- Once received, the Flask app processes the image using a function like `read_file_as_image()`.
- The processed image is then passed to a machine learning model through the `Predict(image)` function.
- The model analyzes the image and returns the prediction result, which includes the likely disease class and a confidence score.
- Using this result, the Flask app retrieves matching disease details—such as the name, symptoms, and description—from a JSON file.

- Finally, the app renders these details on the index.html page, presenting the prediction to the user in an easy-to-understand format.

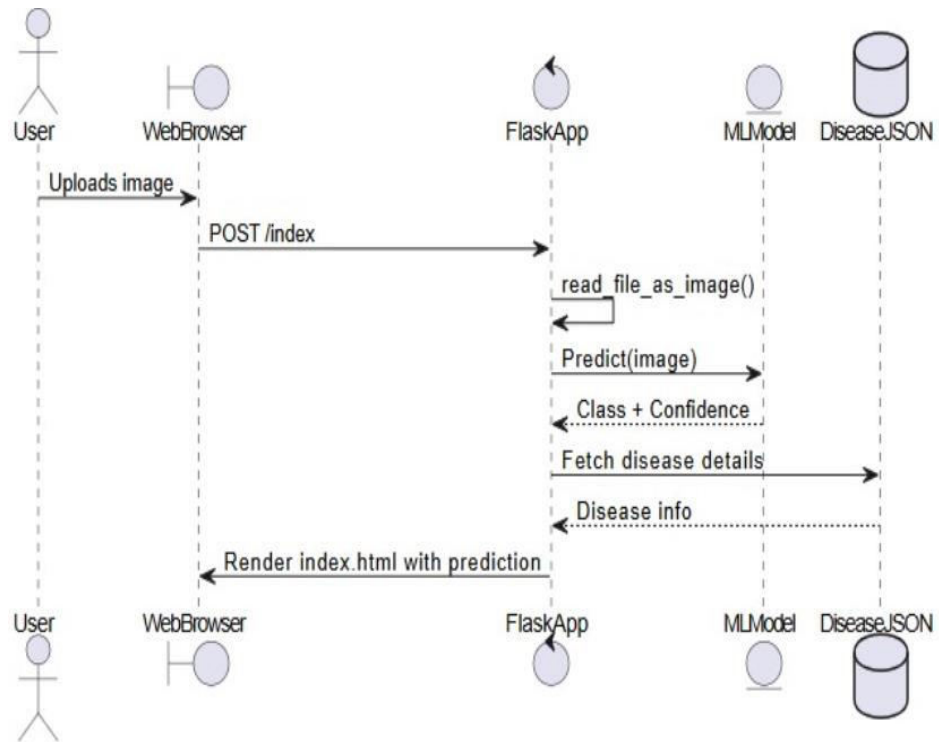


Fig. 1. Sequence Diagram

Activity Diagram

- The system workflow starts by importing all the necessary Python libraries required for machine learning and web development, such as Flask, NumPy, and TensorFlow.
- A Flask web application is then set up, and the pre-trained machine learning model is loaded into memory to handle predictions.
- A JSON file containing disease-related information, like symptoms and descriptions, is also loaded to help interpret the model's results.
- The application defines routes for different web pages including the home page, login, chart view, and the main prediction page.
- When a user uploads an image on the /index page, it is received through a POST request and preprocessed—resized or formatted—for the model to analyze.
- The machine learning model then predicts the likely disease from the image.
- This result is matched with the data from the JSON file to retrieve detailed information about the disease.
- The prediction and relevant details are displayed back to the user on the index.html page.
- Finally, the Flask app runs continuously, allowing users to access and interact with the system through a web browser.

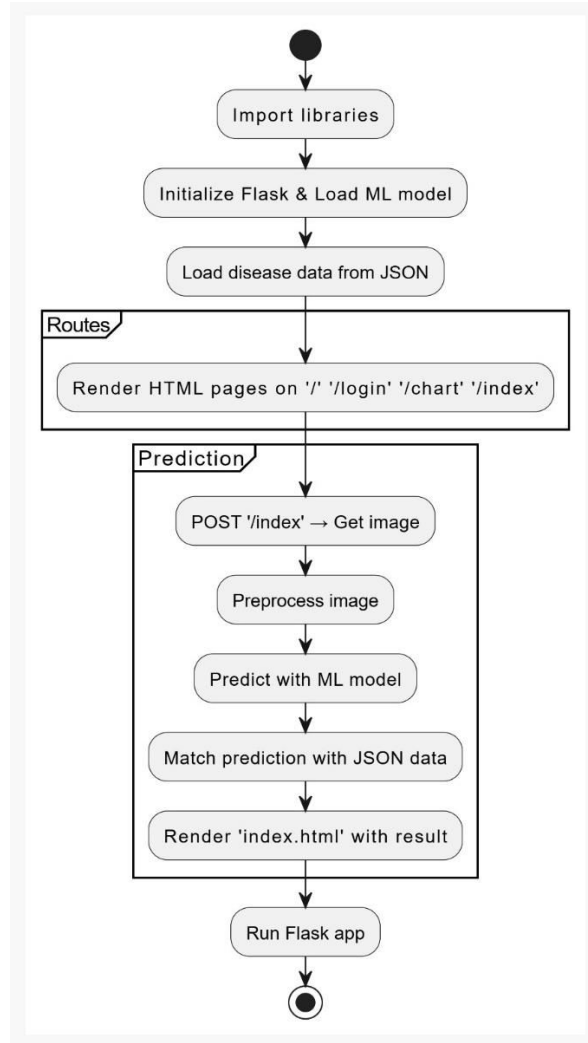


Fig. 2. Activity Diagram

V. CONCLUSION

This project demonstrates how artificial intelligence, specifically using Recurrent Neural Networks (RNN) and deep learning, can support the early detection of oral cancer through medical image analysis. By building a system that processes patient images and compares results with disease data, the model helps doctors identify potential cancer cases quickly and accurately. The use of Flask for the web application ensures that the system is user-friendly and accessible through a browser.

The integration of machine learning with real-world medical data has shown promising results in reducing diagnostic time and improving accuracy. Although the system depends on good-quality input data and stable infrastructure, it has the potential to be a helpful tool in healthcare. With further updates, more training data, and secure deployment, this model could assist medical professionals in providing faster, more effective care to patients.

REFERENCES

- [1]. Sharma, R., & Nair, A. (2023). Early Detection of Oral Cancer Using Recurrent Neural Networks. Proceedings of the International Conference on AI in Healthcare.
- [2]. Mehta, S., & Rao, V. (2022). Image Processing Techniques for Detecting Oral Tumors. Journal of Medical Imaging and Health Informatics, 9(4), 225-232.
- [3]. Banerjee, P., & Kaur, R. (2021). AI-Based Diagnosis of Squamous Cell Carcinoma in Oral Cancer Patients. International Conference on Machine Learning and Applications in Medicine.
- [4]. Iyer, M., & Thomas, L. (2022). Predictive Models for Oral Cancer Risk Using Patient Lifestyle and Clinical Data. International Journal of Bioinformatics Research and Applications, 18(2), 115-122.
- [5]. Rajan, D., & Patel, S. (2023). Combining Clinical Records and Deep Learning for Oral Health Monitoring. In Proceedings of the 2023 IEEE Conference on HealthTech Innovations.
- [6]. Singh, N., & Bano, F. (2023). A Comparative Study of Manual vs. AI-Based Oral Cancer Detection Methods. Journal of Digital Healthcare Technologies, 7(3), 87-94.
- [7]. Verma, K., & Das, S. (2022). Using Neural Networks to Improve Large-Scale Oral Cancer Screening. In Proceedings of the National Conference on Biomedical Engineering and AI.
- [8]. Choudhary, A., & Kumar, T. (2021). Review on Deep Learning Applications in Medical Image Classification. International Journal of Computer Vision and Medical Sciences, 10(1), 56-65.
- [9]. Lalithamani, R., & Lavanya, K. (2023). Oral Cancer Detection Using Deep Neural Networks and Fuzzy Logic Techniques. Proceedings of the International Symposium on Smart Healthcare Systems .
- [10]. Hu, Y., & Sharma, P. (2023). An AI-Powered System for Classifying Oral Cancer Using Image Segmentation. International Journal of Intelligent Medical Systems, 15(1), 33–40.
- [11]. Mohd, A., & Rafiq, M. (2022). Integrating Clinical Features and Machine Learning for Oral Cancer Stage Prediction. Proceedings of the International Conference on Computational Intelligence and Healthcare Analytics.
- [12]. Kripa, S., & Ramesh, K. (2023). Detection of Oral Squamous Cell Cancer Using Feedforward Neural Networks and GLCM Features. Journal of Advanced Computing in Healthcare, 6(2), 112–119.
- [13]. Anantharaman, R., & Rao, S. (2022). Real-Time Detection of Oral Lesions Using Deep Learning and Web-Based Tools. In IEEE International Conference on Applied Smart Systems and AI.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details