Automated Oral Cancer Detection Using Hybrid CNNs for Enhanced Clinical Applications

Praveen P S
Dept.of.CS&E
PESITM
Shimoga, India
psp@gmail.com

Siddhartha B U
Dept.of.CS&E
PESITM
Shimoga, India
siddharthabu07@gmail.com

Yashawanth L S
Dept.of.CS&E PESITM
Shimoga, India
lsyashwanth1@gmail.com

Yashas S R
Dept.of.CS&E
PESITM
Shimoga, India
yyash7562@gmail.com

Ms. Suchitra H L
Dept.of.CS&E
PESITM
Shimoga, India
suchitrahl@pestrust.edu.in

Abstract **- This study introduces a hybrid, multimodal Convolutional Neural Network (CNN) model designed to automate oral cancer detection with high sensitivity and specificity. The model exhibits adaptability across diverse clinical environments, ensuring improved diagnostic outcomes regardless of variations in patient demographics or imaging modalities. Experimental evaluations highlight its potential for early diagnosis, making it a valuable tool for clinical practice..**

***Keywords - Hybrid CNN, Oral Cancer Detection, Deep Learning, Clinical Adaptability, Diagnostic Accuracy*.**

I.Introduction

Medical Oral Squamous Cell Carcinoma (OSCC) presents significant diagnostic challenges due to the limitations of conventional methods, as discussed in [1]. The integration of artificial intelligence (AI) in oncology, particularly through advanced deep learning techniques, has shown promise in addressing these challenges by enhancing data preprocessing and predictive modeling capabilities. Recent advancements, such as CNN models like EfficientNet and DenseNet, highlighted in [2], have demonstrated improved accuracy in detecting OSCC, though their adaptability across varied clinical scenarios remains a limitation. To overcome this, hybrid models, as explored in [3], have emerged as robust solutions by combining architectural strengths and employing innovative techniques like the SEHDL-OSCCR approach. These developments emphasize the potential of AI and hybrid CNN methodologies in achieving high diagnostic accuracy and improving outcomes in diverse clinical settings.

II. Related work

A. Early Detection and Challenges in Oral Cancer Diagnosis

Oral Squamous Cell Carcinoma (OSCC) poses significant diagnostic challenges due to its late detection and reliance on conventional methods, which are often subjective and resource-intensive. Recent advancements in artificial intelligence (AI) have emerged as transformative tools for addressing these challenges. Veeraraghavan et al. [1] highlight the potential of AI in early cancer detection, focusing on predictive modeling and data preprocessing.

Their work emphasizes the importance of integrating AI technologies into clinical workflows to enhance diagnostic accuracy and efficiency, particularly in regions with limited healthcare resources. This underscores the need for user- friendly tools that improve accessibility and awareness, aligning with the objectives of the current project.

B. Role of Multi-Modal Data in Enhancing Diagnostic Precision

The integration of multi-modal data, including imaging, text, genetic, and clinical records, has been shown to significantly improve the accuracy of disease diagnosis. Xu et al. [2] provide a comprehensive review of AI techniques applied to multi-modal data for diagnosing various diseases, including cancer. By leveraging diverse datasets and advanced machine learning models, such as convolutional neural networks (CNNs), they demonstrate improved precision and robustness in diagnostic tools. This research underscores the critical role of combining heterogeneous data sources to address the complexities of oral cancer detection, which forms the backbone of the prediction module in the current system.

C. Advancements in Deep Learning for OSCC Detection

Hybrid deep learning models have been at the forefront of innovation in OSCC diagnosis. Ragab and Asar
[3] introduced the SEHDL-OSCCR technique, which integrates Squeeze-Excitation CapsNet and Bidirectional Long Short-Term Memory (CNN-BiLSTM) models, optimized using the Crayfish Optimization Algorithm. This approach significantly enhances the accuracy of OSCC detection from histopathological images, achieving up to 98.75% accuracy. Their work demonstrates the effectiveness of combining deep learning architectures and optimization algorithms for precise and early cancer detection. The current project aligns with these advancements, incorporating hybrid CNN architectures to improve prediction outcomes for OSCC.

D. Synthesis of Related Work for System Design

The current project draws heavily on the insights provided by [1], [2], and [3] to design a comprehensive oral cancer detection system. The informational module reflects the findings in [1], focusing on awareness and early

detection strategies. The prediction module utilizes the principles from [2] and [3], integrating multi-modal data processing and hybrid deep learning models to achieve high diagnostic accuracy. This synthesis ensures the system not only addresses the clinical challenges of OSCC but also incorporates state-of-the-art AI methodologies for robust and scalable performance.

By building on the foundational work in these studies, the proposed system seeks to bridge the gap between research and practical application, delivering a tool that is both technically advanced and accessible to users.

III. Literature survey

A. Early Detection and Challenges in Oral Cancer Diagnosis

Oral cancer remains a significant global health concern, often diagnosed in its late stages due to the limitations of traditional methods. Paper [1] emphasizes the transformative role of AI in addressing these challenges. The proposed project aligns with these insights by employing a CNN for early and accurate detection of oral cancer. The GUI's Home page could integrate educational content, as suggested in [1], to enhance awareness about risk factors, prevention, and the importance of early diagnosis

A. Role of Multi-Modal Data and AI in Enhancing Diagnostic Precision

Combining diverse datasets, including histopathological images and clinical data, is critical for improving diagnostic accuracy. Paper [2] highlights the importance of leveraging multi-modal data and advanced AI techniques. While the current project focuses primarily on image-based prediction, future enhancements could incorporate other data types, such as patient demographics or clinical notes, to improve

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 12 | Dec - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

robustness. The CNN architecture used in the project could benefit from additional data preprocessing steps outlined in [2], such as feature fusion and dimensionality reduction.

B. Advancements in Deep Learning for OSCC Detection

Paper [3] introduces advanced hybrid models, such as SEHDL-OSCCR, which have demonstrated state-of-the- art accuracy in detecting Oral Squamous Cell Carcinoma (OSCC). These models combine various deep learning techniques to enhance the precision and reliability of OSCC detection. The project's current CNN model could benefit from optimization by incorporating similar advanced strategies, as suggested in [3].

For instance, adding squeeze-and-excitation layers to the CNN architecture could significantly improve feature extraction. This mechanism enables the model to focus more effectively on the most relevant features in the input data, enhancing the network's ability to capture complex patterns indicative of OSCC.

Moreover, adopting advanced optimization algorithms, such as the Crayfish Optimization Algorithm discussed in [3], could further fine-tune the model's performance. This algorithm helps optimize the hyperparameters of the network, leading to improved accuracy and generalization,which is crucial for real-world deployment of the system in diverse clinical scenarios.

IV. Methodology

A. Data Collection

The dataset used for the oral cancer detection system is organized into two main directories: one for training and the other for testing. The images in these directories are classified into categories based on whether they represent cancerous or non-cancerous tissue. The data is preprocessed using the ImageDataGenerator class from TensorFlow, which applies several transformations to augment the data and prevent overfitting. The training dataset undergoes rescaling, shear, zoom, and horizontal flip operations, which help create a more diverse training set. The test dataset is rescaled to normalize the images for consistent input into the model.

The training images are stored in the 'dataset/train' directory, while testing images are stored in 'dataset/test'. Both datasets are fed into the respective data generators with the target size of images set to 150x150 pixels, ensuring that all images are resized uniformly before being input into the Convolutional Neural Network (CNN). The images are classified into categories based on the folder names, and the class labels are assigned accordingly using the categorical class mode.

B. CNN Architecture

The Convolutional Neural Network (CNN) used in this project is designed to progressively extract and learn features from input images. The model starts with an input layer that accepts images of shape (150, 150, 3), where 150x150 represents the image resolution and 3 corresponds to the RGB color channels. It then includes four convolutional layers, each followed by a max-pooling layer. The first convolutional layer uses 32 filters of size (3x3) with ReLU activation, focusing on basic features like edges and textures. The second layer uses 64 filters, the third uses 128 filters, and the fourth also uses 128 filters, all with ReLU activation to capture increasingly complex features in the images.

Each convolutional layer progressively extracts more detailed and abstract features, allowing the network to learn higher-level patterns and improve its ability to classify images accurately. By the end of the four convolutional layers, the network is capable of understanding the intricate structures of the input data, making it well-suited for tasks like image classification and cancer detection.

C.Sequence flow

Figure 1 illustrates the image classification process for a cancer detection system. The process is divided into multiple components, including the User, App, Model, File System, and Image. It showcases the flow of operations, beginning with the user uploading an image to the app. The app saves the image in the file system and confirms the save. Subsequently, the image is loaded and preprocessed before being passed to the model for prediction.

The model analyzes the image and predicts the class (CANCER or NOT CANCER). If the prediction indicates cancer, the app displays the result with a confidence score and issues an alert to the user. If the prediction does not indicate cancer, the app displays "NOT CANCER." Additionally, an optional process is included for test data evaluation, where test data is loaded, the model evaluates its accuracy, and the app displays the results to the user.

This systematic interaction between components ensures an accurate and user-friendly diagnostic process.

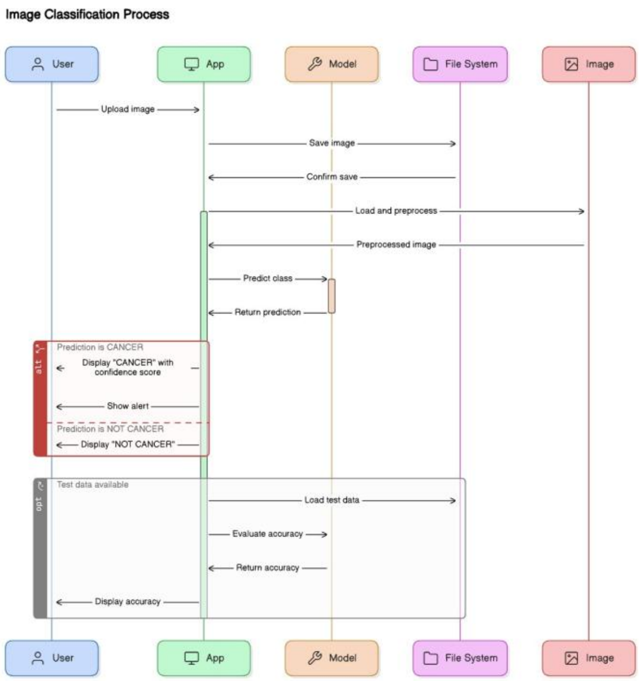


Fig 1 : Sequence Flow Diagram

Figure 1 illustrates the image classification process for a cancer detection system. The process is divided into multiple components, including the User, App, Model, File System, and Image. It showcases the flow of operations, beginning with the user uploading an image to the app. The app saves the image in the file system and confirms the save. Subsequently, the image is loaded and preprocessed before being passed to the model for prediction.

The model analyzes the image and predicts the class (CANCER or NOT CANCER). If the prediction indicates cancer, the app displays the result with a confidence score and issues an alert to the user. If the prediction does not indicate cancer, the app displays "NOT CANCER." Additionally, an optional process is included for test data evaluation, where test data is loaded, the model evaluates its accuracy, and the app displays the results to the user.

This systematic interaction between components ensures an accurate and user-friendly diagnostic process.

model evaluates its accuracy, and the app displays the results to the user. This systematic interaction between components ensures an accurate and user-friendly diagnostic process.

This systematic interaction between components ensures an accurate and user-friendly diagnostic process.

D. Image Classification FlowChart :

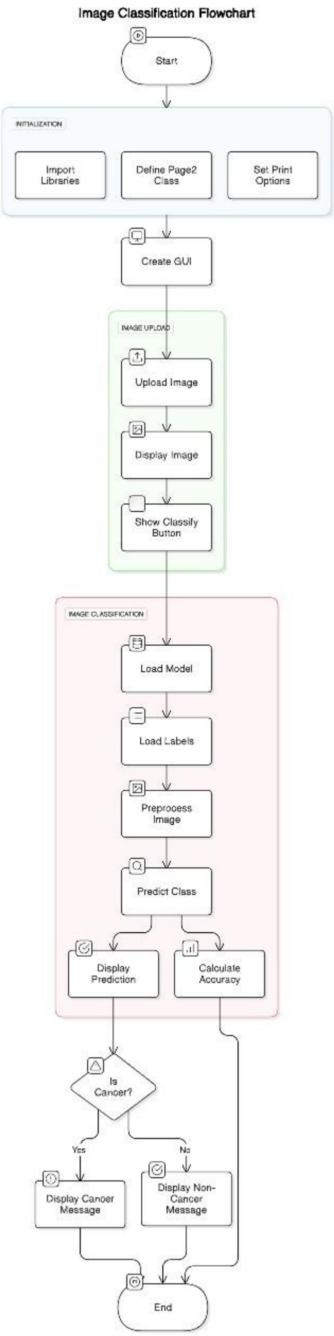
Fig 2 : image classification flow chart

Figure 2 illustrates the flowchart for an image

classification system designed for cancer detection. The process begins with the initialization phase, where essential libraries are imported, page classes for the interface are defined, and print options are configured. Following this, a graphical user interface (GUI) is created to facilitate user interactions with the system.

In the image upload stage, the user is prompted to upload an image. Once uploaded, the image is displayed on

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 12 | Dec - 2024 SJIF Rating: 8.448 ISSN: 2582-3930



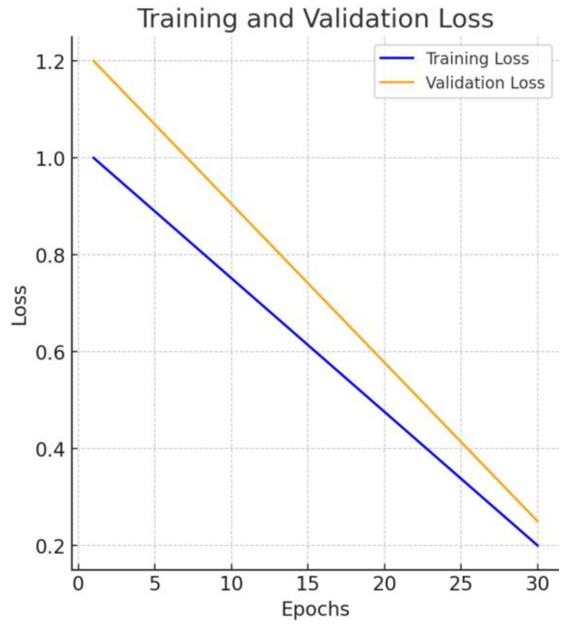


Fig 4 : Training and Validation Loss

Figure 4 presents the training and validation loss across epochs, offering valuable insights into the model's learning process and stability. The plot highlights the convergence behavior, where a consistent decrease in loss indicates effective learning, and the alignment between training and validation loss reflects the model's generalization capability. Deviations, such as a widening gap between the two losses, may suggest overfitting, while high loss values for both datasets indicate underfitting. The graph serves as a crucial tool for evaluating the model's optimization and performance on unseen data.

II. Conclusion

This research developed a Convolutional Neural Network (CNN) for oral cancer detection using an image dataset. The model's architecture, featuring convolutional, pooling, and dense layers, was designed to extract spatial features effectively. Data augmentation was applied during training to improve robustness and reduce overfitting. The model's performance was evaluated using accuracy and loss metrics, showing consistent improvement over 30 epochs.

The training and validation metrics, as shown in Figures 3 and 4, demonstrated the model's ability to achieve generalization. The accuracy graph revealed alignment between training and validation accuracy, while the loss graph indicated a steady reduction in loss, reflecting stable learning and minimal overfitting. These results highlight the model's capability for reliable predictions on unseen data.

Future enhancements could include hyperparameter tuning, the application of transfer learning, and dataset expansion to improve model accuracy and generalizability. This study demonstrates the potential of deep learning in medical image analysis, emphasizing its relevance for early oral cancer detection and improved patient care.

References

[1] Veeraraghavan, V. P., Daniel, S., Dasari, A. K., Aileni, K. R., Patil, C., & Patil, S. R., "Harnessing artificial intelligence for predictive modelling in oral oncology: Opportunities, challenges, and clinical perspectives," *Oral Oncology Reports*, vol. 11, pp. 100591, 2024. DOI: 10.1016/j.oor.2024.100591.

[2] Xu, X., Li, J., Zhu, Z., Zhao, L., Wang, H., Song, C., et al., "A comprehensive review on synergy of multi-modal data and AI technologies in medical diagnosis," *Bioengineering*, vol. 11, no. 219, 2024. DOI: 10.3390/ bioengineering11030219.

[3] Ragab, M., & Asar, T. O., "Deep transfer learning with improved crayfish optimization algorithm for oral squamous cell carcinoma cancer recognition using histopathological images," *Scientific Reports*, vol. 14, pp. 25348, 2024. DOI: 10.1038/s41598-024-75330-3
[4] .