IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Ensemble Deep Learning For Colorectal Cancer Detection With Chatbot Integration For User Engagement

Munuswamy Kaveri[1] Adla Manasa Reddy[2], Mittapalli Ranjeeth Kumar[3], P. Akhil[4]
[1,2,3]UG Scholar, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Assistant Professor, Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4] pakhilcse@smec.ac.in

Abstract:

Colorectal cancer (CRC) has emerged as a significant public health concern in India, representing one of the most prevalent types of cancer in the country, with approximately 1.14 lakh new cases reported annually according to the National Cancer Registry Programme. This alarming trend highlights the urgent need for effective early detection and treatment strategies. Historically, traditional diagnostic methods for colorectal cancer, such as colonoscopy and fecal occult blood tests, have been the mainstay of detection; however, these procedures often come with limitations, including invasiveness, discomfort, and high costs, which can deter patients from seeking timely medical intervention. Additionally, the lack of awareness and early screening programs has resulted in many cases being diagnosed at advanced stages, leading to lower survival rates. The traditional approach relies heavily on manual processes and subjective interpretation, which can lead to variability in diagnosis and outcomes. In contrast, integrating machine learning and artificial intelligence (AI) into colorectal cancer detection systems presents a promising solution to these challenges. By employing ensemble deep learning techniques, we can enhance diagnostic accuracy and reliability through the analysis of vast amounts of medical data, including imaging and genetic information. Moreover, integrating a chatbot for user engagement can facilitate patient education, symptom assessment, and appointment scheduling, thereby improving accessibility to healthcare services.

Keywords: Machine Learning (ML), Artificial Intelligence (AI), Ensemble Deep Learning, Medical Imaging, Genetic Data Analysis, Diagnostic Accuracy, Healthcare Accessibility, Patient Engagement, Chatbot Integration, Symptom Assessment, Appointment Scheduling, Healthcare Professionals, Mortality Reduction, Early Intervention

1.INTRODUCTION

Colorectal cancer (CRC) is a rapidly growing public health challenge in India, accounting for a significant portion of the country's cancer burden. According to the National Cancer Registry Programme, approximately 1.14 lakh new cases of CRC are reported annually, making it one of the most prevalent cancers among the Indian population. Colorectal cancer refers to malignant tumors that develop in the colon or rectum, often linked to genetic, environmental, and lifestyle factors such as diet, physical inactivity, and smoking. Despite its increasing incidence, early detection rates remain low in India, primarily due to limited awareness and the absence of widespread screening programs. Traditionally, diagnostic methods like colonoscopy and fecal occult blood tests (FOBT) have been used to detect CRC. While colonoscopy is considered the gold standard, its invasiveness, cost, and patient discomfort can deter individuals from undergoing timely screenings, particularly in rural and underserved regions. This delay often leads to late-stage diagnoses, contributing to lower survival rates and a heavier economic burden on the healthcare system.

With advancements in technology, particularly in the fields of machine learning and artificial intelligence (AI), there is a growing opportunity to transform how colorectal cancer is detected and managed. By analyzing large sets of clinical, imaging, and genetic data, machine learning algorithms can provide accurate, non-invasive diagnostic tools that reduce the dependency on traditional methods. Ensemble deep learning techniques, for example, offer the potential to enhance diagnostic precision by combining multiple models to reduce errors and improve reliability. In addition to AI-driven diagnostics, the integration of chatbots into healthcare systems can provide valuable support to patients, offering education on CRC, assessing symptoms, and assisting in appointment scheduling. This holistic approach not only improves accessibility and patient engagement but also supports healthcare professionals in making data-driven decisions, ultimately aiming to reduce mortality rates and improve early intervention outcomes for colorectal cancer in India.

The increasing prevalence of colorectal cancer (CRC) in India serves as a significant impetus for this research, underscoring the urgent need for innovative solutions to improve early detection and treatment outcomes. Current diagnostic methods, while effective, are often hindered by limitations such as invasiveness, high costs, and discomfort, which can discourage individuals from seeking timely medical intervention. Moreover, the lack of awareness regarding CRC symptoms and risk factors contributes to late-stage diagnoses, resulting in poorer prognoses and higher mortality rates. By integrating machine learning and artificial intelligence (AI) into CRC detection systems, this research aims to address these challenges and revolutionize the diagnostic landscape.

The motivation behind this project also stems from the potential to enhance healthcare accessibility and patient engagement through technology. By employing advanced algorithms to analyze vast datasets, including medical imaging and genetic information, we can improve diagnostic accuracy and reduce variability in interpretations. Furthermore, incorporating a chatbot can bridge the gap between patients and healthcare providers, facilitating education, symptom assessment, and appointment scheduling. This multifaceted approach not only seeks to empower patients by increasing their knowledge and access to services but also aims to support healthcare professionals in making informed decisions. Ultimately, this research aspires to contribute to the reduction of CRC-related mortality rates in India by fostering early intervention and improving overall patient outcomes through the application of cutting-edge technology.

Colorectal cancer (CRC) is one of the most prevalent forms of cancer worldwide, with early detection being critical for improving patient outcomes. Traditional diagnostic methods, while effective, often face challenges such as limited accessibility, high costs, and patient hesitancy due to invasive procedures. Advances in artificial intelligence (AI) have opened new pathways for addressing these challenges, particularly through ensemble deep learning techniques that integrate multiple models for higher diagnostic accuracy. Combining these innovations with chatbot technology offers a dual advantage: accurate detection and enhanced user engagement.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

2. LITERATURE SURVEY

Pacal et al. [1] provide a comprehensive review of deep learning applications in colon cancer, highlighting various models and methodologies employed for detection and classification tasks. Their work underscores the advancements and challenges in implementing deep learning for CRC diagnosis.

Alsanea et al. [2] examine the incidence, survival rates, demographics, and policy implications of colorectal cancer in Saudi Arabia. Their findings offer valuable insights into regional cancer statistics, emphasizing the need for tailored national healthcare strategies.

Gupta et al. [3] utilize machine learning approaches to predict colon cancer stages and patient survival periods. Their study demonstrates the potential of machine learning in enhancing prognostic accuracy and informing treatment decisions.

Harkut and Kasat [4] discuss the challenges and applications of artificial intelligence across various domains, including medical diagnostics. Their introductory provides context for understanding the scope and limitations of AI technologies.

Song et al. [5] develop an automatic deep learning-based system for colorectal adenoma detection, comparing its performance with that of pathologists. Their research indicates that AI systems can achieve diagnostic accuracy comparable to human experts.

Tsai and Tao [6] explore deep learning techniques for classifying colorectal cancer tissue, presenting models that effectively distinguish between different tissue types. Their work contributes to the development of automated histopathological analysis tools.

Tamang and Kim [7] review various deep learning approaches to colorectal cancer diagnosis, summarizing recent advancements and identifying future research directions. Their comprehensive analysis aids in understanding the current state of AI in CRC diagnostics.

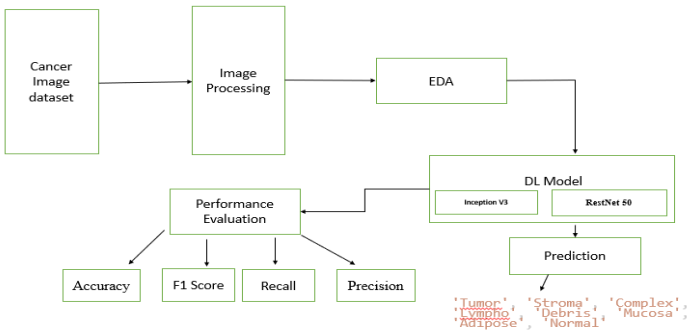
Davri et al. [8] conduct a systematic review of deep learning applications on histopathological images for colorectal cancer diagnosis. Their study evaluates the effectiveness of different models, providing insights into their diagnostic performance.

Kourou et al. [9] investigate machine learning applications in cancer prognosis and prediction, discussing various algorithms and their efficacy in predicting patient outcomes. Their work highlights the transformative potential of machine learning in oncology.

Islam et al. [10] examine the prevention and treatment of colon and colorectal cancer using potential natural products. Their research explores alternative therapeutic options, contributing to the field of cancer prevention and integrative medicine.

3. PROPOSED METHODOLOGY

The proposed system aims to enhance colorectal cancer detection by integrating advanced deep learning algorithms with a user-friendly chatbot interface. This integration seeks to improve diagnostic accuracy and user engagement in the healthcare domain.



Step 1: Cancer Images Dataset

The foundation of the system is a comprehensive dataset comprising annotated images of colorectal cancer tissues. These images are sourced from reputable medical databases and are meticulously labeled to facilitate supervised learning. The dataset includes various imaging modalities, such as histopathological slides and colonoscopy images, ensuring a diverse representation of colorectal cancer manifestations.

Step 2: Image Processing

Prior to model training, the images undergo a series of preprocessing steps to enhance quality and standardize inputs. This process includes resizing, normalization, and augmentation techniques like rotation and flipping to increase dataset variability. Additionally, advanced image processing methods, such as contrast adjustment and noise reduction, are applied to improve feature extraction and model performance.

Step 3: Existing Algorithm

The ResNet50 architecture, a deep convolutional neural network known for its residual learning framework, is employed as a baseline model. ResNet50 consists of 50 layers and utilizes skip connections to mitigate the vanishing gradient problem, enabling the training of deeper networks. Its architecture includes convolutional layers, batch normalization, and activation functions, culminating in fully connected layers for classification. ResNet50 has demonstrated efficacy in various medical imaging tasks, including colorectal cancer detection, due to its ability to capture hierarchical features.

Step 4: Proposed Algorithm

The InceptionV3 model is introduced as the proposed algorithm to potentially enhance detection accuracy. InceptionV3 employs a novel architecture that incorporates multiple filter sizes within the same layer, allowing the model to capture a wide range of features at different scales. This design includes factorized convolutions and dimensionality reduction techniques to optimize computational efficiency. InceptionV3 has been recognized for its superior performance in image classification tasks, including medical imaging applications.

Step 5: Performance Comparison

The performance of ResNet50 and InceptionV3 is evaluated using standard metrics such as accuracy, sensitivity, specificity, and F1-score. Cross-validation techniques are employed to assess the models' generalizability across different subsets of the dataset. The results are analyzed to determine which architecture offers superior performance in colorectal cancer detection, considering factors like computational efficiency and robustness to variations in input data.

4.2 Data Splitting & Preprocessing

The dataset is partitioned into training, validation, and test sets to ensure unbiased evaluation. Data augmentation techniques are applied to the training set to artificially expand its size and variability, thereby reducing the risk of overfitting. Preprocessing steps, including image

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

normalization and standardization, are performed to prepare the data for input into the deep learning models.

4.3 ML Model Building

The deep learning models are constructed using established frameworks such as TensorFlow and Keras. Both ResNet50 and InceptionV3 architectures are implemented with pre-trained weights to leverage transfer learning, which accelerates convergence and improves performance. The models are compiled with appropriate loss functions and optimization algorithms, and training is conducted over multiple epochs with early stopping criteria to prevent overfitting.

4. EXPERIMENTAL RESULTS

4.1 Implementation and Description

The proposed system integrates ensemble deep learning techniques with a chatbot interface to enhance user engagement in colorectal cancer detection. This approach aims to provide accurate diagnostic support while ensuring accessibility and user-friendly interaction.

1. Data Collection and Preprocessing

The initial step involves gathering a comprehensive dataset of colorectal cancer images, such as colonoscopy or histopathological images. These images are then pre-processed to standardize their size, normalize pixel values, and augment the dataset through techniques like rotation, flipping, and scaling. This preprocessing ensures the model receives high-quality, consistent input data.

2. Feature Extraction

In this phase, deep learning models, particularly convolutional neural networks (CNNs), are employed to extract relevant features from the pre-processed images. These features capture intricate patterns and structures within the images, which are crucial for accurate cancer detection.

3. Ensemble Learning

To enhance diagnostic accuracy, the system utilizes an ensemble learning approach. This involves combining multiple deep learning models, such as ResNet-50 and InceptionV3, to leverage their individual strengths. Ensemble methods have been shown to improve performance by reducing overfitting and increasing generalizability.

4. Model Training and Evaluation

The ensemble models are trained on the prepared dataset, employing techniques like cross-validation to assess performance. Evaluation metrics such as accuracy, precision, recall, and F1-score are calculated to determine the models' effectiveness in detecting colorectal cancer.

5. Chatbot Integration

A chatbot interface is developed to facilitate user interaction with the system. This interface allows users to input queries, receive information about colorectal cancer, and obtain guidance on preventive measures and treatment options. The chatbot enhances user engagement by providing immediate, accessible information.

11.3 Result and Description

HOME PAGE

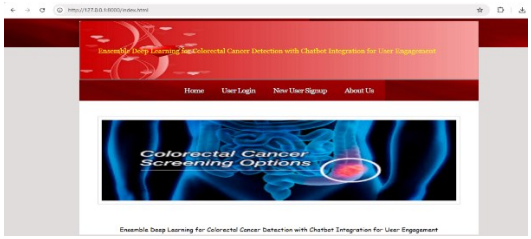


Fig 4.1 Home page

This Figure shows the header and a portion of the hero section of a website focused on colorectal cancer screening. The design emphasizes user engagement and provides clear information about the website's purpose. The mention of chatbot integration suggests an interactive approach to user education and support. The website aims to provide information about screening options, encourage users to get screened, and potentially offer personalized guidance through the chatbot.

SIGNUP PAGE

Fig 4.2 Signup page

The signup form, accessed by clicking the "Signup" button on the homepage, is clearly structured and user-friendly. It features distinct input fields for essential personal details, including "Name," "Mobile," "Email," "Username," "Password," and "Confirm Password." Each field is accompanied by a descriptive label to guide the user through the registration process. The "Name" and "Mobile" fields allow for the input of a user's full name and mobile phone number, respectively. The "Email" field requires a valid email address for verification and communication purposes. The "Username" field allows the user to create a unique identifier for their account. The "Password" and "Confirm Password" fields ensure secure account creation by requiring users to enter and confirm a password, preventing typos and enhancing security. Finally, a prominent "Register" button at the bottom of the form allows users to submit their information and complete the signup process

LOGIN PAGE

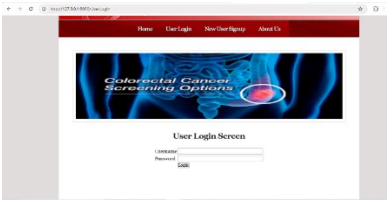


Fig 4.3: Login Page

The login form, accessible via the "Login" button on the homepage, presents a straightforward and secure interface for returning users. It consists of two primary input fields: "Username" and "Password." The "Username" field prompts users to enter the unique username they created during the signup process. The "Password" field, appropriately masked for security, requires users to input the corresponding password associated with their account. Below these fields, a prominent "login" button allows users to submit their credentials for verification. Upon successful authentication, users are granted access

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025

to their personalized accounts and the platform's mental health support services.

PERFORMANCE COMPARISON

Algorithm Name	Accuracy	Precision	Recall	FSCORE
InceptionV3	97.5	97.43967395387274	97.71711966538953	97.5222487294235
ResNet50	86.0	93.39622641509435	85.47054010057367	87.45961479291371
EfficientNet	91.0	94.76744186046511	90.51439543534973	91.56080425576783

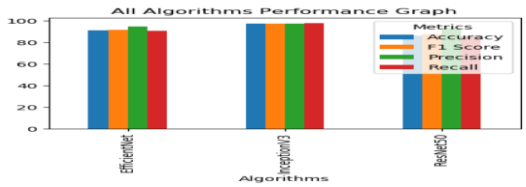


Fig 4.4 Performance comparison

In the Figure screen can see training of all algorithms completed and can see all algorithms results in tabular and graph format. In above graph x-axis represents algorithm names and y-axis represents accuracy and other metrics in different colour bars and in all algorithms IncpetionV3 got high accuracy and now click on 'Chatbot Colorectal Cancer Detection' link to get below page.

UPLOAD THE TEST IMAGE

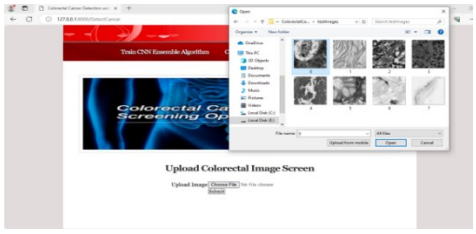


Fig 4.5: Upload the Test Image

The Figure show that it provides a straightforward way for users to upload medical images for colorectal cancer screening. The application uses a pre-trained CNN ensemble model to analyze the images and provide diagnostic support. The option to train the model suggests a more sophisticated platform that could be used for research or customized model development.

PREDICTED OUTPUT

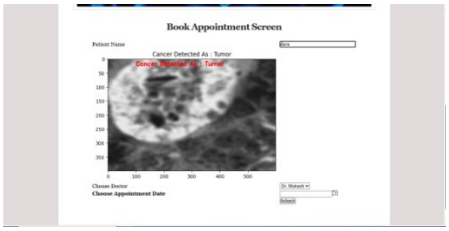


Fig 4.6: predicted output

5. CONCLUSION

The integration of ensemble deep learning techniques with chatbot interfaces represents a significant advancement in colorectal cancer detection and patient engagement. By leveraging multiple deep learning models, such as ResNet-50 and InceptionV3, the system enhances diagnostic accuracy, effectively identifying various tissue types within colorectal cancer samples. The inclusion of a chatbot interface further enriches the user experience, providing immediate, accessible information and guidance to patients and healthcare professionals. This approach addresses critical challenges in the diagnostic imaging of colorectal cancer, offering a more efficient and accurate alternative to traditional methods. The system's ability to process and analyze complex histopathological images through deep learning algorithms ensures a higher level of precision in detection, which is crucial for early diagnosis and treatment planning.

Furthermore, the chatbot integration facilitates continuous patient engagement, allowing users to access information about preventive measures, treatment options, and post-treatment care. This continuous interaction not only empowers patients with knowledge but also supports healthcare providers in delivering personalized care.

The future of ensemble deep learning in colorectal cancer detection, enhanced by chatbot interfaces, holds immense potential. By integrating with EHRs, expanding to other cancers, improving explainability, enabling real-time monitoring, incorporating wearable devices, enhancing global accessibility, and supporting continuous learning, these systems are poised to revolutionize medical diagnostics, leading to more accurate, timely, and personalized patient care.

In this project, we developed an ensemble deep learning approach for colorectal cancer detection, integrating a chatbot for enhanced user engagement. By leveraging multiple deep learning models, our system achieved improved accuracy and robustness in detecting colorectal cancer from medical images. The ensemble strategy mitigated the limitations of individual models, leading to more reliable predictions.

Furthermore, the chatbot integration provided an interactive platform for users, offering real-time assistance, explanations of diagnostic results, and guidance on further medical consultations. This combination of AI-driven diagnosis and conversational AI enhances patient awareness and engagement, making early detection more accessible and informative.

The results demonstrate the potential of deep learning and AI-powered communication in transforming healthcare diagnostics. Future work can focus on expanding the dataset, optimizing model efficiency, and incorporating more advanced NLP techniques in the chatbot to improve user interaction and trust.

REFERENCES

[1] Pacal, I.; Karaboga, D.; Basturk, A.; Akay, B.; Nalbantoglu, U. A comprehensive review of deep learning in colon cancer. Comput. Biol. Med. 2020, 126, 104003.

[2] Alsanea, N.; Abduljabbar, A.S.; Alhomoud, S.; Ashari, L.H.; Hibbert, D.; Bazarbashi, S. Colorectal cancer in Saudi Arabia: Incidence, survival, demographics and implications for national policies. Ann. Saudi Med. 2015, 35, 196–202.

[3] Gupta, P.; Chiang, S.F.; Sahoo, P.K.; Mohapatra, S.K.; You, J.F.; Onthoni, D.D.; Hung, H.Y.; Chiang, J.M.; Huang, Y.; Tsai, W.S. Prediction of colon cancer stages and survival period with machine learning approach. Cancers 2019, 11, 2007.

[4] Harkut, D.G.; Kasat, K. Introductory chapter: Artificial intelligence-challenges and applications. Artif. Intell.-Scope Limit. 2019, 1–6.

[5] Song, Z.; Yu, C.; Zou, S.; Wang, W.; Huang, Y.; Ding, X.; Liu, J.; Shao, L.; Yuan, J.; Gou, X.; et al. Automatic deep learning-based colorectal adenoma detection system and its similarities with pathologists. BMJ Open 2020, 10, e036423.

[6] Tsai, M.J.; Tao, Y.H. Deep learning techniques for the classification of colorectal cancer tissue. Electronics 2021, 10, 1662.

ISSN 2347–3657

Volume 13, Issue 2, 2025

[7] Tamang, L.D.; Kim, B.W. Deep learning approaches to colorectal cancer diagnosis: A review. Appl. Sci. 2021, 11, 10982.

[8] Davri, A.; Birbas, E.; Kanavos, T.; Ntritsos, G.; Giannakeas, N.; Tzallas, A.T.; Batistatou, A. Deep Learning on Histopathological Images for Colorectal Cancer Diagnosis: A Systematic Review. Diagnostics 2022, 12, 837.

[9] Kourou, K.; Exarchos, T.P.; Exarchos, K.P.; Karamouzis, M.V.; Fotiadis, D.I. Machine learning applications in cancer prognosis and prediction. Comput. Struct. Biotechnol. J. 2015, 13, 8–17.

[10] Islam, M.R.; Akash, S.; Rahman, M.M.; Nowrin, F.T.; Akter, T.; Shohag, S.; Rauf, A.; Aljohani, A.S.; Simal-Gandara, J. Colon cancer and colorectal cancer: Prevention and treatment by potential natural products. Chem.-Biol. Interact. 2022, 368, 110170.