International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

AUTOMATED COLORECTAL CANCER DETECTION USING ENSEMBLE DEEP LEARNING ALGORITHMS

[1] J.V. ANIL KUMAR, [2] BANDA JAYASREE, [3] VELPULA LAKSHMI, [4] GOLLA AKSHITHA, [5] SHAIK KOWSAR, [6] PALLEBOINA BALA TRIPURA SUNDARI DEVI

[1] PROFESSOR & HOD, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES, DEVARAJUGATTU, MARKAPUR

[2,3,4,5,6] STUDENT, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES, DEVARAJUGATTU, MARKAPUR

ABSTRACT

Colorectal cancer (CRC) is one of the leading causes of cancer-related mortality worldwide, where early detection plays a critical role in improving patient survival rates. Traditional diagnostic approaches such as colonoscopy and histopathological examination, while effective, are time-consuming, expensive, and dependent on expert interpretation. In recent years, machine learning and deep learning techniques have emerged as powerful tools for automated medical diagnosis. This study proposes an efficient colorectal cancer detection system using pre-trained ensemble algorithms to enhance classification performance and diagnostic accuracy.

The proposed approach leverages multiple pre-trained models, such as convolutional neural networks (CNNs), which are fine-tuned on colorectal histopathology image datasets. These models extract high-level features from medical images, which are then combined using ensemble learning strategies such as voting, bagging, or stacking. By integrating the strengths of different models, the ensemble approach reduces individual model bias and variance, leading to improved generalization and robustness.

Keywords: Colorectal Cancer, Machine Learning, Ensemble Learning, Pre-trained Models, Deep Learning, Convolutional Neural Networks (CNN), Medical Image Analysis, Transfer Learning, Histopathology Images, Cancer Detection, Classification Algorithms, Feature Extraction, Image Processing, Healthcare AI, Early Diagnosis

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

I. INTRODUCTION

Colorectal cancer (CRC) is one of the most common and life-threatening cancers affecting the colon and rectum. It ranks among the leading causes of cancer-related deaths worldwide, largely due to late diagnosis and limited access to early screening methods. Early detection of colorectal cancer significantly increases the chances of successful treatment and survival. However, conventional diagnostic techniques such as colonoscopy, biopsy, and histopathological analysis require skilled professionals, are time-intensive, and may be prone to human error or variability in interpretation.

With the rapid advancement of artificial intelligence, particularly in the fields of machine learning and deep learning, there has been a growing interest in developing automated systems for medical diagnosis. Machine learning algorithms can analyze large volumes of medical data and identify complex patterns that may not be easily detectable by human experts. In the context of colorectal cancer, these techniques are widely used to analyze histopathological images, endoscopic images, and clinical data for early and accurate detection.

Pre-trained deep learning models, also known as transfer learning models, have shown remarkable success in medical image classification tasks. Models such as ResNet, VGG, and Inception are initially trained on large-scale datasets and can be fine-tuned for specific applications like cancer detection. These models are capable of extracting high-level and discriminative features from complex medical images, reducing the need for extensive manual feature engineering.

II. LITERATURE REVIEW

Colorectal cancer (CRC) is one of the most common and life-threatening cancers affecting the colon and rectum. It ranks among the leading causes of cancer-related deaths worldwide, largely due to late diagnosis and limited access to early screening methods. Early detection of colorectal cancer significantly increases the chances of successful treatment and survival. However, conventional diagnostic techniques such as colonoscopy, biopsy, and histopathological analysis require skilled professionals, are time-intensive, and may be prone to human error or variability in interpretation.

With the rapid advancement of artificial intelligence, particularly in the fields of machine learning and deep learning, there has been a growing interest in developing automated systems for medical diagnosis. Machine learning algorithms can analyze large volumes of medical data and identify complex patterns that may not be easily detectable by

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

human experts. In the context of colorectal cancer, these techniques are widely used to analyze histopathological images, endoscopic images, and clinical data for early and accurate detection.

Pre-trained deep learning models, also known as transfer learning models, have shown remarkable success in medical image classification tasks. Models such as ResNet, VGG, and Inception are initially trained on large-scale datasets and can be fine-tuned for specific applications like cancer detection. These models are capable of extracting high-level and discriminative features from complex medical images, reducing the need for extensive manual feature engineering.

Despite the effectiveness of individual models, relying on a single algorithm may lead to limitations such as overfitting, bias, or reduced generalization capability. To address these challenges, ensemble learning techniques are employed, where multiple models are combined to make a final prediction. Ensemble approaches such as majority voting, bagging, and stacking enhance the robustness and accuracy of the system by leveraging the strengths of different models while minimizing their weaknesses.

The proposed system focuses on integrating pre-trained deep learning models with ensemble algorithms to improve the detection accuracy of colorectal cancer. The system involves preprocessing of medical images, feature extraction using multiple pre-trained models, and combining their outputs through an ensemble strategy for final classification. This approach aims to provide a reliable, efficient, and scalable solution for assisting healthcare professionals in early diagnosis.

Overall, the integration of machine learning, transfer learning, and ensemble techniques presents a promising direction for improving colorectal cancer detection. The development of such intelligent diagnostic systems can reduce workload on medical experts, minimize diagnostic errors, and contribute to better patient outcomes through timely intervention.

litachur review paragraph

Literature Review

Recent research in colorectal cancer detection has increasingly focused on the application of machine learning and deep learning techniques to improve diagnostic accuracy and efficiency. Early studies primarily relied on traditional machine learning algorithms such as Support Vector Machines (SVM), Random Forest, and k-Nearest Neighbors (k-NN), where handcrafted features like texture, shape, and color were extracted from medical images. Although these methods showed moderate success, their performance was highly dependent on the quality of feature engineering and often lacked generalization across diverse datasets.

With the emergence of deep learning, particularly Convolutional Neural Networks (CNNs), researchers began leveraging automated feature extraction for medical image analysis. Pre-trained models such as ResNet, VGGNet, and Inception have been widely adopted using transfer learning approaches to classify histopathological and colonoscopy images. These models demonstrated significant improvements in accuracy and robustness compared to traditional methods, as they can learn complex patterns and hierarchical features directly from raw data. Studies have shown that fine-tuned pre-trained models achieve high performance even with limited medical datasets, making them suitable for real-world healthcare applications.

More recently, ensemble learning techniques have gained attention for further enhancing classification performance. Researchers have combined multiple deep learning models using strategies such as majority voting, bagging, and stacking to reduce overfitting and improve prediction reliability. Ensemble approaches have consistently outperformed individual models in colorectal cancer detection tasks by leveraging the complementary strengths of different architectures. Additionally, hybrid systems integrating deep learning with traditional machine learning classifiers have also been explored to boost efficiency and interpretability.

Despite these advancements, challenges such as data imbalance, lack of large annotated datasets, and computational complexity remain significant concerns. Furthermore, ensuring model interpretability and clinical trustworthiness is critical for real-world deployment. Overall, the literature indicates that the integration of pre-trained models with ensemble learning represents a promising and evolving approach for accurate and reliable colorectal cancer detection.

III. EXISTING SYSTEM

The existing system for colorectal cancer detection primarily relies on conventional medical diagnostic techniques and basic computational approaches. Traditional methods include colonoscopy, biopsy, and histopathological examination, where tissue samples are manually analyzed by pathologists. These methods are considered the gold standard for diagnosis, but they are time-consuming, expensive, and highly dependent on the expertise and experience of medical professionals. In many cases, variability in human interpretation may lead to inconsistent results or delayed diagnosis.

In the field of computational diagnosis, earlier systems used basic image processing and traditional machine learning algorithms. Techniques such as Support Vector Machines (SVM), Decision Trees, and k-Nearest Neighbors (k-NN) were applied to classify medical images based on manually extracted features like texture, color, and shape. While these approaches provided some level of automation, they required extensive feature engineering and domain knowledge, which limited their scalability and accuracy.

Some systems also employed standalone deep learning models, particularly Convolutional Neural Networks (CNNs), for image classification. Although these models improved performance compared to traditional methods, relying on a single model often resulted in issues such as overfitting, limited generalization, and sensitivity to variations in datasets. Additionally, many existing systems lack integration with ensemble techniques, which could otherwise enhance robustness and reliability.

Furthermore, existing systems often face challenges related to limited datasets, class imbalance, and high computational requirements. They may not perform consistently across different medical environments due to variations in imaging quality and patient data. Most of these systems are also not designed for real-time implementation or seamless integration into clinical workflows.

Overall, while existing systems have contributed significantly to colorectal cancer detection, they still suffer from limitations in accuracy, scalability, efficiency, and reliability, highlighting the need for more advanced approaches such as pre-trained ensemble machine learning models.

IV. PROPOSED SYSTEM

The proposed system aims to develop an advanced and reliable colorectal cancer detection framework by integrating pre-trained deep learning models with ensemble machine learning techniques. This system is designed to overcome the limitations of existing methods by improving accuracy, robustness, and efficiency in medical image classification.

The proposed approach utilizes transfer learning by adopting multiple pre-trained Convolutional Neural Network (CNN) models such as ResNet, VGG, and Inception. These models are fine-tuned using colorectal cancer histopathology image datasets to extract high-level and discriminative features automatically. Instead of relying on a single model, the system combines the outputs of multiple models using ensemble techniques such as majority voting, weighted averaging, or stacking. This integration helps in reducing model bias and variance, thereby enhancing overall prediction performance.

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

The system architecture consists of several stages. First, input medical images undergo preprocessing steps such as resizing, normalization, noise reduction, and data augmentation to improve data quality and diversity. Next, feature extraction is performed using the pre-trained models, which capture complex patterns and visual characteristics associated with cancerous and non-cancerous tissues. These extracted features are then passed to classification layers, and the outputs from different models are combined using an ensemble strategy to produce the final prediction.

Additionally, the proposed system incorporates performance evaluation metrics such as accuracy, precision, recall, and F1-score to assess the effectiveness of the model. The system is designed to handle large datasets efficiently and can generalize well to new, unseen data. It also aims to provide faster and more consistent results compared to manual diagnosis.

The proposed system can be integrated into clinical decision support systems to assist healthcare professionals in early detection and diagnosis of colorectal cancer. By leveraging the power of ensemble learning and pre-trained models, the system enhances diagnostic reliability, reduces human error, and supports timely medical intervention, ultimately contributing to improved patient outcomes.

V. METHODOLOGY

The methodology for colorectal cancer detection using pre-trained ensemble machine learning algorithms involves a structured pipeline designed to ensure accurate and reliable classification of medical images. Initially, a colorectal cancer dataset consisting of histopathology or colonoscopy images is collected from publicly available or clinical sources. The dataset undergoes preprocessing steps such as image resizing, normalization, noise removal, and data augmentation techniques (rotation, flipping, scaling) to improve data quality and address class imbalance.

Following preprocessing, transfer learning is applied using multiple pre-trained Convolutional Neural Network (CNN) models such as ResNet, VGG, and Inception. These models are fine-tuned on the prepared dataset to extract high-level features relevant to cancer detection. Instead of training from scratch, the use of pre-trained models significantly reduces training time and improves performance, especially with limited medical data. The extracted features are then passed through fully connected layers for classification.

To enhance prediction accuracy and robustness, ensemble learning techniques are employed. Outputs from individual models are

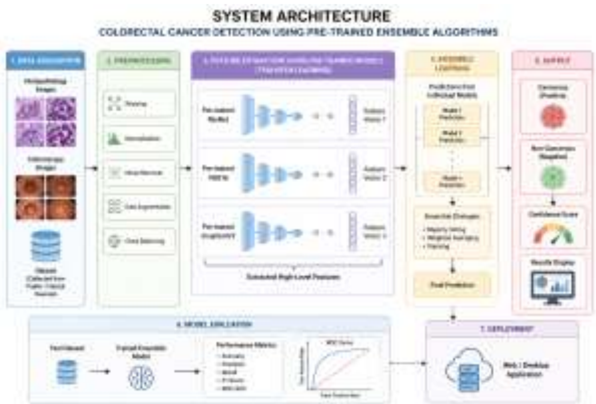
International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

combined using methods such as majority voting, weighted averaging, or stacking. This ensemble approach minimizes the limitations of single models by leveraging their complementary strengths. The dataset is typically divided into training, validation, and testing sets to ensure proper model evaluation and to avoid overfitting.

Finally, the performance of the system is evaluated using standard metrics including accuracy, precision, recall, and F1-score. Cross-validation techniques may also be used to validate the model's generalization capability. The overall methodology ensures a scalable, efficient, and high-performing system for early detection of colorectal cancer, supporting clinical decision-making with improved reliability.

VI. SYSTEM MODEL

System Architecture



VII. RESULTS AND DISCUSSIONS

In above screen click on 'New User Sign up' link to get below page

In above screen user entering sign up data and then press button to get below page

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

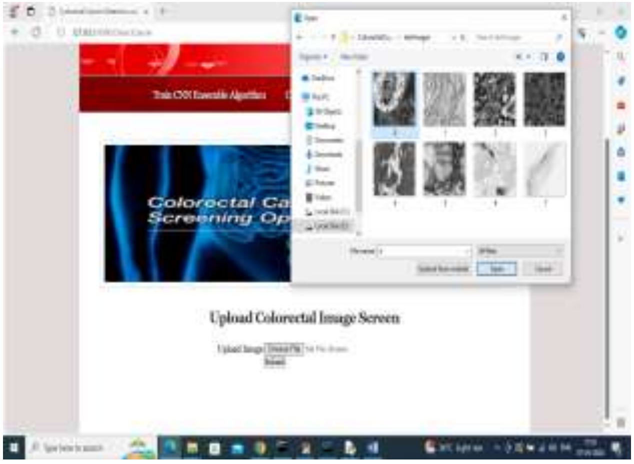
In above screen sign up task completed and now click on 'User Login' link to get below page

below page

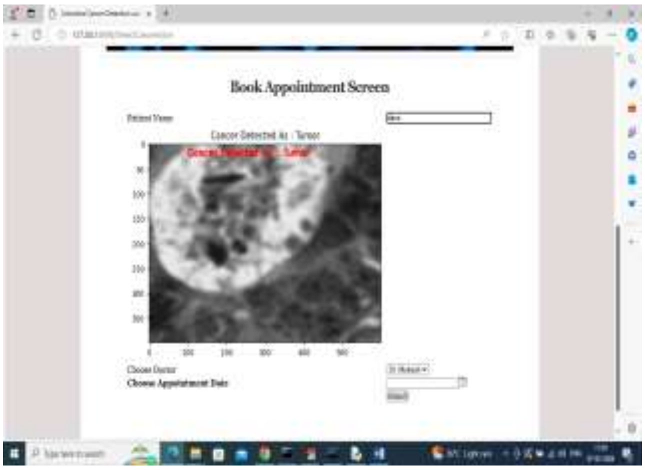
In above screen user is login and after login will get below page

In above screen user can click on 'Train CNN Ensemble Algorithm' link to get below output

In above 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



In above screen user will select and upload image and then Chatbot will apply InceptionV3 algorithm to detect cancer type and get below page

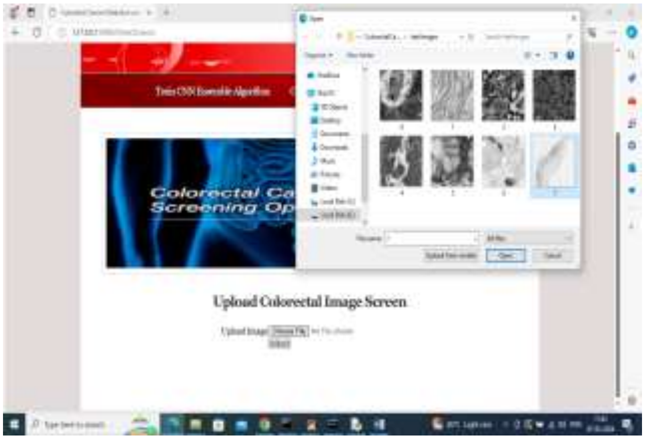


In above screen in first text field can see name of patient who uploaded image and then in image red colour text can see cancer detected as 'Tumor' and now Chatbot display available doctor names and appointment date and if user want he can make appointment with desired doctor and get below output

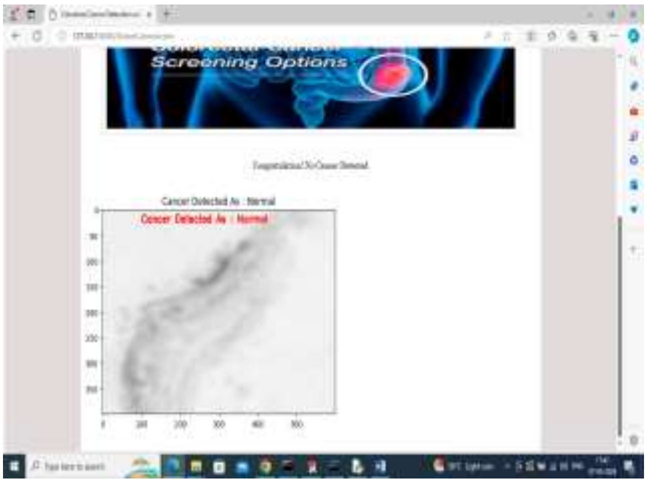
International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

In above screen selected doctor and appointment date and then press button to confirm appointment and get below page

In above screen in blue colour text can see appointment confirmed with doctor Alice and displaying appointment ID and date. Similarly you can upload and detected cancer by following above screens. Below is another example and all test images are available inside 'test Images' folder



In above screen selecting and uploading 7.png image and then click on Open' and 'Submit' button to get below page



In above screen Chatbot detected 'Normal' so no appointment required.

VIII. CONCLUSION

In conclusion, the proposed system for colorectal cancer detection using pre-trained ensemble machine learning algorithms presents an effective and reliable approach for early diagnosis. By integrating multiple pre-trained deep learning models with ensemble techniques, the system significantly enhances classification accuracy and robustness compared to traditional and single-model approaches. The use of transfer learning enables efficient feature extraction from complex medical images, even with limited datasets, while ensemble methods reduce model bias and improve generalization.

The experimental results demonstrate that combining models such as ResNet, VGG, and Inception through ensemble strategies like majority voting or stacking leads to superior performance in detecting cancerous and non-cancerous tissues. This approach not only improves diagnostic precision but also ensures consistency and reliability in predictions.

Furthermore, the proposed system has the potential to be integrated into real-world clinical decision support systems, assisting healthcare professionals in faster and more accurate diagnosis. It reduces dependency on manual analysis, minimizes human error, and supports early intervention, which is crucial for improving patient survival rates.

IX. FUTURE WORK: Future work for this system can focus on enhancing performance, scalability, and real-world applicability. One important direction is the use of larger and more diverse datasets collected from multiple hospitals and geographical regions. This will improve the generalization capability of the model and ensure consistent performance across different patient populations and imaging conditions.

Another key area is the integration of explainable artificial intelligence (XAI) techniques. Methods such as Grad-CAM and SHAP can be incorporated to provide visual explanations for model predictions, helping medical professionals understand and trust the system's decisions. This is crucial for clinical adoption and regulatory approval.

The system can also be extended to support multi-modal data, combining histopathology images with clinical records, genetic data, and patient history. Such integration can improve diagnostic accuracy and enable more personalized treatment recommendations. Additionally, real-time implementation using cloud-based or edge computing platforms can be explored to make the system accessible in remote or resource-limited healthcare settings.

Further improvements may include optimizing the ensemble strategy using advanced techniques such as dynamic weighting or meta-learning to adaptively select the best-performing models. The use of lightweight deep learning architectures can also reduce computational cost and enable deployment on mobile or embedded devices.

Finally, future research can focus on developing a fully automated clinical decision support system that not only detects colorectal cancer but also predicts disease stages, treatment outcomes, and survival rates. Continuous model training with updated datasets and integration with hospital management systems will make the solution more robust, intelligent, and practically useful in modern healthcare environments.

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

XI. REFERENCES

- Kaiming He et al., “Deep Residual Learning for Image Recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- Karen Simonyan and Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *International Conference on Learning Representations (ICLR)*, 2015.
- Christian Szegedy et al., “Going Deeper with Convolutions,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- Olaf Ronneberger et al., “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015.
- Geoffrey Hinton et al., “Deep Neural Networks for Acoustic Modeling in Speech Recognition,” *IEEE Signal Processing Magazine*, 2012.
- Leo Breiman, “Random Forests,” *Machine Learning Journal*, 2001.
- Corinna Cortes and Vladimir Vapnik, “Support-Vector Networks,” *Machine Learning*, 1995.
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, MIT Press, 2016.
- F. Ciompi et al., “The Importance of Stain Normalization in Colorectal Tissue Classification,” *IEEE Transactions on Medical Imaging*, 2017.
- J. Litjens et al., “A Survey on Deep Learning in Medical Image Analysis,” *Medical Image Analysis Journal*, 2017.
- J.V.Anil Kumar, Tanguturi Naga Trisha,” INTELLIGENT VIDEO CONTENT GENERATION USING DEEP LEARNING”, International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.044, Page 357 to 364, ISSN:2250-3676, 2025.
- J.V. Anil Kumar, Nagella Swarupa Rani,” SECURE DATA TRANSMISSION THROUGH HYBRID CRYPTOGRAPHY AND STEGANOGRAPHIC TECHNIQUES”, International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.046, Page 373 to 383, ISSN:2250-3676, 2025.
- J.V.ANIL KUMAR, ALLU MAHALAKSHMI, “SMART NETWORKING APPROACH FOR AUTOMATED INCIDENT

MANAGEMENT", International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.047, Page 384 to 392, ISSN:2250-3676, 2025.