

Colon Cancer Detection using Deep Learning Algorithm

1st E.N. Srivani

Electronics and Communication Engineering
SJC institute of Technology

Department of Electronics and Communication Engineering
REVA University
Bengaluru, India
srivani1598.ece@sjcit.ac.in

2nd G. Seshikala

School of Electronics and communication Engineering
REVA university

Bengaluru, India
seshikala.g@reva.edu.in

Abstract—Currently, deep learning techniques are utilized to diagnose and prognosis the tumors' localization in histopathology images. Globally, Colon cancer (CRC) is the third leading cause of cancer-related death. An intelligent computer-based colon cancer diagnosis is developed by using various Deep Learning (DL) techniques. Early detection of colon tumors is vital for good treatment and diagnosis. Early detection of cancer activities before treatment is significantly important in medical trials or personalizing tumor treatments. Hereditary colorectal cancers can be generally divided into two types such as hereditary non-polyposis colon cancer as well as familial anomalous polyposis. The early detection of colon cancer can increase the patient's survival rate. This research focused on predicting and analyzing colorectal cancer data using deep learning methods. This comprehensive research supports researchers in accomplishing an effective solution for colorectal cancer. At present, deep learning methodologies are commonly employed to enhance the precision of tumor localization in histopathology image analysis and to optimize cancer classification.

Keywords—Colon Cancer, Deep Learning, Histopathology Image Analysis, Localization of tumor, Supervised Machine Learning.

I. INTRODUCTION

Cancer is a broad term used to describe a range of illnesses that have the potential to harm various organs within the human body. As these cells continue to divide and spread throughout the organs, they can become cancerous, potentially leading to death. Cancer is one of the leading causes of death, following cardiovascular diseases. Globally, Colorectal cancer is still the third leading frequently analyzed cancer and the third major cause of cancer cancer-related deaths. The use of custom-made medication is a hot topic in cancer treatment, as it holds the potential to optimize therapy effectiveness based on the individual characteristics of each patient's cancer [1]. The reproducibility of histologic assessment of colon cancer was examined. However, the inter-spectator divergence elimination by using conventional techniques is complex [2]. Although colon cancer can be caused by both genetic and acquired factors, research has shown that colon cancer developed due to genetic factors is significantly higher than that due to acquired factors [3]. Deep learning methods are utilized in medical image processing to resolve pattern recognition [4]. Microarray-based gene expression is majorly utilized in the recognition of colon cancer and diagnosis. Early detection of colon tumors is vital for good treatment as well as diagnosis [5].

However, there are still several obstacles in the early detection of medical-stage tumors using fluid biopsy tests [6]. The implemented deep learning method allowed us to estimate the significance of variables for every toxicity [7]. Data

expansion is a method utilized to enhance the number of models in a dataset by unnaturally making novel samples and merging them with new information [8]. Colorectal tumor is a difficult disease with a lot of dangerous factors, like lifestyles, dietary habits, and genetics. One vital aspect that increases the risk of colorectal tumors among family members is genetic mutations passed down in families [9]. The significant diversity of medical circumstances resulted in poor recitals for patients with metastatic disease [10]. This paper offers an analysis of Colon cancer detection by using deep learning algorithms. Moreover, this paper specifies the limitations of colon cancer.

The overall organization of the manuscript is given as follows: an overview of colon cancer detection using a deep learning algorithm provided in Section 2. The literature survey about existing techniques with their merits and limitations is provided in Section 3. At last, a conclusion is made in Section 4.

II. OVERVIEW OF COLON CANCER DETECTION

Colorectal tumor is a kind of cancer that begins in the large intestine (Colorectal). The colon is an ending part of the digestive zone. Colorectal tumor generally affects older adults, though it can occur at any age. A colon polyp is a small growth that occurs on the inner lining of the colon. While the majority of colon polyps are harmless, some can eventually develop into colon cancer. Doctors often recommend colonoscopy screening to remove polyps and prevent colorectal cancer. A variety of treatments are available for managing colorectal cancer including medication, radiation, surgery, chemotherapy, immunotherapy and targeted therapy. Colorectal cancer is called colon cancer, which differs from rectal cancer, as it starts in the rectum.

A. Symptoms and signs of CRC

- Change in stool steadiness such as narrow stools, loose stools
- Constant abdominal uneasiness, such as pain, cramps, or gas
- A modification in bowel habits.
- Bright red blood in the stool
- Constipation or Diarrhea
- Fatigue or weakness
- Sudden weight loss
- Stool color change

Most of these symptoms may also cause serious conditions. Hence, it is advisable to consult a doctor for early diagnosis.

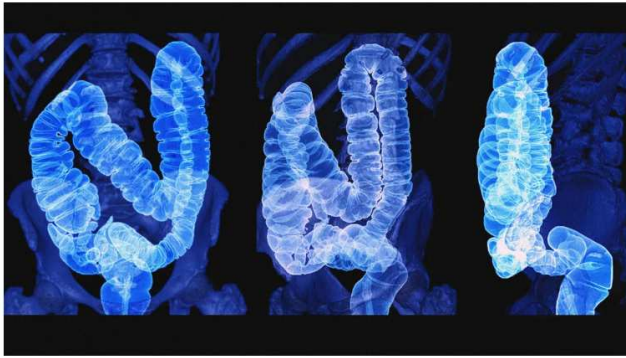


Fig. 1. Symptoms of CRC

B. Types of colon cancer

- Carcinoids
- Hymphomas
- Gastrointestinal stromal tumors
- Sarcomas

C. Treatment options

Treatment for colorectal cancer varies based on various stages and categories. A doctor will also consider a person's overall health, age, and certain other uniqueness while choosing the best healing treatment. There are various treatment options for CRC including radiation therapy, chemotherapy, and surgery. The treatment aims to prevent cancer from spreading, eliminate any uncomfortable symptoms, and eradicate cancer.

- Surgery
- Chemotherapy
- Radiation Therapy

D. Machine Learning

The implementation of computer-based diagnostic devices is becoming more popular. ML is a term used in artificial intelligence to describe devices that can learn and improve themselves by using data from their experiences [5]. There are 2 major types of ML approaches: unsupervised and supervised learning.

1) *Supervised learning*: In this approach, labeled data (matching outputs and inputs) is utilized to generate a model and teaches the computer to efficiently classify unidentified information. Major ordinary supervised learning methods utilized in medicinal research are logistic regression and linear.

2) *Unsupervised learning*: Unsupervised learning is a training method that involves recognizing patterns. Grouping

is a popular unsupervised learning technique, that creates clusters based on general characteristics. The employment of unsupervised ML in medicinal applications. Here, four DL methods provide better performance in colon cancer detection [14].

E. Neural Networks

Neural networks are subsections of the DL, which are designed for problem-solving and tackling the classification problems for recognizing patterns in various data types such as image or voice and identifying tumors. They are collected from many layers, all of which include an unreliable count of neurons [15]. Leftmost or the first one is known as the input layer and the final one is the output layer. Layers along with outputs and inputs are known as unseen layers of the network. Deep learning and Artificial Neural Networks (ANNs) are used to detect colon tumors.

Neural networks, integral components of deep learning, are designed to mimic the human brain's problem-solving capabilities. They excel in tackling intricate classification problems, recognizing patterns in various data types such as images or voices, and identifying tumors.

1) *Feed-forward Neural Networks*: In terms of the direction of data flow, neural networks are surrounded by various categories. Once the result value is acquired from input and the next middle executions, a network is called feed-forward. When a network incorporates the feedback loops, information from the output is fed back to the current outcome based on past results. In recurrent networks, activation of certain neurons causes activation of various group neurons, reasoning cascade stimulations. These kinds of NN have been few familiar than feed-forward networks because of the low level of competence of their learning approaches. In terms of the direction of data flow, neural networks encompass diverse categories.

2) *Training of Neural Network*: The training of a neural network involves adjusting the bias and weight scores until achieving the desired optimal performance. The optimal performance is obtained by resolving optimization issues: reducing cost function. In regression issues, generally, the cost function is taken as the mean squared error, reduced by the gradient descent technique [17]. There are many optimization techniques for NN, like Bayesian regularization, Levenberg–Marquardt, scaled conjugate gradient, and gradient descent with momentum. In patients with metastatic colorectal cancer (CRC), a microarray meta-analysis was employed to identify differentially expressed genes (DEG) in both non-responders and individuals responding to FOLFOX treatment.

III. LITERATURE SURVEY

The detailed evaluation of certain contributions to existing literature is described in Table 1.

TABLE I. LITERATURE SURVEY

Author	Methodology	Advantage	Limitation	Performance measure
Ghosh. et al. [11]	The ensemble deep neural network was used for tumor detection in Colorectal Histology figures through multi-class tissue elements.	Here, Colorectal tumor histopathological figures were investigated significantly to reduce model variance and improve effective functioning.	However, this method was performed less efficiently by using a smaller amount of test and training figures.	Accuracy, precision, and sensitivity.

Lu. W et al. [12]	In metastatic CRC patients, the microarray meta-analysis was used to recognize the differentially expressed genes (DEG) in both non-responders and individuals for FOLFOX treatment.	Here, this method was used to calculate therapeutic replies of CRC patients in the direction of some chemotherapy.	However, this method does not identify the statistically vital negative prophetic factors.	Sensitivity, specificity, PLR, and NLR.
Shayesteh. S et al. [13]	The original scheme was expanded, which employed only worldwide labels to obtain whole-side image localization of carcinoma and classification by merging features as of various magnifications of whole slide images.	Here, the feebly supervised learning approach was utilized to resolve classification problems of colon cancer with few annotation details.	However, the cell level framework method was unsuccessful in studying these features of blood cells while being trained and the blood cells were not bright.	Accuracy, precision, recall, and f1 score.
Masud. M et al. [14]	The classification framework was used to distinguish the middle of five categories of colon and lung tissues by examining their histopathological figures.	Here, the computer-based classification technique in the health check centers allowed pathologists to diagnose further colon and lung tumor cases in less cost, time, and effort.	Moreover, this method does not consider the segmentation of nuclei.	Accuracy, precision, recall, and f-measure.
Lorenzovici. N et al. [15]	The clever computer-assisted colon tumor diagnosis method was designed by utilizing various deep learning methods such as deep neural networks, shallow, and classifications.	Here, the novel technique identified further possible biomarkers to detect colorectal cancer.	However, the binary classification issues were turned into regression issues.	Accuracy, sensitivity, specificity, and precision.
Gupta. P et al. [16]	Varieties of machine learning methods functioned for tumor stage calculation of colon cancer by considering Tumor Aggression Score as a predictive factor.	Here, the various machine learning techniques were used to evaluate test accuracy and train the method with various cross-validations by using autonomous test data sets.	However, the hyperparameters were not accurately validated, and the parameters were not initialized in a professional manner, which led to poor accuracy results.	Accuracy, precision, recall, and f-measure.
Mozdiak. E et al. [17]	This approach investigated the viability of a urinary volatile organic compounds (VOC)-based screening tool for colon cancer.	Here, Data were examined by utilizing the ML technique to evaluate test accuracy for accurate classification of colon cancer over gastrointestinal pathology and adenoma.	In this approach, classification between adenomas and management was insufficient.	Sensitivity, specificity, PPV, and NPV.
Hamida. A.B et al. [18]	The employ of ML structural designs to highlight colon cancer regions and classify in lightly annotated histopathological information background.	Here, the method was used to identify the most appropriate network and effectual training strategy for colon cancer segmentation.	Moreover, the amplification of a full explanation for huge data is a difficult task.	Accuracy, specificity, sensitivity, and dice.
Nadimi. E.S et al. [19]	The convolutional neural network was utilized to characterize the size, encode the image morphology and location of polyps, and RNN with concentration devices to make a description of histological features.	The hybrid and a self-sufficient end-to-end machine learning method were used to detect colon cancer of any size and morphology videos.	However, the Capsule Networks are based on the integration of attention mechanisms and ensemble learning that are harder to train.	Accuracy, specificity, and sensitivity.
Sarker. M.M.K et al. [20]	The machine learning techniques had addressed detection and cell segmentation challenges in immunohistochemistry slides. This serves as the basis for quantifying nuclear discoloration biomarkers, including an Inducible T-cell COSTimulator.	Here, the deep learning method was used to develop predictive and prognostic methods by using robust expansion processes.	However, immune histochemistry is not suitable for examining huge numbers of benchmark communicating cells.	Accuracy, precision, sensitivity, specificity, dice, AJI, and recall.

IV. PROBLEM STATEMENT

Colorectal cancer detection faces various problems during implementation. Some open research issues that occurred while implementing the CRC are specified as follows:

- The colorectal Histology figures through multi-class tissue elements were performed poorly when using smaller amounts of test and training figures.
- Colon cancer is a vastly treatable and frequently curable disease. Surgery is the main form of healing. The recurrence after surgery represents a significant concern and often causes the ultimate mortality.
- Colon cancer detection is the use of radiation. For that reason; it is not suitable for pregnant women.

V. CONCLUSION

In recent days, colorectal cancer has been the third leading frequently analyzed cancer the worldwide and third most common cause of cancer-related deaths. Various deep-learning algorithms are utilized for computer-aided colon cancer detection. In this paper, colon cancer detection achieved better performance by using deep learning methods. The supervised learning technique was utilized to resolve classification issues of colon cancer figures with few annotation details. Various machine learning techniques are used to evaluate the methods and train the method with various cross-validations by using autonomous test data sets. These methods might be competent and similar to machine learning methods. This research is very useful in determining the CRC detection's current trend and future things. Still, there is a great deal to be done on the early-stage CRC patients. This research paper will help those who are looking to understand the existing methods in the process of DL for colorectal cancer detection. This research paper serves as a valuable resource

for those seeking to comprehend the current methodologies employed in the application of deep learning for colorectal cancer detection.

REFERENCES

- [1] Abdallah, N., Hage Chehade, A., Oueidat, M. Chauvet, P., and Marion, J.M., 2022. Lung and colon cancer classification using medical imaging: A feature engineering approach. *Physical and Engineering Sciences in Medicine*, 45(3), pp.729-746.
- [2] Ali, M.S., Hasan, M.I., Islam, M.K. and Rahman, M.H., 2022. Automated detection and characterization of colon Cancer with deep convolutional neural networks. *Journal of Healthcare Engineering*, 2022.
- [3] J.H. Bae, M. Kim, J.S. Lim, and Z.W. Geem, 2021. Feature selection for colon cancer detection using k-means clustering and modified harmony search algorithm. *Mathematics*, 9(5), p.570.
- [4] Arasi, M.A., Obayya, M., Alsini, R., Alruwais, N., Yaseen, I. and Mohamed, A., 2023. Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model. *IEEE Access*.
- [5] Molla, M.M., Shafi, A.S.M., Rahman, M.M. and Jui, J.J. 2020. Detection of colon cancer based on microarray dataset using machine learning as a feature selection and classification techniques. *SN Applied Sciences*, 2(7), pp.1-8.
- [6] Kalaivani, S., Gupta, S., Ameta, G.K., Rajasundaram, A., Dugbakie, B.N. and Oleiwi, A.K., 2022. Prediction performance of deep learning for colon cancer survival prediction on SEER data. *BioMed Research International*, 2022.
- [7] Insausti, A., Oyaga-Iriarte, E., Aldaz, A. and Sayar, O., 2019. Prediction of irinotecan toxicity in metastatic colorectal cancer patients based on machine learning models with pharmacokinetic parameters. *Journal of pharmacological sciences*, 140(1), pp.20-25.
- [8] Kareem, S.A., Mulenga, M., Seera, M., Sabri, A.Q.M., Samudi, C., Mohamad, S.B. and Govind, S., 2021. Feature extension of gut microbiome data for deep neural network-based colorectal cancer classification. *IEEE Access*, 9, pp.23565-23578.
- [9] Dong, C., Liu, J., Lu, X., Jiang, G., Wu, H. and Liu, Y., 2020. Transcription factor expression as a predictor of colon cancer prognosis: a machine learning practice. *BMC medical genomics*, 13(9), pp.1-10.
- [10] Aslan, M.F., Sabanci, K. and Attallah, O., 2022. A framework for lung and colon cancer diagnosis via lightweight deep learning models and transformation methods. *Diagnostics*, 12(12), p.2926.
- [11] Bandyopadhyay, A., Ghosh, S., Ghosh, R., Sahay, S., Santosh, K.C. and Kundu, I., 2021. Colorectal histology tumor detection using ensemble deep neural network. *Engineering Applications of Artificial Intelligence*, 100, p.104202.
- [12] Fu, D., Lu, W., Huang, Z., Kong, X., Zhu, Y., Hwang, M., Jiang, K., Chen, L., Wu, Y., Li, J. and Li, X., 2020. FOLFOX treatment response prediction in metastatic or recurrent colorectal cancer patients via machine learning algorithms. *Cancer Medicine*, 9(4), pp.1419-1429.
- [13] Nazari, M., Shayesteh, S., Sandoughdaran, S., Salahshour, A., Khateri, M., Hajianfar, G., Jozian, F., Yaghobi Joybari, A., Arabi, H., Shiri, I. and Fatehi Feyzabad, S.H., 2021. Treatment response prediction using MRI-based pre-, post-, and delta-radiomic features and machine learning algorithms in colorectal cancer. *Medical physics*, 48(7), pp.3691-3701.
- [14] Sikder, N., Masud, M., Bairagi, A.K., AlZain, M.A. and Nahid, A.A., 2021. A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. *Sensors*, 21(3), p.748.
- [15] Dulf, E.H., Lorenzovici, N., Mocan, L. and Mocan, T., 2021. Artificial Intelligence in Colorectal Cancer Diagnosis Using Clinical Data: Non-Invasive Approach. *Diagnostics*, 11(3), p.514.
- [16] Chiang, S.F., Gupta, P., Mohapatra, S.K., Sahoo, P.K., Onthoni, D.D., You, J.F., Chiang, J.M., Hung, H.Y., Tsai, W.S. and Huang, Y., 2019. Prediction of colon cancer stages and survival period with machine learning approach. *Cancers*, 11(12), p.2007.
- [17] Wicaksono, A.N., Mozdiak, E., Arasaradnam, R.P. and Covington, J.A., 2019. Colorectal cancer and adenoma screening using urinary volatile organic compound (VOC) detection: early results from a single-centre bowel screening population (UK BCSP). *Techniques in coloproctology*, 23(4), pp.343-351.
- [18] Devanne, M., Hamida, A.B., Truntzer, C., Weber, J., Ghiringhelli, F., Derangère, V., Wemmert, C. and Forestier, G., 2021. Deep learning for colon cancer histopathological images analysis. *Computers in Biology and Medicine*, 136, p.104730.
- [19] Buijs, M.M., Nadimi, E.S., Kroijer, R., Herp, J., Nielsen, E., Kobaek-Larsen, M., Blanes-Vidal, V., Baatrup, G. and Pedersen, C.D., 2020. Application of deep learning for autonomous detection and localization of colorectal polyps in wireless colon capsule endoscopy. *Computers & Electrical Engineering*, 81, p.106531.
- [20] Makhoulf, Y., Sarker, M.M.K., Humphries, M.P., Craig, S.G., James, J.A., Loughrey, M., O'Reilly, P., Maxwell, P. and Salto-Tellez, M., 2021. A Means of Assessing Deep Learning-Based Detection of ICOS Protein Expression in Colon Cancer. *Cancers*, 13(15), p.3825.