Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230573

Deep Learning in Colorectal Cancer Classification: A Scoping Review

Rafaa ALALWANI[a], Augusto LUCAS[a], Mahmoud AlZUBAIDI[a], Hurmat Ali SHAH[a], Tanvir ALAM[a], Zubair SHAH[a] and Mowafa HOUSEH[a,1]
[a] *College of Science and Engineering, Hamad Bin Khalifa University, Doha 34110, Qatar*

Abstract. Colorectal cancer (CRC) is one of the most common cancers worldwide, and its diagnosis and classification remain challenging for pathologists and imaging specialists. The use of artificial intelligence (AI) technology, specifically deep learning, has emerged as a potential solution to improve the accuracy and speed of classification while maintaining the quality of care. In this scoping review, we aimed to explore the utilization of deep learning for the classification of different types of colorectal cancer. We searched five databases and selected 45 studies that met our inclusion criteria. Our results show that deep learning models have been used to classify colorectal cancer using various types of data, with histopathology and endoscopy images being the most common. The majority of studies used CNN as their classification model. Our findings provide an overview of the current state of research on deep learning in the classification of colorectal cancer.

Keywords. Deep learning, Colorectal Cancer, classification

1. Introduction

Colorectal cancer (CRC), also known as colon cancer, is a condition where the cells in the colon or rectum grow in an uncontrollable manner [1]. Colorectal cancer was the third most common cancer worldwide in 2020 with over 1.9 million new cases [2]. Classification of CRC types is essential for appropriate risk assessment and follow-up recommendations [3], it is also necessary for proper diagnosis and understand the prognosis [4]. CRC can be classified by examining colorectal polyps or analyzing histological tissues [5,6]. As the CRC cases is growing high, it poses a great challenge for pathologists and imaging specialists to provide better and proper healthcare facilities. To address this issue, researchers are exploring the use of artificial intelligence (AI) to classify colorectal biopsy tissue and endoscopy images, improving diagnosis accuracy and speed while maintaining quality of care [7,8]. In this scoping review paper, we explored the utilization of deep learning technology for the classification of colorectal cancer mentioned in literature.

[1] Corresponding Author: Mowafa Househ, E-mail: mhouseh@hbku.edu.qa.

2. Methodology

This scoping review was conducted following the guidelines from Joanna Briggs Institute (JBI) scoping review method and Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA-ScR) [9,10]. We searched five databases, namely IEEE Xplore, Google Scholar, PubMed, ACM, and ScienceDirect, for articles that focused on the target population (colorectal cancer), target intervention (deep learning), and target outcome (classification). The search terms used are detailed in Appendix 1. The study selection process involved two phases: screening and eligibility. The study selection and data extraction were performed independently by two authors (RA and AL). The Cohen Kappa was calculated to measure the inter-rater reliability between the two reviewers which was 0.832. Refer to Appendix 2 for the data extraction sheet.

3. Results

3.1. Selected Studies

Our search yielded a total of 1004 studies, from which 19 duplicates were identified and removed. After reviewing the remaining 958 unique full-text articles, we excluded 869 articles due to irrelevant population, intervention, study design, outcome, or type of publication. We further assessed the full text of 116 articles, leading to the exclusion of an additional 71 articles due to irrelevant intervention, outcome, population, or study design. Finally, 45 studies met the inclusion criteria and were included in this scoping review. Refer to Appendix 3-A for the PRISMA chart.

3.2. Nature of Studies Included

In this scoping review, most of the included papers were conference papers (n=34, 75.56%). The included studies were from 21 countries, with China having the highest number of papers (n=8, 17.78%). We recorded the country of the publication based on the country of the first author. The majority of the papers were published between 2019 and 2021(n=28, 62.22%). Refer to Appendix 4 for the general characteristics.

3.3. Dataset Characteristic

Public datasets were the most used (n=22, 48.88%), followed by private datasets (n=19, 42.22 %), and some papers used a combination of both (n=4, 8.89 %). Most of the studies used images (n=44, 97.78%). Histopathology images were used in 24 of the studies, followed by endoscopy images in 17 of the included studies. One study used colonography or virtual colonoscopy images, and another used hyperspectral imaging. One study included DNA methylation array data in the form of text files, and another used a combination of endoscopy images and histology reports in the form of text. All the included studies were able to classify different classes of CRC from the available datasets. Some studies used two classes (n=19, 42.22 %) for polyps or tissue types, while others (n=26, 57.77 %) used more detailed classifications that ranges from 3 and up to 8 classes. Endoscopy image classification studies have primarily focused on distinguishing between adenomatous or hyperplastic polyps, and identifying whether they are cancerous or normal. Histopathology image classification studies have taken different approaches, with some classifying nuclei and tissue types such as Tumor, Stroma, Complex, Lympho, Debris, Mucosa, Adipose, and Empty. Other studies have classified samples as malignant or benign, or determined whether the tissue is in a metastatic state. The sizes of the datasets used varied, with the smallest dataset consisting of 63 samples and the largest dataset consisting of 20,000 samples. Refer to Appendix 5 for more dataset properties.

3.4. Deep Learning Models Characteristic

Convolutional neural network (CNN) was the most widely used classification model (n=41, 91.11%). General Adversarial Networks were used in two studies, while Graph Neural Networks and Transformers were each used once. In terms of training optimizers, Adaptive Moment with Estimation was the most frequently used (n=23, 51.11%). Stochastic Gradient Descent was used in 11 studies. Regarding validation methods, the train-test split was the most used method (n=17, 37.78 %). Cross-validation was used in 12 studies. Refer to Appendix 6 for more details on the deep learning characteristics used.

Table 1. Analysis of data types based on models used and classification metric.

Type of File	Number of reported	Models	Average Accuracy	Average Specificity	Average Sensitivity
Endoscopy	17	CNN, GAN	88.22%	86.11%	87.14%
Biopsy-histpathology	24	CNN, GAN, GNN, Transformers	91%	84%	92%
Hyperspectral imaging	1	CNN	N/R	78%	88%
DNA-Mythalon	1	CNN	96.17%	95.83%	96.65%
Colonscopy	1	CNN	87%	N/R	N/R
Endoscopy+Histology+report	1	CNN	81.10%	N/R	N/R

3.5. Evaluation Metrics.

The studies included in this review employed various metrics to evaluate the performance of their machine learning models, such as accuracy, specificity, sensitivity, AUC, F1-score, and precision. Accuracy was recorded in 34 studies, while precision was reported in 10 studies. The study utilizing DNA-Methylation arrays achieved the highest accuracy with 96.16 %, while the study that combined endoscopy and histology recorded the lowest accuracy at 81.10 % as shown in table 1.0. We also looked at the evolution of the classification metrics values over time. From 2019 to 2022, accuracy, sensitivity, AUC and F1- score increased steadily while specificity and precision experienced fluctuations. Refer to Appendix 7 for a comprehensive analysis of the evaluation metrics.

4. Discussion

This scoping review has both practical and research implications. From a practical standpoint, the review highlights the potential of deep learning to aid pathologists in the faster and more accurate diagnosis and classification of CRC. This review also summarizes that deep learning could facilitate the connection of images with DNA biomarkers, leading to more personalized treatment. On the research side, the review emphasizes the need for further studies to compare traditional techniques with deep learning and for improved data collection schemes to enhance the accuracy and reliability of deep learning in medicine. Additionally, the review identifies the need for more public datasets to advance the development of deep learning in medical applications.

5. Conclusion

Our scoping review shows that deep learning models have been successfully used for the classification of CRC using various types of data, including histopathology and endoscopy images. The results of the included studies demonstrate the potential of AI technology to improve the accuracy and speed of CRC classification, which can lead to better patient outcomes. However, more research is needed to validate the use of deep learning models in clinical practice and to develop standardized approaches for data collection, model development, and evaluation. The findings of this scoping review can serve as a basis for future research in this area and provide useful insights for clinicians and researchers working in the field of colorectal cancer classification.

All Appendices including the full references sheet are available online on GitHub link: https://github.com/ICT660/Colorectal_Cancer_Scoping_Review

References

[1] CDC. What Is Colorectal Cancer? [Internet]. Centers for Disease Control and Prevention. 2019. Available from: https://www.cdc.gov/cancer/colorectal/basic_info/what-is-colorectal-cancer.htm

[2] World Cancer Research Fund International. Worldwide cancer data | World Cancer Research Fund International [Internet]. WCRF International. 2020. Available from: https://www.wcrf.org/cancer-trends/worldwide-cancer-data/

[3] Gupta S, Lieberman D, Anderson JC, Burke CA, Dominitz JA, Kaltenbach T, et al. Spotlight: US Multi-Society Task Force on Colorectal Cancer Recommendations for Follow-up After Colonoscopy and Polypectomy. Gastroenterology. 2020 Mar;158(4):1154.

[4] Sawayama H, Miyamoto Y, Ogawa K, Yoshida N, Baba H. Investigation of colorectal cancer in accordance with consensus molecular subtype classification. Annals of Gastroenterological Surgery. 2020 Jul 21;4(5):528–39.

[5] Tanaka S, Hayashi N, Oka S, Chayama K. Endoscopic Assessment of Colorectal Cancer with Superficial or Deep Submucosal Invasion Using Magnifying Colonoscopy. Clinical Endoscopy. 2013;46(2):138.

[6] Fleming M, Ravula S, Tatishchev SF, Wang HL. Colorectal carcinoma: Pathologic aspects. Journal of gastrointestinal oncology [Internet]. 2012 [cited 2019 Apr 2];3(3):153–73. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22943008

[7] Kudo S, Mori Y, Misawa M, Takeda K, Kudo T, Itoh H, et al. Artificial intelligence and colonoscopy: Current status and future perspectives. Digestive Endoscopy [Internet]. 2019 Feb 27;31(4):363–71. Available from: https://onlinelibrary.wiley.com/doi/10.1111/den.13340

[8] Leifler KS. AI can help doctors work faster – but trust is crucial [Internet]. liu.se. 2022 [cited 2023 Mar 27]. Available from: https://liu.se/en/news-item/ai-kan-hjalpa-lakare-arbeta-snabbare---digital-patologi

[9] Peters M, Godfrey C, McInerney P, Munn Z, Trico A, Khalil H. Chapter 11: Scoping Reviews. JBI Manual for Evidence Synthesis. 2020;

[10] Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine. 2018 Sep 4;169(7):467–73.