

Deep Learning-Based Crohn's Disease Prediction: A Comprehensive Examination and Future Perspectives

Zhong Li *

Department of Artificial Intelligence, Beijing University of Technology, Beijing, China

* Corresponding Author Email: zhongli@emails.bjut.edu.cn

Abstract. This study explores deep learning's (DL) role in enhancing Crohn's Disease (CD) diagnosis and treatment, aiming to overcome current diagnostic challenges and improve patient outcomes through more accurate, efficient, and personalized medical interventions. This comprehensive review scrutinized a plethora of studies focusing on the utilization of machine learning (ML) and DL methodologies for diagnosing CD. The investigation spanned various Artificial Intelligence (AI) techniques. This endeavor aimed to illustrate the transformation from traditional ML methods, which necessitate labor-intensive data preprocessing and expert analysis, to DL approaches that autonomously decipher intricate patterns from voluminous datasets. Special attention was accorded to research that leveraged these technologies for distinguishing CD from ulcerative colitis (UC), anticipating disease complications, and pinpointing diagnostic markers. This review elucidates the progression from traditional machine learning techniques, requiring substantial data preparation and expert knowledge, to deep learning algorithms capable of learning directly from raw data. This shift promises to automate and refine the diagnostic process significantly. An examination of various studies showcased how AI applications in CD diagnosis are evolving, underscoring the potential of these technologies to transform the diagnostic landscape by enhancing accuracy, reducing time, and paving the way for personalized treatment strategies. The integration of AI in CD diagnosis has shown significant promise, with DL and ML models achieving higher diagnostic accuracy and efficiency compared to traditional methods.

Keywords: Inflammatory bowel disease, machine learning, Crohn's disease prediction.

1. Introduction

Inflammatory Bowel Disease (IBD) in humans, including Crohn's disease and ulcerative colitis, affects over 10 million people worldwide, according to statistics. In China, the prevalence rate of ulcerative colitis is about 11.6 per 100,000, and Crohn's disease is about 2.29 per 100,000, with an expected rise to 1.5 million by 2025 [1]. Inflammatory bowel diseases have evolved from being rare to common conditions. Crohn's disease may impact any section of the gastrointestinal tract, from the mouth to the anus, with the most frequent sites being the terminal part of the small intestine (ileum) and the start of the colon. It can occur in any part of the digestive tract and may lead to various complications such as intestinal obstruction, abscesses, and even perforation or significant bleeding. Moreover, distinguishing between Crohn's disease symptoms and those of ulcerative colitis or other types of enteritis can be challenging, requiring experienced doctors for accurate diagnosis. Previously considered a rare disease, experienced physicians in this field were scarce. Furthermore, the identification process is also time-consuming. Deep learning can effectively process large-scale data sets through automated feature extraction and accuracy improvement in disease image analysis, significantly speeding up the diagnosis speed and improving the consistency of results. This technology supports multi-task learning and cross-modal analysis, providing doctors with comprehensive diagnostic information and promoting the development of personalized medicine. Therefore, employing deep learning to recognize Crohn's disease through training on large datasets is essential, using sensitive and cost-effective systems to differentiate between UC and CD and also determine the therapeutic effect and further develop the treatment strategy. This not only can catch the disease early and prevent it from getting worse, but also increase the precision and effectiveness of diagnosing and treating conditions.

Artificial Intelligence (AI) has made significant strides in various domains over recent years, including natural language processing, computer vision, autonomous driving, and healthcare. In medicine, a plethora of artificial intelligence algorithms have been applied, such as deep neural networks and reinforcement learning. For instance, Google's DeepMind team collaborated with the Cancer Research UK London Centre to study lung cancer prediction using AI, employing deep learning algorithms, particularly CNNs, to analyze lung CT scans [2]. This analysis aims to predict the survival rates and treatment responses of lung cancer patients. While many teams are exploring this field, most focus on common diseases. Research on Crohn's disease, however, has not received sufficient attention, making exploration in this area particularly valuable. Diego et al. proposed a specially designed convolutional neural network aimed at identifying images affected by lesions through the classification of capsule endoscopy images. Their research shows that this network significantly improves accuracy and processing speed in image classification problems compared to other state-of-the-art deep learning reference architectures [3]. Maurício et al. have provided a scientific contribution to identifying types of inflammatory bowel disease. They collected images obtained through colonoscopy and video capsule endoscopy and applied five CNNs and six vision transformers (ViTs) to determine which deep learning architecture performed best. Furthermore, they optimized the computational requirements of these pretrained deep learning models, achieving a lightweight architecture with 25 times fewer parameters through knowledge distillation, while maintaining the same accuracy as the original architecture. They evaluated the models before and after knowledge distillation both qualitatively and quantitatively to understand the information learned by the two architectures and whether both could achieve higher accuracy [4]. Ruan et al. developed and validated a deep learning diagnostic system trained on a large dataset of colonoscopy images. This system can differentiate between UC and CD, outperforming experienced endoscopists in recognition accuracy [5]. Zheng et al. established a workflow for diagnosis and mining differential diagnostic features using electronic medical records. This process involves fine-tuning pretrained language models into a lightweight and efficient TextCNN model, interpreting neural networks, selecting discriminative attribution features, then applying manual feature inspection and debias training [6]. Pei et al. employed four machine learning models to assess the diagnostic and predictive value of PBRP for UC and CD [7]. Given the importance of this field and the breakthroughs made in recent years, a comprehensive review of these advancements is necessary.

The remainder of this paper is organized as follows. First, this paper will discuss the methods in Section 2, which involve how to use machine learning and deep learning algorithms for recognition. Then, in Section 3, the advantages and disadvantages of these methods as well as limitations and future prospects will be provided. Finally, Section 4 will summarize this paper and present the conclusions drawn from the articles discussed here.

2. Method

2.1. Traditional Machine Learning-based Methods

2.1.1 Predicting Future Complications in Pediatric Crohn's Disease

Chen et al. investigate the effectiveness of machine learning ML in forecasting outcomes for CD by analyzing gene expression data from 101 biopsies taken from patients who have not undergone treatment and from control subjects [8]. They assessed the likelihood of developing complications such as strictures and fistulas, as well as the need for surgical intervention, using ML algorithms that integrate gene expression in the colon with clinical data. Their models successfully predicted these outcomes with AUROC scores of 0.84 for the development of strictures, 0.83 for achieving remission, and 0.75 for the necessity of surgery. Notably, it identified prognostically significant genes overlooked in single-gene analyses, showcasing the potential of colonic gene expression in guiding precise treatment strategies to avert severe CD progression.

2.1.2 Distinguishing Between UC and CD

Pei et al. investigates using ML to differentiate UC from CD through blood tests [7]. Using machine learning techniques on data collected from 414 patients with IBD, 423 healthy subjects, and 344 individuals experiencing intestinal problems unrelated to IBD, it was found that the Multi-Layer Perceptron Artificial Neural Network (MLP-ANN) model was particularly effective in accurately diagnosing conditions. It excelled in distinguishing IBD from non-IBD conditions, UC from CD, and identifying disease activity stages. With a training/testing split ratio of 7:3 and external validation, the study highlights MLP-ANN's potential as a non-invasive diagnostic tool, advocating for further exploration with larger cohorts and diverse models.

2.1.3 Identifying Subtypes and Markers Related to Platelets for Diagnosis

In the study on pediatric Crohn's disease (PCD), Tang et al. combined bioinformatics and machine learning approaches to pinpoint subtypes and markers linked to platelets. Through the analysis of data from the GEO database, they discovered differentially expressed and platelet-related genes via WGCNA and immune infiltration studies. By applying machine learning techniques such as SVM-RFE, LASSO, GBM, XGBoost, and Random Forest, they identified five crucial markers (GNA15, PIK3R3, PLEK, SERPINE1, STAT1) and used them to develop models for predicting platelet-related conditions. The research identified molecular subtypes related to platelets in PCD by using consensus clustering, providing valuable information for tailoring treatment approaches [9].

2.1.4 Identifying Hub Genes

Zheng et al. sought to investigate the impact of genes associated with angiogenesis on CD and to develop models for diagnosing CD and predicting non-responsiveness to infliximab treatment [6]. They used microarray data from the GEO database to perform unsupervised consensus clustering, grouping CD samples into clusters based on the expression levels of angiogenesis-related genes. Through Weighted Gene Co-expression Network Analysis (WGCNA), they identified a module linked to angiogenesis, and machine learning techniques were applied to pinpoint key genes and create a diagnostic model for the disease. Further, genes associated with infliximab non-response were extracted to construct a predictive model.

2.2. Deep Learning-based Methods

2.2.1 CNN and ViTs in Diagnosing the Type of IBD

Maurício et al. focuses on differentiating Crohn's disease and ulcerative colitis using deep learning models, contributing to the accurate and early diagnosis of inflammatory bowel diseases [4]. This approach leverages images from publicly available datasets (LIMUC HyperKvasir and CrohnIPI) to train five CNNs and six ViTs, with the goal of creating a tool to assist physicians in identifying the specific type of Inflammatory Bowel Disease (IBD). Recognizing the challenge of deploying complex models in a hospital setting, the study employs knowledge distillation to generate simpler architectures that maintain the original models' precision. This process includes evaluating and interpreting both the pre-distilled and distilled architectures to ensure performance consistency and learning efficacy. Ultimately, the study successfully reduces the model's complexity by 25 times, maintains good performance, and decreases inference time, while also providing interpretability insights into diagnosing ulcers, bleeding, and lesions indicative of IBD.

2.2.2 CNN Designed to Detect Crohn's Disease in Wireless Capsule Endoscopic Images

Marin-Santos et al. unveils a groundbreaking CNN tailored for detecting Crohn's disease in capsule endoscopic images, showcasing a bespoke architecture that excels in both accuracy and efficiency [3]. Its design, focused on feature extraction and classification, is uniquely suited to the nuances of endoscopic imagery. Trained on a robust dataset of 15,972 images, the CNN outperforms traditional models with a significantly higher ROC curve area, delivering unparalleled accuracy, sensitivity, and specificity. Moreover, its streamlined complexity allows for real-time image analysis, setting a new standard for automatic Crohn's disease detection in clinical settings.

2.2.3 Utilizing a 'Thick Data'-Enhanced Siamese Neural Network for Detecting Crohn's Disease in High-Definition Endoscopic Videos

Sawyer et al. emphasize the development of a Siamese Neural Network (SNN) tailored to recognize Crohn's Disease in high-definition endoscopic videos, utilizing a novel "Thick Data" technique. This approach supplements conventional machine learning methods with detailed heuristic annotations, including bounding boxes and segmentation masks, to improve the model's accuracy in detecting specific disease characteristics. The SNN, with its custom architecture, excels in leveraging heuristic data to pinpoint Crohn's Disease markers accurately. Data augmentation techniques spotlight essential visual cues in the images, bolstering the diagnosis process. A user-friendly prototype enables physicians to easily upload videos for analysis, promising a transformative tool for Crohn's Disease diagnosis and potential applications in other conditions through advanced heuristic learning [10].

2.2.4 ResNeXt-101 for Improved Diagnosis of IBD Using Colonoscopy Images

Wang et al. introduces a cutting-edge CNN model, leveraging the ResNeXt-101 architecture, to distinguish CD, UC and normal conditions using 15,330 colonoscopy images [11]. The ResNeXt-101 model, noted for its outstanding performance, adeptly classifies these images, demonstrating a superior accuracy of 92.04% on a per-image basis and 90.91% on a per-patient basis, outshining the diagnostic accuracy of six clinicians with varied experience levels. This advancement highlights the potential of CNNs, particularly the ResNeXt-101, as a valuable tool for improving IBD diagnosis accuracy in clinical practice, offering essential support especially to less experienced clinicians.

2.2.5 Deep CNN Diagnostic System for UC and CD Differentiation

Ruan et al. created and tested a deep learning-based diagnostic tool designed to differentiate ulcerative colitis (UC) from Crohn's Disease (CD), utilizing a dataset comprising 49,154 images from colonoscopies performed on 1,772 individuals. [5]. Utilizing a CNN, the system was trained and validated across multiple centers, demonstrating high accuracy in identifying UC and CD compared to experienced endoscopists. The deep model outperformed trainee and competent endoscopists in both per-patient (99.1% vs. 78.0% and 92.2%) and per-lesion (90.4% vs. 59.7% and 69.9%) analyses. Additionally, the model significantly reduced the mean reading time for diagnosis (6.20 s vs. 2425.00 s for endoscopists), showcasing its potential to enhance diagnostic efficiency in clinical settings. This system presents a valuable resource for both medical training and clinical application, targeting enhanced precision in diagnosing and managing IBD.

2.2.6 CS-MIL Framework for Crohn's Disease Diagnosis

Deng et al. presents an innovative Cross-Scale Attention Guided Multi-Instance Learning (CS-MIL) framework for accurately diagnosing CD from pathological images [12]. Utilizing a unique cross-scale attention mechanism, the model adeptly aggregates features from various resolutions, closely mirroring the meticulous approach of pathologists who analyze biopsies at multiple scales. This enables a deeper, scale-aware analysis, setting it apart from conventional models with its nuanced feature integration and differential multi-scale attention visualizations. These visualizations illuminate the diagnostic rationale by highlighting lesion patterns across scales. Achieving an impressive AUC score of 0.8924, the CS-MIL framework marks a breakthrough in digital pathology, significantly enhancing CD diagnosis accuracy and providing valuable insights into the model's decision-making process.

2.2.7 A Novel Loss Function CDW-CE for Estimating the Severity of UC from Colonoscopy Images

Polat et al. present an innovative Class Distance Weighted Cross-Entropy (CDW-CE) loss function designed for precise assessment of ulcerative colitis (UC) severity using colonoscopy images [13]. Addressing the ordinal nature of UC severity, CDW-CE incorporates class distances into its calculations, offering a more precise reflection of disease progression. It outperforms traditional loss functions, enhancing model accuracy in severity estimation and providing interpretable visual

insights through Class Activation Maps (CAMs). Additionally, the study advances CD diagnosis and treatment strategy formulation by analyzing angiogenesis-related genes, employing machine learning to identify key diagnostic markers and predict treatment responses. This dual approach significantly enriches clinical understanding and management of inflammatory bowel diseases, marking a notable advancement in medical AI applications.

3. Discussion

In the evolving field of medical diagnostics, particularly for complex diseases like CD, the advent and integration of AI technologies, specifically ML and DL, have marked a significant shift towards more accurate, efficient, and personalized care [14-16].

The distinction between ML and DL in the context of CD diagnosis encapsulates a transition from traditional, feature-engineered models to data-driven, automatically learned feature hierarchies. Traditional ML methods, while pioneering, often necessitate extensive preprocessing and expert knowledge to identify relevant features within the data. In contrast, DL architectures, particularly CNNs and ViTs, excel in automatically extracting complex patterns and features from vast datasets, including endoscopic and histological images. This ability to learn from raw data without explicit feature engineering heralds a new era of diagnostic accuracy and efficiency, as evidenced by the highlighted studies. Despite these advancements, the deployment of DL in clinical settings encounters several challenges. Deep neural networks' opaque characteristics lead to difficulties in interpretability, presenting challenges for medical professionals in comprehending how AI models arrive at their decisions. This lack of transparency is particularly problematic in healthcare, where diagnostic and treatment decisions require clear rationale and confidence. Additionally, the privacy and security of patient data, essential in the medical field, raise concerns about the ethical use and potential misuse of sensitive information. Another critical challenge is the generalizability and applicability of AI models across diverse populations, given the variability in disease presentation and progression, influenced by genetic, environmental, and lifestyle factors.

Addressing these challenges requires a multifaceted approach. Enhancing model interpretability through explainable AI techniques can bridge the gap between AI predictions and clinical decision-making. Federated learning presents a promising avenue for privacy-preserving AI, allowing models to be trained across multiple institutions without directly sharing patient data. Moreover, transfer learning and domain adaptation techniques can ameliorate issues of model generalizability by adapting models trained on one dataset or population to perform accurately on another. Looking forward, the integration of AI in CD diagnosis and management is poised to benefit significantly from interdisciplinary collaboration, involving clinicians, data scientists, and ethicists, to ensure that technological advancements align with ethical standards and clinical needs. Continuous efforts in developing and validating AI models against diverse and representative datasets will be crucial in realizing the full potential of AI in healthcare. Furthermore, regulatory frameworks and guidelines need to evolve alongside technological advancements to ensure the safe, ethical, and effective implementation of AI tools in clinical practice.

4. Conclusion

This review provides the comprehensive overview of the application of DL in diagnosing CD, showcasing its potential to revolutionize diagnostic accuracy, speed, and personalization. This paper discussed the progresses from traditional machine learning to sophisticated DL models that automate feature extraction and analysis, offering a leap towards automated, precise, and patient-centric diagnostics. These advancements highlight DL's ability to decode complex patterns from vast datasets, propelling medicine towards actionable insights with less expert input. Despite progress, challenges like model interpretability, data privacy, and cross-population applicability persist, emphasizing the need for innovation balanced with ethical and clinical considerations. Looking forward, the further

study aims to refine DL models for wider use, focusing on federated learning, interpretability, and privacy, aiming to bridge technology and medicine for improved healthcare solutions.

References

- [1] Bioon. Access the document 2023, <https://www.bioon.com/article/5f58e215829a.html>.
- [2] UCL News. Artificial intelligence helps simplify lung cancer risk prediction. 2023, <https://www.ucl.ac.uk/news/2023/oct/artificial-intelligence-helps-simplify-lung-cancer-risk-prediction>
- [3] Marin-Santos D, et al. Automatic detection of Crohn disease in wireless capsule endoscopic images using a deep convolutional neural network, 53: 12632-12646. *Applied Intelligence* 2023.
- [4] Maurício J, Domingues I. Distinguishing between Crohn's disease and ulcerative colitis using deep learning models with interpretability, 27: 1. *Pattern Analysis and Applications* 2024.
- [5] Ruan G, et al. Development and Validation of a Deep Neural Network for Accurate Identification of Endoscopic Images from Patients With Ulcerative Colitis and Crohn's Disease, 9: 854677. *Frontiers in Medicine* 2022.
- [6] Zheng C, et al. Constructing models for Crohn's disease diagnosis and prediction of infliximab non-response based on angiogenesis-related genes, 15: Article 1239496. *Frontiers in Immunology* 2024.
- [7] Pei J, Wang G, Li Y, Li L, Li C, Wu Y, Liu J, Tian G. Utility of four machine learning approaches for identifying ulcerative colitis and Crohn's disease, 10: e23439. *Heliyon* 2024.
- [8] Chen KA, et al. Linking gene expression to clinical outcomes in pediatric Crohn's disease using machine learning, 14: 2667. *Scientific Reports* 2024.
- [9] Tang D et al. Identification of platelet-related subtypes and diagnostic markers in pediatric Crohn's disease based on WGCNA and machine learning, 15: Article 1323418. *Frontiers in Immunology* 2024.
- [10] Sawyer D. Detecting Crohn's Disease from High Resolution Endoscopy Videos: The Thick Data Approach. Master's thesis, Lakehead University, Department of Computer Science 2023.
- [11] Wang L et al. Development of a Convolutional Neural Network-Based Colonoscopy Image Assessment Model for Differentiating Crohn's Disease and Ulcerative Colitis, 9: 789862. *Frontiers in Medicine* 2022.
- [12] Deng R et al. Cross-scale Attention Guided Multi-Instance Learning for Crohn's Disease Diagnosis with Pathological Images, 13594: 24–33. *Multiscale Multimodal Medical Imaging* 2022.
- [13] Polat G, Ergenc I, Kani HT, Alahdab YO, Atug O, Temizel A. Class Distance Weighted Cross-Entropy Loss for Ulcerative Colitis Severity Estimation. *arXiv:2202.05167v2 [eess.IV]* 2022.
- [14] Qiu Y, Wang J, Jin Z, et al. Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. *Biomedical Signal Processing and Control*, 2022, 72: 103323.
- [15] Lambert B et al. Trustworthy clinical AI solutions: a unified review of uncertainty quantification in deep learning models for medical image analysis. *Artificial Intelligence in Medicine*, 2024: 102830.
- [16] Waljee A K et al. Development and validation of machine learning models in prediction of remission in patients with moderate to severe Crohn disease. *JAMA network open*, 2019, 2(5): e193721-e193721.