INDIVIDUAL PROJECT INTERIM REPORT

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Tackling Crohn's disease using deep learning

Author:
Feifan Fan

Supervisor:
Dr. Bernhard Kainz

January 25, 2023

Submitted in partial fulfillment of the requirements for the MEng Mathematics and Computer Science of Imperial College London

Contents

Chapter 1

Introduction

1.1 Crohn's Disease

Crohn's Disease [1, 2] is known as one of the two main types of Inflammatory Bowel Disease (IBD) [3]. Patients with IBD will suffer from severe and chronic inflammation of the gastrointestinal (GI) tract.

The main concern raised from such symptoms is that chronic and long-term inflammation can lead to damage in the GI tract, along with further possible complications, including bowel cancer. Recent research from the University of Nottingham [4] states that more than half a million people in the UK live with Crohn's Disease and Ulcerative colitis, another primary type of IBD.

However, unlike Ulcerative Colitis, which only affects the colon and rectum, Crohn's disease can develop lesions in any part of the GI tract, from the oesophagus to the anus. This fact poses two main problems. Since inflammation can occur in any part of the GI tract, the patient's digestive process can be affected. Also, due to the diversity of affected locations, the symptom of this disease will vary from patient to patient, including abdominal pain, diarrhoea, fatigue, and weight loss.

1.2 Clinical Challenges

Although the medical community has invested much research into Crohn's disease [5, 6], our clinical experts have not identified the exact cause of Crohn's disease yet. Therefore Crohn's disease is not curable today.

Fortunately enough, early diagnosis coupled with appropriate treatment from clinical professionals can significantly reduce the patient's suffering and improve the quality of their lives. Clinicians typically make a specific diagnosis through various approaches such as enteroclysis, endoscopy, colonoscopy, and radiographic diagnosis, including barium contrast X-rays, Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). The MRI diagnosis has become popular since it is a non-invasive approach and produces more detailed images than its alternatives, such as computerised tomography (CT).

The examiners will look particularly into the terminal ileum, and right colon, where the areas are most involved in Crohn's disease, for symptoms such as bowel wall thickening, small bowel strictures, and mural hyperenhancement [7] to confirm the existence of Crohn's disease. Nevertheless, human experts still need to go through the MRI scans slice by slice, which is time-consuming and inefficient to diagnose one patient.

1.3 Motivation

The prevalence of Machine Learning and Deep Learning, especially convolutional neural networks (CNN), allows us to automatically learn the features from input imaging data and help the human expert with diagnosis. Given that the terminal ileum is essential in Crohn's Disease diagnosis, in 2019, Holland et al. [8] proposed a residual network focusing on the terminal ileum to detect Crohn's Disease from MRI scans automatically. They concluded that the framework's performance depends on the level of localisation in preprocessing. Consequently, they recommend that terminal ileal ground-truth segmentations also be collected for localising the terminal ileum and enhancing the performance of automated detection.

A follow-up study by Abidi et al. [9] in 2022 developed an enhanced deep learning tool using a nnU-Net architecture to automatically locate critical areas (particularly terminal ileum) where radiologists typically perform examinations to make a diagnosis. This study resolves the high dependence on the level of localisation in preprocessing in [8], and the model significantly outperforms the results of the prior work done by Xu et al. [10] in 2021 for terminal ileum segmentation. Furthermore, [9] establishes a strong foundation for a multi-class terminal ileum segmentation algorithm.

Inspired by studies of [8, 10, 9], this individual project aims to develop a multi-class segmentation model to differentiate normal and abnormal terminal ileum regions, and the clinicians will be significantly aided in the diagnosis of Crohn's disease with the success of this project.

1.4 Machine learning Challeges

The performance of a deep learning model is strongly related to the training data's quality, quantity and diversity. Unfortunately, our biggest challenge in this project is the scarcity of training data. Our training dataset is relatively small (only 233 patient cases available) compared with other industry-leading deep learning systems. Another fact we need to face is that developing gold-standard labels or point-wise centerlines for patient data requires manual segmentations by clinical experts, which is also a complex, time-consuming and inefficient process. Therefore, the acquisition of patient data with great quality and quantity is indeed a challenging process.

To alleviate these problems, we introduce a proxy task with weak supervision before the target training task to get comparable results with learning a large number of data with weak labels that are less spatially informative. However, from prior work [9], it is suggested that training from scratch using the nnU-Net framework in the proxy task is inefficient since it takes a long time to converge and exhibits unstable performance. Applying transfer learning to proxy learning [11] or adding a related pre-training task may resolve this situation.

1.5 Objectives

The primary objectives of our work include the following:

1. Replicating the results of the prior work [9] for segmenting the terminal ileum from MRI scans.

2. Developing a multi-class segmentation model of the terminal ileum based on the prior work and employing the trained model as our baseline model.

3. Developing a sophisticated, related pre-training task for the multi-class terminal ileum segmentation that leverages the industrial leading transfer learning techniques to imporve the performance of the proxy model and associated stability.

4. Improving the results of the baseline segmentation model and developing a user-friendly GUI as an extension to this project.

Chapter 2

Related Work

2.1 Automatic Detection and Segmentation of Crohn's Disease Tissues from Abdominal MRI

We need to go back to a decade ago to see the initial accomplishments of applying deep learning techniques to medical image analysis. In 2013, Mahapara et al. [12] first proposed a machine learning-based method for segmenting parts of the bowel for detecting Crohn's disease tissues on MRI scans.

The pipeline firstly oversegmented the given MR image test volume into supervoxels and applied Random Forest (RF) classifiers to classify supervoxels containing diseased tissues, where these supervoxels define the Volume of Interest (VOI). Another set of RF classifiers is used on the test volume to generate a probability map, which indicates the probability of each voxel being diseased tissue, normal tissue or background.

The experimental results showed that the pipeline achieved a decent segmentation performance, with a dice metric value of 0.90 ± 0.04 and a Hausdorff distance of 7.3 ± 0.8 mm. However, this approach is computationally expensive and inefficient since each instance takes a long time for testing, each pipeline stage requires fine-tuning, and the architecture leaves lots of room for improvement.

2.2 Automatic Detection of Bowel Disease with Residual Networks

Holland et al. [8] proposed an end-to-end residual network [13] with a soft attention layer [14] to automate the detection of Crohn's disease. The detection is not based on any of the semantic segmentation work but only seeking for the terminal ileum. Consequently, they revealed the feasibility of using a deep learning algorithm for identifying terminal ileum Crohn's Disease from abdominal MRI.

From experimental results, the best result with localised data and a semi-automatic setting achieved a weighted-f1 score 0.83, which is strongly correlated with the MaRIA [15] score, the current clinical standard. Also, the performance is slightly inefficient when a fully automatic invariant is included.

They suggested that the performance of the model is highly dependent on the level of localisation in the training data and argued that gathering gold-standard segmentations of terminal ilium will be beneficial as a predecessor task if a better performance of the automatic detection needs to achieve.

2.3 Machine learning methods to predict the presence of intestine damage in patients with Crohn's disease

Enchakalody et al. [16] proposed two classifiers for aiding improvements of the reliability of Crohn's Disease in diagnosis and monitoring by looking into cross-sectional views of a small-intestine segment and detecting diseased issues. The classifiers are implemented based on Random Forest (RF) with ensembling and Convolutional Neural Network (CNN).

From the experimental result, both techniques can differentiate diseased and normal small bowel tissue with similar performance, even compared to the case with expert radiologists present. However, this study mainly focused on CT-enterography data, which slightly deviated from our cases with MRI data. And until the day of writing this report, there is still no fully automatic approach for diagnosing Crohn's disease on cross-sectional imaging that reaches the level of expert radiologists.

Chapter 3

Background

3.1 Image Segmentation

3.2 Threshold-based Method

3.3 Region-based Method

3.4 Edge Detection-based Method

3.5 Deep Learning Methods

Bibliography

[1] Daniel C Baumgart and William J Sandborn. Crohn's disease. *The Lancet*, 380(9853):1590–1605, 2012.

[2] NHS. Crohn's disease - nhs. URL https://www.nhs.uk/conditions/crohns-disease/.

[3] Centers for Disease Control and Prevention. What is inflammatory bowel disease? (ibd) | ibd. URL https://www.cdc.gov/ibd/what-is-IBD.htm.

[4] University of Nottingham. Rates of crohn's and colitis have been vastly underestimated for decades, says new study. URL https://www.nottingham.ac.uk/news/rates-of-crohns-and-colitis-have-been-vastly-underestimated-for-decades-says-new-study.

[5] Gautier Hoarau, PK Mukherjee, C Gower-Rousseau, C Hager, J Chandra, MA Retuerto, Christel Neut, Séverine Vermeire, J Clemente, Jean-Frederic Colombel, et al. Bacteriome and mycobiome interactions underscore microbial dysbiosis in familial crohn's disease. *MBio*, 7(5):e01250–16, 2016.

[6] Joseph D Feuerstein, Edith Y Ho, Eugenia Shmidt, Harminder Singh, Yngve Falck-Ytter, Shanaz Sultan, Jonathan P Terdiman, Shahnaz Sultan, Benjamin L Cohen, Karen Chachu, et al. Aga clinical practice guidelines on the medical management of moderate to severe luminal and perianal fistulizing crohn's disease. *Gastroenterology*, 160(7): 2496–2508, 2021.

[7] David H Bruining, Ellen M Zimmermann, Edward V Loftus Jr, William J Sandborn, Cary G Sauer, Scott A Strong, Mahmoud Al-Hawary, Sudha Anupindi, Mark E Baker, David Bruining, et al. Consensus recommendations for evaluation, interpretation, and utilization of computed tomography and magnetic resonance enterography

in patients with small bowel crohn's disease. *Gastroenterology*, 154(4): 1172–1194, 2018.

[8] Robert Holland, Uday Patel, Phillip Lung, Elisa Chotzoglou, and Bernhard Kainz. Automatic detection of bowel disease with residual networks. In *International Workshop on PRedictive Intelligence In MEdicine*, pages 151–159. Springer, 2019.

[9] Ali Abidi. Tackling crohn's disease using deep learning. Master's thesis, Imperial College London, 2022.

[10] Ke Xu. Tackling crohn's disease using deep learning. Master's thesis, Imperial College London, 2021.

[11] Jong-Hwan Jang, Tae Young Kim, and Dukyong Yoon. Effectiveness of transfer learning for deep learning-based electrocardiogram analysis. *Healthcare informatics research*, 27(1):19–28, 2021.

[12] Dwarikanath Mahapatra, Peter J Schüffler, Jeroen AW Tielbeek, Jesica C Makanyanga, Jaap Stoker, Stuart A Taylor, Franciscus M Vos, and Joachim M Buhmann. Automatic detection and segmentation of crohn's disease tissues from abdominal mri. *IEEE transactions on medical imaging*, 32(12):2332–2347, 2013.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[14] Jo Schlemper, Ozan Oktay, Michiel Schaap, Mattias Heinrich, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention gated networks: Learning to leverage salient regions in medical images. *Medical image analysis*, 53:197–207, 2019.

[15] Jordi Rimola, Sonia Rodríguez, Orlando García-Bosch, Ingrid Ordás, Edgar Ayala, Montserrat Aceituno, Maria Pellisé, Carmen Ayuso, Elena Ricart, Lluís Donoso, et al. Magnetic resonance for assessment of disease activity and severity in ileocolonic crohn's disease. *Gut*, 58 (8):1113–1120, 2009.

[16] Binu E Enchakalody, Brianna Henderson, Stewart C Wang, Grace L Su, Ashish P Wasnik, Mahmoud M Al-Hawary, and Ryan W Stidham. Machine learning methods to predict presence of intestine damage in patients with crohn's disease. In *Medical Imaging 2020: Computer-Aided Diagnosis*, volume 11314, pages 742–753. SPIE, 2020.