UPPSALA
UNIVERSITET

M STER THESIS SPECIFIC TION

Machine-learning Assisted Endoscopic Detection of Pre-malignant Lesions in Patients with In ammatory Bowel Disease

Student
Darian André Krummrich
darian-
andre.krummrich.4611@student.uu.se
Signature: _____

Supervisors
Anne Håkansson
annehak@kth.se
Signature: _____

Peter Thelin Schmidt
peter.thelin.schmidt@medsci.uu.se
Signature: _____

Subject Reviewer
Name: _____
Email: _____
Signature: _____

Department of Information Technology

November 23, 2023

Contents

1 Abstract

Colorectal cancer (CRC) has a high mortality rate which can be decreased by early detection and intervention. Patients with inflammatory bowel disease (IBD) are at heightened risk of developing CRC and thus undergo regular surveillance, but recognition of premalignant (precancerous) changes is difficult and inaccurate due to inflammation. These issues can further be exacerbated based on colonoscopy operator di erences, such that many changes are missed.

Recognition of pre-cancerous lesions can reasonably be improved by machine learning techniques by highlighting suspicious lesions for endoscopists during screenings. Novel imaging techniques such as chromoendoscopy further provide superior detection rates, and may be used in combination with AI.

It shall now be investigated whether ML techniques can also be utilized in the support of colonoscopies in IBD patients, increasing accuracy and improving treatment outcomes.

2 Background

Colorectal cancer (CRC) is one of the leading causes of death worldwide [1], where early detection plays a major factor in treatment outcomes [2, 3, 4]. The risk for patients with inflammatory bowel disease (IBD) to develop CRC is increased so that it accounts for approximately 15% of all deaths in IBD patients [5]. As such, patients with IBD undergo regular colonoscopy surveillance. However, the quality of colonoscopies is lower for detecting pre-cancerous tissues in IBD patients as the changes can arise in mucosa that is visually indistinct from surrounding mucosa [6]. Further, due to its molecular biology, altercations may occur more rapidly or in an unconventional sequence compared to patients without IBD [6].

In addition, quality of colonoscopy varies by operator [7]. Especially novice endoscopists overestimate their accuracy of clinical competence [8] and may benefit from targeted interventions.

One approach to increase detection rates is the utilization of AI-based diagnosis. Initial studies have shown that colonoscopy with AI increases the detection rate of adenomas and polyps compared to colonoscopy without AI [9, 10, 11], and that this improvement is especially strong in non-experts in the field of colorectal polyp classification in comparison to experts [12].

In another approach, chromoendoscopy is used to highlight subtle mucosal changes during colonoscopy by dyeing tissues with indigo blue colour. This technique improves the sensitivity of detecting neoplasia in IBD patients [6], and o ers the potential to improve specificity as well. By enhancing endoscopic characterization of lesions, fewer and more targeted biopsies may be performed.

It shall now be investigated whether these improvements achieved through the application of chromoendoscopy and AI support can also be applied to patients with IBD, improving the detection rate of pre-malignant changes at colonoscopy. Video data will be

analyzed by means of deep learning classification to identify pre-cancerous or cancerous tissues.

The research question is: "Can the use of AI support at endoscopy increase detection rates of pre-malignant lesions in patients with IBD?"

3 Description of Tasks

Video datasets of individuals undergoing IBD surveillance colonoscopy are being acquired in a cooperation between Ersta sjukhus, Akademiska sjukhus, and Göteborg sjukhus. The primary objective of the thesis is to propose a deep learning model classification algorithm to recognize pre-cancerous changes, which can be used during colonoscopies to support endoscopists, thereby improving detection accuracy.

It is expected to fulfill following objectives through the project to be implemented:

T1. Conduct a critical review of existing machine learning algorithms used in colonoscopy and in IBD surveillance

T2. Perform an in-depth study of technologies used to develop and compare performances of di ering deep learning models

T3. Design a pre-processing pipeline filtering frequently misclassified images such as those containing endoscopic instruments and blood from the dataset provided to the models

T4. Construct one or several deep learning architectures and tune hyperparameters

T5. Evaluate overall performance of the proposed system

T6. Complete the report

4 Methods

The pre-processing of data as well as the machine learning model training and evaluation will be done in python. Literature research will be conducted using PubMed, IEEE, Scopus, and Google scholar. The structure of the literature review shall follow the PRISMA guidelines [13]:

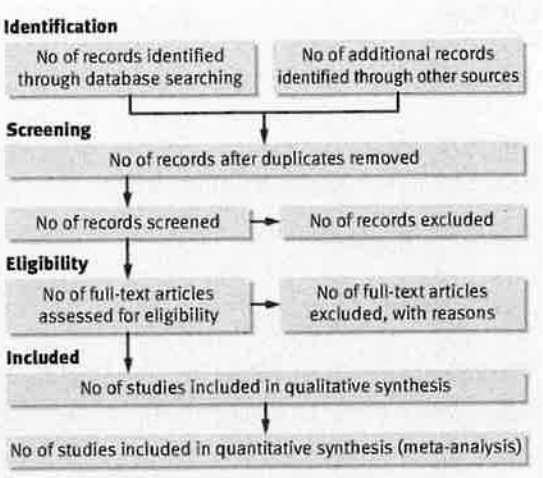


Figure 1: The PRISMA process for literature review [13]

5 Relevant Courses

Courses at Uppsala University (Master's level)

1DT054 - Project CS

1DL073 - Natural Computation Methods for Machine Learning

1TD396 - Computer-Assisted Image Analysis I

Courses at TU Braunschweig (Bachelor's level)

Bachelor's thesis: Identification and Application of Artificial Intelligence Methods for Automated Annotation of Seismocardiographic Datasets

Generation of Biosignals and Medical Images

Introduction to Machine Learning

Medicine 2 (Pathology)

Medicine 1 (Anatomy)

Can be found listed in transcript of records, not in the bachelor's certificate, as it was taken in addition to the credit requirements of the bachelor's studies.

6 Delimitations

This project may include...

M1. Privacy-preserving ML methods such as federated or split learning depending on available data storage and processing environment

M2. Real-time image stream analysis

This project shall not include...

X1. A Grad-CAM (Gradient-weighted Class Activation Mapping) to visualize regions strongly impacting the decision-making process of the ML

X2. A graphical user interface intended for non-technical users

7 Time Plan

This thesis work is expected to be done in 20 weeks, beginning in the fourth week of January 2024. The completion is aimed to be in the first week of June 2024, upon a public presentation regarding the result as well as revision of the report based on the comments by the audience and reviewer. The time allocation spent will be as follows:

(T1, T2) Literature review (2 weeks)

(T3) Data preparation and pipeline construction (2 weeks)

Design of an algorithm to extract unwanted images

Design of an image sampling and loading pipeline

(T4, T5) Training and evaluation of models (8 weeks)

Data validation and hyperparameter tuning

(T4, T5) Perform interpretability analysis to see if the models can be improved (4 weeks)

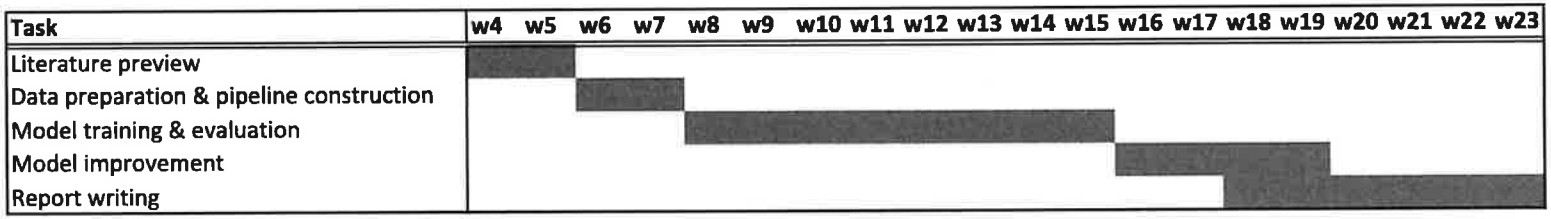
Explore the possibility of using multi-output systems to return feature-based classification as well as image-based classification

(T6) Writing the report (6 weeks)

The graphical representation of the timeline can be found below:

Master's thesis 2024

Start: Week 4, 2024



References

[1] "Global, regional, and national burden of colorectal cancer and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019," *The Lancet. Gastroenterology Hepatology*, vol. 7, no. 7, pp. 627–647, Apr. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9192760/

[2] A. Wiegering, S. Ackermann, J. Riegel, U. A. Dietz, O. Götze, C.-T. Germer, and I. Klein, "Improved survival of patients with colon cancer detected by screening colonoscopy," *International Journal of Colorectal Disease*, vol. 31, no. 5, pp. 1039–1045, May 2016. [Online]. Available: https://doi.org/10.1007/s00384-015-2501-6

[3] R. G. S. Meester, A. G. Zauber, C. A. Doubeni, C. D. Jensen, V. P. Quinn, M. Helfand, J. A. Dominitz, T. R. Levin, D. A. Corley, and I. Lansdorp-Vogelaar, "Consequences of Increasing Time to Colonoscopy Examination After Positive Result From Fecal Colorectal Cancer Screening Test," *Clinical Gastroenterology and Hepatology*, vol. 14, no. 10, pp. 1445–1451.e8, Oct. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S154235651630204X

[4] I. Alonso-Abreu, O. Alarcón-Fernández, A. Z. Gimeno-García, R. Romero-García, M. Carrillo-Palau, D. Nicolás-Pérez, A. Jiménez, and E. Quintero, "Early Colonoscopy Improves the Outcome of Patients With Symptomatic Colorectal Cancer," *Diseases of the Colon Rectum*, vol. 60, no. 8, pp. 837–844, Aug. 2017.

[5] P. Munkholm, "Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease," *Alimentary Pharmacology Therapeutics*, vol. 18, no. s2, pp. 1–5, 2003. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2036.18.s2.2.x

[6] T. L. Zisman and D. T. Rubin, "Colorectal cancer and dysplasia in inflammatory bowel disease," *World Journal of Gastroenterology : WJG*, vol. 14, no. 17, pp. 2662–2669, May 2008. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2709054/

[7] M. Jiang, M. J. Sewitch, A. N. Barkun, L. Joseph, and R. J. Hilsden, "Endoscopist specialty is associated with colonoscopy quality," *BMC Gastroenterology*, vol. 13, no. 1, p. 78, May 2013. [Online]. Available: https://doi.org/10.1186/1471-230X-13-78

[8] M. A. Scaffidi, S. C. Grover, H. Carnahan, R. Khan, J. M. Amadio, J. J. Yu, C. Dargavel, N. Khanna, S. C. Ling, E. Yong, G. C. Nguyen, and C. M. Walsh, "Impact of experience on self-assessment accuracy of clinical colonoscopy competence," *Gastrointestinal Endoscopy*, vol. 87, no. 3, pp. 827–836.e2, Mar. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016510717324392

[9] I. Barua, D. G. Vinsard, H. C. Jodal, M. Løberg, M. Kalager, Holme, M. Misawa, M. Bretthauer, and Y. Mori, "Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis," *Endoscopy*, vol. 53, no. 3, pp. 277–284, Mar. 2021, publisher: Georg Thieme Verlag KG. [Online]. Available: http://www.thieme-connect.de/DOI/DOI?10.1055/a-1201-7165

[10] H. Xu, R. S. Y. Tang, T. Y. T. Lam, G. Zhao, J. Y. W. Lau, Y. Liu, Q. Wu, L. Rong, W. Xu, X. Li, S. H. Wong, S. Cai, J. Wang, G. Liu, T. Ma, X. Liang, J. W. Y. Mak, H. Xu, P. Yuan, T. Cao, F. Li, Z. Ye, Z. Shutian, and J. J. Y. Sung, "Artificial Intelligence–Assisted Colonoscopy for Colorectal Cancer Screening: A Multicenter Randomized Controlled Trial," *Clinical Gastroenterology and Hepatology*, vol. 21, no. 2, pp. 337–346.e3, Feb. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1542356522006735

[11] C. Hassan, M. Spadaccini, A. Iannone, R. Maselli, M. Jovani, V. T. Chandrasekar, G. Antonelli, H. Yu, M. Areia, M. Dinis-Ribeiro, P. Bhandari, P. Sharma, D. K. Rex, T. Rösch, M. Wallace, and A. Repici, "Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis," *Gastrointestinal Endoscopy*, vol. 93, no. 1, pp. 77–85.e6, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016510720345235

[12] Y. Xu, W. Ding, Y. Wang, Y. Tan, C. Xi, N. Ye, D. Wu, and X. Xu, "Comparison of diagnostic performance between convolutional neural networks and human endoscopists for diagnosis of colorectal polyp: A systematic review and meta-analysis," *PLOS ONE*, vol. 16, no. 2, p. e0246892, Feb. 2021, publisher: Public Library of Science. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246892

[13] A. Liberati, D. G. Altman, J. Tetzla , C. Mulrow, P. C. Gøtzsche, J. P. A. Ioannidis, M. Clarke, P. J. Devereaux, J. Kleijnen, and D. Moher, "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration," *BMJ*, vol. 339, p. b2700, Jul. 2009, publisher: British Medical Journal Publishing Group Section: Research Methods & Reporting. [Online]. Available: https://www.bmj.com/content/339/bmj.b2700