STUDY PROTOCOL

Classifying colorectal cancer or colorectal polyps in endoscopic setting using convolutional neural network: protocol for a systematic review and meta-analysis [version 1; peer review: 1 approved with reservations]

Kamyab Keshtkar ¹, Abbas Keshtkar ², Alireza Safarpour ³

¹Faculty of Engineering, University of Tehran, Tehran, Tehran, Iran

²Department of Health Sciences Education Development, School of Public Health, Tehran University of Medical Sciences, Tehran, Tehran, Iran

³Gastroenterohepatology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

V1 First published: 02 Sep 2020, 9:1086
<https://doi.org/10.12688/f1000research.25548.1>
Latest published: 02 Sep 2020, 9:1086
<https://doi.org/10.12688/f1000research.25548.1>

Abstract

Background: Colorectal cancer (CRC) is the third most common cancer worldwide. Although colonoscopy screening has been proven as an effective strategy for preventing CRC unfortunately, even conventional colonoscopy by expert gastroenterologists can miss adenomas or pre-cancerous lesions in up to 25% of cases. This systematic review aimed to classify colorectal polyps (CRP) or CRC in endoscopic clinic settings using a new machine learning method, convolutional neural network (CNN).

Methods: We will search PubMed/MEDLINE, Scopus, Web of Science, IEEE, Inspec, ProQuest, Google Scholar, Microsoft Academic Search, ScienceOpen, arXiv, and bioRxiv from 1st January 2010 to the 31th of July 2020. Our search will not be restricted based on language or geographical area. The primary studies will be selected that have observational design (cross-sectional, case control or cohort); the study subjects will be adult patients (≥ 18 years old) referred to colonoscopy clinics; and the results of their colonoscopy evaluation will be available in the form of images or videos.

The extracted data will be combined using meta-analysis of prediction models. The primary data synthesis will be performed based on area under curve-receiver operating characteristic curve and/or accuracy measures. We will use Stata version 14.2 (Statacorp; College Station, TX) for primary and secondary data synthesis.

Conclusion: The inferences of our secondary research will provide evidence to evaluate the prognostic role of CNN in discriminating CRP or CRC in colonoscopy settings.

Open Peer Review

Approval Status ?

1

version 1

02 Sep 2020

?

[view](#)

1. **Iman Tahamtan** , University of Tennessee, Knoxville, USA

Any reports and responses or comments on the article can be found at the end of the article.

Keywords

Colorectal Polyp, colorectal cancer, convolutional neural network, prediction model, computer aided diagnosis

This article is included in the **Oncology** gateway.

Corresponding author: Kamyab Keshtkar (kkeshtkar@gmail.com)

Author roles: **Keshtkar K:** Conceptualization, Investigation, Methodology; **Keshtkar A:** Conceptualization, Methodology, Supervision, Writing – Review & Editing; **Safarpour A:** Conceptualization, Methodology

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2020 Keshtkar K *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Keshtkar K, Keshtkar A and Safarpour A. **Classifying colorectal cancer or colorectal polyps in endoscopic setting using convolutional neural network: protocol for a systematic review and meta-analysis [version 1; peer review: 1 approved with reservations]** F1000Research 2020, 9:1086 <https://doi.org/10.12688/f1000research.25548.1>

First published: 02 Sep 2020, 9:1086 <https://doi.org/10.12688/f1000research.25548.1>

Introduction

Colorectal cancer (CRC) is the third most common cancer worldwide. Among the causes of cancer-related death, CRC ranks second¹. Although the biological pathways that transform the normal colonic into malignant tissue are different, polyps are presumed to be the precursor lesions for malignant tumors in all cases². Colorectal polyps (CRP) can be seen in various forms or shapes on colon endoscopy. CRP are histologically classified into two main categories based on growth pattern, namely serrated and adenomatous (adenomas) polyps².

A new systematic review and meta-analysis combined 70 primary population-based cross-sectional studies that used colonoscopy for assessing different colorectal neoplastic lesions. This study reported the worldwide prevalence of adenoma, advanced adenoma and CRC as 25.9%, 5.2% and 0.6% in patients older than 50 years, respectively³. Brenner et al. investigated almost 850,000 colonoscopies of German national screening of CRC. They projected that 10-year risk of progression of advanced adenomas to CRC varied between 25% and 40% in men and women older than 50 years of age⁴.

The role of colonoscopy screening has been proven as an effective strategy for preventing CRC development. This intervention can affect CRC mortality by two pathways; first, polypectomy or removing adenomas and second, detecting early stages of CRC⁵. Because removal of CRP does not reduce the risk of CRC to zero, current guidelines recommend a 5–10-year rescreening interval after a negative colonoscopy result^{6,7}. Recently, a large-scale multi-center cohort study followed 200,000 persons who underwent baseline colonoscopies for more than eight years. This large cohort study showed that high-risk and low-risk adenomas resulted in 2.6 and 1.3 fold increase, respectively, in the risk of CRC development when compared to persons without adenoma⁸.

During the last two decades, despite efforts made to screen CRC/CRP in target groups (e.g., population aged more than 50 years old), studies have shown that even conventional colonoscopy by expert gastroenterologists can miss adenomas or pre-cancerous lesions in up to 25% of cases^{9,10}. Over the last few decades, several measures have been developed and recommended to evaluate the quality of colonoscopy in the diagnosis of polyps and colon cancers¹¹. An important quality measure is the Adenoma Detection Rate (ADR). A recent large-scale investigation on more than 300,000 colonoscopies performed by 136 colonoscopists during 13 years showed that an 1.0% increase in the ADR was associated with a 3.0% decrease in the risk of CRC¹².

In recent years, several technologies have been combined with colonoscopy, leading to new endoscopic methods for improving detection rate of colorectal lesions. Image-Enhanced Endoscopy (IEE), cap-assisted colonoscopy, the Third Eye® Retroscope®, wide-angle colonoscope, Endocuff® device, and water-assisted colonoscopy are just a few of these new methods of colonoscopy that are available to augment the detection, diagnosis, and treatment of these subtle lesions^{13,14}.

In the last few years, several adjunct techniques or devices are under investigation for improving ADR in colonoscopy settings. These are methods categorized under the generic term “Computer Aided Diagnosis” (CAD)¹⁵. The collection of models such as “Artificial Intelligence”, “Deep Neural Networks” and “Machine Learning” are under CAD. Recent investigations have shown that CAD-methods, along with colonoscopy data, have advantages; higher CRC/CRP detection, better histopathologic differentiation, decreasing overall healthcare costs, and decreasing operator independency¹⁵.

Convolutional neural networks (CNNs) are a class of deep neural networks that are highly effective at performing image and video analyses. CAD–CNN models for colonoscopy could assist endoscopists in detecting polyps and performing optical diagnosis. CNNs are trained using thousands of colonoscopy images to identify and differentiate between hyperplastic and adenomatous polyps¹⁶.

Since 2018, several systematic reviews have been published to diagnose different clinical outcomes using CNN methods. The assessed outcomes were in skin cancer¹⁷, breast cancer¹⁸, hepatocellular carcinoma or hepatic mass¹⁹, and ischemic brain strokes²⁰.

Based on our knowledge, no systematic review or meta-analysis to evaluate the diagnostic accuracy for discrimination of CRC from CRP has been published. Unfortunately, nearly all systematic reviews assessing the diagnostic accuracy of CNNs in detecting or differentiating different health outcomes, especially different cancers that have been published so far have some limitations. First, the papers had no study protocols and no evidence was provided regarding registration of the protocols. Second, all the published systematic reviews were performed without doing a meta-analysis. Third, in the systematic review methods, risk of bias assessment or critical appraisal was not designed. Therefore, we will aim to design and conduct a systematic review and meta-analysis using a standard and high-quality method to evaluate diagnostic accuracy for discrimination of CRC/CRP by CNN.

Protocol

This protocol was first designed based on the “piori” approach and then registered in Open Science Framework²¹.

Primary objective

Assessing the accuracy of CNN model (Intervention: I) for discrimination malignancy (cancer) and/or polyp (Outcome: O) from normal colorectal tissue (Comparison: C) in CRC or CRP probable patients (Participants: P) who attended colonoscopic clinics or centers, based on colonoscopic images or videos.

Eligibility criteria of primary studies

Type of primary studies. Primary studies will be included as follows: first, all of the study subjects have colonoscopy videos or images (at least one per subject); second, all of the study subjects have pathological or cytological diagnoses (as the gold standard or reference standard); third, the study design is

cross-sectional, prospective, or retrospective approach. Therefore, all observational studies with case-control, cohort, or cross-sectional design will be included. Interventional studies (trials, experimental or quasi-experimental), reviews (secondary research), editorials, letters, and similar articles will be excluded.

Type of participants. Since CRPs are usually diagnosed in adults older than 50 years of age, all studies conducted on adult populations of either gender (i.e., patients older than 18 years) will be eligible to be included.

Reference standard. The histopathologic examination results of the colorectal lesions (CRC or CRP) will be used as the reference standard.

Index test (the output of prediction model). CNNs are a supervised learning method. They can learn and find the relationship between the input (images or videos) and the class labels. CNN layers are generally divided into two categories: hidden layers and fully connected layers. The task of the hidden layers is to extract the features. The fully connected layers are used for the classification and detection object in the input images, at the end of the CNN network. The different class labels of all the assessed images or videos should be reported.

Search strategy

We will search the following bibliographic databases: PubMed/MEDLINE, SCOPUS, Web of Science, IEEE (Institute of Electrical and Electronics Engineers), Inspec, ProQuest, Google Scholar, Microsoft Academic Search, ScienceOpen, arXiv, bioRxiv. Moreover, we will assess relevant conferences for content (Conference on Computer Vision and Pattern Recognition, International Conference of Computer Vision, European Conference on Computer Vision). We will search *Gastroenterology*, *Pattern Recognition*, *Scientific Reports* as key journals by hand-searching. The search time interval will cover the 1st of Jan 2010 up to the 31th of July 2020. Our search will not be restricted based on language or geographical area. The PubMed search syntax is provided as *Extended data*²¹.

Screening and selection processes

After searching in the sources mentioned above, we will screen all the primary studies based on titles or abstracts. A screening checklist will be developed using 4 to 6 criteria. The criteria will be selected based on the most common components reported in the abstracts of primary research. We will select included or potentially included studies for further assessment. Two reviewers will independently evaluate the considered studies based on full-text papers or documents. We will resolve any disagreement between the two reviewers by consensus.

Quality assessment (risk of bias assessment)

Two reviewers will independently assess the methodological quality (risk of bias) of the included studies. The risk of bias checklist will be selected based on **PROBAST tool**²¹. The checklist has two main domains: risk of bias (ROB) and applicability domains. The ROB domains contain four items: participants,

predictors, outcome, analysis. The applicability domain has three items: participants, predictors, and outcome. All seven items will be assessed according to one of three options: low risk of bias, high risk of bias and unclear risk of bias. The overall quality status (overall risk of bias status) will be determined based on defined guidelines (PROBAST guideline). We will resolve any disagreement between the two reviewers by consensus.

Data extraction and data synthesis

We will design an extraction form based on the study objectives. After testing of the extraction form (at least one primary study), we will finalize the data extraction form. Two reviewers will independently extract the required data from all the primary studies (papers or documents). We will resolve any disagreement in the extracted data between the two reviewers by consensus.

The quantitative primary measures are one of two performance indicators: area under curve-receiver operating characteristic curve (AUC-ROC) (C or Concordance Statistics) or accuracy measure. As primary data synthesis (meta-analysis), we will combine the AUC and/or accuracy measures.

The secondary data are the CNN architecture model (VGG²³, AlexNet²⁴, GoogLeNet²⁵, etc.) or the features of CNN architecture, such as the number of layers, the size of layers (kernel (filter) and stride size), kind of pooling layers (max or average) and so on, the CNN model sensitivity, specificity, and diagnostic odds ratio.

The subgroup variable data will be transfer learning (existing or absence), learning rate, the features of CNN architecture (number of layers, size of the kernel and stride, and so on).

Combining the primary and secondary data will be performed based on Debray et al guidelines²⁶. We will use Stata version 14.2 (StataCorp. College Station, TX) and R 4.0.0 for conducting the meta-analysis method.

We will use a forest plot for presenting the accuracy measure pooling and I² (inconsistency) measure and Q Cochrane test for assessing heterogeneity²⁷. We will use the subgroup analysis based on the variables mentioned above. The subgroup analysis or meta-regression will be utilized for finding the potential sources of heterogeneity. The funnel plot, Begg's or Egger's tests²⁸, and Fill and Trim methods²⁹ will be used for evaluating publication or reporting bias.

We will apply sensitivity analysis for assessing the relation between the primary study quality and the accuracy amount, and a one-out remove method will be used too³⁰.

The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) tool to assess the quality of evidence and strength of recommendations was originally developed for interventional systematic review and meta-analysis, but the standard tool was changed later for diagnostic and predictive systematic reviews. However, we will use the modified version of GRADE³¹ for prediction model systematic review.

Dissemination

We will present the systematic review and meta-analysis findings at relevant conferences. Moreover, we will submit the manuscript to relevant scholarly peer-reviewed journal.

Study status

This systematic review is currently in the search phase for information sources.

Discussion

CNN is a promising advanced method that can be useful in differentiating CRC from benign lesions, such as polyps. The results of this systematic review and meta-analysis will further our knowledge about the role of CNN in predicting colorectal lesions. As far as we know, no comprehensive systematic review has been published so far elaborating the role of CNN in predicting colorectal lesions.

We will try to conduct the study strictly adhering to the methods described. Any changes in the study conduction will be clarified in the final manuscript. Since heterogeneity is a common phenomenon in prediction model studies, we will try to address this concern by appropriate methods such as meta-regression and/or subgroup analysis.

Data availability

Underlying data

No data are associated with this article.

Extended data

Open Science Framework: Classification of different polyps in colorectal Endoscopy using Convolutional Neural Network; A Systematic Review Protocol Review title and timescale, <https://doi.org/10.17605/OSF.IO/QJ7EU>²¹.

This project contains the following extended data:

- PubMed search syntax.
- Protocol

Reporting guidelines

Repository name: PRISMA-P checklist³² ‘Classifying colorectal cancer or colorectal polyps in endoscopic setting using Convolutional Neural Network (CNN): protocol for a systematic review and meta-analysis’, <https://doi.org/10.17605/OSF.IO/QJ7EU>²¹.

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

References

1. Keum NN, Giovannucci E: **Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies.** *Nat Rev Gastroenterol Hepatol.* Nature Research, 2019; **16**(12): 713–32.
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Øines M, Helsingen LM, Bretthauer M, et al.: **Epidemiology and risk factors of colorectal polyps.** *Best Pract Res Clin Gastroenterol.* Bailliere Tindall Ltd, 2017; **31**(4): 419–24.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Wong MCS, Huang J, Huang JLW, et al.: **Global Prevalence of Colorectal Neoplasia: A Systematic Review and Meta-Analysis.** *Clin Gastroenterol Hepatol.* W.B. Saunders, 2020; **18**(3): 553–561.e10.
[PubMed Abstract](#) | [Publisher Full Text](#)
4. Brenner H, Hoffmeister M, Stegmaier C, et al.: **Risk of progression of advanced adenomas to colorectal cancer by age and sex: Estimates based on 840,149 screening colonoscopies.** *Gut.* 2007; **56**(11): 1585–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. Zauber AG, Winawer SJ, O'Brien MJ, et al.: **Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths.** *N Engl J Med.* 2012; **366**(8): 687–96.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Lieberman DA, Rex DK, Winawer SJ, et al.: **Guidelines for colonoscopy surveillance after screening and polypectomy: A consensus update by the us multi-society task force on colorectal cancer.** *Gastroenterology.* W.B. Saunders, 2012; **143**(3): 844–57.
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Winawer SJ, Zauber AG, Fletcher RH, et al.: **Guidelines for Colonoscopy Surveillance after Polypectomy: A Consensus Update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society.** *CA Cancer J Clin.* 2006; **56**(3): 143–59; quiz 184–5.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Lee JK, Jensen CD, Levin TR, et al.: **Long-term Risk of Colorectal Cancer and Related Death After Adenoma Removal in a Large, Community-based Population.** *Gastroenterology.* 2020; **158**(4): 884–894.e5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. McCarty TR, Aihara H: **Role of image-enhanced endoscopy: how to improve colorectal polyp detection rates in the coming decade.** *Gastrointest Endosc.* Mosby Inc, 2020; **91**(1): 113–4.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Van Rijn JC, Reitsma JB, Stoker J, et al.: **Polyp miss rate determined by tandem colonoscopy: A systematic review.** *Am J Gastroenterol.* 2006; **101**(2): 343–50.
[PubMed Abstract](#)
11. Lund M, Trads M, Njor SH, et al.: **Quality indicators for screening colonoscopy and colonoscopist performance and the subsequent risk of interval colorectal cancer: A systematic review.** *JBIR Database System Rev Implement Rep.* Lippincott Williams and Wilkins, 2019; **17**(11): 2265–300.
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Corley DA, Jensen CD, Marks AR, et al.: **Adenoma detection rate and risk of colorectal cancer and death.** *N Engl J Med.* 2014; **370**(14): 1298–306.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Lee TJW, Nair S, Beintaris I, et al.: **Recent advances in colonoscopy [version 1; peer review: 2 approved].** *F1000Res.* 2016; **5**: F1000 Faculty Rev-328.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. McGill S, Soetikno R, Kaltenbach T: **Image-enhanced endoscopy in practice.** *Can J Gastroenterol.* 2009; **23**(11): 741–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Chao WL, Manickavasagan H, Krishna SG: **Application of artificial intelligence in the detection and differentiation of colon polyps: A technical review for physicians.** *Diagnostics (Basel).* MDPI AG, 2019; **9**(3): 99.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Alagappan M, Brown JRG, Mori Y, et al.: **Artificial intelligence in gastrointestinal endoscopy: The future is almost here.** *World J Gastrointest Endosc.* 2018; **10**(10): 239–49.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Brinker TJ, Hekler A, Utikal JS, et al.: **Skin cancer classification using convolutional neural networks: Systematic review.** *J Med Internet Res.* 2018; **20**(10): e11936.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Gardezi SJ, Elazab A, Lei B, et al.: **Breast cancer detection and diagnosis using mammographic data: Systematic review.** *J Med Internet Res.* 2019; **21**(7): e14464.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Azer SA: **Deep learning with convolutional neural networks for**

- identification of liver masses and hepatocellular carcinoma: A systematic review.** *World J Gastrointest Oncol.* 2019; **11**(12): 1218–30.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Murray NM, Unberath M, Hager GD, *et al.*: **Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: A systematic review.** *J Neurointerv Surg.* 2020; **12**(2): 156–64.
[PubMed Abstract](#) | [Publisher Full Text](#)
 21. Keshtkar K: **Classification of different polyps in colorectal Endoscopy using Convolutional Neural Network; A Systematic Review Protocol Review title and timescale.** 2020.
<http://www.doi.org/10.17605/OSF.IO/QJ7EU>
 22. Wolff RF, Moons KGM, Riley RD, *et al.*: **PROBAST: A tool to assess the risk of bias and applicability of prediction model studies.** *Ann Intern Med.* 2019; **170**(1): 51–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
 23. Simonyan K, Zisserman A: **Very deep convolutional networks for large-scale image recognition.** arXiv Prepr arXiv14091556. 2014.
[Reference Source](#)
 24. Krizhevsky A, Sutskever I, Hinton GE: **Imagenet classification with deep convolutional neural networks.** In: Advances in neural information processing systems. 2012; 1097–105.
[Publisher Full Text](#)
 25. Szegedy C, Liu W, Jia Y, *et al.*: **Going deeper with convolutions.** In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015; 1–9.
[Publisher Full Text](#)
 26. Debray TPA, Damen JAA, Snell KIE, *et al.*: **A guide to systematic review and meta-analysis of prediction model performance.** *BMJ.* 2017; **356**: i6460.
[PubMed Abstract](#) | [Publisher Full Text](#)
 27. Higgins JPT, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** *Stat Med.* 2002; **21**(11): 1539–58.
[PubMed Abstract](#) | [Publisher Full Text](#)
 28. Egger M, Smith GD, Schneider M, *et al.*: **Bias in meta-analysis detected by a simple , graphical test measures of funnel plot asymmetry.** *BMJ.* 1997; **315**(7109): 629–34.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 29. Duval S, Tweedie R: **Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis.** *Biometrics.* 2000; **56**(2): 455–63.
[PubMed Abstract](#) | [Publisher Full Text](#)
 30. **The Handbook of Research Synthesis and Meta-Analysis - Google Scholar.** [cited 2020 Jul 5].
[Reference Source](#)
 31. Iorio A, Spencer FA, Falavigna M, *et al.*: **Use of GRADE for assessment of evidence about prognosis: Rating confidence in estimates of event rates in broad categories of patients.** *BMJ.* 2015; **350**: h870.
[PubMed Abstract](#) | [Publisher Full Text](#)
 32. Moher D, Shamseer L, Clarke M, *et al.*: **Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement.** *Syst Rev.* 2016; **20**(2): 148–60.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status: ?

Version 1

Reviewer Report 15 September 2020

<https://doi.org/10.5256/f1000research.28194.r70722>

© 2020 Tahamtan I. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Iman Tahamtan

School of Information Sciences, College of Communication and Information, University of Tennessee, Knoxville, TN, USA

- This is an interesting protocol that aims to evaluate the diagnostic accuracy for discrimination of colorectal cancer (CRC) and colorectal polyps (CRP) using a convolutional neural network (CNN).
- The authors have very well explained previous systematic reviews (and studies) related to the study's topic and their rationale for conducting this study. One rationale is mentioned to be methodological issues of previous research, such as not doing the risk of bias assessment in previous systematic reviews or studies being conducted without including a meta-analysis.
- I ask the authors to be unique and explicit in stating the objective of the study. In the 'abstract' of the protocol, they mention that the aim of the study is classifying CRP and CRC using CCN, while it is mentioned to be evaluating the diagnostic accuracy for discrimination of CRC and CRP using CNN. Please also explain, in the objectives (in both the abstract and text) that this study aims to investigate the accuracy of CNN in classifying CRC and CRP compared to its accuracy in identifying the normal colorectal tissue (Comparison: C).
- In the last paragraph of the introduction, the authors mention that one limitation of previous studies is 'not registering their protocols'. Although this may seem to be a limitation, does this justify the need for conducting this research? If the answer is yes, please explain how? However, the second and third provide a strong rationale for the need for conducting this research.
- Although it is clear to experts what type of primary studies will be included in this study, please make some language modifications to the first sentence to make it more transparent. For instance, you could change it as follows: "The primary studies with all the following features will be included in this systematic review and meta-analysis: ...".

- Type of participants: it is not clear why when CRPs are usually diagnosed in adults older than 50 years of age, the authors mention that 'patients older than 18 years will be eligible to be included' in this systematic review. Shouldn't it have been 50 instead of 18?
- Index test (the output of prediction model): please explain what you mean by 'class labels'? What are the features that the hidden layer will extract?
- Please clarify the following sentence: "The fully connected layers are used for the classification and detection object in the input images, at the end of the CNN network". Detection of which objects in the input images? It is not clear what the authors mean in 'at the end of the CNN network'. Does this mean the final thing that the CNN does is the detection of objects in the input images?
- If you have already designed the screening checklist, please mention which criteria it includes?
- Please explain whether you will include all studies with any risk of bias or if those with low and unclear risk of bias will be excluded.
- I also recommend that the authors explain CNN in non-jargon language, for example, by providing an example so that non-experts in machine learning would better understand how CNN works.
- Briefly explain the Debray *et al.* guidelines for combining the primary and secondary data.
- In the following sentence, in parentheses specify the variables again: "We will use the subgroup analysis based on the variables mentioned above".

Is the rationale for, and objectives of, the study clearly described?

Yes

Is the study design appropriate for the research question?

Yes

Are sufficient details of the methods provided to allow replication by others?

Yes

Are the datasets clearly presented in a useable and accessible format?

Yes

Competing Interests: I have already contacted the journal and received permission for reviewing this protocol.

Reviewer Expertise: Systematic review, designing search strategies in academic databases

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research