Explainable AI For Colorectal Lesion Classification Using Deep Learning Models With Attention Mechanism

Muhammad Muzzammil Auzine*
mmuzzammil.auzine@gmail.com
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Maleika Heenaye-Mamode Khan
m.mamodekhan@uom.ac.mu
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Sunilduth Baichoo
sbaichoo@uom.ac.mu
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Preeti Bissoonauth-Daiboo
preetidaiboo@gmail.com
Department of Software and Information Systems, University of Mauritius
Reduit, Mauritius

Zaid Heetun
zaidheetun@gmail.com
Center for Gastroenterology and Hepatology, Dr Abdool Gaffoor Jeetoo Hospital
Port Louis, Mauritius

Xiaohong. Gao
x.gao@mdx.ac.uk
Department of Computer Science, Middlesex University
London, United Kingdom

ABSTRACT

In the context of identifying colorectal lesions in colonoscopy images, previous AI models have predominantly focused on identifying abnormalities without providing comprehensible explanations for their predictions. To address this limitation, this paper introduces a novel framework that utilises two customised deep convolutional neural networks which we have developed, namely: ColoRecNet and Attention-BasedColoRecNet . These models aim to enhance human comprehension by providing interpretable explanations for their predictions. The framework also incorporates Grad-Cam, a visualisation technique, to highlight the distinctive features associated with each lesion class. The performance evaluation of the models on the testing set demonstrates that the Attention-BasedColoRecNet model surpasses the ColoRecNet model in terms of overall accuracy (95.67%), precision (96.02%), and recall (92.47%). Furthermore, the visual explanations generated by Grad-CAM heatmaps serve as additional validation, reinforcing that the Attention-BasedColoRecNet model possesses improved discriminative power and feature representations, resulting in superior classification performance.

CCS CONCEPTS

• **Computing methodologies → Interest point and salient region detections**.

KEYWORDS

Deep Learning, Medical Imaging, Attention Model, Explainable AI

ICAAI 2023, October 13–15, 2023, Istanbul, Turkiye

ACM ISBN 979-8-4007-0898-5/23/10…$15.00
https://doi.org/10.1145/3633598.3633617

ACM Reference Format:
Muhammad Muzzammil Auzine, Maleika Heenaye-Mamode Khan, Sunilduth Baichoo, Preeti Bissoonauth-Daiboo, Zaid Heetun, and Xiaohong. Gao. 2023. Explainable AI For Colorectal Lesion Classification Using Deep Learning Models With Attention Mechanism. In *2023 The 7th International Conference on Advances in Artificial Intelligence (ICAAI) (ICAAI 2023), October 13–15, 2023, Istanbul, Turkiye.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3633598.3633617

1 INTRODUCTION

Colorectal cancer(CRC) is a significant global health concern, ranking among the leading causes of cancer-related deaths worldwide. In 2020, it accounted for approximately 10% of all cancer cases and 9.4% of cancer-related deaths according to the World Health Organization (WHO) [18]. Colorectal cancer, also known as bowel cancer, originates in the colon or rectum and typically develops from precancerous polyps. It is characterised by the uncontrolled growth of abnormal cells that can invade nearby tissues and metastasize to other parts of the body. Globally, the incidence of colorectal cancer has been steadily increasing. In 2020 alone, there were an estimated 1.8 million new cases reported [23]. Timely detection is critical, as colorectal cancer often exhibits minimal symptoms in its early stages. Hence, accurate and efficient classification methods are necessary to enable early intervention and personalised treatment plans.

Deep learning models have gained significant popularity in medical research due to their ability to learn complex patterns from large datasets, especially in the analysis of medical images. By utilizing neural networks with multiple layers, these models can automatically extract relevant features and capture intricate patterns that may be challenging for traditional machine learning approaches. In the context of colorectal cancer classification, deep learning models have shown promise in improving accuracy and diagnostic performance. Numerous studies have reported high classification accuracies, ranging from 85% to 95%, using deep learning models on colorectal cancer datasets [3, 10]. These models have the potential to assist healthcare professionals in accurately identifying cancerous regions, enabling timely intervention and reducing the burden of manual analysis.

In addition to achieving high accuracy, interpretability of AI models is crucial in the medical field, where decisions can significantly impact patient care. Explainable AI techniques aim to provide insights into the decision-making process of AI models, allowing clinicians and researchers to understand and trust the predictions made by these models [13]. In the context of colorectal cancer classification, explainable AI functions can help unravel the black-box nature of deep learning models by providing valuable explanations for their predictions. This interpretability assists clinicians in validating the decisions made by AI models, fostering trust and facilitating the integration of these models into clinical practice [7].

2 LITERATURE REVIEW

2.1 Colorectal Cancer

Colorectal cancer (CRC) is the cancer that affects the colon and rectum. Colon cancer or rectal cancer are often grouped together due to their common characteristics. The Colon and rectum make up the large intestine in the human body's gastrointestinal (GI) tract and plays a vital role in the body's ability to process waste. The colon and rectum consist of five sections as depicted on Figure 1(a).

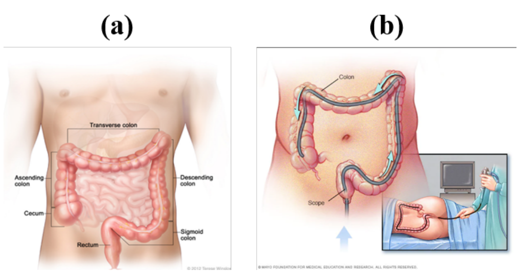

Figure 1: (a):Anatomy of the large intestine [9]; (b): Colonoscopy procedure [5]

Among the various visual examinations techniques available for lower gastrointestinal cancer, colonoscopy is widely regarded as the most effective screening technique when compared to CT and MRI. It offers the unique advantage of providing a comprehensive examination of the entire gastrointestinal tract and allows for the removal of polyps, if necessary, all in a single session [6]. Colonoscopy procedure as illustrated on Figure 1(b) involves a direct visual examination of the entire colon and rectum and is typically performed by a gastroenterologist or surgeon.

2.2 Artificial intelligence in medical research

In recent years, there has been a significant increase in the utilization of artificial intelligence (AI) in the diagnosis and treatment of gastrointestinal diseases [14], particularly in the lower gastrointestinal tract. Various studies have demonstrated the effectiveness of AI in assessing colorectal lesions such as colon polyps, adenomas, cancers, ulcerative colitis, and intestinal motor diseases. While research specifically focused on AI applications for colorectal cancer (CRC) is limited, the growing presence of AI in the medical field suggests that AI will eventually play a crucial role in CRC diagnosis and treatment [22].

However, the implementation of AI-assisted systems for diagnosis in healthcare institutions faces a major challenge due to their reputation as "black boxes" in terms of their computational processes. The lack of transparency in decision-making raises concerns and inhibits the widespread adoption of AI-driven systems, despite their proven effectiveness [2, 26]. To address this issue, Explainable Artificial Intelligence (XAI) has emerged as a field that offers potential solutions by providing explanations for model predictions. XAI techniques can be categorised into various approaches [24], including dimension reduction, feature importance, attention mechanism, knowledge distillation, and surrogate representations. [19] further categories explanation methods into three types: visual explanations, textual explanations, and example-based explanations These techniques offer insights into the reasoning behind AI-based medical image analysis, facilitating transparency and bridging the gap between AI capabilities and the need for explainability in healthcare applications.

2.3 Related Works

Colonoscopy can be employed to directly identify lesions in the intestinal wall, and colonoscopist can use image processing and screening to evaluate whether lesions are associated to CRC. A recent study by [11] introduced a deep learning-based system based on the ResNet50 architechture that enables real-time automated diagnosis of colorectal cancer invasion. The authors highlighted the significant potential of this system in assisting with CRC diagnosis during clinical practice.

The development of a dedicated classification model based on the Caffenet architechture for colorectal polyps was undertaken by[25] . Their approach showcased remarkable performance, surpassing that of endoscopists' visual inspection in terms of recall and accuracy, while preserving a comparable level of precision.

In their research, [4] developed and evaluated a computer-aided diagnosis system called DNN-CAD based on the InceptionV3 architecture for the analysis of narrow-band images of small colorectal polyps. The dataset consisted of 1476 neoplastic polyp images and 681 hyperplastic polyp images, alongside corresponding histologic findings. DNN-CAD achieved a sensitivity of 96.3% and specificity of 78.1% in accurately identifying neoplastic or hyperplastic polyps smaller than 5 mm in the test set. The system surpassed both expert and novice endoscopists in terms of diagnostic time, demonstrating faster classification of polyps. Additionally, DNN-CAD showed high accuracy and intra-observer agreement, indicating its potential as a valuable tool for endoscopic image recognition and other medical image analyses.

[12] aimed to evaluate the use of an artificial intelligence (AI) assisted image classifier for determining the feasibility of curative endoscopic resection of large colonic lesions based on non-magnified endoscopic images. The dataset consisted of 8,000 endoscopic images used to train the AI image classifier, and an independent validation set of 567 endoscopic images from 76 colonic lesions. Their main finding was that the trained AI image classifier based on non-magnified images accurately predicted the probability of curative

resection for large colorectal lesions and performed better than junior endoscopists.

[27] aimed to address the inherent subjectivity and potential diagnostic errors associated with histopathological grading of colorectal cancers. To achieve this, the researchers developed a computer-aided diagnosis (CAD) method named HCCANet. Employing a dataset consisting of 630 histopathology images, the authors introduced a novel attention mechanism called MCCBAM, which was integrated into the CNN-based HCCANet model. The experimental findings demonstrated the superior performance of HCCANet over advanced deep learning and classical machine learning techniques

[21] presented interpretable real-time deep neural networks with Shapley additive explanations For polyp identification, polyp categorization, and polyp segmentation under colonoscopy. The suggested technique demonstrates better real-time performance. The experimental results show that the proposed technique outperforms existing deep learning methods. Furthermore, they achieved gratifying operational efficiency and interpretable feedback, which fulfilled the requirements of the colorectal surgeon.Similarly, [1] utilised the Gradient-weighted Class Activation Mapping technique to visualize the features responsible for classifying an image into a specific leukemia subtype.

Upon reviewing the existing literature on artificial intelligence (AI) for the detection and classification of colorectal cancer, it was identified that further investigation in this research domain holds promise. The prevailing focus of AI models has been on identifying abnormalities in medical images. Nevertheless, despite recent attention from researchers, there have been limited efforts in developing interpretable models that can provide explanations for the predictions made by the models, thereby facilitating human understanding.

3 METHODS

This research paper introduces an innovative framework designed for the identification of colorectal lesions in colonoscopy images. To accomplish this, we have developed two custom deep convolutional neural networks (DCNNs) known as ColoRecNet and Attention-BasedColoRecNet. These networks are constructed based on the Densenet169 and ECA-Net architectures, respectively. Subsequently, To enhance interpretability, an XAI (Explainable Artificial Intelligence) technique has been employed to visualize the distinctive features associated with each lesion class. The schematic representation of the proposed framework can be observed in Figure 2.

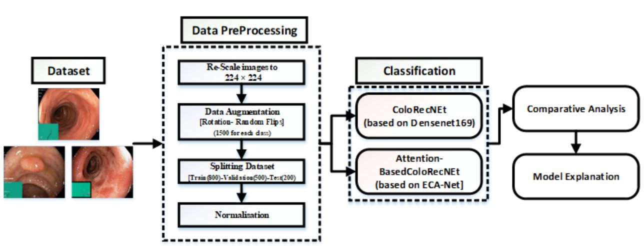

Figure 2: Proposed Colorectal Cancer detection framework using ColoRecNet

3.1 Dataset

The advancement of numerous computing domains, particularly those involving deep learning applications, relies heavily on the availability and quality of datasets. For our research, we utilised the Kvasir dataset, which was developed by [16] in 2017. This dataset consists of a comprehensive collection of images that have been carefully annotated and validated by experienced endoscopists, who are medical experts in the field. The Kvasir dataset encompasses various classes of endoscopic images, including anatomical landmarks and pathological or endoscopic findings within the gastrointestinal (GI) tract. we specifically focused on the Polyps and Ulcerative Classes from the pathological findings subset of the Kvasir dataset. Additionally, we acquired Normal colon images from the Kaggle platform [15]. Figure 3 illustrates a sample for three classes namely:

(a) **Normal Colon** - Healthy normal colon.
(b) **Polyps** - lesions that can be identified as mucosal outgrows.
(c) **Ulcerative Colitis** - a long termed inflammatory condition that affects the colon and rectum.

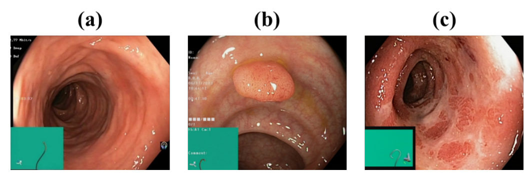

Figure 3: Sample images from dataset. (a): Normal Colon, (b): Polyps, (c): Ulcerative Colitis

3.1.1 *Dataset Preprocessing.*

For each class image bilinear interpolation was applied and images scaled to 224 × 224 pixels. Moreover, image transformations processes were performed including random flipping and rotation. This augmentation process resulted in 1500 images for each class. The dataset were then splitted into Training set, Validation set and Test set in the ratio of 8:5:2. Before feeding images to the network, images were transformed into properly preprocessed floating-point tensors by using normalization.

3.2 Model development

3.2.1 *ColoRecNet.*

ColoRecNet is a deep convolutional network, adapted from the architecture of DensNet169 [8]. ColoRecNet incorporates dense connectivity, where each layer is connected to all previous layers, enabling feature reuse and better gradient flow and combines components such as dense blocks with multiple layers, transition layers for dimensionality reduction, and global average pooling to summarize feature maps in order to create a powerful CNN architecture for accurate image classification. Additionally, an output layer consisting of three neurons with softmax activation function was appended to accommodate the multiclass classification task. The softmax activation function in the output layer facilitates the interpretation of the outputs as probabilities assigned to each class. It was specifically trained on the dataset we developed.

3.2.2 *Attention-BasedColoRecNet.*

Attention-BasedColoRecNet is an extension of ColoRecNet, which is constructed by integrating the

Efficient Channel Attention Network (ECA-Net) architecture. ECA-Net is a convolutional neural network architecture proposed by [20] specifically designed to effectively model the interdependencies among channels in an image. Its fundamental principle involves the seamless integration of a channel attention mechanism within the convolutional layers of the network. This attention mechanism selectively amplifies informative channels while suppressing less relevant ones, thereby facilitating the network's ability to capture long-range dependencies across channels.

3.3 Visual explanation

We employed the Grad-CAM algorithm to identify the image regions that contributed the most to the predictions made by our classification model. Grad-CAM, short for Gradient-weighted Class Activation Maps, was initially introduced by [17]. It serves as a technique to generate "visual explanations" that provide insights into how a convolutional neural network (CNN)-based model arrived at a particular decision, thereby enhancing the transparency of the model's decision-making process. Grad-CAM achieves this by utilizing the gradients of the predicted outcomes, which flow into the output convolutional layer, to generate a localised map that highlights the important regions in the image that influenced the final prediction. By applying Grad-CAM, we were able to gain a better understanding of the model's reasoning and identify the specific image areas that played a crucial role in the classification process.

4 RESULT ANALYSIS

The primary aim of this study was to investigate and apply explainable Artificial Intelligence (XAI) techniques in the field of medical image analysis. To achieve this we have developed two models: ColoRecNet, a Deep Convolutional Neural Network (CNN), and attentionbasedColoRecNet, which incorporated an XAI technique using an attention mechanism, as categorised by a [24]. Both models were then trained for 100 epochs, with a batch size of 32. To optimize the performance of these models, we employed the Adam optimization algorithm, utilizing categorical crossentropy as the loss function. Finally, to accommodate the multiclass classification task, the Softmax activation function was applied. Following the model training, a comparative analysis was conducted, evaluating the models' performance based on metrics such as accuracy, precision, and recall, as illustrated in Table 1.

Table 1: Model Comparative based on Performance Metrics

Model	Set	Acc(%)	Preci(%)	Rec(%)
ColoRecNEt	Training	96.33	96.33	96.33
	Validation	87.73	88.58	87.73
	Testing	88.33	89.47	88.33
Attention-Based ColoRecNEt	Training	99.63	99.63	99.63
	Validation	92.47	93.03	92.20
	Testing	95.67	96.02	92.47

Acc: Accuracy; Preci: Precision; Rec: Recall.

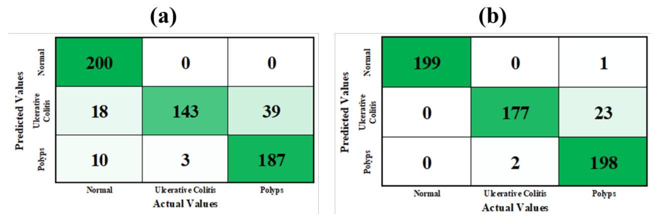

Figure 4: Confusion matrix from the Test Set. (a):ColoRecNET,(b): Attention-BasedColoRecNET.

The analysis of Table 1 reveals that the Attention-BasedColoRecNet model has outperformed the ColoRecNet model. The results demonstrate that the Attention-BasedColoRecNet has shown improvements in terms of overall accuracy, precision, and recall. Specifically, on the testing set, there has been an increase of 8.31% in accuracy, 7.32% in precision, and 4.69% in recall when compared to the ColoRecNet model.

Furthermore, the visualization of the confusion maps from Figure 4 indicates that the Attention-BasedColoRecNet has significantly enhanced the classification task, as it only misclassified 26 samples, compared to the ColoRecNet's 70 misclassified samples. These findings highlight the superior performance and effectiveness of the Attention-BasedColoRecNet model in comparison to the ColoRecNet model.

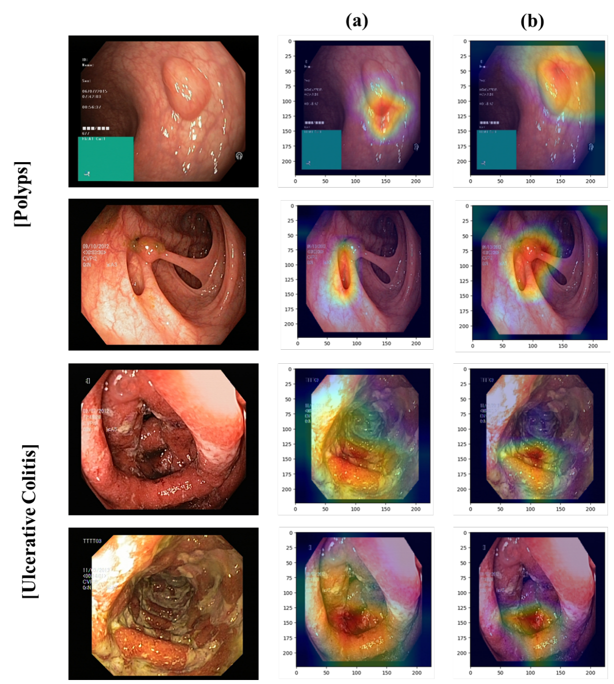

Figure 5: GradCam Visualisation on Polyps and Ulcerative Colitis. (a):ColoRecNET, (b): Attention-BasedColoRecNET.

Visual explanations, categorised as an explanation method by [19], were employed in this study. Specifically, Grad-CAM was applied to the model to elucidate deterministic features associated with a typical class. By examining the Grad-CAM heatmaps, it was possible to gain insights into the learned features and better comprehend the inner workings of complex deep learning models. The intensity of the heatmap signifies the importance of each region in the image. Higher intensity indicates a strong contribution of the corresponding region to the network's prediction for the target class, while regions with lower intensity have less influence. Figure 5 depicts the GradCam Visualisation for Polyps and Ulcerative Colitis classes. Notably, an interesting finding from Figure 5 is that while ColoRecNet correctly classified some samples into their respective classes, the Attention-BasedColoRecNet, enhanced by the ECA module, exhibited improved discriminative power in the network, leading to enhanced feature representations. This improvement is particularly evident in the polyps class in Figure 5.

5 CONCLUSION

This research study emphasizes the potential of deep learning models and explainable AI techniques in the classification of colorectal cancer. By developing custom deep convolutional neural networks and incorporating Grad-CAM, the study has demonstrated notable improvements in accuracy, precision, and recall for identifying colorectal lesions in colonoscopy images. Particularly, the Attention-BasedColoRecNet model, enhanced with the ECA module, has showcased superior performance compared to the ColoRecNet model. The visualization of Grad-CAM heatmaps has provided valuable insights into the learned features, thereby facilitating a better understanding of the underlying mechanisms of the deep learning models. The findings of this research contribute to the advancement of interpretable AI models for colorectal cancer classification, which can assist medical professionals in making accurate diagnoses and informed decisions about patient care. Future research endeavors could focus on further validating the effectiveness and potential impact of these models in larger datasets and real-world clinical settings.

REFERENCES

[1] Arjun Abhishek, Rajib Kumar Jha, Ruchi Sinha, and Kamlesh Jha. 2023. Automated detection and classification of leukemia on a subject-independent test dataset using deep transfer learning supported by Grad-CAM visualization. *Biomedical Signal Processing and Control* 83 (2023), 104722.

[2] Julia Amann, Alessandro Blasimme, Effy Vayena, Dietmar Frey, Vince I. Madai, and the Precise4Q consortium. 2020. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making* 20, 1 (Nov. 2020), 310. https://doi.org/10.1186/s12911-020-01332-6

[3] Dmitrii Bychkov, Nina Linder, Riku Turkki, Stig Nordling, Panu E. Kovanen, Clare Verrill, Margarita Walliander, Mikael Lundin, Caj Haglund, and Johan Lundin. 2018. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific Reports* 8, 1 (Feb. 2018), 3395. https://doi.org/10.1038/s41598-018-21758-3 Number: 1 Publisher: Nature Publishing Group.

[4] Peng-Jen Chen, Meng-Chiung Lin, Mei-Ju Lai, Jung-Chun Lin, Henry Horng-Shing Lu, and Vincent S. Tseng. 2018. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 154, 3 (Feb. 2018), 568–575. https://doi.org/10.1053/j.gastro.2017.10.010

[5] Mayo clinic. 2011. Colonoscopy - Mayo Clinic. https://www.mayoclinic.org/tests-procedures/colonoscopy/about/pac-20393569 Publication Title: Colonoscopy.

[6] Nisa Desai and Loren Rabinowitz. 2022. Colon cancer screening decisions: What's the best option and when? https://www.health.harvard.edu/blog/colon-cancer-screening-decisions-whats-the-best-option-and-when-202206152762 Section: Staying Healthy.

[7] Finale Doshi-Velez and Been Kim. 2017. Towards A Rigorous Science of Interpretable Machine Learning. *arXiv: Machine Learning* (2017).

[8] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. 2017. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* 4700–4708.

[9] National Cancer Institute. 2011. Definition of colon - NCI Dictionary of Cancer Terms - NCI. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/colon Archive Location: nciglobal,ncienterprise.

[10] Jakob Nikolas Kather, Alexander T. Pearson, Niels Halama, Dirk Jäger, Jeremias Krause, Sven H. Loosen, Alexander Marx, Peter Boor, Frank Tacke, Ulf Peter Neumann, Heike I. Grabsch, Takaki Yoshikawa, Hermann Brenner, Jenny Chang-Claude, Michael Hoffmeister, Christian Trautwein, and Tom Luedde. 2019. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nature Medicine* 25, 7 (July 2019), 1054–1056. https://doi.org/10.1038/s41591-019-0462-y Number: 7 Publisher: Nature Publishing Group.

[11] Zihua Lu, Youming Xu, Liwen Yao, Wei Zhou, Wei Gong, Genhua Yang, Mingwen Guo, Beiping Zhang, Xu Huang, Chunping He, Rui Zhou, Yunchao Deng, and Honggang Yu. 2022. Real-time automated diagnosis of colorectal cancer invasion depth using a deep learning model with multimodal data (with video). *Gastrointestinal Endoscopy* 95, 6 (June 2022), 1186–1194.e3. https://doi.org/10.1016/j.gie.2021.11.049

[12] Thomas KL Lui, Kenneth KY Wong, Loey LY Mak, Michael KL Ko, Stephen KK Tsao, and Wai K Leung. 2019. Endoscopic prediction of deeply submucosal invasive carcinoma with use of artificial intelligence. *Endoscopy international open* 7, 04 (2019), E514–E520.

[13] Scott M. Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. 2019. Explainable AI for Trees: From Local Explanations to Global Understanding. arXiv:1905.04610 [cs.LG]

[14] Jun Ki Min, Min Seob Kwak, and Jae Myung Cha. 2019. Overview of Deep Learning in Gastrointestinal Endoscopy. *Gut and Liver* 13, 4 (Jan. 2019), 388–393. https://doi.org/10.5009/gnl18384

[15] Francis Jesmar Montalbo. [n. d.]. WCE Curated Colon Disease Dataset Deep Learning. https://www.kaggle.com/datasets/francismon/curated-colon-dataset-for-deep-learning

[16] Konstantin Pogorelov, Kristin Randel, Carsten Griwodz, Thomas de Lange, Sigrun Eskeland, Dag Johansen, Concetto Spampinato, Duc Tien Dang Nguyen, Mathias Lux, Peter Schmidt, Michael Riegler, and Pål Halvorsen. 2017. *KVASIR: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection.* https://doi.org/10.1145/3083187.3083212

[17] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision.* 618–626.

[18] Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* 71, 3 (2021), 209–249.

[19] Bas H. M. van der Velden, Hugo J. Kuijf, Kenneth G. A. Gilhuijs, and Max A. Viergever. 2022. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis* 79 (July 2022), 102470. https://doi.org/10.1016/j.media.2022.102470

[20] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. 2020. ECA-Net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* 11534–11542.

[21] Sutong Wang, Yunqiang Yin, Dujuan Wang, Zehui Lv, Yanzhang Wang, and Yaochu Jin. 2021. An interpretable deep neural network for colorectal polyp diagnosis under colonoscopy. *Knowledge-Based Systems* 234 (Dec. 2021), 107568. https://doi.org/10.1016/j.knosys.2021.107568

[22] Yutong Wang, Xiaoyun He, Hui Nie, Jianhua Zhou, Pengfei Cao, and Chunlin Ou. 2020. Application of artificial intelligence to the diagnosis and therapy of colorectal cancer. *American journal of cancer research* 10 (Nov. 2020), 3575–3598.

[23] WHO. 2022. Cancer. https://www.who.int/news-room/fact-sheets/detail/cancer

[24] Guang Yang, Qinghao Ye, and Jun Xia. 2022. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Information Fusion* 77 (Jan. 2022), 29–52. https://doi.org/10.1016/j.inffus.2021.07.016

[25] Ruikai Zhang, Yali Zheng, Tony Wing Chung Mak, Ruoxi Yu, Sunny H. Wong, James Y. W. Lau, and Carmen C. Y. Poon. 2017. Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE journal of biomedical and health informatics* 21, 1 (Jan. 2017), 41–47. https://doi.org/10.1109/JBHI.2016.2635662

[26] Yiming Zhang, Ying Weng, and Jonathan Lund. 2022. Applications of Explainable Artificial Intelligence in Diagnosis and Surgery. *Diagnostics* 12, 2 (Feb. 2022), 237. https://doi.org/10.3390/diagnostics12020237 Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[27] Panyun Zhou, Yanzhen Cao, Min Li, Yuhua Ma, Chen Chen, Xiaojing Gan, Jianying Wu, Xiaoyi Lv, and Cheng Chen. 2022. HCCANet: histopathological image grading of colorectal cancer using CNN based on multichannel fusion attention mechanism. *Scientific Reports* 12, 1 (2022), 15103.