International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

Colorectal polyp detection: Elevation of convolutional neural network-based models

Almo'men Bellah Alawnah *

Department of Industrial Engineering, Jordan University of Science and Technology, Irbid, Jordan.

International Journal of Science and Research Archive, 2025, 14(03), 987-993

Publication history: Received on 08 February 2025; revised on 16 March 2025; accepted on 19 March 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.14.3.0754

Abstract

Colorectal polyp detection plays a crucial role in the early diagnosis and prevention of colorectal cancer. This study evaluates multiple convolutional neural network (CNN) models for binary classification of polyp images using a public dataset of 3,000 images (1,500 polyp and 1,500 non-polyp). We compare deep learning architectures based on ResNet101, ResNet50, VGG16, VGG19, Xception, R-CNN, and training time based on accuracy, recall, precision, F1-score, and training time. The results indicate that both ResNet101 and our CNN model achieved the highest performance metrics, with ResNet101 reaching an accuracy of 99.96% and our CNN model achieving 100% accuracy. However, considering computational efficiency, our CNN model demonstrated the shortest training time (1,302 MS), making it the optimal solution for balancing performance and time efficiency. These findings highlight the potential of CNN-based models for real-time polyp detection, which can enhance early diagnosis and improve clinical outcomes.

Keywords: Deep Learning; Colorectal Polyps; Binary Classification; And Detection of Disease-Based Endoscopic Images

1. Introduction

The gastrointestinal tract is essential for human health, controlling nutrient absorption and digestion [1]. A delayed diagnosis of colorectal polyps can heighten the risk of developing colorectal cancer, the second leading cause of death worldwide, according to the International Agency for Research on Cancer (IARC) [2]. Detecting and treating colorectal polyps early can significantly lower the incidence and mortality rates of colorectal cancer [3]. Identifying and removing colonic polyps can help prevent cancer from metastasizing. Colorectal polyps, which include various types of intestinal tumors, are among the most observed gastrointestinal tract disorders [4]. Most colorectal cancers originate from initially benign conditions [5]. Several factors contribute to the development of colorectal polyps, including smoking, obesity, genetics, alcohol consumption, and aging [6]. Many malignant polyps, including small and flat ones, are often missed during routine examinations [7]. Implementing deep learning techniques has effectively reduced error rates while enhancing polyp detection in colonoscopy images. Accurate and timely detection of gastrointestinal diseases is essential for effective treatment and improved patient outcomes. Traditional diagnostic methods may lack sensitivity, highlighting the need for innovative approaches. Additionally, conventional diagnostic methods overlook around 25% of colorectal polyps [8]. Diagnosing these polyps without image analysis is often difficult due to factors such as error-prone interpretation, awkward positioning, blind areas, and small sizes.

1.1. Problem Statement

The medical field generates vast amounts of data from various hospital sources, including video, text, audio, and images, which require analysis and classification. When artificial intelligence is applied, processing and interpreting this data becomes easier, enabling automatic diagnoses. Recently, artificial intelligence has been used for image recognition in procedures like colonoscopy, upper endoscopy, and wireless capsule endoscopy (WCE).

* Corresponding author: Al-Mo'men Bellah Alawnah Email: akalawnah19@eng.just.edu.jo

Detecting features in these images is crucial, but traditional screening methods struggle with the large volume of images, requiring significant time and effort to interpret. Understanding and classifying these images is particularly challenging when dealing with large datasets. In traditional methods, analyzing medical images involved considerable physical and psychological effort, from data collection to understanding disease structures, their spread, and potential treatments.

1.2. Contribution

- Develop Convolutional Neural Networks (CNNs), a deep learning model, to detect and classify colorectal polyps in endoscopic images.
- Train the CNNs using annotated datasets of endoscopic images, ensuring the model learns to identify various polyp characteristics and categorize them accurately.
- Evaluate the model's performance using accuracy, sensitivity, specificity, and F1 score metrics to assess its effectiveness in real-world applications.
- Optimize the CNNs to handle challenges such as small, flat, or obscured polyps, often overlooked in traditional diagnostic methods.
- Continuously refine the model based on feedback and new data to improve detection rates and reduce errors.
- Elevation Of Convolutional Neural Network-Based Models.

1.3. Related Work

According to [9], PolypEffNetV1 is a multifunctional model that combines U-Net for disease segmentation and EfficientNetB5 for classification. The model was trained using the KVASIR and CVC datasets, consisting of 1612 images, which include 1696 polyp regions and 760 non-polyp inflammatory regions. The segmentation process utilized publicly available, open-source KVASIR and CVC datasets. A detailed analysis of the model's metrics showed excellent performance, with PolypEffNetV1 achieving 97% testing accuracy, 84% Jaccard index, 91% Dice coefficient, and 89% F1-score. The classifier demonstrated a validation accuracy of approximately 99%, with specificity at 98% and sensitivity at 99%. For binary classification—distinguishing between polyps and non-polyps—a CNN model was introduced in research [10]. The testing, validation, and training data were collected from three merged databases, and all images were resized to 224x224 pixels, with patient information removed. The dataset contained 4,808 images, of which 3,000 were non-polyp images, and the remaining images contained polyps. The data was then split into training and validation sets. The architecture employed were MobileNetV2, VGG16_bn, and ResNet (34, 50, and 152), all pre-trained using the ImageNet dataset. The models achieved an average accuracy of 95.70%. This study also aimed to analyze the effect of layer depth on accuracy, with ResNet 34, 50, and 152 representing different layer depths. Among these, the ResNet-152 model achieved the highest accuracy at 96.60%, while the VGG16_bn model achieved the lowest accuracy at 94.00%. Research [11] proposed an image-trained model to classify polyps based on their types using 1,991 images gathered from Taipei Medical University. Among these images, 938 depicted adenomatous and adenocarcinomas, while the remaining 1,053 represented hyperplastic polyps. The Deep Convolutional Neural Network (DCNN) algorithm autonomously extracted textural features from each image, combined them with a classifier, and generated features for classification. Inception-V3, ResNet-101, and DenseNet-201 were DCNN architectures used. When AlexNet was trained from scratch, it achieved the highest performance accuracy of approximately 96.4%, outperforming textural and transfer learning. Additionally, an accuracy of 75.6% was achieved by utilizing features extracted from the GLCM texture on the blue channel. Their study [12] implemented the DP-CNN model to classify colonoscopy images as polyps or non-polyps, with SIGMOD as the classifier. Public datasets from the CVC ClinicDB were used to develop the proposed model, which was tested on the CVC ColonDB and ETIS-Larib databases. The model achieved an accuracy of 90.81% on the ETIS-Larib database and 99.60% on the CVC ColonDB database. In terms of precision, the model achieved a perfect score of 100% on the CVC ColonDB database. For the CVC ColonDB dataset, the recall, F1 score, and F2 score were 99.20%, 99.60%, and 99.83%, respectively, with a precision of 89.81%. On the ETIS-Larib database, the recall, F1 score, and F2 score were 89.91%, 92.85%, and 91.00%, respectively. The proposed model achieved 8,737 LP (Low Power) values, making it suitable for real-time applications due to its simplicity and lower cost compared to existing models. The study's authors [13] proposed five different approaches for classifying gastrointestinal (GI) tract disorders into various categories. Three methods employed Convolutional Neural Networks (CNNs) with transfer learning, and two focused on global feature extraction. The authors introduced a technique integrating transfer learning with pre-trained CNN and global features. They found that combining two neural networks—ResNet-152 and DenseNet-161—and an additional Multi-layer Perceptron (MLP) yielded the best results. This integrated approach achieved impressive performance, with an accuracy of 95.80%, precision of 95.87%, and an F1-score of 95.80% on the validation and test datasets provided by the task organizers. In their recent publication [14], the authors introduced an integrated methodology combining PCA (Principal Component Analysis), a modified deep residual network, and AdaBoost ensemble learning to differentiate between endoscopic images with and without polyps. The deep residual network architecture was enhanced by modifying the ResNet-50 model to reduce computation time. Three datasets were used to train the proposed model: ETIS-LaribPolypDB, Kvasir, and CVC-ClinicDB.

The model achieved impressive results, with an MCC of 98.19%, accuracy of 99.1%, sensitivity of 98.82%, precision of 99.37%, and specificity of approximately 99.38%. The authors of [15] developed a polyp detection method based on ResNet50. The data used for the study was gathered from a public dataset consisting of 1,000 images, with 500 images containing polyps and 500 normal (non-polyp) images. The model was evaluated using 10-fold cross-validation, where the 1,000 images were divided into 10 equal sets. Nine sets were used for training, while the remaining one was used for testing. The proposed model achieved an accuracy of 96%. The authors of [16] used approximately 8,641 images from 2,000 patients to train a D-CNN model. Twenty colonoscopy recordings, totaling around five hours, were processed to evaluate the proposed model, and nine archived videos were used for testing. By comparing the CNN model's output to the results of qualified colonoscopies, the study demonstrated the effectiveness of using the CNN model as a benchmark. The model achieved an accuracy of approximately 96.4% with a 7% false-positive rate.

1.4. Deep Learning

Deep Learning (DL) is a subset of machine learning (ML) that focuses on learning representations of data. Unlike traditional machine learning algorithms, which rely on explicitly defined features, deep learning algorithms automatically learn features from the data [17], [18].

In contrast to hand-crafted features, where humans manually design information, deep learning enables models to learn from raw data without needing predefined features. This approach, known as learning to act, maps actions to their effects, eliminating the need for human-designed features and often delivering better results. Moreover, deep learning allows faster adaptation to new tasks, requiring less time and human effort [18]. However, finding an appropriate representation for a specific problem can be time-consuming, as deep learning models may need to learn from multiple levels or types of data features. Despite this challenge, DL provides outstanding robustness and flexibility, effectively tackling complex real-world problems.

2. Dataset and Methodology

2.1. Dataset

The Kvasir dataset, consisting of images captured from within the gastrointestinal (GI) system, was used for analysis. These images are classified and annotated by physicians and experienced endoscopists. Kvasir is a crucial resource for assessing computer-aided detection of individual and complex gastrointestinal disorders. The dataset includes authenticated and annotated endoscopic images of the GI tract, carefully selected by certified endoscopists.

In contrast, released publicly in the autumn of 2017 as part of the MediaEval Medical Multimedia Challenge, this dataset supports scientific objectives for benchmarking purposes. Both deep learning and machine learning models can leverage the extensive collection of images. The pathological findings in the dataset include polyps, ulcerative colitis, and esophagitis. Image resolutions range from 720x576 to 1920x1072 pixels, and the dataset contains 1,500 images labelled as normal and 1,500 images showing colorectal polyps. Figure 1. shows the dataset used in the paper.

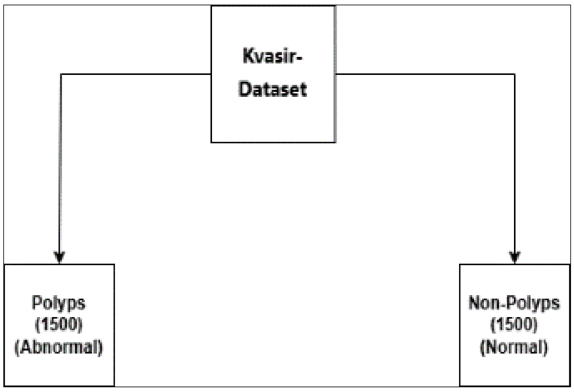


Figure 1 Description of the dataset used in the paper.

2.2. Methodology

This paper's methodology used CNN models, which include R-CNN, CNN, VGG 16, VGG 19, Resent 50, Resent 101, InceptionV3, and Xception. Using hyperparameter as shown in table 1. The data was split into 75 % and 25 % for training and testing using batch size 256, number of epochs 10, and Learning Rate 0.001. and using vertical and horizontal augmentation for all images. Figure 2 shows the paper's methodology.

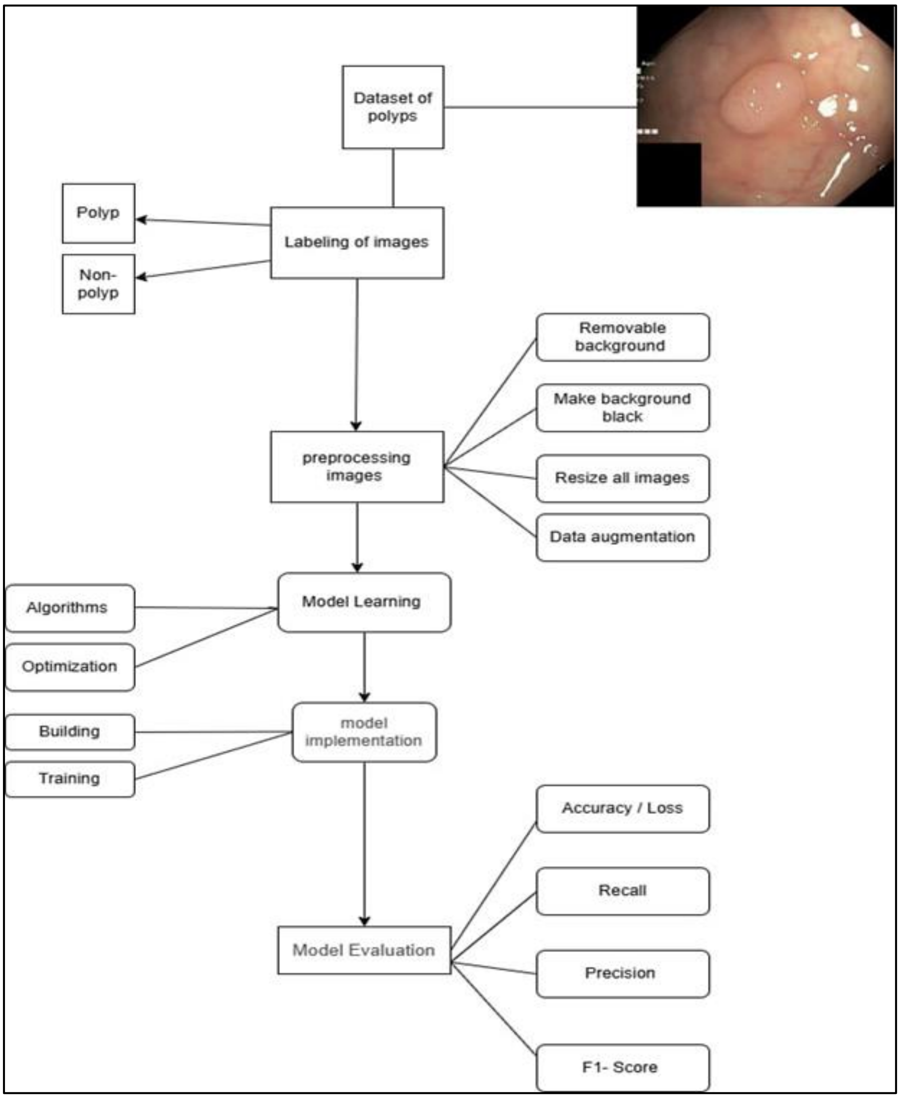


Figure 2 Research paper methodology.

Table 1 The best hyperparameters for each model.

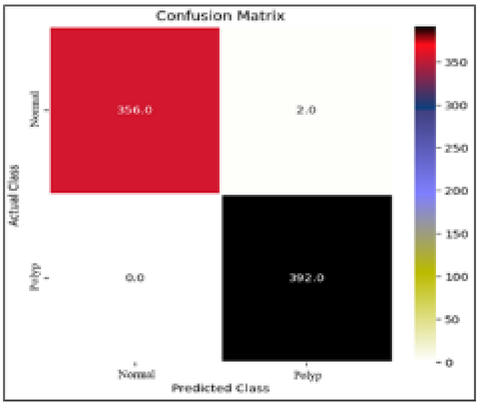
Model	Batch Size	Number of epochs	Learning Rate (LR)
VGG 16	256	10	0.001
VGG 19	256	10	0.001
ResNet 50	256	10	0.001
ResNet 101	256	10	0.001
InceptionV3	256	10	0.001
Xception	256	10	0.001
R-CNN	256	10	0.001
CNN	256	10	0.001

3. Result and Discussion

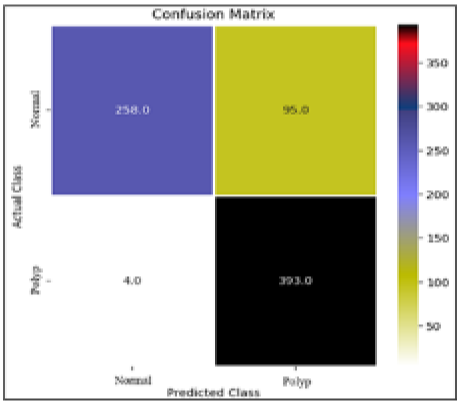
After building and processing the models, the results are presented in Table 2. The best model can be selected based on performance metrics. Also, Figure 3 shows a confusion matrix for results.

Table 2 Summary of algorithm results on prediction class (normal vs polyp).

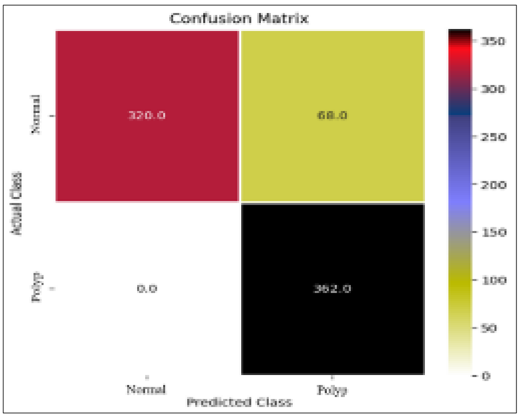
Algorithm	# Epoch	Batch	Accuracy	Learning Rate	Recall	Precision	F1-score	Training Time(ms)
VGG19	10	256	0.9933	0.001	0.99	0.99	0.99	32977
VGG16	10	256	0.9942	0.001	0.99	0.99	0.99	7904
Resnet101	10	256	0.9996	0.001	1	1	1	8219
Resnet50	10	256	1	0.001	0.91	0.92	0.91	4459
Xception	10	256	0.9924	0.001	0.86	0.90	0.86	4128
R-CNN	10	256	0.9991	0.001	0.98	0.98	0.98	1889
CNN	10	256	1	0.001	1	1	1	1302
Inception	10	256	0.9702	0.001	0.73	0.83	0.71	2547



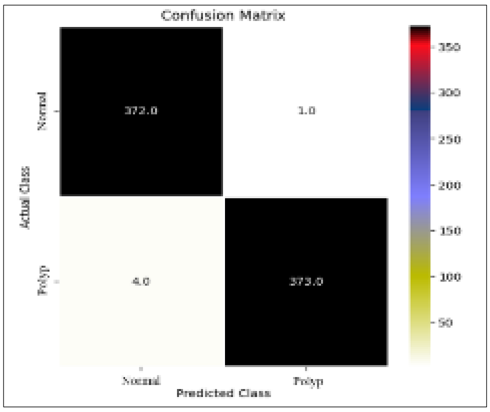
ResNet 101



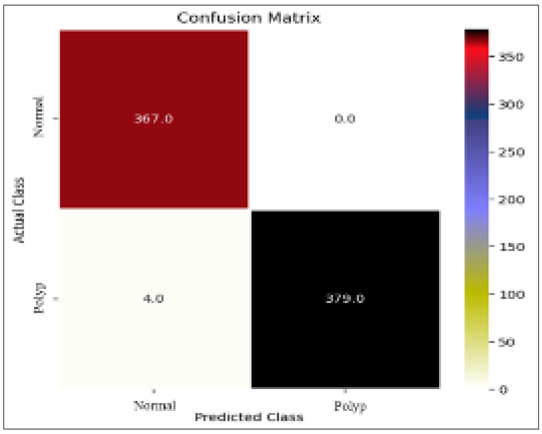
Xception



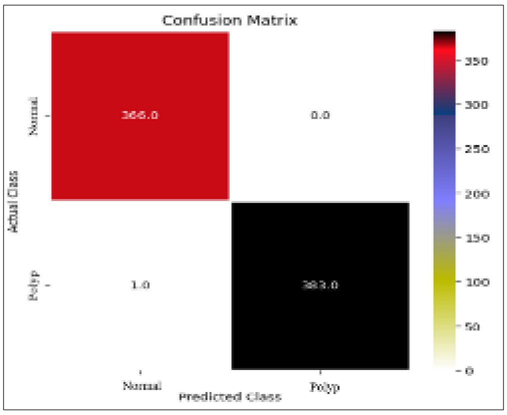
ResNet 50



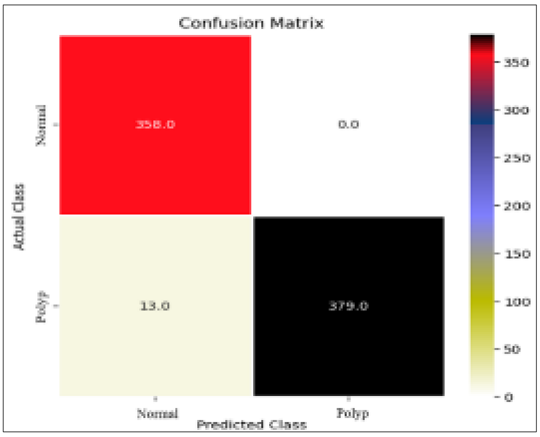
VGG 19



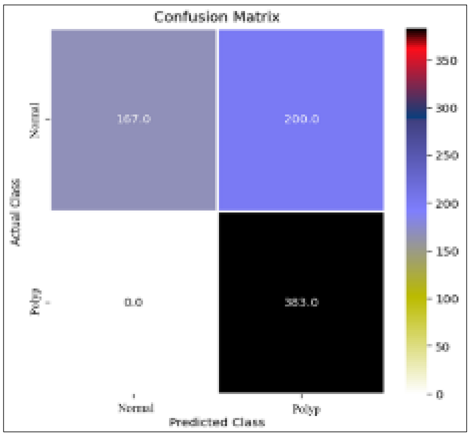
VGG 16



CNN



R-CNN



Inceptionv3

Figure 3 Confusion matrix for (normal & polyp) classification**.**

The results presented in Table 2 highlight the performance of various deep-learning models in classifying colorectal polyp images. Among the tested architectures, ResNet50 achieved the highest accuracy (100%), closely followed by ResNet101 (99.96%) and CNN (100%). However, accuracy alone cannot determine the best model, as computational efficiency is crucial in the real world. Training time analysis reveals that the CNN model has the lowest training time (1,302 MS), significantly outperforming ResNet101 (8,219 MS) and ResNet50 (4,459 MS). This makes CNN the most efficient model in terms of computational cost. Furthermore, CNN maintains perfect recall, precision, and F1-score, making it a highly reliable choice for colorectal poly detection.

Other architectures, such as VGG19 and VGG16, demonstrated competitive accuracy (99.33% and 99.42%, respectively) but required substantially higher training times, particularly VGG19 (32,977 MS). Similarly, Xception and R-CNN performed well but did not surpass CNN's efficiency. The Inception model exhibited the weakest performance, with an accuracy of 97.02% and a relatively lower F1-score (0.71), suggesting it may not be the best choice for this task.

4. Conclusion

This study evaluates multiple CNN-based architectures for colorectal polyp detection using a public dataset of 3,000 images. While ResNet50 and ResNet101 achieve the highest accuracy, the CNN model is optimal considering classification performance and computational efficiency. With its perfect accuracy, recall, precision, and shortest training time, the CNN model is the best candidate for real-time polyp detection applications. These findings emphasize the potential of deep learning in medical image analysis, particularly for aiding in the early diagnosis of colorectal cancer. Future work can explore further optimizations, including transfer learning techniques and real-time deployment in clinical settings to enhance the model's applicability in endoscopic examinations.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest.

References

[1] Cheng LK, O'Grady G, Du P, Egbuji JU, Windsor JA, Pullan AJ. Gastrointestinal system. IEEE Wiley Interdiscip Rev Syst Biol Med. 2010;2(1):65–79.

[2] Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. IEEE Int J Cancer. 2019;144(8):1941–53.

[3] Abraham A, Simpson AL, Raju A, Foo J, Tan S, Goh S, Ngo R, Lim A. Comparative analysis of machine learning models for image detection of colonic polyps vs. resected polyps. IEEE J Imaging. 2023;9(10):215.

[4] Huang Y, Liu J, Wang L, Xu T, Zhao X, Chen K. Establishment of clinical predictive model based on the study of influence factors in patients with colorectal polyps. IEEE Front Surg. 2023;10:1077175.

[5] Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with colorectal polyps. Practice Parameters Committee of the American College of Gastroenterology. IEEE Am J Gastroenterol. 2000;95(11):3053–63.

[6] Hao Y, Wang Y, Qi M, He X, Zhu Y, Hong J. Risk factors for recurrent colorectal polyps. IEEE Gut Liver. 2020;14(4):399–411.

[7] Elkarazle K, Raman V, Then P, Chua C. Improved colorectal polyp segmentation using enhanced MA-NET and modified Mix-ViT Transformer. IEEE Access. 2023; 11:69295–309.

[8] Kuiper T, Marsman WA, Boot H, Dekker E, Bruno MJ, Fockens P. Accuracy for optical diagnosis of small colorectal polyps in nonacademic settings. IEEE Clin Gastroenterol Hepatol. 2012;10(9):1016–e79.

[9] Sadagopan R, Ravi S, Adithya SV, Vivekanandhan S. PolyEffNetV1: A CNN-based colorectal polyp detection in colonoscopy images. IEEE Proc Inst Mech Eng Part H J Eng Med. 2023;237(3):406–18.

[10] Ribeiro J, Nóbrega S, Cunha A. Polyps detection in colonoscopies. IEEE Procedia Comput Sci. 2022; 196:477–84.

[11] Lo CM, Yeh YH, Tang JH, Chang CC, Yeh HJ. Rapid polyp classification in colonoscopy using textural and convolutional features. IEEE Healthcare (Basel). 2022;10(8):1494.

[12] Nisha JS, Gopi P, Palanisamy P. Automated colorectal polyp detection based on image enhancement and dual-path CNN architecture. IEEE Biomed Signal Process Control. 2022.

[13] Aish MA, Abu-Naser SS, Abu-Jamie TN. Classification of pepper using deep learning. In: IEEE INOCON Conf Proc. 2022.

[14] Liew WS, Tang TB, Lin CH, Lu CK. Automatic colonic polyp detection using integration of modified deep residual convolutional neural network and ensemble learning approaches. IEEE Comput Methods Programs Biomed. 2021; 206:106114.

[15] Eun SJ. Development of polyp detection technology by analyzing deep-learning. IEEE J Next-Gener Converg Inf Serv Technol. 2020;9(2):139–47.

[16] Urban G, Tripathi P, Alkhatib A, Lindsey M, Bharadwaj S, Reiter A, Ou A, George R, Singh R, Wang F. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. IEEE Gastroenterology. 2018;155(4):1069–78. e8.

[17] Alkhatib AJ, Alkhattib K, Alawnah AB, Alzoubi A, Alaiad A, Abu Aqoulah A, Alharoun M. Diagnosing brain tumors from MRI images through a multi-fused CNN with auxiliary layers. IEEE Sustainable Mach Intell J. 2024;6(2):1–10.

[18] Alzoubi A, Alkhattib K, Alawnah AB, Abu Aqoulah A, Alaiad A, Alharoun M. Detection of depression from Arabic tweets using machine learning. IEEE Sustainable Mach Intell J. 2024;6(3):1–7.

[19] Alzoubi A, Alaiad A, Alkhattib K, Alkhatib AJ, Abu Aqoulah A, Alharoun M, Alawnah AB. The impact of COVID-19 in depression in Arab Twitter users. Inf Sci Appl. 2024; 2:19–32.