International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Performance Analysis of Transfer Learning Framework for the Detection of Polyps in Colorectal Cancer

[1]Yogesh Chaudhari, [2]Ashish Jani, [3]Harshal A. Sanghvi, [4]Abhijit S. Pandya, [5]Vipin Gupta, [6]Darshee Baxi

Submitted: 23/12/2023 **Revised:** 29/01/2024 **Accepted:** 07/02/2024

Abstract: Colorectal cancer (CRC) begins in the colon or rectum, gastrointestinal tract organs. It is a common cancer that causes many cancer deaths worldwide. CRC usually starts with a polyp, a benign growth that can become cancerous. CRC prevention, treatment, and control require early detection and treatment. In this study, we reviewed various, pertinent research based on CRC diagnostic techniques, colonoscopy, and the use of AI screening. We performed various quantitative and qualitative comparative analyses of diagnostic techniques based on numerous features. Colonoscopy and sigmoidoscopy allow doctors to examine the colon and rectum for abnormalities. Deep learning (DL) techniques in medical imaging and Artificial Intelligence (AI) have improved CRC diagnosis, particularly polyp detection. We discussed the present and possible use of AI, DL in CRC diagnosis. A sigmoidoscopy, a minimally invasive procedure, shows the potential in terms of reducing the number of incidences and mortality. Colonoscopy was the most invasive technique and possesses the risk of morbidity. The Markov model demonstrated that cost per life can be saved for a colonoscopy performed once in 10 years. Thus, colonoscopy certainly proves to be a golden standard with highest sensitivity with the capability of biopsy during diagnosis. The proposed pre-trained VGG19 model confirmed 97% accuracy in polyp detection when applied with the approach of Transfer Learning (TL). The model is not overfitting and is proven to be more accurate than the recommended Adenoma Detection Rate (ADR).

Keywords: *Deep Learning, Medical Imaging, Transfer Learning, Computer Aided Diagnosis, Colorectal Cancer*

1. Introduction

CRC is the third most frequently detected and the second highest contributor to cancer-related fatalities on a global scale. As a result, it presents a significant challenge to the health overall population of the world. In most cases, this cancerous growth originates from the inner epithelial layer of the colon or rectum, and it frequently begins as polyps, which are benign at initial stage. To emphasize the critical role that early detection plays in the prevention, treatment, and effective management of CRC, it is important to note that the transformation from benign polyps to cancerous growth occurs gradually over a period of fifteen to twenty years. Because CRC is so prevalent all over the world, its significance has been brought to light, which has prompted the medical community to investigate more advanced diagnostic techniques[1].

When it comes to identifying abnormalities in the colon and rectum, conventional diagnostic methods, such as colonoscopy and sigmoidoscopy, have been extremely helpful in the past. Nevertheless, these procedures have inherent limitations and risks associated with them, which highlights the need for diagnostic methods that are more effective and less invasive. Over the past few years, the utilization of AI and DL approaches in the field of medical image processing has emerged as a potentially fruitful strategy for enhancing the diagnosis of CRC.

The development of CRC occurs gradually over a period of 15-20 years, beginning with harmless polyps and progressing to malignancy. The extended timeline underscores the importance of early diagnosis, as it offers a vital opportunity for preventive actions, curative treatments, and efficient disease control. Colonoscopy and sigmoidoscopy are traditional diagnostic techniques that have played a crucial role in identifying and diagnosing colorectal abnormalities. Although these techniques have demonstrated efficacy, they are not without limitations and inherent risks. Colonoscopy, for example, is a procedure that invades the body and has the potential to cause illness, so it is important to investigate other diagnostic methods. [2].

[1]Department of Computer Science and Engineering Navrachana University Vadodara Gujarat, India
yogeshc@nuv.ac.in
[2]Department of Computer Science and Engineering Navrachana University Vadodara Gujarat, India
ashish.jani@nuv.ac.in
[3]Department of Computer Science and Engineering Florida Atlantic University Boca Raton, USA
hsanghvi2020@fau.edu
[4]Department of Computer Science and Engineering Florida Atlantic University Boca Raton, USA
pandya@fau.edu
[5]Specialty Gastro Center and Broward Health Light House Point and W. Sample Road, FL, USA
vipinbusy@yahoo.com
[6]Department of Biomedical and Life Sciences Navrachana University Vadodara Gujarat, India
darshee.b@nuv.ac.in

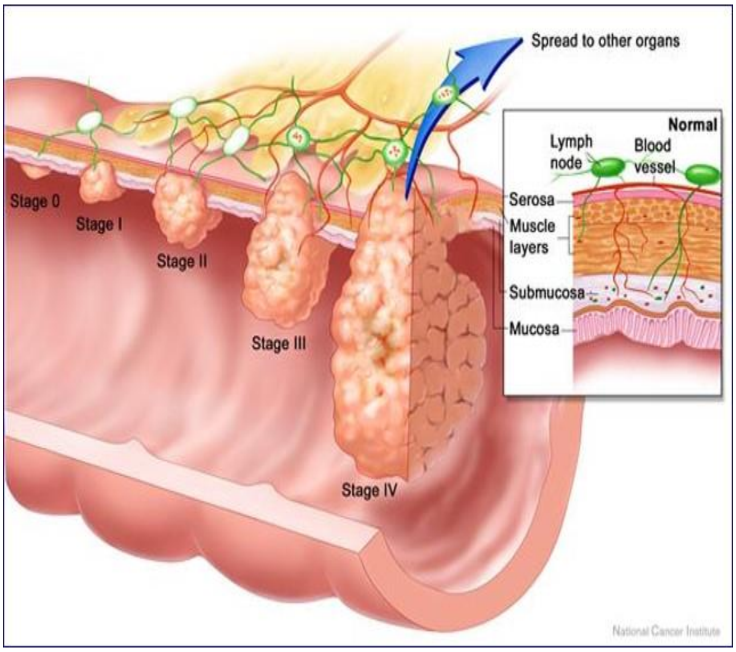

Fig 1: Various stage of CRC (Source-colorectal cancer alliance)

A revolutionary change has occurred in the field of CRC diagnosis as a result of the development of AI, which has taken into account the requirement for diagnostic procedures that are both more efficient and less invasive. AI, specifically deep learning, has demonstrated potential in the field of medical imaging, presenting new opportunities to improve diagnostic precision and effectiveness [3].

The present work aims to overcome the constraints of current diagnostic methods by conducting a thorough examination and comparison of conventional techniques used for diagnosing CRC. The objective is to comprehensively comprehend their capabilities, limitations, and areas requiring enhancement, thereby establishing the foundation for the subsequent investigation of AI-driven and Computer Aided Diagnosis (CAD) solutions. A comprehensive examination is undertaken to investigate the role of AI, specifically in the field of medical imaging, for the diagnosis of CRC. AI has the capacity to transform the field by offering more precise and efficient diagnostic capabilities, which can have a substantial impact on patient outcomes. This study presents a new framework that utilizes the VGG19 model to classify polyps, making it a valuable addition to the current advancements in this field. The VGG19 model is a Convolutional Neural Network (CNN) that has proven to be highly effective in tasks involving the classification of images. Within the context of CRC, this framework is designed to attain an impressive 97% accuracy in identifying polyps, surpassing the recommended ADR. The model's non-overfitting nature enhances its credibility.

The anticipated effect of this proposed framework is significant. Enhancing the precision and effectiveness of CRC diagnosis greatly aids in the early detection, treatment, and successful control of the disease. The decrease in mortality rates associated with CRC is a concrete result that demonstrates the potential of AI solutions in revolutionizing healthcare practices and improving patient care. This study represents a leading position in the transition towards more sophisticated and patient-oriented diagnostic approaches in the field of CRC.

What is Transfer Learning?

TL is an approach in which a pre-existing model created for a specific task is utilized as the initial foundation for a model designed for a different task. TL utilizes the acquired knowledge from solving one problem to enhance the performance on a related yet distinct problem, rather than starting the model training process from the beginning.

Let θ_{source}= "parameters of the pre-trained model on the source task", θ_{target}= "parameters of the model for the target task", $\mathcal{L}_{source}$= "Loss on the source task", $\mathcal{L}_{target}$= "Loss on the target task". The idea is to minimize the loss on the target task by θ_{target} while utilizing the knowledge gained from the pre-trained model on the source task is represented by eq.1

$$min\theta_{target}\left(\mathcal{L}_{target}\left(\theta_{target}\right)\right) + \lambda.\mathcal{L}_{source}(\theta_{source})) \quad \text{……………………….. (eq. 1)}$$

where, λ= "hyperparameter that controls the influence of the source task on the target task". The model is fine-tuned on the target task with a smaller learning rate to prevent large updates that may omit useful features learned during pre-training. This is implemented using gradient descent which is calculated wrt θ_{target} and the weights are updated as eq.2:

$$\theta_{target} \leftarrow \theta_{target} - \alpha.\nabla\theta_{target}(\mathcal{L}_{target}(\theta_{target}) + \lambda.\mathcal{L}_{source}(\theta_{source})) \;\ldots\ldots\text{(eq. 2)}$$

where, α= "learning rate".

The process helps leverage knowledge gained from a related task even when there might be limited labeled data is available for the target task. TL is essential in improving the efficiency and effectiveness of DL models when applied to CRC detection. In order to effectively train a Deep Neural Network (DNN) for medical image analysis, specifically for tasks like detecting polyps in colorectal images, a substantial quantity of annotated data is typically necessary. Nevertheless, the process of gathering such comprehensive datasets can poses difficulties like privacy apprehensions, ethical deliberations, and limitations in resources.

TL overcomes this constraint by enabling the utilization of pre-trained models, typically trained on extensive datasets for generic image recognition tasks, to be adjusted or customized for the particular task of CRC detection. The initial layers of these pre-trained models have already acquired the ability to identify fundamental characteristics such as edges, textures, and shapes, which are pertinent in diverse image recognition assignments. TL in CRC detection allows the model to utilize the knowledge acquired from analyzing a variety of images to enhance its understanding and classification of features that are unique to colorectal images. TL enhances the accuracy, generalization, and robustness of model.

The Role of DL in Polyp Detection

DL is crucial in the field of medical imaging, specifically in identifying polyps related to CRC. Within the realm of CRC, polyps serve as initial precursors that have the potential to develop into malignancies, highlighting the critical significance of precise and prompt identification. Convolutional neural networks (CNNs), a type of DL technique, have shown impressive effectiveness in automatically detecting polyps from medical images, such as those obtained from endoscopic and colonoscopy examinations.

DL is applied in polyp detection by training models, such as VGG19, within the proposed framework. This training is done using extensive datasets of annotated medical images. These models acquire complex patterns and distinctive characteristics that are characteristic of both typical and atypical tissue, allowing them to accurately differentiate polyps from healthy tissue.

DL models possess the ability to engage in ongoing learning and adjustment, enabling them to enhance their performance as they encounter a wider range of datasets. DL plays a transformative role in polyp detection within the context of CRC, providing a more efficient and accurate alternative to conventional diagnostic methods. DL utilizes neural networks to improve the early detection of CRC, leading to prompt intervention and better patient outcomes.

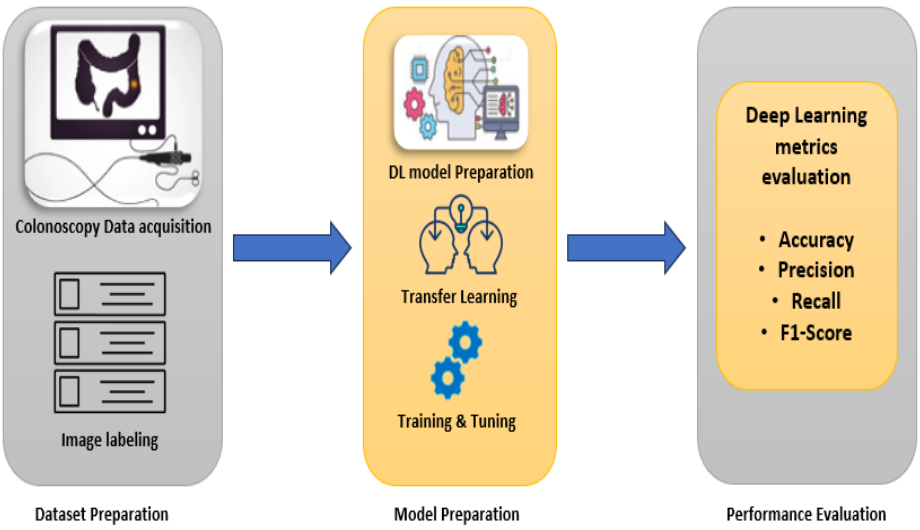

Fig 2: Role of Deep learning in CRC

A number of studies have investigated the use of TL for polyp detection. In one study, researchers trained a CNN on the ImageNet dataset, which contains over 1 million images from 1,000 different classes. The CNN was then fine-tuned on a dataset of "colonoscopy images", and it achieved an accuracy of 96.41% in detecting polyps. Another study used a pre-trained CNN called ResNet50 to detect polyps in "colonoscopy images". ResNet50 was first trained on the ImageNet dataset, and then it was fine-tuned on a dataset of "colonoscopy images". The study achieved an accuracy of 93.8% in detecting polyps. In addition to CNNs, other types of AI models have also been used for polyp detection using transfer learning. For example, one study used a Recurrent Neural Network (RNN) to detect polyps in colonoscopy videos. The RNN was first trained on a dataset of natural language processing (NLP) tasks, and then it was fine-tuned on a dataset of colonoscopy videos. The study achieved an accuracy of 92.7% in detecting polyps.

TL has a number of advantages for polyp detection in colonoscopy. First, it allows AI models to be trained on large datasets of general images, which can lead to better performance than training on smaller datasets of "colonoscopy images" alone. Second, TL is relatively fast and efficient, as it does not require the AI model to be trained from scratch on a dataset of "colonoscopy images".

TL offers several advantages in the context of polyp detection:

- Improved Performance: By leveraging "pre-trained models", TL can enhance the performance of polyp detection algorithms. The pre-trained models have already learned generic features from large-scale datasets, which can be highly relevant for polyp detection.
- Reduced Training Time: Training DL models from scratch requires a significant amount of time and computational resources. TL reduces the training time by starting with a "pre-trained model" and "fine-tuning" it on a smaller dataset of polyp images or videos.
- Enhanced Generalization: TL helps the model generalize well to new, unseen data. The "pre-trained models" have learned generic features that can be utilized for a wide variety of different activities, including polyp detection.
- Overcoming Data Limitations: Obtaining a large annotated dataset of polyp images or videos can be challenging. TL allows the utilization of pre-existing large-scale datasets, making it possible to train accurate polyp detection models even with limited data.

2. Literature Review

Worldwide 9,30,000 fatalities and 19,00,000 cases of CRC were reported in 2020 [4]. In Europe, CRC ranks as the second most common cancerous disease in incidences and deaths both with 5,19,820 new cases and 2,44,824 deaths in 2020 [4]. These numbers indicate that CRC is associated with higher rates of morbidity and mortality among different types of oncological diseases. The most well-known types of polyps are those in the colon, while they can also occur in the "ear canal, cervix, stomach, nose, uterus, bladder, and throat". A polyp is a little growth of extra tissue that forms on the colon's lining. They are common, even though they are unpleasant. Around 25% of population, including men and women, 50 years or older, have colon and rectal polyps. [5] These polyps are categorized into two types, based on the shape as they mature.

Boonsim 2023 presents an optimized DL technique using the InceptionResnetV2 model and Faster R-CNN framework, achieving high precision, recall, and F1-Measure in polyp detection [6]. Tzavara 2021 explores "transfer learning in polyp and endoscopic tool segmentation", showing improved performance with pre-training of models [7]. Luca 2019 provides an overview of recent contributions in automatic polyp detection, including the application of DL [8]. Tang 2021 focuses on computer-aided colon polyp detection using transfer learning, demonstrating the potential for real-time detection and classification with improved accuracy [9]. These findings collectively highlight the effectiveness of TL in enhancing polyp detection during colonoscopy.

TL has gained attention in polyp research as a way to address the challenge of limited data for training large-scale segmentation networks. It allows for the use of pre-trained models from a source task, reducing the need for extensive data collection and labeling. The Poly-SAM model, a finetuned version of the Segment Anything Model (SAM), achieved state-of-the-art performance on multiple datasets, with dice scores above 88% [10]. TL also enables fast prototyping of machine learning models by leveraging pre-trained models, which is particularly beneficial when training on millions of images is time-consuming and resource-intensive [11]. Overall, TL has demonstrated its potential in improving polyp segmentation and accelerating research in this field [12-13].

The purpose of this study is to develop a software program that will assist medical professionals in the fight against colon cancer, both during the detection phase and after the surgery has been performed [14-15]. TL has been employed to create algorithms for the automated identification and categorization of colorectal polyps in "colonoscopy images". Multiple investigations have been undertaken regarding the identification and categorization of colorectal polyps through the utilization of DL techniques. These studies have demonstrated encouraging outcomes in relation to precision and responsiveness. The "Z-line, pylorus, and cecum" serve as anatomical reference points to determine the extent of the colonoscopy procedure F6. The proposed model demonstrates a 92% accuracy in detecting and categorizing colorectal polyps from colonoscopy images [16]. A novel method for classifying polyp images, "Network-in-Network" (NIN), has been developed utilizing transfer learning. [17]. The efficacy of utilizing DL for real-time identification of colon polyps during colonoscopy has been confirmed through the validation of four separate datasets. The text is accompanied by a reference number [18]. A novel approach to colorectal polyp classification using deep ensemble learning has been introduced that incorporates optimized network parameters and includes crucial information on hyperparametric settings for model optimization [19].

A recent research study by Noppakun Boonsim and Saranya Kanjaruek from Khon Kaen University in Thailand focused on optimizing the parameters for polyp detection using a DL technique called InceptionResnetV2. The researchers trained the model on a dataset of polyp and non-polyp images and used the Faster R-CNN framework for precise polyp detection. The proposed method achieved remarkable results, with a precision of 92.9%, recall of 82.3%, F1-Measure of 87.3%, and F2-Measure of 54.6% on a public ETIS-LARIB dataset. This optimized TL approach significantly reduced the chances of missing polyps during clinical inspections and improved the detection of multiple polyps in colon images. [6]

Another research study by Ji Young Lee et al, focused on the real-time detection of colon polyps using deep learning. The researchers developed a deep-learning algorithm based on YOLOv2 and validated it using four independent datasets. The algorithm exhibited high sensitivity (96.7%) and accuracy in polyp detection. All 38 polyps detected by the endoscopists were identified by the algorithm in addition to discovery of seven polyps that were missed by the human observers. The algorithm's performance was further tested on 15 unaltered colonoscopy videos, where it achieved a per-image sensitivity of 89.3% and a low false positive rate of 8.3%. [20] The polyps are also categorized based on their carcinogenic risk, as outlined in Table 1.

Table 1 Polyp types, occurrences, and associated cancer risk [21]

Sr. No	Type of polyp	How common	Cancer risk
1.	Inflammatory	Typically seen in people with “ulcerative colitis” or “Crohn's disease”, “inflammatory bowel illnesses”.	Low; most growths are benign
2.	Hamartomatous	Discovered in individuals who have polyposis syndromes, such Peutz-Jaeger, Cowden, or Juvinille Polyposis.	Commonly non-cancerous
3.	Hyperplastic	Usually tiny and typically situated at the end of the colon and the rectum.	Considered as being lower risk
4.	Adenomatous (tubular adenoma)	These type of colon polyps are most common and make almost 70% of polyps.	The majority of these polyps do not progress to cancer, although larger polyps possess a higher threat.
5.	Villous or tubulovillous adenoma	Make about approximately 15% of polyps.	The majority of these polyps do not progress to cancer, although larger polyps possess a higher threat.
6.	Serrated Adenoma	Make about 10 to 15% of all polyps.	These polyps result into 20–30% of colon cancer cases.
7.	Adenocarcenoma	Make about 2% of all polyps.	These polyps are 100% cancerous.

Present Diagnostic Techniques:

Multiple diagnostic techniques are used by physicians to find and to diagnose cancer. These tests are also useful in analyzing the growth and spread of cancer that indicate the stage and metastasis, respectively. These tests further help in planning the course of treatment and types of medication to be prescribed to the patient. Currently, a biopsy is the only reliable method for confirmation tests with many types of cancers. In cases where a biopsy is not possible, doctors may recommend other tests for the accurate diagnosis. Various diagnostic techniques are used to identify CRC and abnormalities in the colon. A few of these important tests and techniques are considered for study

- Colonoscopy [22-24].
- Computed tomographic colonography (CTC) [25-29].
- Double contrast barium enema (DCBE) [30-31].
- Flexible sigmoidoscopy (FS) [32-34].
- “Fecal immunochemical test” (FIT) [35-40].

- "High-sensitivity guaiac-based fecal occult blood test" (HSgFOBT) [41-42].
- "Multi-target stool DNA" (MT-sDNA) [43].

Present use of AI in Diagnostic techniques

AI is revolutionizing the entire medical field. Gastroenterology practice is no exception. The importance and utilization are increasing for use of AI in gastroenterology with acceptance from practitioners and government bodies. AI has proven its role in early detection and prevention of CRC, thereby, ultimately reducing the time and cost of treatment and increasing the year expectancy of the patient. In this section, the use of AI is explained various diagnostic techniques [44-47] like colonoscopy, CTC [42] and FIT [43].

Table 2 Summary of diagnostic techniques and corresponding application of AI

Sr. No	Diagnostic Technique	Present scope and objectives based on AI
1	Colonoscopy	To detect polyps, to classify polyps, to characterize the polyp. To reduce unnecessary removal of non-neoplastic polyps To detect deficient coverage in colonoscopies [24]
2	CTC	To minimize the patient's exposure to radiation through optimization. Automated segmentation of organs to identify and describe pathology. To lessen the likelihood of performance and interpretation errors [29].
3	FIT	For prioritization of colonoscopy [40]. To automate the interpretation of results. To automate the wider BCSP – to deliver results and forwarding of patient information where required.
4	HSgFOBT, DCBE, FS, MT-sDNA	Lack of research articles that have documented the utilization of AI.

3. Methodology

VGG-19 has demonstrated impressive results in diverse image classification tasks and is widely recognized as a standard architecture in the field of computer vision. Despite its computational expense due to a large number of parameters, VGG-19 has significantly influenced the advancement of DNN architectures for image recognition tasks. It is pre-trained on ImageNet dataset with more than a thousand classes and "CIFAR 100" dataset [48]. VGG19 has three additional fully connected layers compared to VGG16 and plays a significant role in extracting low-level spatial information [49]. The accuracy of VGG19 in disease classification is reported to be 94% [50]. In the context of crime prediction, VGG19 is used for feature extraction and object comparison, achieving 81% accuracy [51]. For rice plant disease detection, a Deep CNN TL method based on VGG19 is proposed, achieving an accuracy of 97% [52].

Dataset:

The open-source dataset, CVC-Clinic DB, has been investigated for the experimental setup [53]. Annotation was used to produce the positive and negative images of polyps. For training, testing, and validation purposes, the cropped photographs were organized in separate sets. The images are divided into positive and negative subgroups within each of these sets.

Table 3 Summary of Dataset used in experiment.

Sr. No.	Particular	Details
1	True colour (RGB) images	612 frames extracted from colonoscopy videos
2	Polyp mask (Monochrome Images)	612 images identical to the RGB images in dimensions. Black and white represents a polyp positive and negative respectively.
3	Sequences	29
4	Dimensions	384 by 288
5	True Image size (in kB)	324 kB

6	Polyp mask size (in kB)	108 kB
7	Size of dataset	258 MB

The dataset used in this study was "CVC-ClinicDB", created for the Endoscopic Vision Challenge and contains frames that were taken from videos of colonoscopies. Two kinds of images were included in the "CVC-ClinicDB" database: "colored Red Green Blue" (RGB) images, and corresponding "monochrome polyp masks". Table 3 contains a description of the dataset. [54] Figure 3 has 3 randomly selected images from the dataset included, as well as each image's associated monochrome mask.

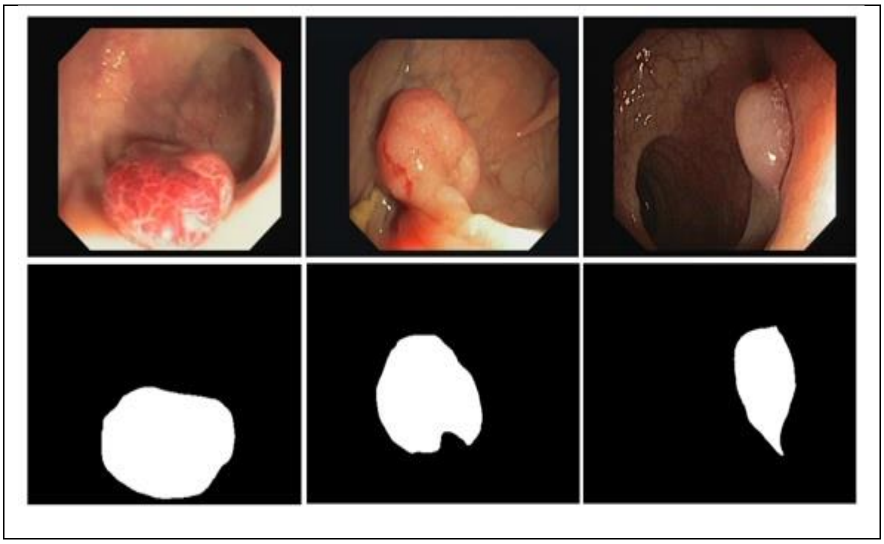

Fig 3: RGB polyp frames are in the top row and corresponding monochrome polyp masks are in the bottom row.

Algorithm for Proposed methodology

ALGORITHM 1: POLYP DETECTION USING VGG19

Data Preprocessing:

- *Resize RGB images* → 96x96.
- Apply image augmentation techniques (blur, flip, rotation, brightness) ←generate 10,010 images.
- *Split the data* ←training, validation, and testing sets.
- Create positive and negative subsets within each set.

Feature Extraction:

- Load pre-trained ***VGG19*** model.
- Freeze ← convolutional layers.
- Extract features from the last convolutional layer of the VGG19 model

Fine-tuning:

- Add a new fully-connected layer with **1024** neurons followed by a "***ReLU"*** activation function.
- Add another fully-connected layer with 2 neurons followed by a "***Softmax"*** activation function.
- Train the newly added layers on the training set using the "***Adam optimizer"*** and categorical cross-entropy loss function.

Evaluation:

- Evaluate the model using metrics → ***Accuracy, Precision, Recall, F1-Score***

Prediction:

Use the trained model to predict the class (polyp positive or negative) for each image in the testing set.

Analysis:

Analyze ← results and identify areas for improvement.

Visualize the results using techniques like heatmaps to understand which parts of the image are most. influential in the prediction

Convolutional Neural Network Architecture: VGG-19

The architecture of VGG-19is a specific type of CNN that has been specifically created for the purpose of classifying images. The model is a member of the Visual Geometry Group (VGG) family and is known for its straightforward and consistent architecture. VGG-19 model is developed by the Visual Geometry Group at the University of Oxford that consists of 19 processing layers, which comprise both convolutional and fully linked layers.

Overview of Architecture:

- **Input Layer:** The input layer of the network is designed to receive images with dimensions of 224×224×3, where 3 represents the RGB color channels.
- **Convolutional Blocks:** Comprises four sets of convolutional layers, each thereafter accompanied by a max-pooling layer.
  - Convolutional layers employ 3x3 filters with a stride of 1.
  - The max-pooling layers utilize a window size of 2x2 and a stride of 2.
- **Fully connected layers:** Following the convolutional blocks, there are three fully linked layers, each consisting of 4096 neurons.
  - The output of each fully linked layer is subjected to Rectified Linear Unit (ReLU) activation functions.
- **Output Layer:** The last layer consists of a fully connected neuron that use a sigmoid activation function for binary classification or softmax for multi-class classification.

The following eq.3 to eq.7 represents the operations within one convolutional block depicts the sequential transformations that allows VGG-19 to capture features for image classification (Polyp) task.

$$Z_1 = Conv(X, W_1) + b_1 \quad \text{……… (eq. 3)}$$

$$A_1 = ReLU(Z_1) \quad \text{………………..(eq. 4)}$$

$$Z_2 = Conv(A_1, W_2) + b_2 \quad \text{……..(eq. 5)}$$

$$A_2 = ReLU(Z_2) \quad \text{…………………….(eq. 6)}$$

$$Output = Maxpool(A_2) \quad \text{……(eq. 7)}$$

where, $\boldsymbol{W_1}$ & $\boldsymbol{W_2}$= "convolutional filters", $\boldsymbol{b_1}$& $\boldsymbol{b_2}$= "bias terms", ReLU= "REctified Linear Unit activation function".

Experimental Setup

The images that have been annotated were further processed with image augmentation with objective to increase dataset. In order to reduce the dimensionality of the photos, they were scaled down from 384 by 288 to 96 by 96. The initial dataset of 910 images was augmented by applying 11 transformation operations to generate 10,010 images for processing. The various operation performed on selected images for data augmentation were – blur, flip, rotation, brightness, etc. Figure 4 illustrates how these photos were divided into three sets—training, validation, and testing—each with two subsets: positive and negative.

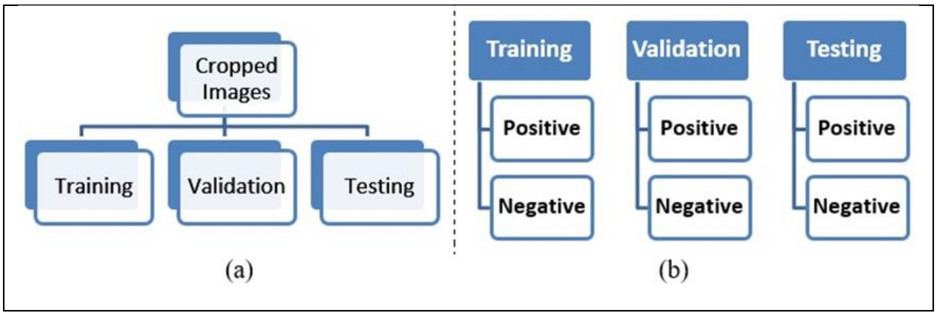

Fig 4: Data hierarchy – (a) Cropped dataset grouped in 3 sub datasets and (b) each sub dataset has two groups of datasets – positive and negative samples.

The Keras library contains several pre-trained CNN models that were trained using the ImageNet dataset. These already-trained models may be used for feature extraction, prediction, and fine-tuning. Few of these well-known pre-trained models include VGG16, ResNet50, Inception V3, Xception, and GoogLeNet.

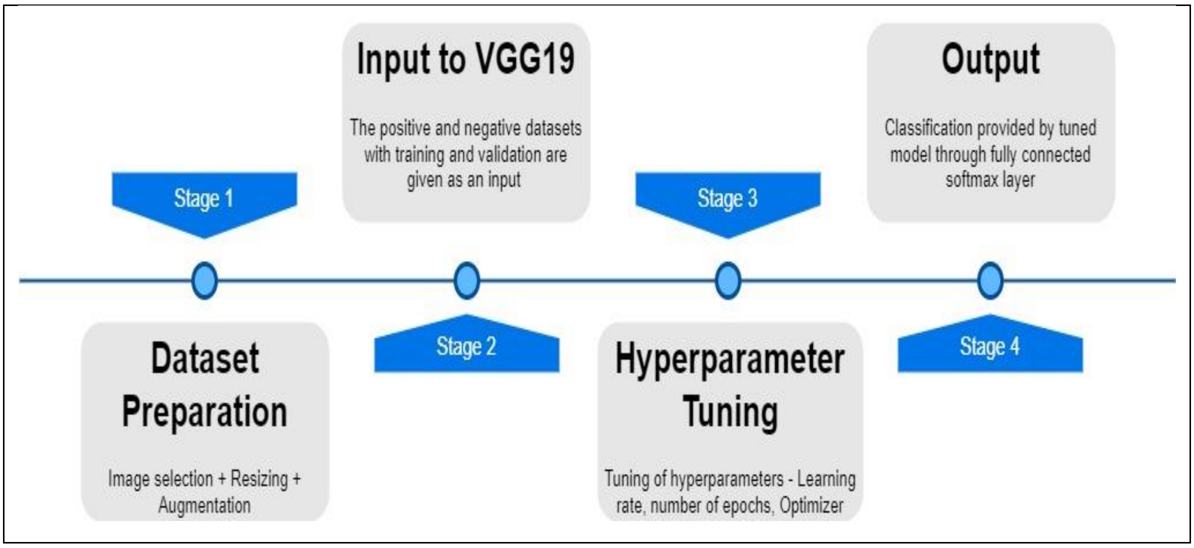

Fig 5: Proposed Deep Learning Architecture for Polyp Detection

VGG19 outperformed but with larger memory requirements than VGG16. The 16 and 19 layers in the VGG16 and VGG19 models, respectively, are made up of convolution layers, layers with maximum pooling, and fully connected layers. VGG16 is investigated and also modified [33] with reference to the work published previously [32] to examine the effectiveness for the diagnosis of gastrointestinal polyps. Figure- 4 displays a diagrammatic representation of the VGG16 architecture. [50] The major stages in the experiment are as shown in Figure-5,6.

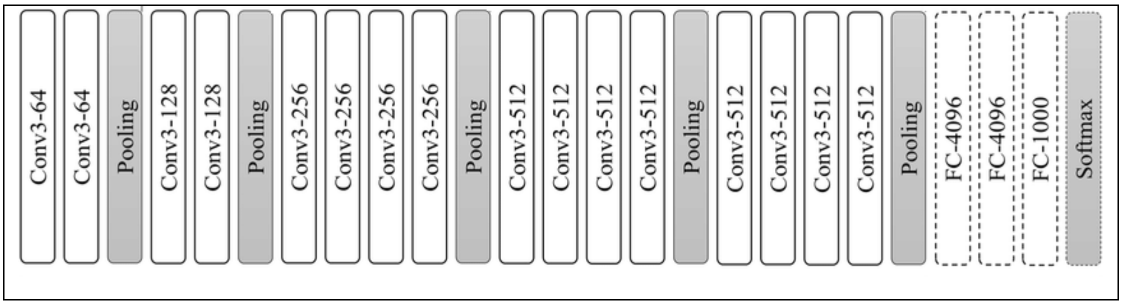

Fig 6: Layered architecture of VGG19

The layers are called "3 by 3 conv, X" layers, where "conv" denotes convolutional layers with 3 by 3 kernels and "X" denotes the quantity of filters. The MaxPool layer implemented the reduction of dimensions by a factor of 2. Fully Connected layer had 1000 and 4096 units at the very end. Softmax generates one of a thousand classifications using labeled data from ImageNet.

4. Results

In this present work, the VGG16 model was executed for polyp detection i.e., ‘polyp detected’ and ‘polyp not detected’. In the implementation, the VGG16 model weights were assigned from ‘imagenet’. As we are aimed at binary classification, the output layer was flattened to 1. The ‘relu’ and ‘sigmoid’ activation functions were used with 1024 hidden units and a drop rate of 0.4. After execution of this model, the outcomes are depicted in Figure-7 (a).

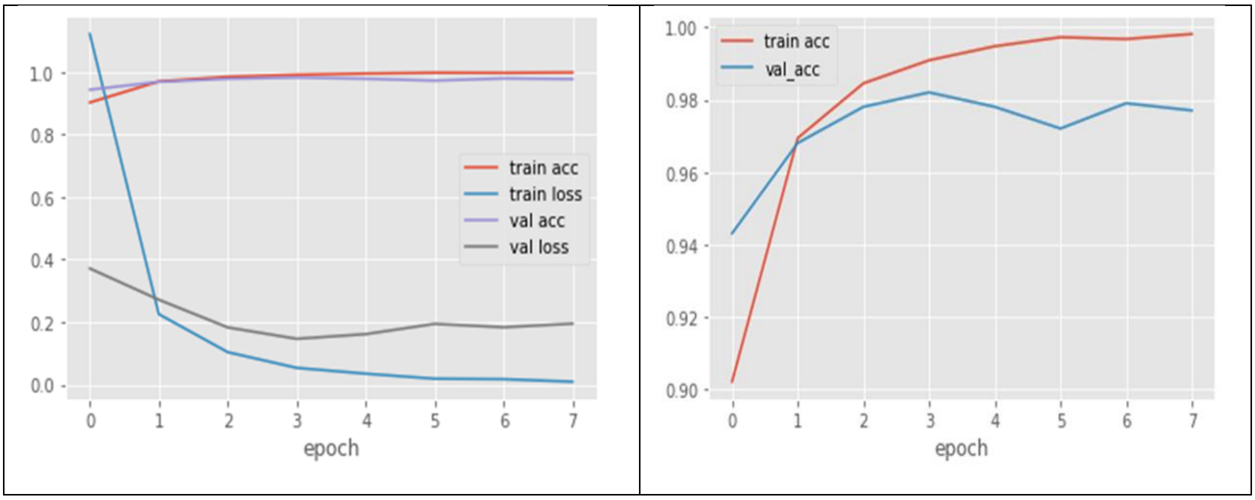

Fig 7: (a) The plot of accuracy and loss in training and validation and (b) Training accuracy and validation accuracy zoomed from 0.90 to 1.00 graph.

The graph in 7 (a) is magnified further from 0.90 to 1.00 on y-axis to plot the details of validation accuracy and training accuracy as shown in Figure 7(b). The scores obtained in the table 4 are visualized using a graph as shown in Figure-8 (a). To test the implementation, 445 samples from each set i.e., positive, and negative were considered with class labels as 'polyp not present' and 'polyp is present' as shown in Figure-8 (b).

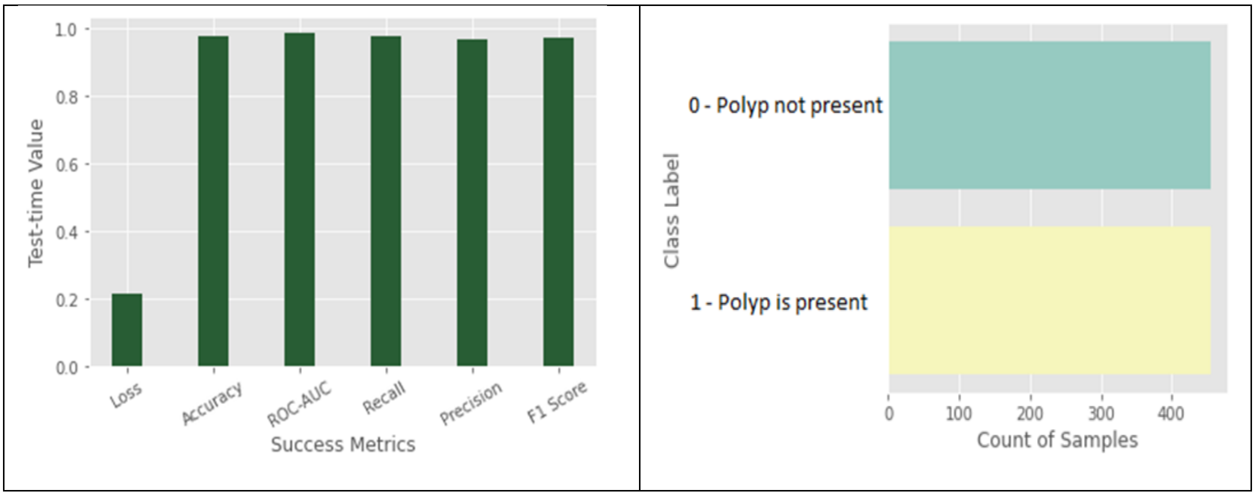

Fig 8: (a) Bar graph to plot the multiple scores and (b) The test dataset of positive and negative images

To measure the performance of classification for the implemented system, confusion matrix is used. The score value for the "Receiver Operating Characteristic" (ROC) curve's "Area Under Curve" (AUC) is also obtained. These obtained scores are as given in table 4.

Table 4 Performance measurement scores

Sr. No	Measurement Parameter	Value
1	Test Loss	0.214212
2	Test Accuracy	0.976024
3	AUC-ROC	0.986348
4	Recall	0.978022
5	Precision	0.967391
6	F1 Score	0.972679

From the dataset displayed in Figure-4 (b), we selected 10 random samples and presented the results as shown in Figure-9. In these 10 images, '0' represents negative sample, i.e. polyp not present and '1' represents positive sample, i.e., polyp present.

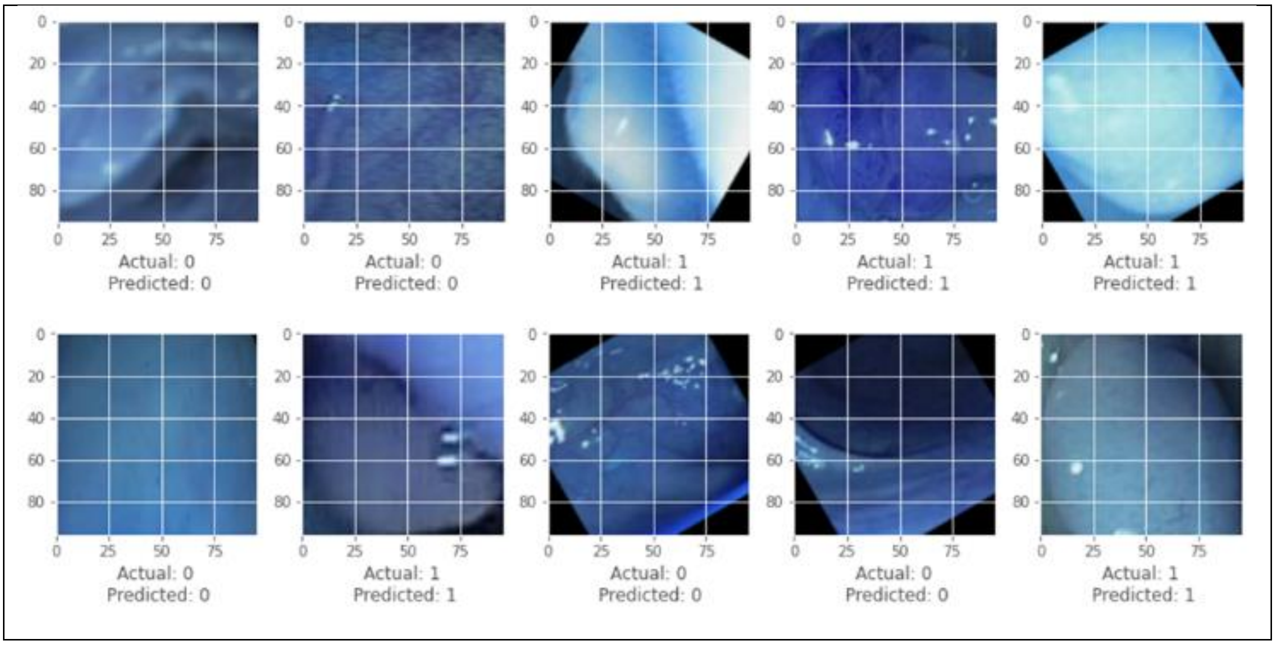

Fig 9: Visual test results obtained for 10 randomly selected test images

The 'actual' means the class of test image, i.e. positive or negative and 'predicted' is the output generated by the implemented system. The results can be obtained for any given specific input images. The result obtained for a specific input is shown in Figure-10.

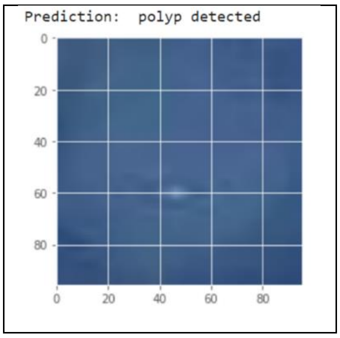

Fig 10: Polyp detection in a specific image.

5. Discussion

The colonoscopy continues to proves to be the golden standard with highest-sensitivity, specificity, and accuracy with capability of biopsy. Screening tests demonstrate the potential to prevent CRC at an early detection stage, hence, nations should implement regular population-based CRC screening programs [56-61]. The success of population-based programs depends on factors, such as (a) screening costs, (b) number of skilled colonoscopist/endoscopist per thousand population, (c) ergonomic burden of screening procedures on professionals, (d) patients' participation in programs, (e) need of social distancing in pandemic situations like COVID19, (f) training and long learning curve to achieve to achieve satisfactory experience by practitioners (g), increase in number of "need-to-screen" individuals due to increasing population, changes in lifestyle and environmental factors like pollution and climate change, (h) standardization of instruments and procedures, and (i) diagnostic methods used. The acceptance and approval of ColonFlag, GI-Genius, and Cologuard defines the success of AI in endoscopy.

The application of AI techniques in polyp detection has not yet been fully integrated into the cyber-physical domain. There is a limited amount of research that was conducted on the utilization of cyber-physical systems for the automated diagnosis of polyps. In the current context, it is possible to enhance the capabilities of cameras used in colonoscopy by integrating processing units and communication systems that adhere to standard protocols.

6. Conclusion

The use of TL in polyp detection during colonoscopy is a game-changer in CRC screening. It significantly improves the performance of DL algorithms. Studies have shown that optimized TL approaches can enhance polyp detection sensitivity, reduce false positives, and increase the chances of detecting multiple polyps. Hence TL is promising candidate in CAD. At the same time, challenge in polyp detection using TL is the performance variation based on dataset for fine-tuning. Another challenge is that TL can be susceptible to overfitting, which can lead to poor performance on new data. Transfer learning is a promising approach for improving the accuracy and efficiency of polyp detection in colonoscopy. However, more research is needed to address the challenges of transfer learning, such as dataset variability and overfitting.

Future research on transfer learning for polyp detection in colonoscopy could focus on the following areas:

- Developing more robust AI models that are less susceptible to overfitting.
- Developing methods for combining transfer learning with other AI techniques, such as deep reinforcement learning.
- Developing methods for integrating transfer learning-based CAD systems into clinical practice.

By addressing these challenges, transfer learning has the potential to revolutionize the way that polyps are detected during colonoscopy.

References

[1] Kudo SE, Mori Y, Misawa M, Takeda K, Kudo T, Itoh H, Oda M, Mori K, Artificial intelligence and colonoscopy: Current status and future perspectives, Digestive Endoscopy, Jul;31(4), 2019, 363-37

[2] Seong Ho Park , Kyunghwa Han, Methodological Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction, Radiology journal of Radiological Society of North America, Jan Volume 286, 2018, 800-809

[3] June-Goo Lee, Sanghoon Jun, Young-Won Cho, Hyunna Lee, Guk Bae Kim, Joon Beom Seo, Namkug Kim, Deep Learning in Medical Imaging: General Overview, Korean Journal of Radiology, 2017 Jul-Aug;18(4):570-584

[4] Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol. 2021 Oct;14(10):101174. doi: 10.1016/j.tranon.2021.101174. Epub 2021 Jul 6. PMID: 34243011; PMCID: PMC8273208.

[5] Johnson CD, ChenM-H, Toledano AYet al (2008) Accuracy of CT colonography for detection of large adenomas and cancers. N Engl J Med 359(12):1207–1217

[6] N. Boonsim and S. Kanjaruek, “Optimized transfer learning for polyp detection”, ECTI-CIT Transactions, vol. 17, no. 1, pp. 73–81, Feb. 2023.

[7] Nefeli Panagiota Tzavara and Bjørn-Jostein Singstad, “Transfer Learning in Polyp and Endoscopic Tool Segmentation from Colonoscopy Images” Nordic Machine Intelligence, MedAI 2021 https://doi.org/10.5617/nmi.9132

[8] M. Luca, A. Ciobanu and V. Drug, "Deep Learning and Automatic Polyp Detection in Colonoscopies: a Review of Recent Contributions and Future Outlook," 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania, 2019, pp. 1-4, doi: 10.1109/EHB47216.2019.8970041.

[9] Tang CP, Chen KH, Lin TL. Computer-Aided Colon Polyp Detection on High Resolution Colonoscopy Using Transfer Learning Techniques. Sensors (Basel). 2021 Aug 6;21(16):5315. doi: 10.3390/s21165315. PMID: 34450756; PMCID: PMC8402119.

[10] Noppakun, Boonsim., Saranya, Kanjaruek. (2023). Optimized transfer learning for polyp detection. ECTI Transactions on Computer and Information Technology, doi: 10.37936/ecti-cit.2023171.250910

[11] Yuheng, Li., Mingzhe, Hu., Xiaofeng, Yang. (2023). Polyp-SAM: Transfer SAM for Polyp Segmentation. arXiv.org, doi: 10.48550/arXiv.2305.00293

[12] (2023). Polyp-SAM: Transfer SAM for Polyp Segmentation. doi: 10.48550/arxiv.2305.00293

[13] Ricardo, Ribani., Mauricio, Marengoni. (2019). A Survey of Transfer Learning for Convolutional Neural Networks. doi: 10.1109/SIBGRAPI-T.2019.00010

[14] Hosna, A., Merry, E., Gyalmo, J. et al. Transfer learning: a friendly introduction. J Big Data 9, 102 (2022). https://doi.org/10.1186/s40537-022-00652-w

[15] Weiss, K., Khoshgoftaar, T.M. & Wang, D. A survey of transfer learning. J Big Data 3, 9 (2016). https://doi.org/10.1186/s40537-016-0043-6

[16] Hosna A, Merry E, Gyalmo J, Alom Z, Aung Z, Azim MA. Transfer learning: a friendly introduction. J Big Data. 2022;9(1):102. doi: 10.1186/s40537-022-00652-w. Epub 2022 Oct 22. PMID: 36313477; PMCID: PMC9589764.

[17] Iman M, Arabnia HR, Rasheed K. A Review of Deep Transfer Learning and Recent Advancements. Technologies. 2023; 11(2):40. https://doi.org/10.3390/technologies11020040

[18] R. Ribani and M. Marengoni, "A Survey of Transfer Learning for Convolutional Neural Networks," 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), Rio de Janeiro, Brazil, 2019, pp. 47-57, doi: 10.1109/SIBGRAPI-T.2019.00010.

[19] Yancong Zhou, Xuemei Zhang , Yan Wang , Bo Zhang Journal of Physics: Conference Series. 1920 012058, Transfer learning and its application research, 2021, doi:10.1088/1742-6596/1920/1/012058

[20] Lee JY, Jeong J, Song EM, Ha C, Lee HJ, Koo JE, Yang DH, Kim N, Byeon JS. Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. Sci

Rep. 2020 May 20;10(1):8379. doi: 10.1038/s41598-020-65387-1. PMID: 32433506; PMCID: PMC7239848.

[21] Macari M., Johnson C.D. (2006) CT Colonography. In: Hodler J., Von Schulthess G.K., Zollikofer C.L. (eds) Diseases of The Abdomen and Pelvis. Springer, Milano. https://doi.org/10.1007/88-470-0508-6_12

[22] Ciuti G, Skonieczna-Żydecka K, Marlicz W, Iacovacci V, Liu H, Stoyanov D, Arezzo A, Chiurazzi M, Toth E, Thorlacius H, Dario P, Koulaouzidis A. Frontiers of Robotic Colonoscopy: A Comprehensive Review of Robotic Colonoscopes and Technologies. Journal of Clinical Medicine. 2020; 9(6):1648. https://doi.org/10.3390/jcm9061648

[23] Sornapudi S, Meng F, Yi S. Region-Based Automated Localization of Colonoscopy and Wireless Capsule Endoscopy Polyps. Applied Sciences. 2019; 9(12):2404. https://doi.org/10.3390/app9122404

[24] Daniel Freedman, Yochai Blau, Liran Katzir, Amit Aides, Ilan Shimshoni, Danny Veikherman, Tomer Golany, Ariel Gordon, Greg Corrado, Yossi Matias, Ehud Rivlin. Detecting Deficient Coverage in Colonoscopies. Computer Vision and Pattern Recognition, Cornell University.

[25] Pickhardt PJ, Choi JR, Hwang I et al (2003) Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med 349(23):2191–2200

[26] Atkin W, Dadswell E, Wooldrage K et al (2013) Computed tomographic colonography versus colonoscopy for investigation of patients with symptoms suggestive of colorectal cancer (SIGGAR): a multicentre randomised trial. Lancet 381(9873):1194–1202

[27] Halligan S, Wooldrage K, Dadswell E et al (2013) Computed tomographic colonography versus barium enema for diagnosis of colorectal cancer or large polyps in symptomatic patients (SIGGAR): a multicentre randomised trial. Lancet 381(9873): 1185–1193

[28] Obaro, A.E., Plumb, A.A., North, M.P. et al. Computed tomographic colonography: how many and how fast should radiologists report? Eur Radiol 29, 5784–5790 (2019). https://doi.org/10.1007/s00330-019-06175-y

[29] Scalise P, Mantarro A, Pancrazi F, Neri E. Computed tomography colonography for the practicing radiologist: A review of current recommendations on methodology and clinical indications. World J Radiol. 2016;8(5):472-483. doi:10.4329/wjr.v8.i5.472

[30] Rollandi GA, Biscaldi E, DeCicco E. Double contrast barium enema: technique, indications, results and limitations of a conventional imaging methodology in the MDCT virtual endoscopy era. Eur J Radiol. 2007 Mar;61(3):382-7. doi: 10.1016/j.ejrad.2006.07.032. Epub 2006 Dec 11. PMID: 17161931.

[31] Ashfaq Chandio, Ayub Ali, Ali Javaid, Syed Mustafa, Ruxandra Pietrosanu, Fuad Aftab, Muhammmed Ibrahim and Syed Naqvi. The Diagnostic Yield of Double Contrast Barium Enema. International Journal of Surgery Research and Practice, ISSN: 2378-3397

[32] Adebogun AO, Berg CD, Laiyemo AO. Concerns and challenges in flexible sigmoidoscopy screening. Colorectal Cancer. 2012;1(4):309-319. doi:10.2217/crc.12.33

[33] Ko CW, Doria-Rose VP, Barrett MJ, Kamineni A, Enewold L, Weiss NS (2019) Screening colonoscopy and flexible sigmoidoscopy for reduction of colorectal cancer incidence: A case-control study. PLoS ONE 14(12): e0226027. https://doi.org/10.1371/journal.pone.0226027

[34] Cross AJ, Wooldrage K, Robbins EC, Pack K, Brown JP, Hamilton W, Thompson MR, Flashman KG, Halligan S, Thomas-Gibson S, Vance M, Saunders BP, Atkin W. Whole-colon investigation vs. flexible sigmoidoscopy for suspected colorectal cancer based on presenting symptoms and signs: a multicentre cohort study. Br J Cancer. 2019 Jan;120(2):154-164. doi: 10.1038/s41416-018-0335-z. Epub 2018 Dec 19. PMID: 30563992; PMCID: PMC6342953.

[35] Nauzer Forbes, Robert J. Hilsden, Steven J. Heitman. The appropriate use of fecal immunochemical testing. CMAJ 2020 January 20;192:E68. doi: 10.1503/cmaj.190901

[36] Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Intern Med. 2014;160(3):171. doi:10.7326/M13-1484

[37] Niedermaier T, Balavarca Y, Brenner H. Stage-Specific Sensitivity of Fecal Immunochemical Tests for Detecting Colorectal Cancer: Systematic Review and Meta-Analysis. Am J Gastroenterol. 2020 Jan;115(1):56-69. doi: 10.14309/ajg.0000000000000465. PMID: 31850933; PMCID: PMC6946106.

[38] Cubiella J, Castro I, Hernandez V, et al. Diagnostic accuracy of fecal immunochemical test in average- and familial-risk colorectal cancer screening. United European Gastroenterol J. 2014;2(6):522-529. doi:10.1177/2050640614553285

[39] Liles EG, Perrin N, Rosales AG, Smith DH, Feldstein AC, Mosen DM, Levin TR. Performance of a quantitative fecal immunochemical test for detecting advanced colorectal neoplasia: a prospective cohort study. BMC Cancer. 2018 May 2;18(1):509. doi: 10.1186/s12885-018-4402-x. PMID: 29720130; PMCID: PMC5932873

[40] Ruth M. Ayling, Stephen J. Lewis, Finbarr Cotter. Potential roles of artificial intelligence learning and faecal immunochemical testing for prioritisation of colonoscopy in anaemia. British Journal of HAEMATOLOGY. 03 February 2019 https://doi.org/10.1111/bjh.15776

[41] Soin, S., Akanbi, O., Ahmed, A. et al. Use and abuse of fecal occult blood tests: a community hospital experience. BMC Gastroenterol 19, 161 (2019). https://doi.org/10.1186/s12876-019-1079-9

[42] Lee MW, Pourmorady JS, Laine L. Use of Fecal Occult Blood Testing as a Diagnostic Tool for Clinical Indications: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2020 May;115(5):662-670. doi: 10.14309/ajg.0000000000000495. PMID: 31972617.

[43] Eckmann JD, Ebner DW, Kisiel JB. Multi-Target Stool DNA Testing for Colorectal Cancer Screening: Emerging Learning on Real-world Performance. Curr Treat Options Gastroenterol. 2020 Jan 21. doi: 10.1007/s11938-020-00271-5. Epub ahead of print. PMID: 31965446.

[44] Ferlizza E, Solmi R, Sgarzi M, Ricciardiello L, Lauriola M. The Roadmap of Colorectal Cancer Screening. Cancers (Basel). 2021;13(5):1101. Published 2021 Mar 4. doi:10.3390/cancers13051101

[45] Niederreiter, M., Niederreiter, L., Schmiderer, A. et al. Colorectal cancer screening and prevention—pros and cons. memo 12, 239–243 (2019). https://doi.org/10.1007/s12254-019-00520-z

[46] Goyal H, Mann R, Gandhi Z, et al. Scope of Artificial Intelligence in Screening and Diagnosis of Colorectal Cancer. J Clin Med. 2020;9(10):3313. Published 2020 Oct 15. doi:10.3390/jcm9103313

[47] Strul H, Arber N. Screening techniques for prevention and early detection of colorectal cancer in the average-risk population. Gastrointest Cancer Res. 2007;1(3):98-106.

[48] (2023). Plant Disease Classification using Deep Learning Approach (VGG19). doi: 10.1109/icaccs57279.2023.10112823

[49] R., C., Meena., Sunil, Joshi., Sandeep, K., Raghuwanshi. (2022). VGG19-based Transfer learning for Rice Plant Illness Detection. doi: 10.1109/ICECA55336.2022.10009303

[50] (2023). Style transfer with VGG19. Applied and Computational Engineering, doi: 10.54254/2755-2721/6/20230752

[51] Marcel, Simon., Erik, Rodner., Joachim, Denzler. (2016). ImageNet pre-trained models with batch normalization.. arXiv: Computer Vision and Pattern Recognition,

[52] Marcel, Simon., Erik, Rodner., Joachim, Denzler. (2016). ImageNet pre-trained models with batch normalization.. arXiv: Computer Vision and Pattern Recognition,

[53] Bernal, J., Sánchez, F. J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., & Vilariño, F. (2015). WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. Computerized Medical Imaging and Graphics, 43, 99-111.

[54] Yogesh Chaudhari, R. Senthil Kumar, Darshee Baxi. AI based automated diagnosis of polyps in colonoscopy: An introduction for researchers and Physicians in healthcare. Journal of Engineering Research and Application, ISSN: 2248-9622 Vol. 9, Issue 12 (Series - II) December 2019, pp 06-12

[55] Bansal, M., Kumar, M., Sachdeva, M. et al. Transfer learning for image classification using VGG19: Caltech-101 image data set. J Ambient Intell Human Comput 14, 3609–3620 (2023). https://doi.org/10.1007/s12652-021-03488-z

[56] Qaseem, A. Crandall, C.J. Mustafa, R.A. Hicks, L.A.Wilt, T.J. Screening for colorectal cancer in asymptomatic average-risk adults: A guidance statement from the American college of physicians. Ann. Intern. Med. 2019, 171, 643–654.

[57] Lacy, B.E. Chan, J.L. Physician burnout: The hidden health care crisis. Clin. Gastroenterol. Hepatol. 2018, 16, 311–317.

[58] Siau, K. Anderson, J.T. Ergonomics in endoscopy: Should the endoscopist be considered and trained like an athlete? Endosc. Int. Open 2019, 7, E813–E815.

[59] Krigel, A. Chen, L. Wright, J.D. Lebwohl, B. Substantial increase in anesthesia assistance for outpatient colonoscopy and associated cost nationwide. Clin. Gastroenterol. Hepatol. 2019, 17, 2489–2496.

[60] Robertson, D.J. Ladabaum, U. Opportunities and challenges in moving from current guidelines to personalized colorectal cancer screening. Gastroenterology 2019, 156, 904–917.

[61] Repici, A. Maselli, R. Colombo, M. Gabbiadini, R. Spadaccini, M. Anderloni, A. Carrara, S. Fugazza, A. Di Leo, M. Galtieri, P.A. et al. Coronavirus (COVID-19) outbreak: What the department of endoscopy should know. Gastrointest. Endosc. 2020, in press