IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

A Synergistic Approach to Colon Cancer Detection: Leveraging EfficientNet and NSGA-II for Enhanced Diagnostic Performance

NOUSHIN SABA[1], AFIA ZAFAR[1], MOHSIN SULEMAN[1], KAINAT ZAFAR[1],SHAHNEER ZAFAR[1], ADIL ALI SALEEM[2], HAFEEZ UR REHMAN SIDDIQUI[2], MUHAMMAD IQBAL[3]* and SYED SAJID ULLAH[4]*

[1]Department of computer science, National university of Technology, Islamabad, Pakistan (e-mail:noushin.saba@nutech.edu.pk; Afiazafar@nutech.edu.pk; mohsin.suleman@nutech.edu.pk; Kainatz@nutech.edu.pk; Shahneerzafar@nutech.edu.pk)
[2]Intitute of Computer Science, Khwaja Fareed Univeristy of Engineering and Information Technology, Rahim Yar Khan, Punjab, Pakistan. (e-mail:adilalisaleem@gmail.com; hafeez@kfueit.edu.pk.)
[3]School of Interdisciplinary Engineering and Sciences (SINES), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan (e-mail:muhammad.iqbal@sines.nust.edu.pk)
[4]Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway. (e-mail:syed.s.ullah@uia.no)

Corresponding author: Muhammmad Iqbal, Syed Sajid Ullah (e-mail: muhammad.iqbal@sines.nust.edu.pk, syed.s.ullah@uia.no).

ABSTRACT Colon cancer remains a leading cause of cancer-related mortality globally, necessitating early and accurate diagnosis to improve patient outcomes. Traditional diagnostic methods rely heavily on manual interpretation by pathologists, which can result in inaccuracies and delays in treatment. This study proposes an innovative, automated approach to colon cancer diagnosis by integrating advanced machine learning techniques with deep learning architectures. We employed EfficientNet, a state-of-the-art convolutional neural network, to extract intricate features from histopathological images, alongside the Non-dominated Sorting Genetic Algorithm II for optimal feature selection. This hybrid approach significantly enhances diagnostic performance while reducing computational complexity. The model was evaluated using five diverse datasets: Colon Cancer Histopathological Images, Kvasir, Kvasir-SEG, Hyper-Kvasir, and Endotect. The results indicate that our method outperforms traditional models such as CNN, AlexNet, ResNet, and GoogleNet, achieving an accuracy of 99.97% on the Colon Cancer Histopathological Images dataset. These findings suggest that this novel approach can substantially enhance early detection and diagnosis of colon cancer, providing a scalable solution to current diagnostic challenges. Ultimately, our study lays the groundwork for future advancements in automated cancer diagnostics, contributing to improved patient outcomes and more efficient healthcare delivery. The code and dataset for reproducing these results are publicly accessible at https://github.com/Noushin-Saba/ColonCancerDetectionandDiagnosis.

INDEX TERMS CNN, Colon Disease, Cancer, EfficientNet, EfficientNet-NSGA-II, NSGAII

I. INTRODUCTION

CANCER remains one of the most prevalent and complex diseases worldwide, affecting millions of lives each year. Early detection is crucial for improving survival rates, particularly for cancers like colon cancer, where treatment is more effective when the disease is identified at an early stage [1]. However, conventional diagnostic methods, such as biopsies carry health risks, including bleeding and infection, and are not always ideal for detecting early-stage cancers. Furthermore, these methods can be invasive, time-consuming, and costly, underscoring the urgent need for alternative, non-invasive diagnostic tools [2]. Colon cancer, characterized by the uncontrolled growth of cells in the colon, poses a significant challenge to global health. It often results from genetic instability and the accumulation of various molecular alterations, including abnormal activation of genes that regulate cell growth and mitosis. The delay in diagnosis and treatment remains a critical issue, contributing to cancer surpassing heart disease as the leading cause of death across all age groups. According to the World Health Organization, colon cancer is the second leading cause of cancer-related deaths globally and the third most frequently diagnosed cancer. In 2020, over 1.9 million new cases were reported, resulting in more than 930,000 deaths, accounting for 9.4% of all cancer-

related fatalities. Projections indicate that by 2040, there will be 3.2 million new cases and 1.6 million deaths annually [3]. In Pakistan, the incidence rate for colon cancer is 4.4 per 100,000 men and 3.9 per 100,000 women, making it the fourth most common cancer in the country and a major contributor to years of life lost adjusted for disability [4]. Figure 1 shows the colon with cancer and without cancer.

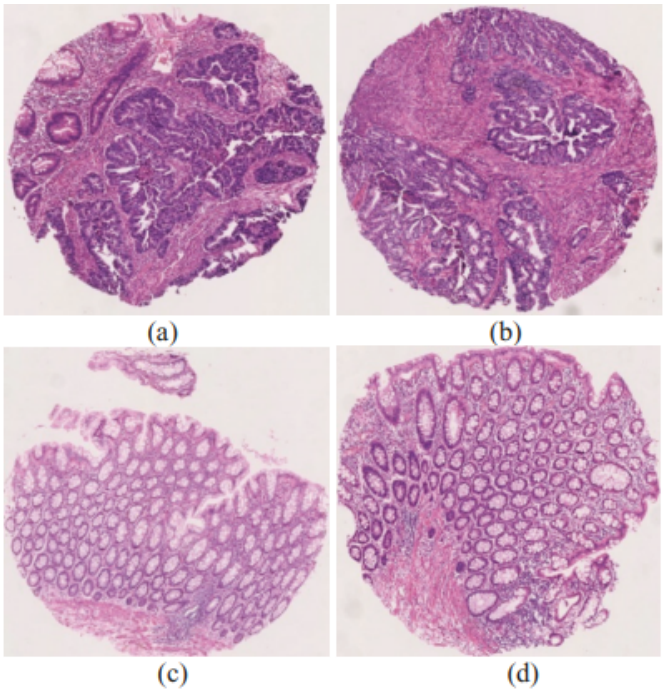


FIGURE 1. **(a) and (b) Images of colon cancer; (c) and (d) Images without colon cancer.**

While image processing techniques are pivotal in diagnosing colon diseases, careful analysis is crucial as improper handling, such as denoising, can degrade feature quality and lead to inaccurate classifications. Filters used during preprocessing may inadvertently remove essential features, thereby reducing their effectiveness in diagnosis [5]. Current methods for colorectal cancer detection, particularly those employing machine learning and deep learning approaches, have made strides but still suffer from notable gaps. Many existing techniques, especially those relying on CNN architectures, face challenges related to feature redundancy, high computational costs, and the black-box nature of deep learning models, which limits their interpretability in clinical settings. Moreover, there is a lack of optimization in the feature selection process, leading to suboptimal performance due to the presence of irrelevant or redundant features. Existing studies often focus on limited datasets, reducing the generalizability and robustness of the proposed models across diverse clinical environments. Additionally, many previous works fail to adequately address the balance between achieving high accuracy and maintaining computational efficiency, which is crucial for real-time clinical application. Therefore, there is a need for advanced, automated methods that can enhance the accuracy and efficiency of colon cancer diagnosis.

To address these challenges, this paper proposes an integrated approach that combines machine learning with neural network training to improve feature extraction and disease detection. The proposed study aims to bridge these gaps by introducing a novel approach that leverages the strengths of EfficientNet, a state-of-the-art CNN architecture, for feature extraction, and NSGA-II, a multi-objective genetic algorithm, for optimal feature selection. By integrating these two powerful techniques, the study not only enhances diagnostic accuracy but also reduces computational complexity, making the model more suitable for practical deployment in clinical settings. This hybrid model addresses the limitations of existing approaches by minimizing feature redundancy, improving interpretability, and delivering superior performance across multiple datasets. The innovation lies in the combination of EfficientNet's efficient learning capabilities with NSGA-II's robust optimization framework, which together ensure that the proposed system can generalize well across diverse imaging conditions and provide reliable early diagnosis for colorectal cancer. The main contribution of this paper is listed below.

- Introduced an innovative approach that integrates EfficientNet for advanced feature extraction and NSGA-II for optimal feature selection, addressing challenges like feature redundancy and high computational costs. This hybrid model demonstrates a significant leap in colon cancer detection performance.
- Achieved lower computational complexity compared to traditional models, making the system more efficient and suitable for clinical applications in colon disease detection.
- Conducted extensive evaluations on five publicly available datasets (Colon Cancer Histopathological Images, Kvasir, Kvasir-SEG, Hyper-Kvasir, and Endotect), showcasing the model's scalability, generalizability, and robustness across various imaging conditions.
- Provided a scalable and robust diagnostic tool, offering a foundation for future research to improve model interpretability and clinical acceptance.

II. LITERATURE REVIEW

In this section, we explore a range of feature extraction and classification techniques. This includes both traditional handcrafted methods and advanced deep learning algorithms that have been employed for the classification of images in the context of colon disease. Colorectal cancer is one of the leading causes of cancer related mortality worldwide, making its early detection crucial for reducing death rates. Traditionally, physicians have relied on manual methods such as visual inspection during colonoscopy, biopsy, histopathological examination, and fecal occult blood tests (FOBT) to detect and diagnose colorectal cancer [6]. While these methods remain the gold standard, they are highly dependent on the expertise of the physician and are prone to errors due to human fatigue, subjective interpretation, and variability

among pathologists [7]. Moreover, traditional approaches often require invasive procedures, which can be uncomfortable for patients and carry risks such as bleeding or perforation [8]. These challenges have driven the demand for automated and non-invasive techniques that offer high sensitivity and specificity while reducing human error and improving the consistency of results [9]. In response to the limitations of conventional methods, computer-aided detection (CAD) systems have emerged as promising tools for enhancing the accuracy and efficiency of colorectal cancer screening. These systems leverage image processing algorithms to assist clinicians in detecting and characterizing polyps more reliably. Among these methods, Convolutional Neural Networks (CNNs) have become particularly prominent due to their ability to automatically learn complex features from large datasets, enabling precise detection and classification of polyps in colonoscopy images [10].

CNNs have revolutionized medical image analysis [11], [12] by providing state-of-the-art performance in detecting and segmenting polyps. For example, recent studies have shown that CNNs can achieve higher accuracy in polyp detection compared to traditional image processing methods by learning robust features directly from data [13], [14]. However, CNNs are not without limitations: they require large amounts of labeled data for training, are prone to overfitting, and often suffer from high computational costs [15]. Moreover, interpreting the features learned by CNN models remains a significant challenge, as these models are often perceived as "black boxes," with limited transparency into the decision-making process. The lack of interpretability can undermine clinical trust and hinder the adoption of deep learning-based approaches in medical practice [16]. Furthermore, current approaches do not effectively address the issue of feature redundancy and irrelevant data, which can lead to decreased model performance. Most deep learning models extract a high number of features, some of which may not be relevant for the classification task. The presence of redundant or irrelevant features increases computational complexity, reduces model interpretability, and can adversely impact the accuracy and efficiency of the diagnostic process [17]. These gaps underscore the need for an approach that leverages the strengths of deep learning for feature extraction while incorporating robust feature selection techniques to enhance model performance and clinical applicability.

Despite the progress made, existing CADx systems and hybrid models still face several challenges. Many studies have been conducted on separate datasets, leading to difficulties in directly comparing performance outcomes. Moreover, the lack of standardized datasets can result in models that are not generalizable across different patient populations or clinical settings [18], [19]. In addition, there is often a lack of explainability in these systems, which hinders their acceptance among healthcare professionals [20]. Previous approaches to colorectal cancer detection have several notable limitations. Many earlier methods depend on a fixed set of handcrafted features, requiring an in-depth understanding of specific image characteristics [21]. These approaches often utilize texture analysis, where a limited number of local descriptors extracted from images are fed into classifiers such as Support Vector Machines (SVM) or Random Forests. Although some studies have achieved moderate accuracy, these techniques generally suffer from poor generalization and limited transferability across different datasets, reducing their effectiveness when faced with inter-dataset variability [22]. Additionally, many studies have tested their models on a narrow range of classes or less diverse datasets, which diminishes the robustness and applicability of their findings [23]. For example, research by Smith et al. (2020) was conducted on a dataset lacking sufficient diversity, limiting its generalizability to real-world scenarios. Furthermore, a considerable number of existing algorithms rely heavily on endoscopic and histological data, which can restrict their practical utility [24].

The reliance on histological data, in particular, is not always be feasible due to the invasive nature of the procedures or the specific clinical context, further limiting the applicability of these methods [25]. In recent years, several novel approaches have been explored to address these challenges. Oliveira et al. [26] proposed a hybrid model combining transfer learning with optimization algorithms to enhance feature selection in colorectal cancer detection, demonstrating improved accuracy and computational efficiency. Similarly, Kumar et al. [27] introduces CRCCN-Net, a lightweight convolutional neural network for automated classification of colorectal histopathological images, achieving comparably good performances. However, the model's reliance on specific datasets may limit its generalizability to diverse clinical settings, and its performance on out-of-distribution samples requires further investigation. These studies highlight the increasing focus on optimizing deep learning models to balance accuracy and efficiency, particularly through the use of feature selection techniques and hybrid models. However, many of these approaches still face limitations. For instance, while Alboaneen et al. [28] showed improved results, their model's performance deteriorated when tested on highly imbalanced datasets, a common issue in medical imaging. Additionally, interpretability remains a significant barrier for clinical adoption, as highlighted by recent work from Xu et al. [29], who emphasized the need for transparency in AI-driven medical systems.

Despite significant advancements, existing methods face persistent gaps, including feature redundancy, computational inefficiencies, and lack of interpretability, as highlighted in studies like [24] and [28]. Furthermore, the reliance on limited datasets, as observed in [27], often compromises generalizability. Building on these insights, proposed study addresses critical limitations of existing models by introducing a hybrid approach that combines EfficientNet with NSGA-II for optimized feature extraction and selection. Unlike previous studies that have primarily focused on CNN-based approaches without robust feature selection, our model reduces feature redundancy, improves interpretability, and enhances clinical applicability by selecting only the most relevant features

through NSGA-II. While prior research explored hybrid models, few have incorporated evolutionary algorithms such as NSGA-II to refine the feature selection process comprehensively [30], [31]. This study also enhances model scalability and efficiency, leveraging EfficientNet, which has demonstrated superior performance to traditional CNNs in medical imaging tasks. Additionally, incorporating NSGA-II allows us to address feature relevance and model transparency, aligning with Fang et al.'s [32] recommendation for feature selection techniques that improve model interpretability in diagnostic contexts. By critically evaluating recent advancements and challenges in colorectal cancer detection, this study extends the current body of knowledge by introducing a hybrid, interpretable approach to deep learning. The presented model not only aims to improve accuracy but also addresses generalizability and interpretability, responding to the broader demand for clinically viable, AI-based diagnostic solutions in colorectal cancer detection.

III. METHODOLOGY

In this study, a hybrid approach is proposed that combines EfficientNet and the nondominated classification genetic algorithm II (NSGA-II) for the early diagnosis of colon cancer. Histopathological images from five diverse datasets—Colon Cancer Histopathological Images, Kvasir, Kvasir-SEG, Hyper-Kvasir, and Endotect—were processed to extract complex features using EfficientNet, a state-of-the-art convolutional neural network. These features were then optimized via NSGA-II to reduce computational complexity and improve model performance. The effectiveness of the proposed method was evaluated against conventional models, including CNN, AlexNet, ResNet, and GoogleNet, using metrics such as accuracy, precision, recall, F1 score, and AUC.

A. DATASETS

This study utilized five diverse datasets summarized in Table 1, Lung and Colon Cancer Histopathological Images [33], Kvasir Dataset [34], Kvasir-SEG Dataset [35], Hyper-Kvasir Dataset [36], and Endotect Dataset [37] to comprehensively evaluate the performance of the proposed approach.These datasets, consisting of images generated by different colonoscopy devices and prepared in various laboratories, provide a broad and varied foundation for assessment. By training, validating, and testing the system on these heterogeneous image sets, the goal was to enhance its robustness and ensure adaptability across a wide range of imaging conditions.

TABLE 1. Summary of the datasets used in this study.

Dataset	Images	Classes	Modality	Data Source
Lung and Colon Cancer	25,000	Lung benign tissue, Lung adenocarcinoma, Lung squamous cell carcinoma, Colon adenocarcinoma, Colon benign tissue	Histopathological Images	Histopathology repositories
Kvasir Dataset	8,000	Dyed lifted polyps, Dyed resection margins, Esophagitis, Normal cecum, Normal pylorus, Normal z-line, Polyps, Ulcerative Colitis	Endoscopic Images	Collected at Vestre Viken Health Trust, Norway
Kvasir-SEG Dataset	1,000	Polyps	Endoscopic Images	Collected at Vestre Viken Health Trust, Norway
Hyper-Kvasir Dataset	110,079	Barretts, Bbps-0-1, Bbps-2-3, Dyed lifted polyps, Dyed resection margins, Hemorrhoids, Ileum, Impacted stool, Normal cecum, Normal pylorus, Normal z-line, Oesophagitis-a, Oesophagitis-b-d, Polyp, Retroflex-rectum, Retroflex-stomach, Short-segment-barretts, Ulcerative-colitis-0-1, Ulcerative-colitis-1-2, Ulcerative-colitis-2-3, Ulcerative-colitis-grade-1, Ulcerative-colitis-grade-2, Ulcerative-colitis-grade-3	Endoscopic Images	Collected at Vestre Viken Health Trust, Norway
Endotect Dataset	1,200	Polyps	Endoscopic Images	Images from Hyper-Kvasir

1) Lung and Colon Cancer Histopathological Images

The dataset [33] used in this study consists of 25,000 histopathological images across five classes, with each class containing 5,000 images. The images are 768 x 768 pixels in size and are in JPEG format. Originally, the dataset included 750 validated and HIPAA-compliant images of lung tissue (250 benign, 250 adenocarcinomas, and 250 squamous cell carcinomas) and 500 images of colon tissue (250 benign and 250 adenocarcinomas). To expand the dataset, the Augmentor package was employed, increasing the total number of images to 25,000. The five classes in the dataset are: lung benign tissue, lung adenocarcinoma, lung squamous cell carcinoma, colon adenocarcinoma, and colon benign tissue. For this research, only the colon adenocarcinoma and colon benign tissue classes were used for model training and evaluation. Sample images of the data set, including benign tissue of

the colon (denoted colonn) and adenocarcinoma of the colon (denoted colonca), are displayed in Figure 2. These images provide a visual representation of healthy versus cancerous colon tissue, highlighting the key differences captured in the histopathological analysis.

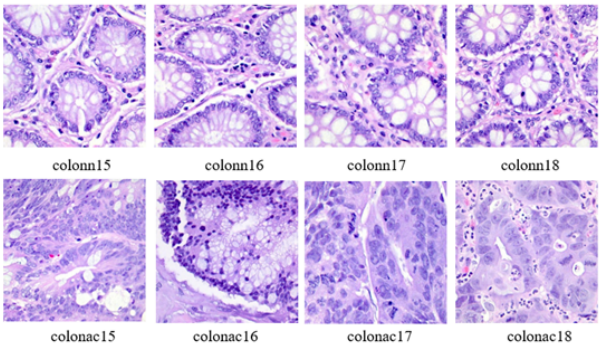


FIGURE 2. Sample histopathological images from the dataset. Colon benign tissue (colonn) and Colon adenocarcinoma (colonca).

2) Kvasir Dataset

The dataset comprises 8,000 images categorized into eight distinct classes, with 1,000 images per class. These images, varying in resolution from 720x576 to 1920x1072 pixels, are organized into folders based on their content. The data set includes both anatomical landmarks and pathological findings, along with images related to the removal procedures of the lesion. Anatomical landmarks such as the Z-line, pylorus, and cecum are featured, while pathological findings include esophagitis, polyps, and ulcerative colitis. In addition, lesion removal images, such as "dyed and lifted polyps" and "dyed resection margins", are also provided. Sample images from the dataset are shown in Figure 3, illustrating the variety of content and image quality.

3) Kvasir-SEG dataset

The Kvasir-SEG dataset, with a total size of 46.2 MB, comprises 1,000 polyp images along with their corresponding ground truth annotations sourced from the Kvasir Dataset v2. The images vary in resolution from 332x487 to 1920x1072 pixels and are encoded in JPEG format to facilitate efficient online browsing. Each image and its associated mask are organized into separate folders, with identical filenames to ensure easy reference. Sample images from the dataset are shown in Figure 4.

4) Hyper-Kvasir Dataset

The Hyper-Kvasir dataset, publicly available and developed using Olympus and Pentax imaging devices at Vestre Viken Hospital Trust, Norway, comprises 110,079 images and 374 videos. Verified by Vestre Viken Hospital, the Cancer Registry of Norway, and Karolinska University Hospital in Sweden, this comprehensive dataset includes a broad collection of images categorized into four distinct parts: labeled image data, unlabeled image data, segmented image data, and

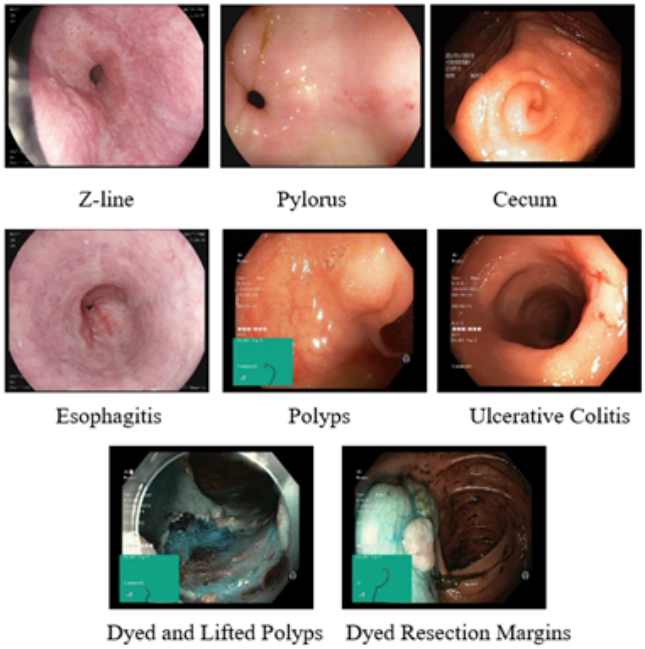


FIGURE 3. Sample images from the Kvasir dataset. Anatomical landmarks: Z-line, pylorus, and cecum. Pathological findings: esophagitis, polyps, and ulcerative colitis. Lesion removal images: dyed and lifted polyps, and dyed resection margins.

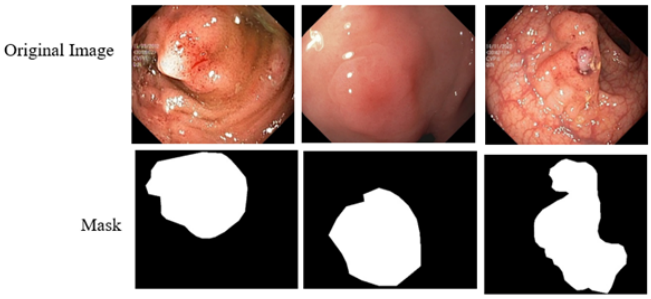


FIGURE 4. Sample images from the Kvasir-SEG dataset. Polyp images with varying resolutions.

annotated video data. For this study, 800 images of polyps were utilized, with expert review classifying them into hyperplastic and adenomatous categories. Although the dataset does not include subclass information for polyps, it provides segmentation masks for 1,000 polyp images, which were also verified by experts. Sample images from the dataset are shown in Figure 5, illustrating anatomical landmarks (cecum, ileum, retroflex-rectum) and pathological findings (hemorrhoids, polyps, and ulcerative colitis).

5) Endotect Dataset

The dataset is divided into four distinct categories: labeled image data, unlabeled image data, segmented image data, and annotated video data. It includes 110,079 images and 373 videos, capturing a range of anatomical landmarks, as well as pathological and normal findings. In total, the dataset en-

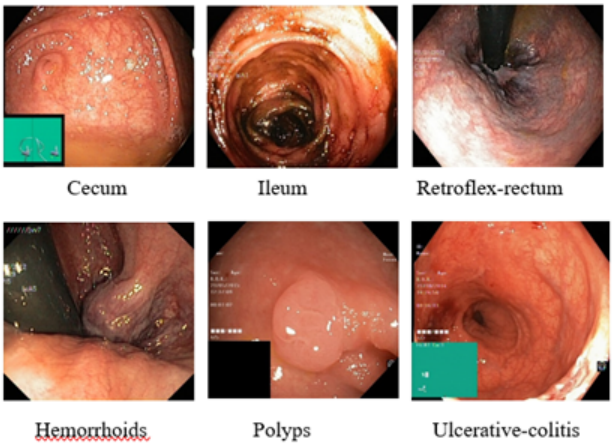


FIGURE 5. **Sample images from the Hyper-Kvasir dataset. Anatomical landmarks: cecum, ileum, and retroflex-rectum. Pathological findings: hemorrhoids, polyps, and ulcerative colitis.**

compasses over 1.1 million images and video frames. Sample images from the dataset are shown in Figure 6.

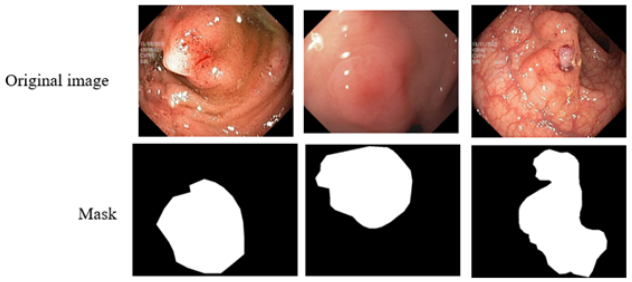


FIGURE 6. **Sample images from the Endotect Dataset.**

B. ARCHITECTURE OF PROPOSED SYSTEM

The proposed system introduces a novel hybrid approach combining the strengths of EfficientNet for feature extraction and NSGA-II for optimal feature selection, tailored specifically to the task of colon cancer detection. The process begins with the input of histopathological images, which undergo a comprehensive preprocessing stage. This involves converting images to RGB format, resizing them to a standardized dimension of 244x244 pixels to ensure consistency across the dataset, and converting them to tensor format. The images are then normalized using the mean and standard deviation to enhance the model's ability to learn meaningful patterns while minimizing the effects of variability in the input data. Next, the preprocessed images are fed into the EfficientNet model, a highly efficient convolutional neural network architecture known for its capability to achieve superior performance with fewer parameters. EfficientNet is employed to automatically extract high-level features that represent the complex texture, structure, and morphological characteristics of both healthy and cancerous colon tissues. This step is crucial for capturing subtle differences that are essential for accurate classification.

To further enhance the classification accuracy, the NSGA-II is employed for feature selection. NSGA-II is a robust multi-objective evolutionary algorithm designed to select the most relevant features while eliminating redundant or non-informative ones. This step optimizes the feature set, improving the classifier's efficiency by reducing the dimensionality of the data while maintaining its discriminative power. The selected features are then processed by a SVM classifier, which is chosen for its effectiveness in handling high-dimensional data and its ability to maximize the margin between different classes. The SVM classifier categorizes the images into "Healthy" or "Colon Cancer" categories, providing an initial diagnosis based on the extracted and optimized features. Finally, the performance of the proposed system is evaluated using several critical metrics, including the confusion matrix, accuracy, precision, recall, F1 score, area under the curve (AUC) and receiver operating characteristic curve (ROC). These metrics provide a comprehensive evaluation of the system's diagnostic capabilities, ensuring that it not only achieves high accuracy but also performs consistently across various test scenarios. The integration of EfficientNet and NSGA-II within this architecture presents a promising approach for early and accurate detection of colon cancer, potentially enhancing clinical decision-making and patient outcomes. Figure 7 shows the architecture of the proposed system.

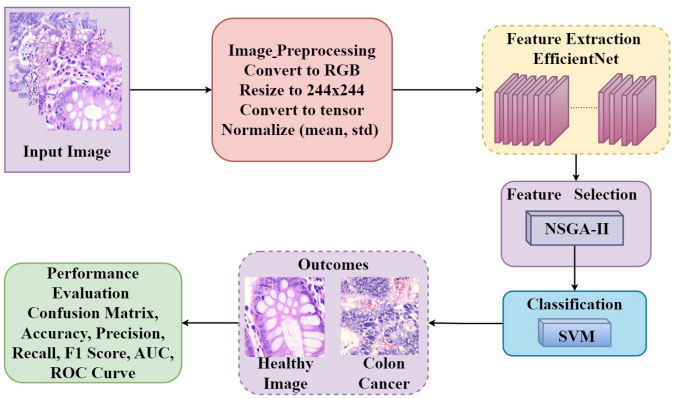


FIGURE 7. **Architecture of the proposed system.**

C. EFFICIENTNET

In this section, the proposed model's approach to learning and representing endoscopic features of colon diseases is described. The EfficientNet architecture is utilized for feature extraction in the proposed model, capitalizing on its ability to capture intricate patterns and representations from images without the need for fine-tuning. EfficientNet is renowned for its highly efficient scaling and performance, achieved through a compound scaling method that uniformly scales all dimensions of depth, width, and resolution. In this approach, the pretrained EfficientNet model, which has been trained on a large dataset, is leveraged to extract features from the endoscopic images of colon diseases. By using the output

from the model before the final classification layer as feature vectors, the rich, hierarchical representations learned during its initial training are effectively utilized. The pretrained features provide a sufficiently meaningful representation of the images for this specific task, enabling effective classification without necessitating additional fine-tuning. The Colorectal Disease Classification Using EfficientNet-NSGA-II architecture comprises seven blocks, which are fundamental components of EfficientNet and are based on Mobile Inverted Bottleneck Convolutions (MBConv). The process begins with an initial convolutional layer that applies a 2D convolution with 32 filters, a kernel size of 3x3, and a stride of 1. This layer extracts initial low-level features from the image. The subsequent step normalizes the output from the convolutional layer to accelerate training and enhance stability. Finally, the SiLU (Sigmoid Linear Unit) activation function is applied to introduce non-linearity. The SiLU activation function is mathematically defined as:

$$\text{SILU}(x) = x \cdot \frac{1}{1 + e^{-x}} \tag{1}$$

Each MBConv block is composed of several layers, including depthwise separable convolutions, batch normalization, activation functions, and Squeeze-and-Excitation (SE) layers. Specifically, MBConv1 applies depthwise convolution with a depthwise filter, which emphasizes spatial information. Due to a stride of 2, the output size changes, and the feature maps are normalized. The SiLU (Sigmoid Linear Unit) activation function is then applied to the normalized output. Additionally, the block adjusts channel weights through a squeeze operation (global average pooling) followed by an excitation step, which involves two fully connected layers with sigmoid activation. The output of MBConv1 is thus refined through these processes to capture essential features effectively.

$$\text{Squeeze} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X(i,j) \tag{2}$$

$$\text{Excitation} = \sigma\left(W_2 \cdot \text{ReLU}\left(W_1 \cdot Z\right)\right) \tag{3}$$

In these equations, H and W represent the height and width of the feature maps, $X(i,j)$ denotes the input feature map, and W_1 and W_2 are the weights of the fully connected layers. MBConv6 has a similar structure to MBConv1, but includes a higher expansion ratio, using six times more filters for depthwise convolution to capture more complex patterns. The stride and kernel sizes are varied across layers to progressively reduce the spatial dimensions. The output feature map generated by the model consists of 1,280 channels, representing a high-dimensional feature space that captures various learned attributes of the input image. This feature map is then processed by the final 1×1 convolutional layer, which reduces each spatial position across these channels into a single vector. As a result, this layer consolidates the information, producing a feature vector with 1,280 values. Thus, for each input image, the model generates a 1,280-dimensional feature vector, effectively summarizing the most critical aspects of the image in a compact form. The detailed working of the proposed EfficientNet architecture is presented in Figure 8.

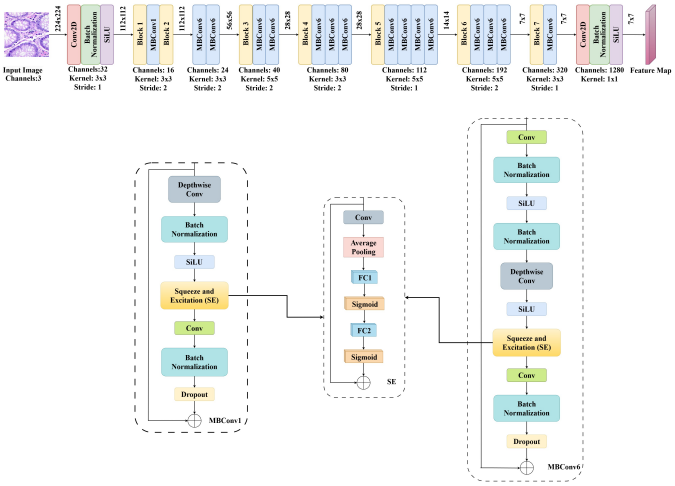


FIGURE 8. The proposed architecture of EfficientNet.

D. NON-DOMINATED SORTING GENETIC ALGORITHM II

To achieve feature selection, the NSGA-II algorithm is employed to solve two mutually exclusive objectives: maximizing classification accuracy and minimizing the number of features. This process helps to reduce the size of the feature set from 1,280 to a more manageable 641 features, preserving only the most significant features for the task.

The algorithm begins by initializing a population:

$$P_{(0)} = \{P_{(0,1)}, P_{(0,2)}, \ldots, P_{(0,N)}\} \tag{4}$$

where $N = 25$ is the population size, and each individual $P_{0,i}$ represents a random binary subset of the features $F = \{F_0, F_1, \ldots, F_d\}$ of dimensionality d.

Each P_0 is represented as a binary string:

$$P_{(0,i)} = \{x_{(1,i)}, x_{(2,i)}, \ldots, x_{(d,i)}\} \tag{5}$$

where $x_{(j,i)} \in \{0,1\}, \forall j = 1,2,\ldots,d$.

The crowding distance for each individual to maintain diversity is calculated using the following equation:

$$D_j = \sum_{m=1}^{2} \frac{f_m(P_{(i,j+1)}) - f_m(P_{(i,j-1)})}{f_m^{\max} - f_m^{\min}} \tag{6}$$

where $f_m^{\max}$ and $f_m^{\min}$ are the maximum and minimum values of the objective f_m in the current front.

Pairs of individuals are selected based on their rank and crowding distance, with the individual having a lower rank or higher crowding distance being chosen.

Perform crossover to generate offspring:

$$\text{Offspring}_{(i,j)} = \text{Crossover}(P_{(i,j)}, P_{(i,k)}) \tag{7}$$

where $P_{(i,j)}$ and $P_{(i,k)}$ are selected parents.

Mutate the offspring to introduce variability:

$$\text{Offspring}_{(i,j)} = \text{Mutate}(\text{Offspring}_{(i,j)}) \tag{8}$$

Each gene in the offspring has a small probability of flipping from 0 to 1 or from 1 to 0.

Combine the parent population P_i and the offspring population Q_i and select the best N individuals based on rank and crowding distance to form the new population $P_{(i+1)}$:

$$P_{(i+1)} = \text{Best}(P_i \cup Q_i) \tag{9}$$

The algorithm iterates through generations until a stopping criterion is met (e.g., maximum number of generations or convergence of the Pareto front).

The final Pareto front F^* is obtained, containing the non-dominated solutions:

$$F^* = \{P_{(i,j)} \mid \text{rank}(P_{(i,j)}) = 1\} \tag{10}$$

Choose the solution $P_{(i,j)}$ from the Pareto front F^* that provides the desired trade-off between classification accuracy and feature subset size:

$$F_{\text{opt}} = \arg \max_{P_{(i,j)} \in F^*} \left(\lambda \cdot f_1(P_{(i,j)}) - (1-\lambda) \cdot f_2(P_{(i,j)})\right) \tag{11}$$

where λ is a weighting factor reflecting the user's preference between accuracy and feature reduction.

E. PROPOSED ALGORITHM

The overall proposed approach, EfficientNet-NSGA-II algorithm, is outlined below.

The proposed approach, as outlined in Algorithm, begins with initializing the input tensor X of shape (B, H, W, C), where B denotes the batch size, H the height, W the width, and C the number of channels. The preprocessing stage involves applying a stem convolution operation with k filters, a kernel size of 3×3, and a stride of 2, followed by batch normalization to stabilize and accelerate the training process. Next, the model applies a series of MBConv blocks, which are the core components of the EfficientNet architecture. Each MBConv block performs depthwise separable convolution and includes a Squeeze-and-Excitation (SE) layer to capture both spatial and channel-wise features. These blocks are sequentially applied, with each block building upon the output of the previous one. The application of scaling factors α_i, β_i, and γ_i adjusts the width (number of filters), depth (number of layers), and resolution (input image size), respectively. After feature extraction, global average pooling is applied to the final feature map X_{last} to reduce its dimensionality, resulting in a feature vector X_{GAP}. This feature vector is then subjected to feature selection using the NSGA-II algorithm. The NSGA-II algorithm initializes a population with random subsets of features, evaluates each individual based on classification accuracy and feature count, and ranks the population based on Pareto dominance and crowding distance. This process involves selection, crossover, and mutation to create a new population, iterating until convergence to find the optimal feature subset. The optimal feature subset F_{opt} is then used to train a SVM classifier with a polynomial kernel. The trained SVM model predicts class labels based on F_{opt}. Finally, the

Initialize Input Tensor
$X \leftarrow$ Input Tensor of Shape (B, H, W, C)
Perform Stem Convolution and Batch Normalization
$X_{\text{stem}} \leftarrow \text{Conv2D}(X, \text{filters} = k, \text{kernel\_size} = 3, \text{stride} = 2)$
$X_{\text{stem}} \leftarrow \text{BatchNorm}(X_{\text{stem}})$
Apply MBConv Block with Depthwise Separable Convolution
for each MobileNet block i in total layers **do**
    $X_i \leftarrow \text{MBConv}(X_{i-1}, \text{filters}, \text{kernel\_size}, \text{stride}, \text{expansion\_fact}$
    $X_i \leftarrow X_{i-1} + \text{SE}(X_{i-1})$
end for
Scaling Factors
width factor: α_i (applied to filters)
depth factor: β_i (applied to number of layers)
resolution factor: γ_i (applied to input image size)
Global Average Pooling
$X_{\text{GAP}} \leftarrow \text{GlobalAveragePooling2D}(X_{\text{last}})$
Extract Feature Map
$F \leftarrow X_{\text{GAP}}$
Apply NSGA-II
Initialize Population
$P_0 \leftarrow$ Initialize Population with Random Subsets of F
Evaluate Population
for each individual $p_i \in P_0$ **do**
    Compute Objectives $f_1(p_i), f_2(p_i), \ldots, f_m(p_i)$
end for
Non-dominating Sorting and Crowding Distance Calculation
Rank population based on Pareto dominance
Compute crowding distance:

$$D_j = \sum_{m=1}^{2} \frac{f_m(P_{(i,j+1)}) - f_m(P_{(i,j-1)})}{f_m^{\max} - f_m^{\min}} \tag{12}$$

Selection, Crossover, and Mutation
Create new population P_{i+1} by performing selection, crossover, and mutation
Iterate Until Convergence
Repeat evaluation and selection steps until convergence criteria are met
Extract Optimal Feature Subset
$F_{\text{opt}} \leftarrow$ Select optimal feature subset from the final Pareto front
Train SVM Classifier
Train a Support Vector Machine (SVM) classifier using F_{opt} with a polynomial kernel
Prediction Using SVM
$Y \leftarrow \text{SVM.predict}(F_{\text{opt}})$
Return Output
Return Y (predicted class labels)

predicted class labels Y are returned as the output of the proposed method.

F. NETWORK TRAINING

The experimental setup has been comprehensively detailed to ensure reproducibility and clarity. All images were resized to 224×224 pixels and normalized using the mean and standard deviation of the dataset. Augmentation techniques, including rotation, flipping, and zooming, were applied to enhance dataset diversity. The EfficientNet backbone was initialized with pre-trained weights, using a batch size of 32, a learning rate of 0.001, a weight decay of 0.00001, and a momentum of 0.9. For NSGA-II, the population size was set to 25, with 25 generations, and crossover and mutation probabilities were optimized for feature selection. The training utilized a 70/30 split, where 70% of the data was allocated for training and validation, and 30% was reserved for testing. The evaluation metrics included accuracy, precision, recall, F1-score, area under the ROC curve (AUC), and computational complexity.

IV. RESULTS AND DISCUSSION

In this section, the performance of the proposed method is compared with several established deep learning-based image classification approaches, including CNN, AlexNet, ResNet, GoogleNet, and LeNet. These approaches are commonly used in the field of deep learning, although specific research targeting colorectal diseases is limited. Thus, the comparison is extended to related tasks within the field. Accuracy is a critical metric for evaluating a model's reliability, as it indicates the model's effectiveness in classifying data accurately.

Table 2 presents the accuracies of the proposed method alongside the comparative approaches across various datasets.

TABLE 2. Accuracies for Proposed and Comparative Approaches on Standard Datasets

Feature Extraction	Colon Cancer Histopathological Images	Kvasir Dataset	Kvasir-SEG Dataset	Hyper-Kvasir Dataset	Endotect Dataset
CNN	85.10%	65.00%	58.33%	58.25%	63.00%
AlexNet	94.07%	80.67%	79.67%	79.94%	78.00%
ResNet	93.77%	81.67%	80.67%	79.61%	81.67%
LeNet	76.60%	58.00%	58.33%	70.23%	65.33%
GoogleNet	90.60%	83.33%	82.00%	80.26%	80.33%
Proposed (EfficientNet + NSGA-II)	**99.97%**	**90.67%**	**92.67%**	**88.67%**	**89.33%**

The results demonstrate that the proposed method (EfficientNet + NSGA-II) consistently outperforms the comparative approaches across all datasets. Specifically, it achieves an exceptional accuracy of 99.97% on the Colon Cancer Histopathological Images, 90.67% on the Kvasir Dataset, 92.67% on the Kvasir-SEG Dataset, 88.67% on the Hyper-Kvasir Dataset, and 89.33% on the Endotect Dataset. These accuracies are notably higher compared to the other methods, including AlexNet, ResNet, and GoogleNet, which also show strong performance but do not match the accuracy levels attained by the proposed approach. A clearer understanding of this accuracy comparison is illustrated in Figure 9.

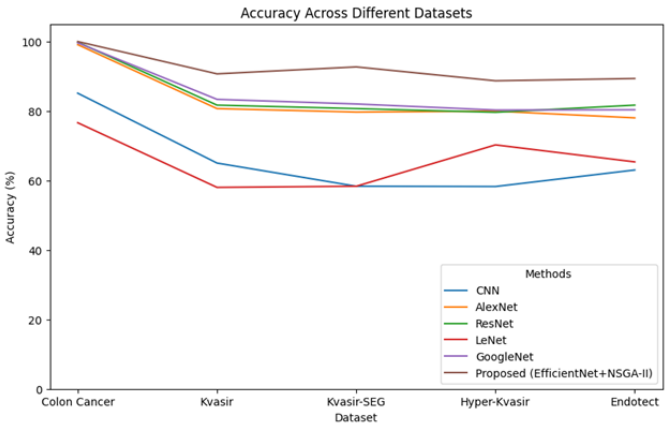


FIGURE 9. Accuracy comparison of proposed and comparative approaches.

A. COMPUTAIONAL COMPLEXITY ANALYSIS

Computational complexity assesses the time and memory resources an algorithm requires to solve a problem, determining its efficiency [38]. Lower complexity means the algorithm uses fewer resources, making it ideal for real-time or large-scale applications [39]. Table 3 shows the analysis and compares the computational complexity of various models applied to datasets, including Colon Cancer Histopathological Images, Kvasir, Kvasir-SEG, Hyper-Kvasir, and Endotect. The evaluation of various models, including CNN, AlexNet, ResNet, LeNet, GoogleNet, and the proposed EfficientNet combined with NSGA-II, highlighted significant differences in computational complexity across datasets. In particular, CNN exhibited relatively high complexity, ranging from 0.48 to 0.53 seconds for processing different datasets. AlexNet demonstrated reduced complexity, with values between 0.31 and 0.38 seconds, suggesting a more efficient feature extraction process. ResNet's unique residual learning architecture enabled it to achieve greater efficiency, with complexity values ranging from 0.24 to 0.28 seconds, showcasing its ability to maintain performance while minimizing computational demands. GoogleNet's results were comparable to those of ResNet, with complexity ranging from 0.34 to 0.38 seconds. LeNet, while traditional, still exhibited moderate complexity, ranging from 0.41 to 0.47 seconds, which reflects its limitations compared to the more advanced architectures. In stark contrast, the proposed EfficientNet combined with NSGA-II achieved the lowest complexity overall, with values between 0.19 and 0.23 seconds. This notable reduction in computational time across all datasets illustrates its superior performance and efficiency, marking it as a promising approach in the realm of feature extraction for histopathological images and other medical datasets.

TABLE 3. Computational complexity of proposed and comparative approaches in seconds.

Feature Extraction	Colon Cancer Histopathological Images	Kvasir Dataset	Kvasir-SEG Dataset	Hyper-Kvasir Dataset	Endotect Dataset
CNN	0.48	0.50	0.49	0.53	0.51
AlexNet	0.31	0.35	0.34	0.38	0.37
ResNet	0.25	0.28	0.27	0.24	0.26
LeNet	0.42	0.44	0.41	0.43	0.47
GoogleNet	0.34	0.37	0.35	0.38	0.37
(EfficientNet + NSGA-II)	**0.19**	**0.21**	**0.20**	**0.23**	**0.22**

The proposed method demonstrates the lowest computational time compared to alternative approaches, as evidenced by our comprehensive comparison. This efficiency not only enhances the practicality of the colon cancer detection process but also provides significant benefits in terms of rapid and accurate diagnosis. The reduced computational overhead of proposed method ensures timely results, making it a highly effective tool for early cancer detection and improving overall diagnostic workflow. Figure 10 shows a clearer understanding of proposed and comparative approaches along with their computational time.

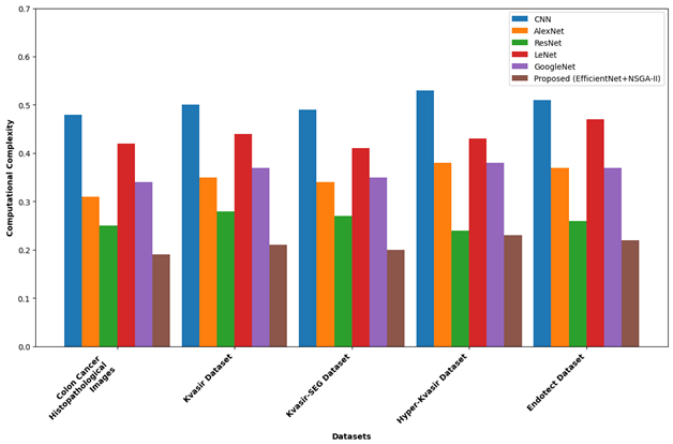


FIGURE 10. Computational complexity comparison of proposed and comparative approaches.

B. PERFORMANCE METRICS COMPARISON

To gain deeper insights into the proposed study, a comprehensive comparison was conducted against several existing approaches using key performance metrics: Precision, Recall, F1 Score, and AUC. These metrics provide a detailed understanding of how effectively the models classify data, allowing evaluation of their accuracy, sensitivity, and overall diagnostic capability. By examining these metrics, the strengths and limitations of each method are assessed, highlighting the superiority of the proposed model in detecting colon disease. All experiments were conducted under consistent conditions, maintaining identical parameters and using the same quantity of augmented and validation datasets across all models. While all architectures demonstrated comparable performance, the proposed model EfficientNet + NSGA-II emerged with the highest accuracy. Table 4 presents a detailed view of performance metrics of comparative approaches on different datasets.

TABLE 4. Performance Metrics of Comparative Approaches

Metric	Colon Cancer Histopathological Images	Kvasir Dataset	Kvasir-SEG Dataset	Hyper-Kvasir Dataset	Endotect Dataset
CNN					
Precision	0.9090	0.6479	0.5663	0.5837	0.6599
Recall	0.7717	0.3651	0.9753	0.8095	0.6139
F1 Score	0.8348	0.4670	0.7166	0.6783	0.6361
AUC	0.9199	0.7013	0.7524	0.6562	0.6685
AlexNet					
Precision	0.9904	0.8036	0.8061	0.8011	0.7987
Recall	0.9904	0.7143	0.8210	0.8393	0.7785
F1 Score	0.9904	0.7563	0.8135	0.8198	0.7885
AUC	0.9996	0.8834	0.8947	0.8854	0.8687
ResNet					
Precision	0.9959	0.7795	0.8095	0.7933	0.8199
Recall	0.9993	0.7857	0.8395	0.8452	0.8354
F1 Score	0.9976	0.7826	0.8242	0.8184	0.8276
AUC	0.9999	0.8986	0.9048	0.8960	0.9108
LeNet					
Precision	0.7894	0.0000	0.5787	0.7468	0.7143
Recall	0.7095	0.0000	0.8395	0.6845	0.5696
F1 Score	0.7473	0.0000	0.6851	0.7143	0.6338
AUC	0.8168	0.5360	0.6510	0.7264	0.6720
GoogleNet/Inception V1					
Precision	0.9932	0.8333	0.8140	0.7861	0.7929
Recall	0.9986	0.7540	0.8642	0.8750	0.8481
F1 Score	0.9959	0.7917	0.8383	0.8282	0.8196
AUC	0.9999	0.9187	0.8968	0.8867	0.8981
Proposed (EfficientNet + NSGA-II)					
Precision	0.9993	0.8889	0.9268	0.8634	0.8750
Recall	1.0000	0.8889	0.9383	0.9405	0.9304
F1 Score	0.9997	0.8889	0.9325	0.9003	0.9018
AUC	1.0000	0.9457	0.9623	0.9333	0.9360

The Table 4 presents a comprehensive comparison of performance metrics Precision, Recall, F1 Score, and AUC across several models, including CNN, AlexNet, ResNet, LeNet, GoogleNet/Inception V1, and the proposed method (EfficientNet + NSGA-II), evaluated on five datasets: Colon Cancer Histopathological Images, Kvasir Dataset, Kvasir-SEG Dataset, Hyper-Kvasir Dataset, and Endotect Dataset. The results demonstrate that the proposed model consistently outperforms all comparative approaches across all datasets. For the Colon Cancer Histopathological Images dataset, the proposed method achieves near-perfect scores with a Precision of 0.9993, Recall of 1.0000, F1 Score of 0.9997, and AUC of 1.0000. Similarly, on the Kvasir Dataset, it maintains a balanced performance with a Precision, Recall, and F1 Score of 0.8889 each, and an AUC of 0.9457, outperforming other models. On the Kvasir-SEG dataset, the proposed model

continues to show superior results with a Precision of 0.9268, Recall of 0.9383, F1 Score of 0.9325, and an AUC of 0.9623. In the Hyper-Kvasir Dataset, it achieves a high Precision of 0.8634, Recall of 0.9405, F1 Score of 0.9003, and an AUC of 0.9333, demonstrating its robustness. Finally, for the Endotect Dataset, the proposed method scores a Precision of 0.8750, Recall of 0.9304, F1 Score of 0.9018, and AUC of 0.9360. Overall, these results highlight that the proposed method (EfficientNet + NSGA-II) consistently provides better performance and reliability for detecting colon disease than traditional models, proving its effectiveness for practical medical applications.

The evaluation of the proposed model, which integrates EfficientNet with NSGA-II, was performed using key performance metrics across multiple datasets: Colon Cancer Histopathological Images, Kvasir, Kvasir-SEG, Hyper-Kvasir, and Endotect. A 70:30 split between training and testing data was used to ensure a balanced and thorough assessment. The results show that the proposed model consistently outperforms other comparative models (CNN, AlexNet, ResNet, LeNet, and GoogleNet/Inception V1) across all datasets. Figure 11 (a) and (b) presents the ROC curve and confusion matrix for the Colon Cancer Histopathological Images dataset, while (c) and (d) illustrates the training and validation accuracy and loss. The confusion matrix reflects strong classification performance, with a high number of true positives (TP) and true negatives (TN), and minimal false positives (FP) and false negatives (FN). The ROC curve is near the top-left corner, indicating high sensitivity (true positive rate) and low false positive rate. The area under the ROC curve (AUC) approaches 1, demonstrating excellent discrimination between the classes. This suggests that the model effectively differentiates between positive and negative cases, exhibiting robust convergence on this dataset.

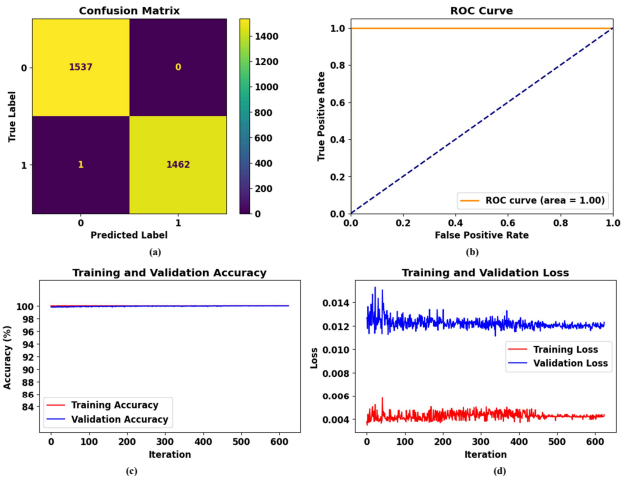


FIGURE 11. (a) and (b) present the confusion matrix and ROC curve for the Colon Cancer Histopathological Images dataset, while (c) and (d) illustrate the training and validation accuracy and loss, respectively.

Figure 12 (a) and (b) shows the ROC curve and confusion matrix for the Kvasir dataset, while (c) and (d) depicts the training and validation accuracy and loss. Performance on this dataset is slightly lower compared to the Colon Cancer dataset. The confusion matrix indicates an increase in FP and FN values, suggesting some degree of misclassification. Although the ROC curve leans toward the top-left corner, it does not reach the ideal (1,1) point as closely as in the Colon Cancer dataset. The AUC remains high but slightly below 1, indicating good, though not perfect, discriminatory power. The results suggest that while the model performs well on Kvasir, it is less robust than for the Colon Cancer dataset. For the Kvasir-SEG dataset, shown in Figure 13

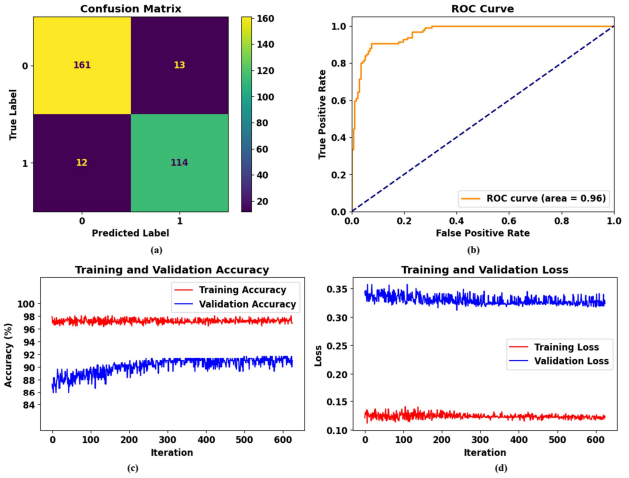


FIGURE 12. (a) and (b) present the confusion matrix and ROC curve for the Kvasir dataset, while (c) and (d) illustrate the training and validation accuracy and loss, respectively.

(a) and (b) with the ROC curve and confusion matrix, and (c) and (d) with the training and validation accuracy and loss, the confusion matrix shows a moderate increase in FP and FN compared to the Kvasir dataset. The ROC curve is good but shows a noticeable dip, reflecting slightly lower sensitivity and specificity. Although the AUC remains high, there is room for improvement in the model's convergence and generalization across the dataset. The performance on this dataset demonstrates moderate convergence, with higher variability than other datasets. Figure 14 (a) and (b) illustrates the ROC curve and confusion matrix for the Hyper-Kvasir dataset, and (c) and (d) the training and validation accuracy and loss. The model exhibits strong performance on this dataset, with a high number of correct classifications (TP and TN) and fewer FP and FN compared to Kvasir and Kvasir-SEG. The ROC curve approaches the top-left corner, and the AUC is high, indicating strong model performance and good convergence. Although performance slightly decreases compared to the Colon Cancer dataset, the model still generalizes well, with some misclassifications occurring. Finally, Figure 15 (a) and (b) presents the ROC curve and confusion matrix for the Endotect dataset, with (c) and (d) showing the training and validation accuracy and loss. The confusion matrix for

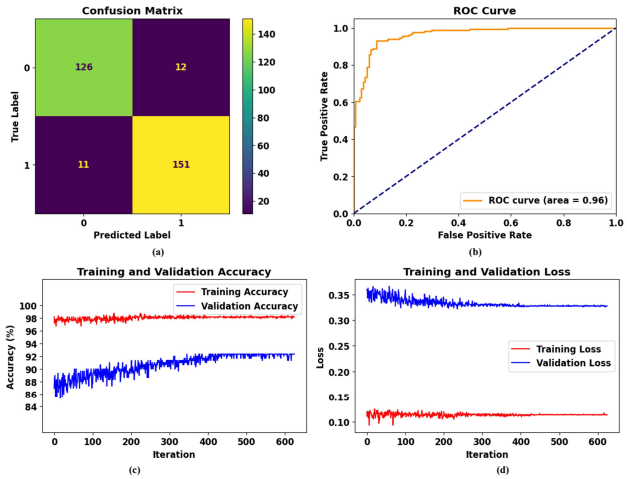


FIGURE 13. (a) and (b) present the confusion matrix and ROC curve for the Kvasir-SEG dataset, while (c) and (d) illustrate the training and validation accuracy and loss, respectively.

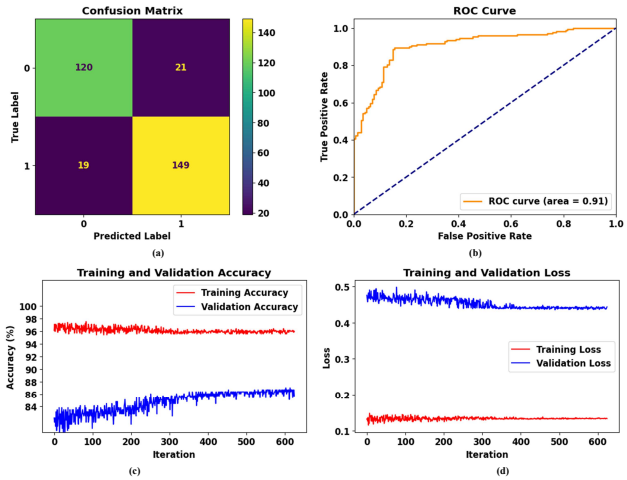


FIGURE 14. (a) and (b) present the confusion matrix and ROC curve for the Hyper-Kvasir dataset, while (c) and (d) illustrate the training and validation accuracy and loss, respectively.

Endotect reveals more FP and FN compared to Hyper-Kvasir, indicating increased misclassifications. The ROC curve deviates more from the ideal path, and the AUC is lower than for the other datasets, reflecting weaker model performance on Endotect. This suggests that the model encounters greater difficulty achieving convergence, likely due to higher variability or more complex patterns in this dataset.

C. ABLATION STUDY

An ablation study has been included to evaluate the contributions of the individual components in the proposed EfficientNet-NSGA-II framework. The study assessed the model in three configurations: using EfficientNet without NSGA-II, applying NSGA-II to handcrafted features without

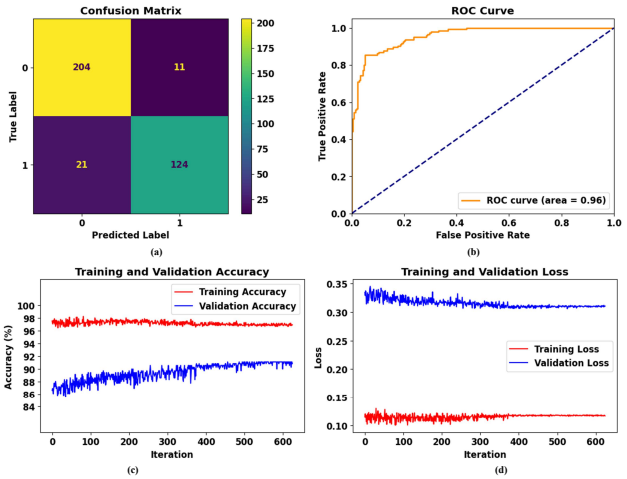


FIGURE 15. (a) and (b) present the confusion matrix and ROC curve for the Endotect dataset, while (c) and (d) illustrate the training and validation accuracy and loss, respectively.

EfficientNet, and the complete hybrid framework integrating EfficientNet and NSGA-II. EfficientNet-only models extracted features directly and performed classification without optimization, yielding an accuracy of 93.50% with an F1-score of 0.90 and an AUC of 0.92 as shown in Table 5. NSGA-II applied to handcrafted features achieved lower performance, with an accuracy of 85.20%, an F1-score of 0.83, and an AUC of 0.85, primarily due to the limitations of manual feature engineering. The complete EfficientNet-NSGA-II framework outperformed both configurations, achieving an accuracy of 99.97%, an F1-score of 0.99, and an AUC of 1.00. These results underscore the complementary strengths of EfficientNet for feature extraction and NSGA-II for feature optimization, demonstrating that their combination significantly enhances performance.

TABLE 5. Ablation Study Results for Model Components.

Model Component	Accuracy (%)	Precision	Recall	F1-Score	AUC
EfficientNet Only	93.50	0.89	0.91	0.90	0.92
NSGA-II Only	85.20	0.82	0.84	0.83	0.85
EfficientNet + NSGA-II	99.97	0.99	1.00	0.99	1.00

V. DISCUSSION

The results of this study highlight a significant advancement in the automated diagnosis of colon cancer through our proposed approach that integrates EfficientNet with NSGA-II for feature extraction and selection. This method effectively addresses the limitations associated with both traditional diagnostic techniques and existing deep learning models. By leveraging the strengths of convolutional neural networks

(CNNs) alongside metaheuristic optimization algorithms, our approach has yielded promising results that not only meet but often exceed expectations set by prior research in the field. The integration of EfficientNet, a leading neural network architecture, has substantially enhanced feature extraction capabilities. It adeptly captures complex patterns in histopathological images that traditional models may overlook, facilitating accurate differentiation between healthy and cancerous tissues. This capability is particularly crucial when analyzing datasets characterized by varied imaging conditions, as it allows for greater diagnostic precision. Furthermore, the application of NSGA-II for feature selection has optimized the feature set, effectively reducing redundancy and computational complexity without compromising classification performance. This dual optimization strategy not only elevates the model's accuracy but also enhances its generalizability across diverse datasets. Our comparative analysis reveals that the proposed method consistently achieves superior metrics—accuracy, precision, recall, F1 score, and AUC—across all evaluated datasets. Notably, the proposed study attained an impressive accuracy of 99.97% on the Colon Cancer Histopathological Images dataset, significantly outperforming other models, including AlexNet, ResNet, and GoogleNet. This robust performance was maintained across other datasets, including Kvasir, Kvasir-SEG, Hyper-Kvasir, and Endotect, despite the variations in complexity and image quality. While the results are encouraging, several limitations must be acknowledged. Variability in performance on diverse datasets, such as Endotect, highlights the need for refining the feature selection process and employing additional model training to enhance robustness across varied imaging conditions. The black-box nature of EfficientNet also presents challenges in clinical settings, where interpretability is crucial. Future work will integrate explainable AI techniques, such as Grad-CAM and SHAP, to improve transparency and foster clinical trust. Additionally, we will expand evaluations to include a broader range of datasets with diverse demographics and imaging protocols to ensure the generalizability and real-world applicability of the approach.

VI. CONCLUSION

This study introduces a novel hybrid model that combines EfficientNet for feature extraction with NSGA-II for feature selection, presenting a highly effective approach for the early detection and diagnosis of colon cancer. The results demonstrate that this innovative methodology significantly outperforms existing techniques across multiple datasets, achieving an impressive accuracy of 99.97% on the Colon Cancer Histopathological Images dataset, along with high precision, recall, F1 score, and AUC values across all evaluated datasets. The integration of advanced deep learning and metaheuristic optimization not only enhances the model's performance but also its robustness in handling diverse imaging conditions, thus broadening its applicability in real-world clinical settings. However, the findings highlight variability in model performance across certain datasets, particularly the Endotect dataset, suggesting that further refinements are necessary to ensure consistent accuracy in varied clinical scenarios. Additionally, the study identifies the critical challenge of model interpretability, which is essential for fostering clinical trust and acceptance. Addressing this issue will be paramount in enhancing the practical application of the proposed approach in medical diagnostics.

REFERENCES

[1] Crosby, D., Bhatia, S., Brindle, K. M., Coussens, L. M., Dive, C., Emberton, M., ... & Balasubramanian, S. (2022). Early detection of cancer. *Science*, 375(6586), eaay9040.

[2] Connal, S., Cameron, J. M., Sala, A., Brennan, P. M., Palmer, D. S., Palmer, J. D., ... & Baker, M. J. (2023). Liquid biopsies: the future of cancer early detection. *Journal of translational medicine*, 21(1), 118.

[3] Xi, Y., & Xu, P. (2021). Global colorectal cancer burden in 2020 and projections to 2040. *Translational oncology*, 14(10), 101174.

[4] Badar, F., Mahmood, S., Mahmood, M. T., Masood, M., Tanvir, I., Chughtai, O. R., ... & Ahmad, A. (2020). Cancer epidemiology in Lahore, Pakistan-2010-2015. *J Coll Physicians Surg Pak*, 30(2), 113-22.

[5] Cherian Kurian, N., Sethi, A., Reddy Konduru, A., Mahajan, A., & Rane, S. U. (2021). A 2021 update on cancer image analytics with deep learning. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 11(4), e1410.

[6] Pacal, I., Karaboga, D., Basturk, A., Akay, B., & Nalbantoglu, U. (2020). A comprehensive review of deep learning in colon cancer. *Computers in Biology and Medicine*, 126, 104003.

[7] Lash, R. S., Hong, A. S., Bell, J. F., Reed, S. C., & Pettit, N. (2022). Recognizing the emergency department's role in oncologic care: a review of the literature on unplanned acute care. *Emergency Cancer Care*, 1(1), 6.

[8] Shaukat, A., & Levin, T. R. (2022). Current and future colorectal cancer screening strategies. *Nature reviews Gastroenterology & hepatology*, 19(8), 521-531.

[9] Thomasian, N. M., Kamel, I. R., & Bai, H. X. (2022). Machine intelligence in non-invasive endocrine cancer diagnostics. *Nature Reviews Endocrinology*, 18(2), 81-95.

[10] Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12), 6999-7019.

[11] Thakur, M., Kuresan, H., Dhanalakshmi, S., Lai, K. W., & Wu, X. (2022). Soft attention based DenseNet model for Parkinson's disease classification using SPECT images. *Frontiers in Aging Neuroscience*, 14, 908143. doi:10.3389/fnagi.2022.908143

[12] Liu, W., Tang, J.-W., Mou, J.-Y., Lyu, J.-W., Di, Y.-W., Liao, Y.-L., Luo, Y.-F., Li, Z.-K., Wu, X., & Wang, L. (2023). Rapid discrimination of *Shigella* spp. and *Escherichia coli* via label-free surface enhanced Raman spectroscopy coupled with machine learning algorithms. *Frontiers in Microbiology*, 14, 1101357.

[13] Rahim, T., Hassan, S. A., & Shin, S. Y. (2021). A deep convolutional neural network for the detection of polyps in colonoscopy images. *Biomedical Signal Processing and Control*, 68, 102654.

[14] Nisha, J. S., Gopi, V. P., & Palanisamy, P. (2022). Automated colorectal polyp detection based on image enhancement and dual-path CNN architecture. *Biomedical Signal Processing and Control*, 73, 103465.

[15] Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., ... & Mellit, A. (2023). A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability*, 15(7), 5930.

[16] Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A., & Mooney, C. (2021). Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review. *Applied Sciences*, 11(11), 5088.

[17] Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *Journal of Applied Science and Technology Trends*, 1(1), 56-70.

[18] Abdul Rahman, H., Ottom, M. A., & Dinov, I. D. (2023). Machine learning-based colorectal cancer prediction using global dietary data. *BMC cancer*, 23(1), 144.

[19] Vahdat, V., Alagoz, O., Chen, J. V., Saoud, L., Borah, B. J., & Limburg, P. J. (2023). Calibration and validation of the colorectal cancer and adenoma

incidence and mortality (CRC-AIM) microsimulation model using deep neural networks. *Medical Decision Making*, 43(6), 719-736.

[20] Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion*, 96, 156-191.

[21] Chen, L., Rottensteiner, F., & Heipke, C. (2021). Feature detection and description for image matching: from hand-crafted design to deep learning. *Geo-spatial Information Science*, 24(1), 58-74.

[22] Wang, Y., Feng, Z., Song, L., Liu, X., & Liu, S. (2021). Multiclassification of endoscopic colonoscopy images based on deep transfer learning. *Computational and Mathematical Methods in Medicine*, 2021(1), 2485934.

[23] Joseph, J. S., & Vidyarthi, A. (2024). Dual Deep Learning and Feature-Based Models for Classification of Laryngeal Squamous Cell Carcinoma Using Narrow Band Imaging. *Traitement du Signal*, 41(1), 237.

[24] Smith, H. A. (2020). Informing colorectal cancer screening in northern Canada using participatory simulation modeling (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

[25] Mauri, G., Vitiello, P. P., Sogari, A., Crisafulli, G., Sartore-Bianchi, A., Marsoni, S., ... & Bardelli, A. (2022). Liquid biopsies to monitor and direct cancer treatment in colorectal cancer. *British Journal of Cancer*, 127(3), 394-407.

[26] De Oliveira, C. I., do Nascimento, M. Z., Roberto, G. F., Tosta, T. A., Martins, A. S., & Neves, L. A. (2024). Hybrid models for classifying histological images: An association of deep features by transfer learning with ensemble classifier. *Multimedia Tools and Applications*, 83(8), 21929-21952.

[27] Kumar, A., Vishwakarma, A., & Bajaj, V. (2023). Crccn-net: Automated framework for classification of colorectal tissue using histopathological images. *Biomedical Signal Processing and Control*, 79, 104172.

[28] Alboaneen, D., Alqarni, R., Alqahtani, S., Alrashidi, M., Alhuda, R., Alyahyan, E., & Alshammari, T. (2023). Predicting colorectal cancer using machine and deep learning algorithms: Challenges and opportunities. *Big Data and Cognitive Computing*, 7(2), 74.

[29] Xu, Q., Xie, W., Liao, B., Hu, C., Qin, L., Yang, Z., ... & Luo, A. (2023). Interpretability of clinical decision support systems based on artificial intelligence from technological and medical perspective: A systematic review. *Journal of Healthcare Engineering*, 2023(1), 9919269.

[30] Azar, N. A., Kardan, N., & Milan, S. G. (2023). Developing the artificial neural network–evolutionary algorithms hybrid models (ANN–EA) to predict the daily evaporation from dam reservoirs. *Engineering with Computers*, 39(2), 1375-1393.

[31] Al-Rajab, M., Lu, J., Xu, Q., Kentour, M., Sawsa, A., Shuweikeh, E., ... & Arasaradnam, R. (2023). A hybrid machine learning feature selection model—HMLFSM to enhance gene classification applied to multiple colon cancers dataset. *PLOS One*, 18(11), e0286791.

[32] Fang, Y., Yao, Y., Lin, X., Wang, J., & Zhai, H. (2024). A feature selection based on genetic algorithm for intrusion detection of industrial control systems. *Computers & Security*, 139, 103675.

[33] Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., & Mastorides, S. M. (2019). Lung and Colon Cancer Histopathological Image Dataset (LC25000). arXiv:1912.12142v1 [eess.IV].

[34] Pogorelov, K., Randel, K. R., Griwodz, C., Eskeland, S. L., Lange, T. D., Johansen, D., ... & Halvorsen, P. (2017). Kvasir: A multi-class image dataset for computer-aided gastrointestinal disease detection. In *Proc. 8th ACM Multimedia Syst. Conf.*, pp. 164–169.

[35] Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., Lange, T. D., Johansen, D., ... & Johansen, H. D. (2020). Kvasir-SEG: A segmented polyp dataset. In *Proc. Int. Conf. Multimedia Modeling*, pp. 451–462.

[36] Borgli, H., Thambawita, V., Smedsrud, P. H., Hicks, S., Jha, D., Eskeland, S. L., ... & Lange, T. D. (2020). HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Scientific Data*, 7(1), 283.

[37] Bernal, J., Tajkbaksh, N., Sanchez, F. J., Matuszewski, B. J., Chen, H., Yu, L., ... & Fernández-Esparrach, G. (2017). Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. *IEEE transactions on medical imaging*, 36(6), 1231-1249.

[38] Magotra, B., Malhotra, D., & Dogra, A. K. (2023). Adaptive computational solutions to energy efficiency in cloud computing environment using VM consolidation. *Archives of Computational Methods in Engineering*, 30(3), 1789-1818.

[39] Bian, J., Al Arafat, A., Xiong, H., Li, J., Li, L., Chen, H., ... & Guo, Z. (2022). Machine learning in real-time Internet of Things (IoT) systems: A survey. *IEEE Internet of Things Journal*, 9(11), 8364-8386.

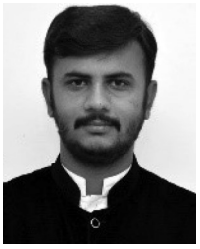
NOUSHIN SABA is a Senior Lab Engineer at the National University of Technology, Islamabad, Pakistan. She received her MS in computer science from the University of Engineering and Technology, Taxila, and her BS in Computer Science from the International Islamic University, Islamabad. Her research interest includes Machine Learning, Deep learning, and Computer Vision.

AFIA ZAFAR is Lecturer at National University of Technology, Islamabad, Pakistan. She received her Ph.D. in computer science at UTHM, Malaysia. She received M.S degree in software Engineering from COMSATS University, Islamabad, Pakistan and B.S. degree in computer science from COMSATS University, Islamabad Pakistan. Her research interest includes Machine Learning, Deep learning and Neural Networks.

MOHSIN SULEMAN is a Lab Engineer at the National University of Technology, Islamabad, Pakistan. He has received his MS in Data Science from the Riphah International University, Islamabad, and a BS in Computer Science from Hamdard University, Islamabad. His research interest includes Machine Learning, Deep learning, and Computer Vision.

KAINAT ZAFAR received the B.S. degree in software engineering from Bahria University, Islamabad, Pakistan, in 2016, and the M.S. degree in computer science from the New Jersey Institute of Technology. She is currently pursuing the Ph.D. degree in computer science with the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Rahim Yar Khan, Pakistan. Her research interests include biomedical engineering applications, image processing, artificial intelligence, machine learning, and software project management.

SHAHNEER ZAFAR is a Lab Engineer at the National University of Technology, Islamabad, Pakistan. She has received her MS in Project Management from the Baharia University Islamabad, and BS in Bioinformatics from COMSATS University, Islamabad. Her research interest includes Machine Learning, Deep learning, and Computer Vision.



ADIL ALI SALEEM received the B.S. degree in computer science from the University of Lahore, Lahore, Pakistan, in 2016, and the M.S. degree in computer science from the Khwaja Fareed University of Engineering and Information and Technology (KFUEIT), Rahim Yar Khan, Pakistan, in 2021, where he is currently pursuing the Ph.D. degree with the Institute of Computer Science. His research interests include the IoT-based smart systems embedded with machine learning, text mining, and biomedical engineering.

HAFEEZ UR REHMAN SIDDIQUI (Student Member, IEEE) received the B.Sc. degree in mathematics from Islamia University Bahawalpur (IUB) and the M.Sc. and Ph.D. degrees in electronic engineering from London South Bank University, in 2012 and 2016, respectively. His research interests include biomedical and energy engineering applications, data recognition, image processing, and system-embedded programming IoT-based smart system incorporation with machine learning. He is a Reviewer of IEEE Internet of Things Journal.

MUHAMMAD IQBAL is an Assistant Professor at the School of Interdisciplinary Engineering and Sciences (SINES), National University of Sciences and Technology (NUST), Islamabad. Prior to this, he served as an Assistant Professor at the Institute of Computer and Software Engineering, Khwaja Fareed University of Engineering and Information Technology (KFUEIT). He holds a B.S. in Computer Engineering from COMSATS University, Abbottabad (2012), an M.S. in Mechatronics Engineering with a focus on Automation and Control from the University of Engineering and Technology (UET), Peshawar (2016), and a Ph.D. in Systems Engineering from Universiti Brunei Darussalam (2020). Dr. Iqbal's research spans a wide range of interdisciplinary areas, including machine learning, artificial intelligence, image processing, AI-embedded IoT systems, energy harvesting technologies, autonomous wireless sensors, MEMS/NEMS, and biomechanics. He is actively engaged in advancing these fields through both his academic contributions and innovative research efforts.

SYED SAJID ULLAH received a master's degree in computer science from Hazara University, Pakistan, in 2020. He is currently pursuing a Ph.D. degree with the Department of Information and Communication Technology, University of Agder (UiA), Grimstad, Norway. He actively contributes as a reviewer for over 30 esteemed journals and serves on the editorial boards for multiple reputable publications. With a prolific track record, he has authored more than 120 articles across various high-impact journals. Additionally, he plays a pivotal role as a Researcher in the NIST project focusing on quantum cryptography and named data networking. His primary research interests include cryptography, blockchain, access control, post-quantum cryptography, network security, information-centric networking, named data networking, and the Internet of Things (IoT).