

Se-Resnet: A Novel Method for Gastrointestinal (GI) Diseases Classification from Wireless Capsule Endoscopy (WCE) Images



Panguluri Padmavathi^{ID}, Jonnadula Harikiran^{ID}*

School of Computer Science and Engineering, VIT-AP University, Amaravathi 522237, Andhra Pradesh, India

Corresponding Author Email: harikiran.j@vitap.ac.in

<https://doi.org/10.18280/ts.400404>

ABSTRACT

Received: 14 February 2023

Revised: 26 July 2023

Accepted: 10 August 2023

Available online: 31 August 2023

Keywords:

Gastrointestinal tract diseases, feature selection, wireless capsule endoscopy, disease classification, dataset

The GI tract can develop some medical issues that may require a doctor to evaluate them. These consist of growth anomalies, tissue inflammations, and gastrointestinal issues. In this work, we propose a novel deep-learning (DL) technique to classify the categories of Gastrointestinal Diseases from Wireless Capsule Endoscopy (WCE) images. It has five steps to evaluate. Initially, utilizing the mean filter to remove the noise from given input images. Then extract the features such as shape and position from wireless capsule endoscopy images using the DenseNet-121 technique. To select the features, we utilize the Enhanced Whale Optimization Algorithm (EWOA). Finally, to classify the eight classes of gastrointestinal diseases, we propose a SE-ResNet technique to classify the GI diseases into Ulcerative-colitis, Normal-cecum, Dyed-resection-margins, Esophagitis, Normal-pylorus, Dyed-lifted-polyps, Normal-z-line, Polyps categories with Bald Eagle Search optimization technique to get better accuracy of classification outcomes. In our experiments, we used the Kvasir v2 dataset, and the experiments performed well in terms of recall, precision, accuracy, and f1-score. The performance of the classification technique achieves 99.66% accuracy. The proposed method detects GI disorders on WCE images better than "state-of-the-art" methods while also classifying the items.

1. INTRODUCTION

Gastrointestinal (GI) diseases refer to a group of disorders that affect the organs and structures of the digestive system. The digestive system is responsible for breaking down food, absorbing nutrients, and eliminating waste from the body. GI diseases can affect any part of the digestive tract, including the esophagus, stomach, small intestine, large intestine (colon), rectum, and anus. At the moment, gastrointestinal tract (GI) infections such as cancer, bleeding, ulcer polyps, and Crohn's are relatively common but ulcers and bleeding are common disorders. An American researcher reported that since 2017, 135,430 new cases of gastrointestinal tract infections have been recorded, and since 2011, more than 200,000 new cases have been reported year globally [1-3]. If identified and treated at an early stage, this digestive system infection is treatable.

Food that has been consumed travels to the anus from the mouth through the GI tract, which is a tubular channel. The GI tract is split into two main sections anatomically. The "lower" GI tract is made up of the section from the tiny intestine to the anus, while the "upper" GI tract extends from the mouth to the duodenum [4, 5]. The digestive system's upper portion is primarily in charge of swallowing and food digestion. Before the food is sent to the lower GI tract, it is digested in this area by gastric acids and enzymes. The remainder turns into solid waste, which is then saved in the rectum and expelled from the body through the anus.

The GI tract can develop some medical issues that may require a doctor to evaluate them. These consist of growth anomalies, tissue inflammations, and digestive disorders. For instance, inflammations may be brought by an aberrant

immune response, esophageal lining changes may result from acid reflux, and polyps on the colon lining may form when dividing cells clump together [5-7]. As in the case of sores or ulcers, these abnormalities may be severe in and of themselves. Another danger is the subsequent development of difficulties, such as polyps that could grow into cancer. The GI tract's organs must be examined visually by medical personnel. Endoscopy is a procedure that is typically used to examine the digestive tract. These procedures entail the use of an endoscope and a flexible long tube that is typically connected to a fiber-optic camera and inserted into an opening. A medical professional can now see the GI system due to this, and when the endoscope moves through the GI tract. An endoscope is introduced via the anus during a colonoscopy to examine the colon or large intestine [8, 9]. Narrow band imaging (NBI) and wireless capsule endoscopy (WCE) are further types of endoscopic methods.

The main purpose of WCE, a noninvasive procedure is to give small intestine diagnostic imaging. In the WCE procedure, the patient ingests a capsule with a camera placed inside it. This device travels through the digestive tract, takes pictures, and sends those pictures to an outside receiver. The WCE technology, developed over time by numerous companies to become more precise and effective was first introduced in 1989 by a research group [10, 11]. Because traditional endoscopy is extremely uncomfortable and upsetting for patients, this technique replaced it and simplified the diagnosis for both the patient and the examiner. A camera within the patient constantly records images for eight hours as part of the WCE procedure, sending around 60000 images to the receiver. To determine which frames or pictures contain the virus or

disease, a doctor looks at these images. Because there are so many pictures, it takes a long time for a medical specialist to examine them [12]. This adds more work for the practitioner and could result in a wrong diagnosis of the contaminated area of the intestine. Consequently, during the past ten or so years, the recognition of pictures that include infected regions utilizing different statistical and ML techniques has become a popular research topic. For the specialist to quickly review only the frames that have visual contents of the damaged area and begin the necessary corrective steps, an automatic recognition system analyses the huge volume of pictures and detects the infected frames [13-15]. To overcome issues, we propose a novel DL method to classify the many types of gastrointestinal diseases using WCE images. Initially, removing noise from input pictures using the mean filter. Then, using the DenseNet-121 approach, extract the features from WCE pictures. Then, we employ the Enhanced Whale Optimization Algorithm for feature selection. Finally, to categorize the eight categories of gastrointestinal diseases, we propose the SE-ResNet technique with the Bald Eagle Search optimization technique. The key contributions are:

- To reduce the noises from given input images, we utilize the mean filter.
- Then extract the features such as shape and position from WCE images using the DenseNet-121 technique.
- To select the features, we utilize the Enhanced Whale Optimization Algorithm. Finally, to classify gastrointestinal diseases into eight classes, we propose a deep learning approach of SE-ResNet with Bald Eagle Search optimization technique to get better accuracy of results.
- The performance of experiments using the Kvasir v2 dataset for good performance with precision, accuracy, f1-score, and recall parameters.

The rest of the article is organized as follows. Section 2 includes a listing of relevant material to the paper. Section 3 provides the problem statement. Section 4 discusses the proposed method. The outcomes are presented in Section 5. Finally, Section 7 summarizes the findings.

2. LITERATURE BACKGROUND

For the categorization of biomedical pictures, DL and ML techniques have been widely utilized. These classification

techniques support medical professionals in making precise diagnoses and treatment recommendations. The methods currently used to categorize GI illnesses are reviewed in this section (Table 1). An AI approach is suggested by Öztürk and Özkaya [16] for classifying GI tract picture datasets with a limited amount of labeled information. The foundation of the suggested AI technique is the CNN structure, which is acknowledged as the most effective automatic categorization method currently in use. This method assumes that to categorize unbalanced datasets robustly, a shallowly trained CNN structure needs to be backed by a powerful classifier. The features from every pooling layer in the CNN structure are sent to an LSTM layer for this reason. By integrating all LSTM layers, a categorization is created.

Sharif et al. [17] suggested a method based on the merging of geometric and deep CNN features. Initially, a method called contrast-enhanced color features is utilized to extract disease zones from provided WCE images. Segmented disease regions are used to extract geometric features. Then, based on the Euclidean Fisher Vector, a special VGG19, and VGG16 deep CNN feature fusion is carried out. Unique features and geometric features are combined, and the best features are then chosen using a conditional entropy technique. K-Nearest Neighbor is employed to categorize the final features that have been chosen.

Using endoscopic images, Mohapatra et al. [18] suggested a hybrid EWT and CNN technique for GI illness identification. The primary goal is to identify abnormal GI tract disorders, specifically Ulcerative Colitis, Barrett's, Polyps, Esophagitis, and Hemorrhoids. As a result, only abnormal and normal pictures from the dataset are considered. The initial step includes some procedures, including image scaling, picture augmentation, and picture pre-processing, which aid in preparing the dataset and resolving some issues before other processing. The use of EWT, which helps to break down the pictures into their IMF, which serves as a picture feature pattern extractor is the next stage. The deep network is trained using these extracted patterns. The system is tested and trained on two different levels. Classifying the abnormal and normal pictures is the aim of the first stage of classification. Deliberately removing the normal pictures, aids in the subsequent phase's concentration on the diseased class pictures. Only the aberrant pictures from level one's second level are further categorized into the previously mentioned five distinct illness classifications.

Table 1. The merits and demerits of previous studies on GI

Reference	Method	Merits	Demerits
Öztürk and Özkaya [16]	LSTM	It not only improves the classification performance but also decreases the classification time.	The computation cost is high.
Sharif et al. [17]	KNN	The suggested EFV attributes fusion approach achieved enhanced classification accuracy on individual attributes.	The performance time was increased, because lack of feature selection.
Mohapatra et al. [18]	EWT + CNN	It increased the efficacy of the classification accuracy.	The development of a time-saving, more accurate real-time assistant model for gastroenterologists to identify digestive tract disorders and deliver appropriate treatment to patients.
Ramamurthy et al. [19]	EfficientNet B0	The suggested feature fusion network performed well in terms of inter-class metrics.	Categorization of GI disorders was undertaken using limited datasets.
Haile et al. [20]	SVM	They utilized deep attribute concatenation as a single attribute vector by merging VGGNet and InceptionNet systems as attribute extractors, followed by SVM classification.	The methodologies provided here are intended to classify a small amount of GI disorders in a specific region of the human GI tract.

To categorize various gastrointestinal disorders, Ramamurthy et al. [19] suggested an automated categorization method based on DL. The suggested models are trained using the HyperKvasir tagged pictures dataset. To increase the number of samples for improved generalization, the input images are initially enhanced. Two separate networks, EfficientNet B0 and the suggested Effimix network received these modified samples as input. Dropout regularization and feature concatenation were both used after the features from these two networks were joined. The input gastrointestinal picture sets are divided into 23 classes by the suggested approach.

To create a model for the diagnosis of gastrointestinal diseases, Haile et al. [20] suggested concatenating the retrieved attributes of the InceptionNet and VGGNet networks. To extract features from the provided endoscopic pictures, the DCNN's InceptionNet and VGGNet are trained and deployed. Then, using machine-learning classification methods, these retrieved features are concatenated (KNN, SVM, Random Forest, and Softmax). Utilizing the provided standard dataset, SVM is one of these approaches, outperforming the others in terms of performance.

3. PROBLEM STATEMENT

From the above existing papers, they have some problems

to solve. The size of the experimental dataset is small, so classification performance can't be accurate. The size of the experimental dataset is small. Then, without feature selection performance, it obtained more time to evaluate, and the computation time of simulation performance was higher. To overcome these problems, propose a deep learning technique to classify GI disease accurately. Here, it uses a feature selection technique to select essential features to reduce the computation time.

4. PROPOSED METHODOLOGY

Pathologies that affect the digestive system collectively are referred to as gastrointestinal (GI) illnesses [21]. Polyps, malignancies, infection, Crohn's disease, and diverticulitis are a few of the most prevalent GI conditions. These illnesses are the common causes of death in humans, based to the World Health Organization. The major approaches for examining GI tract illnesses at the moment are diagnostic imaging reviewed by clinicians and laboratory tests. In this paper, we propose a new DL technique to classify the categories of Gastrointestinal Diseases using WCE pictures. It has five steps to perform. Initially, utilizing the mean filter to remove the noise from input pictures. To extract the features from WCE pictures, using the DenseNet-121 technique. Then to select the features, we utilize the Enhanced Whale Optimization Algorithm.

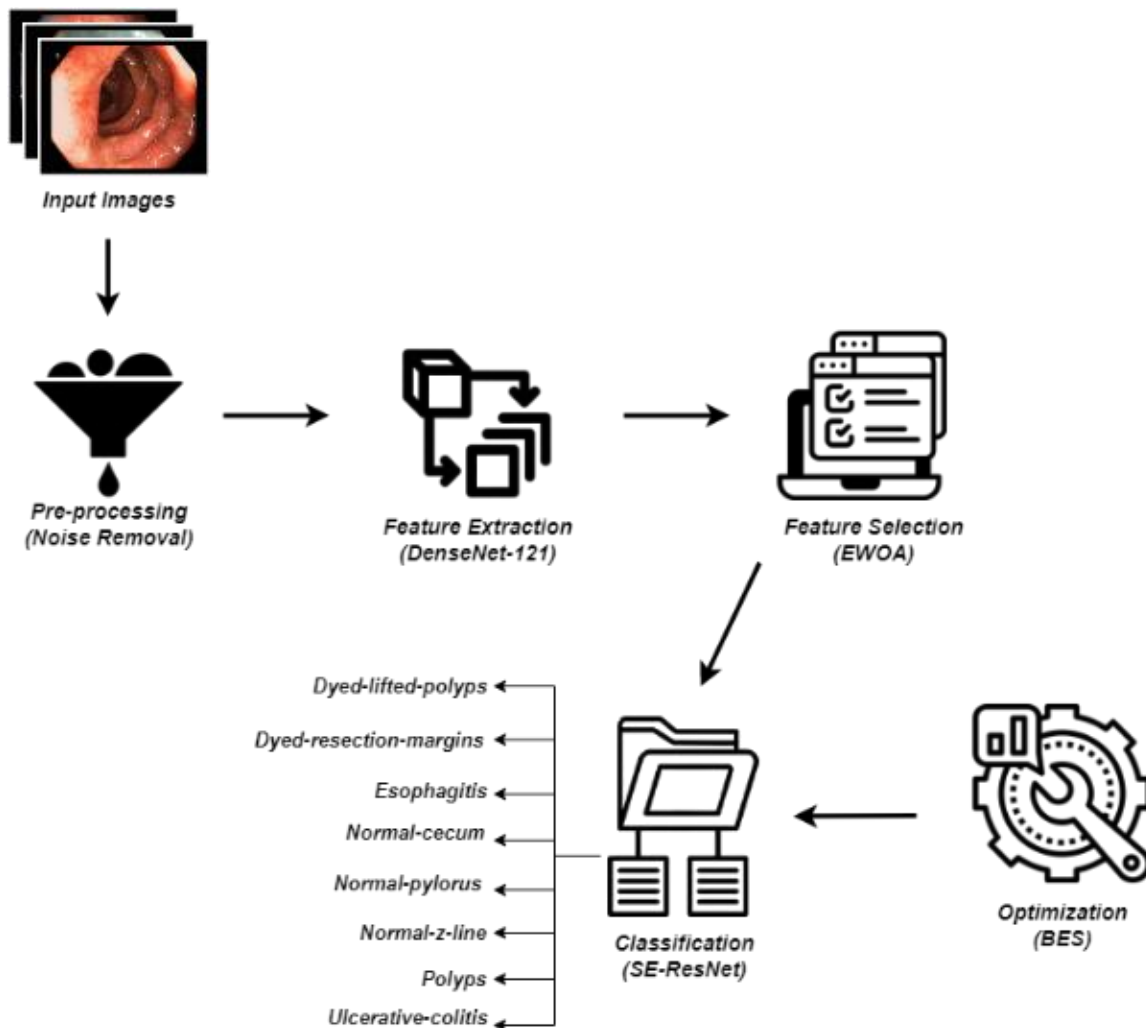


Figure 1. The proposed methodology architecture

Finally, to classify the eight classes of gastrointestinal diseases, utilize a deep learning technique of SE-ResNet with Bald Eagle Search optimization technique to get better accuracy of results. Figure 1 illustrated the architecture of the proposed technique.

4.1 Pre-processing

During the pre-processing stage, reduce the noise from the pictures. We remove noise from the input pictures to enhance the diagnosis of GI disease. Here, mean filtering is used to initially reduce the noise in the input pictures. Using numerous picture smoothing templates for picture convolution processing is a standard method for picture de-noising in the spatial domain to reduce or reduce noise. The fundamental idea behind mean filtering is to replace a pixel's single grey value with the total of the grey values of all of the pixels around it. The picture is considered as $g(x, y)$ after smoothing and mean filtering, and it is determined for a pixel point (x, y) in a given picture with $f(x, y)$, where its neighborhood S contains M pixels, by the following Eq. (1):

$$G(x, y) = \frac{1}{M} \sum_{(i,j) \in S} f(x, y)(x, y) \notin S \quad (1)$$

4.2 Feature extraction

After noise reduction is completed, we use the DenseNet-121 method to extract the features. This research proposes to utilize the DenseNet-121 approach to extract features and categorize GI diseases. Convolution and pooling layers from the DenseNet-121 method are used to extract features like shape and position from input wireless capsule endoscopy images.

Each layer is linked to every layer below it by a CNN named DenseNet-121. A Dense Block in CNNs is a module that connects all layers directly. Use a dense connection strategy to achieve feature extraction as the input for the subsequent layer, where every layer is connected layer and to each layer before it on the channel dimension [22]. A dense convolution neural network of the type DenseNet-121 is the one being used in this situation. The DenseNet design objective is to make DL networks significantly more profound while simultaneously making them simpler to develop by employing fewer interactions between the layers. Every layer in the DenseNet CNN is connected to any non-subsequent levels that are deeper in the hierarchy. This is carried out to facilitate the greatest data stream between organizational levels. To maintain the feedforward nature of the system, each layer compiles inputs from all preceding layers and offers its component direction to all layers that will follow it.

As a result, all of the initial convolutional blocks in the " i^{th} " layer has highlight guides and "I" information sources. The element maps for all succeeding "I-I" levels are provided. This presents an alternative to the simple "I" links found in typical profound learning models relationships within the company and it represent in Eq. (2).

$$(I * (I + 1))/2 \quad (2)$$

Since it is unnecessary to learn trivial element maps, it employs fewer boundaries than traditional convolutional neural networks.

DenseNet features two vital elements in addition to the essential pooling and convolutional layers. The levels are called Transition Blocks and Dense Blocks. When the amount of filters varies between Dense Blocks, the elements of the channel are enlarged. The development rate (k) helps to generalize the l^{th} layer. Each layer's data addition is controlled by this. In a nutshell, it is the amount of channels that a dense block outputs. This suggests that a dense layer's (l 's) share of features with the dense layer (l 's) preceding it is $k[l]$. This is referred to as the growth rate because, after every layer, $k[l]$ channel characteristics are synthesized and supplied as input to the subsequent layer. The two convolutional operations that make up a dense layer are 3×3 CONV and 1×1 CONV. The dense block of the DenseNet-121 is made up of six of these thick layers. The outcome of every dense block has a depth equal to its rate of expansion.

DenseNet is built on a core convolution and pooling layer. Following a dense block layer, there is an average pooling layer, followed by a classification layer, a change layer, a progress layer, another dense block layer, and lastly, a progress layer. Each dense block contains two convolutions, each with estimated sections of 1×1 and 3×3 . The repetition of this occurs frequently in dense block 1, multiple times in dense block 2, 24 times in dense block 3, and lastly 16 times in dense block 4. A standard convolution (1×1) is used to extract the features from each thick layer, and a 3×3 convolution is used to reduce the feature depth/channel count. After every dense block, the total amount of feature maps equals the number of input features.

4.3 Feature selection

After feature extraction, we utilize Enhanced Whale Optimization Algorithm (EWOA) to select the features. The humpback whales' bubble-net feeding technique, served as the model for the whale optimization algorithm. Three strategies were used by the whale optimization algorithm to mathematically represent this behavior: spiral bubble-net assaulting, encircling prey, and looking for prey. Consider a population of whales called $X^t = (X_1^t, X_2^t, \dots, X_N^t)$ where each vector $X_i^t = (X_{i,1}^t, X_{i,2}^t, \dots, X_{i,D}^t)$ specifies the location of the i th whale in iteration t . A matrix X^1 is constantly initialized in the issue space in the first iteration ($t = 1$) and for the subsequent iterations ($t > 1$), X^t is updated utilizing the three strategies of surrounding prey, searching for prey, and spiral bubble-net attacking. Additionally, the coefficient vector A_i^t is considered when a whale is deciding between two approaches to hunting and circling its target. Using these ideas X_i^{t+1} , Eq. (3) is used to determine the i^{th} humpback whale's position. The probability rate ρ_i^t is a random value between intervals (0, 1), the coefficient vector A_i^t is summed utilizing Eq. (4), and a_i^t is computed utilizing Eq. (5) to decrease linearly from 2 to 0 for iterations.

$$X_i^{t+1} = \begin{cases} \text{Encircling prey} & (\rho_i^t < 0.5) \text{ and } (|A_i^t| < 1) \\ \text{Search for prey} & (\rho_i^t < 0.5) \text{ and } (|A_i^t| \geq 1) \\ \text{Spiral bubble net attacking} & \rho_i^t \geq 0.5 \end{cases} \quad (3)$$

$$A_i^t = 2 \times a_i^t \times \text{rand} - a_i^t \quad (4)$$

$$a_i^t = 2 - t \times \left(\frac{2}{\text{MaxIt}} \right) \quad (5)$$

4.3.1 E-WOA (Enhanced whale optimization algorithm)

The E-WOA, which adds a pooling mechanism and 3 efficient search tactics called migrating, preferential selection, and enriched surrounding prey, enhances the evaluation of the conventional whale optimization technique.

Definition 1

(Pooling mechanism): Given a matrix $Pool = (P_1, P_2, \dots, P_k)$ of size with members $P_i = (P_{i,1}, P_{i,2}, \dots, P_{i,D})$ that are formed after each iteration using Eq. (6), X_{brnd}^t is computed utilizing Eq. (9) to construct a random position near the best humpback whale X_{best}^t , and X_{worst}^t is the worst finding found [23]. In Eq. (6), \bar{B}_i^t is its reverse vector and B_i^t is a binary random vector, where the equivalent values of non-zero components in B_i^t are equal to zero in \bar{B}_i^t and the corresponding values of zero components are equal to ones in \bar{B}_i^t . To improve diversity, the pooling mechanism uses a crossover operator to combine the unfavorable and promising solutions. A new solution is swapped out for an existing Pool member once the size of the Pool is finished.

$$P_i^t = B_i^t \times X_{brnd}^t + \bar{B}_i^t \times X_{worst}^t \quad (6)$$

Migrating search strategy: Using Eq. (7), this search approach randomly divides a section of the humpback whale to cover previously unexplored areas and improve exploration. Utilizing Eq. (8), where $rand$ is a shared random amount between 0 and 1, δ_{min} and δ_{max} are the upper and lower bounds of the issue, it is possible to calculate the X_{rnd}^t , a random point in the range of the search space. When utilizing Eq. (9) to determine the best humpback whale, X_{best}^t , the X_{brnd}^t is a random position in the vicinity of X_{best}^t , where δ_{best_min} and δ_{best_max} are the upper and lower bounds of X_{best}^t .

$$X_i^{t+1} = X_{rnd}^t - X_{brnd}^t \quad (7)$$

$$X_{rnd}^t = rand \times (\delta_{max} - \delta_{min}) + \delta_{min} \quad (8)$$

$$X_{brnd}^t = rand \times (\delta_{best_max} - \delta_{best_min}) + \delta_{best_min} \quad (9)$$

Preferential selecting search strategy: The canonical WOA's search for prey method has improved exploration potential thanks to the preference selection strategy. The formula for this method is Eq. (10), where X_i^t is the i^{th} whale's actual position, P_{rnd1}^t , and P_{rnd2}^t are two random numbers drawn from the Pool matrix in iteration t , C_i^t is specified, and A_i^t is sampled using the Cauchy distribution with the parameter settings. Since the preferential selection search approach is suggested to enhance the whale optimization algorithm exploration capability, sizable step size is required to distribute the whales across many search space regions to find a wide variety of answers. Because of the greater likelihood of creating bigger values, this technique makes advantage of the heavy-tailed Cauchy distribution.

$$X_i^{t+1} = X_i^t + A_i^t \times (C_i^t \times P_{rnd1}^t - P_{rnd2}^t) \quad (10)$$

Enriched encircling prey search strategy: Eq. (11), where P_{rnd3}^t is chosen at random from the matrix Pool and D^{it} is derived using Eq. (12), is utilized to enhance the encircling prey technique used in the WOA.

$$X_i^{t+1} = X_{best}^t - A_i^t \times D^{it} \quad (11)$$

$$D^{it} = |C_i^t \times X_{best}^t - P_{rnd3}^t| \quad (12)$$

The selected features are utilized to classify GI diseases.

4.4 Classification

We propose the new deep-learning SE-ResNet approach for categorizing gastrointestinal disorders. CNN has proven to be the best at computer visual tasks. The SE technique automatically determines each channel's weight through learning and increases the important components. The SE method's architecture is depicted in Figure 2.

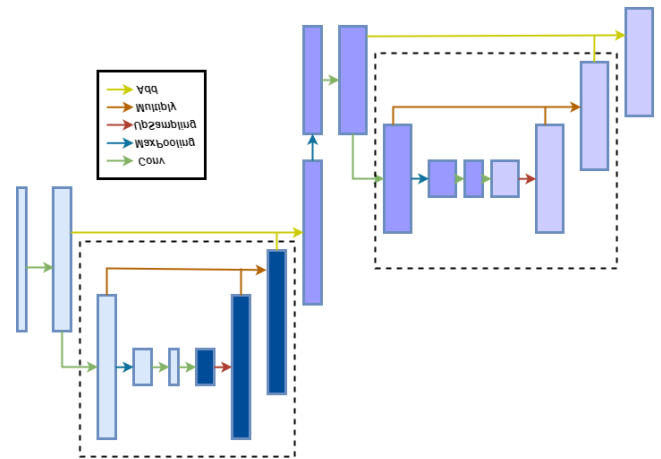


Figure 2. The structure of two operators, excitation and squeeze, make up one block and two SE blocks. Various operations are denoted by variously colored arrows

While CNN-based networks are popular because they outperform alternative architectures in terms of performance, they have drawbacks in that they are difficult to converge and require more memory to train. As a result, ResNets have been proposed as a training method for very deep CNNs. ResNet is a stacked block-wise architecture of the same shape, with each block including direct connections between the result of a lower layer and the inputs of a higher layer. SE-ResNet introduces additional layers called "Squeeze-and-Excitation" blocks within each residual block. These blocks focus on channel-wise feature recalibration, allowing the network to adaptively emphasize important features and suppress less relevant ones. The main components of the Squeeze-and-Excitation block are as follows, the first step is a global average pooling layer, which aggregates information globally from each channel in the feature map. It reduces the spatial dimensions of the feature map to a single value per channel. The pooled features are then passed through two fully connected (dense) layers. These layers learn to model the interdependencies between channels, enabling the network to capture the relative importance of each channel. The output of the excitation step is a set of channel-wise scaling factors that represent the importance of each channel. These scaling factors are used to reweight the original feature map, emphasizing more critical channels and suppressing less important ones.

By using global average pooling, the squeeze operator compresses the given input data to produce a statistical channel. Around the nonlinearity, the excitation processing

has a ReLU and 2 completely connected layers. To create a weighted channel, the excitation operator weights the input information [24]. The squeeze-and-excite operators keep the feature channel's size constant. It has frequently been used as the primary classification framework in the task of picture recognition and classification because the residual block of ResNet can effectively leverage shallow categories to acquire more key values. As a result, ResNet was the primary structure utilized in this paper's classification section. Combining the ResNet model and the SE block creates the SE-ResNet module, and Figure 3 depicts the computation method for this module.

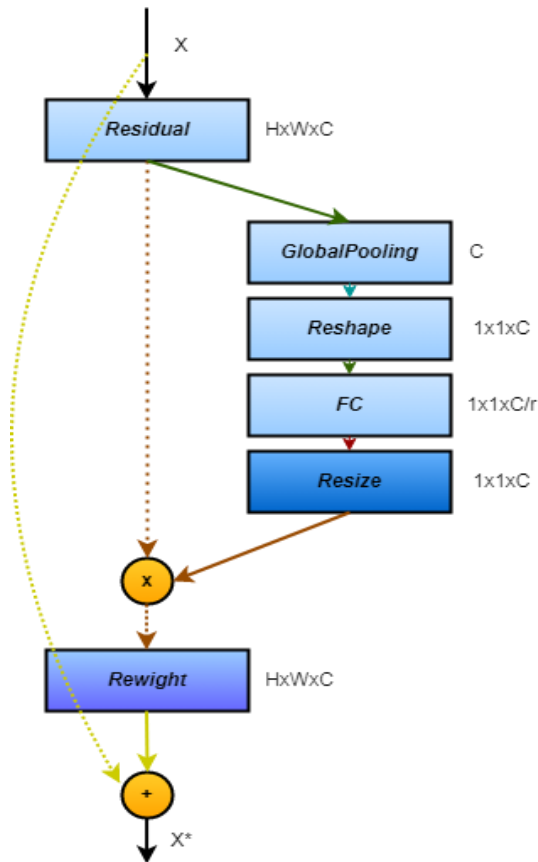


Figure 3. The SE-ResNet module architecture

Excitation and squeeze occur before the summing operation, as illustrated in Figure 3. The SE framework, which can both fully utilize shallow features and additionally reweight every channel of absorb features to improve the classification, is integrated into the residual block of ResNet. The SE-ResNet result is obtained by us as Eq. (13).

$$y = F(f_{se}(x), (\omega_i)) + x \quad (13)$$

where, x and y are the SE-input ResNet's outcome, $f_{se}(\cdot)$ is the method of the SE block, and ω_i is the network's weight for the i^{th} input. However, we must specify the scale of the feature picture during the squeeze operation because this has a significant impact on the reweighting value. This research indicates a changeable scale based on the size of the feature channel because the size of every input class picture is not the same. We define the j^{th} SE-ResNet block's output as in Eq. (14).

$$y_j = F(f_{se}(x_j), (\omega_{ij})) + x_j \quad (14)$$

y_j is the result of the j^{th} SE-ResNet architecture, where. We use an optimization approach from the Bald Eagle Search Optimization technique to achieve greater classification accuracy.

4.5 Optimization

For optimization, the bald eagle search optimization algorithm is used. A new meta-heuristic optimization technique called the bald eagle search (BES) was presented in 2020. Bald eagles are at the top of the food chain due to their size. They are sporadic hunters. They can eat any simple, readily accessible protein-rich diet. They prefer to eat fish, particularly salmon, either dead or alive. Because bald eagles have an outstanding vision and the ability to gaze in two directions simultaneously, they can locate fish from a great distance. Their clever social behavior in their hunting method served as the primary source of inspiration for BES. Bald eagles divide their hunting strategy into 3 levels. Choosing space, looking in space, and swooping are these phases. The eagle chooses the area with the most prey throughout the process of choosing the place. The eagle begins looking for prey inside the chosen space during the searching-in-the-space phase [25]. The eagle begins swinging from its ideal position in the swooping phase, which comes last. The best hunting location is then identified. All of the eagle's movements from this point forward are focused on it.

The following is a definition of the bald eagle's hunting system in terms of mathematics:

- **Selecting-Space Phase.** During this stage, the bald eagle chooses the best location based on the availability of food. Mathematically, this activity is denoted in Eq. (15):

$$X_{new} = X_{best} + \alpha \times r(X_{mean} - X_i) \quad (15)$$

where, X_{best} denotes the search is chosen based on the best eagle's location, X_{mean} denotes the average distance between all bald eagle locations (the population mean), X_i denotes the presence of an eagle location, r denotes a random parameter created in the range [0-1], and α denotes a constant parameter.

- **Searching-in-Space Phase.** In this phase, the bald eagle moves in various directions within the predetermined spiral zone from the previous section in search of prey. Additionally, the ideal location for swooping and prey hunting is chosen. In this stage, the eagle position is updated based on Eq. (16):

$$X_{new} = X_i + z(i) \times (X_i - X_{i+1}) + p(i) \times (X_i - X_{mean}) \quad (16a)$$

$$p(i) = \frac{pr(i)}{\max(|pr|)}, z(i) = \frac{zr(i)}{\max(|zr|)} \quad (16b)$$

$$pr(i) = r(i) \times \cos(\theta(i)), zr(i) = r(i) \times \sin(\theta(i)) \quad (16c)$$

$$\theta(i) = \alpha \times \pi \times r1 \quad (16d)$$

$$r(i) = \theta(i) + R \times r2 \quad (16e)$$

where, R is a constant parameter that accepts values between 0.5 and 2, α is a constant parameter that has a value between [0.5, 2], and $r1$ and $r2$ are two random parameters.

• **Swooping Phase.** All bald eagles begin swinging to their discovered prey in this step from the ideal position they achieved in the previous step. At this time, the eagles begin to swing from the best search posture towards their prey, as described in Eq. (17):

$$X_{new} = r3 \times X_{best} + pl(i) \times (X_i - t1 \times X_{mean}) + zl(i) \times (X_i - t2 \times X_{best}) \quad (17a)$$

$$pl(i) = \frac{pr(i)}{\max(|pr|)}, zl(i) = \frac{zr(i)}{\max(|zr|)} \quad (17b)$$

$$pr(i) = r(i) \times \sin h(\theta(i)), zr(i) = r(i) \times \cos h(\theta(i)) \quad (17c)$$

$$\theta(i) = \alpha \times \pi \times r3, r(i) = \theta(i) \quad (17d)$$

where the two constant variables are $t1$ and $t2$.

The earlier stages are thought to be essential for quickly arriving at a good solution.

5. RESULTS AND DISCUSSION

The first portion of this section categorizes GI disorders utilizing the dataset's analysis and method for retrieving disease features, comparing our approach to "state-of-the-art" methodologies.

5.1 Dataset description

The proposed framework's performance is assessed using the Kvasir v2 dataset. Anatomical landmarks, disease abnormalities, or GI endoscopic operations are all depicted in

the Kvasir v2 dataset. There are 8000 pictures in total and eight classes in this collection. Z-line, cecum, pylorus, ulcerative colitis, esophagitis, polyps, dyed resection margins, and dyed and lifted polyps are some of these classifications. There are 1000 pictures in each category. One image from every class in the Kvasir dataset can be found in Figure 4.

The Z-line, which lies between the stomach and the esophagus, provides important illness symptoms. Another class in this research, pylorus, enters the beginning of the small intestine. The colonoscopy starts at the cecum, which is the area closest to the large intestine. One of the three significant pathological signs, esophagitis, shows up on the Z-line and is connected to reflux. Intestinal lesions are polyps that can be recognized from healthy mucosa by their color and texture. Finally, by seeing how ulcerative colitis affects the large intestine, the condition can be identified.

5.2 Quantitative metrics

This research proposes a new technique to classify gastrointestinal diseases into eight categories. Here, first, we use the mean filter to reduce the noise from given input images to get feature extraction good. After the pre-processing step, we extract the features such as the shape and position of the disease in wireless capsule endoscopy images utilizing by DenseNet-121 method. The heatmap function is performed in the feature extraction step. Then using the EWOA to select the features. Finally, classify the diseases such as Normal-cecum, Normal-z-line, Dyed-lifted-polyps, Esophagitis, Dyed-resection-margins, Ulcerative-colitis, Normal-pylorus, and Polyps utilizing by SE-ResNet technique. To improve and get higher classification accuracy, we propose a Bald Eagle Search optimization method. Figure 5 demonstrated the performance of the proposed methodology.



Figure 4. Samples from the Kvasir dataset

Original Image	Pre-Processing (Noise Removed)	Feature Extraction (Heat Map)	Classification
			Dyed-lifted-polyps
			Dyed-resection-margins
			Esophagitis
			Normal-cecum
			Normal-pylorus
			Normal-z-line
			Polyps
			Ulcerative-colitis
(a)	(b)	(c)	(d)

Figure 5. The performance outcome is, (a) original images, (b) noise removed, (c) heatmap images, and (d) gastrointestinal tract diseases classification

5.3 Evaluation metrics

The accuracy and loss function of the proposed concatenated framework is assessed. Using a categorical cross-entropy loss function, which is frequently employed for multi-class issues, we calculated the difference between the outcome predictions and the desired outcome. The learning rate and other parameters were changed utilizing the bald eagle search optimization method to improve the model's accuracy and reduce loss. The starting learning rate was set to 0.001. According to the classification's performance criteria, the classed classes are contrasted. The performance parameters are calculated using the confusion matrix. The following is how every performance measure parameter is calculated:

● **Accuracy:** The accuracy is calculated as the number of accurate predictions divided by the whole sample size. The accuracy is represented in Eq. (18):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

● **Precision:** It is a metric that shows the percentage of people who have had a condition diagnosed. People who are projected to have a disease are TP and FP, while those who already have a condition are TP, as follows in Eq. (19):

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

● **Recall:** It is a metric that shows the percentage of patients who were identified by the model as having an illness and who have that disease. The people diagnosed by the method as having a condition are TP, and the real positives (those with a disease are TP and FN) are as follows in Eq. (20):

$$Recall = \frac{TP}{TP + FN} \quad (20)$$

● **F1-score:** It evaluates the similarity rate between ground-truth regions and predicted and reflects a harmonic average of precision and sensitivity. It is determined with Eq. (21):

$$F1 - score = \frac{2 * precision * recall}{precision + recall} \quad (21)$$

5.4 Performance evaluation

The proposed methodology provides the highest categorization accuracy when compared to other existing methods in experimental performance. Table 2 displays the accuracy, recall, precision, and f1-score results for LSTM [16], KNN [17], EWT + CNN [18], EfficientNet B0 [19], SVM [20], and the proposed SE-ResNet on the wireless capsule endoscopy pictures from the Kvasir v2 dataset. It is consistent with experimental outcomes for the eight classes. The proposed methodology has greater classification accuracy values than other existing approaches, according to the results. So, we showed the comparative outcomes as a graph, both with and without the use of an optimization algorithm. In comparison to other methods already being used, Figure 6 showed the accuracy analysis of the proposed method without optimization. Additionally, Figure 7 compares the classification outcomes of the proposed approach to those of other previous techniques without the use of optimization

algorithm graphs. With less computation time, the classification accuracy was increased by our proposed models.

Table 2. Calculate Precision, F1-score, Accuracy, and Recall (%) using the proposed and existing methods without optimization

Approaches	Accuracy (%)	Recall (%)	F1-score (%)	Precision (%)
LSTM [16]	98.05	97.84	98.05	98.02
NN [17]	99.24	99.10	99.18	99.21
EWT + CNN [18]	94.25	75.92	80.90	93.53
EfficientNet B0 [19]	97.99	98.03	97.94	97.89
SVM [20]	98.06	98.02	97.87	97.92
Proposed (SE-ResNet)	99.31	99.24	99.28	99.32

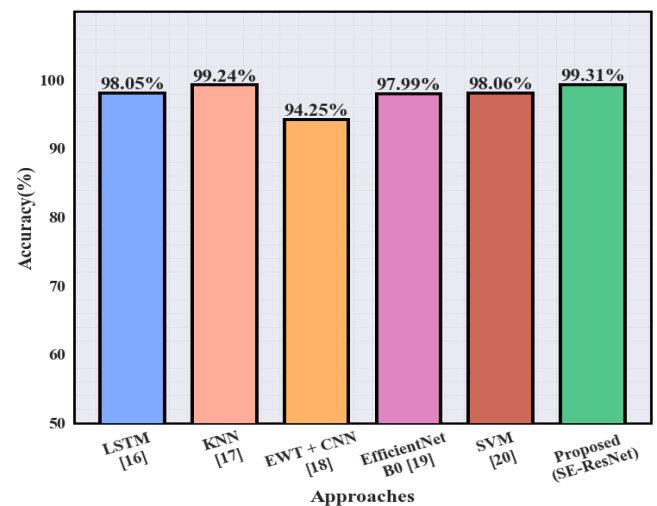


Figure 6. Analysis of accuracy using various methods

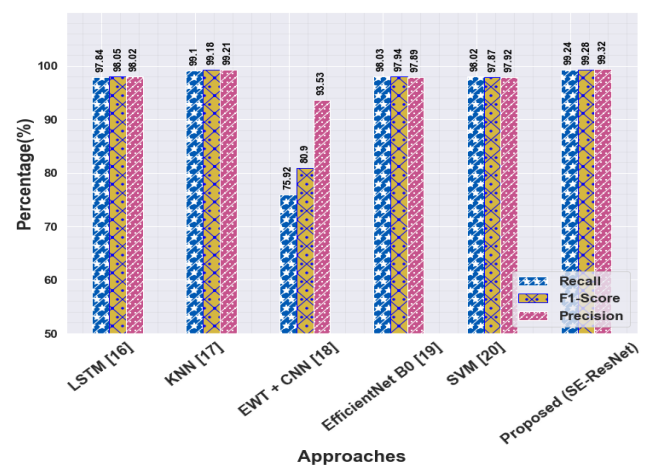


Figure 7. Comparison of the categorization results of the proposed method with those of previous methods without optimization process

Table 3 displays our results using an optimization technique based on precision, f1-score, accuracy, and recall. The accuracy analysis of the proposed approach with optimization accuracy in comparison to other methods is shown in Figure 8. Additionally, Figure 9 compares the classification outcomes of the proposed method with those of other previous

techniques using graphs of the optimization algorithm. In terms of f1-score, accuracy, recall, and precision, Table 3 compares the performance of LSTM [16], KNN [17], EWT + CNN [18], EfficientNet B0 [19], SVM [20], and the proposed SE-ResNet with BES method. The proposed methodology has greater classification accuracy values than other deep learning methods already being used, according to the results.

Table 3. Calculate Precision, F1-score, Accuracy, and Recall (%) using the proposed and existing methods with optimization

Approaches	Accuracy (%)	Recall (%)	F1-Score (%)	Precision (%)
LSTM [16]	98.05	97.84	98.05	98.02
KNN [17]	99.24	99.10	99.18	99.21
EWT + CNN [18]	94.25	75.92	80.90	93.53
EfficientNet B0 [19]	97.99	98.03	97.94	97.89
SVM [20]	98.06	98.02	97.87	97.92
Proposed (SE-ResNet)	99.66	99.57	99.61	99.52

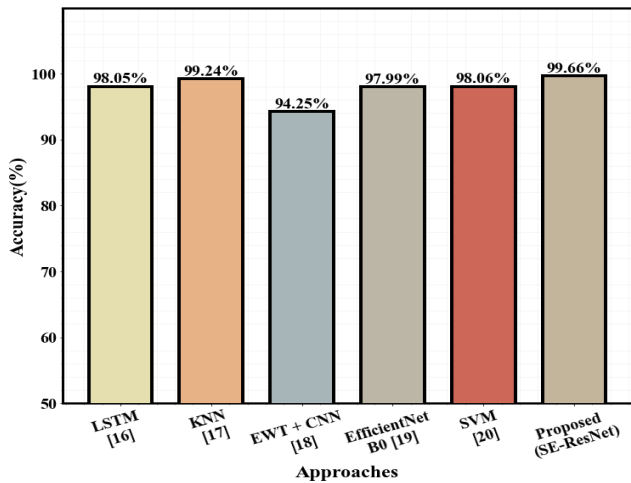


Figure 8. Analysis of accuracy using various methods

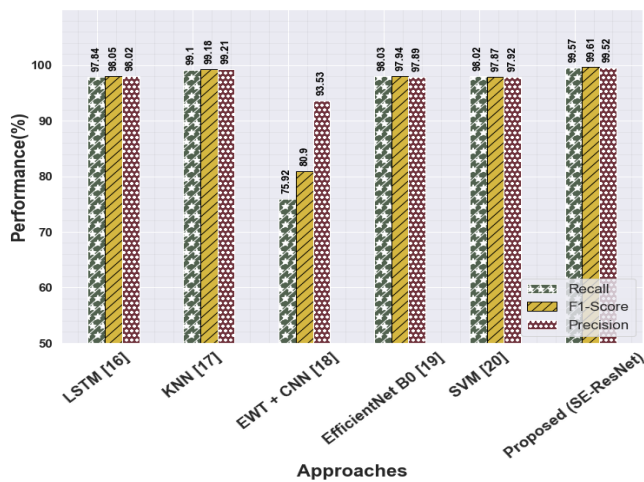


Figure 9. Comparison of the categorization results of the proposed method with those of previous methods with the optimization process

The achieved higher recall for the proposed approach is 99.57%, compared to 97.84% for LSTM [16], 99.10% for KNN [17], 75.92% for EWT + CNN [18], 98.03% for EfficientNet B0 [19], and 98.02% for SVM [20]. The proposed approach's specificity is also better than other existing approaches. The accuracy rate for EfficientNet B0 [19] is 97.99 percent, which is the lowest. Figure 9 shows the comparison of recall, precision, and f1-score outcomes for the proposed technique classification with those from other methods using optimization algorithm graphs. With the optimization approach, the proposed SE-ResNet technology achieves higher accuracy when compared to other existing methods. According to the analysis of the experiment, implementing the proposed framework increased classification performance and reduced computing time for training the pictures. The proposed method achieved higher classification accuracy compared to other existing methods. Also, I have done the evaluation performance with less computation time.

5.5 Computation time

Another concept that is discussed is the concept of computation time. Deep learning techniques aim to make computations simpler. Table 4 contrasts our proposed SE-ResNet technique's computation times with those of other previously employed methods. With little computing work, categorization accuracy is improved. Figure 10 displays the computation time utilizing the proposed model and the most recent approaches.

Table 4. Utilizing the proposed and existing methods

Methods	Computation Time
LSTM [16]	0.23
KNN [17]	0.19
EWT + CNN [18]	0.20
EfficientNet B0 [19]	0.26
SVM [20]	0.17
Proposed (SE-ResNet)	0.14

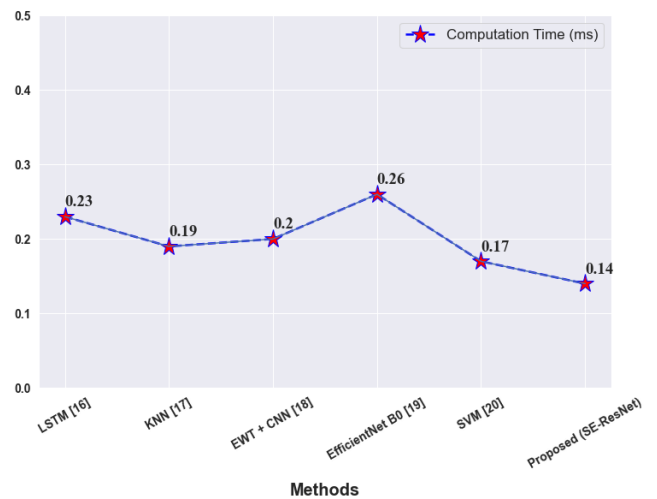


Figure 10. The computation time for the proposed methodology and existing methodologies

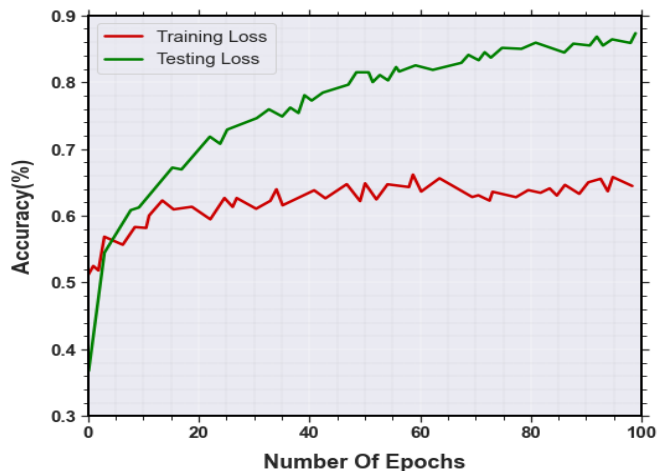


Figure 11. Training vs validation accuracy

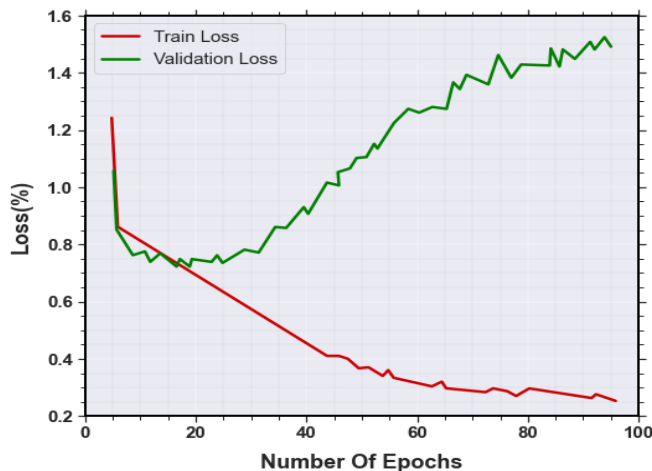


Figure 12. Training vs validation loss

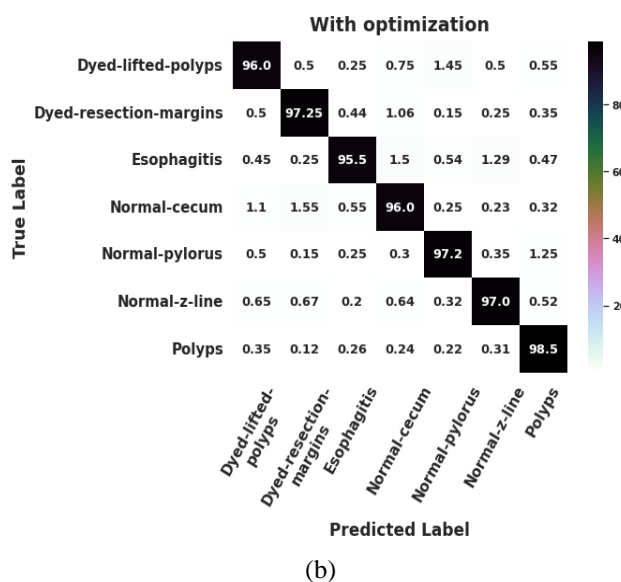
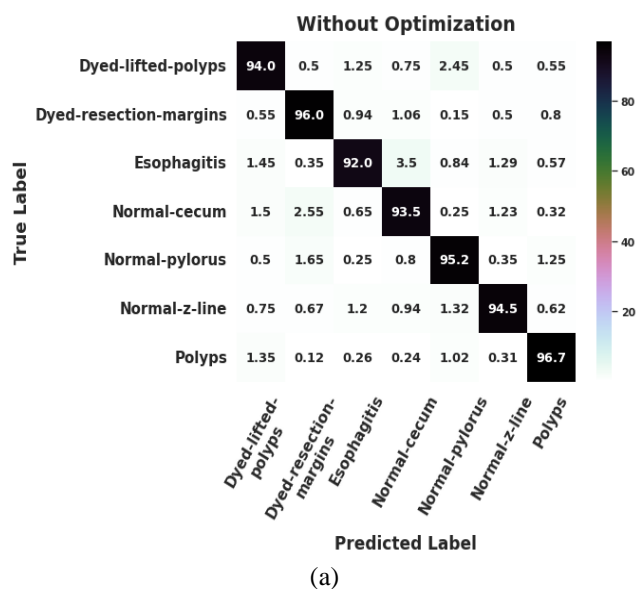


Figure 13. Confusion matrix of face objects (a) without optimization, and (b) with optimization

5.6 Evaluation of training results

Our accuracy after 100 epochs was 99.66%, which is notable provided that accuracy curves often converge. The accuracy of validation and training is displayed in Figure 11. The validation loss curves rapidly rise and fall. It implies that more test findings may be beneficial. The validation and training loss is displayed in Figure 12.

Figures 11 and 12 show the loss and accuracy of the training. SE-ResNet provides higher accuracy and loss forecasts. In both the training and validation stages of the categorization process for GI illnesses, our system beats earlier approaches.

The confusion matrix is the method that is most frequently used to evaluate classification errors. The confusion matrix for the SE-ResNet proposed model was developed based on the explanations for the confusion matrix that were provided. Figure 13 displays the classification cross-validation test's obtained confusion matrix.

6. MANAGERIAL IMPLICATIONS

Managerial implications for gastrointestinal (GI) diseases are essential for healthcare organizations, medical

professionals, and policymakers to address the challenges associated with the prevention, diagnosis, treatment, and management of these conditions. Adequate healthcare infrastructure is crucial for efficiently managing GI diseases. Managers should ensure that hospitals, clinics, and healthcare facilities have the necessary equipment, specialized units, and qualified staff to handle GI-related cases effectively. Managerial efforts should be directed towards raising public awareness about GI diseases, their risk factors, and early symptoms. Educational campaigns can help promote healthy lifestyle choices and encourage individuals to seek timely medical attention. Developing screening programs for high-risk populations can aid in the early detection of GI diseases, such as colorectal cancer. Timely detection can improve treatment outcomes and reduce healthcare costs. Managers can support research and development initiatives to advance the understanding and treatment of GI diseases. Encouraging clinical trials and collaboration with academic institutions can lead to innovative therapies and improved patient outcomes.

7. CONCLUSION AND FUTURE WORKS

Bacteria, viruses, and parasites can all cause gastrointestinal

infections (GIs). Most instances resolve in a few days, but severe symptoms like high fever, vomiting, or blood in the stool call for medical intervention. But the diagnosis of GI diseases, there are challenges to detect. To overcome the issues, propose a novel deep learning technique to classify GI diseases. First, collect the WCE images from the dataset, using the mean filter to remove the noises in input images. Then, extract the features using the DenseNet-121 technique. After feature extraction, select the essential features using the Enhanced Whale Optimization Algorithm. Finally, we propose a SE-ResNet technique to categorize the eight classes of gastrointestinal diseases with the Bald Eagle Search optimization method to get better accuracy of classification results. Our experiments, which utilized the Kvasir v2 dataset, were successful in terms of recall, precision, accuracy, and f1-score. The classification technique performs with 99.66% accuracy with less computation time. But this proposed method also has some limitations. The availability of high-quality datasets for training classification models can be limited. Variability in data collection, inconsistent reporting, and limited sample sizes can hinder the development of robust and generalizable classification models. In the future, we need to evaluate the performance on a few more datasets to accurate the classification process. And increase the number of WCE images for training purposes.

REFERENCES

[1] Hmoud Al-Adhaileh, M., Mohammed Senan, E., Alsaade, W., Aldhyani, T.H.H., Alsharif, N., Alqarni, A., Uddin, M.I., Alzahrani, M.Y., Alzain, E.D., Jadhav, M.E. (2021). Deep learning algorithms for detection and classification of gastrointestinal diseases. *Complexity*, 2021: 1-12.

[2] Mohapatra, S., Nayak, J., Mishra, M., Pati, G.K., Naik, B., Swarnkar, T. (2021). Wavelet transform and deep convolutional neural network-based smart healthcare system for gastrointestinal disease detection. *Interdisciplinary Sciences: Computational Life Sciences*, 13: 212-228. <https://doi.org/10.1007/s12539-021-00417-8>

[3] Holland, A.M., Bon-Frauches, A.C., Keszthelyi, D., Melotte, V., Boesmans, W. (2021). The enteric nervous system in gastrointestinal disease etiology. *Cellular and Molecular Life Sciences*, 78(10): 4713-4733. <https://doi.org/10.1007/s00018-021-03812-y>

[4] Khan, M.A., Khan, M.A., Ahmed, F., Mittal, M., Goyal, L.M., Hemanth, D.J., Satapathy, S.C. (2020). Gastrointestinal diseases segmentation and classification based on duo-deep architectures. *Pattern Recognition Letters*, 131: 193-204. <https://doi.org/10.1016/j.patrec.2019.12.024>

[5] Öztürk, Ş., Özkaya, U. (2021). Residual LSTM layered CNN for classification of gastrointestinal tract diseases. *Journal of Biomedical Informatics*, 113: 103638. <https://doi.org/10.1016/j.jbi.2020.103638>

[6] Aziz, I., Simrén, M. (2021). The overlap between irritable bowel syndrome and organic gastrointestinal diseases. *The Lancet Gastroenterology & Hepatology*, 6(2): 139-148. [https://doi.org/10.1016/s2468-1253\(20\)30212-0](https://doi.org/10.1016/s2468-1253(20)30212-0)

[7] Matisz, C.E., Gruber, A.J. (2022). Neuroinflammatory remodeling of the anterior cingulate cortex as a key driver of mood disorders in gastrointestinal disease and

disorders. *Neuroscience & Biobehavioral Reviews*, 133: 104497. <https://doi.org/10.1016/j.neubiorev.2021.12.020>

[8] Mineshige, T., Inoue, T., Yasuda, M., Yurimoto, T., Kawai, K., Sasaki, E. (2020). Novel gastrointestinal disease in common marmosets characterised by duodenal dilation: a clinical and pathological study. *Scientific Reports*, 10(1): 1-10. <https://doi.org/10.1038/s41598-020-60398-4>

[9] Vuik, F.E., Nieuwenburg, S.A., Moen, S., Schreuders, E.H., Pool, M.D.O., Peterse, E.F., Spada, C., Epstein, O., Fernández-Urién, I., Hofman, A., Kuipers, E.J., Spaander, M.C. (2022). Population-based prevalence of gastrointestinal abnormalities at colon capsule endoscopy. *Clinical Gastroenterology and Hepatology*, 20(3): 692-700. <https://doi.org/10.1016/j.cgh.2020.10.048>

[10] Soffer, S., Klang, E., Shimon, O., Nachmias, N., Eliakim, R., Ben-Horin, S., Kopylov, U., Barash, Y. (2020). Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. *Gastrointestinal Endoscopy*, 92(4): 831-839. <https://doi.org/10.1016/j.gie.2020.04.039>

[11] Muhammad, K., Khan, S., Kumar, N., Del Ser, J., Mirjalili, S. (2020). Vision-based personalized wireless capsule endoscopy for smart healthcare: taxonomy, literature review, opportunities and challenges. *Future Generation Computer Systems*, 113: 266-280. <https://doi.org/10.1016/j.future.2020.06.048>

[12] Saito, H., Aoki, T., Aoyama, K., et al. (2020). Automatic detection and classification of protruding lesions in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointestinal Endoscopy*, 92(1): 144-151. <https://doi.org/10.1016/j.gie.2020.01.054>

[13] Jain, S., Seal, A., Ojha, A., Yazidi, A., Bures, J., Tacheci, I., Krejcar, O. (2021). A deep CNN model for anomaly detection and localization in wireless capsule endoscopy images. *Computers in Biology and Medicine*, 137: 104789. <https://doi.org/10.1016/j.combiomed.2021.104789>

[14] Wang, G.B., Xuan, X.W., Jiang, D.L., Li, K., Wang, W. (2022). A miniaturized implantable antenna sensor for wireless capsule endoscopy system. *AEU-International Journal of Electronics and Communications*, 143: 154022. <https://doi.org/10.1016/j.aeue.2021.154022>

[15] Muruganatham, P., Balakrishnan, S.M. (2022). Attention aware deep learning model for wireless capsule endoscopy lesion classification and localization. *Journal of Medical and Biological Engineering*, 42(2): 157-168. <http://dx.doi.org/10.1007/s40846-022-00686-8>

[16] Öztürk, Ş., Özkaya, U. (2021). Residual LSTM layered CNN for classification of gastrointestinal tract diseases. *Journal of Biomedical Informatics*, 113: 103638. <https://doi.org/10.1016/j.jbi.2020.103638>

[17] Sharif, M., Attique Khan, M., Rashid, M., Yasmin, M., Afza, F., Tanik, U.J. (2021). Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. *Journal of Experimental & Theoretical Artificial Intelligence*, 33(4): 577-599. <https://doi.org/10.1080/0952813X.2019.1572657>

[18] Mohapatra, S., Pati, G.K., Mishra, M., Swarnkar, T. (2023). Gastrointestinal abnormality detection and

- classification using empirical wavelet transform and deep convolutional neural network from endoscopic images. *Ain Shams Engineering Journal*, 14(4): 101942. <https://doi.org/10.1016/j.asej.2022.101942>
- [19] Ramamurthy, K., George, T.T., Shah, Y., Sasidhar, P. (2022). A novel multi-feature fusion method for classification of gastrointestinal diseases using endoscopy images. *Diagnostics*, 12(10): 2316. <https://doi.org/10.3390/diagnostics12102316>
- [20] Haile, M.B., Salau, A.O., Enyew, B., Belay, A.J. (2022). Detection and classification of gastrointestinal disease using convolutional neural network and SVM. *Cogent Engineering*, 9(1): 2084878. <https://doi.org/10.1080/23311916.2022.2084878>
- [21] Mohapatra, S., Nayak, J., Mishra, M., Pati, G.K., Naik, B., Swarnkar, T. (2021). Wavelet transform and deep convolutional neural network-based smart healthcare system for gastrointestinal disease detection. *Interdisciplinary Sciences: Computational Life Sciences*, 13: 212-228. <https://doi.org/10.1007/s12539-021-00417-8>
- [22] Solano-Rojas, B., Villalón-Fonseca, R., Marín-Raventós, G. (2020). Alzheimer's disease early detection using a low cost three-dimensional densenet-121 architecture. In *The Impact of Digital Technologies on Public Health in Developed and Developing Countries: 18th International Conference, ICOST 2020, Hammamet, Tunisia, June 24–26, 2020, Proceedings 18* (pp. 3-15). Springer International Publishing. https://doi.org/10.1007/978-3-030-51517-1_1
- [23] Nadimi-Shahraki, M.H., Zamani, H., Mirjalili, S. (2022). Enhanced whale optimization algorithm for medical feature selection: A COVID-19 case study. *Computers in Biology and Medicine*, 148: 105858. <https://doi.org/10.1016/j.combiomed.2022.105858>
- [24] Jiang, Y., Chen, L., Zhang, H., Xiao, X. (2019). Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PloS One*, 14(3): e0214587. <https://doi.org/10.1371/journal.pone.0214587>
- [25] Sayed, G.I., Soliman, M.M., Hassanien, A.E. (2021). A novel melanoma prediction model for imbalanced data using optimized SqueezeNet by bald eagle search optimization. *Computers in Biology and Medicine*, 136: 104712. <https://doi.org/10.1016/j.combiomed.2021.104712>