IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 12 February 2023, accepted 1 March 2023, date of publication 10 March 2023, date of current version 17 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3256084

RESEARCH ARTICLE

Modified Salp Swarm Algorithm With Deep Learning Based Gastrointestinal Tract Disease Classification on Endoscopic Images

MARWA OBAYYA[1], FAHD N. AL-WESABI[2], MASHAEL MAASHI[3], ABDULLAH MOHAMED[4], MANAR AHMED HAMZA[5], SUHANDA DRAR[5], ISHFAQ YASEEN[5], AND MOHAMED IBRAHIM ALSAID[5]

[1]Department of Biomedical Engineering, College of Engineering, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[2]Department of Computer Science, College of Science and Art at Mahayil, King Khalid University, Abha 62217, Saudi Arabia
[3]Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[4]Research Centre, Future University in Egypt, New Cairo 11845, Egypt
[5]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam Bin Abdulaziz University, Al-Kharj 16242, Saudi Arabia

Corresponding author: Fahd N. Al-Wesabi (falwesabi@kku.edu.sa)

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Large Groups Project under grant number (10/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R203), Princess Nourah bint Abdulrahman University, Riyadh, Research Supporting Project number(RSP2023R787), King Saud University, Riyadh, Saudi Arabia. Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444).

ABSTRACT Nowadays, the analysis of gastrointestinal (GI) tract disease utilzing endoscopic image classification becomes an active research activity from the biomedical sector. The latest technology in medical imaging is Wireless Capsule Endoscopy (WCE) for diagnosing gastrointestinal diseases namely bleeding, ulcer, polyp, and so on. Manual diagnoses will be time taking and tough for the medical practitioner; thus, the authors have designed computerized approaches for classifying and detecting such diseases. Many research groups presented various machine learning (ML) and image processing methods for classifying GI tract diseases in recent times. Conventional data augmentation and image processing methods are integrated with adjusted pre-trained deep convolutional neural networks (CNNs) for classifying diseases in the GI tract from WCI images. This study presents a Modified Salp Swarm Algorithm with Deep Learning based Gastrointestinal Tract Disease Classification (MSSADL-GITDC) on Endoscopic Images. The presented MSSADL-GITDC technique mainly focuses on the examination of WCE images for GIT classification. To accomplish this, the presented MSSADL-GITDC technique applies median filtering (MF) technique for image smoothening. The presented MSSADL-GITDC technique designs improved capsule network (CapsNet) model for feature extraction where the CapsNet model is modified by the class attention layer (CAL). Moreover, MSSA based hyperparameter tuning process is performed to improve the efficiency of the improved CapsNet model. For GIT classification, deep belief network with extreme learning machine (DBN-ELM) was used. Finally, backpropagation is applied for supervised fine tuning of the DBN-ELM model. The experimental validation of the MSSADL-GITDC technique takes place on Kvasir-V2 database reported the betterment of the MSSADL-GITDC technique on GIT classification with maximum accuracy of 98.03%.

INDEX TERMS Medical imaging, gastrointestinal tract diseases, deep learning, metaheursitics, fine-tuning, salp swarm algorithm.

I. INTRODUCTION

Gastrointestinal (GI) diseases are increasingly common in the human digestive system. Some of the common factors of mortality are colorectal cancer, Stomach cancer, and esophageal cancer [1], [2], [3]. Generally, Endoscopy is necessary to diagnose diseases and it is the initial step in identifying GI tract diseases [4]. These endoscopic examinations even improve the analysis of the clinical characteristics of lesions for determining their type and severity and making proper

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Zuo.

diagnoses [5]. Differences in the knowledge of medical practitioner cause errors in certain cases, particularly with regard to problematic aspects of videos and images from endoscopy [6], [7]. These inconsistencies may result in negative impact and misdiagnoses on patient care. Automated disease classification effectively solves this issue by offering doctors reliable and objective identification of various GI endoscopic images, hence improving prognosis, economizing valuable time of clinicians, and reducing the misdiagnosis rate [8]. The researchers are still working on automated GI disease classification for achieving better classification accuracy and lesion detection.

Recently, several computerized methods are provided for medical disease classification and detection [9], [10]. They concentrated on renowned medical imaging modalities like Electroencephalogram Signals, mammography for breast cancer, carcinoma, and pathology like deep learning (DL) exhibits great enhancement in medical image processing. DL can be referred to as a robust machine learning (ML) algorithm for the classification of automatic medical infections into respective categories [11]. Convolutional Neural Networks (CNNs) is a main cast of DL to extract deep high-level features. CNN processed input data into different forms such as multi-dimensional, images, videos, and signals. A simple CNN method has numerous layers like fully connected (FC), convolutional, classification, and pooling [12]. The commonly used pre-trained CNN techniques are ResNet, AlexNet, GoogleNet, and VGG.

GI-tract disease classification of artifacts, and multi-class diseases, from GI endoscopic imageries utilizing attention-guided CNN, was not implemented before [13], [14]. Artefact classification and multi-class disease generalization were necessary not only for diagnosis and even for avoiding training biases. Optimal usage of the DL method for the classification of automated GI disease is limited by a lack of data. Dissimilar to traditional ML-related techniques like support vector machine (SVM) are leveraged for extracting features, CNN has revealed better efficiency in extracting features [15]. The effective utility of CNN has enhanced classification and image recognition-based tasks.

This study presents a Modified Salp Swarm Algorithm with Deep Learning based Gastrointestinal Tract Disease Classification (MSSADL-GITDC) on Endoscopic Images. The presented MSSADL-GITDC technique employs improved capsule network (CapsNet) model with class attention layer (CAL). To improve the performance of the improved CapsNet model, MSSA based hyperparameter tuning process is performed. For GIT classification, deep belief network-extreme learning machine (DBN-ELM) appraoch with backpropagation is used. The experimental validation of the MSSADL-GITDC technique takes place on Kvasir-V2 dataset. In short, the key contributions of the study are given below.

- An automated MSSADL-GITDC technique encompasses image pre-processing, improved CapsNet feature extraction, MSSA hyperparameter tuning, and DBN-ELM-BP classification is presented for GTI diagnosis. To the best of our knowledge, the MSSADL-GITDC system never existed in the literature.
- Design an improved CapsNet model for feature extraction where the CAL is used for capturing the discriminative class-specific features for handling the class dependency.
- Hyperparameter tuning using MSSA helps to enhance the performance of the CapsNet model, which is designed by the incorporation of oppositional based learning (OBL) concept with traditional SSA.
- Validate the performance of the MSSADL-GITDC technique on Kvasir-V2 dataset and the outcomes are examined under different measures.

II. RELATED WORKS

Su et al. [16] introduced a new and practical approach for identifying gastrointestinal (GI) disease in wireless capsule endoscopy (WCE) image with CNN. The projected technique employs 3 backbone networks improved and optimally tuned by transfer learning (TL) as extracting features, and combined classification utilizing ensemble learning was trained for detecting GI diseases. Khan et al. [17] examined an automatic structure for GIT disease segmentation and classifier dependent upon deep feature map and Bayesian optimum DL selective feature. The presented structure was composed of some key steps, pre-processed to classifier. During the subsequent stage, the authors projected a deep saliency map to segment infected areas. The segmentation area is then utilized for training a pre-training fine-tuned system termed MobileNetV2 utilizing TL. The fine-tuning approach hyperparameter was initialization employing Bayesian optimization (BO).

In [18], an effectual classifier approach was projected to a GIT classifier task which comprises a smaller count of labeled datasets and takes an instance count of imbalance betwixt classes. Based on this method, utilizing an effectual technique at last the CNN infrastructure generates the selected efficiency while the CNN infrastructure could not strongly be trained. Therefore, an extremely effectual long short term memory (LSTM) infrastructure was planned and additional to the resultant of CNN. Haile et al. [19] introduce a concatenated neural network (NN) approach by concatenating the extracting features of InceptionNet and VGGNet for developing a GI disease analysis method. These extracting features are then concatenated and categorized utilizing ML classifier approaches.

In [20], a novel approach was executed dependent upon the fusion of geometric and CNN features. Primarily, disease areas can be extracted in provided WCE images utilizing a novel system termed as contrast-enhanced color feature. The geometric feature can be extracted in segmentation disease part. So, unique VGG-16 and VGG-19 deep CNN features fusion can be executed dependent upon Euclidean Fisher Vector. The feature selection can be finally classified by k-nearest

neighbor (KNN). Ramamurthy et al. [21] presented a new approach to classifier of endoscopy images by concentrating on feature mining with CNN. This projected method was created by integrating a recent structure (viz., EfficientNetB0) with custom-built CNN infrastructure termed Effimix.

Biradher and Aparna [22] offered a novel approach to classifier betwixt bleeding and nonbleeding classes of WCE images. The presented system makes utilizes an easy Deep CNN that comprises 6 convolution layers interchanged with max-pooling layers and this approach was related to present ones with respect to distinct performance metrics. Cao et al. [23] examined an effectual approach for classifying distinct lesion images developed by WCE. Primarily, the authors attain feature map of similar resolution by executing a max-pooling function on distinct convolution layers, afterward calculate the pooled feature map with trainable weighted parameters, and lastly, one-by-one convolutional kernels can be utilized for merging the integrated quantized feature maps. The authors improve the efficiency of extracting features by integrating multi-level convolution features, comprising either low- or high-level features.

Wang et al. [24] examine a 2-phase endoscopic image classifier approach that is effectual integrate complementary benefits of mid-level CNN feature and capsule network (CapsNet). In detail, the main problem our technique is a lesion-aware CNN extraction feature element that is encoding suitably detailed data of lesions in mid-level CNN feature and in-turn allow the subsequent capsule classifier network for efficiently learning deformation-invariant connections amongst image entities. In [25], a CapsNet improved with radon transform for extraction feature was presented for improving the possibility of colorectal cancer detection. The impact of this work lies in the integration of radon transforms from the presented method for improving the recognition of polyps by executing effectual extraction of tomographic features.

Khan et al. [26] introduce a DL and Moth-Crow optimizer-based approach for GI disease classifier. Afterward, utilizing transfer learning (TL), 2 pre-training SL techniques can be fine-tuning and training on GI disease image. Features are extraction in the middle layer utilizing both fine-tuned DL approaches (average pooling). On both extraction deep feature vectors, a hybrid Crow-Moth optimizer technique was presented and implemented. Afriyie et al. [14] examines a less sophisticated yet effectual pre-processed system to identify endoscopic images termed as denoising CapsNets (Dn-CapsNets). Furthermore, the authors created activation maps (AM) utilizing the feature representation for visualizing the outcomes.

Though several ML and DL models for for GIT classification are available in the literature, it is still needed to enhance the classification performance. Owing to continual deepening of the model, the number of parameters of DL models also increases quickly which results in model overfitting. At the same time, different hyperparameters have a significant impact on the efficiency of the CNN model. Particularly, the hyperparameters such as epoch count, batch size, and learning rate selection are essential to attain effectual outcome. Since the trial and error method for hyperparameter tuning is a tedious and erroneous process, metaheuristic algorithms can be applied. Therefore, in this work, we employ MSSA for the parameter selection of the CapsNet model.

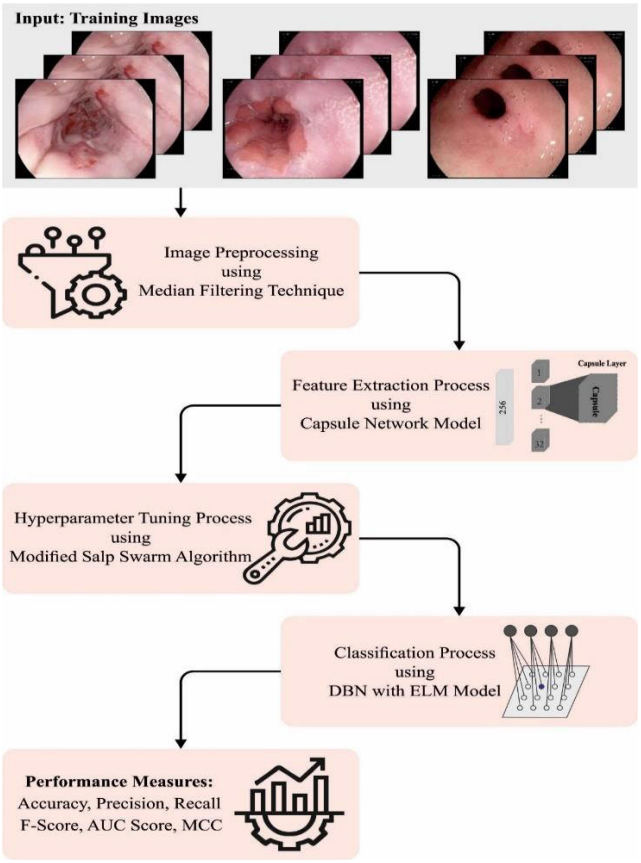

FIGURE 1. Overall process of MSSADL-GITDC system.

III. THE PROPOSED MODEL

This study has developed a new MSSADL-GITDC approach for the examination of WCE images for GIT classification. The presented MSSADL-GITDC approach encompasses several stages of operations namely median filter (MF) based pre-processing, improved CapsNet feature extraction, MSSA hyperparameter tuning, and DBN-ELM-BP classification. Fig. 1 depicts the workflow of MSSADL-GITDC approach.

A. IMAGE PRE-PROCESSING

The renowned order-statistics filter is the MF, which replaced the values of pixel with median of the gray levels in neighborhood of that pixel:

$$\hat{f}(x, y) = \{g(s, t)\} \tag{1}$$

The actual valued of the pixel can be added to the median computation. MF is common for some types of random noise they grant exceptional noise reduction abilities, with less blurring compared to linear smoothing filters of the same size.

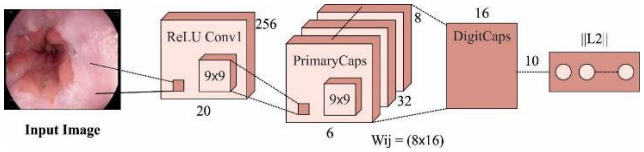

FIGURE 2. Structure of CapsNet.

B. IMPROVED CapsNet BASED FEATURE EXTRACTION

In the presented model, the improved CapsNet model is used for feature extraction. CapsNet architecture is established to maintain the location of an object and its features in image and model hierarchical relationship [27]. In the CNN model, valuable information arrives before the pooling layer. Furthermore, the CNN technique produces a scalar value in neural output. The CapsNet produces vector output of same size but with different routings since the capsule comprises numerous neurons. The CapsNets utilize a vector activation function termed squashing. In contrast, CNN exploits scalar input activation functions as given in the following;

$$v_j = \frac{||S_j||^2}{1 + ||S_j||^2} \frac{S_j}{||S_j||} \tag{2}$$

In Eq. (2), v_j is capsule output j, and s_j is total capsule input.

Excepting the final layer of CapsNets, the total input values of the s_j capsules can be established by weighted sum of the predictive vector ($U_{j|i}$) in capsule as located in the lower layer and ($U_{j|i}$) can be evaluated by multiplying the capsule in lower layer with weight matrix (W_{ij}) and output (O_i).

$$S_j = \sum\nolimits_i b_{ij} u_{j|i} \tag{3}$$

$$u_{j|i} = W_{ij} O_i \tag{4}$$

where, b_{ij} signifies co-efficient defined by dynamic routing method and calculated by the following expression;

$$b_{ij} = \frac{exp(a_{ij})}{\sum_k exp(a_{ik})} \tag{5}$$

In Eq. (5), a_{ij} shows the *log*-likelihood. The amount of relation coefficients amongst capsules i, and capsule in the topmost layer is 1 and the preceding *log* probability can be defined using Softmax. In CapsNet, a margin loss was introduced to define whether the object of specific class is existing and evaluated by Eq. (6);

$$L_k = T_k \max(O, m^+ - \|v_k\|)^2 + \lambda(1 - T_k max(O, \|v_k\| - \mathrm{m}))^2 \tag{6}$$

The T_k value is 1 when class k is presented. In addition, $m^+ = 0.9$ and $m^- = 0.1$ signifies hypervariable and the down weight of the loss. The vector length can be evaluated in the CapsNet and represents the probability, when vector direction encompasses the variable datasets such as texture, size, color, position, etc. Fig. 2 represents the architecture of CapsNet.

By using the extracting feature in pre-trained, CapsNet is at a high level and directly fed to FC layer for generating multi-label predictions, it can be difficult for learning the high-order probabilistic dependence via recurrent feeding with related features. Thus, a CAL is exploited to explore features regarding each class. With size $W \times W \times K$, and $1v_l$ represents the l-th convolution filter in CAL layer and it can be accomplished as follows:

$$M_l = X * w_l, \tag{7}$$

In Eq. (7), l extent from 1 to numerous classes and $*$ signifies convolution function. Then, a CAL layer, in which the filter count is equal to class, is added to produce class specific feature depictions regarding entire classification. With enough training, they can learn class-wise attention mapping. It can be monitored that class attention mapping highlights distinctive regions for different classifications and determines nearly no activation regarding missing classes. Subsequently, class attention mapping M_l are transformed to class wise feature vector v_l of W^2 dimension through vectorized.

C. HYPERPARAMETER TUNING USING MSSA

To optimize the performance of improved CapsNet model, the MSSA hyperparameter tuning technique is used. The SSA goes to the family of Salpidae which has tubular and transparent body structure [28]. The mathematical modeling of a Salp chain can be divided into follower and leader groups. The leader leads and directs the group, and the follower follows one after the other (and the leader indirectly or directly). Like other approached, the position of Salps can be described in n-dimension searching space, whereas n refers to the count of parameters. A $2D$ matrix named x stores the position of each Salp. Also, a food source named F is considered a target of collection in the search space. In order to update the location of leader, the subsequent equation has been used:

$$x_i^1 = F_i + c_1 \left((ub_i - lb_i) c_2 + lb_j \right) c_3 \geq 0 F_i - c_1 \left((ub_i - lb_i) c_2 + lb_i \right), c_3 < 0 \tag{8}$$

Let, x_i^1 be the initial Salp location (leader) at i^{th} parameter, F_i refers to the food source position at i^{th} dimensional, ub_i and lb_i signify the upper as well as lower limits of i^{th} parameter, c_1, c_2, and c_3 represent the random number. A leader upgrades the location with respect to the food source, as demonstrated by Eq. (8). In this work, co-efficient c_1 is crucial since it gives a tradeoff amongst search and exploitation:

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \tag{9}$$

In Eq. (9), l indicates the existing iteration and L shows an overall amount of iterations. c_2 and c_3 are uniformly generated random numbers in [1, 0]. The direction of j^{th} variable following location movement to negative or positive infinity and step size are defined as c_2 and c_3. Newton's law of motion can be utilized for updating the follower's location as follows.

$$X_j^i = \frac{1}{2} a t^2 + v_0 t \tag{10}$$

If $i \geq 2, x_j^i$ shows the i^{th} follower location at j^{th} variable, t indicates time, v_0 represent the initial velocity and $a = \frac{v_{final}}{v_0}$

is determined where we have $v = \frac{x-x0}{t}$. Meanwhile, time is constant, and the difference between them is 1, such that Eq. (7) can be formulated in the following, consider $v_0 = 0$.

$$X_j^i = \frac{1}{2}at^2 + v_0 t \tag{11}$$

If $i \geq 2$, x_j^i is the Salps position of i^{th} follower in j^{th} dimension. Then, simulate Salp chain using Eqs. (8) and (11). The transient response of processes can be represented by 2 significant factors: the closeness of output to the input and the speed of response. The error signal can be formulated as follows:

$$e(t) = u(t) - y(t) \tag{12}$$

By using OBL concept, the MSSA can be derived. It is an optimization algorithm that is used for enhancing the quality of initial population solution by differentiating the solution. It involves opposite and original solutions [29]. Eventually, the OBL approach gets the optimum solution from each solution.

- Opposite number: x represent a real number over interval $x \in [lb, ub]$. The opposite number of x can be indicated as $\tilde{x}$ and it is represented as follows:

$$\tilde{x} = lb + ub - x \tag{13}$$

Eq. (13) is generalized for employing it in a search space with multi-dimensional. Consequently, each searching-agent location and thier opposite location would be characterized by the subsequent Eqs. (14) & (15):

$$x = [x_1, x_2, x_3, \ldots, x_D] \tag{14}$$

$$\tilde{x} = \left[\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots, \tilde{x}_D\right] \tag{15}$$

The values of each element in $\tilde{x}$ is expressed by:

$$\tilde{\chi} = lb_j + ub_j - \chi_j \; where \; j = 1, 2, 3, \ldots, D \tag{16}$$

If $f(\tilde{x})$ fitness value of the opposite solution is better than $f(x)$ of novel solution x, then $x = \tilde{x}$; otherwise ; $x = x$.

Algorithm 1 Pseudocode of SSA

```
Initialize Salps population with respect to lb, ub        while (stop-
ping condition is met)
    Compute the objective function for all agents
    Choose the optimal search factor
    Upgrade c1 by (9)
        for each salp
            if (i == 1)
                Upgrade the leader Salps's location by (10)
            else
                Upgrade follower location by (11)
            end
        end
    Check the Salps position is in the range of lb, ub        End
Return optimal salps
```

- The process for incorporating OBL with SSA is discussed below:

a) Initialize the salp position X as x_i whereas ($i = l, 2, \ldots n$).
b) Define the opposite position of salp population OX as $\tilde{x}$ whereas ($i = l, 2, \ldots n$).
c) Selecting the n best salps in $X \cup OX$ and signifies the early population of *SSA*.

The MSSA approach derived a fitness function (FF) to have a better classifier outcome. It determined positive values for signifying superior outcomes of the candidate solutions. In this article, the reduction of classifier error rate will be treated as the FF, as follows.

$$\begin{aligned} fitness(x_i) &= ClassifierErrorRate(x_i) \\ &= \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \end{aligned} \tag{17}$$

D. IMAGE CLASSIFICATION USING DBN-ELM MODEL

In this study, the DBN-ELM model is employed for GIT classification. Since Hinton et al. developed the DBN that is effectively used in dimensionality reduction, classification, regression, etc [30]. The conception of unsupervised pretraining was introduced for DBN which aims at the problem of falling into local minima and slow convergence caused by random initialization of parameters, and simultaneously the problems of setting labeled training samples are resolved. Once unsupervised pretraining was utilized and there exist training instances, generalization and training errors are considerably decreased. In the presented method, BP and pretraining initial parameters to fine-tune the network for increasing the efficiency and accuracy of classification. Generally speaking, DBN-ELM comprises DBN for ELM and feature extraction as classifier. Assume DBN involves $n - th$ hidden layer, n-l layer is initialized by greedy traind, offset and weight from $n - l$ to $n - th$ hidden layers and from n to following layers.

$$\sum_{i=1}^{m} \beta_i g(W_i \cdot O_{n-1} + b_i) = y_j, j = 1, \ldots, l \tag{18}$$

The better prediction result is attained by minimal output error:

$$\sum_{i=1}^{m} \|y_j - t_j\| = 0 \tag{19}$$

Furthermore, a β_i is attained, making the subsequent formula true, as:

$$\sum_{i=1}^{m} \beta_i O_{n,j} = t_j, j = 1, \ldots, l \tag{20}$$

The abovementioned equation is rehabilitated to:

$$O_n \beta = H \tag{21}$$

where, O_n signifies the output from n-lth to $n - th$ layers that is formulated as follows:

$$O_n(W_{1,\ldots,W_m}, b_1, \ldots, b_m, O_{n-1,1}, \ldots, O_{n-1,i})$$

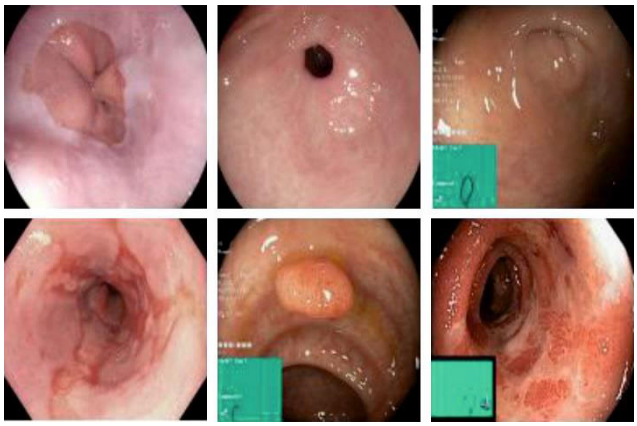

FIGURE 3. Sample images.

$$= \left[g(W_1 \cdot O_{n-1,l} + b_1) \ldots g(W_m \cdot O_{n-1,l} + b_m) \vdots \right.$$
$$\vdots$$
$$\left. \vdots\, g(W_1 \cdot O_{n-1,l} + b_i) \ldots g(W_m \cdot O_{n-1,l} + b_m) \right] \quad (22)$$

$$\beta = \left[\beta_1^T \vdots H_m^T \right], H = \left[H_1^T \vdots H_m^T \right] \quad (23)$$

Currently, the network is trained to achieve a $\hat{W}$, b_i, $\hat{\beta}$, such that:

$$\|o_m(\hat{W}_i, \hat{b}_i)\hat{\beta} - H\| = \|o_n(\hat{W}_i, \hat{b}_i)\hat{\beta} - H\| \quad (24)$$

W_i and β_i parameters are selected randomly in the last layer. Once the two variables are defined, the output H can be exclusively attained by:

$$\hat{\beta} = O_m^T H \quad (25)$$

BP model divide learning method into two phases: Initially, signals are broadcasted forward, and the input was processed by hidden, input and output layers; then, error BP: when there is an error betwixt the actual and expected output, then the error is divided to unit of every layer for correcting the neuron weight of all, and the layers error signal will go back together previous route. Such processes are reiterated until the network error satisfies the requirement.

IV. PERFORMANCE VALIDATION

In this section, the experimental validation of the MSSADL-GITDC approach is tested using the Kvasir-V2 dataset. Pathological observations, polyp removal, and anatomical landmarks are among the 8 groups that make up the database with 1000 images as given in Table 1. The images in the dataset range in resolution in 720 × 576 to 1920 × 1072 pixels. The dataset holds 8000 images. Few sample images are shown in Fig. 3.

The proposed model is simulated using Python 3.6.5 tool on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

TABLE 1. Details of dataset.

Labels	Class	No. of Instances
C-1	Dyed lifted polyps	1000
C-2	Dyed resection margins	1000
C-3	Esophagitis	1000
C-4	Normal-cecum	1000
C-5	Normal-pylorus	1000
C-6	Normal-Z-line	1000
C-7	Polyps	1000
C-8	Ulcerative colitis	1000
Total Number of Instances		**8000**

TABLE 2. GIT classification outcome of MSSADL-GITDC system on 80:20 of TR/TS databases.

Labels	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	AUC_{score}	MCC
Training Phase (80%)						
C-1	97.62	90.17	91.48	90.82	95.01	89.46
C-2	97.86	92.71	89.79	91.22	94.39	90.02
C-3	97.28	91.40	86.38	88.82	92.61	87.32
C-4	97.53	89.68	90.58	90.12	94.55	88.72
C-5	97.98	91.22	92.71	91.96	95.72	90.81
C-6	97.75	92.07	89.78	90.91	94.33	89.63
C-7	97.92	89.23	94.72	91.90	96.55	90.76
C-8	98.45	93.39	94.21	93.80	96.63	92.92
Avg.	**97.80**	**91.23**	**91.21**	**91.19**	**94.97**	**89.95**
Testing Phase (20%)						
C-1	98.31	91.26	93.82	92.52	96.35	91.58
C-2	98.12	94.47	90.82	92.61	95.02	91.56
C-3	97.75	93.16	88.50	90.77	93.79	89.53
C-4	97.19	89.55	88.24	88.89	93.37	87.28
C-5	98.31	91.94	95.10	93.49	96.94	92.54
C-6	98.00	94.15	89.39	91.71	94.30	90.61
C-7	98.00	89.81	95.10	92.38	96.76	91.28
C-8	98.56	92.92	96.10	94.48	97.51	93.68
Avg.	**98.03**	**92.16**	**92.13**	**92.11**	**95.50**	**91.01**

The GIT classification performance of the MSSADL-GITDC model is presented in the form of confusion matrix under training (TR) and testing (TS) databases in Fig. 3. The results indicated that the MSSADL-GITDC model has recognized WCE images into 8 class labels. For example, on 80% of TR database, the MSSADL-GITDC model has recognized 752 into C-1, 712 into C-2, 691 into C-3, 721 into C-4, 738 into C-5, 720 into C-6, 754 into C-7, and 749 into C-8. Eventually, on 20% of TS database, the MSSADL-GITDC approach has recognized 167 into C-1, 188 into C-2, 177 into C-3, 180 into C-4, 194 into C-5, 177 into C-6, 194 into C-7, 197 into C-8. Meanwhile, on 70% of TR database, the MSSADL-GITDC methodology has recognized 670 into C-1, 633 into C-2, 608 into C-3, 639 into C-4, 606 into C-5, 584 into C-6, 683 into C-7, 656 into C-8.

In Table 2 and Fig. 5, overall GIT classification outcomes of the MSSADL-GITDC model under 80% of TR and 20% of TS databases are given in terms of different measures such as accuracy ($accu_y$), precision ($prec_n$), recall ($reca_l$), F-score (F_{score}), AUC score (AUC_{score}), and Mathew Correlation Coefficient (MCC). The outcomes signified that the MSSADL-GITDC model has accurately recognized eight

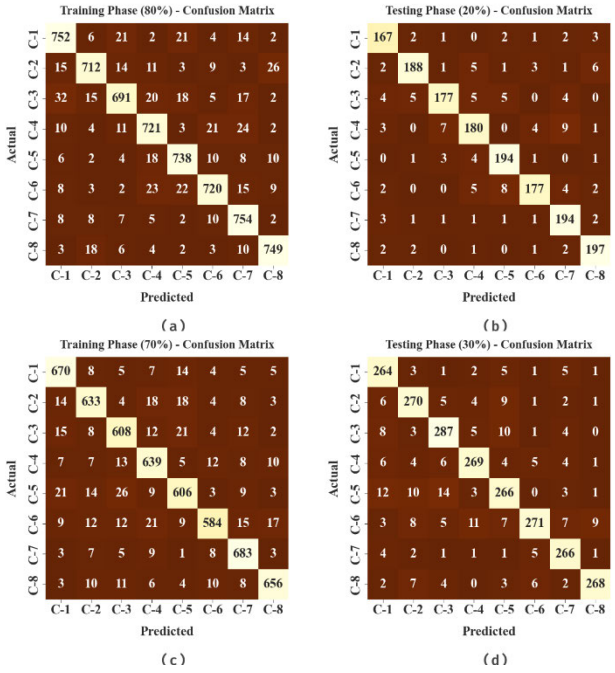

FIGURE 4. Confusion matrices of MSSADL-GITDC system (a-b) TR and TS databases of 80:20 and (c-d) TR and TS databases of 70:30.

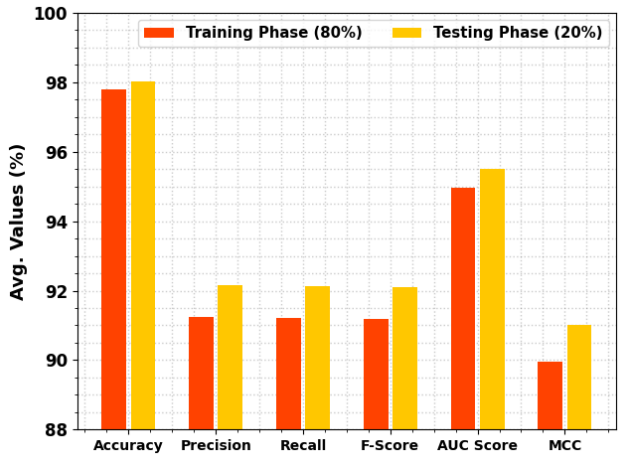

FIGURE 5. Average outcome of MSSADL-GITDC system on 80:20 of TR/TS databases.

classes under all aspects. For example, on 80% of TR database, the MSSADL-GITDC method has offered average $accu_y$ of 97.80%, $prec_n$ of 91.23%, $reca_l$ of 91.21%, F_{score} of 91.19%, AUC_{score} of 94.97%, and MCC of 89.95%. Moreover, on 20% of TS database, the MSSADL-GITDC approach has offered average $accu_y$ of 98.03%, $prec_n$ of 92.16%, $reca_l$ of 92.13%, F_{score} of 92.11%, AUC_{score} of 95.50%, and MCC of 91.01%.

In Table 3 and Fig. 6, an overall GIT classification outcome of the MSSADL-GITDC methodology under 70% of TR and 30% of TS databases is given. The experimental values implied that the MSSADL-GITDC approach has accurately recognized 8 classes under all aspects. For instance, on 70% of TR database, the MSSADL-GITDC approach has offered average $accu_y$ of 97.67%, $prec_n$ of 90.72%, $reca_l$ of

TABLE 3. GIT classification outcome of MSSADL-GITDC system on 70:30 of TR/TS databases.

Labels	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	AUC_{score}	MCC
Training Phase (70%)						
C-1	97.86	90.30	93.31	91.78	95.92	90.57
C-2	97.59	90.56	90.17	90.36	94.41	88.99
C-3	97.32	88.89	89.15	89.02	93.80	87.49
C-4	97.43	88.63	91.16	89.87	94.74	88.41
C-5	97.20	89.38	87.70	88.53	93.12	86.94
C-6	97.50	92.85	86.01	89.30	92.55	87.97
C-7	98.20	91.31	94.99	93.12	96.83	92.10
C-8	98.30	93.85	92.66	93.25	95.89	92.28
Average	**97.67**	**90.72**	**90.64**	**90.65**	**94.66**	**89.34**
Testing Phase (30%)						
C-1	97.54	86.56	93.62	89.95	95.84	88.64
C-2	97.29	87.95	90.60	89.26	94.42	87.72
C-3	97.21	88.85	90.25	89.55	94.26	87.94
C-4	97.67	91.19	89.97	90.57	94.36	89.24
C-5	96.58	87.21	86.08	86.64	92.11	84.69
C-6	97.12	93.45	84.42	88.71	91.75	87.21
C-7	98.25	90.78	94.66	92.68	96.69	91.72
C-8	98.42	95.04	91.78	93.38	95.56	92.50
Average	**97.51**	**90.13**	**90.17**	**90.09**	**94.38**	**88.71**

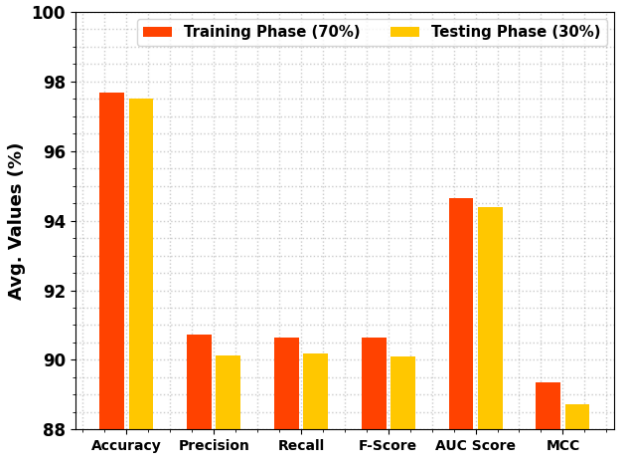

FIGURE 6. Average outcome of MSSADL-GITDC system on 70:30 of TR/TS databases.

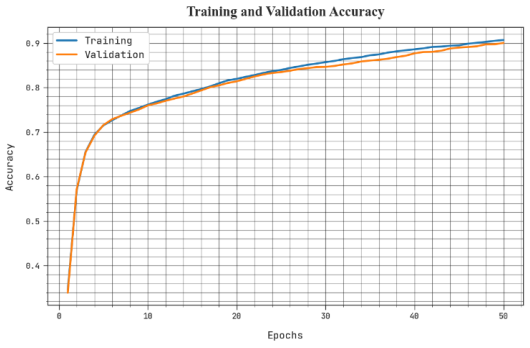

FIGURE 7. TACC and VACC analysis of MSSADL-GITDC system.

90.64%, F_{score} of 90.65%, AUC_{score} of 94.66%, and MCC of 89.34%. Furthermore, on 30% of TS database, the MSSADL-GITDC system has attained average $accu_y$ of 97.51%, $prec_n$ of 90.13%, $reca_l$ of 90.17%, F_{score} of 90.09%, AUC_{score} of 94.38%, and MCC of 88.71%.

FIGURE 8. TLS and VLS analysis of MSSADL-GITDC system.

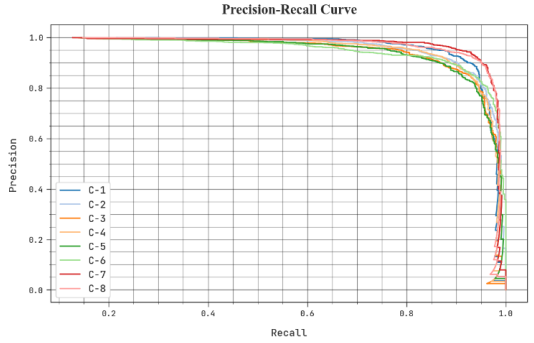

FIGURE 9. Precision-recall analysis of MSSADL-GITDC system.

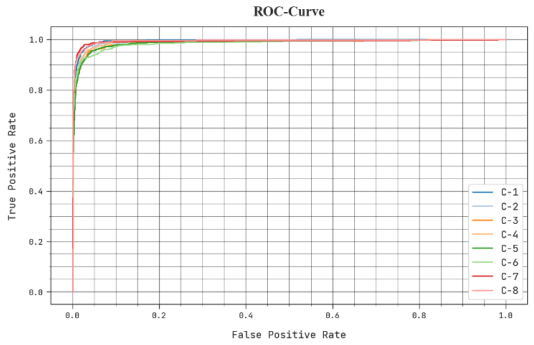

FIGURE 10. ROC curve analysis of MSSADL-GITDC system.

The TACC and VACC of the MSSADL-GITDC approach are investigated on GIT classification performance in Fig. 7. The figure demonstrated that the MSSADL-GITDC system has exhibited superior performance with maximal values of TACC and VACC. It can be observable that the MSSADL-GITDC algorithm has obtained higher TACC outcomes.

The TLS and VLS of the MSSADL-GITDC approach are tested on GIT classification performance in Fig. 8. The figure pointed out that the MSSADL-GITDC system has revealed better performance with least values of TLS and VLS. The MSSADL-GITDC system has resulted in lesser VLS outcomes.

A noticeable precision-recall study of the MSSADL-GITDC system in the test database is demonstrated in Fig. 9. The figure revealed that the MSSADL-GITDC approach has led to maximal values of precision-recall values.

A ROC study of the MSSADL-GITDC algorithm in the test database was portrayed in Fig. 10. The outcome implied that the MSSADL-GITDC system has outperformed their capability in classifying in several class labels.

TABLE 4. Comparative analysis of MSSADL-GITDC methodology with other algorithms [8], [31].

Methods	$Prec_n$	$Reca_l$	$Accu_y$	F_{score}
MSSADL-GITDC	92.16	92.13	98.03	92.11
ECA–Net	89.00	91.24	92.68	88.42
DL–OCT	90.09	90.29	90.20	89.00
LSMT–CNN	90.14	91.75	92.82	91.98
Attention-Guided CNN	91.44	90.32	93.25	91.05
VGG16	92.01	90.13	96.02	91.24
ResNet-18	78.55	79.22	78.86	78.80
LR Tree	87.04	89.17	94.13	91.53

To assure the better performance of the MSSADL-GITDC model, a wide range of comparison study is made in Table 4. The experimental values indicated that the ResNet-18 model has shown least performance while the ECA-Net and LR Tree models have accomplished closer performance. Although the other DL models have exhibited reasonable performance, the MSSADL-GITDC model has shown its superior performance with $prec_n$ of 92.16%, $reca_l$ of 92.13%, $accu_y$ of 98.03%, and F_{score} of 92.11%. These results reported that the MSSADL-GITDC model has accomplished maximum performance on GIT classification.

V. CONCLUSION

This study has developed a new MSSADL-GITDC system for the examination of WCE images for GIT classification. The presented MSSADL-GITDC technique comprises several stages of operations such as MF based pre-processing, improved CapsNet feature extraction, MSSA hyperparameter tuning, and DBN-ELM-BP classification. In the presented model, the improved CapsNet model is used for feature extraction where the CAL is used for capturing the discriminative class-specific features for handling the class dependency. At the same time, hyperparameter tuning using MSSA helps to optimize the performance of the CapsNet model, which is designed by the incorporation of the OBL concept with traditional SSA. The experimental validation of the MSSADL-GITDC approach occurs on Kvasir-V2 dataset. The simulation results reported the betterment of the MSSADL-GITDC technique on GIT classification. Thus, the MSSADL-GITDC technique can be employed for accurate classification of GIT on WCE images. In the future, the performance of the MSSADL-GITDC technique was enhanced by deep ensemble fusion models.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated during the current study.

ETHICS APPROVAL

This article does not contain any studies with human participants performed by any of the authors.

CONSENT TO PARTICIPATE

Not applicable.

FUNDING DETAILS

None.

INFORMED CONSENT

Not applicable.

REFERENCES

[1] V. Raut, R. Gunjan, V. V. Shete, and U. D. Eknath, "Gastrointestinal tract disease segmentation and classification in wireless capsule endoscopy using intelligent deep learning model," *Comput. Methods Biomech. Biomed. Eng., Imag. Visualizat.*, vol. 2021, pp. 1–17, Aug. 2022, doi: 10.1080/21681163.2022.2099298.

[2] S. Yang, C. Lemke, B. F. Cox, I. P. Newton, I. Näthke, and S. Cochran, "A learning-based microultrasound system for the detection of inflammation of the gastrointestinal tract," *IEEE Trans. Med. Imag.*, vol. 40, no. 1, pp. 38–47, Jan. 2021.

[3] R. T. Sutton, O. R. Zaiane, R. Goebel, and D. C. Baumgart, "Artificial intelligence enabled automated diagnosis and grading of ulcerative colitis endoscopy images," *Sci. Rep.*, vol. 12, no. 1, pp. 1–10, Feb. 2022.

[4] J. Yogapriya, V. Chandran, M. G. Sumithra, P. Anitha, P. Jenopaul, and C. S. G. Dhas, "Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–12, Sep. 2021, doi: 10.1155/2021/5940433.

[5] M. M. Auzine, P. Bissoonauth-Daiboo, M. H. M. Khan, S. Baichoo, X. Gao, and N. G. Sahib, "Classification of artefacts in endoscopic images using deep neural network," in *Proc. 3rd Int. Conf. Next Gener. Comput. Appl. (NextComp)*, Oct. 2022, pp. 1–5, doi: 10.1109/NextComp55567.2022.9932202.

[6] F. J. P. Montalbo, "Fusing compressed deep ConvNets with a self-normalizing residual block and alpha dropout for a cost-efficient classification and diagnosis of gastrointestinal tract diseases," *MethodsX*, vol. 9, 2022, Art. no. 101925, doi: 10.1016/j.mex.2022.101925.

[7] F. Mohammad and M. Al-Razgan, "Deep feature fusion and optimization-based approach for stomach disease classification," *Sensors*, vol. 22, no. 7, p. 2801, Apr. 2022.

[8] S. Parasa, M. Wallace, U. Bagci, M. Antonino, T. Berzin, M. Byrne, H. Celik, K. Farahani, M. Golding, S. Gross, and V. Jamali, "Proceedings from the first global artificial intelligence in gastroenterology and endoscopy summit," *Gastrointestinal Endoscopy*, vol. 92, no. 4, pp. 938–945, Oct. 2020.

[9] M. H. Al-Adhaileh, E. M. Senan, F. W. Alsaade, T. H. H. Aldhyani, N. Alsharif, A. Abdullah Alqarni, M. I. Uddin, M. Y. Alzahrani, E. D. Alzain, and M. E. Jadhav, "Deep learning algorithms for detection and classification of gastrointestinal diseases," *Complexity*, vol. 2021, pp. 1–12, Oct. 2021, doi: 10.1155/2021/6170416.

[10] Y. Djenouri, A. Belhadi, A. Yazidi, G. Srivastava, P. Chatterjee, and J. C. W. Lin, "An intelligent collaborative image-sensing system for disease detection," *IEEE Sensors J.*, vol. 23, no. 2, pp. 947–954, Jan. 2023, doi: 10.1109/JSEN.2022.3202437.

[11] I. Iqbal, K. Walayat, M. U. Kakar, and J. Ma, "Automated identification of human gastrointestinal tract abnormalities based on deep convolutional neural network with endoscopic images," *Intell. Syst. Appl.*, vol. 16, Nov. 2022, Art. no. 200149, doi: 10.1016/j.iswa.2022.200149.

[12] M. A. Khan, M. A. Khan, F. Ahmed, M. Mittal, L. M. Goyal, D. J. Hemanth, and S. C. Satapathy, "Gastrointestinal diseases segmentation and classification based on duo-deep architectures," *Pattern Recognit. Lett.*, vol. 131, pp. 193–204, Mar. 2020, doi: 10.1016/j.patrec.2019.12.024.

[13] X. Luo, J. Zhang, Z. Li, and R. Yang, "Diagnosis of ulcerative colitis from endoscopic images based on deep learning," *Biomed. Signal Process. Control*, vol. 73, Mar. 2022, Art. no. 103443.

[14] Y. Afriyie, B. Weyori, and A. Opoku, "Gastrointestinal tract disease recognition based on denoising capsule network," *Cogent Eng.*, vol. 9, no. 1, 2022, Art. no. 2142072.

[15] S. Igarashi, Y. Sasaki, T. Mikami, H. Sakuraba, and S. Fukuda, "Anatomical classification of upper gastrointestinal organs under various image capture conditions using AlexNet," *Comput. Biol. Med.*, vol. 124, Sep. 2020, Art. no. 103950.

[16] Q. Su, F. Wang, D. Chen, G. Chen, C. Li, and L. Wei, "Deep convolutional neural networks with ensemble learning and transfer learning for automated detection of gastrointestinal diseases," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106054, doi: 10.1016/j.compbiomed.2022.106054.

[17] M. A. Khan, N. Sahar, W. Z. Khan, M. Alhaisoni, U. Tariq, M. H. Zayyan, Y. J. Kim, and B. Chang, "GestroNet: A framework of saliency estimation and optimal deep learning features based gastrointestinal diseases detection and classification," *Diagnostics*, vol. 12, no. 11, p. 2718, Nov. 2022.

[18] S. Özkaya and U. Özkaya, "Gastrointestinal tract classification using improved LSTM based CNN," *Multimedia Tools Appl.*, vol. 79, nos. 39–40, pp. 28825–28840, Oct. 2020.

[19] M. B. Haile, A. O. Salau, B. Enyew, and A. J. Belay, "Detection and classification of gastrointestinal disease using convolutional neural network and SVM," *Cogent Eng.*, vol. 9, no. 1, Dec. 2022, Art. no. 2084878.

[20] M. Sharif, M. A. Khan, M. Rashid, M. Yasmin, F. Afza, and U. J. Tanik, "Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images," *J. Experim. Theor. Artif. Intell.*, vol. 33, no. 4, pp. 577–599, Jul. 2021.

[21] K. Ramamurthy, T. T. George, Y. Shah, and P. Sasidhar, "A novel multi-feature fusion method for classification of gastrointestinal diseases using endoscopy images," *Diagnostics*, vol. 12, no. 10, p. 2316, Sep. 2022.

[22] S. Biradher and P. Aparna, "Classification of wireless capsule endoscopy bleeding images using deep neural network," in *Proc. IEEE Delhi Sect. Conf. (DELCON)*, Feb. 2022, pp. 1–4, doi: 10.1109/DELCON54057.2022.9753487.

[23] Y. Cao, W. Yang, K. Chen, Y. Ren, and Q. Liao, "Capsule endoscopy image classification with deep convolutional neural networks," in *Proc. IEEE 4th Int. Conf. Comput. Commun. (ICCC)*, Dec. 2018, pp. 1584–1588, doi: 10.1109/CompComm.2018.8780859.

[24] W. Wang, X. Yang, X. Li, and J. Tang, "Convolutional-capsule network for gastrointestinal endoscopy image classification," *Int. J. Intell. Syst.*, vol. 37, no. 9, pp. 5796–5815, Sep. 2022, doi: 10.1002/int.22815.

[25] M. A. Ayidzoe, Y. Yu, P. K. Mensah, J. Cai, E. Y. Baagyere, and F. U. Bawah, "SinoCaps: Recognition of colorectal polyps using sinogram capsule network," *J. Intell. Fuzzy Syst.*, vol. 44, no. 2, pp. 3079–3091, 2023, doi: 10.3233/JIFS-212168.

[26] M. A. Khan, K. Muhammad, S. H. Wang, S. Alsubai, A. Binbusayyis, A. Alqahtani, A. Majumdar, and O. Thinnukool, "Gastrointestinal diseases recognition: A framework of deep neural network and improved moth-crow optimization with DCCA fusion," *Hum.-Centric Comput. Inf. Sci.*, vol. 12, p. 25, May 2022.

[27] H. Fang, J.-Q. Liu, K. Xie, P. Wu, X.-Y. Zhang, C. Wen, and J.-B. He, "MR-CapsNet: A deep learning algorithm for image-based head pose estimation on CapsNet," *IEEE Access*, vol. 9, pp. 141245–141257, 2021, doi: 10.1109/ACCESS.2021.3119615.

[28] M. Tubishat, N. Idris, L. Shuib, M. A. M. Abushariah, and S. Mirjalili, "Improved SALP swarm algorithm based on opposition based learning and novel local search algorithm for feature selection," *Exp. Syst. Appl.*, vol. 145, May 2020, Art. no. 113122, doi: 10.1016/j.eswa.2019.113122.

[29] L. Tightiz, S. Mansouri, F. Zishan, J. Yoo, and N. Shafaghatian, "Maximum power point tracking for photovoltaic systems operating under partially shaded conditions using SALP swarm algorithm," *Energies*, vol. 15, no. 21, p. 8210, Nov. 2022.

[30] H. Qi, S. Xie, Y. Chen, C. Wang, T. Wang, B. Sun, and M. Sun, "Prediction methods of common cancers in China using PCA-ANN and DBN-ELM-BP," *IEEE Access*, vol. 10, pp. 113397–113409, 2022, doi: 10.1109/ACCESS.2022.3215706.

[31] Z. M. Lonseko, P. E. Adjei, W. Du, C. Luo, D. Hu, L. Zhu, T. Gan, and N. Rao, "Gastrointestinal disease classification in endoscopic images using attention-guided convolutional neural networks," *Appl. Sci.*, vol. 11, no. 23, p. 11136, Nov. 2021, doi: 10.3390/app112311136.