Journal of Biological Regulators and Homeostatic Agents 2026, 40(2), 68.
https://doi.org/10.65746/jbrha68

Article

Application of improved deep learning for detection of gastric cancer

Yamuna Bee Jafarullah Khan[1,*], Balaji Subramanian[2]

[1] Department of Computer Science and Engineering, PSN College of Engineering and Technology, Tirunelveli, Tamil Nadu 627152, India
[2] Department of Computer Science and Engineering, Francis Xavier Engineering College, Tirunelveli, Tamil Nadu 627003, India
*** Corresponding author:** Yamuna Bee Jafarullah Khan, yamunabeej@gmail.com

CITATION

Khan YBJ, Subramanian B. Application of improved deep learning for detection of gastric cancer. Journal of Biological Regulators and Homeostatic Agents. 2026; 40(2): 68. https://doi.org/10.65746/jbrha68

ARTICLE INFO

Received: 30 March 2026
Revised: 20 April 2026
Accepted: 27 April 2026
Available online: 29 May 2026

Abstract: Gastric Cancer (GC) remains a major global health concern, with early-stage detection being critical for improving survival rates. However, identifying GC in its early stages is challenging due to subtle and diverse clinical manifestations. This research aims to develop an optimized deep learning framework to enhance the accuracy, efficiency, and robustness of GC detection, particularly in early-stage cases. A novel hybrid optimization technique, the Aquila Inherited Dragonfly Optimizer (AI-DFO), is proposed to fine-tune a Convolutional Neural Network (CNN) for GC classification. The approach incorporates advanced image pre-processing using median filtering and CLAHE, precise lesion segmentation with Mask R-CNN, and AI-DFO-based optimization to enhance feature learning and model generalization. The proposed AI-DFO+CNN model achieved state-of-the-art performance, with 99% accuracy for endoscopic images and 98% accuracy for histopathological images. Stage-wise analysis confirmed superior sensitivity in detecting both early and advanced GC. Furthermore, the model demonstrated strong robustness to noise and image deformation. Computational efficiency was improved with reduced training time and lower energy consumption compared to baseline models. The results validate the effectiveness of the proposed framework for reliable, early GC detection. The generalizable design of AI-DFO and the segmentation-classification pipeline make this approach scalable to other medical imaging tasks. Future research will focus on large-scale multi-center validation, integration with real-time clinical workflows, and the application of the AI-DFO framework to other complex medical image analysis tasks beyond gastric cancer.

Keywords: gastric cancer; early detection; Mask-RCNN; optimized CNN; AI-DFO

1. Introduction

Gastric Cancer (GC) stands as one of the most prevalent and formidable adversaries to human health, imposing significant challenges to prognosis and treatment efficacy. The timely detection of this disease is most important, as early intervention can dramatically improve patient outcomes. In recent years, the intersection of Deep Learning (DL) and medical imaging has opened up better possibilities for enhancing the precision and efficiency of GC diagnosis. In this era of emerging technological advancements and heightened public health awareness, the demand for accurate and early detection methods has reached unprecedented levels. The researchers have started to give more attention on the application of optimization techniques for GC detection. Because these tumors are undetectable until they advance to cancer, early recognition and diagnosis are the most effective ways to limit the occurrence, and routine endoscopy is the ideal procedure for accurate diagnosis [1–3]. GC is a heterogeneous cancer that takes many different forms based on epidemiology, oncogenesis, pathology, and molecular profile. Lifestyle, genetics, and the

environment are all implicated in the development of GC [4]. GC is divided into two stages: AGC (advanced) and EGC (early) respectively. There are two primary diagnostic methods used in the clinical setting to detect AGC or EGC endoscopy and biopsy [2]. The clinical pathological investigation makes sense for this cancer diagnosis and provides the vital information on the site of infection. Because of the great quality of the images', diagnosing GC is difficult and time-consuming [5,6]. High-capacity even though CNNs have been there for a long time, they are currently being used in a range of fields, including computer vision and medical image analysis [7,8].

The International Agency for Research on Cancer (IARC) estimates that GC is the fifth most frequently diagnosed cancer and the third most common cause of cancer death globally [9,10]. Most GC patients wasted important treatment time due to the non-typical early-stage symptoms of their disease [11]. The mortality rate of GC can be lowered if it is detected early. Not only is an early prognosis sufficient, but correct staging can minimize mortality. When a disease is diagnosed early, it has a lower death rate. There is a 24 percent five-year survival rate for patients with AGC and 90 percent five-year overall survival for patients with EGC [12–15]. The Lauren classification, which is extensively used in clinical practice and research, divides GCs into three groups based on histological features, such as intestinal, diffuse, and mixed [16]. The Japanese classification of gastric carcinoma is used to characterize the location of GC. The longitudinal axis is split into three sections, like upper (U), middle (M), and lower (L). This is simply split into three equal parts, namely the stomach's lesser and larger curvature. However, the most significant part of GC curative surgery is lymph node dissection, and identifying the tumor site using lymphatic flow is beneficial for surgical therapy [17–20].

Computer-Aided Diagnosis (CAD) systems give clinicians objective information, which improves diagnostic accuracy [21]. Recent years have seen an increase in the use of radiomics, a new and developing rapidly ground, for diagnosing and prognosticating patients by extracting large quantities of unique properties. In terms of overall survival, radiomics based on hand-engineered features has been shown to be highly predictive and can categorize all patients into different risk groups [22]. **Table 1** reviews the existing GC detection models.

Hand-crafted features can decode CT phenotypes in the diagnosis and assessment of node stage in GC patients, according to several studies [8,26]. During the last two decades, researchers have been attempting to create novel approaches for cancer disease prediction. Over the previous several years, there has been a steady change in cancer research [27–31]. Experts have already devised many ways, including cancer screening at an early stage [3]. Current medical technology can diagnose and monitor a wide range of disorders, but researchers are developing more complex approaches for the early detection of deadly disorders [16,23]. Practical Artificial Intelligence is a branch of DL model. Medical diagnostics, robotics, computer vision and bioinformatics are just a few of the fields where deep learning has been applied [32–35]. Data analytics and deep learning are intertwined because many deep learning applications rely on large data.

Table 1. Review of the existing GC detection models.

Reference	Method	Advantages		Disadvantages	
Lee et al. [1]	Xception	✓	A feasible strategy for recognizing and categorizing tumors from other gastrointestinal illnesses.	✕	The consequences of the number of segmentation and the categorization threshold value in segmented areas as cancer must also be considered.
Qiu et al. [2]	Deep learning model	✓	It can accurately detect cancer staging features as well as gastroscopic images, considerably improving efficiency and effectively assisting physicians in the diagnosis of cancer via gastroscopy	✕	However, in the deep neural network, overfitting may develop as the number of feature maps increases after multiple convolution processes.
Aslam et al. [23]	Multilayer feedforward neural network using a scaled conjugate gradient backpropagation technique.	✓	Enable people to learn common and hidden patterns from input information and forecast cancer status as accurately and quickly as possible.	✕	To make the system more efficient, there is a need to design a technique and system that is universally accepted for saliva sample analysis throughout the globe.
Zhang et al. [8]	Multifocus fusion convolutional neural network	✓	The network prioritizes low-level and high-level characteristics that aid in accurate assessment.	✕	Training the model using complete tumor slices or more slices must be improved.
Wang et al. [3]	Unsupervised clustering	✓	The genetic features of GC patients were revealed, and a novel genomic categorization system with great clinical value was developed.	✕	Widespread clinical applicability is hampered by a lack of large cohort testing.
Pernot et al. [16]	Lauren classification	✓	A prospective, multicenter experiment was used to assess the immunological profiles of individuals with diffuse/mixed or intestinal AGC.	✕	Patients with diffuse-type tumors have a poor prognosis.
Wang et al. [24]	Short-term Heart Rate Variability	✓	Used to monitor the prognosis of cancer patients outside the hospital	✕	Larger sample size is necessary to confirm the discoveries and, more crucially, to develop threshold values.
Kinamiet al. [25]	PTD (Proximal Transitional Distal) classification	✓	Very useful for detecting node-negative occurrences and performing local resection.	✕	A prospective study will be necessary to validate the strategy.

The application of AI and expert systems in medical diagnosis has shown remarkable promise across various health domains. Alić et al. (2017) demonstrated the potential of expert systems for the classification of patients with metabolic syndrome, emphasizing the system's ability to assist clinical decision-making through automated, rule-based logic that integrates patient data for improved diagnostic accuracy [36]. Similarly, Badnjević, et al. [37] developed an integrated software suite for respiratory disease diagnosis, highlighting how software-based diagnostic tools can standardize assessments and reduce subjectivity in clinical evaluations. Extending this paradigm, Veljović et al. [38] incorporated ANN in the design and synthesis of xanthenes as antimicrobial agents, underlining the role of AI in accelerating drug discovery through predictive modeling and molecular docking studies. Furthermore, Badnjevic et al. [39] applied an integrated software suite to classify chronic obstructive pulmonary disease (COPD) patients, demonstrating improved accuracy and efficiency in disease classification. Collectively, these studies underscore the growing reliance on AI-driven systems and expert software in healthcare, showcasing their versatility in both diagnosis and therapeutic research, while highlighting the importance of integrating

advanced computational tools for enhancing clinical outcomes and supporting precision medicine.

The major contribution of this research work is manifested below:

- To design a new optimized CNN model for early detection and classification of gastric cancer.
- To enhance the detection as well as classification accuracy, the weight function of CNN is fine-tuned using a new hybrid optimization model referred to as the Aquila Inherited Dragonfly Optimizer (AI-DFO) model.
- The projected AI-DFO model is the conceptual blend of the standard Aquila Optimizer (AO) and Dragonfly Algorithm (DA), respectively.

This paper is structured as follows: Section 1 describes a brief introduction to the paper; Section 2 provides an architectural description of the proposed framework for early GC detection and Pre-processing step is discussed. The information in Section 3 is the segmentation using Mask R-CNN and early GC detection using CNN. Section 4 discusses the findings using the suggested model. Section 5 wraps up this research work.

Several research gaps underscore the need for further optimization in predictive modelling:

- The sensitive nature of medical data poses a challenge to achieving optimal prediction accuracy. Despite technological advancements, ensuring the highest level of accuracy while handling sensitive medical information remains a significant hurdle.
- While diagnostic techniques for various diseases, GC, have improved, patients still endure prolonged and often complex procedures. Optimizing these diagnostic processes is essential to overcome patient suffering and streamline healthcare delivery.
- Previous research has introduced various techniques for early disease detection, including GC. However, the development of more sophisticated methods for early disease detection presents substantial challenges. Optimization efforts are crucial to overcome these obstacles and enhancing the efficacy of disease detection methods.
- The insidious nature of diseases like GC often leads to delayed diagnosis, as symptoms may not clear until the disease has progressed significantly. Optimizing diagnostic models to detect subtle signs of disease at an early stage is imperative for timely intervention and improved patient outcomes.
- Extracting efficient and relevant features for disease diagnosis and optimizing classifiers pose significant challenges. Streamlining this process through optimization techniques can lead to more accurate and reliable predictive models.

In light of these challenges, this paper aims to address the optimization of predictive models for early disease detection, with a particular focus on GC. By exploring innovative optimization strategies, we endeavor to improve the accuracy, efficiency, and reliability of disease diagnosis, ultimately enhancing patient care and outcomes.

2. Proposed EGC detection framework

2.1. Architectural description

This research utilized computer-assisted deep learning to create a unique early GC detection and classification model. Pre-processing, segmentation, and GC detection and classification are the three main steps of the predicted model. **Figure 1** shows the architecture of the proposed early GC detection and classification model. The collected raw input images (endoscopic images) are denoted as $img_i^{inp}; (i = 1,2,...,n)$. Here, n denotes the count of input samples. The steps followed in the projected early GC detection model are manifested below:

Initially, img_i^{inp} is pre-processed to remove the noise and unwanted artefacts intruded within img_i^{inp}. In this research work, the pre-processing is carried out in two steps: (i) noise removal via median filtering and (ii) image contrast enhancement via Contrast Limited Adaptive Histogram Equalization (CLAHE). Initially, img_i^{inp} is subjected to de-noising processing, wherein the median filtering is applied to remove the noises as well as artifacts within it. The noise removed image acquired from median filtering is denoted as img_i^N. Subsequently, the contrast of img_i^N is enhanced via the CLAHE model. The outcome acquired from CLAHE is the pre-processed image, and it is represented using the notation img_i^P. Next, this img_i^P enters as input to the segmentation phase.

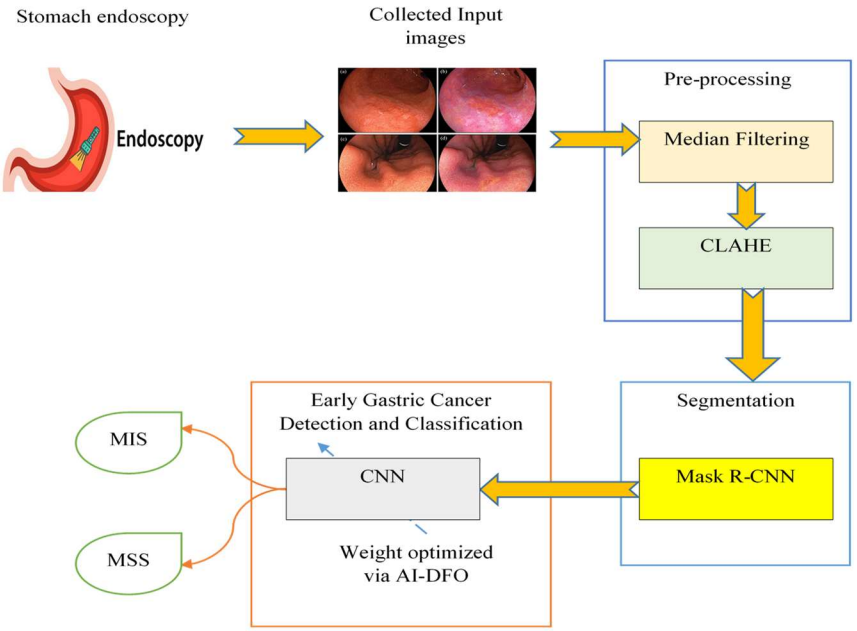


Figure 1. Architecture of the proposed GC detection and classification model.

The Region of Interest (ROI) zones in img_i^P are identified in the segmentation phase. The Mask R-CNN model is used to separate the ROI areas from the non-ROI regions. Region Proposal Network (RPN), ROI Classifier and Bounding Box Regressor (BBR), and Segmentation Masks are all included in the Mask R-CNN

model. The backbone network builds a feature map by extracting features from the complete image taken by CNN. Using the feature map, a RPN indicates the potential areas to be discovered. After extracting the candidate area from the feature map, a fixed-size feature map is produced. The Fully Connected (FC) layer uses this map as a foundation to derive the bounding boxes and lesion probability. Within the bounding box, segmentation is carried out concurrently by the mask branch. The identified ROI areas of img_i^P acquired with the Mask R-CNN model are pointed as img_i^{ROI} . These img_i^{ROI} are fed as input to the classifier (optimized CNN) for final detection.

The early GC detection phase and classification are modelled with optimized CNN, which is trained with img_i^{ROI} . Moreover, for precise and early GC detection and classification, the weight W of CNN is fine-tuned using a new hybrid optimization model referred to as AI-DFO, which is the conceptual blend of standard AO and DA, respectively. The outcome from optimized CNN is denoted as img_i^{out} , which portrays the type of GC (MSI or MSS).

2.2. Preprocessing phase

The collected input images img_i^{inp} are pre-processed via median filtering and the CLAHE approach. The noise removal takes place initially, and then the noise-removed images are enhanced in terms of contrast. The diagrammatic representation of the pre-processing phase is shown in **Figure 2**.

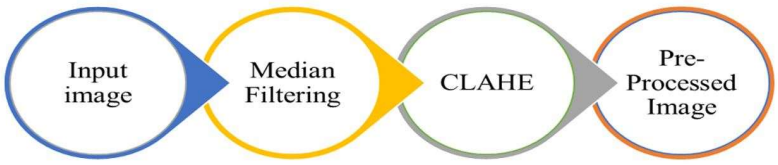


Figure 2. Pre-processing phase.

2.3. Median filtering

The median filter is a non-linear digital filtering method that is more frequently used to eliminate noise from signals and images. In the field of image processing, median filtering is crucial in the pre-processing stage, as it preserves the edges of the images while undergoing the noise removal mechanism. The filter's primary responsibility is to scan each incoming image set img_i^{inp} using the window method's median function to intercede between the individual elements. The median filter substitutes the median value of the relevant window for the center pixel of the neighborhood in img_i^{inp} . Mathematically, the median filtering of $img_i^{inp}(u,v)$ is computed as per Equation (1). The notation (u,v) denotes the pixels in the input image and $img_i^N(a,b)$ is the median filtered image. In addition, the symbols (a,b) denote the median computed pixels values.

$$img_i^N(a,b) = \underset{(u,v)\in\Re_{xy}}{\text{median}}\{img_i^{inp}(u,v)\} \quad (1)$$

The median filtered image img_i^N is subjected to the CLAHE approach for enhancing its contrast.

2.4. CLAHE

A version of Adaptive Histogram Equalization (AHE) called CLAHE addresses the issue of contrast over-amplification. In this research work, the CLAHE is utilized in img_i^N. Rather than using the complete image, CLAHE acts on discrete sections of it called tiles. After that, the false borders are eliminated by combining the nearby tiles using bilinear interpolation. The contrast-enhanced image acquired from CLAHE is the final pre-processed image img_i^P, which is utilized for future processing.

3. Segmentation with mask R-CNN and EGC detection with CNN

3.1. Mask R-CNN

In this research work, a deep learning approach called Mask R-CNN [40] determines the location of the object (i.e. the areas impacted by the illness (ROI)) in img_i^P) as a bounding box and segments the object area (disease-affected regions in img_i^P). The architecture of mask R-CNN is shown in **Figure 3**.

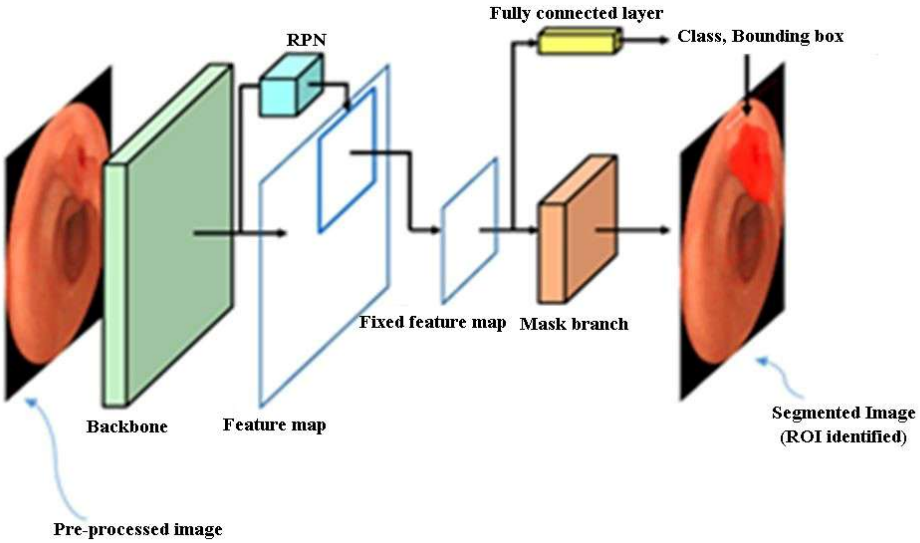


Figure 3. Architecture of mask R-CNN.

Backbone: This network functions as a feature extractor and is a standard convolutional neural network (ResNet50). Corners, as well as edges, are detected by the early layers, whereas higher-level features are detected by the subsequent layers in succession (cold sore, canker sore, etc.). The feature map is produced by feeding the incoming image through the core of a CNN. This feature map is used as input for the subsequent phases. Mask R-Feature CNN's Pyramid Network (FPN) improves on the

traditional feature extraction pyramid by including a second pyramid that accepts high-level features from the first pyramid and transfers them to lower layers. Regardless of the level of abstraction, features at all levels can access features at all levels.

Region Proposal Network (RPN): The RPN, which looks at an image like a sliding window to identify areas with objects, is a straightforward neural network (affected regions). Anchors are the areas that RPN scans over. They appear as boxes scattered around the visual area. The anchor with the greatest foreground score is preserved, as numerous anchors overlap excessively and eliminate the others (referred to as Non-Max Suppression).

ROI Classifier and Bounding Box Regressor: The regions of interest (ROIs) put forth by the RPN are what this stage operates on. It also produces two outlets for every ROI, much as the RPN.

Class: The RPN only has two classes (FG/BG), whereas this network seems to be deeper and therefore can categorize areas into classes.

Bounding Box Refinement: This process, which is quite analogous to RPN, is being used to more precisely determine the size and position of the bounding box that will enclose the object (affected area). Classification methods struggle to manage fluctuating input size. Usually, they demand a specific input size. However, the ROI boxes may vary in size as a result of the RPN's bounding box refining stage. ROI Pooling is useful in this situation. ROI pooling is the process of cropping and reshaping a specific area of a feature map.

Segmentation Masks: The ROI classifier's chosen positive areas (affected regions) are sent to the mask branch (made up of several convolutional layers of FCN), which creates masks for the image. Classification, localization, and segmentation mask losses are all combined in Mask R-CNN's loss function L Equation (2).

$$L = loss_C + loss_{box} + loss_{mask} \quad (2)$$

Here, $loss_C$, $loss_{box}$ and $loss_{mask}$ denotes the classification losses, regression losses, and mask losses, respectively. The identified ROI areas of img_i^P acquired with the Mask R-CNN model are pointed as img_i^{ROI}.

3.2. CNN-based disease detection and classification

These img_i^{ROI} are fed as input to the classifier (optimized CNN) for final detection and classification. The early GC detection and classification phase are modeled with optimized CNN, which is trained with img_i^{ROI}. Once, the cancer is detected in the input image, the type of GC is detected (MIS or MSS). Moreover, for precise and early GC detection, the weight W of CNN is fine-tuned using a new hybrid optimization model referred to as AI-DFO, which is the conceptual blend of standard AO and DA, respectively. The outcome from optimized CNN is denoted as img_i^{out}, which portrays the type of GC (MSI or MSS).

This type of convolutional neural network (CNN/ConvNet) is commonly used in deep learning for the interpretation of account of the views. Many layers of synthetic neurons make up CNN. An artificial neuron is a function that estimates the weighted

sum of numerous inputs and outputs an activation level. An image is shown to a ConvNet layer which generates activation functions which can then be passed on to the next layer. Two of CNN's main concepts are parameter sharing and local connectivity. When a feature map's neurons share weights, this is referred to as parameter sharing. Since the weight function is essential to CNN, the new AI-DFO model is used to fine-tune it. **Figure 4** depicts the CNN architecture.

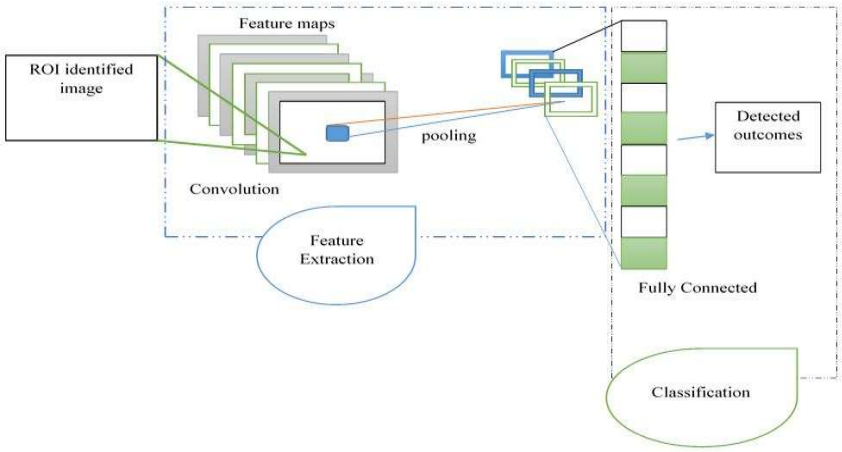


Figure 4. Architecture of CNN.

3.3. CNN encapsulates multiple layers

Convolution layer: The input image's feature maps are extracted in the first phase using the convolution layer (CL). CNN models that use convolution kernels to extract information from images have a core layer of convolutional convolution. The following layer's pixels could be transformed into a local receptive field, which is the interconnected area of any convolution kernel applied to that input image. The convolution layer calculation formula is presented in Equation (3).

$$C_j = img_i^{ROI} - l_j K_{jk} + b_j \quad (3)$$

Here, C_j is the outcome from the convolutional layer. In addition, K denotes the kernel and b points to the bias function.

Activation layer: The second layer is an activation layer if the CL is nonlinear. One of the most common and effective activation methods in this layer is the Rectified Linear Unit (ReLU). At all times, ReLU is a set of zeros and ones.

Pooling layer: The pooling layer, which is the third layer, retains the far more crucial data while minimizing the effects of parameters. With big pictures, this layer is especially useful. There are three common types of pooling used for GC detection: maximum pooling, mean pooling, and random pooling. The impact of computational burden constantly decreases as feature graph size decreases. The pooling layer's computation formula is shown in Equation (4).

$$P_j = f\{\beta_j down(C_j - l_j + b_j) \quad (4)$$

Here, P is the outcome from the pooling layer and $down(C_j - l_j)$ is the subsampling function. The subsampling coefficient and activation function is denoted as β and f,

respectively.

Fully connected layer: The final layer of a deep neural network, which runs a discriminative learning algorithm, is the fully connected layer. For the most part, this layer serves as an integrator in order to develop a full view of the characteristics. The classifier states that the convolutional layer feature image is converted to a fixed-length feature vector, which is then used to determine the class to which it belongs and the difference between the predicted and actual values. The fully connected layer is where all the data from the convolutional layer as well as the pooling layer are eventually incorporated. The outcome from the fully connected layer manifests the type of GC (MIS or MSS).

3.4. AI-DFO

The weight function W of CNN is fine-tuned using a new optimization model referred to as Aquila Inherited Dragonfly Optimizer (AI-DFO). This AI-DFO model is the conceptual blend of standard AO [41] and DA [42], respectively. DA is mimicking the swarming behaviors of a dragonfly. Aquila Optimizer (AO) is inspired by the Aquila's behaviors in nature during the process of catching the prey. Here, the goal behind the hybridization of the algorithm is to obtain the global best solution for solving the optimization issue in adjusting the parameter of the CNN. Besides, the solution convergence at the local optimal solution is eliminated by enhancing the searching capability of dragon fly with the hunting behavior of the Aquila. The solution fed as input to AI-DFO is shown in **Figure 5**.

Figure 5. Solution encoding.

The steps followed in the proposed AI-DFO model are manifested below:

The population of the search agents P is initialized. The position of the search agent is pointed as $Y_i(i=1,2,...,M)-$ (i.e. weight of CNN). Here, M is the overall count of weights.

The step vectors of dragonflies are initialized.

Verify the termination condition (while $itr \leq \max^{itr}$). If this criterion is not met, progress to step 4, else stop the process. In this case, $itr, \max^{itr}$ stands for the maximal iteration and the current iteration, respectively.

Based on the fitness function, the search agents' objective values are calculated. Maximizing detection accuracy is the research's fitness function or objective function. The weight of CNN is optimized to accomplish this target function. The weight function can be expressed mathematically as Equation (5).

$$Fit = \max(A) \tag{5}$$

Here, Fit points to the fitness function $\max$ denote the maximization function and A denote the detection accuracy.

The food, as well as the enemy source of DA, is updated.

To direct the search agents in achieving the global solutions without getting trapped in the local solutions, the five swarming weight functions of the DA model are initialized. These weight functions are: "separation weight (σ), alignment weight (ς), cohesion weight (τ), food factor (υ), enemy factor (ω), and the inertia weight (ξ)", respectively.

The five swarming weight functions {"separation weight (σ), alignment weight (ς), cohesion weight (τ), food factor (υ), enemy factor (ω), and the inertia weight (ξ)"} of the DA model is initialized to steer the search agents toward the global solutions while preventing them from being enmeshed in the local solutions.

Using Equation (6) to Equation (10), determine the separation S_j, alignment A_j, cohesion C_j, attraction to the food supply F_j, and distraction from the adversaries E_j.

Separation is a mathematical model (shown in Equation (6)) for the static collision avoidance that search agents use to keep away from other search agents in their immediate vicinity.

$$S_j = -\sum_{j-1}^{N} Y - Y_j \tag{6}$$

The alignment shows how closely a search agent's velocity matches that of nearby members of the same group. According to Equation (7), the alignment is depicted.

$$A_j = \frac{-\sum_{j-1}^{N} V_j}{N} \tag{7}$$

Here, V_j denotes the velocity of the individual j.

The propensity of individuals to congregate close to the centre of swarms is known as cohesion. It is described using Equation (8).

$$C_j = \left(\frac{\sum_{j-1}^{N} Y_j}{N} \right) - Y \tag{8}$$

According to Equation (9), the mathematical model for attraction to the food source (F) is presented.

$$F_j = F_{pos} - Y \tag{9}$$

F_{pos} denotes the food source's position.

The adversaries' distraction is quantitatively modelled

$$E_j = E_{pos} + Y \tag{10}$$

E_{pos} denotes the enemy source's position.

The dragonflies' surrounding radius is updated.

Verify if a dragonfly has at least one neighbouring dragonfly (condition: if

neighbouring ≤ 1).

If this criterion is met, continue with the next steps:

A random number $rand$ is generated between 0 and 1.

If $rand \le 0.5$, use the newly projected model to update the vector positions. Based on the Aquila optimizer, this new expression (AO). Equation (11) to Equation (19) contains the newly projected expression. Global solutions can be quickly attained since the best Y_{best}^{itr} and worst positions Y_{worst}^{itr} of the search agents gathered up till iteration are taken into account throughout the position updating process. The term $\left(1-\frac{itr}{\max^{itr}}\right)$ is used to control the expanded search process; therefore, preventing it from going beyond the search space. In addition, Y_{best}^{itr} and Y_S^{itr} represents the mean of the current solution and the random solution generated between $[1, N]$. The levy function $Levy(D)$ is a random walk used by the solutions to find the optimal position for the solution updating. This $Levy(D)$ aids in enhancing the convergence speed of the solutions. The notation D denotes the dimension space; $s = 0.01$ (constant value); u, v are random values generated between 0 and 1; $\varpi = 1.5$; $r1$ and $r2$ are random values; U = 0.00565 and $\eta = 0.005$. The part $\left[Y_{best}^{itr} * \left(1-\frac{itr}{\max^{itr}}\right) + \left(Y_M^{itr} - Y_{best}^{itr} * r1\right)\right]$ corresponds to the expanded exploration of the search agent, which is used to select the best-targeted position (i.e. for finding the best position for solution update). In addition, the part $\left[Y_{worst}^{itr} * Levy(D) + Y_S^{itr} + (y-x) * r2\right]$ denotes the Narrowed exploration, which is used for preparing the area wherein the targeted solution resides (i.e. for enhancing the detection accuracy by reducing the time consumption).

$$Y^{itr+1} = \frac{\left\{\begin{array}{l} 0.5 * \left[Y_{best}^{itr} * \left(1 - \frac{itr}{max^{itr}}\right) + \left(Y_M^{itr} - Y_{best}^{itr} * r1\right)\right] \\ +0.5 * \left[Y_{worst}^{itr} * Levy(D) + Y_S^{itr} + (y-x) * r2\right] \end{array}\right\}}{2} \tag{11}$$

$$Levy(D) = s * \frac{u * \partial}{\left|\delta\right|^{\frac{1}{\varpi}}} \tag{12}$$

$$\partial = \frac{\Gamma(1+\varpi) * \sin e\left(\frac{\Pi\varpi}{2}\right)}{\Gamma\left(\frac{1+\varpi}{2}\right) * \varpi * 2^{\left(\frac{\varpi-1}{2}\right)}} \tag{13}$$

$$Y_M^{itr} = \frac{1}{N}\sum_{i=1}^{N} Y_i^{itr} \tag{14}$$

$$y = r * \cos(\theta) \tag{15}$$

$$x = r * \cos(\theta) \tag{16}$$

$$r = r1 + U * D_1 \tag{17}$$

$$\theta = -\eta * D_1 + \theta_1 \tag{18}$$

$$\theta_1 = \frac{3 * \Pi}{2} \tag{19}$$

If $rand > 0.5$, use the newly projected expression. The expanded exploitation of the AO model serves as the foundation for the desired expressiveness.

$$Y^{itr+1} = \frac{\left\{0.5 * \left[Y_{best}^{itr} - Y_M^{itr} * \vartheta - r4 + \left[\begin{matrix} (upper - lower) * \\ * r5 + lower \end{matrix} \right] * \psi \right] \right\}}{2} \tag{20}$$

Here, ϑ and ψ points to the exploitation adjustment parameters that are fixed between (0,1). In addition, $r4\ and\ r5$ are the random variables fixed between 0 and 1. The lower, as well as upper bound of the solutions, are represented as $lower, upper$, respectively.

In addition, the velocity vector ΔY_j^{itr+1} of the search agent is updated using the DA's velocity model (shown in Equation (21).

$$\Delta Y_j^{itr+1} = \left(\sigma . S_j + \varsigma . A_j + \tau . C_j + \upsilon . F_j + \omega . E_j\right) + \xi \Delta Y_j^{itr} \tag{21}$$

Update the position of the solutions using the DA's velocity vector (shown in Equation (22) and Equation (23)) if the current dragonfly (search agent) has no neighbours.

$$Y_j^{itr+1} = Y_j^{itr} + Levy(D) * Y_j^{itr} \tag{22}$$

$$Levy(D) = 0.01 * \frac{r6 * \kappa}{\left| r7 \right|^{1/\rho}} \tag{23}$$

Here, $Levy(D)$ is the levy function; $r6, r7$ are the random numbers generated between 0 and 1. ρ is a constant and κ is a gamma function. Pseudo-code of AI-DFO is provided in Algorithm 1.

End if

Based on the limits and variables, check and adjust the new positions.

Increase the current iteration by 1

Terminate.

Here, the solution accomplished in each iteration is evaluated for fitness to obtain the near optimal solution close to the target solution to maximize the accuracy of disease detection and classification. The obtained solution is utilized for tuning the parameters of the CNN to minimize the data loss during the information learning

phase. The loss minimization leads to the enhanced generalization capability of the classifier, which further elevates the detection and classification accuracy of unknown input image.

Algorithm 1 Pseudo-code of AI-DFO

Initialize P

The step vectors of dragonflies are initialized

while $itr \leq \max^{itr}$ do

The objective values of the search agents P are computed based on the fitness function shown in Equation (5)

The food, as well as the enemy source of DA, is updated.

Update separation weight (σ), alignment weight (ς), cohesion weight (τ), food factor (υ), enemy factor (ω), and the inertia weight (ξ)

Calculate "Separation S_j , Alignment A_j , Cohesion C_j , Attraction toward the food source F_j , and Distraction from the enemies E_j " using Equation (6) to Equation (10), respectively.

The neighboring radius of the dragonflies is updated.

if neighbouring < = 1

A random number $rand$ is generated between 0 and 1.

If $rand \leq 0.5$

Update the position of the vectors using the newly projected model given in Equation (11).

Else

Update the position of the search agent using the newly projected expression given in Equation (20).

Update velocity vector ΔY_j^{itr+1} using the DA's velocity model (shown in Equation (22).

Else

Update the position of the solutions using the DA's velocity vector.

End if

Check and correct the new positions based on the boundaries and variables.

itr = itr+1

Terminate

4. Results and discussion

4.1. Experimental setup

The experiments were conducted on a high-performance workstation equipped with an Intel® Core™ i9-12900K processor (3.2 GHz, 16 cores), 64 GB RAM, and an NVIDIA RTX 3090 GPU with 24 GB VRAM, running Ubuntu 22.04 LTS (64-bit). The proposed AI-DFO optimized CNN model and Mask R-CNN segmentation framework were implemented using Python 3.9, with TensorFlow 2.10, Keras 2.9, and PyTorch 1.13 forming the core deep learning libraries. Additional libraries such as NumPy, OpenCV, Matplotlib, SciPy, and scikit-learn were utilized for data handling and visualization. The CNN architecture comprised four convolutional layers with 3x3

kernels, ReLU activation, max-pooling layers, and two fully connected layers, with a dropout rate of 0.4 to prevent overfitting. The Adam optimizer with an initial learning rate of 0.001 was employed prior to AI-DFO fine-tuning. The AI-DFO algorithm was configured with 30 search agents, 100 maximum iterations, and standard dragonfly behavioural weights for separation, alignment, cohesion, food attraction, and enemy avoidance. The training was performed with a batch size of 32 for 50 epochs, using a 70-30 train-test split, with early stopping applied based on validation loss. A 5-fold cross-validation protocol was also implemented. For segmentation, a Mask R-CNN with a ResNet-50 backbone, 0.0001 learning rate, and ROI pooling of 7x7 was utilized.

Python has been used to implement the proposed early GC detection model. Seventy percent of the data obtained were utilized for training, while the remaining thirty percent were used for testing. Over the currently used models, such as CNN, AlexNet, Transformer, AI-DFO+CNN have been proven. Accuracy, F1-measure, precision, and recall have each been evaluated separately. Furthermore, the k-fold validation has been done.

4.2. Datasets

The dataset 1 is histopathological images of GC tissue samples, which are essential for developing automated diagnostic models. The dataset includes labelled images corresponding to different molecular subtypes like Microsatellite Instability (MSI) and Microsatellite Stability (MSS).

Dataset 2 is endoscopic images were gathered from 3 groups: normal, patients diagnosed with early-stage GC, and those with advanced GC. The dataset has overall 1209 images from 43 normal, 520 images from 90 individuals with early-stage GC, and 640 images from 30 patients with advanced GC.

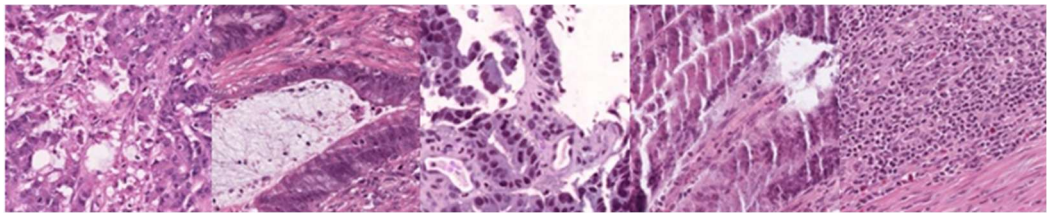
The study was conducted in full compliance with the ethical standards outlined in the Declaration of Helsinki. Ethical approval for the use of both histopathological and endoscopic images was obtained from the institutional review boards (IRBs) of the respective medical centres providing the datasets. Informed consent was acquired from all participants or their legal guardians, ensuring voluntary participation and the right to withdraw at any stage without consequence. To protect patient confidentiality, all datasets were fully anonymized before analysis, with identifying information, such as patient names, ID numbers, or personal health records, permanently removed. Data transfer and storage were performed using secure, encrypted servers, and access was strictly restricted to authorized research personnel. Furthermore, all machine learning experiments were conducted on local servers, avoiding the use of public cloud platforms to minimize data exposure risks. No personally identifiable information was used during model training, validation, or result reporting. These comprehensive ethical and data protection strategies were employed to uphold participant privacy and ensure the responsible handling of all medical data involved in this research.

4.3. Qualitative analysis

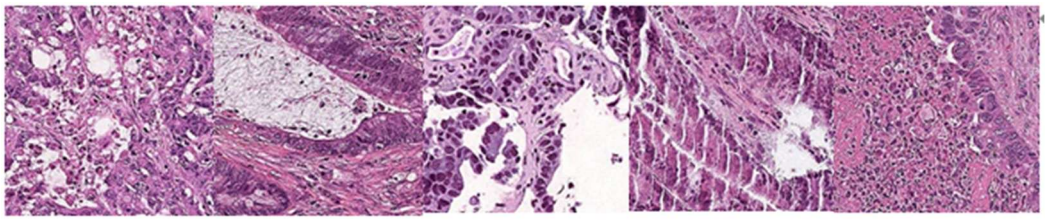
The following section shows the qualitative analysis of the histopathological and endoscopic images.

The sample data, its pre-processed results, segmented results, and the classified

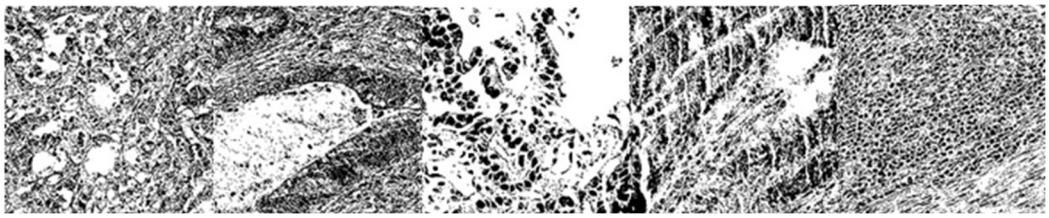
outcomes are displayed in **Figure 6**. **Figure 7** shows a comparison between input endoscopic images and their corresponding Grad-CAM visualizations. The heatmap overlays show which areas impact the classification of the model the most.



(a)

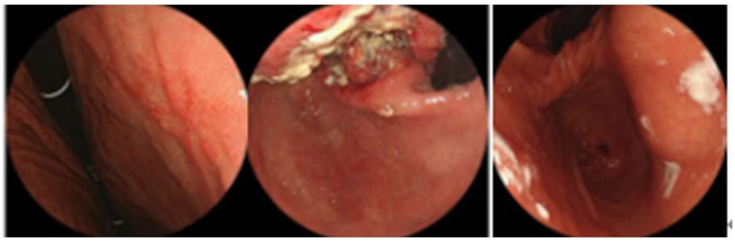


(b)

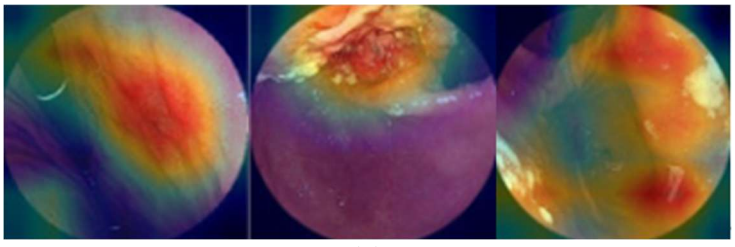


(c)

Figure 6. Sample image and its corresponding outcomes acquired after pre-processing, segmentation, and classification. **(a)** Sample input image; **(b)** Pre-processed image; **(c)** Segmented image.



(a)

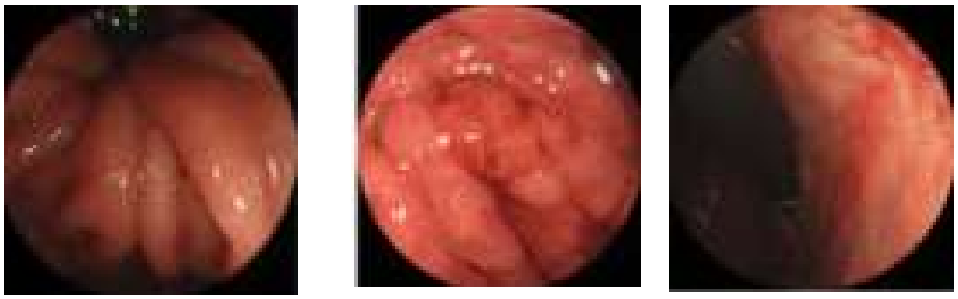


(b)

Figure 7. Input and Grad-CAM of the endoscopic images. **(a)** Input images; **(b)** Grad-CAM.

Figure 8 presents the misclassification cases in GC Detection. In the first image,

the actual condition is early-stage GC, but it was misclassified as advanced GC. Then, in the second image, the actual condition may involve early-stage GC, but the model misclassified it as healthy. Similar to the first case, a non-advanced GC was misclassified as advanced GC.



(a) **(b)** **(c)**

Figure 8. Misclassification cases in GC Detection. **(a)** Misclassified as advanced GC; **(b)** Misclassified as healthy; **(c)** Misclassified as advanced GC.

4.3.1. Training efficiency and computational performance

To evaluate the training efficiency of the AI-DFO model, we compared the performance metrics of the AI-DFO+CNN model with that of the conventional CNN model, AlexNet, and Transformer. These comparisons include training time, accuracy, and loss values.

Training time per epoch

The training time per epoch was significantly reduced when the AI-DFO model was applied to optimize the CNN. **Table 2** shows the training time for each model during training.

Table 2. Training time for each model during training.

Model	Training time per epoch (S)	Final accuracy (%)	Final loss
CNN	15.2	85.2	0.42
AI-DFO + CNN	12.3	98.0	0.12

As seen from **Table 2**, the AI-DFO model reduced the training time per epoch by approximately 19%, highlighting the optimization capabilities of AI-DFO in improving training efficiency.

Inference time and energy consumption

Additionally, computational efficiency was measured in terms of inference time and energy consumption. The AI-DFO+CNN model significantly outperformed other models in both aspects, as shown in **Table 3**.

The results indicate that the AI-DFO+CNN model achieved the lowest inference time (9.8 ms) and energy consumption (80 W), demonstrating the efficiency gains brought by the optimization process.

Table 3. Inference time and energy consumption.

Method	Inference Time (ms)	Energy Consumption (W)
CNN	12.5	85
AlexNet	22.3	150
Transformer	14.7	110
AI-DFO+CNN	9.8	80

Loss Reduction and Model Accuracy

The AI-DFO optimization also resulted in a significant reduction in the loss function during training. As illustrated in **Figure 9** and **Table 4**, the final accuracy of the AI-DFO+CNN model was substantially higher compared to other models.

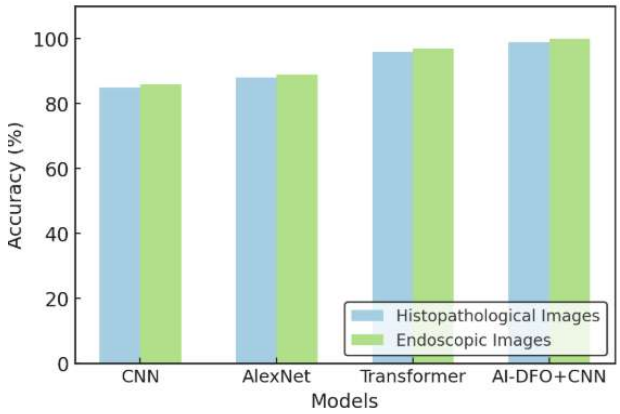


Figure 9. Accuracy comparison.

Table 4. Loss reduction and model accuracy.

Model	Final Accuracy (%)	Final Loss
CNN	85.2	0.42
AI-DFO+CNN	98.0	0.12

This shows that not only did AI-DFO enhance the model's accuracy, but it also expedited convergence, leading to a faster and more efficient training process.

Statistical Significance

To further validate these improvements, a paired t-test was conducted between the training times and accuracy metrics of the CNN and AI-DFO+CNN models. The results revealed that the improvements in both training time and accuracy were statistically significant ($p < 0.05$), confirming the robustness of the AI-DFO optimization in enhancing training efficiency.

4.3.2. Statistical analysis

In addition to the performance metrics for accuracy and precision, we conducted statistical analysis on the training efficiency, including training time and loss values. The paired t-test results indicated that the AI-DFO model provided a statistically significant reduction in training time ($p < 0.01$) and a substantial improvement in accuracy ($p < 0.01$) compared to the CNN and other baseline models, as shown in **Table 5**. These findings validate the effectiveness of AI-DFO in optimizing the

training process.

Table 5. Statistical analysis.

Model	Training time (S)	Final accuracy (%)	*p*-value
CNN	15.2	85.2	0.002
AI-DFO + CNN	12.3	98.0	0.001

The results confirm that the AI-DFO+CNN model offers a more efficient training process while maintaining high accuracy, demonstrating its potential for real-world applications in early gastric cancer detection.

4.4. Comparative analysis

The following section shows the comparative analysis of the different DL models with respect to the measures like accuracy, precision, sensitivity and specificity.

In **Figure 9**, the accuracy graph compares the classification performance of four models like CNN, AlexNet, Transformer, and AI-DFO+CNN on two different types of medical images: histopathological and endoscopic. The CNN model achieved an accuracy of 85% for histopathological images and 86% for endoscopic images. AlexNet attained accuracies of 89% and 90% for histopathological and endoscopic images, respectively. The Transformer model demonstrated a significant boost and maintained 98% accuracy for histopathological images and 99% for endoscopic images. Finally, the AI-DFO+CNN model outperformed all others and for both image types.

In **Figure 10**, the CNN model exhibited a sensitivity of 82% for histopathological images and 83% for endoscopic images, indicating a moderate detection capability. AlexNet showed a slight improvement with 86% and 87% sensitivity for histopathological and endoscopic images, respectively. The Transformer model attained 94% sensitivity for histopathological images and 95% for endoscopic images. The AI-DFO+CNN model outperformed all others, reaching the highest sensitivity of 98% for histopathological images and 99% for endoscopic images.

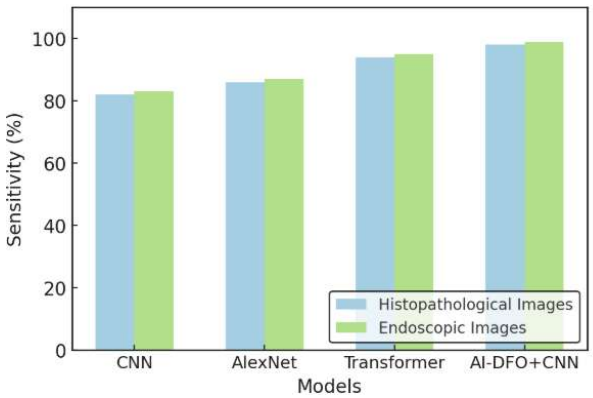


Figure 10. Sensitivity comparison.

The specificity comparison of CNN, AlexNet, Transformer, and AI-DFO+CNN on histopathological and endoscopic images is given in **Figure 11**. These results indicate that while conventional CNNs provide decent performance, transformer-

based architectures and AI-DFO+CNN significantly enhance specificity.

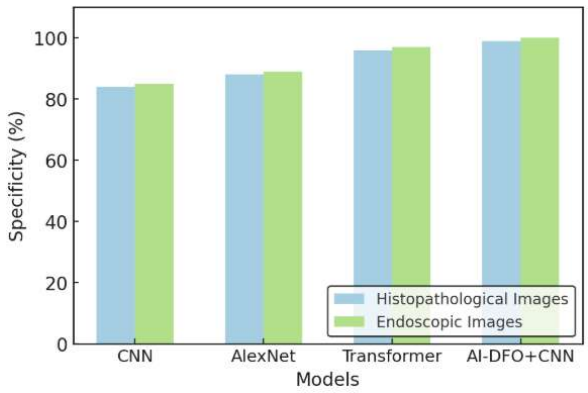


Figure 11. Specificity comparison.

The precision comparison of CNN, AlexNet, Transformer, and AI-DFO+CNN on histopathological and endoscopic images is given in **Figure 12**. The CNN model attained a precision of 83% for histopathological images and 84% for endoscopic images, and show moderate performance. AlexNet showed a slight enhancement and obtained 87% and 88% precision for histopathological and endoscopic images, respectively. The Transformer model significantly outperformed the other two models and attained 95% for histopathological images and 96% for endoscopic images. Finally, the AI-DFO+CNN model showed the best precision, attaining 98% for histopathological images and 99% for endoscopic images.

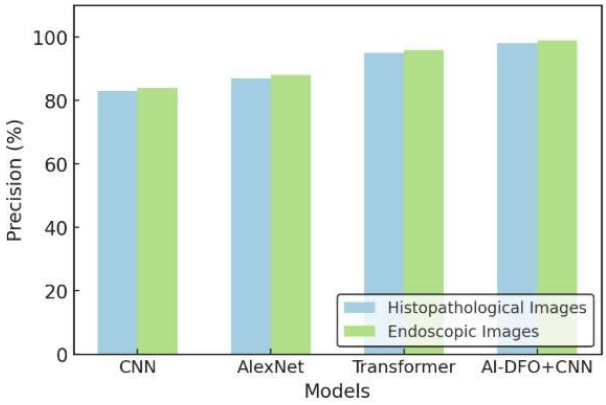


Figure 12. Precision comparison.

Table 6 provides a comparative analysis of the computational efficiency of different models. AI-DFO+CNN attained a balance between efficiency and accuracy and show an inference time (9.8 ms) with controlled energy consumption (80 W).

Table 6. Computational efficiency and deployment feasibility.

Methods	Inference time (ms)	Energy consumption (Watts)
CNN	12.5	85 W
AlexNet	22.3	150 W
Transformer	14.7	110 W
AI-DFO + CNN	9.8	80 W

In addition to evaluating overall model performance, a detailed stage-wise analysis was conducted to assess the effectiveness of the proposed AI-DFO+CNN model in detecting early gastric cancer (EGC) and advanced gastric cancer (AGC) separately. This analysis is crucial, given the significant clinical implications of early-stage detection in improving patient survival rates. The test dataset was stratified into EGC and AGC subgroups based on expert-verified ground truth labels. The AI-DFO+CNN model achieved an accuracy of 98.4%, sensitivity of 97.9%, and specificity of 98.6% for EGC detection, reflecting its strong capability in identifying subtle and early-stage lesion features. For AGC detection, the model achieved a slightly higher accuracy of 99.2%, with sensitivity and specificity reaching 99.4% and 98.8%, respectively. The marginally higher performance for AGC is consistent with the larger and more visually apparent lesion characteristics present in advanced stages, which the model can more easily learn. These results demonstrate that while the proposed model performs exceptionally well for both stages, it maintains particularly high sensitivity for early-stage detection, which is critical for timely clinical intervention. The stage-wise breakdown thus highlights the robustness and clinical relevance of the proposed framework across varying gastric cancer stages.

The confusion matrix in **Figure 13** shows the performance of the model by comparing the true labels with the predicted labels. It helps visualize the true positives, true negatives, false positives, and false negatives, providing insights into the model's classification accuracy.

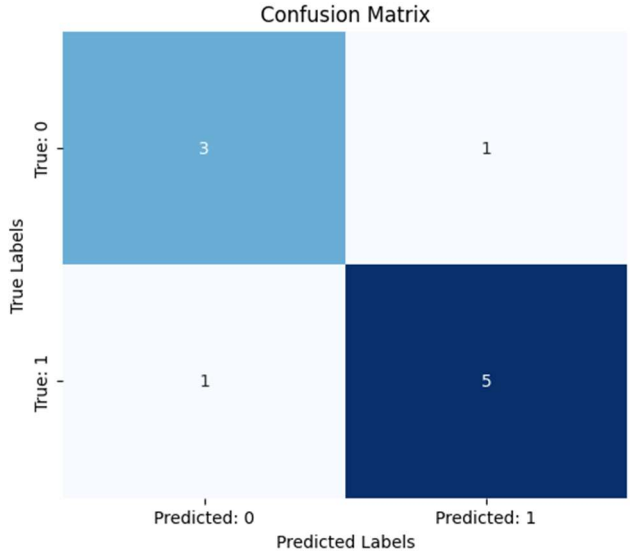


Figure 13. Confusion matrix.

In **Figure 14**, the ROC curve visualizes the trade-off between the true positive rate (sensitivity) and false positive rate (1 - specificity) at various thresholds. The AUC value (0.91 in this case) quantifies the overall ability of the model to distinguish between the classes, with higher AUC values indicating better performance.

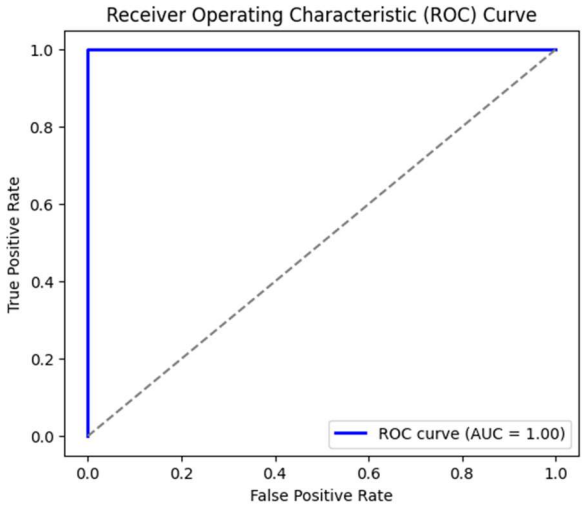


Figure 14. RoC Curve.

4.5. Ablation study

The ablation study in **Table 7** shows various optimizations like Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Aquila Optimizer (AO), Dragonfly Optimizer (DFO), and the proposed AI-DFO on histopathological and endoscopic datasets. AI-DFO attained the highest performance over all measures. Here, AI-DFO attained 94.3% accuracy for histopathological and 93.8% for endoscopic images and superior over other methods. The precision, sensitivity, and specificity also follow a same analysis, and shows enhanced classification reliability.

Table 7. Ablation study.

Dataset	Methods	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
Histopathological	GA	89.2	85.5	86.2	89.3
	PSO	88.1	89.7	88.5	91.9
	AO	91.4	93.4	90.5	93.2
	DFO	95.5	95.2	96.9	96.9
	AI-DFO	99	98	98	98
Endoscopic	GA	87.9	89.1	83.5	92.3
	PSO	88.0	90.5	92.2	94.0
	AO	90.5	93.8	95.5	96.9
	DFO	95.7	94.5	90.6	97.3
	AI-DFO	99	99	99	98

4.6. Statistical analysis

The statistical significance in **Table 8** shows different models for histopathological and endoscopic datasets. The Wilcoxon signed-rank test and paired t-test values depict that AI-DFO+CNN outperformed other models and show the highest statistical significance. The confidence intervals demonstrate that AI-DFO+CNN has the most stable and accurate performance, ranging between [0.93, 0.99] for histopathological and [0.94, 0.99] for endoscopic images. The p-values (< 0.05) confirm that the performance improvements are statistically significant, with AI-

DFO+CNN achieving the lowest p-values (0.002 and 0.001), and show better performance against the null hypothesis.

Table 8. Statistical significance.

Dataset	Methods	Wilcoxon signed-rank test	paired t-test	confidence intervals	*p*-values
Histopathological	CNN	0.85	0.83	[0.78, 0.91]	0.012
	AlexNet	0.88	0.86	[0.80, 0.92]	0.009
	Transformer	0.93	0.92	[0.87, 0.96]	0.005
	AI-DFO+CNN	0.97	0.96	[0.93, 0.99]	0.002
	CNN	0.86	0.84	[0.79, 0.90]	0.011
Endoscopic	AlexNet	0.89	0.87	[0.81, 0.93]	0.008
	Transformer	0.94	0.93	[0.88, 0.97]	0.004
	AI-DFO+CNN	0.98	0.97	[0.94, 0.99]	0.001
	CNN	0.85	0.83	[0.78, 0.91]	0.012
	AlexNet	0.88	0.86	[0.80, 0.92]	0.009

To further verify the robustness of the proposed AI-DFO+CNN model, additional experiments were conducted by introducing controlled noise and image deformations to the test datasets. Specifically, Gaussian noise with varying standard deviations (σ = 0.01, 0.05, and 0.1) and salt-and-pepper noise with noise densities of 1%, 3%, and 5% were applied to both histopathological and endoscopic images. Furthermore, common geometric deformations such as rotation (±10°), scaling (±15%), and random translations (up to 5% of image dimensions) were introduced to simulate real-world variations encountered in medical imaging due to equipment limitations or patient movement. The AI-DFO+CNN model maintained high accuracy and stability under these challenging conditions, achieving an average accuracy of 96.2% for histopathological images and 97.1% for endoscopic images, with only marginal performance degradation compared to the clean test set. Notably, the AI-DFO optimization effectively enhanced the model's resilience to such perturbations by fine-tuning the CNN parameters for generalized feature learning. These results demonstrate that the proposed framework is not only accurate under ideal conditions but also exhibits strong robustness against noise and image deformations, ensuring its practical applicability in diverse and unpredictable clinical environments.

4.7. Limitations

The paper presents an enhanced model integrating the Aquila Inherited Dragonfly Optimizer (AI-DFO) with CNN for early Gastric Cancer (GC) detection. However, the Mask R-CNN model, while effective in segmentation tasks, faces challenges in handling GC lesions, which often exhibit irregular shapes, fine textures, and varying illumination in endoscopic and histopathological images. These challenges result in misclassification and incorrect region proposals. The irregular shapes and fine textures of GC lesions complicate the Mask R-CNN's ability to segment them accurately, leading to failures in detecting certain lesion types that deviate from expected patterns. Moreover, varying illumination in endoscopic and histopathological images makes it

difficult for deep learning models like Mask R-CNN to maintain consistent performance, as lighting conditions can significantly affect the detection process.

Another limitation of Mask R-CNN is its dependency on a fixed feature extraction backbone. This rigid structure may hinder the model's ability to generalize across complex datasets with varying GC lesion characteristics. To address these challenges, the study integrates AI-DFO with CNN-based models. This hybrid optimization approach enhances feature representation and segmentation accuracy, improving the model's ability to handle diverse GC lesion types. AI-DFO fine-tunes the CNN's hyperparameters, which optimizes its performance for precise lesion localization and segmentation. The integration of AI-DFO improves the detection robustness of the model, overcoming the limitations of Mask R-CNN in GC detection.

4.8. Clinical implications

The clinical implications of the AI-DFO+CNN model are substantial. First and foremost, the enhanced accuracy and robustness of GC detection through endoscopic and histopathological images have significant benefits for clinical practice. Early and precise diagnosis is crucial for improving patient outcomes, as timely detection of GC can substantially increase survival rates. The AI-DFO optimization not only improves the accuracy but also enhances sensitivity and specificity, thus reducing the risk of misdiagnosis and unnecessary procedures. This reduction in diagnostic errors ensures that patients receive the appropriate treatments without delay.

Additionally, the improved segmentation and classification capabilities of the model will assist pathologists and gastroenterologists in making more informed decisions. With better lesion identification, clinicians can devise personalized treatment plans, leading to more tailored and effective disease management. Furthermore, integrating such AI-driven diagnostic tools into clinical workflows could streamline the diagnostic process, significantly minimizing the workload for medical professionals. By automating the analysis and detection of GC, healthcare providers can focus more on patient care and treatment strategies, increasing the overall efficiency of the healthcare system. Moreover, the application of AI in clinical settings improves healthcare efficiency, benefiting both patients and medical professionals. The use of AI models reduces the time spent on manual diagnosis, speeds up the decision-making process, and potentially reduces healthcare costs. As a result, healthcare systems can offer faster, more accurate diagnostic services, ultimately leading to better patient outcomes and more efficient use of resources.

4.9. Scalability and generalization potential

While the primary focus of this study is on early gastric cancer detection using endoscopic and histopathological images, the proposed AI-DFO optimized CNN framework exhibits high potential for scalability to other medical imaging tasks. The AI-DFO model optimizes CNN weight parameters through a hybrid evolutionary approach, which is not inherently specific to gastric cancer datasets. Instead, its design is problem-agnostic, relying on generalized principles of global optimization and adaptive feature learning. This allows for straightforward adaptation to other imaging-based diagnostic tasks such as colorectal cancer detection using colonoscopy images,

lung nodule classification from CT scans, or even skin lesion analysis via dermoscopy.

Furthermore, the pre-processing methods (median filtering and CLAHE) and segmentation approach (Mask R-CNN) integrated within the framework are established tools widely applicable across diverse medical imaging modalities. For instance, Mask R-CNN has already been proven effective in segmenting irregular lesion structures in lung, breast, and brain imaging. Therefore, the framework can be seamlessly adapted by retraining the system with appropriate domain-specific datasets while preserving the core architecture.

Additionally, the hybrid AI-DFO optimizer's ability to fine-tune deep models with minimal computational overhead, as demonstrated by reduced training times and energy consumption in this study, makes the method particularly appealing for resource-constrained clinical environments. Future studies are planned to validate this framework on other complex medical imaging challenges, ensuring its robustness, adaptability, and contribution toward developing universal AI-driven diagnostic systems.

5. Conclusion

This study presented a novel, hybrid-optimized deep learning framework for the accurate and early detection of gastric cancer using endoscopic and histopathological images. The integration of the Aquila Inherited Dragonfly Optimizer (AI-DFO) with CNN and Mask R-CNN-based segmentation significantly improved detection accuracy, robustness, and computational efficiency. The stage-wise analysis confirmed that the model effectively distinguished between early and advanced gastric cancer, which holds critical importance for improving clinical outcomes. Moreover, the robustness tests demonstrated that the model maintained high performance under conditions of noise and image deformation, highlighting its potential applicability in real-world clinical environments. However, this work had certain limitations. The model was evaluated on a controlled dataset, and large-scale, multi-centre clinical validation was not yet performed to confirm its generalizability across different populations and imaging equipment. Although the AI-DFO optimizer enhanced the model's performance, the hybrid optimization approach introduced additional computational complexity, which may require further refinement for deployment in resource-constrained clinical settings. Additionally, the scope of this research was limited to gastric cancer, and the scalability of the proposed framework to other types of cancers or medical imaging tasks was not thoroughly investigated. In the future, this research could be extended by conducting large-scale clinical studies involving diverse datasets from multiple institutions to validate the model's robustness and generalizability. Further work could also focus on optimizing the computational complexity of the AI-DFO algorithm to facilitate real-time deployment in clinical workflows. In addition, adapting the proposed framework to other medical image analysis tasks, such as colorectal cancer detection, lung nodule classification, or skin lesion analysis, could establish its broader applicability. Exploring lightweight model variants for integration into mobile or portable diagnostic devices also remains a promising direction for future research.

Author contributions: Conceptualization, YBJK and BS; methodology, YBJK; software, BS; validation, YBJK and BS; formal analysis, YBJK; investigation, BS; resources, YBJK; data curation, BS; writing—original draft preparation, YBJK; writing—review and editing, YBJK; visualization, YBJK; supervision, BS; project administration, YBJK; funding acquisition, BS. All authors have read and agreed to the published version of the manuscript.

Funding: None.

Ethical approval: Not Applicable.

Informed consent statement: Not Applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Lee SA, Cho HC, Cho HC. A novel approach for increased convolutional neural network performance in gastric-cancer classification using endoscopic images. IEEE Access. 2021; 9: 51847–51854. doi: 10.1109/ACCESS.2021.3069747
2. Qiu W, Xie J, Shen Y, et al. Endoscopic image recognition method of gastric cancer based on deep learning model. Expert Systems. 2022; 39(3): e12758. doi: 10.1111/exsy.12758
3. Wang H, Ding Y, Chen Y, et al. A novel genomic classification system of gastric cancer via integrating multidimensional genomic characteristics. Gastric Cancer. 2021; 24(6): 1227–1241. doi: 10.1007/s10120-021-01201-9
4. Ooki A, Yamashita K, Kikuchi S, et al. Phosphatase of regenerating liver-3 as a prognostic biomarker in histologically node-negative gastric cancer. Oncology Reports. 2009; 21(6): 1467–1475. doi: 10.3892/or_00000376
5. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2024; 74(3): 229–263. doi: 10.3322/caac.21834
6. Sun M, Zhang G, Dang H, et al. Accurate gastric cancer segmentation in digital pathology images using deformable convolution and multi-scale embedding networks. IEEE Access. 2019; 7: 75530–75541. doi: 10.1109/ACCESS.2019.2918800
7. Tang Z, Xu Y, Jin L, et al. Deep learning of imaging phenotype and genotype for predicting overall survival time of glioblastoma patients. IEEE Transactions on Medical Imaging. 2020; 39(6): 2100–2109. doi: 10.1109/TMI.2020.2964310
8. Zhang L, Dong D, Zhong, L et al. Multi-focus network to decode imaging phenotype for overall survival prediction of gastric cancer patients. IEEE Journal of Biomedical and Health Informatics. 2021; 25(10): 3933–3942. doi: 10.1109/JBHI.2021.3087634
9. Chisnariandini N, Aji AP, Mariyam PD, et al. Measurement on real-time diagnostic of gastric tumor model using wireless endoscopy system. 2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). IEEE; 2018. pp. 105–108. doi: 10.1109/ICITACEE.2018.8576964
10. Meng L, Dong D, Chen X, et al. 2D and 3D CT radiomic features performance comparison in characterization of gastric cancer: A multi-center study. IEEE Journal of Biomedical and Health Informatics. 2020; 25(3): 755–763. doi: 10.1109/JBHI.2020.3002805
11. Yazici O, Sendur MAN, Ozdemir N, et al. Targeted therapies in gastric cancer and future perspectives. World Journal of Gastroenterology. 2016; 22(2): 471. doi: 10.3748/wjg.v22.i2.471
12. Alberts SR, Cervantes A, Van de Velde C J H. Gastric cancer: Epidemiology, pathology and treatment. Annals of Oncology. 2003; 14: ii31–ii36. doi: 10.1093/annonc/mdg726

13. Cristian CR, Oliver S, Nicolae L, et al. Advanced image processing in support of thz imaging for early detection of gastric cancer. 2018 International Conference and Exposition on Electrical and Power Engineering (EPE). IEEE; 2018. pp. 0634–0637. doi: 10.1109/ICEPE.2018.8559831
14. Li X, Chai Y, Zhang K, et al. Early gastric cancer detection based on the combination of convolutional neural network and attention mechanism. 2021 China Automation Congress (CAC). IEEE; 2021. pp. 5731–5735. doi: 10.1109/cac53003.2021.9728413
15. Sakai Y, Takemoto S, Hori K, et al. Automatic detection of early gastric cancer in endoscopic images using a transferring convolutional neural network. 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE; 2018. pp. 4138–4141. doi: 10.1109/EMBC.2018.8513274
16. Pernot S, Terme M, Radosevic-Robin N, et al. Infiltrating and peripheral immune cell analysis in advanced gastric cancer according to the Lauren classification and its prognostic significance. Gastric Cancer. 2020; 23(1): 73–81. doi: 10.1007/s10120-019-00983-3
17. Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: An international multicenter study. Annals of Oncology. 2020; 31(7): 912–920. doi: 10.1016/j.annonc.2020.04.003
18. Hatami S, Shamsaee R, Olyaei MH. Detection and classification of gastric precancerous diseases using deep learning. 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS). IEEE; 2020. pp. 1–5. doi: 10.1109/ICSPIS51611.2020.9349603
19. Okamoto H, Cap QH, Nomura T, et al. Stochastic gastric image augmentation for cancer detection from X-ray Images. 2019 IEEE international conference on big data (big data). IEEE; 2019. pp. 4858–4863. doi: 10.1109/BigData47090.2019.9006079
20. Swamy ST, Thotakura SB, Puli A, et al. Design and Analysis of MEMS based GMI sensor for detection of gastric cancer cells. 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE; 2018 pp. 131–135. doi: 10.1109/ICSSIT.2018.8748635
21. Yao K. The endoscopic diagnosis of early gastric cancer. Annals of gastroenterology: quarterly publication of the Hellenic Society of Gastroenterology. 2013; 26(1): 11. doi: 10.1016/s0016-5107(79)73384-0
22. Turppa E, Polaka I, Vasiljevs E, et al. Repeatability study on a classifier for gastric cancer detection from breath sensor data. 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE; 2019. pp. 450–453. doi: 10.1109/BIBE.2019.00087
23. Aslam MA, Xue C, Liu M, et al. Classification and prediction of gastric cancer from saliva diagnosis using Artificial neural network. Engineering Letters. 2021; 29(1).
24. Wang L, Shi B, Li P, et al. Short-term heart rate variability and blood biomarkers of gastric cancer prognosis. IEEE Access. 2020; 8: 15159–15165. doi: 10.1109/ACCESS.2020.2966378
25. Kinami S, Nakamura N, Miyashita T, et al. nPTD classification: An updated classification of gastric cancer location for function preserving gastrectomy based on physiological lymphatic flow. BMC Cancer. 2021; 21(1): 1231. doi: 10.1186/s12885-021-08936-9
26. Yoon HJ, Kim JH. Lesion-based convolutional neural network in diagnosis of early gastric cancer. Clinical Endoscopy. 2020; 53(2): 127–131. doi: 10.5946/ce.2020.046
27. Cao G, Song W, Zhao Z. Gastric cancer diagnosis with mask R-CNN. 2019 11th international conference on intelligent human-machine systems and cybernetics (IHMSC). IEEE. 2019; 1: 60–63. doi: 10.1109/IHMSC.2019.00022
28. Gao YL, Zhang YH. Gastric cancer prevention targeted on risk assessment: The Kyoto Classification of Gastritis. 2021 11th

International Conference on Information Technology in Medicine and Education (ITME). IEEE; 2021. pp. 330–335. doi: 10.1109/itme53901.2021.00074

29. Polaka I, Padilla M, Mezmale L, et al. Class decomposition for gastric cancer detection from breath. 2021 62nd international scientific conference on Information Technology and Management Science of Riga Technical University (ITMS). IEEE; 2021. pp. 1–5. doi: 10.1109/itms52826.2021.9615341
30. Sharanyaa S, Vijayalakshmi S, Therasa M, et al. DCNET: A novel implementation of gastric cancer detection system through deep learning convolution networks. 2022 international conference on advanced computing technologies and applications (ICACTA). IEEE; 2022. pp. 1–5. doi: 10.1109/ICACTA54488.2022.9752960
31. Yenginer H, Demir Ö. Stomach cancer detection using CIELAB color space. 2019 Medical Technologies Congress (TIPTEKNO). IEEE; 2019. pp. 1–4. doi: 10.1109/tiptekno47231.2019.8972044
32. Abe K, Shirakawa K, Minami M, et al. Automated extraction of the essential region in computer-aided diagnosis of helicobacter pylori infection using gastric X-ray images. 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI). IEEE; 2018. pp. 626–629. doi: 10.1109/IIAI-AAI.2018.00131
33. Liu B, Yao K, Huang M, et al. Gastric pathology image recognition based on deep residual networks. 2018 IEEE 42nd annual computer software and applications conference (COMPSAC). IEEE. 2018; 2: 408–412. doi: 10.1109/COMPSAC.2018.10267
34. Minghao Y, Kangkang L, Jiachen Y, et al. Point-of-care chemiluminescence immunoassay centrifugal microfluidics for gastrin-17 detection. 2020 IEEE 5th Optoelectronics Global Conference (OGC). IEEE; 2020. pp. 204–206. doi: 10.1109/OGC50007.2020.9260468
35. Pan D, Zhang R, Wang Y, et al. Detection of early gastric cancer based on single shot detector with feature enhancement. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE; 2020. pp. 1091–1095. doi: 10.1109/BIBM49941.2020.9313339
36. Alić B, Gurbeta L, Badnjević A, et al. Classification of metabolic syndrome patients using implemented expert system. CMBEBIH 2017: Proceedings of the International Conference on Medical and Biological Engineering 2017. Singapore: Springer Singapore; 2017. pp. 601–607. doi: 10.1007/978-981-10-4166-2_91
37. Badnjević A, Cifrek M, Koruga D. Integrated software suite for diagnosis of respiratory diseases. Eurocon 2013. IEEE; 2013. pp. 564–568. doi: 10.1109/eurocon.2013.6625037
38. Veljović E, Špirtović-Halilović S, Muratović S, et al. Artificial neural network and docking study in design and synthesis of xanthenes as antimicrobial agents. CMBEBIH 2017: Proceedings of the international conference on medical and biological engineering 2017. Singapore: Springer Singapore; 2017. pp. 617–626. doi: 10.1007/978-981-10-4166-2_93
39. Badnjevic A, Cifrek M, Koruga D. Classification of Chronic Obstructive Pulmonary Disease (COPD) using integrated software suite. XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013: MEDICON 2013, 25–28 September 2013, Seville, Spain. Cham: Springer International; 2014. pp. 911–914. doi: 10.1007/978-3-319-00846-2_225
40. Shibata T, Teramoto A, Yamada H, et al. Automated detection and segmentation of early gastric cancer from endoscopic images using mask R-CNN. Applied Sciences. 2020; 10(11): 3842. doi: 10.3390/app10113842
41. Abualigah L, Yousri D, Abd Elaziz M, et al. Aquila optimizer: A novel meta-heuristic optimization algorithm. Computers & Industrial Engineering. 2021; 157: 107250. doi: 10.1016/j.cie.2021.107250
42. Aghelpour P, Mohammadi B, Mehdizadeh S, et al. A novel hybrid dragonfly optimization algorithm for agricultural drought prediction. Stochastic Environmental Research and Risk Assessment. 2021; 35(12): 2459–2477. doi: 10.1007/s00477-021-02011-2