AVE Trends in Intelligent Health Letters

Identification of Gastric Cancer Using Deep Learning Techniques with the Help of Breath Samples in the IoT Environment

G. Rajesh[1,*], A. B. Gurulakshmi[2], B. Manjunatha[3], Thirumalraj Karthikeyan[4]

[1]Department of Electronics and Communication Engineering, New Horizon College of Engineering, Bengaluru, Karnataka, India.
[2,3]Department of Mechanical Engineering, New Horizon College of Engineering, Bengaluru, Karnataka, India.
[4]Department of Artificial Intelligence, Trichy Research Labs, Quest Technologies, Tiruchirappalli, Tamil Nadu, India.
rajesh.gundlapalli@gmail.com[1], dr.gurulakshmiab@newhorizonindia.edu[2], manjunatha.princi@gmail.com[3], thirumalraj.k@gmail.com[4]

*Corresponding author

Abstract: Deep learning is a new method that is frequently utilised in the medical industry for illness diagnosis. For the analysis of (EGC), a new research avenue has been devised. Because of their efficacy, computer-aided diagnostic (CAD) systems minimise mortality rates. We proposed a new approach for feature extraction, utilising a Deep Q Network (DQN), for this study. To identify stomach cancer from breath samples, an upgraded Swin Transformer network was coupled with the suggested feature extraction approach. An Improved Attention Swin with feature pyramid network (IAS-FPN) architecture aims to identify different phases of GC proficiently by adding context information to the fundamental Swin Transformer. Furthermore, to prevent small-cell information loss, an enhanced weighted bidirectional network is provided by incorporating skip connections with enhanced Dwarf Mongoose Optimisation (IDMO). To be more specific, we found fifty to differentiate EGC, Advanced Gastric Cancer (AGC), and healthy people. By learning the characteristics and preserving the construction of the input tasters, this CAD scheme lowers the distance between the input and output. The traits were derived from unlabeled breath-sample data. The suggested network design achieves good results for advanced gastric cancer classification using breath data, with an overall accuracy of 94.07%. Furthermore, the created model achieves high accuracy and F1 score, making it ideal for scientific use.

Keywords: Deep Learning; Breath Analysis; Computer-Aided Diagnostic (CAD); Deep Q Network (DQN); Advanced Gastric Cancer (AGC); F-Score Value; Fundamental Swin Transformer.

Cite as: G. Rajesh, A. B. Gurulakshmi, B. Manjunatha, and T. Karthikeyan, "Identification of Gastric Cancer Using Deep Learning Techniques with the Help of Breath Samples in the IoT Environment," *AVE Trends in Intelligent Health Letters*, vol. 2, no. 2, pp. 94–107, 2025.

Journal Homepage: https://avepubs.com/user/journals/details/ATIHL

Received on: 10/09/2024, **Revised on:** 28/12/2024, **Accepted on:** 23/02/2025, **Published on:** 05/06/2025

DOI: https://doi.org/10.64091/ATIHL.2025.000169

1. Introduction

Currently, the Industrial Internet of Things (IIoT) is one of the fastest-emerging networks capable of gathering and transmitting massive volumes of data utilising sensors in the healthcare sector [19]. In the therapeutic sector, IoT is regarded as an expert

application and is frequently referred to as medical IoT (mIoT) or (IoMT) [7]. The IoMT is a networked infrastructure of medical applications, equipment, health systems, and services. It is used to analyse the physical properties of sensor nodes that collect data from the patient's body, utilising sophisticated, portable devices. IoMT enables wireless and remote devices to safely interact over the Internet, while integrating AI methodologies provides quick, flexible medical data analysis and diagnosis [2]. IoT devices manage various unknown aspects when transmitting data to the cloud, including network topology, energy transfer, and processing capabilities [20]. Caregivers, health care professionals, and patients have successfully adopted telehealth services for remote patient monitoring, illness identification, and effective treatment [3]. All of these gadgets and services contribute to the evolution of healthcare toward the Healthcare Industry 4.0 [23]. (GC) is the deadliest, fourth most predominant, and foremost reason for cancer-related deaths globally. According to Leja and Linē [12], it ranked among the five highest GC rates in 2018. GC is linked to a variety of factors, including lifestyle and genetics. GC is divided into two phases.: (1) AGC, (2) EGC.

Endoscopy is the most often utilised procedure in clinics for detecting GC, which is a particularly violent kind of cancer that is problematic to diagnose in its initial stages due to the ambiguous signs of EGC [10]. Only a million fatalities were observed worldwide due to the dismal forecast of GC. Because GC indications in the early phases are not typical, most GC-affected roles waste valuable therapy time [8]. The mortality rate of stomach cancer can be lowered if it is detected early. Not only is an early prognosis sufficient, but correct staging can minimise mortality. The disease's early prognosis determines the death rate. AGC patients have a 2-year survival rate of 24%, whereas EGC patients have a 5-year survival rate [9]; [10]. Ancient Greek physicians used odour in the breath to identify many ailments. Previous research has established that breath gas is a complex mixture including over 3000 VOC indicators [26]. Because these VOCs change characteristics throughout metabolism, they can be employed for cancer detection. Breath analysis has been used to diagnose lung cancer and GC. Previous research has concentrated on the early detection of certain malignancies. Breath analysis is gaining popularity in disease diagnosis because it can yield reliable, repeatable findings with no harm to the patient during analysis [22]. To distinguish between healthy and ill patients, VOC levels in exhaled air are measured. Salivary volatile organic compounds (VOCs) are elevated when a person is unwell. These volatile organic compounds are the root cause of the breath analysis for sickness. Medical diagnostics, robotics, audio/voice identification, and industrial applications are just some of the domains where deep learning has been applied.

The Transformer network has recently shown signs of improved visual presentation. (ViT) network shows that a model based solely on attention can achieve better results than one based on CNNs [24]. These volatile organic compounds are the root cause of the breath analysis for sickness. Deep learning is a type of artificial intelligence. There are several recognition and industrial uses for deep learning techniques. Transformer networks have recently shown exceptional performance in image identification. It's worth noting that, compared to a perfect CNN-based model, the (ViT) network displays superior performance [25]. However, such effort is now absent from transformer-based copies. Inspired by these efforts, we proposed a novel and intuitive approach to incorporate an attention operation into the core of the Swin Transformer to further enhance its foreground-enhancing mechanism [13] further. Second, the vertically oriented target detection architecture relies heavily on the neck. It takes the relevant features gleaned from the spine and reprocesses them so that the head may take the next logical step in learning the task at hand. Nowadays, people typically use their necks for BiFPN [17]. Strong semantic feature representations are created by downsampling high-resolution data and blending them with low-resolution features [18]. However, the feature map loses significant detail due to the sequential downsampling step [24]. This makes it more difficult to find very small objects. An improved FPN with actions to skip the connection and strike out duplicate production nodes was developed to fix this problem. In a nutshell, this essay adds four things to the conversation:

- To include a block to provide more contextual info and learn more recognised characteristics, as well as to examine its performance in cancer discovery tasks. Furthermore, the IASB unit we established in the mainstay may hypothetically be incorporated into already existing.
- We also present a simple network for fusing feature maps across multiple encoder stages. The revised dwarf-mongoose optimisation technique is used to minimise the FPN's weight. This optimisation approach modifies the fundamental algorithm (DMO) in three effective ways [21]. To begin, the IDMO alpha selection differs from the DMO in that calculating fitness incurs only a computational cost, which is added to the number of members.
- The IDMO's exploration and exploitability are increased by selecting the healthiest dwarf mongoose as the alpha and introducing a new-fangled operator to manage the alpha's movement. Second, the scout group's motions are randomised to inject variation into the search process and to investigate previously unexplored locations [27]. Finally, the exchange criterion for babysitters is changed so that when the criterion is met [29].

2. Related Works

(WSIs) Human stomach cancer tissue slices are presented by Liu et al. [15] to demonstrate their multi-task technique for identifying PNI areas relevant to diagnosis. To achieve the goal of recognising PNI during gastric cancer segmentation, the proposed framework comprises a stomach cancer model, a neural network for refining findings, and a PNI executive module.

Adequate comparative trials were performed on our GC-PNI dataset of patients with gastric cancer. Our proposed model has been tested, and the results show a sensitivity of 0.933 for diagnosing PNI. Zhu et al. [30] set a standard for the pleasant and painless prediction of stomach cancer. Using a specialised tongue image-capturing device, 703 images of the tongue from persons with and without stomach cancer were taken. To minimise noise introduced by non-tongue parts of the image when extracting features from the tongue, the Deeplabv3+ network was used. Using statistical methods and deep learning, researchers examined nine tongue characteristics for their links to stomach cancer. A system for predicting the development of stomach cancer has been created. The trial results showed that the proposed framework can detect anomalies with a high degree of accuracy, established by combining various arithmetic means. Lee et al. [11] attempted to use only robots to determine whether MSI was present in GC tissue slides. Compared to the TCGA's 0.893 and 0.902 AUCs, the AUC for classifiers trained on the Asian FFPE external validation cohort was 0.874. However, the morphologic properties of MSI-H tissues vary, making the DL look mismatched with malignancies from diverse organs. Histo-morphological analysis of MSI-H GC tissue reveals two distinct subtypes: moderately to poorly differentiated intestine-type tumours and diffuse-type mucinous tumours.

Even though the DL-based classifier has exceptional performance, it cannot be fully explained by the detected morphological features. This study's findings support the idea that DL can be trained to identify the most important elements for establishing MSI status in GC tissue slides. This study demonstrates how a DL-based MSI can effectively address challenging issues. Ma et al. [16] provide a deep learning (DL)-based, fully automated EGC diagnostic approach. To automate the analysis of EGC images, this research uses a novel annotated endoscopic image dataset compiled from a single source. A guided-attention deep network based on ResNet-50 was implemented to estimate the pixel-level segmentation of EGC issue areas. This was conducted with the express purpose of producing reliable EGC score forecasts. In tiny and micro fuzzy scenarios of cancer cell detection in the stomach cancer slice dataset, Guo et al. [6] propose a small, micro fuzzy pathology identification approach for the entire pathology. The law of feature distribution helps the network learn features on micro and nanotargets. Changes in channels and locations might cause adjustments in network emphasis and node prominence. When the attention mechanism is applied to the deep blur scenario, the mAP is 0.611, suggesting an improvement in recognition capacity that may be useful for clinical decision making. By providing a second opinion, the computer-assisted diagnosis method proposed by Chae and Cho [5] can improve diagnostic efficiency and eradicate incorrect diagnoses. Using Vision Transformer, a classification model with the best transfer learning performance, we aimed to classify normal lesions and initial stomach cancer.

In addition, we included an MFAA tactic to enhance the model's classification performance with limited medical data. By applying MFAA to medical data, we significantly improved the model's classification performance during training. The F1-score of 0.87 anomalies was attained in experimental testing of the model. Using GES-1 and SNU-16 stomach cancer cell lines, compiled a Raman spectrum library [14]. Complete spectral data, fingerprint-region spectral data, high-wavenumber-region spectral data, Raman-spectral data, and Raman-background spectral data were used to construct five spectral datasets. To improve the detection of stomach cancer cells across multiple datasets, we developed an SELM. Called SL-Raman. For the gastric cancer cells we evaluated, SL-Raman achieved a perfect score of 1.000 for the differentiation of one cell type, and a score of 1.000 or higher for six other cell types. In addition, the separation rate was 100% for stomach cancer cells at two distinct stages of differentiation. These results suggest a new method, Raman spectroscopy combined with SL-Raman, for the early and reliable diagnosis of stomach cancer. To help surgeons avoid possible blind spots during EGD, Zhang et al. [28] developed a deep learning-based technique that tracks the movement of the camera. Our approach can classify 12 individual parts of the oesophagus, stomach, and small intestine, as well as the background, with an accuracy of 87.22%. The majority of our misclassifications were found to be between neighbouring regions, and we discovered that attention block and MoCo pre-training. Our model achieves 89.6% accuracy and 88.1% recall on a dataset of annotated EGD surgical films. This method may worsen GI symptoms.

2.1. Problem Statement

Cancer is a broader term that refers to a range of illnesses that affect numerous organs of the human body, such as the lungs, which are increasingly common in sustainable cities. There are now roughly 17 million cancer-related cases globally, with an estimated increase to nearly 23 million new cases each year by 2035. Even when treated with modern technologies, the patient survival rate is greatly reduced. Even after a diagnosis of lung cancer, a person's maximum life expectancy is 5 years. Aside from that, misdiagnosis is a source of concern. Specialists in the medical field have found that there are cases in which a benign category is misidentified as a malignant kind and vice versa. The patient's life is in jeopardy and doubtful as a result. Therefore, early detection increases the patient's survival probability. Recent developments in computational methods and smart computing have enabled the development of an automated IoHT-based intelligent model for diagnosing stomach cancer. Early detection of cancer-related health problems may be possible with this approach. The IoHT component can monitor patients and record their data in real time. Information about the patient can be gleaned from these records and used in the future. This can be useful for doctors, as it helps them spot cancer earlier in patients. The data samples need to be cleaned of these less critical factors. Heuristic algorithms are useful in situations like this because they help identify and remove irrelevant material from the dataset.

The accuracy and efficiency of a classification system benefit greatly from using a refined, well-tuned dataset. Overfitting is a problem for many classifiers, and researchers try to avoid overfitting by not modifying them too much.

3. Proposed System

Figure 1 depicts the suggested model's workflow, which will describe the entire procedure.

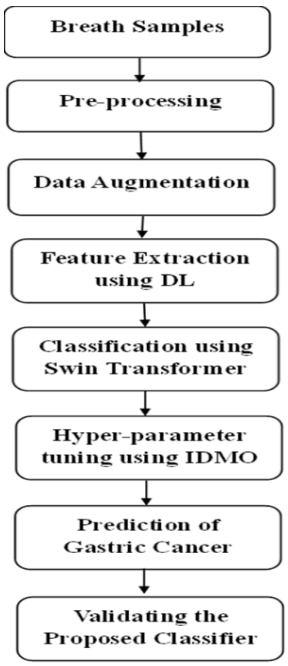

Figure 1: Work flow of the model

3.1. Dataset Description

This study followed the Reporting of Studies (REMARK) standards. All breath samples were collected at Shanghai Tongren Hospital in China [4]. All participants had already been instructed in performing clinical research. The ethics team of Tong University accepted this work. There were 200 participants, including 55 EGC patients, 56 healthy individuals, and 89 AGC-affected relatives. While collecting the breath samples, three criteria were followed: (1) individuals had already been clinically diagnosed with GC primarily diabetes, were excluded. The clinical features of the participants are shown in Table 2. The GC Manual was employed. Because age and gender did not affect EGC patients, AGC patients, or healthy individuals, we omitted this evidence to reduce potential bias. For roughly an hour, all participants were advised to wipe their mouths and refrain from eating and drinking. We utilised 75% of the breath samples for the authentication set and 25% for the test set. For modelling in this work, we selected a spectral range of 400-1500 nm.

3.2. Data Augmentation

A neural network's robustness is based on how well it captures the structure of the input data. The reliability of a neural network improves with the size of its training data set. Large amounts of data are needed to prevent overfitting and underfitting. We used data augmentation to create extra data for this research. The entire spectrum is represented in a single dimension as input to the proposed design. This is a monstrous proportion. Therefore, we used 1200 separate values for each of the 200 spectra after clipping. Breath samples can be shifted to the right or left to enlarge the dataset. In this probe, we shifted the breath samples to the right by 2 cm1. After data enhancement, the number of samples increased to 368, and the total data volume increased to 453,600, with 240,000 spectral standards.

3.3. Data Pre-Processing

The breath samples were collected at the hospital and were affected by noise. To make these breath samples meaningful, data preparation techniques were applied. This noise element might lead to an incorrect categorisation. We remove and decrease irrelevant and random changes in breath tasters. Spikes were produced as the breath sample collided with the detector. These

have limited bandwidth and are caused by differences in sensor positions. Compared to the Raman spectra, the spike's bandwidth is quite modest. Using the premise outlined above, we eliminate spikes from the Raman samples. The noise is high in frequency. Various approaches have been developed to reduce data noise. We employed a median filter to eliminate noise from breath samples in this investigation. The medial filter successfully removes the noise. Baseline adjustment is an essential step in addressing problems caused by negative spectral values. The baseline adjustment does not affect the intensity of the Raman band signal. The Labspec5 program was also used to smooth the spectrum.

3.4. Feature Extraction Using Deep Q Network (DQN)

The input data is fed into DQN to determine the properties of the breath samples. Experience replay is used in the DQN. The agent's involvement u_v = (f_v, r_v, e_v, f_(v+1)) at each time point in the dataset. $D = (u_1, \ldots, u_v)$. Here, D is the dataset technique, the Q learning information is used on statistical samples of experience (f_j, w_j, e_j, f_{j+1}) that dataset D. Lease the episode as $\{1, \ldots, \alpha\}$, time as $\{1, \ldots, \chi\}$, and as the interval essential to appraise the system factor. As a result, the function used to update at repetition x is represented as follows:

$$\beta_2 = \eta\left[\left(e + {}^{\delta max}_{w'}\widehat{N}(f', w'; \theta_x^-) - N(f, w; \theta_x)\right)^2\right] \quad (1)$$

Where e is the reward, is the reduction factor, _x denotes the network limit of the Q network at iteration x, and _x- denotes the network limit used to determine the target. $\kappa_j = e + {}^{\delta max}_{w'}\widehat{N}(f', w'; \theta_x^-)$at the iteration x, correspondingly. Therefore, it is renowned that the term $\kappa j = e + {}^{\delta max}_{w'}\widehat{N}(f', w'; \theta_x) - N(f, w; \theta_x$of The instability in convergence stems from the fact that the aforementioned loss function is a TD error that varies with the value of network parameter _x at iteration x. For better convergence stability, DQN takes into account the neural fitted Q method, where the parameter in j = e +(_w' max)N (f ',w'; _x) is a carefully fitted value. $\kappa j = e + {}^{\delta max}_{w'}\widehat{N}(f', w'; \theta_x^-)$at. The network limit $\hat{N}$ is skilled at each repetition using the projected optimisation procedure.

3.4.1. Training of DQN

The fitness function computes and expresses the best charge of the system parameter as:

$$F = \frac{1}{\vartheta}\sum_{v=1}^{\vartheta}\left[O_v - \widehat{N}_v\right]^2 \quad (2)$$

Where F is for fitness function, N stands for the total sum of samples, and O stands for the desired output. The characteristics of gastrointestinal samples are transformed into a two-dimensional matrix and supplied to the suggested model.

3.5. Overview of the Projected Framework

In this section, we shall introduce IAS-FPN. Our proposed IAS-FPN firstly employs the tried-and-true backbone-neck-head layout. Based on the foundational Swin-tiny, we developed and released IAS as the core of our IAS-FPN. In this context, the central nervous system is called an encoder. We started with the basic FPN in the neck and built an improved version with the skip-strike output node actions added.

3.5.1. IAS in Backbone of IAS-FPN

Partitions of size R44 are used to divide the RGB picture (RHW3). These patches serve as tokens that are fed into the encoder (Swin-tiny) for further processing. To get the features into the correct dimension (labelled C), a linear embedding operation precedes the first transformer layer. The encoder uses a four-stage, stratified process to extract attributes. The encoder is constructed in steps, with each stage requiring the cooperation of a context layer. To better so, the context layer is composed of stacked NF FEABs, while the transformer layer is constructed of blocks. Each regular block features a shift window, sometimes known as a window. To restrict what can be displayed, the mask setting can be adjusted. When compared to the transformer block in ViT, the window focus is equal after relative position encoding is implemented. Tokens within a window are only considered, while shifting window attention takes into account several non-overlapping windows. Due to the temporal linearity of both attention processes, this method excels at representing high-definition visual features. To create a hierarchical feature map with the fewest possible tokens, patch merging is used at the start of each stage, beginning with the second. Each of the four iterations produces a feature map with a different resolution. is $\left\{\frac{H}{4} \times \frac{W}{4} \times C, \frac{H}{8} \times \frac{W}{8} \times 2C, \frac{H}{16} \times \frac{W}{16} \times 4C, \frac{H}{32} \times \frac{W}{32} \times 8C\right\}$. V, in which the Q, K, and V matrices stand for questions, keys, and values. We do the same thing by presenting a relative bias $RPB \in R^{W \times W}$ in the control of self-attention:

$$\text{Attention}(Q, K, V) = \text{SoftMax}\left(\frac{QK^T}{\sqrt{D^k}} + RPB\right)V, \quad (3)$$

where D^k signifies the dimension of the key. The output chin map is sent to the setting layer as input after the last Swin block of the present layer. Target detection relies heavily on having access to detailed context data. Thus, the shadowing context layer is crucial to the encoder and is the paper's main contribution. There are N_F IASB in each tier of context. To direct the area rather than the entire feature map, the IASB boosts foreground features while reducing background features. We partition FEAB's input X' into three parts, labelled E_Q, E_K, and E_V for "context query," "context key," and "context value," respectively, to bring the computational cost down to a linear scale. To linearly project the input features X', we multiply them by the learnable matrices W_Q, W_K, and X_V', which are themselves initialised at random. Normalised by a softmax function, the query and key enhancements E_Q and E_K are then utilised to revise X_V'. In addition, we introduce a smooth fine-tuning scalar constant Q that can be learned after the linear layer. This is because the scalar dampens the converter's magnitude during initialisation. The second advantage is that the learnable scalar we provided decays properly as its weight falls during training, keeping the output within an acceptable range. Here is the equation we developed to focus attention on the foreground:

$$\text{ISAttention}\left(E_Q, E_K, X'_V\right) = \text{SoftMax}\left(F.\, E_Q E_K^T\right)X'_V + \Theta \quad (4)$$

Where the forefront mask at the feature pixel site (x, y) is:

$$F(x,y) = \begin{cases} 1, & \text{if } M(x,y) = 1 \\ 0, & \text{if } M(x,y) = 1 \end{cases} \quad (5)$$

In this case, $M(x,y)2\{0,1\}^{H\times W\times C}$ is a Binarisation purpose whose value is predicted from the (x, y) location of the original input chin X using mask prediction.

3.5.2. Improved FPN as the Neck of IAS

Because breath samples are tiny and dense, the feature illustration power of a single layer of backbone is limited.

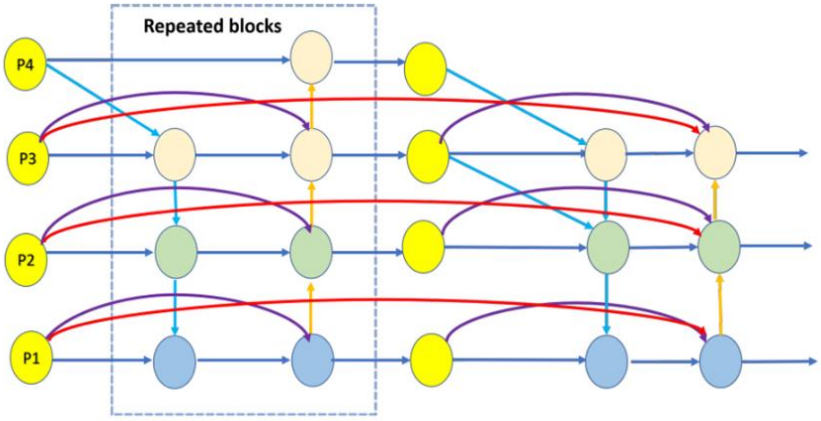

Figure 2: Improved FPN

As a result, we require a sophisticated feature fusion network to efficiently and quickly fuse features at different scales across multiple encoder layers. In general, the feature map output $P^O = f(P_i^I)$, where P_i^I attitudes for the feature map of the ith level, and f stands for the feature synthesis technique. One of the most cutting-edge FPNs now is the weighted network. Cross-scale weighted feature fusion is one of the two areas where BiFPN excels. In some cases, the former enables the aggregation to include a richer set of contextual details. The latter is meant to teach the network how each input quality affects the output. In this paper, we will talk about the weighted feature fusion approach:

$$O = \sum_i \frac{\omega_i}{\epsilon + \sum_j \omega_j} . I_i \quad (6)$$

Where _i is a learnable weight that is larger than or identical to 0 by the ReLu purpose that comes next; To maintain numerical stability, = 0.0001 is utilised. I_i denotes the ith level's input feature map. When our task requires a four-layer input feature map, we should use the IFPN architecture. Unfortunately, this layout makes the output feature maps of certain layers invisible

to all layers after aggregation. Figure 2 depicts the enhanced FPN structure. The backbone provides four levels of characteristics at various scales, which are considered inputs; the first n repeating blocks output, following the IFPN principle. Each aggregated feature map occurs as an input to the RPN in the last block of the original FPN. In this situation, the production at each level directly affects the ultimate result. However, in the final layer, our enhanced approach only accepts a 3-level output. Furthermore, we devised a skip connection method to prevent loss of tiny object information. For instance, the ith-level input feature map of the first layer is denoted P_iF, the feature iterations are denoted P_iT, and the input and output maps are denoted P_iI and P_iO, respectively. The last layer's second-level map is calculated as:

$$P_2^O = \mathrm{Conv}\left(\frac{\omega_1'.P_2' + \omega_2'.P_2^T + \omega_3'.\mathrm{Resize}(P_1^O) + \omega_4'.P_2^F}{\omega_1' + \omega_2' + \omega_3' + \omega_4' + \epsilon}\right) \quad (7)$$

The IDMO method was used in the study to maximise the learnable weight of the FPN model, as explained below.

3.5.3. The Improved Dwarf Mongoose

IDMO's main goal is to improve DMO, as outlined by Agushaka et al. [1]. DMO's solutions for problems F9, F15, and F17 show that it has this constraint, as they are not optimal. Three relatively minor but significant changes to the underlying algorithm (DMO) are at the heart of this optimisation approach. To begin, fitness doesn't improve the quality of the alpha or any other members of the IDMO group, whereas it does in the DMO. The IDMO's ability to explore and exploit new areas is improved when the healthiest dwarf mongoose is chosen to control the alpha's movements. Second, the scout troop's actions. Finally, the exchange criterion for babysitters is tweaked so that the babysitters who are swapped out communicate with the dwarf mongoose and share knowledge of where they can get food and a safe place to sleep. The proposed IDMO uses a three-step optimisation process.

In contrast to DMO, which combines scouting and foraging into a single action, this model shows that scouting is carried out independently of foraging. The nd matrix in Eq. 8 represents independent dwarf mongooses as agents. In the exploration phase, the adapted alpha (Eq. 10) uses the steps outlined in Eq. 11 to guide the team through uncharted territory. Equation 12 improves exploration and exploitation in the IDMO by introducing a new movement operator. Randomisation is employed to change the paths taken by scout groups, as shown in Eq. 13, to inject variety into the search and uncover new areas. In Eq. 14, we see that once the babysitting trade criteria are met, exploitation happens. The acquired answer is optimised in this stage.

3.6. Population Initialisation

As demonstrated in Eq. (8), the IDMO populace is stochastically seeded with a matrix of potential dwarf mongooses (X). The populace vector lies between the upper (U) and lower (L) bounds of the optimisation difficulty.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d-1} & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d-1} & x_{2,d} \\ \vdots & \vdots & x_{i,j} & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d-1} & x_{n,d} \end{bmatrix} \quad (8)$$

Where n is the sum of dwarfs in a knoll, x_(i,j) signifies the site of the ith population's jth dimension, and each x_(i,j) is specified in Eq. (9):

$$x_{i,j} = \mathrm{rand} \times (U - L) + L \quad (9)$$

3.6.1. Alpha Group

This group's population size is of dwarf mongooses. As shown in Eq. 10, the alpha female (a) leads this group and is chosen as the alpha female. In IDMO, alpha selection differs from DMO in that analysing the likelihood value of each fitness does not add to the other memberships:

$$a = \min(fit_1, fit_2, \ldots., fit_n) \quad (10)$$

The alpha female uses the peep's vocalisation to keep the group together. The IDMO moves through the issue space as specified in Eq. 11. It is first characterised as the fittest dwarf mongoose and leads the other members of the family toward a prospective food source. This situation differs significantly from the DMO, in which only the alpha's vocalisation influences the location of the utilised to establish the location of the additional mongoose in IDMO, and a novel worker, described in Eq. 12, regulates the alpha drive, boosting the IDMO's exploitability:

$$X_{i+1} = a + phi * rand * (X_i - X_k) \quad (11)$$

$$\omega = e^{-4*(C_{iter}/Max_{iter})^2} \quad (12)$$

where $phi = \left(\frac{peep}{2}\right) * rand * \omega, X_i$ is the preceding dwarf site, and rand is a consistently distributed random sum $[-1,1]$. X_k It is an arbitrarily designated dwarf mongoose.

3.6.2. Scout Group

Scouts are responsible for finding a suitable sleeping area, semi-nomadic, never returning to their mound. After foraging, the IDMO simulates a scout party scouting for the next sleeping mound. Because dwarves are known to congregate around copious food sources, the fitness of the scouts is regarded as a prospective sleeping mound, and the fittest mound. Eq. 13 is used to model the scouts:

$$X_{i+1} = a + phi * rand * (X_k - X_h)/2 \quad (13)$$

Where rand is a random sum among [0, 1], and X_k, X_h are randomly designated dwarf mongooses.

3.7. The Babysitters

Eq. 14 gives the exchange criterion for the babysitter. Once the criterion is reached, the traded babysitters engage with the dwarf to learn the sleeping mound, perhaps resulting in better-fitted mongooses rather than starting them again as in DMO, and then return to zero. This enhancement is depicted in Eq. 15, where the dwarves that will replace them are chosen at random, and their knowledge is provided to the sitters as stated. L is reset to zero if it reaches zero, and then increased by the present repetition and CF:

$$L = \begin{cases} Roundup\left(0.6 * n * dim * \left(\frac{1}{C_{iter}}\right)\right) \\ L * C_{iter} * CF \quad when\ L < 0 \end{cases} \quad (14)$$

$$X_{i+1} = \left(X_j + rand * \left(1 - \frac{X_h + X_h}{2}\right) * br\right) \quad (15)$$

Where, $CF = \left(1 - \frac{C_{iter}}{Max_{iter}}\right)^{\left(2\frac{C_{iter}}{Max_{iter}}\right)}$ controls the shared volitive drive of the dwarf mongooses, X_j, X_k, X_h are arbitrarily selected gnome mongooses to substitute the sitters, and br is the birthrate.

The IDMO's computational difficulty is greatly reduced because it reduces the amount of alpha selection. The optimisation procedure begins with forage, led by the alpha female. A chosen number, known as the babysitters, are left behind to care for the nest. The quest for abundant food sources resembles the IDMO's exploring phase. Because dwarf mongooses are others, the babysitters are replaced at lunchtime so they can eat. When this occurs, the swapped babysitters return to their previously known food sources to eat quickly. This scenario represents the IDMO's exploitation phase. At the end of the day, scouting for a sleeping mound expands space. The IDMO algorithm, like the DMO, has only one unique limit to fine-tune.

4. Results and Discussion

For the categorisation of gastric cancer, the suggested IAS-FPN was built and evaluated using MATLAB 2017b (Natick, MA, USA). The network was trained on a 2.3 GHz Core i5-2350M CPU using simple approaches, starting at 0.0003. After 1000 epochs, the neural net converged.

4.1. Performance Evaluation

The goal of this research is to develop a classifier that distinguishes between EGC, AGC, and well people. We created IA and associated it with the five existing approaches listed in Section 2. All models employ pathological samples or other imaging approaches, whereas the current models do not collect breath samples. As a consequence, all models are evaluated with our data, and the results are averaged in Table 1 to identify EGC, AGC, and healthy people. GC classification is a three-class problem in this study. There are three possible outcomes for classifiers 0, 1, and 2: healthy. Classification outcomes for each generated model were calculated:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{16}$$

$$\text{Recall} = \frac{\text{TP}}{\text{(TP+FN)}} \tag{17}$$

$$\text{F} - \text{Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision+Recall}} \tag{18}$$

$$\text{Accuracy} = \frac{\text{TN+TP}}{\text{TP+TN+FP+FN}} \tag{19}$$

A successful classifier should achieve high accuracy while maintaining low precision and recall. If any of the criteria as mentioned above are not met, the proposed perfect is inaccurate and cannot be utilised in scientific requests.

Table 1: Performance analysis with various classifiers

Method	Accuracy	Specificity	Recall	Sensitivity	Precision	F1 score
U-Net	91.60	89.07	81.56	96.67	96.67	88.47
YOLO-V5	89.88	88.15	79.75	93.33	93.33	86.01
Vision transformer	90.74	87.41	79.46	97.41	97.41	87.52
MFAA	92.59	90.56	83.56	96.67	96.67	89.69
SELM	91.48	88.33	80.73	97.78	97.78	88.44
IAS-FPN	94.07	91.48	85.21	98.15	98.15	91.22

When the models are tested using F-score analysis, the proposed model achieved 91.22%, U-Net 88%, YOLO-V5 86%, MFAA 89%, SELM 88%, and VT 87%. When comparing all models, YOLO-V5 and VT performed poorly across accuracy, specificity, and recall. For instance, the existing models achieved 89% to 92% accuracy, 93% to 97% sensitivity, 87% to 90% specificity, 79% to 83% recall, and 93% to 97% precision. But the proposed IAS-FPN model achieved 94.07% accuracy, 98.15% sensitivity, 91.48% specificity, 85.21% recall, and 98.15% precision.

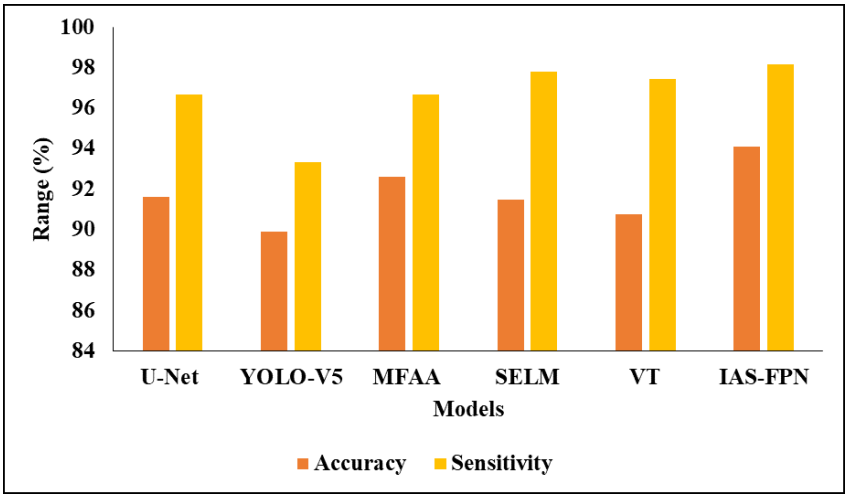

Figure 3: Graphical analysis of various classifiers

Figures 3, 4 and 5 present a graphical analysis of the proposed classifier compared with existing DL models, using various metrics.

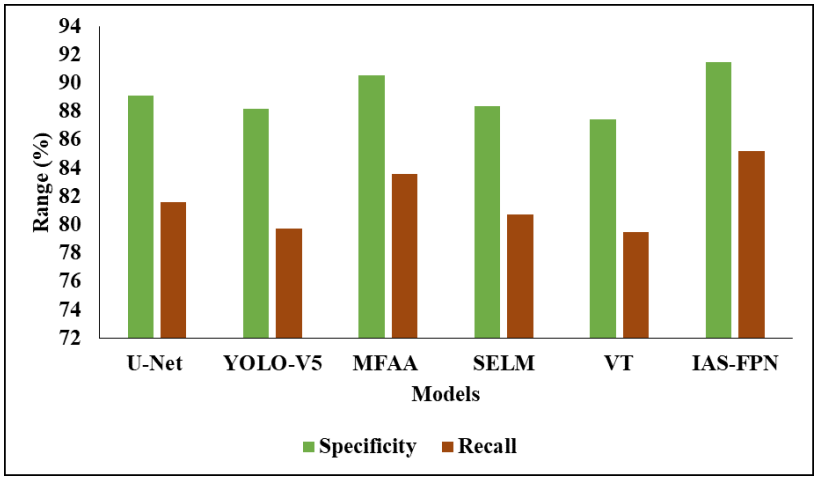

Figure 4: Analysis of IAS-FPN

The existing optimisation techniques, such as Chimp Optimisation Algorithm (COA), Elephant Herd Optimisation (EHO), Cat Swarm Optimisation (CSO), Harris Hawks Optimisation (HHO) and Red Fox Optimisation (RFO), are considered for validation.

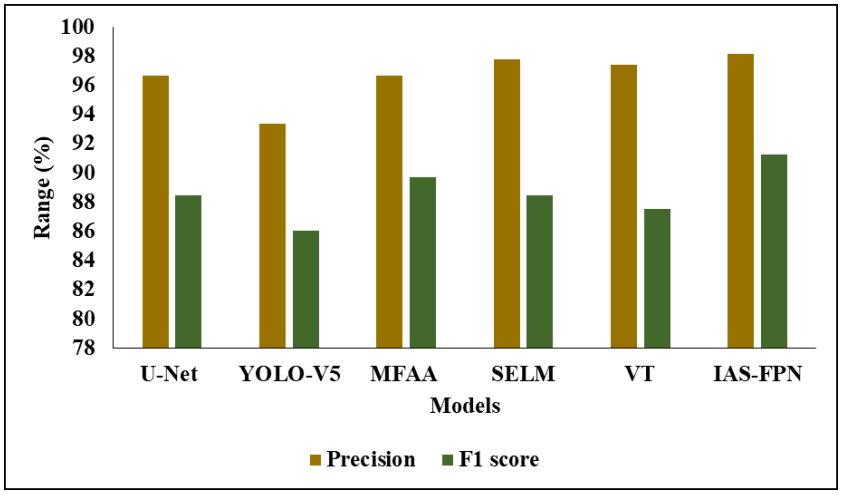

Figure 5: Comparative analysis for different DL

The above techniques are implemented with the improved FPN, and results are averaged to improve the selection of learning weights, as shown in Table 2.

Table 2: Performance analysis with various optimisations

Method	F1 score	Accuracy	Specificity	Recall	Precision	Sensitivity
EHO	88.89	91.98	89.81	82.54	96.30	96.30
CSO	90.00	92.84	90.93	84.19	96.67	96.67
COA	85.62	89.14	85.19	76.61	97.04	97.04
HHO	83.85	87.78	84.07	74.93	95.19	95.19
RFO	75.95	82.35	90.74	80.43	85.50	85.56
IDMO	91.22	94.07	91.48	85.21	98.15	98.15

When the investigation of models is approved for accuracy, the projected model attained 94.07%, COA has 89.14%, EHO has 91.98%, CSO has 92.84%, HHO has 87.78% and RFO has 82.35%. Compared with all models, IDMO achieved 98.15% sensitivity and precision, 91% specificity and F1-score, and 85.21% recall. But the existing models such as COA, EHO, CSO,

HHO and RFO achieved nearly 82% to 92% of accuracy, 85% to 97% of sensitivity, 84% to 90% of specificity, 74% to 84% of recall, 85% to 97% of precision and 75% to 90% of F1-score.

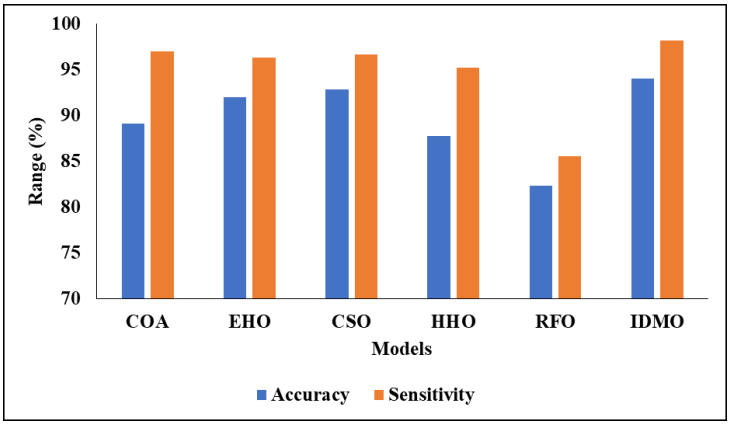

Figure 6: Analysis of different optimisations

From these experiments, it is clear that the proposed IDMO achieved better performance in selecting the optimal weight for the improved FPN model.

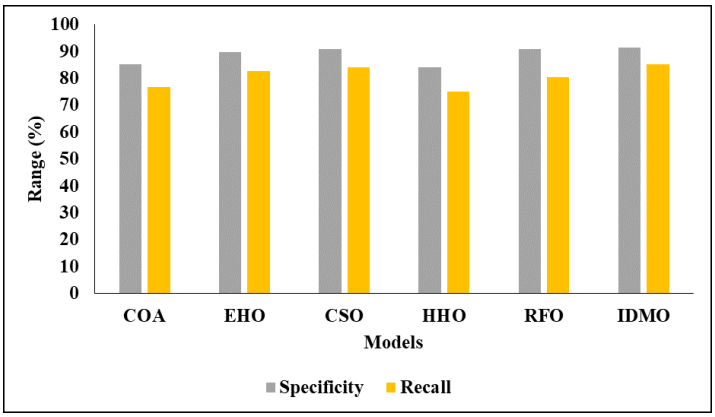

Figure 7: Validation comparison of the proposed IDMO

Figures 6, 7 and 8 present a graphical comparison of the proposed model with existing optimisation models across numerous metrics.

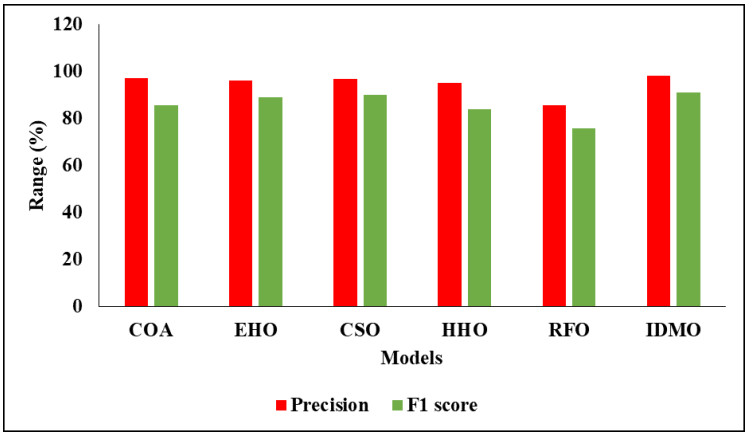

Figure 8: Graphical representation of various optimisation models

In the analysis of training time, the existing models, such as U-Net, YOLO-V5, MFAA, SELM, and Vision Transformer, achieved 1086s, 1121s, 1988s, 2641s, and 2035s, respectively; the proposed model achieved 2205s. When the test samples are run, the proposed model takes only 2.5s, whereas the existing models take 8s, 7s, 6s, 4s, and 3s (Table 3).

Table 3: Assessment based on period ingesting with classification accuracy

Classifiers	Training time (seconds)	Testing time (seconds)	Accuracy (%)
U-Net	1086	8	70
YOLO-V5	1121	7	79
MFAA	1988	6	85
SELM	2641	4	89
Vision transformer	2035	3	91
IAS-FPN	2205	2.5	94.07

The reason is that the proposed model extracts relevant features from pre-processed data, whereas the existing models directly classify the samples, thereby increasing training and testing time.

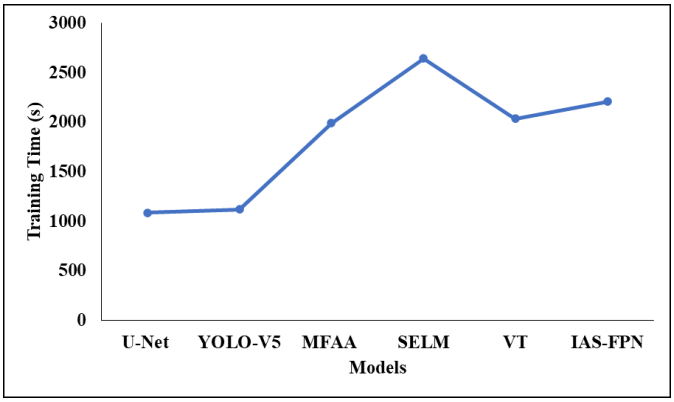

Figure 9: Training time

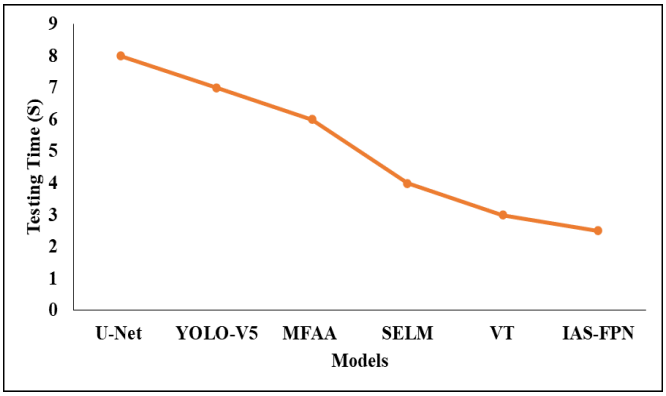

Figure 10: Analysis of various models on testing time

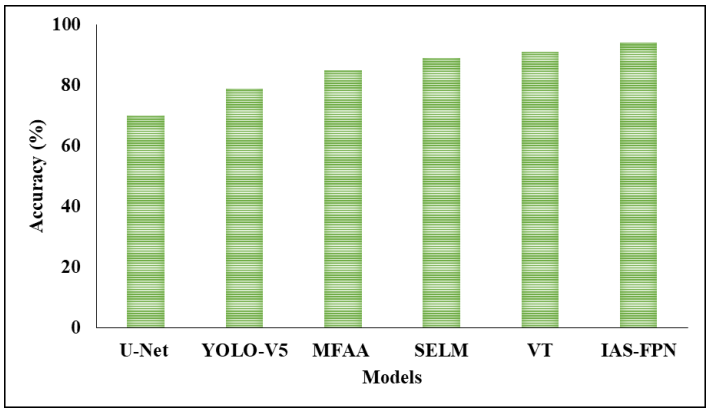

Figure 11: Accuracy comparison

Figures 9, 10 and 11 present a graphical investigation of the performance of different classifiers during training and testing.

5. Conclusion

In this study, we proposed and built a network capable of distinguishing between healthy individuals and cancer patients. The researchers created a new enhanced attention Swin Transformer (IA-Swin) model that leverages contextual information to identify dense cancer lesions accurately. We also developed a network by devising a skip connection technique to retain a large

amount of information about tiny malignant cells. Furthermore, a well-organised neck of the IFPN network was proposed to balance accuracy and efficiency by reducing the number of layers, with IDMO optimising the FPN weights. The proposed approach can also distinguish between AGC and EGC. Because EGC diagnosis is difficult, we developed a CAD system to assist with it. The suggested architecture (IAS-FPN) achieves good results for stomach cancer classification utilising breath analysis, with an overall accuracy of 94%. The algorithm may be learned more quickly. Furthermore, the neural network trained on breath analysis outperformed all other available approaches to date. Although the suggested technique performs well, its accuracy needs improvement to handle real-world scenarios. Other deep learning classifiers will be considered in the future to improve detection efficiency and performance with limited training samples.

Acknowledgement: The authors express their heartfelt gratitude to New Horizon College of Engineering and Quest Technologies for providing essential support and research facilities. They further acknowledge the valuable guidance and encouragement received from the faculty and research teams throughout the study.

Data Availability Statement: The data utilized in this research are available upon reasonable request from the corresponding authors, ensuring transparency and reproducibility of the study.

Funding Statement: The authors confirm that no external funding or financial support was received for the execution of this research or preparation of the manuscript.

Conflicts of Interest Statement: The authors declare that they have no conflicts of interest that could have influenced the research outcomes or the presentation of the findings.

Ethics and Consent Statement: All authors unanimously consent to the dissemination of this publication for academic, research, and educational purposes, ensuring accessibility to interested readers.

Reference

1. A. Agushaka, J. O. Jeffrey, A. E. Ezugwu, and L. Abualigah, "Dwarf mongoose optimization algorithm," *Computer Methods in Applied Mechanics and Engineering*, vol. 391, no. 3, p. 114570, 2022.
2. A. M. Ahmed, T. A. Rashid, and S. A. M. Saeed, "Cat swarm optimization algorithm: A survey and performance evaluation," *Computational Intelligence and Neuroscience*, vol. 2020, no. 1, pp. 1-20, 2020.
3. H. M. Alabool, D. Alarabiat, L. Abualigah, and A. A. Heidari, "Harris hawks optimization: A comprehensive review of recent variants and applications," *Neural Computing and Applications*, vol. 33, no. 15, pp. 8939–8980, 2021.
4. M. A. Aslam, C. Xue, Y. Chen, A. Zhang, M. Liu, K. Wang, and D. Cui, "Breath analysis based early gastric cancer classification from deep stacked sparse autoencoder neural network," *Scientific Reports*, vol. 11, no. 1, p. 4014, 2021.
5. J. W. Chae and H. C. Cho, "Enhanced classification of gastric lesions and early gastric cancer diagnosis in gastroscopy using Multi-Filter AutoAugment," *IEEE Access*, vol. 11, no. 3, pp. 29391–29399, 2023.
6. Q. Guo, W. Yu, S. Song, W. Wang, Y. Xie, L. Huang, J. Wang, Y. Jia, and S. Wang, "Pathological detection of micro and fuzzy gastric cancer cells based on deep learning," *Computational and Mathematical Methods in Medicine*, vol. 2023, no. 1, pp. 1-11, 2023.
7. A. U. Haq, J. P. Li, S. Khan, M. A. Alshara, R. M. Alotaibi, and C. Mawuli, "DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment," *Scientific Reports*, vol. 12, no. 1, p. 15331, 2022.
8. Y. Ikenoyama, T. Hirasawa, M. Ishioka, K. Namikawa, S. Yoshimizu, Y. Horiuchi, A. Ishiyama, T. Yoshio, T. Tsuchida, Y. Takeuchi, S. Shichijo, N. Katayama, J. Fujisaki, and T. Tada, "Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists," *Digestive Endoscopy*, vol. 33, no. 1, pp. 141–150, 2021.
9. M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Systems with Applications*, vol. 149, no. 7, p. 113338, 2020.
10. I. Lansdorp-Vogelaar, R. G. Meester, M. Laszkowska, F. A. Escudero, Z. J. Ward, and J. M. Yeh, "Cost-effectiveness of prevention and early detection of gastric cancer in Western countries," *Best Practice and Research Clinical Gastroenterology*, vol. 50, no. 3, p. 101735, 2021.
11. S. H. Lee, Y. Lee, and H. J. Jang, "Deep learning captures selective features for discrimination of microsatellite instability from pathologic tissue slides of gastric cancer," *International Journal of Cancer*, vol. 152, no. 2, pp. 298–307, 2023.
12. M. Leja and A. Linē, "Early detection of gastric cancer beyond endoscopy—new methods," *Best Practice and Research Clinical Gastroenterology*, vol. 50, no. 3, p. 101731, 2021.
13. J. Li, H. Lei, A. H. Alavi, and G. G. Wang, "Elephant herding optimization: variants, hybrids, and applications," *Mathematics*, vol. 8, no. 9, p. 1415, 2020.

14. X. Li, J. Huang, C. Wang, X. Yu, T. Zhao, C. Huang, and Y. Gao, “Expectation-maximization algorithm leads to domain adaptation for a perineural invasion and nerve extraction task in whole slide digital pathology images,” *Medical and Biological Engineering and Computing*, vol. 61, no. 2, pp. 457–473, 2023.
15. K. Liu, B. Liu, Y. Zhang, Q. Wu, M. Zhong, L. Shang, Y. Wang, P. Liang, W. Wang, Q. Zhao, and B. Li, “Building an ensemble learning model for gastric cancer cell line classification via rapid Raman spectroscopy,” *Computational and Structural Biotechnology Journal*, vol. 21, no. 12, pp. 802–811, 2023.
16. L. Ma, X. Su, L. Ma, X. Gao, and M. Sun, “Deep learning for classification and localization of early gastric cancer in endoscopic images,” *Biomedical Signal Processing and Control*, vol. 79, no. 1, p. 104200, 2023.
17. D. Połap and M. Woźniak, “Red fox optimization algorithm,” *Expert Systems with Applications*, vol. 166, no. 3, p. 114107, 2021.
18. P. K. Balasubramanian, W. C. Lai, G. H. Seng, C. Kavitha, and J. Selvaraj, “Apestnet with mask R-CNN for liver tumor segmentation and classification,” *Cancers*, vol. 15, no. 2, p. 330, 2023.
19. S. Qahtan, K. Y. Sharif, A. A. Zaidan, H. A. AlSattar, O. S. Albahri, B. B. Zaidan, H. Zulzalil, M. H. Osman, A. H. Alamoodi, and R. T. Mohammed, “Novel multi security and privacy benchmarking framework for blockchain-based IoT healthcare industry 4.0 systems,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 6415–6423, 2022.
20. A. Raghuvanshi, U. K. Singh, and C. Joshi, “A review of various security and privacy innovations for IoT applications in healthcare,” *in Advanced Healthcare Systems: Empowering Physicians with IoT-Enabled Technologies, Wiley online library*, New Jersey, United States of America, 2022.
21. H. Sasaki, T. Horiuchi, and S. Kato, “A study on vision-based mobile robot learning by deep Q-network,” *in 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, Kanazawa, Japan, 2017.
22. J. B. Y. So, R. Kapoor, F. Zhu, C. Koh, L. Zhou, R. Zou, Y. C. Tang, P. C. K. Goo, S. Y. Rha, H. C. Chung, J. Yoong, C. T. Yap, J. Rao, C. K. Chia, S. Tsao, A. Shabbir, J. Lee, K. P. Lam, M. Hartman, W. P. Yong, H. P. Too, and K. G. Yeoh, “Development and validation of a serum microRNA biomarker panel for detecting gastric cancer in a high-risk population,” *Gut*, vol. 70, no. 5, pp. 829–837, 2021.
23. K. Thilagam, A. Beno, M. V. Lakshmi, C. B. Wilfred, S. M. George, M. Karthikeyan, V. Peroumal, C. Ramesh, and P. Karunakaran, “Secure IoT healthcare architecture with deep learning-based access control system,” *Journal of Nanomaterials*, vol. 2022, no. 1, pp. 1-8, 2022.
24. C. Tian, Y. Xu, W. Zuo, B. Zhang, L. Fei, and C. W. Lin, “Coarse-to-fine CNN for image super-resolution,” *IEEE Transactions on Multimedia*, vol. 23, no. 6, pp. 1489–1502, 2020.
25. C. Tian, Y. Yuan, S. Zhang, C. W. Lin, W. Zuo, and D. Zhang, “Image super-resolution with an enhanced group convolutional neural network,” *Neural Networks*, vol. 153, no. 9, pp. 373–385, 2022.
26. L. Wu, X. He, M. Liu, H. Xie, P. An, J. Zhang, H. Zhang, Y. Ai, Q. Tong, M. Guo, M. Huang, C. Ge, Z. Yang, J. Yuan, J. Liu, W. Zhou, X. Jiang, X. Huang, G. Mu, X. Wan, Y. Li, H. Wang, Y. Wang, H. Zhang, D. Chen, D. Gong, J. Wang, L. Huang, J. Li, L. Yao, Y. Zhu, and H. Yu, “Evaluation of the effects of an artificial intelligence system on endoscopy quality and preliminary testing of its performance in detecting early gastric cancer: a randomized controlled trial,” *Endoscopy*, vol. 53, no. 12, pp. 1199–1207, 2021.
27. N. Yoshida, H. Doyama, T. Yano, T. Horimatsu, N. Uedo, Y. Yamamoto, N. Kakushima, H. Kanzaki, S. Hori, K. Yao, I. Oda, C. Katada, C. Yokoi, K. Ohata, K. Yoshimura, H. Ishikawa, and M. Muto, “Early gastric cancer detection in high-risk patients: a multicentre randomised controlled trial on the effect of second-generation narrow band imaging,” *Gut*, vol. 70, no. 1, pp. 67–75, 2021.
28. C. Zhang, A. Ding, Z. Fu, J. Ni, Q. Chen, Z. Xiong, B. Liu, Y. Cao, S. Chen, and X. Liu, “Deep learning for gastric location classification: An analysis of location boundaries and improvements through attention and contrastive learning,” *Smart Health*, vol. 28, no. 6, p. 100394, 2023.
29. J. Zhang, Y. Tian, Z. Luo, C. Qian, W. Li, and Y. Duan, “Breath volatile organic compound analysis: an emerging method for gastric cancer detection,” *J. Breath Res.*, vol. 15, no. 4, p.044002, 2021.
30. X. Zhu, Y. Ma, D. Guo, J. Men, C. Xue, X. Cao, and Z. Zhang, “A framework to predict gastric cancer based on tongue features and deep learning,” *Micromachines*, vol. 14, no. 1, p. 53, 2022.