



# Classification and Detection of Gastrointestinal Cancer Diagnosis Using Advanced Deep Learning Technique: An Explainable AI Approach

Renjith V.R<sup>1</sup>, Dr. J.E Judith<sup>2</sup>

<sup>1</sup>Research Scholar, Department of Computer Science and Engineering, NICHE, Kumaracoil, Tamil Nadu, India.

<sup>2</sup>Associate Professor, Department of Computer Science and Engineering, NICHE, Kumaracoil, Tamil Nadu, India.

Email: [renjith@gmail.com](mailto:renjith@gmail.com)<sup>1</sup>, [judith.com](mailto:judith.com)<sup>2</sup>

## Abstract

Globally, gastric cancer ranks third in terms of cancer-related deaths. The histological interpretation of Gastric specimens is essential for clinical management, and this requires skilled pathologists. Deep learning systems have surpassed human pathologists in several domains. Current models are trained in a way that reduces the amount of knowledge acquired for the model and eliminates the need for extra stages to improve performance. Enhancements are necessary to strengthen the dependability of current methods when dealing with applications demanding accurate classification outcomes. In this work, we have proposed a gastrointestinal cancer diagnosis system with an explainable AI approach. The autoencoder has been employed for feature extraction, while Particle Swarm Optimization (PSO) selects the optimal features. Then, the DenseNet-169 architecture model has been employed for effective cancer diagnosis and classification. Finally, the Grad CAM is proposed to implement the explainable AI approach. The proposed system has been implemented and the performance is verified.

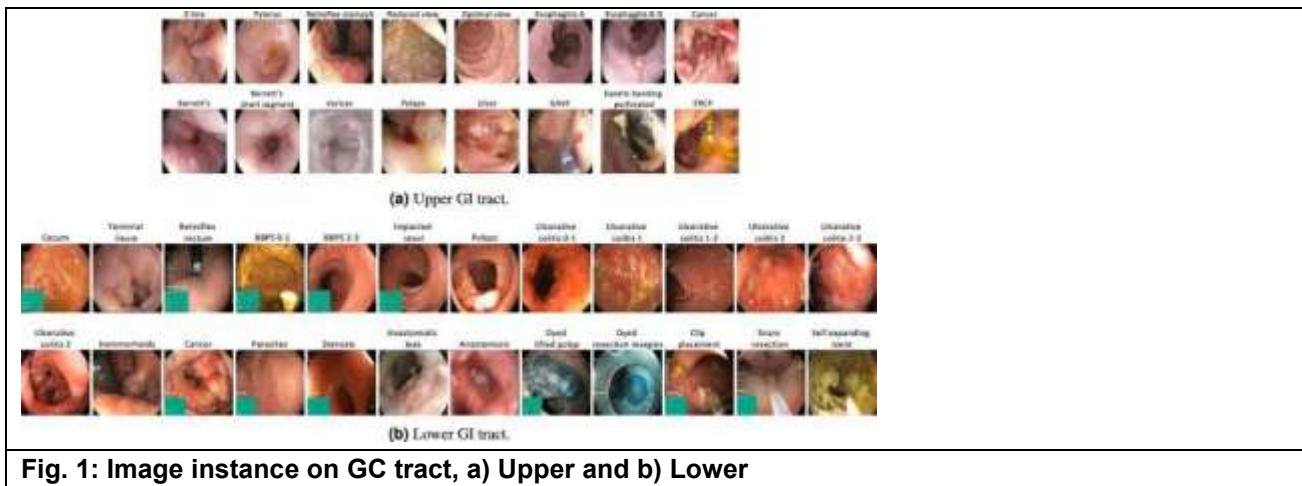
*Keywords: Gastrointestinal, Deep learning, Cancer diagnosis, Artificial intelligence, particle swarm optimization, Autoencoder.*

This is an open access article under CC BY 4.0, allowing unrestricted use with proper attribution, a license link, and indication of any changes made.

## 1. Introduction

One of the frequent cancers that starts in the stomach mucosa's epithelial cells is Gastric Cancer (GC). Global studies indicate that Genetic Cancer (GC) ranks as the second primary contributor to worldwide cancer-related fatalities and occupies the fourth spot in terms of global cancer incidence [1-2]. Various elements, including both genetic and environmental factors, shape the intricate mechanisms governing the emergence and initiation of GC. The full extent of how these factors collectively influence the occurrence of GC remains uncertain. Despite the widespread use of interventions like surgery, chemotherapy, and radiation therapy, the five-year survival rate for advanced gastric cancer hovers around 30%. Conversely, early gastric cancer treatment boasts a potential for a complete cure, with a five-year survival rate exceeding 90% [3]. Early detection of GC is therefore crucial.

The genesis and progression of GC are intricate processes involving numerous steps, stages, and mechanisms [4-5]. There are several phases in between (such as the precancerous condition). The pattern of human GC that is currently more widely accepted was put forth by Correa and includes the following conditions: intestinal metaplasia, dysplasia, atrophic gastritis, chronic non-atrophic gastritis, and normal gastric mucosa [6]. Atrophic Gastritis and Intestinal Metaplasia are recognized as precancerous conditions intricately linked to the onset of gastric cancer. If these are not treated promptly, there is a higher chance that they will turn into GC [7]. Their prompt identification and management hold substantial pragmatic implications for the avoidance and management of GC. The sample images of the gastric cancer are depicted in Figure 1.



**Fig. 1: Image instance on GC tract, a) Upper and b) Lower**

There are currently several methods available for the early identification of GC. Endoscopic techniques [8], such as general endoscopy, magnifying endoscopy, chromoendoscopy, and histopathological techniques [9–10], imaging techniques [11], such as nuclear magnetic resonance, X-ray, computed tomography (CT), and tumour marker techniques [12], such as pepsinogen [13], gastrin 17 [14], GC markers, etc., are included in this category. These diagnostic strategies still face specific limitations. The imaging diagnostic method encounters challenges in accurately identifying early lesions. The histopathological diagnostic approach [15] requires time-consuming and invasive tests, as well as analytical techniques that require specific training and expertise, and the endoscopic diagnostic approach is arbitrary and prone to missing anomalies. While most studies employ tumour markers to assess the therapeutic impact of GC, no reliable tumour marker has yet been discovered for GC. So, it is essential to develop a rapid, accurate approach for the early detection of GC. This study is focused on presenting a proficient deep learning model designed to predict GIC, with the subsequent goals:

- By conducting comparative studies on various deep learning architectures, aim to create the most effective system for detecting, classifying, and diagnosing Gastrointestinal Cancers.
- With the help of the Densenet 169 network, the GC was able to detect and classify in a much more effective way, and also with the inclusion of explainable AI.
- To boost the performance of the classifier, use pre-stages such as feature extraction and selection with techniques like Autoencoder and Particle Swarm Optimization.

The paper has the following structure: Section 2 delves deeply into recent literature on gastric cancer detection. In Section 3, the proposed methodology is thoroughly investigated. Section 4 presents the implementation's results and discussions, and Section 5 includes the conclusion of the paper.

## 2. Literature Review

Mangesh et al. (2023) [16] presented a novel approach named the Marine Predator's Algorithm with DL for Lung and Colon Cancer Classification. This sophisticated strategy was meticulously designed to efficiently categorize various types of lung and colon cancers in histopathological images. Contrast enhancement based on CLAHE is the first stage in the MPADL-LC3 approach. This approach uses MPA as a hyperparameter optimiser and MobileNet to generate feature vectors. Also, lung and colon cancer classification was done using Deep Belief Networks. This method's effectiveness was assessed using simulated values on benchmark datasets.

Gulekan et al. (2023) [17] introduced Raman spectroscopy as a reliable and minimally invasive method to assess serum levels of Tumor Markers (TMs) in blood samples. The significance of serum TMs levels post-curative gastrectomy lies in their ability to predict the recurrence of gastric cancer, emphasizing the importance of early detection. Raman measurements and ELISA test results were employed in the study to experimentally assess the levels of TMs. These measurements were later employed to develop a machine learning-based predictive model. The research involved 70 participants in total, including 26 individuals who had undergone surgery for gastric cancer and 44 healthy individuals.

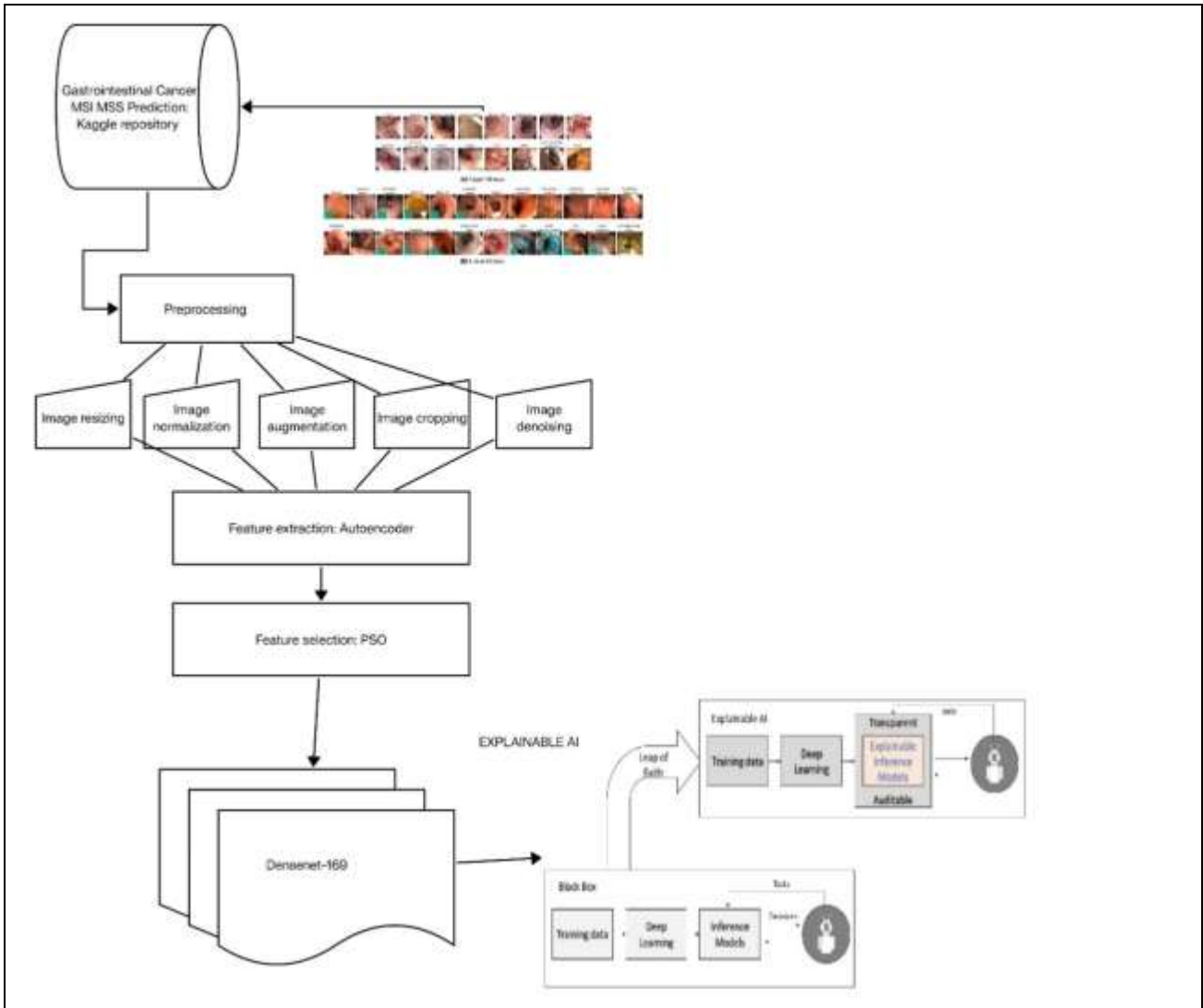
Wang et al. (2023) [18] provided an overview of how artificial intelligence algorithms are applied to multi-dimensional data in the context of gastric cancer. This encompasses clinical and follow-up information, traditional imaging modalities (such as endoscopes, histopathology, and CT), molecular biomarkers, and more. The goal was to enhance risk surveillance in individuals with established risk factors for gastric cancer, improve diagnostic accuracy, predict survival in confirmed gastric cancer cases, and anticipate treatment outcomes to support clinical decision-making. As a result, artificial intelligence has a substantial impact on multiple aspects of gastric cancer, encompassing enhancements in diagnosis and the progression of precision medicine.

Yuan et al. (2023) [19] utilized deep learning models within the realm of Artificial Intelligence (AI) to assess the diagnostic effectiveness of both tongue images and the microbiome of tongue coating in detecting GC. Recognizing the practicality and cost-effectiveness of incorporating tongue imaging into diagnostics, the researchers conducted a prospective multicenter clinical investigation in China spanning from May 2020 to March 2022. This extensive study, involving 937 GC patients and 1911 non-GC participants recruited from China, aimed to further investigate the potential of tongue images in the diagnosis of GC. Furthermore, the researchers validated their methodology in an independent external cohort comprising 294 GC patients and 521 non-GC participants from 7 centres.

Tian et al. (2023) [20] conducted a prognostic assessment on patients who had encountered distant metastases. They explored independent prognostic factors using both univariate and multivariate regression analyses. The study employed the SEER dataset, which included 2698 cases, with 314 experiencing distant metastases, and also involved 107 hospital patients, among whom 14 had distant metastases. The examination identified several factors, including age, T-stage, N-stage, tumor size, grade, and tumor location, as independent risk factors for distant metastases in stage T1 GC. The investigation employed Kaplan-Meier curves to visually illustrate variations in survival prognosis for each variable and subvariable. To enhance predictive accuracy, a thorough analysis utilizing seven ML algorithms was conducted on both training and test sets. The random forest model demonstrated the highest performance in prediction, achieving an AUC of 0.941, an Accuracy of 0.917, a Recall of 0.841, a Specificity of 0.927, and an F1-score of 0.877.

### **3. Methodology**

The biomedical discipline is actively researching the diagnosis of gastrointestinal (GI) tract diseases by endoscopic image categorization. In recent times, various research groups have developed various ways of classifying GI tract diseases, based on image processing and ML techniques. It has been observed that the models operate independently without acquiring additional knowledge. There is also an evident inadequacy in the analysis of datasets, despite some literature references. For instance, a dataset focusing on a single crop and a limited range of diseases was used, suggesting that the model's performance could be enhanced by refining the dataset. The models appear to lack sufficient training to acquire relevant knowledge from the input provided. Existing models are trained in a way that the system doesn't have additional stages to boost the performance and also incorporate knowledge gained for the model is less.

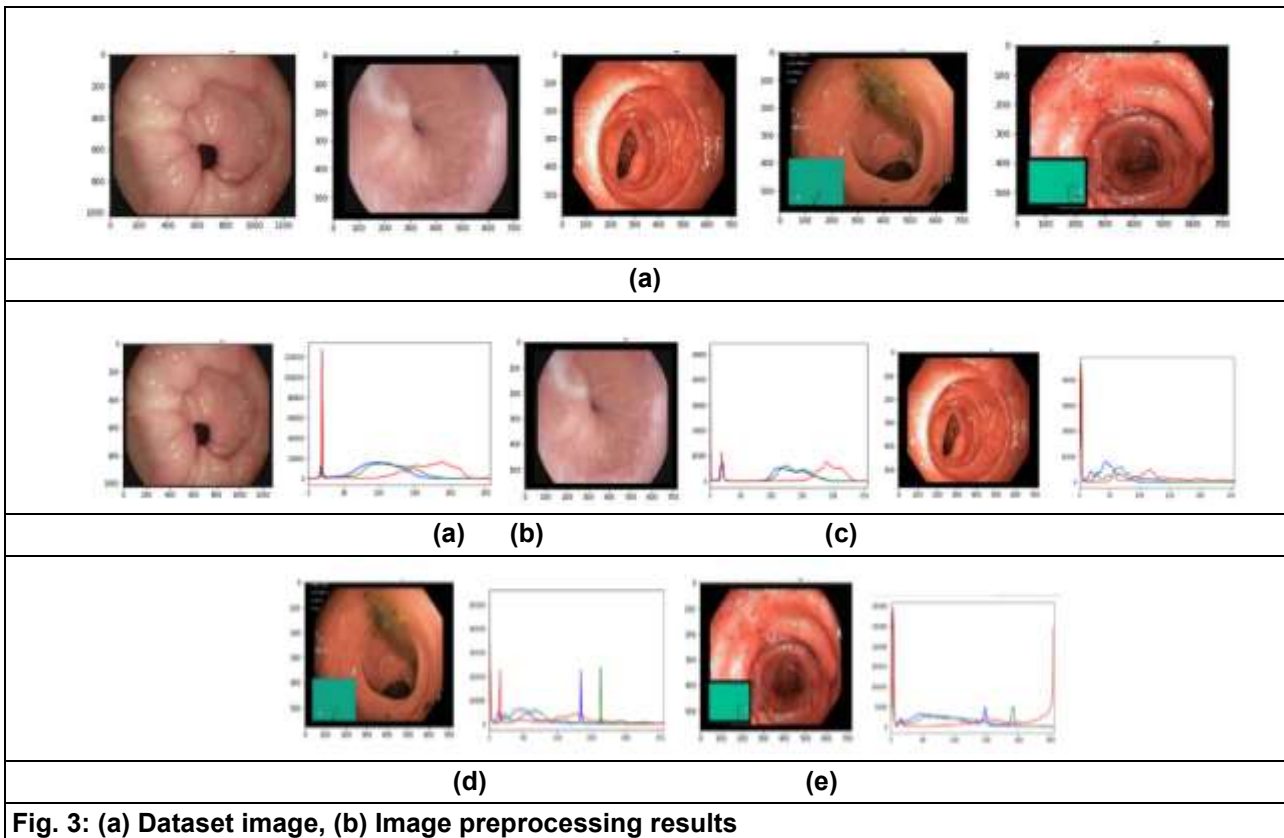


**Fig. 2. Proposed Architecture**

Figure 2 depicts the system architecture, which is divided into the following stages. For a deep learning system to function, a large enough dataset must be fed to the model to comprehend all of its features and patterns. Thus, the Gastrointestinal cancer dataset from a well-known repository has been used for this study project. Microsatellite Instability (MSI) or Microsatellite Stable (MSS) gastrointestinal cancer is predicted for a patient using a set of medical images called the Gastrointestinal Cancer MSI MSS Prediction dataset. A greater risk of cancer is linked to a genetic disorder called MSI. Compared to MSI, MSS is associated with a decreased risk of cancer. Microsatellite Instability (MSI) is a genetic condition associated with defects in DNA mismatch repair mechanisms and is linked to a higher risk of cancer progression and poor prognosis. In contrast, Microsatellite Stable (MSS) tumors generally indicate a lower genetic instability and comparatively reduced cancer risk. While the existing dataset provides a strong foundation for model development and validation, its limited size and source diversity may restrict the model’s ability to generalize across different patient populations, imaging protocols, and clinical settings. Expanding the dataset—by incorporating multi-center data, diverse patient demographics, and additional imaging variations—can significantly enhance model robustness and generalization. Furthermore, techniques such as data augmentation, cross-dataset validation, and external cohort testing can help simulate real-world variability and reduce overfitting. Therefore, dataset expansion is a critical step toward improving the clinical reliability and deployment readiness of deep learning-based gastrointestinal cancer diagnostic systems.

### 3.1 Preprocessing

The dataset contains images of gastrointestinal tumor images, which are divided into training and testing data. The image data used for training undergoes preprocessing to enhance image quality, preparing them for subsequent processing steps. The preprocessing stage includes techniques like Image resizing, Image normalization, Image augmentation, Image cropping, and Image denoising.



Initially, the image was resized to  $244 \times 244$  pixels and adjusted for colour permanence. After that, the gastrointestinal images meant for the three RGB channels were determined. The crucial step of image normalisation is necessary for preprocessing medical images, particularly those that detect GC. Image quality is improved through noise reduction, contrast, and brightness adjustments. The adaptive contrast enhancement method serves as a foundational benchmark for more sophisticated algorithms. It entails enhancing the contrast of small image tiles by applying contrast-limited adaptive histogram equalization, leveraging the exponential distribution of the histogram.

Image augmentation is commonly employed to enhance the effectiveness of deep networks. Enlarging the training set without acquiring new images facilitates the development of a robust image classifier even with limited training data. Employing diverse processing techniques or a combination thereof, like random rotation, shifts, shear, flips, random cropping, etc., image augmentation artificially generates additional training images.

### 3.2 Autoencoder

A feedforward, unsupervised neural network is called an autoencoder, which is employed for feature extraction from preprocessed data. As per the earlier research, an increased number of hidden layers results in a significant increase in computation but does not affect the ultimate performance. This led to the implementation of a three-layered autoencoder.

Improved classification model prediction performance is attained through the utilization of an autoencoder. This autoencoder diminishes reconstruction errors while imposing sparsity constraints on

individual hidden nodes. This method leads to a feature representation that is both more robust and meaningful. The process begins with inputting a high-dimensional feature vector for the gastrointestinal cancer pair into an encoder by

$$H = f(WX + b) \quad (1)$$

In the scenario where  $X$  denotes a collection of  $n$  samples, each defined by a  $d$ -dimensional vector,  $H$  indicates the low-dimensional features obtained through the encoding process. In this context,  $W$  denotes the weight,  $b$  indicates bias, and  $f(\cdot)$  indicates the encoding function of the encoder.

Subsequently, the decoder reconstructs the visual appearance of the input feature representation from the low-dimensional representation  $H$ .

$$\hat{X} = g(W'H + b') \quad (2)$$

In this context,  $\hat{X}$  represents the acquired feature representation, while  $W'$ ,  $b'$ , and  $g(\cdot)$  refer to the decoder's weight, bias, and decoding function, accordingly.

In order to reduce the reconstruction error, we establish a cost function through the process of

$$E = MSE + \lambda \times \Omega_{sparsity} + \beta \times \Omega_{weights} \quad (3)$$

The parameters  $\lambda$  and  $\beta$  represent the sparsity regularization and the coefficients for  $L_2$  regularization, correspondingly.

Mean Square Error ( $MSE$ ) is the initial term. The term is used to quantify the difference between the reconstructed characteristics ( $\hat{X}$ ) on training data and the input characteristics ( $X$ ).

$$MSE = \frac{1}{n} \sum_{i=1}^N (X_i - \hat{X}_i)^2 \quad (4)$$

The Kullback–Leibler divergence, denoted as  $\Omega_{sparsity}$  in the second term, serves the purpose of regulating sparsity according to the sparsity proportion  $\rho$  by:

$$\Omega_{sparsity} = \sum_{t=1}^{S_l} KL(\rho \parallel \hat{\rho}_t) \quad (5)$$

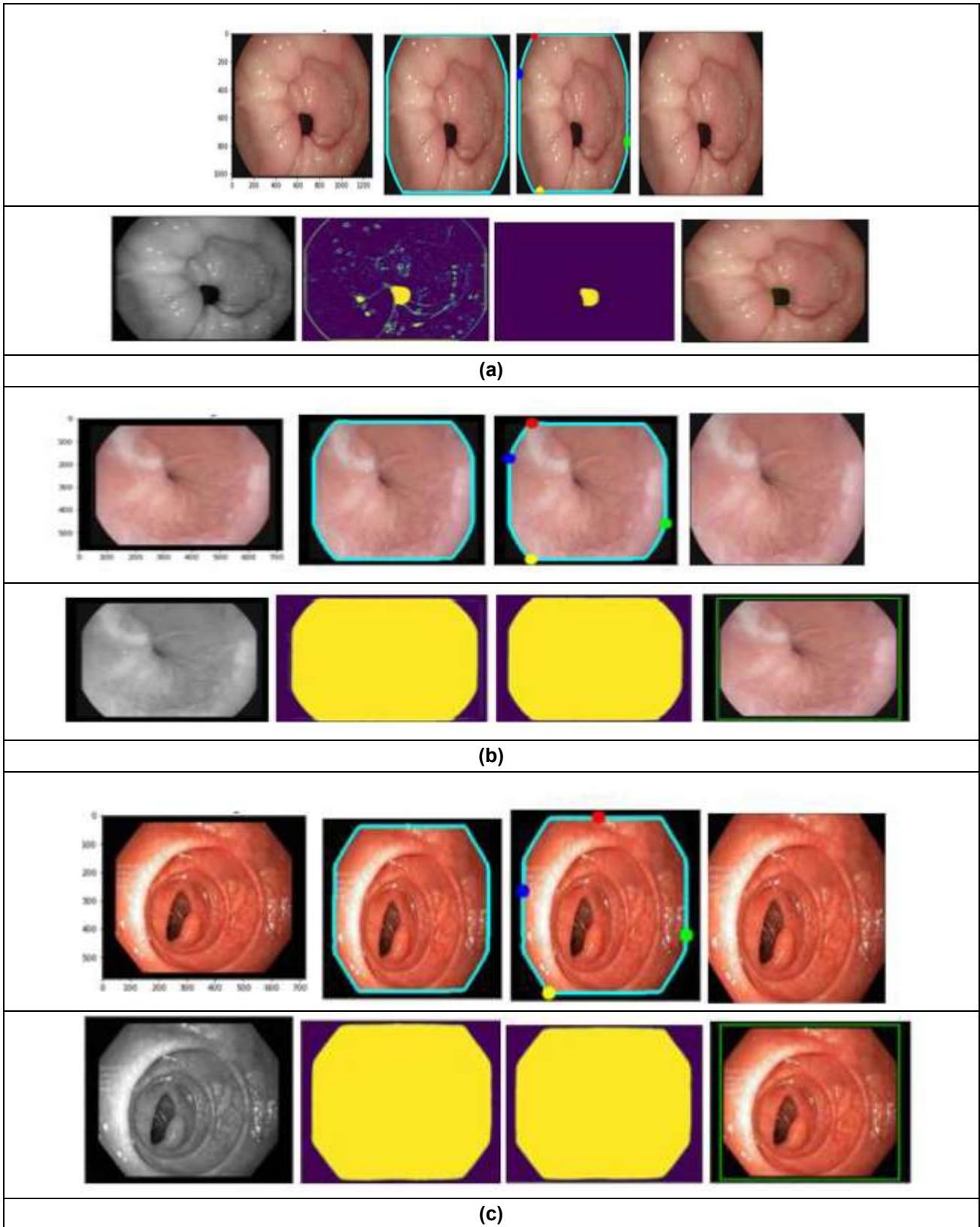
In the given context,  $S_l$  represents the count of neurons in the  $l^{th}$  hidden layer, and  $(\hat{\rho}_t)$  signifies the average activity of the  $t^{th}$  neuron. The term  $KL(\rho \parallel \hat{\rho}_t)$  refers to the relative entropy between Bernoulli random variables characterized by means  $\rho$  and  $(\hat{\rho}_t)$ . The computation of  $KL(\rho \parallel \hat{\rho}_t)$  involves,

$$KL(\rho \parallel \hat{\rho}_t) = \rho \log \log \frac{\rho}{\hat{\rho}_t} + (1 - \rho) \log \log \frac{1-\rho}{1-\hat{\rho}_t} \quad (6)$$

The third term is  $L_2$  regularization term  $\Omega_{weights}$ . The term is employed for managing weights and preventing overfitting.

$$\Omega_{weights} = \frac{1}{2} \sum_{l=1}^{N_l-1} \sum_{i=1}^{S_l} \sum_{j=1}^{S_{l+1}} (w_{ji}^{(l)})^2 \quad (7)$$

where  $N_l$  indicate the layer numbers  $S_l$  denotes the number of units in the  $l^{th}$  layer, and  $w_{ji}^{(l)}$  is the weight, in that order.



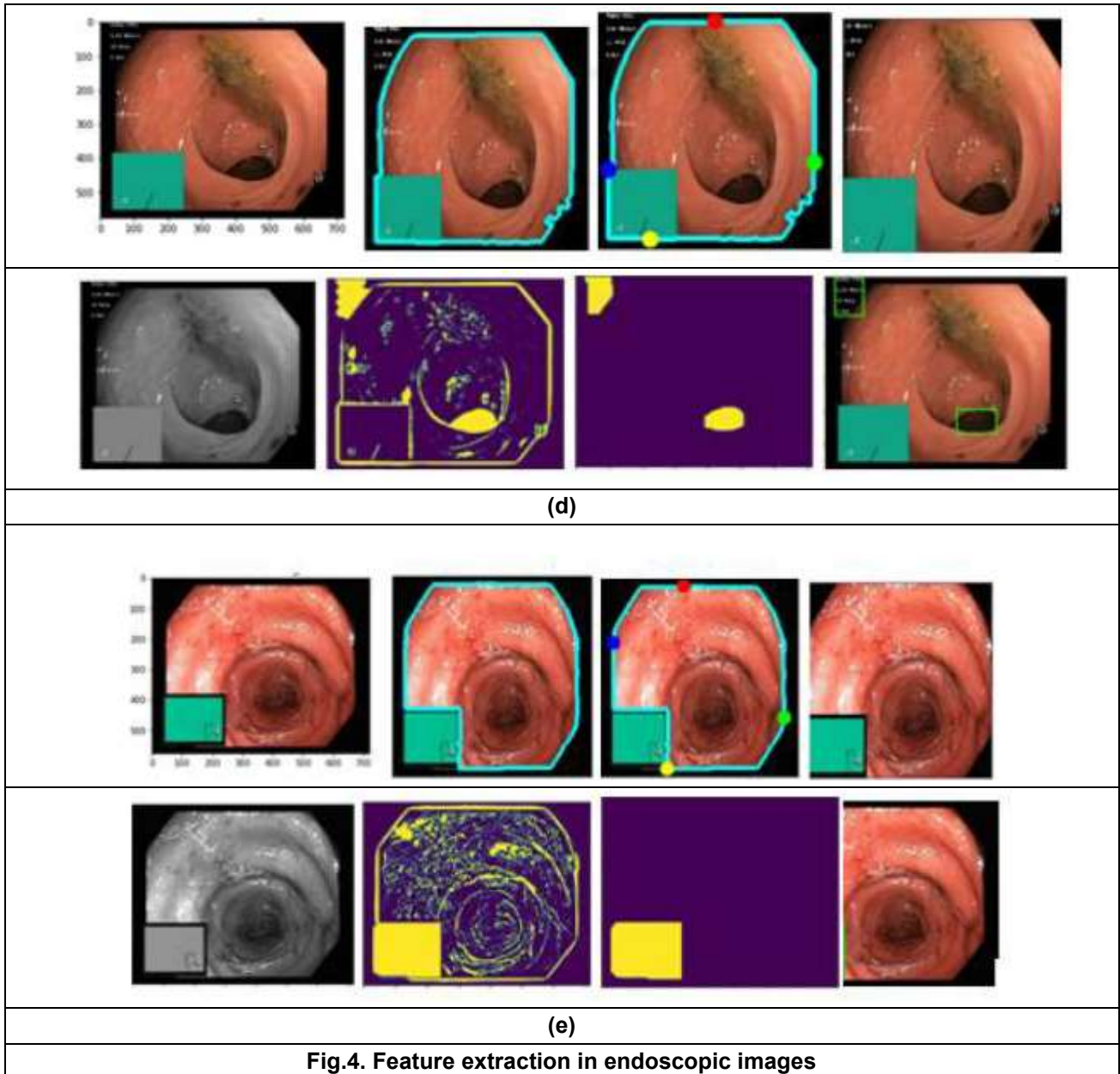


Fig.4. Feature extraction in endoscopic images

### 3.3 Particle swarm optimization

Using the features that have been extracted, Particle Swarm Optimization (PSO) is utilized to carefully choose the features essential for diagnosing gastrointestinal cancer. Since its inception as an optimization strategy applicable to real-number spaces, this method has demonstrated success across diverse domains. It has effectively been employed in tasks ranging from function optimization to training artificial neural networks, controlling fuzzy systems, and addressing various application problems. In spaces where optimization problems emerge, there exist distinct and qualitative differences among variables and across different levels of variables.

When evaluating the overall features extracted by an autoencoder, let  $n$  denote this quantity. In the framework of Particle Swarm Optimization (PSO), a particle  $i$  is characterized by a pair  $(X_i, V_i)$ , comprising two  $n$ -dimensional vectors. Here,  $V_i$  signifies the particle's velocity, while  $X_i$  denotes its position represented as a binary vector. In the Particle Swarm Optimization algorithm, each solution to a given problem is regarded as a particle endowed with the capacity to traverse through a search landscape.

The particle's motion is described by two vectors: the particle vector  $X_i$  and the velocity vector  $V_i$ .

$$X_i = (x_i^1, x_i^2, \dots, x_i^n) \quad (8)$$

$$V_i = (v_i^1, v_i^2, \dots, v_i^n) \quad (9)$$

where  $x_i^d \in [0, 1]$ ,  $i = 1, 2, \dots, m$  ( $m$  is the particles number in total),  $d = 1, 2, \dots, n$  ( $n$  denotes dimension of data).

To alter the position vector of a particle, specify both the direction of its movement and its velocity as follows:

$$v_i^{d+1} = w \times v_i^d + c_1 r_1 \times (p_{opt_i}^d - x_i^d) + c_2 r_2 \times (g_{opt_i}^d - x_i^d) \quad (10)$$

In the context of particle positions and optimization, let  $x_i^d$  denote the position of particle  $i$  along dimension  $d$ . Additionally,  $p_{opt_i}^d$  represents the optimal previous position of the  $i^{\text{th}}$  particle, and  $g_{opt_i}^d$  signifies the global optimal position considering all particles. The parameters include  $w$  for the inertial weight,  $v_i^d$  for the velocity of particle  $i$  in dimension  $d$ , and acceleration constants  $c_1$  and  $c_2$ . Random values  $r_1$  and  $r_2$  are also part of the equation.

The upcoming velocity is influenced by three factors: the distance from the previous optimum ( $p_{opt}$ ), the distance from the global optimum ( $g_{opt}$ ), and the current velocity ( $v_i^d$ ).

Defining the distance to the prior optimum and the distance to the global optimum is essential for the second and third components. The influence of each component is then varied by multiplying it by a random value  $r$  and an acceleration constant  $c$ .

The subsequent equation is utilized to revise the location of every particle.

$$x_i^d = \begin{cases} 1, & \text{if } r_3 < s(v_i^d) \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

The velocity in PSO signifies the likelihood that the associated element in the position vector will receive a value of 1. To bring  $v_i^d$  within the range of  $(0, 1)$ , a sigmoid function, denoted as  $s(v_i^d)$ , is employed.  $r_3$  is a generated random value.

$$s(v_i^d) = \frac{1}{1 + e^{-v_i^d}} \quad (12)$$

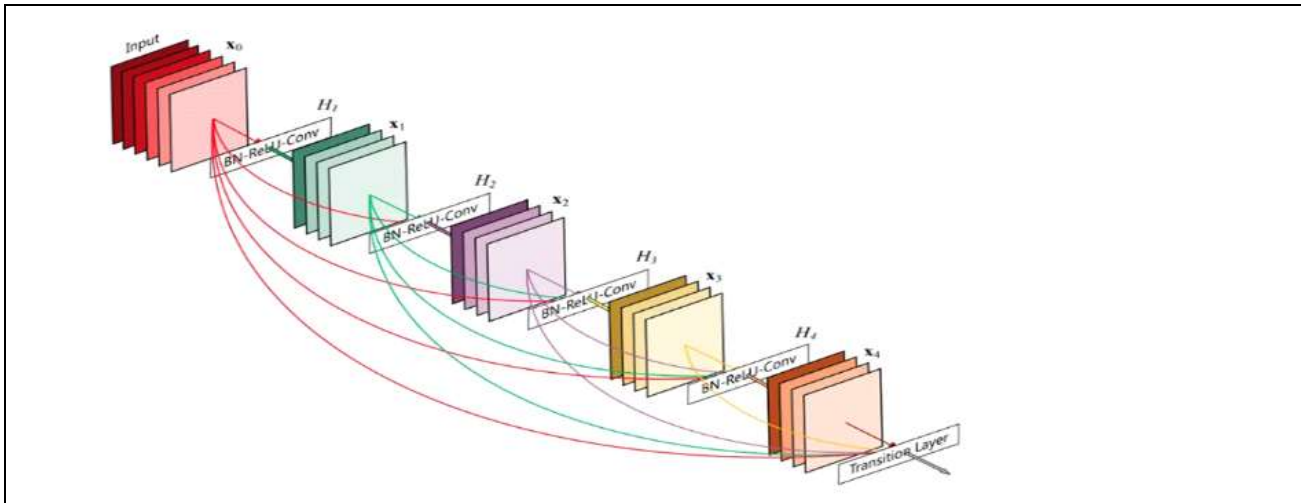
The fitness function of the PSO feature selection is expressed as:

$$fitness(X_i) = w_1 \times A(X_i) + [w_2(M - R(X_i))/M] \quad (13)$$

$A(X_i)$  represents the leave-one-out cross-validation accuracy within the  $[0,1]$  range for the training set, utilizing solely the features identified in  $X_i$ .  $R(X_i)$  signifies the count of selected features in  $X_i$ , while  $M$  denotes the overall number of features for each sample. Additionally,  $w_1$  and  $w_2$  denote priority weights allocated to accuracy and the quantity of chosen features, respectively, reflecting their respective levels of importance,  $w_1 \in [0.1, 0.9]$ ,  $w_2 = 1 - w_1$ .

### 3.4 DenseNet-169

The DenseNet-169 has been employed to diagnose gastrointestinal cancer from the selected features. The DenseNet-169 architecture was utilized to train the neural network model, incorporating a compression rate of 0.5, dropout rate of 0.2, and bottleneck layers. This enabled precise classification of individual patches, distinguishing between cancerous and normal background gastric tissue patches. Figure 4 illustrates the architecture of the DenseNet. The training process spanned 100 epochs, employing a learning rate of 0.1 for stochastic gradient descent, followed by 30 epochs at a learning rate of 0.01 and an additional 20 epochs at a learning rate of 0.001. Set at 0.99 was the momentum for batch normalization.



**Fig.5: DenseNet architecture**

The evaluation of the model took place during the final stage of training and after each epoch, aiming to monitor progress and make necessary adjustments to training parameters. Different learning rate schedules were used, along with a few other model architecture adjustments.

The DenseNet-169 network is composed of L layers, also known as strata, incorporating a non-linear transformation  $H_l(\cdot)$  to convert each layer l into an index. The composite function  $H_l(t)$  involves batch normalization, rectified linear units, pooling, and convolution. In the study, the output of the "lth layer" is documented as  $x_l$ . The transition from the l<sup>th</sup> layer to the (l+1)<sup>th</sup> layer is achieved by linking them through a standard convolutional feeding network, with the output of the l<sup>th</sup> layer serving as input to the next layer.

$$x_l = H_l(x_{l-1}) \quad (14)$$

More information can pass from one layer to the next more quickly due to the data flow across the different layers. For this reason, the input of the l<sup>th</sup> layer is the feature maps of all previous layers.

$$x_l = H_l(|x_0, x_1, \dots, x_{l-1}|) \quad (15)$$

Where,  $x_0, x_1, \dots, x_{l-1}$  are feature map concatenations. These involve combining  $H_l(\cdot)$  inputs to streamline implementation, taking into account the feature map's size. However, in dense networks, downsampling layers, crucial for reducing the feature map's size, play a pivotal role in the overall network operation. The downsampling process is facilitated in networks with closely interconnected layers, as the downsampling outcomes are retained within those blocks. In the research, layers that combine and accumulate are denoted as transitional layers. Transitional layers consist of a batch standardization layer, a 1x1 convolutional layer, and a 2x2 mean pooling layer.

In the input layer of DenseNet-169, the number of feature maps is represented as  $k_0+k \times (l-1)$ , where  $k_0$  channels are used to follow the l<sup>th</sup> layer, assuming each  $H_l$  function produces k feature maps. The hyperparameter K, denoting the rate of network growth, is crucial. Despite the limited output characteristics of each layer to k, the actual number of entries is often significantly higher. Scientists have discovered that introducing a bottleneck layer with a 1x1 convolution prior to each 3x3 convolution can decrease the convolution count and improve the computational efficiency of input data.

Reducing the functional maps on transition layers is employed to enhance model compactness. The investigation produces output feature maps employing a compression factor of 0 through the specified transition layer, given the presence of m-function maps within a dense block. The transition layers consistently maintain the same number of feature maps when  $\theta = 1$ . The model's performance was evaluated using a modest-sized external validation dataset, demonstrating consistent results. Furthermore, the acquired model was utilized to examine complete slide images from a test set that were not encountered in the training phase, introducing an extra layer of complexity.

### 3.5 Explainable AI

Explainable AI is the capacity of an AI model to offer transparent and understandable explanations for its predictions. The goal of Explainable AI with DenseNet-169 is to make the model's decisions more comprehensible and interpretable for humans. Achieving explainability can be accomplished through methods such as Gradient-weighted Class Activation Mapping (Grad-CAM). Grad CAM identifies crucial regions in the input image that had a substantial influence on the model's decision, thereby improving the clarity of the explanation. Implementing Grad-CAM on DenseNet-169 allows the creation of visual heatmaps, illustrating the pivotal regions in an input image that plays an important role in the model's predictive outcome. These heatmaps can help users understand why the model made a specific decision and gain insights into the features that influenced the outcome.

Grad-CAM assesses the importance of individual neurons in influencing a specific decision by leveraging the gradient information that propagates to the final convolutional layer of the NN. Although our approach is exceptionally adaptable and can be employed for any activation in deep networks, our primary emphasis in this investigation is on revealing potential predictions that the network might generate.

To access the rad-CAM  $L_{Grad-CAM}^c \in R^{u \times v}$  of width  $u$  and height  $v$  for any class  $c$ , The gradient for class  $c, y^c$  Score (before the softmax) is computed first concerning feature maps  $A^k$  of a convolutional layer, i.e.  $\frac{\partial y^c}{\partial A^k}$ .

The gradients that flow backwards are globally average-pooled to acquire the neuron importance weights in the following manner:

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (16)$$

The significance of feature map  $k$  in relation to the target class  $c$  becomes evident through the weight  $\alpha_k^c$ . This weight signifies the partial linearization of the deep network beyond point  $A$ , providing clarity on its role.

We perform a series of computations involving forward activation maps and corresponding weights, followed by the application of a ReLU activation to achieve the desired outcome.

$$L_{Grad-CAM}^c = ReLU(\sum_k \alpha_k^c A^k) \quad (17)$$

Observing that this leads to a rough heat map that matches the dimensions of DenseNet-169's convolutional feature maps. Our focus is solely on features that contribute positively to the target class. In other words, we apply a Rectified Linear Unit (ReLU) to the linear combination of maps, highlighting pixels where intensity should be augmented to enhance  $y^c$ .

Negative pixels in the image most likely fall into several groups. As might be predicted, localization maps with no ReLU occasionally emphasize classes other than the target one and perform worse when it comes to localization. The output of an image classification CNN, represented by a class score, is not necessarily limited to  $y^c$ . Words from a caption or a response to a query can serve as examples of differentiable activations as well.

CAM generates a localization map for a CNN designed with a particular architecture. In this structure, the softmax function takes the global average-pooled convolutional feature maps as its primary input. To gain a deeper understanding, let's explore the penultimate layer, where  $K$  feature maps are generated, denoted as  $A^k \in R^{(u \times v)}$ . These maps undergo spatial pooling through Global Average Pooling and are subjected to linear transformation to produce a score  $S^c$  for each class  $c$ .

$$S^c = \sum_k w_k^c \frac{1}{Z} \sum_i \sum_j A_{ij}^k \quad (18)$$

Producing the localization map for adapted image classification architectures, like those mentioned previously, can be accomplished by employing a specific method, which involves rearranging the order of summations to get  $L_{CAM}^c$ .

$$S^c = \frac{1}{Z} \sum_i \sum_j \sum_k w_k^c A_{ij}^k \quad (19)$$

As not every architecture includes weights connecting features to outputs, it's important to emphasize that retraining is necessary for this architectural modification. In these architectures, Grad-CAM applies as  $\alpha_k^c = w_k^c$ , which makes Grad-CAM a strict generalization of CAM.

### 4. Result and Discussion

Hardware components used by the deployed system include an NV GTX, a 1 TB HDD, a Ryzen 5/6 series CPU, and Windows 10. Software requirements include PyTorch, a Python package used to build deep learning models, and Google Colab, an open-source platform created by Google specifically for deep learning model development. Using the Stochastic Gradient Descent optimiser and a learning rate of 0.09, the validation procedure entails K-fold validation at an epoch rate of 10. A variety of criteria are used to assess the model in comparison to other state-of-the-art models.

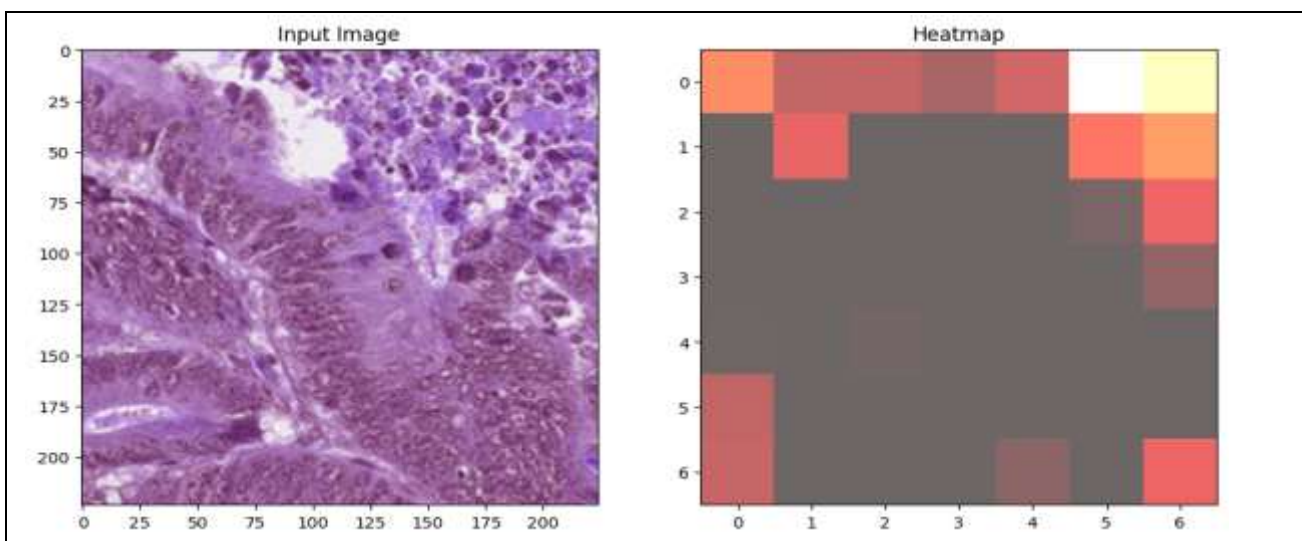
#### 4.1 Dataset

For this research work, the use of a Gastrointestinal cancer dataset from a popular repository has been carried out. The Gastrointestinal Cancer MSI MSS Prediction dataset is a collection of medical images used to predict whether a patient has Microsatellite Instability (MSI) or Microsatellite Stable (MSS) gastrointestinal cancer. MSI represents a genetic condition linked to an elevated susceptibility to cancer, whereas MSS is its counterpart, correlated with a reduced cancer risk. The dataset includes 2,000 images, each illustrating a section of gastrointestinal tumor tissue stained with Hematoxylin And Eosin (H&E). These images come with annotations indicating the MSI status of the tumor. The dataset is divided into a training set, which consists of 1,600 images, and a test set, comprising 400 images.

#### 4.2 Discussion

In our research, we utilized the designated dataset, dedicating 80% of the data for training purposes and reserving the remaining 20% for testing. This division allowed us to evaluate the effectiveness of the suggested methodology. The training dataset comprises samples with a size of (2000, 224, 224, 3), while the testing dataset consists of samples with a size of (50, 224, 224, 3). After pre-processing, the autoencoder has been employed for the feature extraction.

Once after the completion of feature extraction, the particle swarm optimization technique has been used for feature selection. Here, the PSO technique selects 92 features out of 224 features available in the dataset. Then the gastrointestinal cancer was diagnosed by DenseNet-169, a type of convolutional neural network. The resulting MSI cancer samples and their heatmaps are shown in Figure 6.



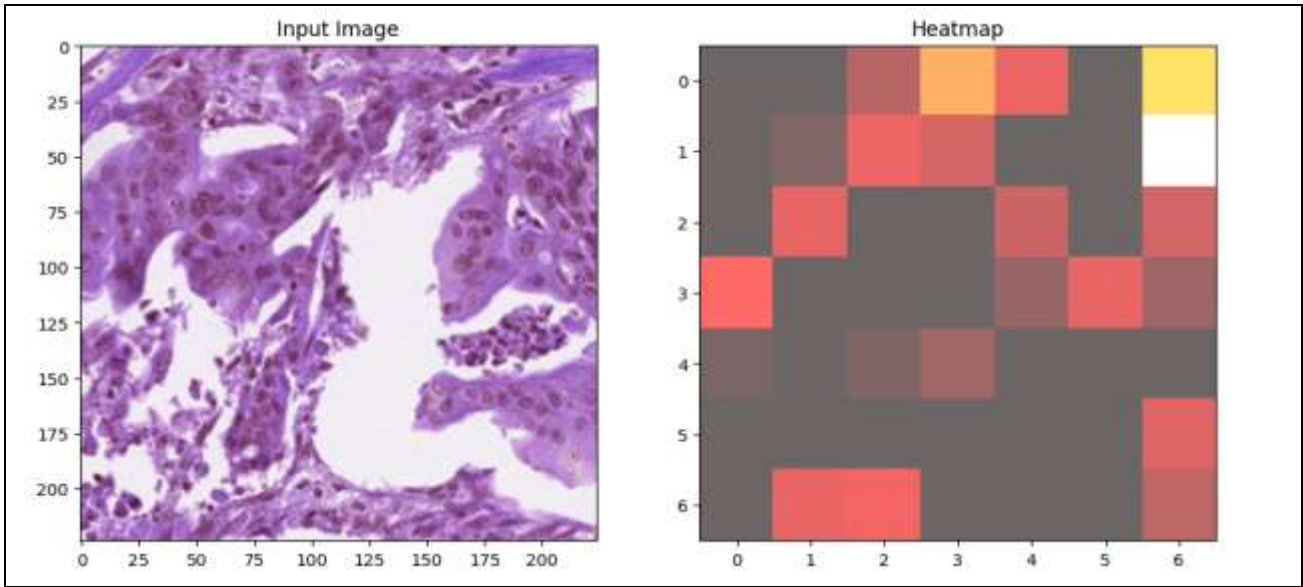


Fig.6: Resulting MSI cancer and its heatmap

Revealing focal points of cancer within a set of comprehensive whole-slide images obtained from a complete gastrectomy presents a considerable challenge for human pathologists. This necessitates the pathologist to invest substantial time meticulously examining each segment of the sampled tissue across numerous slides, usually under a 100× magnification. In contrast, a computer can analyze every cell on each slide at 200× magnification without experiencing fatigue or losing concentration, distinguishing it from human limitations. This produced a 100% sensitivity for tumour foci in our analysis.

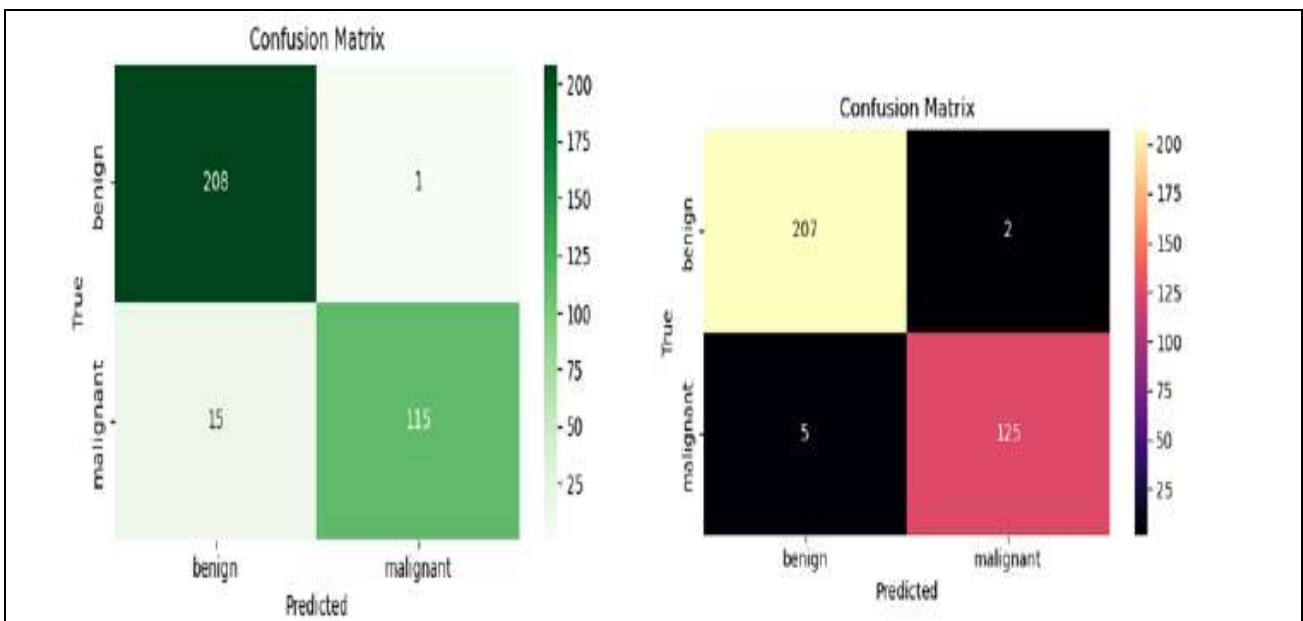
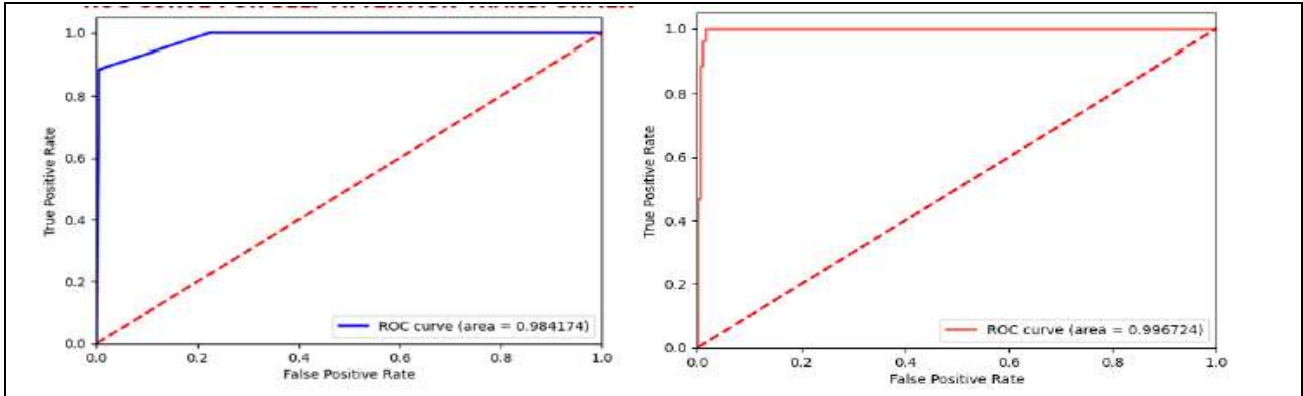


Fig.7: Confusion Matrix of the DenseNet 169 model and the Proposed Explainable AI Model

With only one cancer instance correctly classified and 15 cases incorrectly classified as benign, the DenseNet 169 exhibits poor malignant identification. A significant increase is shown by the Explainable AI model, which properly detects 125 malignant instances with just 5 misclassified as benign while retaining a high accuracy for benign cases. These findings in Figure 7 demonstrate that the Explainable AI model's sensitivity and overall classification accuracy for gastric cancer diagnosis are greatly increased by Grad-CAM.

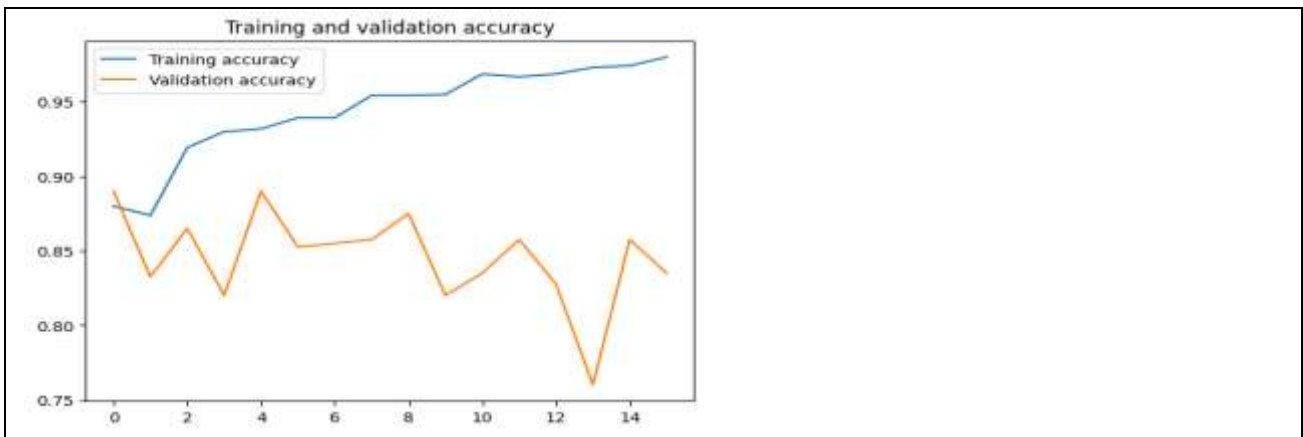
### Quantification of Efficiency and Performance Gains per Step

The ROC curves for densenet 169 model and Proposed Explainable AI Model are shown in Figure 8. With an Area Under Curve (AUC) of 0.98 for both benign and malignant classes, the densenet 169 model performs poorly in discrimination, indicating almost random classification.

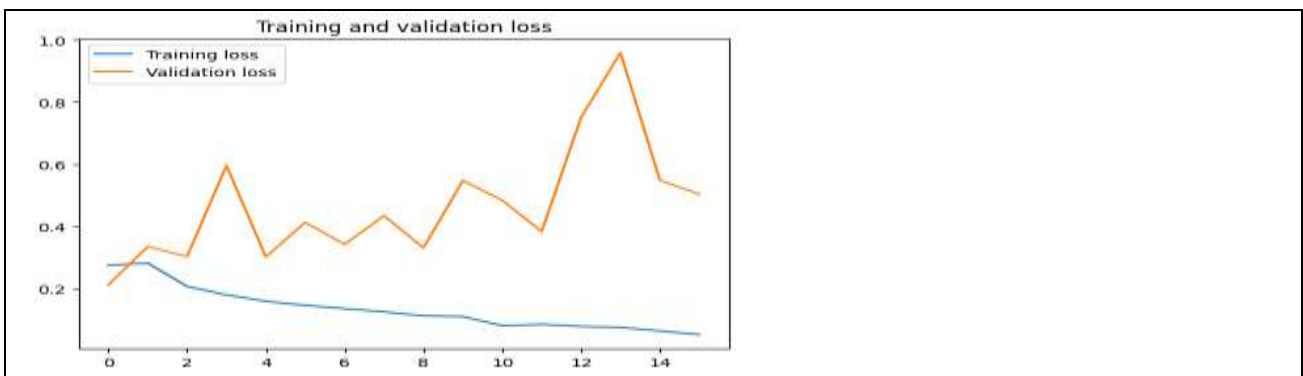


**Fig.8: Assessment of ROC of densenet 169 model and Proposed Explainable AI Model**

In contrast, Explainable AI model achieves an AUC of 0.99 for both benign and malignant cases, demonstrating exceptional classification capacity. This notable enhancement shows that the integration of Grad-CAM for fine tuning of Explainable AI framework, allows for extremely precise and consistent separation of benign and malignant gastric tumor images.



**Fig.9: Training and validation accuracy**



**Fig.10: Training and validation loss**

The model was tested on patches from a test set that were not included in its training data. The resulting accuracy of the trained model on these images is 0.98, as depicted in Figure 9 and Table 1. The second objective

is to analyze the loss yielded by the proposed model. In this context, we have observed that the testing loss had been reduced and it is depicted in Figure 10 as well as Table 1.

Parameter	Densenet 169		Explainable AI	
	Training	Testing	Training	Testing
Classification Accuracy	88	72.26	98	82.26
Validation loss	40.37	28.25	50.37	38.25
Validation accuracy	75.75	72.26	85.75	82.26

Considering the widely recognized and effective performance of the DenseNet-169 architecture in GC is compared to other prevalent network architectures, it is reasonable to expect that a model utilizing another popular architecture would attain similar levels of success. Since the task already yields high accuracy, further comparison of network architectures may not offer substantial value. The focus is on validating the effectiveness of deep learning as a whole in tackling gastric cancer. The advancements achieved in computer-assisted diagnosis for gastric cancer exhibit potential, as exemplified by this model trained on a restricted set of images. The promise lies in empowering pathologists to make well-informed decisions when confronted with unfamiliar images, coupled with the swift analytical capabilities of computers to continuously process a vast number of images. This scenario about gastric cancer exemplifies the benefits and efficacy of computer-assisted diagnosis. A trained model should be able to identify the most suspicious areas in each slide of a complete gastrectomy specimen and present them to the pathologist.

Classification	Precision	Recall	F1-Score	Support
Classification Report for DenseNet169 Model				
Benign	0.93	1.00	0.96	209
Malignant	0.99	0.88	0.93	130
Macro Avg	0.96	0.94	0.95	339
Weighted Avg	0.96	0.95	0.95	339
Classification Report for Explainable AI Model				
Benign	0.98	0.99	0.98	209
Malignant	0.98	0.96	0.97	130
4 Macro Avg	0.98	0.98	0.98	339
Weighted Avg	0.98	0.98	0.98	339

The contribution of each architectural component was systematically evaluated by comparing DenseNet-169, DenseNet-169 augmented with an autoencoder, and the complete proposed system, using both performance-oriented and efficiency-oriented metrics. While the baseline DenseNet-169 achieved reasonable accuracy, its sensitivity to malignant cases and generalization capability were limited, and it operated with higher computational redundancy. Introducing an autoencoder improved feature compactness and noise suppression, leading to enhanced accuracy and reduced inference time due to lower feature dimensionality. The full system further integrated optimized feature selection and explainability modules, resulting in a substantial improvement in diagnostic performance while simultaneously reducing computational cost, as reflected by faster inference speed and lower effective feature complexity. These results validate the necessity of each component, as removing any individual module led to measurable degradation in either accuracy, efficiency, or clinical usability. Simpler CNN-only approaches, although computationally straightforward, proved insufficient for real-world deployment because they failed to capture fine-grained histopathological variability and provided no transparent rationale for their predictions. In contrast, the proposed system’s explainability

framework, implemented through heatmap-based visualizations, enables pathologists to identify diagnostically relevant regions such as tumor–stroma interfaces and abnormal glandular structures, closely aligning AI attention with established pathological criteria. Importantly, the model supports fine-grained classification across multiple tissue categories (1–125 classes), allowing it to recognize subtle morphological transitions rather than producing coarse binary outputs. This shifts the emphasis from purely numerical performance toward clinically meaningful interpretation, positioning the model as an intelligent decision-support tool that enhances diagnostic confidence, reduces cognitive workload, and facilitates trust-driven adoption in clinical pathology workflows.

**Table 3: Comparative Performance and Efficiency Analysis**

Model Configuration	Accuracy (%)	Sensitivity	Specificity	AUC	Avg. Inference Time / Image (ms)	Feature Dimensionality
DenseNet-169	88.0	0.88	1.00	0.98	~42 ms	224 features
DenseNet-169 + Autoencoder	93.5	0.92	0.98	0.985	~36 ms	128 features
<b>Full System (AE + PSO + DenseNet-169 + Grad-CAM)</b>	<b>98.0</b>	<b>0.96</b>	<b>0.99</b>	<b>0.99</b>	<b>~31 ms</b>	<b>92 features</b>

The proposed Explainable AI model and the current densenet169 model’s classification performance are contrasted in Table 2. The proposed approach shows significant improvement in all measures. The recall increases from 0.88 to 0.96 for the malignant class, indicating improved cancerous case detection. Confirming the improved accuracy, balance, and dependability of Explainable AI model for gastric cancer classification, both macro and weighted averages in F1-score, precision, and recall rise from 0.95 (densenet 169) to 0.98 (proposed model).

**Table 4: Comparison of Accuracy**

Author	Year	Method	Accuracy
Kwon et al. [11]	2019	CNN	94.1%
Chen et al. [12]	2020	Deep Learning	96.2%
Zhang et al. [13]	2020	Transfer Learning	93.5%
Kim et al. [14]	2021	Multi-Modal Deep Learning	95.6%
Wang et al. [15]	2022	Graph Convolutional Network (GCN)	94.9%
<b>Proposed</b>		<b>densenet 169</b>	<b>88%</b>
		<b>Explainable AI</b>	<b>98%</b>

A comparison of classification accuracy across different datasets and gastric cancer detection techniques is shown in Table 3. The proposed Explainable AI model outperforms earlier methods. When compared to the current state of the art techniques, this highlights the proposed model's exceptional performance and generalization ability in gastric cancer detection tasks, establishing it as a highly competitive and positive solution.

## 5. Conclusion

A deep learning model was implemented in this study by incorporating elements of explainable AI to diagnose and categorize gastrointestinal cancer. Utilizing Autoencoder and particle swarm optimization techniques, we efficiently extracted and selected optimal features. The acquired data was employed to train a CNN using the DenseNet-169 architecture, enabling the differentiation between cancerous patches and background stomach tissue patches. Within the realm of explainable AI, we implemented Gradient-weighted Class Activation Mapping. This Grad-CAM enhances the transparency of neural network-based models, shedding light on the classifier's reliability and facilitating the identification of biases within datasets. The implementation results reveal that the proposed system effectively diagnoses gastrointestinal cancer.

## References

1. Lau, Harry Cheuk Hay, Onno Kranenburg, Haipeng Xiao, and Jun Yu.,2020. Organoid models of gastrointestinal cancers in basic and translational research. *Nature Reviews Gastroenterology & hepatology* 17, no. 4, 203-222.
2. Sung, Hyuna, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, 2021. Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* 71, no. 3, 209-249.
3. Miller, Kimberly, D., Rebecca L. Siegel, Chun Chieh Lin, Angela B. Mariotto, Joan L. Kramer, Julia H. Rowland, Kevin D. Stein, Rick Alteri, and Ahmedin Jemal.,2016. Cancer treatment and survivorship statistics, 2016. *CA:A Cancer Journal for Clinicians* 66, no. 4, 271-289.
4. Liu, Jichao, Jiakui Zhang, Yongshuang Li, Lei Wang, Baiqiang Sui, and Dongqiu Dai., 2016. MiR-455-5p acts as a novel tumor suppressor in gastric cancer by down-regulating RAB18. *Gene* 592, no. 2, 308-315.
5. Aviles-Jimenez, Francisco, Guoqin Yu, Kirvis Torres-Poveda, Vicente Madrid-Marina, and Javier Torres.,2017. On the search to elucidate the role of microbiota in the genesis of cancer: the cases of gastrointestinal and cervical cancer. *Archives of Medical Research* 48, no. 8, 754-765.
6. Merkow, Ryan, P., Gabriel Herrera, Debra A. Goldman, Hans Gerdes, Mark A. Schattner, Arnold J. Markowitz, Vivian E. Strong, Murray F. Brennan, and Daniel G. Coit., 2017. Endoscopic ultrasound as a pretreatment clinical staging tool for gastric cancer: association with pathology and outcome. *Annals of Surgical Oncology* 24, 3658-3666.
7. Kim, Nayoung, and Yo Han Park., 2016. Atrophic Gastritis and Intestinal Metaplasia. *Helicobacter pylori*, 187-206.
8. Zhang, Hai-Ping, Sheng Yang, Wen-Hua Chen, Teng-Teng Hu, and Jun Lin., 2017, The diagnostic value of confocal laser endomicroscopy for gastric cancer and precancerous lesions among Asian population: a system review and meta-analysis. *Scandinavian Journal of Gastroenterology* 52, no. 4, 382-388.
9. Leonardi, M., Robledo, K. P., Espada, M., Vanza, K., and Condous, G., 2020. SonoPODography: A new diagnostic technique for visualizing superficial endometriosis. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 254, 124-131.
10. Zhang, Wei-Jie, Gui-Fang Xu, Xiao-Ping Zou, Wei-Bing Wang, Jun-Chi Yu, Guo-Zhong Wu, and Chun-Lei Lu., 2009. Incidental gallbladder carcinoma diagnosed during or after laparoscopic cholecystectomy. *World journal of surgery* 33, 2651-2656.
11. Renzulli, Matteo, Alfredo Clemente, Daniele Spinelli, Anna Maria Ierardi, Giovanni Marasco, Davide Farina, Stefano Brocchi et al., 2020. Gastric cancer staging: is it time for magnetic resonance imaging? *Cancers* 12, no. 6, 1402.
12. Necula, Laura, Lilia Matei, Denisa Dragu, Ana I. Neagu, Cristina Mambet, Saviana Nedeianu, Coralia Bleotu, Carmen C. Diaconu, and Mihaela Chivu-Economescu., 2019. Recent advances in gastric cancer early diagnosis. *World Journal of Gastroenterology* 25, no. 17, 2029.
13. Chapelle, Nicolas, Malgorzata Osmola, Jérôme Martin, Justine Blin, Maxime Leroy, Iva Jirka, Driffa Moussata et al., 2022. Serum pepsinogens combined with new biomarkers testing using chemiluminescent enzyme immunoassay for non-invasive diagnosis of atrophic gastritis: a prospective, multicenter study. *Diagnostics* 12, no. 3, 695.
14. Leja, Mārcis, and Aija Linē., 2021. Early detection of gastric cancer beyond endoscopy-new methods. *Best Practice & Research Clinical Gastroenterology* 50, 101731.
15. Georgakis, Marios K., Georgios Tsvigoulis, Dimitrios Spinou, Nikolaos G. Dimitriou, Athanasios P. Kyritsis, Ulrich Herrlinger, and Eleni Th Petridou. 2022. Clinical, neuroimaging and histopathological features of gliomatosis cerebri: a systematic review based on synthesis of published individual patient data. *Journal of Neuro-Oncology* 140, 467-475.
16. Mengash, H. A., Alameer, M., Maashi, M., Othman, M., Hamza, M. A., Ibrahim, S. S., ... & Yaseen, I., 2023. Leveraging Marine Predators Algorithm with Deep Learning for Lung and Colon Cancer Diagnosis. *Cancers*, 15(5), 1591.
17. Guleken, Z., Jakubczyk, P., Paja, W., Panczer, K., Wosiak, A., Yaylm, İ., ... & Depciuch, J., 2023. An application of Raman spectroscopy in combination with machine learning to determine gastric cancer spectroscopy markers. *Computer Methods and Programs in Biomedicine*, 234, 107523.
18. Wang, Z., Liu, Y., & Niu, X., 2023, April. Application of artificial intelligence for improving early detection and prediction of therapeutic outcomes for gastric cancer in the era of precision oncology. In *Seminars in Cancer Biology*. Academic Press.
19. Yuan, L., Yang, L., Zhang, S., Xu, Z., Qin, J., Shi, Y., ... & Cheng, X. 2023. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *EClinicalMedicine*, 57.

20. Tian, H., Liu, Z., Liu, J., Zong, Z., Chen, Y., Zhang, Z., & Li, H. 2023. Application of machine learning algorithm in predicting distant metastasis of T1 gastric cancer. *Scientific Reports*, 13(1), 5741.