Journal of Artificial Intelligence and Technology, (Ahead of Print)
https://doi.org/10.37965/jait.2026.1779

ISTP
RESEARCH ARTICLE

Improved AlexNet–Bi-LSTM Model-Based Gastrointestinal Cancer Classification with Shape and Texture Feature Set

Vanishree Abhay,[1,2] N. Achyutha Prasad,[1,2] N. Sathish Babu,[3] and Rekha Prasad[2,4]
[1]Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bengaluru, India
[2]Department of Computer Science and Engineering, Visvesvaraya Technological University, Belagavi, India
[3]Department of General Surgery, Sri Siddhartha Academy of Higher Education, Tumakuru, India
[4]Department of Computer Science and Engineering, K S Institute of Technology, Bengaluru, India

(Received 31 July 2026; Revised 01 October 2026; Accepted 17 October 2026; Published online 11 November 2026)

***Abstract*: *Background*:** Gastrointestinal (GI) cancer classification is the process of identifying and categorizing cancers that occur in the GI tract based on their morphological, histopathological, and molecular characteristics. Existing GI cancer classification systems often struggle with low interpretability and limited generalizability across diverse datasets. This study aims to improve the accuracy of classifying cancer.
***Methods*:** The Improved AlexNet–Bi-LSTM model-based Gastrointestinal Cancer Classification (IAB-GCC) with a shape and texture feature set is proposed to improve diagnostic accuracy. The input images are from a dataset consisting of 411,890 unique image patches generated from formalin-fixed paraffin-embedded (FFPE) diagnostic slides. Then, the proposed IAB-GCC process begins with input images, which are preprocessed using an Adaptive Weighted Gaussian Filtering (AWGF) technique to reduce noise and enhance image quality. Key features including Weighted Shannon Entropy (WSE)-based features, texture features using Gray Level Co-occurrence Matrix (GLCM), shape, and morphological features are then extracted and fed into the Alex–-Bi-LSTM network, where the extracted features are passed into an Improved AlexNet (IANet) model to train the features, and then the learned feature maps of IANet are subjected to a Bi-directional Long Short-Term Memory (Bi-LSTM) model for classification.
***Results*:** The combination of IANet and Bi-LSTM enables the system to classify cancer cases as stable or unstable, which provides precise and reliable diagnostic results. The proposed IAB-GCC model achieves a higher precision 97% and an accuracy score of 96% in the classification process of GI cancers.
***Conclusion*:** The proposed hybrid IAB-GCC provides accurate, reliable, and interpretable GI cancer classification, demonstrating potential for clinical diagnostic applications.

***Keywords*:** cancer classification; gastrointestinal cancer; Improved AlexNet; shape feature; texture feature

I. INTRODUCTION

In recent decades, our understanding of how cancer originates and develops has grown significantly [1]. In recent years, researchers have conducted extensive studies on the link between gut microbiota and gastrointestinal (GI) cancers that uncover the key role that interactions between host metabolism and microbiota play in tumor development [2]. GI cancers are responsible for over 3 million deaths globally each year [3]. GI cancer is a type of primary tumor that occurs in the digestive tract [4]. It affects both men and women equally, and it is the second most prevalent type of cancer among individuals aged 15 to 44 years in India. One of the primary factors contributing to the severity of this cancer type is that it often goes undetected or cannot be effectively treated in its early stages [5]. The GI tract is a vital and complex system responsible for digestion and nutrient absorption and is essential for maintaining human health [6]. GI cancers encompass tumors of the colon, stomach, rectum, esophagus, pancreas, anus, liver, gallbladder, and bile duct, which rank among the most prevalent cancers and are a leading cause of cancer-related deaths worldwide [7].

High-risk individuals are those with a history of long-term GI diseases or with immediate family members who have had GI tumors [8]. GI tract malignancies include cancers of the stomach, esophagus, colon, small intestine, and anus that are a major cause of illness and death with more than 200,000 new cases and around 80,000 deaths annually in the United States [9]. GI cancer represents more than a quarter of all cancer cases globally and is responsible for over one-third of all cancer-related deaths each year [10]. These cancers can lead to conditions such as esophagitis, polyps, and ulcerative colitis, where each of these differs significantly in diagnosis and treatment [11]. In GI, the microbiota has been shown to influence the effectiveness of radiotherapy, chemotherapy, and immunotherapy, which suggests that the gut microbiota represents a promising target for enhancing antitumor treatments [12]. However, because GI endoscopic procedures like colonoscopy and gastroscopy are not widely accessible, most patients are diagnosed at an advanced stage, which leads to poor post-operative survival outcomes [13]. In recent decades, medical imaging technologies have advanced rapidly, greatly improving both the speed and resolution of imaging devices [14].

Tumor imaging data play a crucial role in uncovering the spatial structure, morphology, tissue composition, and internal organization of tumors, which provides essential information for

Corresponding author: Achyutha Prasad N (e-mail: achyuth001@gmail.com).

© The Author(s) 2026. This is an open access article published under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

cancer diagnosis [15]. In recent years, the application of deep learning to histopathology images has achieved high performance in traditional tasks [16]. Accurately predicting life expectancy is a critical task for clinical decision guidance and improvement for patient care and quality of life [17]. Existing convolutional neural network (CNN)-LSTM and attention-based GI cancer classification models primarily rely on deep feature representations, which often suffer from information loss due to noise, inadequate preservation of subtle tumor boundaries, and insufficient utilization of complementary shape and texture characteristics. Moreover, many attention-based methods focus on salient regions but do not explicitly incorporate spatially informative handcrafted features relevant to GI lesions. To address these limitations, a novel Improved AlexNet–Bi-LSTM model-based Gastrointestinal Cancer Classification (IAB-GCC) with a shape and texture feature set is proposed for the precise diagnosis of cancer. The primary contributions presented in this work include:

- An adaptive weighted Gaussian filter (GF) is proposed to effectively reduce noise and preserve subtle tumor edge information, thereby enhancing lesion representation for subsequent feature extraction and classification.
- A Weighted Shannon Entropy (WSE)-based feature extraction strategy is introduced to emphasize spatially significant regions and capture discriminative information from GI lesions.
- An Improved AlexNet (IANet) architecture is developed to preserve complementary shape and texture features while reducing feature redundancy, enabling more effective lesion representation than conventional CNN-based models.
- Bi-LSTM is used to capture long-range dependencies in CNN-extracted feature sequences, enabling better modeling of global contextual relationships in histopathology images for improved classification. Therefore, the proposed AlexNet–Bi-LSTM improve GI cancer classification performance and diagnostic accuracy.

The remainder of the paper is structured as follows: Section II discusses related research on GI cancer classification. Section III details the proposed IAB-GCC approach. Section IV assesses the method's performance through evaluations with present techniques. Finally, Section V concludes the study.

II. LITERATURE REVIEW

In 2024, Ahmed S. Almasoud *et al.* [18] introduced a GICDC-AVOADL, where the primary objective of this model was to analyze GI images for accurate cancer identification. To accomplish this, the GICDC-AVOADL framework employed an enhanced EfficientNet-B5 model to extract features from input images, while AVOA has utilized to optimize the hyperparameters of the improved EfficientNet-B5. Additionally, a DCAE was applied for GI classification and the detection of cancer. Comprehensive simulations were carried out to assess the model's performance, and the results demonstrated that GICDC-AVOADL outperformed existing methods in GI cancer detection.

In 2023, Olfat M. Mirza *et al.* [19] developed a novel GDDC-HRODL framework to classify medical images for GC. To achieve this, the method begins with preprocessing to enhance image quality. Next, the HybridNet model is applied to generate feature vectors, while hyperparameter optimization is performed using the HRO algorithm. These hyperparameter tuning strategies collectively contribute to improved GC classification performance. Experimental evaluations on medical datasets confirm that the GDDC-HRODL model achieves superior results in GC classification compared to existing methods.

In 2024, Fahdah A Almarshad *et al.* [20] proposed a SOADL-GCC method to analyze GI tract images for cancer detection and classification. In this approach, BF was first applied to remove noise and enhance image quality. Feature extraction was then carried out using a CapsNet with the SOA employed to fine-tune the CapsNet hyperparameters. The final classification was performed using a DBN. The SOADL-GCC framework was evaluated on the Kvasir dataset, where experimental simulations demonstrated superior detection performance.

In 2024, Abeer A. Wafa *et al.* [21] have proposed a model that is structured into four stages: data acquisition, image processing, feature extraction, and classification. For data acquisition, histopathology images were obtained from TCGA and the Kaggle repository, with certain modifications applied. In the image processing stage, operations such as color transformation, normalization, resizing, and labeling were performed to prepare the images for DL models. This methodology demonstrated substantial improvements in GI cancer detection compared to existing techniques, which underscores the promise of DL-based approaches for histopathology analysis. These findings were expected to encourage further research in DL-driven GI cancer detection.

In 2023, Yanmiao Dai *et al.* [22] have identified nine candidate DNA methylation markers and assessed them in plasma and tissue cohorts to determine the most effective markers. A training cohort and a validation cohort, which includes patients with GCs, polyps, benign tumors, and healthy controls, were used to develop and validate a DNA methylation panel. An independent prospective cohort was further evaluated to assess the panel's performance and compared it with conventional blood protein tumor markers.

In 2025, Zhaojun Ye *et al.* [23] proposed an RAA-VLSTM model to analyze comprehensive clinical data. The dataset consists of EHRs from multiple healthcare centers, which covers patient demographics, treatment outcomes, clinical history, and genetic information. Data preprocessing involved normalization, tokenization, and stop-word removal to effectively represent textual data. For feature extraction, advanced word embeddings were applied to improve model performance. This study provided a solid foundation for applying advanced NLP and machine learning methods in GI cancer management.

In 2025, Chien-Jung Huang *et al.* [24] have obtained gene expression profiles of seven GI cancer types from TCGA and analyzed their TME, which classified samples into hot and cold tumor subgroups. They then developed a computational framework to build cancer-specific hot tumor classification models using only a small number of genes. The performance of these few-gene models was validated using external independent datasets and qPCR experiments.

In 2024, Narges Yousefpour *et al.* [25] have developed an intraoperative tumor detection probe and calibrated it using human models of GI cancers, which allows real-time scanning of excised masses. The EGCD system has measured the impedimetric properties of GI masses through a handpiece probe and a needle-based head probe. Key classification parameters, which included IPS and Z1kHz, were extracted. EGCD performance was evaluated on palpable GI masses and compared with histopathology results.

In 2025, Muhammad Zubair *et al.* [26] have suggested a multi-channel attention mechanism-based framework that was capable of dynamically focusing on relevant features and capturing complex relationships in medical data. The framework integrated three distinct attention mechanism channels with CNNs to extract

multi-channel features during the classification process. This approach provided a practical AI solution for overcoming challenges posed by complex medical image characteristics, which demonstrated substantial potential for improving diagnostic accuracy and enhancing treatment outcomes in gastric cancer.

In 2024, Auzine *et al.* [27] have developed an ensemble model consisting of three CNNs: InceptionV3, InceptionResNetV2, and VGG16, which was based on averaging techniques and by combining their respective predictions. These models were trained on the Kvasir dataset, which consists of pathological findings related to GI cancer. The results obtained from this analysis demonstrated a positive and encouraging advancement in the exploration of XAI approaches, specifically in the context of GI cancer detection within the healthcare domain.

In 2025, Khan *et al.* [28] have suggested deep learning framework for classifying and localizing GI diseases from wireless capsule endoscopy (WCE) images. The framework begins with dataset augmentation to enhance training robustness. Bayesian Optimization (BO) is employed for dynamic hyperparameter tuning, and an Entropy-controlled Marine Predators Algorithm (EMPA) selects optimal features. The suggested framework offers improved accuracy, precision, and computational efficiency compared to state-of-the-art models.

Table I presents a summary of different approaches for Gastrointestinal Cancer Classification, which outlines their main characteristics, limitations, and potential directions for future improvement based on multiple studies.

A. RESEARCH GAPS

Existing GI cancer classification approaches face several limitations in achieving robust clinical performance. Many methods show reduced effectiveness in real-time endoscopic analysis and lack adaptability for deployment in practical healthcare environments [18]. Optimization-based models such as ALSTM and ALO exhibit lower performance compared to ensemble learning techniques, indicating limited convergence efficiency [19]. Several studies require large-scale validation across asymptomatic and diverse populations to ensure generalizability [22]. Fixed feature sets and limited gene-based representations reduce cross-subtype applicability [24], while attention mechanisms suffer from noise sensitivity and interpretability issues [26]. Although hybrid CNN-LSTM architectures attempt to capture contextual relationships, they often overlook feature redundancy and fail to explicitly prioritize spatially significant shape and texture information. Existing CNN-LSTM models also fail to effectively handle feature redundancy and weak boundary preservation, necessitating improved adaptive and robust learning frameworks. Moreover, existing automated image-based diagnostic systems are not directly compatible with non-image modalities such as intraoperative probes, NLP-based clinical data, or DNA methylation profiles. To address these limitations, there is a need for a robust GI cancer classification framework that incorporates adaptive edge-preserving preprocessing, discriminative shape and texture feature extraction, and bidirectional contextual learning. The proposed approach aims to overcome these challenges by

Table I. Characteristics and limitations of conventional methods

Authors [citation]	Methodology	Characteristics	Limitations
Ahmed S. Almasoud *et al.* [18]	GICDC-AVOADL	Applied AVOA-driven tuning to optimize hyperparameters effectively.	Weak tumor edges may not be preserved; limited spatial feature prioritization
Olfat M. Mirza *et al.* [19]	ALSTM and ALO	Utilized the ALO algorithm to fine-tune the ALSTM model, enhancing its accuracy.	Focuses on temporal dependencies but ignores fine-grained texture/shape features; small dataset evaluation
Fahdah A Almarshad *et al.* [20]	SOADL-GCC	Employed BF-based preprocessing to enhance the quality of images.	Fixed smoothing may suppress weak edges; deep features only; limited generalizability.
Abeer A. Wafa *et al.* [21]	DL	Highlighted the effectiveness of recurrent layers in modeling temporal relationships in histopathology data.	Ignores spatially significant handcrafted features.
Yanmiao Dai *et al.* [22]	Plasma-based DNA methylation panel	Developed a targeted multi-cancer test capable of detecting several cancers in the same organ system at once.	Requires large-scale testing among asymptomatic and average-risk individuals to verify real-world performance.
Zhaojun Ye *et al.* [23]	NLP-driven predictive modeling	Achieved better performance on long sequences relative to conventional LSTM models.	Not suitable for image-based diagnosis.
Chien-Jung Huang *et al.* [24]	PCA	The few-gene models accurately classified tumors into hot and cold subgroups.	Limited applicability across different tumor types
Narges Yousefpour *et al.* [25]	EGCD	Integrated with intraoperative frozen section analysis for immediate detection during surgery.	Requires specialized setup; not automated image-based classification
Muhammad Zubair *et al.* [26]	Multi-channel attention mechanism	Multi-channel attention mechanism enhances feature extraction and captures complex patterns in GC histopathology images.	It struggles with adaptability, interpretability, and noise sensitivity.
Auzine *et al.* [27]	Ensemble Model with XAI	Developed an ensemble CNN with explainable AI for gastrointestinal cancer classification, improving accuracy and interpretability.	May face limitations in scalability and robustness across diverse datasets.
Khan *et al.* [28]	SC-DSAN	A fused deep learning architecture with a shallow neural network classifier for gastrointestinal cancer classification.	Limited generalizability and requires further validation on diverse datasets.

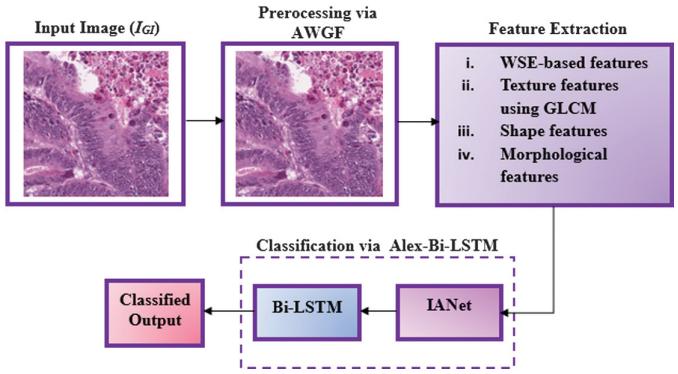


Fig 1. Process of IAB-GCC.

improving classification accuracy, robustness, and generalization capability across diverse datasets and clinical conditions.

III. PROPOSED METHODOLOGY

GI cancer classification is the process of identifying and categorizing cancers that occur in the GI tract based on their morphological, histopathological, and molecular characteristics. This classification helps to distinguish between different cancer stages to support precise diagnosis for personalized treatment planning. Existing GI cancer classification systems often struggle with sensitivity to noise, low interpretability and limited generalizability across diverse datasets. So, this paper proposes an IAB-GCC to improve diagnostic accuracy and facilitate timely treatment decisions.

Figure 1 shows the workflow of the proposed IAB-GCC. The IAB-GCC (as shown in Fig. 1) process begins with input images, which are first preprocessed to enhance quality and reduce noise using an Adaptive Weighted Gaussian Filtering (AWGF) technique. This step ensures that subsequent analysis focuses on meaningful patterns within the histopathological images, which improves the accuracy of feature extraction.

Following preprocessing, key features are extracted, such as WSE-based features, texture features using Gray Level Co-occurrence Matrix (GLCM), shape features, and morphological features. These features are then fed into an AlexNet–Bi-LSTM classification model, which involves an IANet and Bi-LSTM networks. These models influence both spatial and temporal patterns for robust classification. The final output distinguishes between stable and unstable GI cancer cases.

A. PREPROCESSING

Preprocessing is the first stage in IAB-GCC that involves the transformation of images (I_{GI}) (I_{GI}) to enhance the performance of downstream processing. In the IAB-GCC, preprocessing involves applying a GF to the input image (I_{GI}) to produce an enhanced output. A GF [29] is an image processing method that smooths images and reduces noise by applying a Gaussian function. The conventional GF for the kernel value at a pixel is given in Eq. (1):

$$G_{x,y} = \frac{1}{2\pi\delta^2} exp\left[\frac{-x^2+y^2}{2\delta^2}\right] \tag{1}$$

where x and y are the rows and columns from the center, respectively.

However, this conventional GF can blur image edges since it computes a weighted average of all pixels in the kernel neighborhood, which may cause a loss of fine details and result in blurry images. So, this paper proposes an AWGF technique to preserve important edges and fine details better. AWGF enhances histopathological image quality by reducing noise while preserving important tissue structures for accurate analysis. The AWGF is given by the formulation in Eq. (2):

$$G_{x,y} = \frac{1}{2\pi\delta^2} exp\left[\frac{-x^2+y^2}{2\delta^2}\right] * w_{x,y} \tag{2}$$

where $w_{x,y}$ is the weight of x and y that is given in Eq. (3):

$$w_{x,y} = g(x,y,\delta) \times L_n[\delta_n(x,y)] \tag{3}$$

where $g(x,y,\delta) = exp[\frac{x^2+y^2}{2\delta^2}]$, that is, the spatial Gaussian weight and $L_n[\delta_n(x,y)] = \frac{1}{1+\beta.\delta_n(x,y)}$, that is, the local noise adaptive weight. Here, $\delta_n(x,y)$ is the estimated noise level at (x,y) and β controls the strength of noise influence.

Finally, the enhanced image $(I_{GI}(G))$ through AWGF preprocessing technique is obtained.

1). FEATURE EXTRACTION. Feature extraction is the method of detecting and isolating significant patterns from preprocessed images $(I_{GI}(G))$ that are essential for classification. In IAB-GCC, this involves the extraction of features such as WSE-based features, texture features using GLCM, shape features, and morphological features.

2). WSE-BASED FEATURES. Shannon Entropy (SE) [30] measures the randomness of pixel concentration values that indicate how much information is present in the image $(I_{GI}(G))$. The conventional SE is given by the formulation in Eq. (4):

$$S(E^n) = -\sum_{e\in E} b(e^n) log_2 b(e^n) \tag{4}$$

where E^n is the random variable and $b(e^n)$ denotes the dual feature maps of n blocks. This conventional SE does not consider the spatial distribution of pixels. So, this study introduces a WSE that prioritizes spatially important areas and treats all pixels equally. The WSE-based features are used as spatially weighted descriptors and are later concatenated with other handcrafted and deep features to enhance the discriminative representation of GI images. The WSE is given by the formulation in Eq. (5) and Eq. (6):

$$S_w(E^n) = -\sum_{i\in E} b_w(i)\, log_2 b_w(i) \tag{5}$$

$$b_w(i) = \frac{w(i)b(i)}{\sum_j w(j)b(j)} \tag{6}$$

where $w(i)$ is the weight based on tissue relevance, $b(i)$ is the image intensity histogram with bin feature maps, and $i \in E$ represents a bin of intensity values from a finite set E.

The weight map $w(p,q)$ representing tissue relevance is calculated by the following steps:

1. Convert RGB image into stain space (i.e., Hematoxylin-Eosin-DAB) denotes the Hematoxylin Channel intensity at pixel (p,q) as in Eq. (7):
$$I^{he}(p,q) \tag{7}$$
2. Normalize the Hematoxylin Channel to the range that is given in Eq. (8):
$$I^{he}_{norm}(p,q) = \frac{I^{he}(p,q) - min(I^{he})}{max(I^{he}) - min(I^{he})} \tag{8}$$

Using the above steps (1) and (2), the weight map is given in Eq. (9):

$$w(p,q) = 1 - I^{he}_{norm}(p,q) \tag{9}$$

where $w(p,q)$ reflects the relative tissue importance at each pixel based on nuclei staining.

The WSE-based features extracted are denoted as S_w.

3). TEXTURE FEATURE EXTRACTION USING GLCM. GLCM [31] is a texture analysis method that measures how often pairs of pixel intensity values occur at a specific spatial relationship in an image that captures patterns like dissimilarity, contrast, energy, correlation, and homogeneity. The GLCM-based texture features are extracted as second-order statistical descriptors and are integrated with WSE, shape, morphological, and deep CNN features to improve texture representation in the proposed framework.

1. Dissimilarity
   Dissimilarity (T_{dis}) refers to a measure of how different or varied pixel intensities or features are within an image. It is given in Eq. (10):
$$T_{dis} = \sum_{xy} |x - y| q(x,y) \tag{10}$$
2. Contrast
   Contrast (T_C) refers to the measure of the difference in intensity between pixels in an image. It indicates how much the light and dark areas of the image vary, which helps to highlight textures. It is given by the formulation in Eq. (11):
$$T_C = \sum_{x=1}^{c} \sum_{y=1}^{c} (x - y)^2 q_{xy} \tag{11}$$
3. Energy
   Energy (T_E) measures the repetition of pixel intensity pairs in an image. This is given in Eq. (12):
$$T_E = \sum_{x,y} (q_{x,y})^2 \tag{12}$$
4. Correlation
   Correlation (T_{cor}) measures the degree of linear relationship between the gray levels of neighboring pixels in an image. This is given in Eq. (13):
$$T_{cor} = \frac{\sum (r_x - \bar{y}_x)(s_y - \bar{s}_y)}{\sqrt{\sum (r_x - \bar{y}_x)^2} \sqrt{\sum (s_y - \bar{s}_y)^2}} \tag{13}$$
5. Homogeneity
   Homogeneity (T_h) measures the closeness of the distribution of elements to its diagonal. It is given in Eq. (14):
$$T_h = \sum_{x,y} \frac{q_{x,y}}{1 + |x - y|} \tag{14}$$

The texture features extracted using GLCM are combined and denoted as T_F.

Table II. GLCM parameters for texture feature computation

Parameter	Values
Direction	0,45,90,135 (in degrees)
Distance	1,2,3
Quantization	16 gray levels

For texture feature extraction, we utilized the GLCM. Table II shows the key parameters for GLCM computation as follows:

4). SHAPE FEATURES. Shape features [32] capture the ordered properties and physical characteristics of objects within an image $I_{GI}(G)$. They represent attributes such as area, perimeter, and length.

1. Area
   Area (S_A) refers to the total number of pixels that make up a particular region of interest within an image, which is given in Eq. (15):
$$S_A = \sum (r,s) \tag{15}$$
   where (r,s) represent the points located inside the region.
2. Perimeter
   Perimeter (S_P) is the total length of the edge pixels that outline an object of interest in an image, given in Eq. (16):
$$S_P = A + 2B \tag{16}$$
   where A denotes the count of even values and B denotes the count of odd values.
3. Length
   Length is the longest dimension of an object that is typically measured as the maximum distance between two boundary points of the object.
   The shape features extracted, such as area, perimeter, and length, are combined and denoted as S_F.

The extracted shape features are combined with texture (GLCM), WSE, morphological, and deep learning features to provide complementary structural information for improved classification.

5). MORPHOLOGICAL FEATURES. Morphological features [31] capture the structure, shape, and form of objects within an image $I_{GI}(G)$. They include attributes such as eccentricity, solidity, and compactness that provide information about the spatial organization of the objects.

1. Eccentricity
   Eccentricity (M_E) describes how stretched or elongated an object is, and it represents 0 being a perfect circle, and values closer to 1 are more elongated. It is given by the formulation in Eq. (17):
$$M_E = \frac{\sqrt{(r^2 - s^2)}}{r} \tag{17}$$
   where r is the major axis and s is the minor axis.
2. Solidity
   Solidity shows how solid an object stays. A value near 1 means the object is very dense. It is given in Eq. (18):
$$M_S = \frac{O_{area}}{convex_{H_{area}}} \tag{18}$$
   where O_{area} is the object area and $convex_{H_{area}}$ represents the convex hull area.
3. Compactness
   Compactness measures how closely an object's shape resembles a perfect circle. It reflects the "tightness" or "roundness" of a region in a preprocessed image. It is given in Eq. (19):
$$M_C = \frac{4\pi O_{area}}{P^2} \tag{19}$$
   where P is the perimeter of the object.

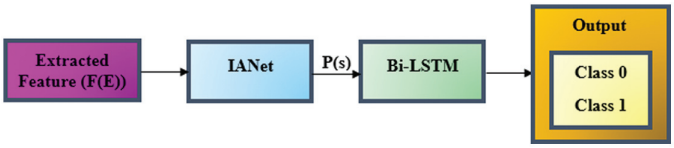


Fig 2. Process of classification.

The morphological features extracted using the eccentricity, solidity, and compactness attributes are combined and denoted as M_F.

Finally, WSE-based features, texture features using GLCM, shape features, and morphological features are combined and represented as $F(E) = [S_w, T_F, S_F, M_F]$.

The morphological features are fused with WSE, GLCM, shape features, and CNN-based deep features to construct a comprehensive feature representation for GI cancer classification.

6). CLASSIFICATION. The classification stage begins with the input of extracted features $(F(E))$, which are derived from the feature extraction phases. These features are fed into the AlexNet–Bi-LSTM classification model that encompasses IANet and Bi-LSTM. The IANet is a deep CNN designed to perform hierarchical feature learning. IANet processes the input $F(E)$ and generates the learned feature map scores, which effectively capture high-level representations of the GI tissue characteristics. Figure 2 shows the process of classification.

From Fig. 3, input is processed through IANet for feature extraction, and the resulting feature maps from Dropout-2 are reshVE aped and passed to Bi-LSTM for sequential learning. Finally, the learned representations are regularized using Dropout (0.5), followed by a fully connected (FC) layer and sigmoid activation for final classification output.

Advantages of Proposed Hybrid Model

- Multi-feature integration: Combines entropy, GLCM texture, shape, and morphological descriptors with deep CNN embeddings, ensuring richer and more discriminative representation of GI cancer tissue.
- Adaptive channel recalibration (ISE-F block): Prevents suppression of important handcrafted features, highlighting subtle lesion cues that conventional CNNs may overlook.
- Multi-scale feature extraction (Improved Inception): Captures both fine-grained local irregularities and broader structural distortions in GI images.
- Sequential dependency modeling (Bi-LSTM): Learns long-range relationships across heterogeneous feature dimensions, improving classification stability and robustness.
- Regularization and generalization: Dropout and batch normalization reduce overfitting, enabling the model to generalize across diverse patient datasets.

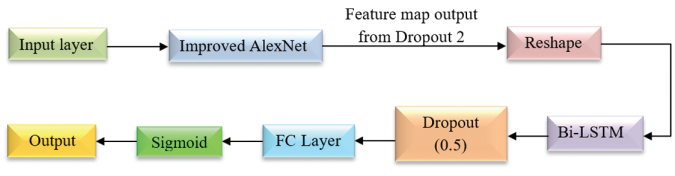


Fig 3. Hybrid model of improved AlexNet–Bi-LSTM.

- Enhances detection of early-stage GI cancer where subtle statistical and morphological cues are critical.

This shows the proposed hybrid model achieves superior GI cancer detection by fusing entropy, texture, shape, and morphological descriptors with deep CNN features, enhanced through ISE-Inception modules and Bi-LSTM sequence modeling. This design improves sensitivity to subtle lesion cues, strengthens boundary recognition, and delivers robust classification performance compared to conventional architectures.

The learned feature maps scores from IANet are then transferred to a Bi-directional Long Short-Term Memory (Bi-LSTM) network, which is capable of learning sequential dependencies and contextual relationships within the feature maps. The Bi-LSTM processes the forward and backward information flow that ensures a more complete understanding of the feature dynamics. Finally, the model outputs the classification result, where a label of 0 corresponds to MSS (Microsatellite Stable) and a label of 1 corresponds to MSI-MUT (Microsatellite Instable or Highly Mutated), thereby distinguishing between stable and unstable cancer types.

7). IMPROVED AlexNet (IANet). AlexNet [33] consists of five convolutional layers followed by three FC layers. AlexNet is selected for its computational efficiency and strong capability in learning low- and mid-level visual features. The network takes an input feature through convolution, pooling, and normalization. As the feature moves through the layers, spatial dimensions shrink while the number of feature maps grows, enabling the network to learn increasingly complex features. The final convolutional layer outputs a feature map, which is flattened and passed through the FC layers that produce class scores for image classification. The depth of this conventional model is very low and struggles to learn features. To overcome the above drawback, an IANet is proposed that efficiently extracts complex features and achieves high accuracy in image classification. The architecture of the proposed system is shown in Fig. 4.

IANet learns deep spatial representations from histopathological images for effective feature extraction and classification. The IANet initiates with an input feature $F(E)$, which first passes through a convolutional layer with 96 filters, followed by batch normalization and max pooling to reduce spatial dimensions and stabilize training. The feature then moves through Inception modules A and B, which are followed by another max pooling layer and allows the network to capture multi-scale features. The conventional Inception modules [34] are computationally intensive and increase the model's complexity. So, we introduce an Improved Inception Module (IIM) to capture multi-scale features with a reduction in the number of parameters.

The IIM processes the feature map of the convolutional layer by splitting it into multiple parallel branches, allowing the network to capture features at different scales simultaneously. For IIM A, the branch filter sizes are 64 for the 1×1 convolution, 128 for the 3×3 convolution, and 32 for the 5×5 convolution, while for IIM B, the sizes are 128 for the 1×1 convolution, 192 for the 3×3 convolution, and 96 for the 5×5 convolution. The four branches include a 1×1 convolution, a Squeeze and Excitation block, followed by a 3×3 convolution, a Squeeze and Excitation block, followed by a 5×5 convolution, and a max-pooling layer, followed by a 1×1 convolution. The Squeeze and Excitation block [35] can increase computational complexity and slow down training. So, this study proposes an Improved Squeeze and Excitation block with Fuzzy (ISE-F) to focus on important feature channels and enhance performance with minimal extra computation.

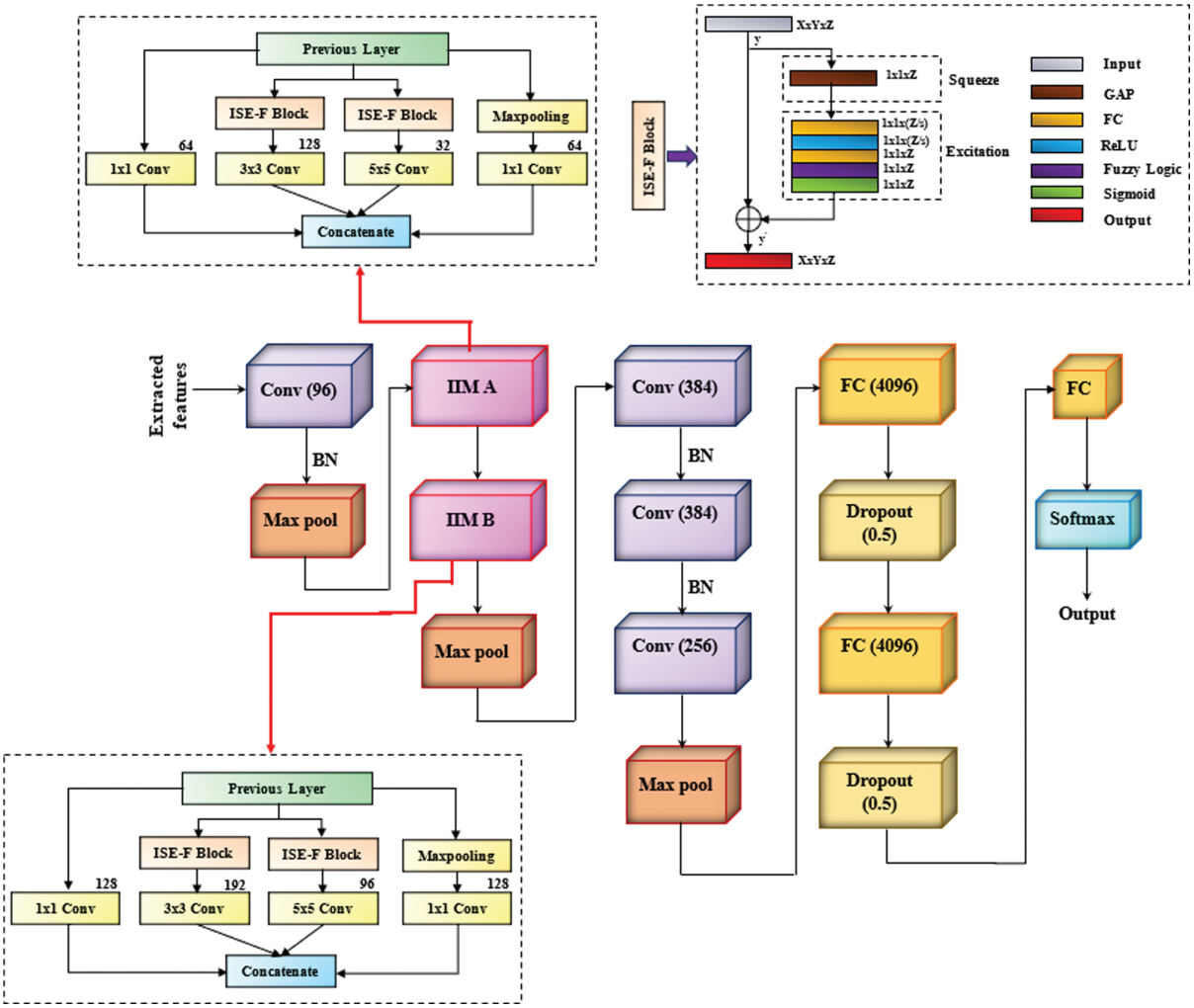


Fig 4. Architecture of IANet.

The ISE-F block within a multi-branch network is aimed at enhancing the network's performance by identifying the most important feature channels. The input feature map is denoted as in Eq. (20):

$$M_i = [m_1, m_2, \ldots, m_z] \in S^{X \times Y \times Z} \tag{20}$$

This M_i is given to the Squeeze step where a Global Average Pooling (GAP) operation condenses each channel of the input feature map into a single value and summarizes the spatial information. The GAP is given in Eq. (21):

$$G_z = \frac{1}{X \times Y} \sum_{p=1}^{X} \sum_{q=1}^{Y} M_z(p,q) \tag{21}$$

where $G \in S^Z$ that is produced by compressing M through $X \times Y$ spatial dimensions.

During Excitation, this compressed vector is passed through a small two-layer FC network with an ReLU activation in between to model inter-channel relationships and generate weights that indicate the importance of each channel. The excitation vector is given in Eq. (22):

$$u = E_2.\lambda(E_1, G) \tag{22}$$

where $u \in S^Z$ is the excitation vector, $\lambda(.)$ is the ReLU activation function, $E_1 \in S^{\frac{z}{s} \times z}$, and $E_2 \in S^{z \times \frac{z}{s}}$.

The refined feature map is then forwarded to the next layer. After the FC layers, a Fuzzy Logic layer can be incorporated to improve decision-making by handling uncertainty and imprecision in the features. In Fuzzy Logic [36], the high and low membership per channel is calculated using Eq. (23) and Eq. (24):

$$h_m = \vartheta(u_z) = \frac{1}{1 + e^{-u_z}} \tag{23}$$

$$l_m = \vartheta(-u_z) = \frac{1}{1 + e^{u_z}} \tag{24}$$

where $\vartheta(.)$ is the sigmoid function, and h_m and l_m are high and low membership, respectively.

Now, the Fuzzy combination of Excitation is given by combining u_z using the Fuzzy membership. It is given in Eq. (25):

$$\tilde{u}_z = h_m.u_z + l_m.(1 - u_z) \tag{25}$$

Therefore, the final gating mechanism with sigmoid activation is given in Eq. (26):

$$r_z = \vartheta(\tilde{u}_z) = \frac{1}{1 + e^{-\tilde{u}_z}} \tag{26}$$

Thus, the final output of the block obtained by rescaling M with activation r_z is given in Eq. (27):

$$\widetilde{y}_z = F_{scale}(m_z, r_z) \tag{27}$$

where $\widetilde{y} = [\widetilde{y}_1, \widetilde{y}_2, \ldots, \widetilde{y}_z]$ and $F_{scale}(m_z, r_z)$ is the channel-wise multiplication between the scalar r_z and the feature map $m_z \in S^{X \times Y}$.

Then these output weights of the ISE-F block are used to scale the input feature map (M_i) and multiply the excitation output with the original input features to emphasize important channels.

Further in IIM, the outputs from all branches are then concatenated into a single feature map. A key component is the 1×1 convolution, which reduces the number of input channels before the larger convolutions and improves their efficiency. This design enables the network to be both wider and deeper without significantly increasing computational cost, which makes the IIM highly effective for multi-scale feature extraction. These features are then passed to the Maxpooling layer, which reduces the spatial dimensions of the feature map.

Next in IANet, subsequent convolutional layers with 384 and 256 filters are accompanied by batch normalization and pooling that further extract complex representations from the image. The extracted features are then passed through three FC layers, where the first two have 4096 neurons each with a dropout rate of 0.5 to prevent overfitting, while the final FC layer is followed by a sigmoid activation to generate the output probability scores ($P(s)$).

8). BI-DIRECTIONAL LONG SHORT-TERM MEMORY (Bi-LSTM). The Bi-LSTM module is employed to capture contextual dependencies and sequential relationships among the extracted deep features generated by IANet. While histopathology images are static, the Bi-LSTM is not utilized to model temporal information. Instead, the deep features extracted by IANet are represented as an ordered feature sequence, where each feature embedding captures distinct spatial and semantic characteristics of the tissue image. The Bi-LSTM learns bidirectional dependencies among these feature representations, enabling the extraction of contextual relationships between local texture patterns and global structural information. Bi-LSTM is used to model long-range dependencies in CNN-extracted feature sequences from histopathology images, capturing global contextual relationships beyond local spatial features. This improves discrimination of complex tissue structures and enhances classification performance. A Bi-LSTM processes sequential input in both forward and backward directions. The output scores ($P(s)$) from IANet are then forwarded to a Bi-LSTM network, which is designed to learn sequential correlations and contextual patterns within the probabilities. By processing the information in both forward and backward directions, the Bi-LSTM detect complex interactions among the features that improve the model's capability to capture subtle variations in tissue morphology.

This refined analysis of probabilities enables more precise classification of GI cancer. The Bi-LSTM generates a final output of 0 for MSS and 1 for MSI-MUT that segregates between stable and unstable cancer types based on the learned probability relationships.

IV. RESULTS AND DISCUSSION

A. SIMULATION PROCEDURE

The Proposed Gastrointestinal Cancer Classification with Shape and Texture Feature Set was simulated by device name SYS-175 using PYTHON. The processor employed was "11th Gen Intel(R) Core (TM) i5-1135G7 @2.40 GHz 2.42 GHz, and the Installed RAM size was 16.0 GB (15.7 GB usable)." Further analysis of Gastrointestinal Cancer Classification was carried out by using the TCGA COAD MSI vs MSS Prediction Dataset [37] and Gastric Cancer Histopathology Tissue Image Dataset (GCHTID) [38]. The python library configurations used to ensure reproducibility is provided in Table III.

Table III. Python libraries for reproducible implementation

Library	Version
numpy	1.21.6
opencv-python	4.11.0.86
scikit-learn	1.0.2
PySimpleGUI	5.0.8.3
matplotlib	3.5.3
scikit-image	0.19.2
pandas	1.2.5
colorama	0.4.4
tensorflow	2.6.1
pytictoc	1.5.2
scipy	1.7.3
phasepack	1.5

B. DATASET DESCRIPTION

TCGA COAD MSI vs MSS Prediction Dataset: Dataset 1

The dataset is built from the formalin-fixed paraffin-embedded (FFPE) diagnostic slides ("DX" category at the GDC data portal), which ensures clinical relevance. From these slides, 411,890 unique image patches were generated, representing both the colorectal and gastric cancer patients. Along with this, it includes 117273 of MSS and 75039 of MSIMUT.

In our implementation, 5000 samples were randomly selected from the dataset to ensure balanced and unbiased representation; although the original datasets contain 117,273 (MSS_JPEG) and 75,039 (MSIMUT) images, we restricted our experiments to 2,500 samples per dataset (total 5,000) to ensure computational feasibility and balanced representation. The class label 0 is pointing to MSS, and 1 implies MSIMUT with 2500 data for each class label. The MSS label refers to tumors with stable microsatellite DNA, indicating proficient DNA mismatch repair, whereas the MSI–MUT label denotes tumors with microsatellite instability, reflecting defective mismatch repair and higher mutation frequency, which is associated with distinct tumor biology and clinical behavior in GI cancers. The dataset split was carefully designed to prevent data leakage and ensure generalizability. Specifically, images assigned to the training set were never repeated in the test set. Patient-level separation was enforced so that no patches from the same patient appeared in both sets, thereby eliminating patch-level leakage. To further strengthen reliability, we employed cross-validation, which allowed the model to be trained and evaluated on multiple folds of the data, reducing bias from any single split. Table IV shows the training and testing data count of Dataset 1.

GCHTID: Dataset 2 (Used only for cross-validation)

This dataset, the GCHTID, is a valuable resource for researchers working on artificial intelligence in medical imaging, specifically for gastric cancer. It contains a large collection of expertly curated histopathology images.

Table IV. Training and testing data count for Dataset 1

Training data %	Number of training data	Number of testing data
60	3000	2000
70	3500	1500
80	4000	1000
90	4500	500

1). MODEL SUMMARY ANALYSIS. The proposed IAB-GCC architecture follows a hybrid deep learning pipeline combining convolutional feature extraction with sequence-based classification. Table V shows the analysis of model summary based on layers and parameters. The input image of size $32 \times 32 \times 1$ is first processed through a Conv2D layer with 96 filters, followed by batch normalization to stabilize feature distributions and improve convergence. A max pooling layer reduces spatial dimensions, and GAP is applied to obtain a compact 96-dimensional feature representation. These extracted deep features are then forwarded to a dense layer (6 units) for feature refinement, followed by a bidirectional layer (256 units) to capture contextual and bidirectional dependencies in the learned representations. A dropout layer (0.5) is used to reduce overfitting and improve generalization. Finally, a FC output layer with two classes performs the final classification using activation function, enabling binary classification of GI cancer patterns.

2). COMPARATIVE PERFORMANCE STUDY. As part of this performance study, it measures how the models are working to classify the GI cancer by using the shape and texture features. Table VI shows the hyperparametr of the proposed and existing method.

The following bar graphs highlight the different aspects of system reliability under various measures. This helps to ensure the cancer cases are detected in early stages. Among this classification, we compare the base models, namely M-CAM [26], ALSTM-ALO [19], Ensemble Model with XAI [27], and SC-DSAN [28]. Under Fig. 5(i), the precision metric validates how many samples that were predicted as cancer are actually cancer cases. The Bi-LSTM model has produced the correct case percentage of 83.6% in 80% of data training and 88% of precision ranges in 90%. Followed by Fig. 5(ii), the accuracy reflects how the model uses the shape and texture feature set to correctly classify the GI cancer cases overall. The Baseline Bi-LSTM models have earned an overall accuracy range of 81% in 70% of the training set and realizing 89% in 90%.

Table V. Analysis of model summary based on layers and parameters

Layers	Parameters
input_1	(None, 32, 32, 1)
conv2d (Conv2D)	(None, 30, 30, 96)
batch_normalization	(None, 30, 30, 96)
max_pooling2d	(None, 14, 14, 96)
global_average_pooling2d	(None, 96)
dense (Dense)	(None, 6)
bidirectional	(None, 256)
dropout_2 (Dropout)	(None, 256)
dense_10 (Dense)	(None, 2)

Table VI. Hyperparameter of the proposed and existing methods

Model	Parameters
PolyNet	optimizer = 'adam', epochs-50 batch size-32
LeNet	optimizer = 'adam' epoch-50 batch size-64
Bi-LSTM	optimizer: rmsprop epoch-50 batch size-64
GoogleNet	optimizer = 'adam', epochs-50 batch size-64
ShuffleNet	optimizer='adam', epochs-50 batch size-32
IAB-GCC	optimizer = 'adam', epochs-50 batch size-64

The IAB-GCC model has used a hybrid deep learning model by integrating the improved AlexNet and Bi-LSTM model to improve the classification process of GI cancers. This is proved via the higher precision and accuracy scores of 97% and 96%, respectively. Among Fig. 5(iii) & (iv), it graphed the score of sensitivity and specificity under the cancer classification of baseline and IAB-GCC model. The IAB-GCC model has accomplished a higher sensitivity of 95% and specificity of 97% to exceed the other baseline models.

The false classification of GI cancers of the baseline model and the IAB-GCC model is calculated under the conditions shown in Fig. 6. The FDR metric tells how often the models predicted patients as having cancer are actually wrong. Continuously, the FPR and FNR reveal the percentage of actual cancer that the models failed to detect and wrongly classified as cancer. The analysis of negative cases, the Bi-LSTM model has marked the higher FDR of 15%, FNR of 9% (in Fig. 6(ii)), and FPR of 9% over 90% (in Fig. 6(iii)). In the same manner, the ShuffleNet model is trained under 60% to prohibit the maximal FDR score of 25% (in Fig. 6(i)), in 60% of data training and noted the 10.1% in 90% of training. By exceeding the baseline models, the IAB-GCC model results in extremely lower negative cases by enhancing the filtering method and using improved entropy-based features. Notably, the FDR metric of 26%, FNR of 4.5% and FPR of 3%, respectively.

The section below presents further validation among the baseline model and the IAB-GCC model by altering the data training percentages from ranging of 60 to 90. Under this estimation, the IAB-GCC model has marked the highest score over the F-measure, MCC, and NPD metrics that were categorized in Fig. 7.

3). ABLATION ASSESSMENT. The ablation assessment is carried out under the classification of GI cancers. These were carried out to know the necessity of each stage and its efficiency levels. Table VII presents the ablation study of proposed IAB-GCC model with conventional hybrid, Gaussian filtering, and entropy-based methods in 80%. In addition, the IAB-GCC is compared without the preprocessing process. Each stage of the process has different functionality under classification. So, if we removed preprocessing, it attains a lesser accuracy of 84.8%, because without this stage, it is

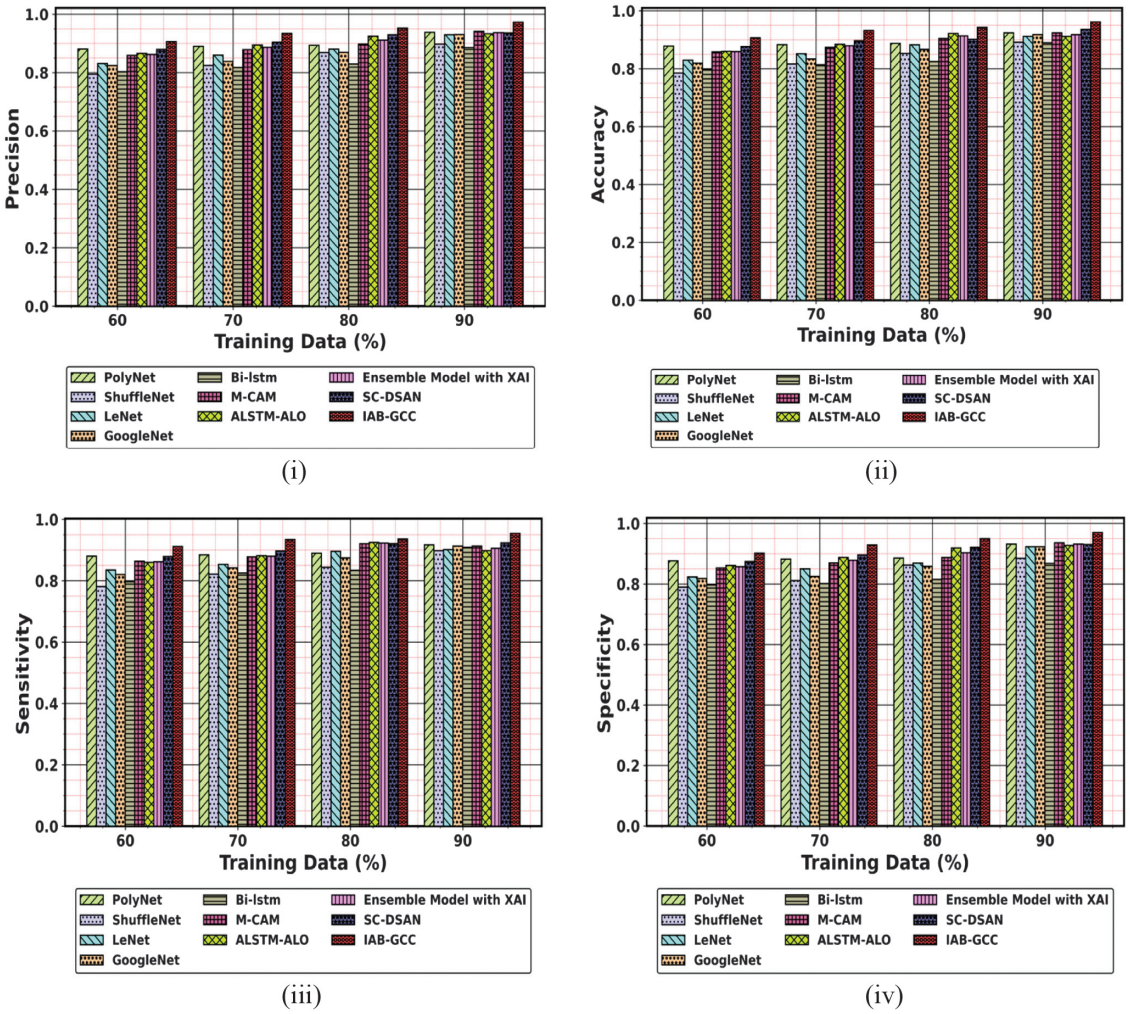


Fig. 5. Comparative performance study under positive case measures: (i) precision, (ii) accuracy, (iii) sensitivity and (iv) specificity.

not able to reduce the noise in images to classify them perfectly. While the IAB-GCC is differentiated from the conventional model frameworks, it accomplished the lesser efficiency. For example, IAB-GCC with conventional Gaussian and conventional entropy has reported the accuracy of 87.6% and 86.4%, respectively. The IAB-GCC model uses an improved Gaussian filtering technique for preprocessing, an improved entropy-based feature for feature extraction to boost their cancer classification process by the higher accuracy of 94.3%, and a lower false estimation of 0.064 and 0.050 under FNR and FPR, respectively.

Table VII. Ablation assessment of the proposed model

Metrics	IAB-GCC	IAB-GCC with conventional hybrid	IAB-GCC without preprocessing	IAB-GCC model with conventional Gaussian	IAB-GCC model with conventional entropy
Accuracy	0.943	0.884	0.848	0.876	0.864
Sensitivity	0.936	0.868	0.849	0.879	0.849
Specificity	0.950	0.902	0.847	0.872	0.881
Precision	0.953	0.909	0.862	0.886	0.889
F-measure	0.945	0.888	0.856	0.883	0.869
MCC	0.886	0.769	0.695	0.751	0.729
NPV	0.933	0.858	0.833	0.865	0.838
FPR	0.050	0.098	0.153	0.128	0.119
FNR	0.064	0.132	0.151	0.121	0.151
FDR	0.047	0.091	0.138	0.114	0.111

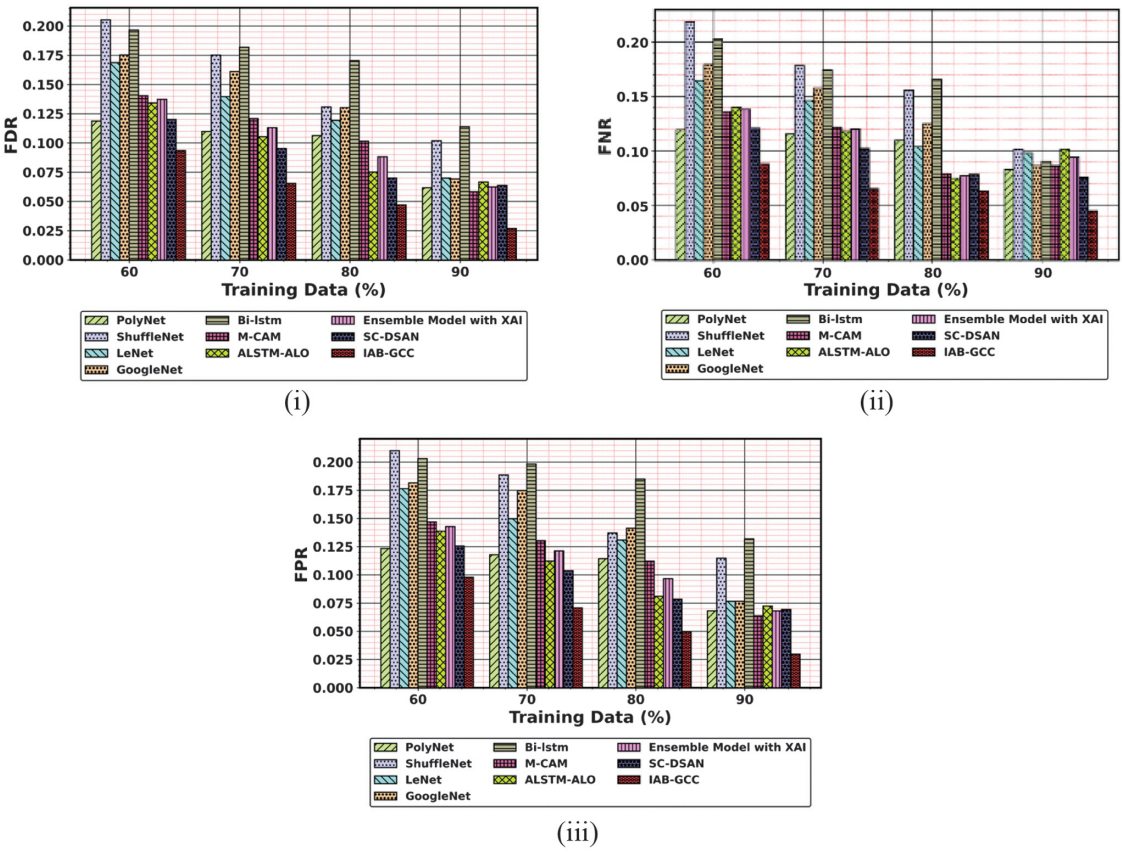


Fig. 6. Comparative performance study under negative case measures: (i) FDR, (ii) FNR, and (iii) FPR.

4). STATISTICAL VALIDATION. The statistical assessment is undertaken regarding accuracy among the baseline, and the IAB-GCC model is displayed in Table VIII. Over the accuracy-based assessment, the higher values reveal that it is better for cancer classification. But the baseliner models, likely Bi-LSTM, ShuffleNet, and LeNet, have produced the mean accuracy range of 83.2%, 83.7%, and 86.9%. These models all showcase lower mean accuracy when compared with the IAB-GCC model (mean accuracy of 93.6%). Furthermore, the median accuracy of the Bi-LSTM model is 81.9%, the minimum accuracy is 79.7%, and the maximum accuracy is 89.0%. Among these observations, the average reliability that is estimated by statistical measures, from this, the IAB-GCC model average score under min, max, and so on, is higher as per the statement of accuracy. This is verified by the maximal accuracy of 96.2% and the minimal accuracy of 90.7%.

5). QUANTITATIVE ANALYSIS. Within this validation, the various categories of filtering techniques are tested with the proposed filtering approach to see which framework gives a higher quality of image to classify the cancer. For calculating this, we use the support of two quality analysis metrics, namely PSNR and SSIM. PSNE metric measures the quality of an image after the image has been processed using the filtering methods. In the same manner, the SSIM metric checks the similarity between the two images (original and predicted) by comparing their structural details:

Table VIII. Statistical assessment

Models	Mean	Median	Std dev	Min	Max
PolyNet	0.893	0.886	0.018	0.879	0.924
ShuffleNet	0.837	0.835	0.040	0.786	0.892
LeNet	0.869	0.867	0.031	0.830	0.912
GoogleNet	0.860	0.851	0.038	0.820	0.918
Bi-lstm	0.832	0.819	0.035	0.797	0.890
M-CAM	0.890	0.890	0.026	0.859	0.924
ALSTM-ALO	0.895	0.898	0.024	0.861	0.922
Ensemble Model with XAI	0.893	0.896	0.024	0.860	0.918
SC-DSAN	0.903	0.899	0.021	0.877	0.936
IAB-GCC	0.936	0.938	0.020	0.907	0.962

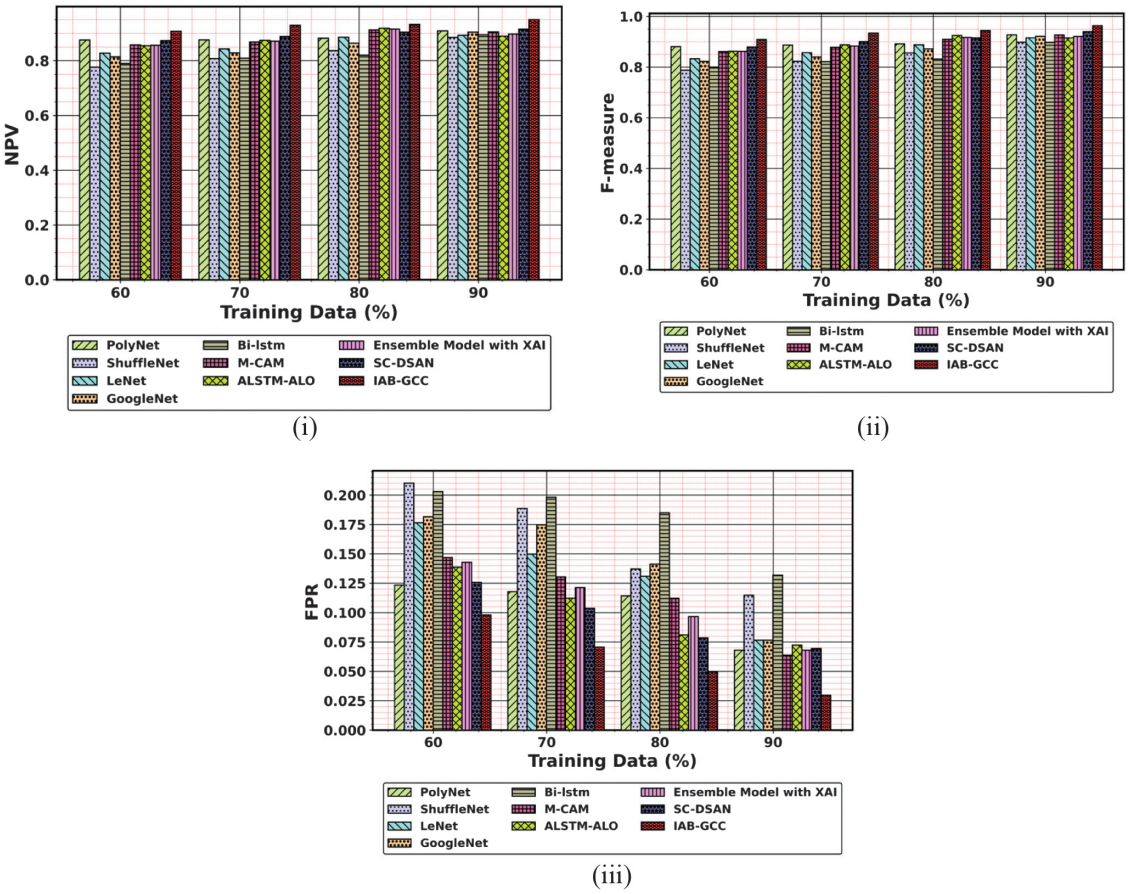


Fig. 7. Comparative performance study under other case measures: (i) NPV, (ii) F-measure, and (iii) MCC.

$$PSNR = 10.\log_{10}\left(\frac{MAX_Y^2}{MSE}\right) \tag{28}$$

$$SSIM(Y,R) = \frac{(2_{\mu Y\mu R} + T_1)(2_{\sigma YR} + T_2)}{(\mu_Y^2 + \mu_R^2 + T_1)(\sigma_Y^2 + \sigma_R^2 + T_2)} \tag{29}$$

Table IX displays the comparison of baseline filtering frameworks such as Gabor, Bilateral, Median, and conventional Gaussian. Along with this, it distinguishes the newly introduced modified GF methods. Conventional Gaussian filtering has lower quality in preprocessed images, leading to some wrong estimation. Because of this, it resulted in the PSNR of 26.873 dB and SSIM of 0.667. The IAB-GCC model has improved the Gaussian filtering (AWGF) with a noise model to effectively reduce the noise by adaptive smoothing strength to dictate the higher clarity and similarity by PSNR of 33.105 dB and SSIM of 0.844.

Table IX. Quantitative validation under different filtering techniques

	SSIM	PSNR [db]
Gabor filtering	0.760	23.744
Bilateral filtering	0.632	23.862
Median filtering	0.749	20.466
Conv Gaussian filtering	0.667	26.873
AWGF	0.844	33.105

6). PREPROCESSING IMAGE RESULT. Figure 8 illustrates the clear preprocessing outcome of conventional filtering techniques and the proposed AWGF.

V. CONFIDENCE INTERVAL

The confidence intervals are shown in Table X, which shows the range of each model's actual performance. These performance intervals are predicted to fall within a certain degree of confidence. The suggested IAB-GCC exhibits the greatest and most consistent performance, ranging from 0.8995 to 0.9725, whereas PolyNet is anticipated to reach performance between 0.8605 and 0.9264. While intervals like ShuffleNet's [0.7636, 0.9100] reflect higher unpredictability in model outputs, narrower intervals like IAB-GCC's suggest more consistent and stable performance. All things considered, these intervals offer a statistical measure of certainty regarding each model's performance, allowing for more accurate comparisons than single-point measures.

VI. STATISTICAL TEST ANALYSIS

Statistical Test Analysis of Model Performance

The t-test, Friedman p-values, and Wilcoxon p-values for each model are shown in Table XI. The statistical study assesses the

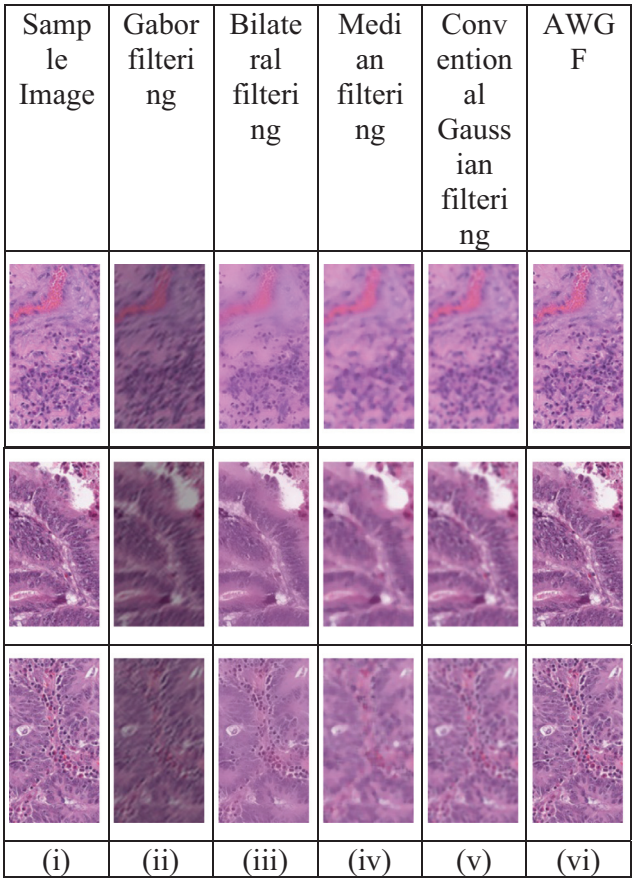


Fig. 8. Preprocessing result of proposed filtering techniques: (i) sample image, (ii) Gabor filtering, (iii) bilateral filtering, (iv) median filtering, (v) conventional Gaussian filtering, and (vi) AWGF.

models' performance differences at $\alpha = 0.05$. In at least one of these tests, models like Bi-LSTM, ShuffleNet, M-CAM, ALSTM-ALO, LeNet, Ensemble Model with XAI, and SC-DSAN have p-values below 0.05, showing notable gains in performance. In certain tests, PolyNet and GoogleNet's p-values are marginally over 0.05, indicating that their improvements are not statistically significant at the 5% level. Overall, by offering inferential support for variations in model performance, this technique enhances descriptive metrics.

Table X. Confidence intervals in existing and proposed models

Models	Confidence interval
PolyNet	[0.8605, 0.9264]
ShuffleNet	[0.7636, 0.9100]
LeNet	[0.8118, 0.9265]
GoogleNet	[0.7901, 0.9292]
Bi-lstm	[0.7668, 0.8962]
M-CAM	[0.8433, 0.9375]
ALSTM-ALO	[0.8506, 0.9390]
Ensemble Model with XAI	[0.8480, 0.9372]
SC-DSAN	[0.8635, 0.9424]
IAB-GCC	[0.8995, 0.9725]

Table XI. Statistical test analysis in various existing models

IAB-GCC Vs	T-test	Wilcoxon *P*-value	Friedman *P*-value
PolyNet	0.052	0.049	0.054
ShuffleNet	0.041	0.029	0.037
LeNet	0.047	0.038	0.047
GoogleNet	0.055	0.055	0.055
Bi-lstm	0.024	0.008	0.007
M-CAM	0.047	0.037	0.046
ALSTM-ALO	0.041	0.028	0.036
Ensemble Model with XAI	0.033	0.030	0.043
SC-DSAN	0.046	0.034	0.037

VII. K-FOLD ANALYSIS

Table XII compares the performance of various models across five K-fold analysis. The proposed IAB-GCC achieved the best overall performance, with scores ranging from 0.952 to 0.970, demonstrating superior accuracy, robustness, and stability across all folds. Compared with other existing models, IAB-GCC consistently outperformed methods such as the Ensemble Model with XAI, ALSTM-ALO, M-CAM, and PolyNet. The remaining models, including SC-DSAN, ShuffleNet, LeNet, GoogleNet, and Bi-LSTM, achieved comparatively lower performance values. These results confirm the effectiveness of the proposed IAB-GCC model in improving predictive capability and ensuring reliable generalization across different validation datasets.

VIII. CROSS-DATASET VALIDATION ANALYSIS

The cross-dataset validation results demonstrate the robustness and generalization capability of the proposed model when trained and tested on different datasets. Table XIII shows the cross-validation of both Dataset 1 and Dataset 2. When trained using Dataset 1 and tested on Dataset 2, the model achieved an accuracy of 0.949, sensitivity of 0.943, and specificity of 0.956, indicating effective

Table XII. K-fold analysis of existing and proposed methods

Models	K-Fold1	K-Fold2	K-Fold3	K-Fold4	K-Fold5
Ensemble model with XAI	0.918	0.927	0.916	0.935	0.934
SC-DSAN	0.877	0.883	0.886	0.890	0.881
PolyNet	0.929	0.932	0.918	0.932	0.92
ShuffleNet	0.891	0.892	0.881	0.891	0.882
LeNet	0.898	0.907	0.91	0.903	0.901
GoogleNet	0.884	0.896	0.887	0.909	0.882
Bi-LSTM	0.89	0.887	0.903	0.873	0.886
M-CAM	0.918	0.923	0.924	0.939	0.929
ALSTM-ALO	0.934	0.922	0.933	0.935	0.921
IAB-GCC	0.958	0.969	0.952	0.959	0.970

Table XIII. Cross-dataset validation performance of the proposed model using Dataset 1 and Dataset 2

Metrics	Trained with Dataset 1 and tested with Dataset 2	Trained with Dataset 2 and tested with Dataset 1
Accuracy	0.949	0.948
Sensitivity	0.943	0.941
Specificity	0.956	0.955
Precision	0.959	0.958
F-measure	0.951	0.949
MCC	0.892	0.890
NPV	0.939	0.937
FPR	0.044	0.045
FNR	0.057	0.059
FDR	0.041	0.042

classification performance. Similarly, training on Dataset 2 and testing on Dataset 1 produced comparable results, with an accuracy of 0.948 and a precision of 0.958. The high F-measure, MCC, and NPV values in both cases confirm the reliability and consistency of the model. Additionally, the low FPR, FNR, and FDR values demonstrate reduced misclassification rates and strong cross-dataset generalization performance. These consistent results further indicate the model's potential for clinical decision-support applications through improved generalization performance.

IX. OPTIMIZATION ANALYSIS

The optimization analysis is performed to evaluate the impact of different optimization algorithms on the convergence behavior and classification performance of the proposed IAB-GCC model. Table XIV shows the Optimization Analysis for the Proposed IAB-GCC Model. By comparing Adam, RMSprop, SGD, and Adagrad, the study identifies the most suitable optimizer for stable and efficient learning. Adam achieves the best overall results with highest accuracy (0.943), precision (0.953), sensitivity (0.936), and lowest error rates (FPR, FNR, and FDR), indicating faster convergence and better generalization. RMSprop and Adagrad show moderate performance, while SGD performs comparatively lower due to slower convergence. This analysis ensures selection of an optimal training strategy for reliable GI cancer classification.

Table XIV. Optimization analysis for the proposed IAB-GCC model

Metrics	Adam	RMSprop	SGD	Adagrad
Accuracy	0.943	0.939	0.935	0.939
Sensitivity	0.936	0.865	0.837	0.887
Specificity	0.950	0.928	0.940	0.935
Precision	0.953	0.871	0.839	0.896
F-measure	0.945	0.868	0.838	0.891
MCC	0.886	0.824	0.797	0.842
NPV	0.933	0.923	0.919	0.926
FPR	0.050	0.072	0.060	0.065
FNR	0.064	0.135	0.163	0.113
FDR	0.047	0.129	0.161	0.104

X. LEARNING RATE ANALYSIS

The learning rate analysis evaluates the sensitivity of the proposed IAB-GCC model to different learning rates, including 0.001, 0.01, 0.05, and 0.1, to identify the most stable and optimal training configuration. Table XV shows the Learning Rate Analysis of the Proposed IAB-GCC Model. The results show that a learning rate of 0.001 achieves the best overall performance, with highest accuracy (0.943), precision (0.953), sensitivity (0.936), and F-measure (0.945), along with the lowest error rates (FPR, FNR, and FDR). As the learning rate increases, performance gradually degrades due to unstable convergence and suboptimal weight updates. Higher learning rates (0.05 and 0.1) significantly reduce classification performance, indicating overshooting during optimization. This analysis confirms that a smaller learning rate ensures stable convergence and better generalization for GI cancer classification.

XI. DISCUSSION

The proposed IAB-GCC model demonstrates strong performance in classifying GI cancers, achieving a high precision of 97%, accuracy of 96%, sensitivity of 95%, and specificity of 97%, outperforming baseline models such as Bi-LSTM, ShuffleNet, and LeNet. Existing GI cancer classification methods often suffer from inadequate noise handling, limited discriminative feature extraction, and insufficient modeling of complex feature relationships, which can reduce classification performance in challenging histopathological images. The proposed IAB-GCC framework addresses these limitations through AWGF-based preprocessing for effective noise suppression and edge preservation, WSE for emphasizing spatially significant regions, and integrated shape and texture features for enhanced feature representation. Furthermore, the IANet–Bi-LSTM architecture captures both local discriminative patterns and contextual feature dependencies, resulting in improved classification accuracy, robustness, and generalization compared with conventional approaches. Statistical validation through confidence intervals and t-test, Wilcoxon, and Friedman tests at $\alpha = 0.05$ confirms that these improvements are generally significant, providing inferential support beyond descriptive metrics. However, closer examination of failure cases reveals that baseline models exhibit substantial false discovery rates up to 25% and false negative or false positive predictions, highlighting challenges in correctly classifying noisy samples. Despite the use of

Table XV. Learning rate analysis of the proposed IAB-GCC model

Metrics	Learning rate (0.001)	Learning rate (0.01)	Learning rate (0.05)	Learning rate (0.1)
Accuracy	0.943	0.933	0.926	0.920
Sensitivity	0.936	0.902	0.893	0.853
Specificity	0.950	0.932	0.918	0.905
Precision	0.953	0.913	0.895	0.894
F-measure	0.945	0.917	0.907	0.897
MCC	0.886	0.864	0.857	0.831
NPV	0.933	0.902	0.909	0.906
FPR	0.050	0.068	0.082	0.095
FNR	0.064	0.098	0.107	0.147
FDR	0.047	0.087	0.105	0.106

AWGF preprocessing and improved entropy-based feature extraction, the model's generalizability limits the relatively small subset of 5,000 training images. Potential biases also exist due to class imbalance in the original dataset and variability in histopathological image quality, which could affect performance across different populations or imaging protocols. Furthermore, the multi-stage IANet-Bi-LSTM architecture improves interpretability and accuracy, which increases computational complexity. Overall, although the IAB-GCC model shows strong quantitative performance, careful consideration of dataset quality, representativeness, and potential biases. Future work will focus on validating the proposed IAB-GCC model on larger and more diverse datasets, reducing computational complexity, and enhance the interpretability and clinical applicability of the proposed framework.

XII. CONCLUSION

This paper proposed an IAB-GCC with a shape and texture feature set to improve diagnostic accuracy. The IAB-GCC process began with input images, which were preprocessed using an AWGF technique to reduce noise and enhance image quality. Key features included WSE-based features, texture features using GLCM, shape and morphological features were then extracted and fed into an IANet and Bi-LSTM model. This combination influenced both spatial and temporal patterns, which enabled the system to classify cancer cases as stable or unstable, providing precise and reliable diagnostic results. To conclude, the IAB-GCC model with the shape and texture feature set is classifying more accurately under all cases than the baseline models. However, the study is limited by the relatively small dataset, potential class imbalance, and computational complexity, which may affect generalizability and real-time deployment. Future work will focus on expanding the dataset, improving robustness to diverse imaging conditions, and optimizing the architecture for real-time clinical deployment, ensuring broader applicability and practical clinical utility.

DATA AVAILABILITY STATEMENT

The data available at https://zenodo.org/records/2530835#.XVPlRHUzYeM

CREDIT AUTHOR STATEMENT

Vanishree Abhay: Conceptualization, Methodology, Software.
Achyutha Prasad N.: Data curation, Writing – Original draft preparation.
Dr. Sathish Babu N: Software, Validation.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

FUNDING STATEMENT

This research did not receive any specific funding.

ETHICS APPROVAL

There are no human or animal participants in this article; therefore ethical approval is not applicable for this work.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] T. Seidlitz and D. E. Stange, "Gastrointestinal cancer organoids—Applications in basic and translational cancer research," *Exp. Mol. Med.*, vol. 53, no. 10, pp. 1459–1470, 2021. DOI: 10.1038/s12276-021-00654-3

[2] C. Cao *et al.*, "Host-gut microbiota metabolic interactions and their role in precision diagnosis and treatment of gastrointestinal cancers," *Pharmacol. Res.*, vol. 207, 2024. DOI: 10.1016/j.phrs.2024.107321

[3] O. Prokopchuk *et al.*, "Novel tissue inhibitor of metalloproteinases-1/liver/cachexia score predicts prognosis of gastrointestinal cancer patients," *J. Cachexia Sarcopenia Muscle*, vol. 12, no. 2, pp. 378–392, 2021. DOI: 10.1002/jcsm.12680

[4] C. Qian *et al.*, "Neutrophil-lymphocyte ratio and circulating tumor cells counts predict prognosis in gastrointestinal cancer patients," *Front Oncol.*, vol. 11, 2021. DOI: 10.3389/fonc.2021.710704

[5] H. P. Gandikota, S. Abirami, and M. S. Kumar, "Bottleneck feature-based U-Net for automated detection and segmentation of gastrointestinal tract tumors from CT scans," *Trait. Sig.*, vol. 40, no. 6, p. 2789, 2023. DOI: 10.18280/ts.400639

[6] M. Fahad *et al.*, "Deep insights into gastrointestinal health: A comprehensive analysis of GastroVision dataset using convolutional neural networks and explainable AI," *Biomed. Signal Process. Control*, vol. 102, 2025. DOI: 10.1016/j.bspc.2024.107260

[7] L.-S. Li, X.-Y. Guo, and K. Sun, "Recent advances in blood-based and artificial intelligence-enhanced approaches for gastrointestinal cancer diagnosis," *World J. Gastroenterol.*, vol. 27, no. 34, 2021. DOI: 10.3748/wjg.v27.i34.5666

[8] L. Wang *et al.*, "Utilization of ultrasonic image characteristics combined with endoscopic detection on the basis of artificial intelligence algorithm in diagnosis of early upper gastrointestinal cancer," *J. Healthc. Eng.*, vol. 2021, no. 1, 2021. DOI: 10.1155/2021/2773022

[9] A. J. Read *et al.*, "Prediction of gastrointestinal tract cancers using longitudinal electronic health record data," *Cancers (Basel)*, vol. 15, no. 5, 2023. DOI: 10.3390/cancers15051399

[10] H. Gou *et al.*, "Gut microbial metabolites: Shaping future diagnosis and treatment against gastrointestinal cancer," *Pharmacol. Res.*, 2024. DOI: 10.1016/j.phrs.2024.107373

[11] A. Pal *et al.*, "Advanced segmentation of Gastrointestinal (GI) cancer disease using a novel U-MaskNet Model," *Life (Basel)*, vol. 14, no. 11, 2024. DOI: 10.3390/life14111488

[12] J. Wang *et al.*, "Global analysis of microbiota signatures in four major types of gastrointestinal cancer," *Front. Oncol.*, vol. 11, 2021. DOI: 10.3389/fonc.2021.685641

[13] B. Hu *et al.*, "Continuous prediction for tumor mutation burden based on transcriptional data in gastrointestinal cancers," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, 2024. DOI: 10.1186/s12911-024-02794-8

[14] X. Huang *et al.*, "The application of artificial intelligence in upper gastrointestinal cancers," *J. Natl. Cancer Center.*, 2024. DOI: 10.1016/j.jncc.2024.12.006

[15] D. Ai *et al.*, "Tumor heterogeneity in gastrointestinal cancer based on multimodal data analysis," *Genes (Basel)*, vol. 15, no. 9, 2024. DOI: 10.3390/genes15091207

[16] N. G. Laleh *et al.*, "Deep learning for interpretable end-to-end survival prediction in gastrointestinal cancer histopathology," In

COMPAY 2021: The third MICCAI workshop on Computational Pathology, 2021.
[17] M. Tang *et al.*, "Near-term prognostic impact of integrated muscle mass and function in upper gastrointestinal cancer," *Clin. Nutr.*, vol. 4, no. 9, pp. 5169–5179, 2021. DOI: 10.1016/j.clnu.2021.07.028
[18] A. S. Almasoud *et al.*, "Gastrointestinal cancer detection and classification using African vulture optimization algorithm with transfer learning," *IEEE Access*, vol. 12, pp. 23122–23131, 2024. DOI: 10.1109/ACCESS.2024.3351773
[19] O. M. Mirza *et al.*, "Computer aided diagnosis for gastrointestinal cancer classification using hybrid rice optimization with deep learning," *IEEE Access*, vol. 11, pp. 76321–76329, 2023. DOI: 10.1109/ACCESS.2023.3297441
[20] F. A Almarshad *et al.*, "An efficient optimal CapsNet model-based computer-aided diagnosis for gastrointestinal cancer classification," *IEEE Access*, vol. 12, 2024. DOI: 10.1109/ACCESS.2024.3442831
[21] A. A. Wafa *et al.*, "Integrating deep learning for accurate gastrointestinal cancer classification: A comprehensive analysis of MSI and MSS patterns using histopathology data," *Neural Comput. Appl.*, pp. 1–33, 2024. DOI: 10.1007/s00521-024-10287-y
[22] Y. Dai *et al.*, "A sensitive and robust plasma-based DNA methylation panel for early detection of target gastrointestinal cancers," *Neoplasia*, vol. 46, 2023. DOI: 10.1016/j.neo.2023.100941
[23] Z. Ye *et al.*, "Advanced NLP-driven predictive modeling for tailored treatment strategies in gastrointestinal cancer," *SLAS Technol.*, 2025. DOI: 10.1016/j.slast.2025.100264
[24] C.-J. Huang *et al.*, "Construction of hot tumor classification models in gastrointestinal cancers," *J. Transl. Med.*, vol. 23, no. 1, 2025. DOI: 10.1186/s12967-025-06230-x
[25] N. Yousefpour *et al.*, "Electrical tumor detection probe calibrated to diagnose gastrointestinal cancer mass in real-time," *J. Clin. Med.*, vol. 13, no. 19, 2021. DOI: 10.3390/jcm13195823
[26] M. Zubair *et al.*, "An interpretable framework for gastric cancer classification using multi-channel attention mechanisms and transfer learning approach on histopathology images," *Sci. Rep.*, vol. 15, no. 13087, pp. 1–34, 2025. DOI: 10.1038/s41598-025-97256-0
[27] M. M. Auzine *et al.*, "Development of an ensemble CNN model with explainable AI for the classification of gastrointestinal cancer," *PLoS One*, vol. 19, Article e0305628, 2024. DOI: 10.1371/journal.pone.0305628
[28] M. A. Khan *et al.*, "A novel network-level fused deep learning architecture with shallow neural network classifier for gastrointestinal cancer classification from wireless capsule endoscopy images," *BMC Med. Inform. Decis. Mak.*, vol. 25, no. 1, p. 150, 2025.
[29] K. V. Ranjitha and T. P Pushphavathi, "Analysis on improved gaussian-wiener filtering technique and GLCM based feature extraction for breast cancer diagnosis," *ScienceDirect*, vol. 235, pp. 2857–2866, 2024.
[30] M. Vila *et al.*, *Analysis of image informativeness measures*. ICIP, 2014, pp. 1086–1090. DOI: 10.1109/DATABIA53375.2021.9650095
[31] F. Y. Manik, D. S. B. Ginting, and M. I. Aldeena, "Extraction analysis of texture and morphological features as characteristics of cabbage images," *IEEE Xplore*, pp. 88–92, 2021. DOI: 10.1109/DATABIA53375.2021.9650095
[32] U. Upadhyay and S. Gupta, "A survey on image feature extraction techniques," *Int. J. Sci. Res. Eng. Trends*, vol. 10, no. 3, pp. 243–255, 2024.
[33] S. Shaees *et al.*, "Facial emotion recognition using transfer learning," *ReseachGate*, vol. 01, pp. 192–196, 2020.
[34] G. Jiang, "Security detection design for laboratory networks based on enhanced LSTM and AdamW algorithms," *Int. J. Inf. Technol. Syst. Approach*, vol. 16, no. 2, pp. 1–13, 2017. DOI: 10.4018/IJITSA.319721
[35] J. Hu *et al.*, "Squeeze-and-excitation networks," arXiv, pp. 1–13, 2019.
[36] M. Irfan *et al.*, "Improved Brain Tumor detection in MRI: Fuzzy sigmoid convolution in deep learning," arXiv, pp. 1–8, 2025.
[37] https://zenodo.org/records/2530835#.XVPlRHUzYeM
[38] https://doi.org/10.6084/m9.figshare.26014469.v