Research article

GastroFuse-Net: an ensemble deep learning framework designed for gastrointestinal abnormality detection in endoscopic images

Sonam Aggarwal[1], Isha Gupta[1], Ashok Kumar[2], Sandeep Kautish[3], Abdulaziz S. Almazyad[4], Ali Wagdy Mohamed[5,6], Frank Werner[7,*] and Mohammad Shokouhifar[8,*]

[1] Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
[2] Model Institute of Engineering and Technology, Jammu, J&K, India
[3] Chandigarh University, Mohali, Punjab 140413 India
[4] Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh 11543, Saudi Arabia
[5] Operations Research Department, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza 12613, Egypt
[6] Applied Science Research Center, Applied Science Private University, Amman 11931, Jordan
[7] Faculty of Mathematics, Otto-von-Guericke University, Magdeburg 39016, Germany
[8] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

* **Correspondence:** Email: frank.werner@ovgu.de, shokouhifar.sbu@gmail.com; Tel: +493916752025, +989121944288.

Abstract: Convolutional Neural Networks (CNNs) have received substantial attention as a highly effective tool for analyzing medical images, notably in interpreting endoscopic images, due to their capacity to provide results equivalent to or exceeding those of medical specialists. This capability is particularly crucial in the realm of gastrointestinal disorders, where even experienced gastroenterologists find the automatic diagnosis of such conditions using endoscopic pictures to be a challenging endeavor. Currently, gastrointestinal findings in medical diagnosis are primarily determined by manual inspection by competent gastrointestinal endoscopists. This evaluation procedure is labor-intensive, time-consuming, and frequently results in high variability between laboratories. To address these challenges, we introduced a specialized CNN-based architecture called GastroFuse-Net, designed to recognize human gastrointestinal diseases from endoscopic images. GastroFuse-Net was developed by combining features extracted from two different CNN models with different numbers of layers, integrating shallow and deep representations to capture diverse aspects of

the abnormalities. The Kvasir dataset was used to thoroughly test the proposed deep learning model. This dataset contained images that were classified according to structures (cecum, z-line, pylorus), diseases (ulcerative colitis, esophagitis, polyps), or surgical operations (dyed resection margins, dyed lifted polyps). The proposed model was evaluated using various measures, including specificity, recall, precision, F1-score, Mathew's Correlation Coefficient (MCC), and accuracy. The proposed model GastroFuse-Net exhibited exceptional performance, achieving a precision of 0.985, recall of 0.985, specificity of 0.984, F1-score of 0.997, MCC of 0.982, and an accuracy of 98.5%.

Keywords: Deep learning; Convolutional Neural Network; medical image analysis; computer-aided diagnosis; artificial intelligence

1. Introduction

Endoscopic examinations are widely recognized as the traditional method for identifying gastrointestinal troubles, owing to their well-established efficacy. During an endoscopy, a patient's inner organs are meticulously examined to identify any underlying issues, enabling the medical team to determine the most effective course of treatment based on the patient's symptoms. An endoscope, a bendy tube with a fixed camera at one end and a light source attached, is used in this approach. The camera takes pictures of the organs, that are analyzed in greater detail. Different types of endoscopy procedures exist, depending on the purpose of the examination of the structures being observed and the equipment used. Endoscopes can be introduced into the body through a surgical incision, the mouth, or the esophagus [1].

The human digestive system is renowned for its diverse array of mucosal traits, presenting a broad spectrum of illnesses ranging from minor maladies to potentially life-threatening diseases. Given that there are 3.5 million reported cases and over 2.2 million deaths related to malignancies each year globally [2], it is essential to prioritize accurate and timely diagnosis for effective treatment and reducing illness and mortality rates [3,4]. Therefore, it is crucial to enhance the performance of clinical examinations and implement systematic screening strategies.

Computer-assisted automatic diagnosis is a recent subject of studies that shows the potential to revolutionize healthcare systems and scientific practice. Within the healthcare domain, deep convolutional neural networks (DCNNs) constitute a promising area to explore, as they possess the capability to aid medical experts in turning in remarkable care on a massive scale. This has been substantiated via numerous studies [5–9]. A vast volume of advanced image data, coupled with superior algorithms, enables the effective utilization of a DCNN-based system in the identification of gastrointestinal lesions inside endoscopic images [10–12]. The accurate and timely identification of ailments holds paramount importance because it directly influences remedy planning and patient monitoring [13–15].

Over 15 years, there has been an active search amongst researchers to analyze the usage of computer algorithms for the purpose of detecting problems within the human gastrointestinal machine via the examination of endoscopic images [16–18]. Recent clinical studies have proven that deep learning models have promising outcomes in identifying irregularities, especially within the realm of gastrointestinal polyp detection [19]. Nevertheless, the diagnostic efficacy of these models is significantly reliant on the quantity and quality of the data at hand.

This research employs the Kvasir dataset [20], which incorporates 8000 endoscopic images meticulously annotated by proficient endoscopists. The dataset consists of 8 unique classes, containing anatomical landmarks, medical outcomes, and gastrointestinal endoscopic approaches. The key goal of this study is to design a revolutionary algorithm based on Convolutional Neural Networks (CNN) tailored for the automated classification of anomalies inside the gastrointestinal system across multiple classes. The proposed deep learning framework aspires to correctly classify endoscopic images of the gastrointestinal tract into various categories with minimal preprocessing and optimized augmentation strategies. This specialized architecture demonstrates the potential to enhance the performance of the multi-class classification procedure, outperforming gastroenterologists in terms of precision and results. The results show that the proposed deep learning-based methodology can help gastroenterologists in categorizing gastrointestinal problems. The following are the contributions to this research:

- Designing a simple yet effective shallow CNN model with a careful choice of fewer layers to save computational resources and maintain competitive endoscopic images.
- Designing a deep CNN model with more layers to learn intricate patterns associated with various gastrointestinal conditions. This hierarchical feature learning significantly boosted the model's ability to discriminate and identify complex patterns.
- Introducing a comprehensive CNN-based architecture (called GastroFuse-Net) that combines knowledge from the designed shallow and deep CNN architectures in a novel way using feature concatenation. This method best utilizes the deep model's proficiency in complex feature extraction and the shallow model's capacity to collect wide context, producing a synergistic and all-encompassing classification framework.
- Assessing the performance of the proposed models using precision, recall, F1-score, Mathew's correlation coefficient (MCC), and accuracy metrics on the test dataset. These evaluations demonstrate the intricate balance between model complexity and performance, offering a comprehensive comprehension of the consequences of layer depth in CNN architectures for classifying gastrointestinal abnormalities.

The remaining sections of the paper are organized in the following manner: In Section 2, we conduct a thorough literature analysis focusing on deep learning models employed for classifying gastrointestinal abnormalities using endoscopy images. In Section 3, we detail the dataset used in this study, outlining its characteristics and the specific preprocessing steps undertaken. Additionally, the proposed novel architectures tailored to the task of gastrointestinal abnormality classification have been explained in this section. Section 4 is dedicated to evaluating the performance of our proposed models based on precision, recall, F1-score, Matthew's correlation coefficient, and accuracy. The research is concluded in Section 5 by summarizing the essential findings and insights derived from the experiments.

2. Related work

A number of countries are making an investment resource into AI (Artificial Intelligence) research right now, with an exclusive emphasis on computer vision and deep learning [21–23]. However, getting medical data might be complicated because of restricted availability imposed by means of regulatory constraints and a scarcity of human expertise needed for manually labeled training data. These limits make building computerized examination structures difficult. In response, distinct

strategies for automating the identification of anomalies inside the human gastrointestinal tract using machine learning and endoscopic images have been developed by the research groups. The time-consuming nature of manually examining large sets of gastric images, necessitating professional knowledge, has been highlighted in recent literature reviews on gastrointestinal anomalies [24,25]. To tackle this trouble, it is viable to develop AI-powered diagnostic tools that may automate the interpretation of large endoscopic data. Existing machine learning and deep learning techniques are presented in the subsequent sub-sections.

2.1. Existing machine learning techniques

In [26], the authors proposed a methodology by hybridizing Haralick texture features and Local Binary Pattern visual descriptors to classify gastrointestinal abnormalities from endoscopic images. The researchers carried out individual training of logistic regression models for each variable and used an ensemble approach to make the very last prediction. This technique obtained an accuracy of 94% with an F1-score of 0.76 and an MCC of 0.73 on testing data. In [27], a computer-aided detection system was designed to reduce missed polyps during video endoscopy. In this method, color wavelet features from endoscopic images were used to train a Support Vector Machine (SVM) classifier. The obtained results demonstrated the effectiveness of this technique with an accuracy of 98.34%, sensitivity of 98.67%, and specificity of 98.23%. In [28], gastrointestinal abnormalities were detected and classified using a variety of machine learning classifiers including logistic regression, decision tree, naïve bayes, SVM, and random forest. Among these methods, logistic regression emerged as the top performer. In [24], the authors detected gastric abnormalities using an SVM classifier achieving an accuracy of 0.86. Another study [29] identified multiple gastrointestinal disease detection using Linear Discriminant Analysis (LDA) and achieved an accuracy of 0.91, precision of 0.87, recall of 0.85, and F1-score of 0.86.

2.2. Existing deep learning techniques

In a study [30], the authors generated function vectors from geometric patterns extracted from images with the use of Inception-v3 and VGG16 networks, and the SVM classifier was trained on those extracted features. Their proposed technique correctly identified anomalies of the gastrointestinal system in endoscopic images with an MCC of 0.826. Another computational model stimulated from the Inception-v1 network to identify gastrointestinal issues and anatomical landmarks correctly was proposed [31]. The model consisted of convolutional layers of various sizes and pooling layers. The researchers employed an extensive range of data augmentation procedures to optimize performance. However, their method faced challenges in distinguishing between Dyed Lifted Polys and Dyed Resection Margins, in addition to discerning between classes of polyps and ulcerative colitis.

Researchers [32] proposed a ResNet-50 architecture to facilitate the automated analysis of endoscopic data. This approach took into account the interrelationships among the extracted variables. In a recent work [33], NASNet, Inception-v4, and Inception-ResNet-v2 architectures have been used to identify anatomical landmarks and detect sick tissue inside the human gastrointestinal system. Preprocessing techniques were additionally used to improve the quality of the images, resulting in an MCC of 0.93. An attention-based model was introduced in a separate study [34] to divide endoscopic images into four categories. During the subsequent phase of this study, an anomaly identification

methodology was employed to discern atypical types inside the initial stage. Researchers [35] presented a series of picture preparation procedures, which are then followed by the implementation of the EWT (Empirical Wavelet Transform). The decomposed images were then used as input for the proposed CNN model for disease classification at two levels. The results indicated an accuracy of 96.65% and MCC of 0.9298 in the first level of classification. In the second level of classification, 94.25% accuracy and 0.810 MCC was achieved. Researchers [36] presented the classification of gastrointestinal tract disorders using different transfer learning models. The best results were achieved using EfficientNetB0 with 98.01% accuracy, a precision of 98%, and a recall of 98%.

We present the GastroFuse-Net model, a new CNN-based architecture designed specifically for automatically classifying anomalies in the gastrointestinal system. This model is innovative and differs from previous studies. GastroFuse-Net combines the knowledge from both shallow and deep CNN architectures by merging their features through concatenation. This innovative strategy maximizes the deep model's ability to extract intricate features and the shallow model's capability to capture extensive context. The objective is to address the constraints identified in previous research, such as difficulties in differentiating between distinct gastrointestinal categories and enhancing the overall accuracy of classification. GastroFuse-Net effectively combines the advantages of shallow and deep models to provide a more complete and synergistic approach to classifying gastrointestinal abnormalities.

3. Material and methods

A deep learning-based solution has been proposed in this study for the automatic diagnosis of gastrointestinal abnormalities using deep learning. The proposed methodology carried out for this research is shown in Figure 1. Details of each block have been discussed in the sub-sections below.

3.1. Dataset description

The dataset utilized in this investigation was acquired from the Kvasir dataset [20]. There are eight distinct types of images stored in the Kvasir Database with 1000 images belonging to each category. Additional classifications of these photos include three for anatomical landmarks, three for pathological stages, and two for lesion eradication. To classify anatomical features, we have the pylorus, the z-line, and the cecum. The pylorus is the area of anatomy that spans the gap between the small intestine's first segment and the stomach. The z-line represents the anatomical boundary, where the esophagus transitions into the stomach. Examining this particular landmark holds significance in disease identification, as the manifestation of Esophagitis is commonly observed at this site. The cecum serves as the initial segment of the large intestine, and the achievement of reaching this anatomical marker is regarded as the culmination of a colonoscopy procedure. Esophagitis, polyps, and ulcerative colitis are the three disease-related conditions. Esophagitis is a medical condition characterized by inflammation of the esophagus, leading to the development of a mucosal rupture at the z-line. Ulcerative colitis is a pathological condition characterized by inflammation of the colon, specifically the large intestine. Polyps are aberrant growths that develop within the large intestine and can potentially progress into a precancerous state. Each of these categories represents a crucial aspect that a gastroenterologist examines during an endoscopic procedure. Two distinct categories are associated with the excision of lesions, namely dyed-lifted polyps and dyed resection margins.
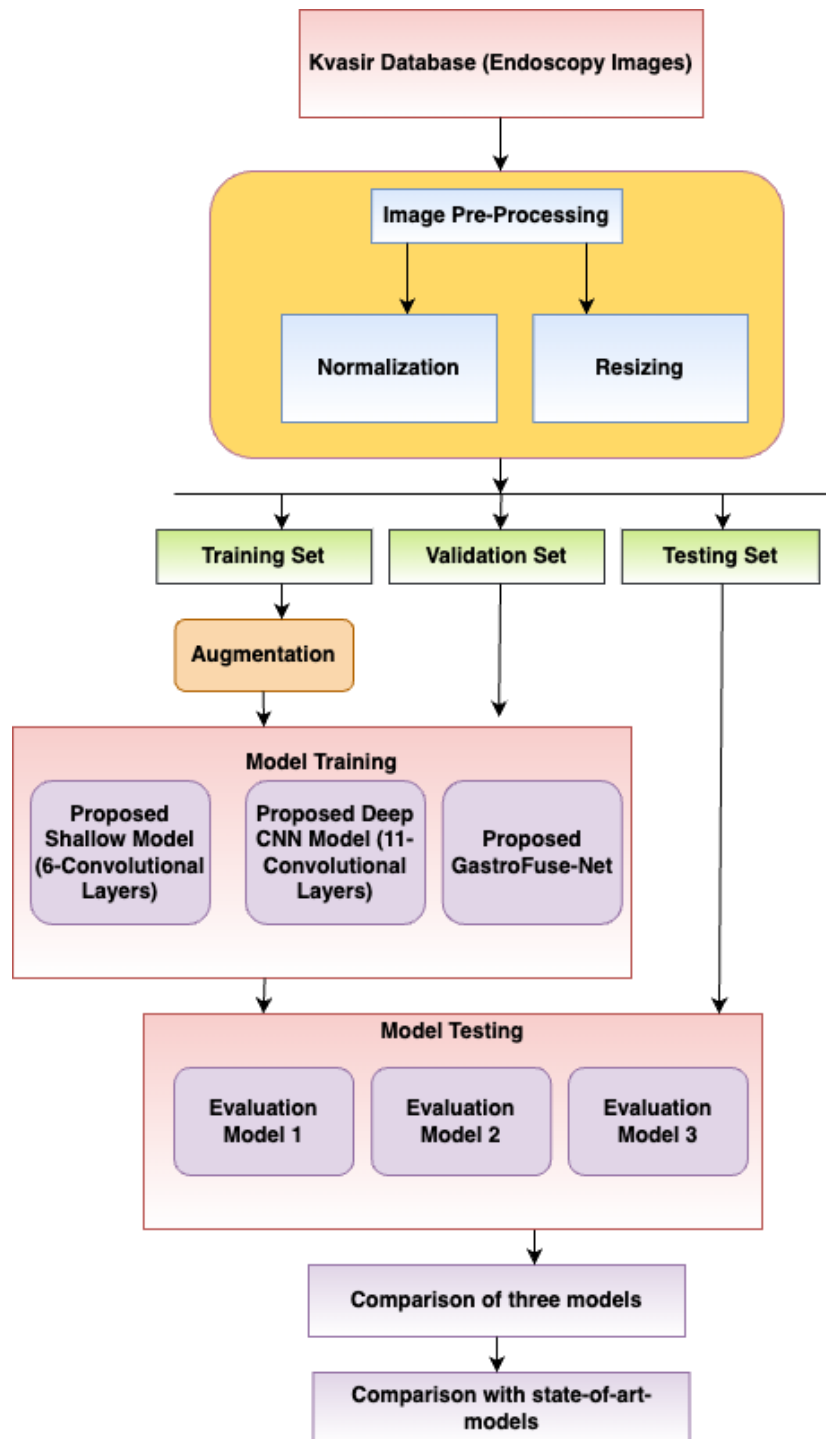
Figure 1. Proposed Methodology.

A dye is used for polyp removal to enhance the visibility of the polyp, and a procedure called "lifting" is used to separate polyps from the surrounding tissue. Before removing a polyp, images are taken of it that have been dyed and lifted, and afterward, the resection margins are stained. The dataset exhibits a range of image resolutions from 720x576 pixels to 1920x1072 pixels. Figure 2 displays image samples for each class in the dataset. Before training the model, the dataset was partitioned into three sets: Training, Test set, and Validation set. A total of 80% of the data were devoted to the training set, while the remaining 20% were reserved for the test set. A subset containing 10% was allocated for

validation among the total training samples. Table 1 provides the distribution of sample sizes across different classes within each data set.

Table 1. Dataset description with Train, Test and Validation Splits.

Class Name	Total Images in the dataset	Train Set	Augmented Train Set	Validation Set	Test Set
Ceum	1000	720	5760	80	200
Pylorus	1000	720	5760	80	200
Z-line	1000	720	5760	80	200
Esophagitis	1000	720	5760	80	200
Ulcerative Colitis	1000	720	5760	80	200
Polyps	1000	720	5760	80	200
Dyed Lifted Polyps	1000	720	5760	80	200
Dyed Resection Margins	1000	720	5760	80	200
Total Images	**8000**	**5760**	**46,080**	**640**	**1600**
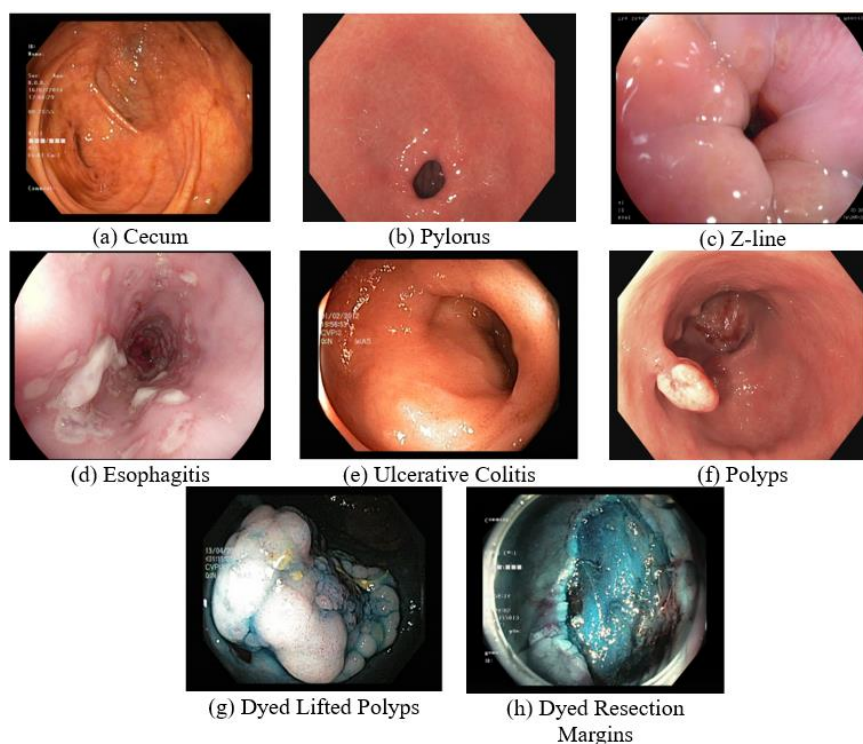
Figure 2. Sample Images from the dataset: (a) Cecum, (b) Pylorus, (c) Z-line, (d) Esophagitis, (e) Ulcerative Colitis, (f) Polyps, (g) Dyed Lifted Polyps, and (h) Dyed Resection Margins

3.2. Data pre-processing

Effective data preparation is crucial for enhancing the efficiency of machine learning and computer vision applications that depend on visual input. Image normalization and scaling are essential stages in this procedure. Image normalization transforms pixel values into a standardized scale from 0

to 1. Normalization is performed to ensure consistent intensity levels across all pictures. This stage is critical because it reduces the potential dominance of some images, that could arise from their more extensive pixel range, during the training process.

Moreover, the normalization procedure is essential in promoting convergence throughout the training of a model. After normalizing the images, they were uniformly reduced to dimensions of 256 by 256 pixels. Image scaling guarantees the standardization of the input size for the model. Resizing helps decrease computational complexity, hence improving the model's capacity to learn more efficiently from the given data. Using the methods of image normalization and scaling to a size of 256 by 256, the preprocessed dataset is efficiently optimized for subsequent machine-learning tasks.

3.3. Data augmentation

Data augmentation techniques are applied to the images within the training set for each class to address the problems associated with data insufficiency and overfitting, raising the training set's image count effectively as a result. Regarding the proficient and accurate implementation of data augmentation, many data transformations have been employed, including vertical and horizontal flipping, rotation at angles of 30, 45, and 60 degrees in a clockwise direction, a shear range of 0.2 factor, cropping by a factor of 0.2, and adjusting brightness by a value of 0.2. During the training process, on-the-fly image augmentation operations were performed. All the data transformations were employed simultaneously for the images to diversify the augmented dataset comprehensively while preserving the original image integrity and variability. The method of enhancing images has been achieved via the OpenCV library. Figure 3 depicts the visual representations of the original and augmented images pertaining to the Ulcerative Colitis and Polyps classes.
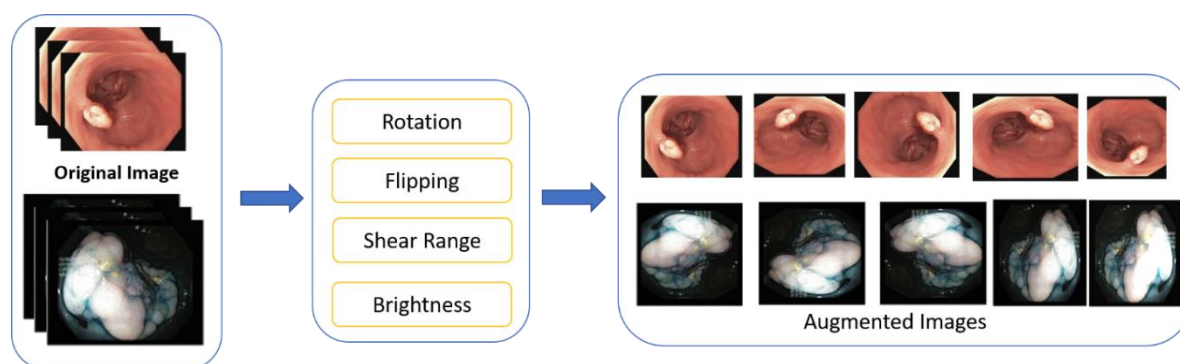
Figure 3. Sample Data Transformations on Training Set.

3.4. Proposed CNN-based architectures

We proposed a CNN-based architecture for identifying gastrointestinal abnormalities, a challenging task even for experienced gastroenterologists. Initially, two CNN-based architectures are presented, namely deep CNN and shallow CNN. These networks are designed to capture different aspects of the input gastrointestinal images and are independently trained on the dataset.

Deep CNN is responsible for capturing intricate features and learning hierarchical representations of the input images. It is composed of four convolutional blocks as shown in Figure 4(a). Three convolutional layers and a max-pooling layer make up each block. Convolutional layers are used for

feature extraction and its operation is given by Eq (1).

$$Z[i,j] = \sum_{m=0}^{F-1} \sum_{n=0}^{F-1} X[i+m, j+n].W[m,n] + b \tag{1}$$

where $Z[i, j]$ is the output feature map as position (i, j), X is the input layer, W is the convolutional filter, F is the filter size and b is the bias term. Convolutional layers pose a challenge due to the generation of a large number of neurons. To address this issue, CNN models employ pooling layers. These layers aggregate groups of pixels into a single representative pixel. Specifically, max pooling layers identify sets of pixels and replace them with the maximum value within that set, as described in Eq (2).

$$P_{max}(i.j) = max_{m,n} X[((i-1)s + m, (j-1)s + n)] \tag{2}$$

where $Pmax(i.j)$ represents the output of maxpooling operation at position (i, j) and s is the stride. Also, a batch normalization layer is applied after each convolutional layer to improve training stability and convergence. Mathematically, batch normalization is given by Eqs (3) and (4):

$$\hat{X}_{i,j} = \frac{X_{i,j} - \mu}{\sqrt{\sigma^2 + \epsilon}} \tag{3}$$

$$Y_{i,j} = \gamma \hat{X}_{i,j} + \beta \tag{4}$$

where $\hat{X}_{i,j}$ is the normalized output, $X_{i,j}$ is the input, μ is the mean, σ^2 is the variance, ϵ is a small constant for numerical stability, γ is a learnable scale parameter, β is a learnable shift parameter, and $Y_{i,j}$ is the final output after batch normalization. Batch normalization ensures that the input to subsequent layers is normalized, facilitating more stable and efficient training of deep CNNs. As shown in Figure 4(a), the filter size is increased by a multiple of 2 in each block, starting with 32 filters and a 3x3 kernel size in the first block doubling the number of filters in subsequent blocks. Following the four convolutional blocks, a global average pooling layer is applied to obtain a fixed-length feature vector. The global average pooling layer computes the average value of each feature map across the entire spatial dimensions. Mathematically, for a given feature map $\bar{X}$ with dimensions H X W X C, the global average pooling operation is defined by Eq (5):

$$Y_c = \frac{1}{H \, X \, W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{i,j,c} \tag{5}$$

The vector obtained from global average pooling layer is fed into a fully connected network of two dense layers. The first dense layer has 256 neurons, and the second dense layer has eight neurons, representing each class of gastrointestinal abnormality. This fully connected network is responsible for performing the final classification.

On the other hand, as shown in Figure 4(b), the shallow CNN focuses on extracting more local and low-level features from the input images. It consists of three convolutional blocks, each containing two convolutional layers followed by a max-pooling layer. To prevent overfitting, a dropout layer is applied after each max-pooling layer. The dropout layer introduces regularization during training by randomly deactivating a fraction of neurons, and it is given by Eq (6).

$$Dropout(x) = \begin{cases} x \ X \ \frac{1}{1-p}, & with \ probability \ 1 - p \\ 0, & with \ probability \ p \end{cases} \quad (6)$$

where p is the dropout probability, and the layer randomly sets a fraction of input values to zero during each forward pass. The kernel size for each convolutional layer is set to 3x3, and the number of filters increases from 32 in the first block to 64 in the second block and then to 86 in the third block, as shown in Figure 4(b).
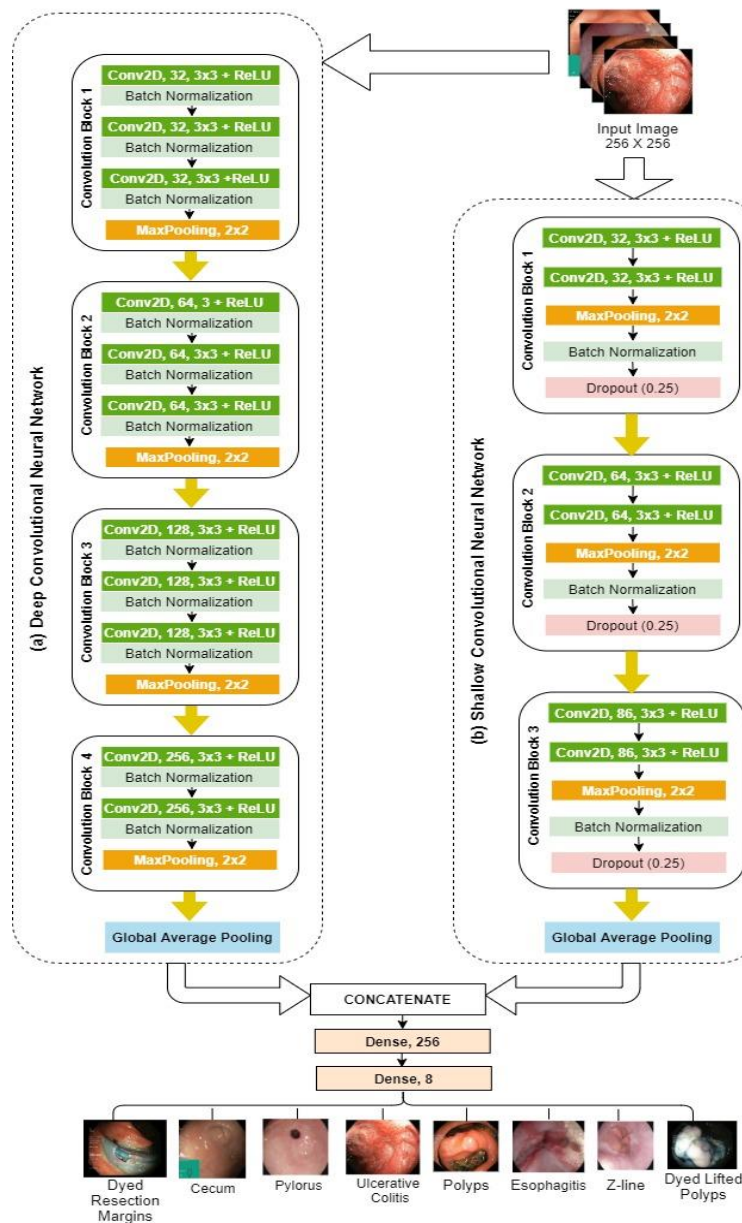
Figure 4. Proposed architecture of GastroFuse-Net consisting of (a) Deep Convolutional Neural Network (Deep CNN) and (b) Shallow Convolutional Neural Network (Shallow CNN).

After obtaining results from two networks, we performed feature fusion by concatenating the extracted features from both networks to leverage the complementary characteristics of the deep and shallow CNNs. Specifically, features from the penultimate layer of the deep CNN were extracted and

concatenated with the corresponding features from the shallow CNN. This process combines the high-level semantic information learned by the deep CNN with the local and low-level details captured by the shallow CNN. After feature fusion, the combined feature set was classified using an artificial neural network, resulting in a novel architecture called GastroFuse-Net. Its integrated feature representation enables the model to exploit richer features, enhancing its ability to differentiate between gastrointestinal abnormalities. The formal definition of GastroFuse-Net is given by Algorithm 1.

Algorithm 1: GastroFuse-Net

1: **Begin:**
2: **Load and preprocess the dataset:**
    Load Kvasir dataset with 8000 images and their corresponding labels
    Preprocess the Dataset (normalize, resize, and augment)
    Split the dataset into Train, Test and Validation data
3: **Data Augmentation:**
    Apply data tranformations (rotation, flip, sheer range, and brightness) on train data
4: **Define deep CNN and shallow CNN architecture**
    Define architecture for deep CNN
    Include convolutional, max-pooling, and batch normalization layers
    Define architecture for shallow CNN
    Include convolutional layers, max-pooling, dropout for regularization.
5: **Feature fusion using Gastro-FuseNet**
    Extract penultimate layer outputs from both Deep CNN and Shallow CNN
    Concatenate features to obtain fused feature representation
    Create an Artificial Neural Network (ANN) for final classification.
6: **Training GastroFuse-Net**
    Initialize hyperparameters, epochs, and batch size.
    for epoch in range(total_epochs):
        Initialize an empty list to store fused features and corresponding labels.
        for batch in training_batches:
            Load batch of preprocessed images and labels.
            Extract features from Deep CNN and Shallow CNN
            Concatenate features to get fused features.
            Append fused features and labels to the list.
        Convert the list to NumPy arrays.
        Train GastroFuse-Net using the fused features and labels
    End Tranining Loop
7: **Evaluate Performance**
    Evaluate the trained GastroFuse-Net on the Test set.
    Generate confusion matrix to analyze model predictions.
    Calculate accuracy, precision, recall, specificity, F1-score, and MCC.
8: **End**

3.5. Experimental setup

The proposed network for classifying gastrointestinal abnormalities was trained using the backpropagation method with the gradient descent algorithm to optimize biases and weights. The convolutional layers were applied with ReLU activation function for introducing non-linearity and capturing complex patterns, while softmax was used for multi-class classification in the output layer. The Adam optimizer with a fixed learning rate of 1e-3 was employed for 50 epochs with 32 batch size. The categorical cross-entropy loss function was used for training to minimize dissimilarity between predicted probabilities and actual labels. Categorical cross entropy is given by Eq (7):

$$L(y, \hat{y}) = -\sum_i y_i . \log(\hat{y}_i) \tag{7}$$

where y represents the true label and $\hat{y}$ represents the predicted label and the sum is taken over all classes. A validation set was utilized to optimize the hyperparameters, adjusting learning rate, batch size, and other relevant parameters for best performance. The implementation of experiments was done using Keras and TensorFlow on an NVIDIA GeForce GTX 1080 graphics card for computational power. The finalized model's performance was evaluated on a separate test set which provided a fair assessment of the model's ability to generalize and classify gastrointestinal abnormalities on unseen data.

3.6. Performance metrics

Different performance criteria have been used to evaluate how well a model can identify gastrointestinal problems from endoscopic pictures. The model's predictions are broken down in depth in the confusion matrix that has been obtained, which is the first achievement. The matrix enables the generation of other indicators for performance, and it offers insights into the types of errors generated by models and by displaying true positives, true negatives, false positives, and false negatives.

The confusion matrix was used to construct some indicators for performance which incorporates Matthews Correlation Coefficient (MCC), accuracy, precision, recall, and specificity. By dividing the total number of correctly classified cases by the size of the test set, accuracy can be understood as the overall correctness of the predictions. The formula of accuracy is given by Eq (8).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

where TP, TN, FP and FN are true positives, true negatives, false positives and false negatives obtained from the confusion matrix. Moreover, accuracy may not be adequate when dealing with imbalanced datasets or classes of varying importance. By calculating the accurate percentage for predicted positive instances relative to all instances predicted as positive, precision measures the model's capacity to prevent false positives. Precision is given by Eq (9).

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

Similarly, recall determines how effectively the model can detect positive instances by determining what positive examples percentage are true positives. Recall is given by Eq (10).

$$Recall = \frac{TP}{TP+FN} \qquad (10)$$

When dealing with imbalanced datasets or trying to lessen false positives and negatives, the F1 score is a good balanced measure because it combines recall and accuracy. The F1-score is given by Eq (11).

$$F1 - score = \frac{2*Precision*Recall}{Precision+Recall} \qquad (11)$$

To evaluate the model's accuracy for avoiding false positives in the negative class, specificity was used in addition to recall. Specificity is given by Eq (12)

$$Specificity = \frac{TN}{TN+FP} \qquad (12)$$

Finally, when it comes to classification problems, Matthews Correlation Coefficient (MCC) shows a balanced performance metric that takes all four metrics into account. MCC is given by Eq (13).

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \qquad (13)$$

4. Results and discussions

In this section, the performance of the proposed architectures is discussed. Accuracy plots, loss plots and confusion matrix along with other performance parameters for each model are presented in subsequent sections.
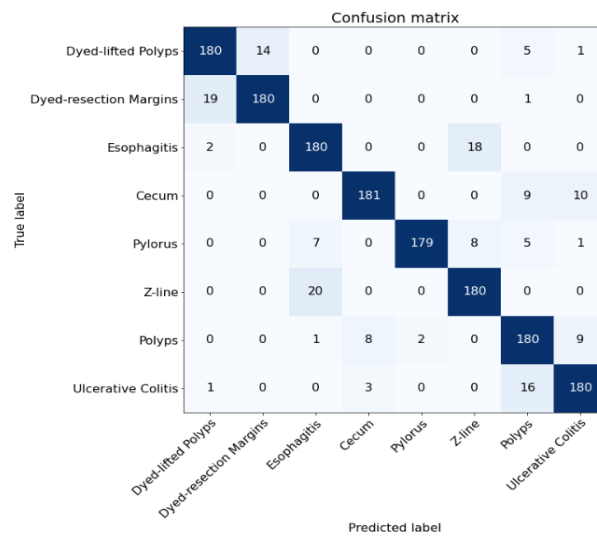
4.1. Performance of shallow CNN

The variations in accuracy and loss over 50 epochs during the training of the proposed shallow CNN is shown in Figures 5(a) and 5(b), respectively. It can be seen in these figures that there are continuously increasing and decreasing fluctuations in both training and validation accuracy and loss. Despite these fluctuations, an overall positive trend of increasing accuracy and decreasing loss indicates that the model is learning and improving over time.

The confusion matrix obtained on the test data set from Shallow CNN is shown in Figure 6 and the performance metrics obtained from it for each class are presented in Table 2. The results demonstrate that the model achieved an impressive overall accuracy of 90.3%, indicating its proficiency in correctly classifying most samples. Notably, the model performed exceptionally well in accurately identifying samples from the Cecum and Pylorus classes. However, some misclassifications were observed where samples from these classes were incorrectly categorized as either Polyps or Ulcerative Colitis.

Additionally, the classes Dyed Lifted Polyps and Dyed Resection Margins were frequently misclassified interchangeably. Furthermore, the class Z-line showed a considerable number of misclassifications, predominantly being classified as Esophagitis. These findings emphasize the model's strengths in distinguishing certain classes effectively. However, it also highlights areas, where the model may require further improvement, especially in differentiating between classes that share similar features.

(a)   (b)
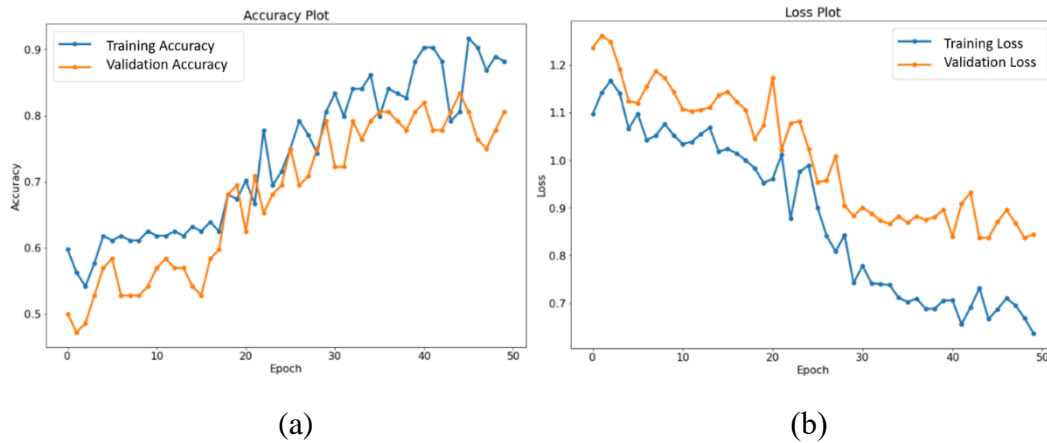
Figure 5. Plots obtained while training of Shallow CNN (a) Accuracy Plot and, (b) Loss Plot.
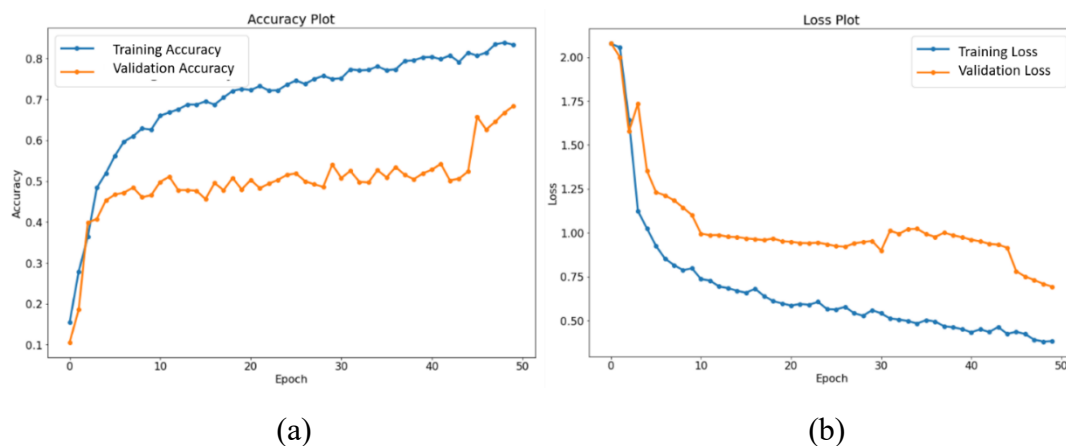

Figure 6. Confusion matrix of the Shallow CNN model.

Table 2. Performance Parameters obtained using Shallow CNN.

Class Name	Precision	Recall	F1-score	Specificity	MCC
Dyed Lifted Polyps	0.891	0.900	0.895	0.984	0.880
Dyed Resection Margins	0.927	0.900	0.913	0.990	0.901
Esophagitis	0.865	0.900	0.882	0.980	0.865
Cecum	0.942	0.905	0.923	0.992	0.913
Pylorus	0.989	0.895	0.939	0.998	0.933
Z-line	0.873	0.900	0.886	0.981	0.870
Polyps	0.833	0.900	0.865	0.974	0.846
Ulcerative Colitis	0.895	0.900	0.897	0.985	0.883
Average	**0.902**	**0.900**	**0.900**	**0.986**	**0.886**

4.2. Performance of Deep CNN

Figures 7(a) and 7(b) illustrate the changes in accuracy and loss during the training of a deep CNN model over 50 epochs. A notable observation is that the deep CNN model exhibits much lower training and validation accuracy and loss fluctuations than the previously evaluated shallow CNN model. However, it is worth noting that after approximately ten epochs, the validation loss and training loss seem to plateau, showing minor fluctuations around a similar value. This observation suggests that the model might have reached a saturation point where further training does not significantly improve its performance. Hence, this could indicate that additional model enhancement techniques could be explored to achieve even better results.

(a)   (b)

Figure 7. Plots obtained while training of Deep CNN (a) Accuracy Plot and, (b) Loss Plot.

Figure 8. Confusion matrix of Deep CNN model.

The confusion matrix of the deep CNN model and its corresponding performance metrics are presented in Table 3 and Figure 8, demonstrating its impressive classification performance. The deep CNN model improved significantly, correctly classifying 96.2% of the samples compared to the

shallow CNN. Analyzing the confusion matrix, we observe that, similar to the shallow CNN, the class "Z-line" has been misclassified as "Esophagitis". However, the number of misclassified samples has reduced, indicating the deep CNN's enhanced ability to discriminate between these classes. Additionally, while some instances of "Dyed Resection Margins" were misclassified as "Dyed Lifted Polyps", the frequency of misclassifications is lower than in the shallow CNN model. The considerable accuracy gain achieved by the deep CNN model and its reduced misclassification errors indicate its superior performance in handling complex patterns and improving overall class separability.

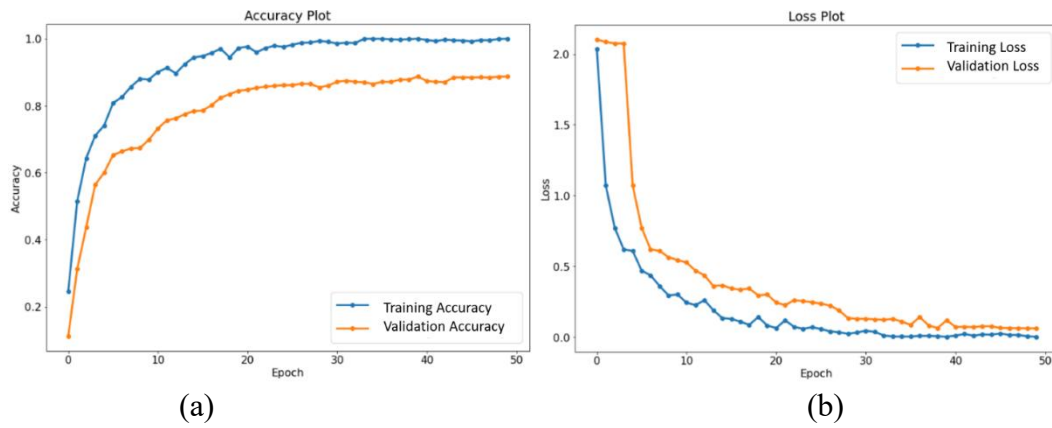**Table 3.** Performance Parameters were obtained using Deep CNN.

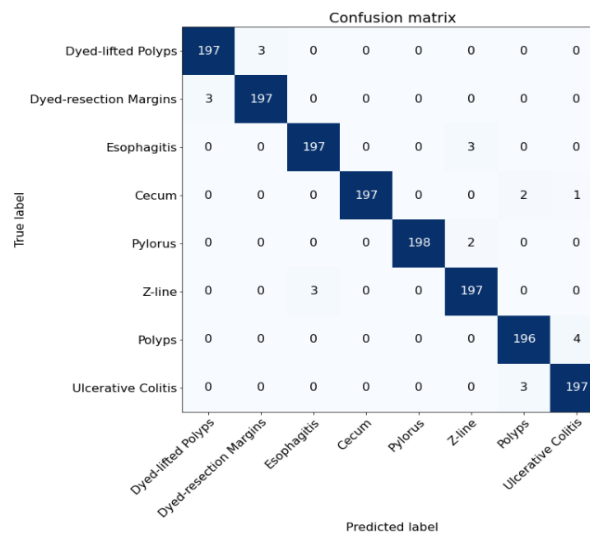| Class Name	Precision	Recall	F1-score	Specificity	MCC
Dyed Lifted Polyps	0.955	0.960	0.957	0.993	0.951
Dyed Resection Margins	0.979	0.960	0.969	0.997	0.965
Esophagitis	0.927	0.960	0.943	0.989	0.935
Cecum	0.989	0.960	0.974	0.998	0.971
Pylorus	0.990	0.990	0.990	0.998	0.988
Z-line	0.949	0.930	0.939	0.992	0.930
Polyps	0.936	0.960	0.948	0.990	0.940
Ulcerative Colitis	0.955	0.960	0.957	0.993	0.951
Average	**0.960**	**0.960**	**0.960**	**0.994**	**0.954**

4.3. Performance of GastroFuse-Net

Figures 9(a) and 9(b) show the change in accuracy and loss during training of GastroFuse-Net. The results demonstrate a smooth and steady progression of accuracy, with a maximum training accuracy of 99.9% and a validation accuracy of 82.4%, which indicates that the model effectively learned from the combined features and achieved remarkable accuracy, surpassing the individual networks. Moreover, the training loss for GastroFuse-Net showed a significant decrease to 0.0342, while the validation loss decreased to 0.235. This decline in loss signifies that the model has gained a better understanding of the underlying patterns in the data, reinforcing its ability to make precise predictions.

The confusion matrix and performance metrics obtained for GastroFuse-Net, as shown in Figure 10 and Table 4, reveal its improved performance, correctly classifying 98.5% of the samples, surpassing the accuracy achieved by both Deep and Shallow CNN models. However, some challenges persist, particularly in distinguishing between the Polyps and Ulcerative Colitis classes, where a few samples are misclassified interchangeably. Additionally, there were three instances, where samples of Esophagitis were incorrectly predicted as Z-line class. Furthermore, minimal misclassifications were observed between the Dyed Lifted Polyps and Dyes Resection Margins classes. Despite these minor misclassifications, GastroFuse-Net demonstrates significant progress in accurately classifying most samples, showcasing its potential for enhanced performance over individual CNN models. These findings underscore the effectiveness of feature fusion in improving classification accuracy and highlight the importance of further addressing specific misclassification patterns to enhance the model's capabilities.

Figure 9. Plots obtained while training of GastroFuse-Net (a) Accuracy Plot and, (b) Loss Plot.

Figure 10. Confusion matrix of GastroFuse-Net.

Table 4. Performance Parameters of GastroFuse-Net.

Class Name	Precision	Recall	F1-score	Specificity	MCC
Dyed Lifted Polyps	0.985	0.985	0.985	0.997	0.982
Dyed Resection Margins	0.985	0.985	0.985	0.997	0.982
Esophagitis	0.985	0.985	0.985	0.997	0.982
Cecum	1.000	0.985	0.992	1.000	0.991
Pylorus	1.000	0.990	0.995	1.000	0.994
Z-line	0.975	0.985	0.980	0.996	0.977
Polyps	0.975	0.980	0.977	0.996	0.974
Ulcerative Colitis	0.975	0.985	0.980	0.996	0.977
Average	**0.985**	**0.985**	**0.984**	**0.997**	**0.982**

4.4. Comparison with the state-of-the-art

The proposed model has been compared with the state-of-the-art shown in Table 5. The proposed model achieved the highest precision of 0.98, recall of 0.98, F1-score of 0.99, specificity of 1.00, MCC of 0.98, and accuracy of 0.985 as compared to the other models in the literature.

Table 5. State-of-the-art comparison of the proposed model.

| Reference	Precision	Recall	F1-score	Specificity	MCC	Accuracy
[30]	0.8475	0.84	0.8475	0.97	0.82	0.96
[20]	0.82	0.82	0.82	0.97	0.80	0.95
[37]	0.97	0.97	0.97	-	-	0.973
[32]	0.89	0.89	0.89	0.99	0.87	0.984
[38]	0.96	0.96	0.96	-	-	0.96
[39]	-	-	-	-	-	0.97
[15]	0.94	-	0.93	-	-	0.94
[35]	-	-	-	-	0.9298	0.96
[36]	0.98	0.98	-	-	-	0.9801
Proposed Shallow CNN	0.902	0.900	0.900	0.986	0.886	0.903
Proposed Deep CNN	0.960	0.960	0.960	0.994	0.954	0.962
Proposed GastroFuse-Net	**0.985**	**0.985**	**0.984**	**0.997**	**0.982**	**0.985**

5. Discussions

The results obtained demonstrate that the proposed models are effective in the classification of gastrointestinal abnormalities. It is worth noting that GastroFuse-Net achieved a substantially higher accuracy rate of 98.5% than both Deep CNN and Shallow CNN. The comparison of the performance parameters obtained for each proposed model is shown in Figure 11. It can be seen in this figure that with feature fusion in GastroFuse-Net, the capability of the model to classify the images increased. For the shallow CNN model, accuracy of 0.9, MCC of 0.89, F1-score of 0.9, and specificity of 0.98. recall of 0.9 and precision of 0.9 are achieved. By increasing layers in the Deep CNN model, accuracy increased to 0.96, MCC to 0.95, F1-score to 0.96, specificity to 0.99, recall to 0.96 and precision to 0.96. However, when the combined features of both models were fused in GastroFuse-Net, the best results were obtained with accuracy increasing to 0.985, MCC increased to 0.98, F1-score increased to 0.9, specificity to 1.00, recall to 0.98 and precision increased to 0.99.

It is worth mentioning that GastroFuse-Net demonstrated an exceptional capacity to differentiate between classes, as indicated by its elevated precision, recall, and specificity. The enhanced performance of GastroFuse-Net can be explained by the strategic concatenation of features, which effectively merges the advantageous attributes of shallow and deep models. The Shallow CNN exhibited competence in comprehending broad contexts, whereas the Deep CNN demonstrated an exceptional performance in extracting intricate features.

Furthermore, a runtime analysis showing the amount of time needed to train each suggested model is given in Table 6. The Deep CNN model had a training duration of 22,926.2 seconds (about 6.4 hours), the Shallow CNN model took around 10,383.4 seconds (approximately 2.88 hours), and GastroFuse-Net had a training time of roughly 16,028.1 seconds (approximately 4.45 hours). Although GastroFuse-

Net demonstrated improved performance, it reached this level of accuracy within a training duration that falls between the more time-consuming Deep CNN model and the quicker Shallow CNN model. This highlights efficacy of GastroFuse-Net as it provides a convincing trade-off between excellent classification accuracy and comparatively short training times in contrast to Deep CNN and Shallow CNN.
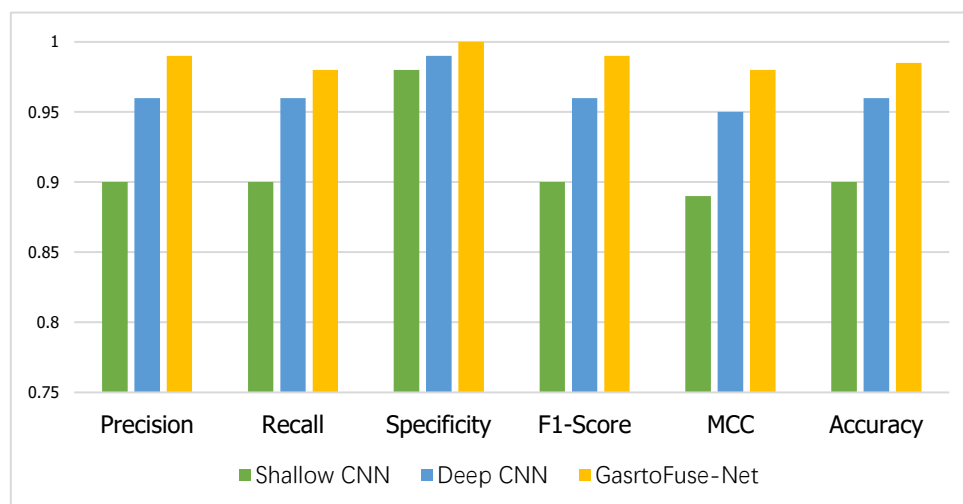
Figure 11. Performance comparison of the proposed models.

Table 6. Run Time analysis of different proposed models.

Model Name	Time taken to train the model
Deep CNN	22,926.2 s
Shallow CNN	10,383.4 s
GastroFuse-Net	16,028.1 s

6. Conclusions

Deep learning has demonstrated its efficacy as a beneficial tool for analyzing medical images in gastroenterology. Medical practitioners frequently depend on their specialized knowledge and years of practice to analyze images, resulting in subjective assessments and possible delays in diagnosing. Deep learning systems can address these constraints by offering unbiased and efficient categorizations of gastrointestinal disorders from endoscopic images.

Prompt and specific diagnosis is critical in medical practice, as it could substantially affect patient results. Mistakes in diagnosis can result in misguided treatments and extended hospital stays. Employing computer-aided diagnosis algorithms can lower the load of gastroenterologists and mitigate the ability for misdiagnosis. Accurate identification of gastrointestinal ailments, inclusive of colon cancer and polyps, is mainly essential because of their great effect on world health.

The GastroFuse-Net is a data-driven deep convolutional neural network (DCNN) method developed to deal with the difficulty of detecting gastrointestinal issues. The model's feature fusion method, which integrates the advantage of shallow and deep CNN models, achieved exceptional performance. It correctly categorized 98.5% of the samples with high precision, recall, specificity, F1-score, and Matthews Correlation Coefficient (MCC) values of 0.985, 0.985, 0.984, 0.997, and 0.982, respectively.

Although the outcomes are promising, there are possibilities for future enhancement and research. Possible improvements include expanding the training dataset, examining higher-decision pictures, and modifying the model for use in different medical imaging domains. Gastric endoscopists might consider the suggested algorithm a second opinion, particularly when it is difficult to get specialized medical support. GastroFuse-Net demonstrates promising abilities in helping healthcare practitioners in the appropriate and quick diagnosis of gastrointestinal issues. Additional clinical data and the exploration of progressive algorithms will improve the successful incorporation of this technology into clinical exercise.

Use of AI tools declaration

The authors have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgment

Conflict of interest

Frank Werner, the Guest Editor for the journal of MBE, was not involved in the editorial review or the decision to publish this article. All authors declare that there is no conflict of interest.

References

1. H. Brenner, M. Kloor, C. P. Pox, Colorectal cancer, *Lancet*, **383**, (2014), 1490–1502. https://doi.org/10.1016/S0140-6736(13)61649-9

2. M. F. Kaminski, J. Regula, E. Kraszewska, M. Polkowski, U. Wojciechowska, J. Didkowska, et al., Quality indicators for colonoscopy and the risk of interval cancer, *N. Engl. J. Med.*, **362** (2010), 1795–1803. https://doi.org/10.1056/nejmoa0907667

3. T. Takahashi, Y. Saikawa, Y. Kitagawa, Gastric cancer: current status of diagnosis and treatment, *Cancers (Basel)*, **5** (2013), 48–63. https://doi.org/10.3390/cancers5010048

4. T. Yada, C. Yokoi, N. Uemura, The current state of diagnosis and treatment for early gastric cancer, *Diagn. Ther. Endosc.*, **2013** (2013), 24132. https://doi.org/10.1155/2013/241320

5. A. Shokouhifar, M. Shokouhifar, M. Sabbaghian, H. Soltanian-Zadeh, Swarm intelligence empowered three-stage ensemble deep learning for arm volume measurement in patients with lymphedema, *Biomed. Signal Process. Control*, **85** (2023), 105027. https://doi.org/10.1016/j.bspc.2023.105027

6. E. J. Topol, High-performance medicine: the convergence of human and artificial intelligence, *Nat. Med.*, **25** (2019), 44−56. https://doi.org/10.1038/s41591-018- 0300-7

7. N. Sharma, S. Gupta, A. Rajab, M. A. Elmagzoub, K. Rajab, A. Shaikh, Semantic Segmentation of Gastrointestinal Tract in MRI Scans Using PSPNet Model With ResNet34 Feature Encoding Network, *IEEE Access*, **11** (2023), 132532−132543. https://doi.org/10.1109/ACCESS.2023.3336862

8. N. Sharma, S. Gupta, D. Koundal, S. Alyami, H. Alshahrani, Y. Asiri, et al., U-Net model with transfer learning model as a backbone for segmentation of gastrointestinal tract, *Bioengineering (Basel)*, **10** (2023), 119. https://doi.org/10.3390/bioengineering10010119

9. J. Yang, M. Shokouhifar, L. Yee, A. A. Khan, M. Awais, Z. Mousavi, DT2F-TLNet: A novel text-independent writer identification and verification model using a combination of deep type-2 fuzzy architecture and Transfer Learning networks based on handwriting data, *Expert Syst. Appl.,* **242** (2024), 122704. https://doi.org/10.1016/j.eswa.2023.122704

10. A. Sharma, R. Kumar, P. Garg, Deep learning-based prediction model for diagnosing gastrointestinal diseases using endoscopy images, *Int. J. Med. Inform.*, **177** (2023), 105142. https://doi.org/10.1016/j.ijmedinf.2023.105142

11. V. Raut, R. Gunjan, V. V. Shete, U. D. Eknath, Gastrointestinal tract disease segmentation and classification in wireless capsule endoscopy using intelligent deep learning model, *Comput. Method. Biomec.*, **11** (2023), 606−622. https://doi.org/10.1080/21681163.2022.2099298

12. K. Zhang, Y. Zhang, Y. Ding, M. Wang, P. Bai, X. Wang, et al., Anatomical sites identification in both ordinary and capsule gastroduodenoscopy via deep learning, *Biomed. Signal Proces.*, **90** (2024), 105911. https://doi.org/10.1016/j.bspc.2023.105911

13. Y. Mori, S. E. Kudo, M. Misawa, Y. Saito, H. Ikematsu, K. Hotta, et al., Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: A prospective study, *Ann. Intern. Med.*, **169** (2018), 357–366. https://doi.org/10.7326/M18-0249

14. C. M. Hsu, C. C. Hsu, Z. M. Hsu, T. H. Chen, T. Kuo, Intraprocedure Artificial Intelligence Alert System for Colonoscopy Examination, *Sensors*, **23** (2023), 1211. https://doi.org/10.3390/s23031211

15. E. Young, L. Edwards, R. Singh, The Role of Artificial Intelligence in Colorectal Cancer Screening: Lesion Detection and Lesion Characterization, *Cancers*, **15** (2023), 5126. https://doi.org/10.3390/cancers15215126

16. D. K. Iakovidis, A. Koulaouzidis, Software for enhanced video capsule endoscopy: challenges for essential progress, *Nat. Rev. Gastroenterol. Hepatol.*, **12** (2015), 172–186. https://doi.org/10.1038/nrgastro.2015.13

17. S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, M. Tzivras, Computer-aided tumor detection in endoscopic video using color wavelet features, *IEEE Trans. Inf. Technol. Biomed.*, **7**, (2003), 141–152. https:// doi.org/10.1109/TITB.2003.813794

18. M. Liedlgruber, A. Uhl, Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review, *IEEE Rev. Biomed. Eng.*, **4** (2011), 73–88. https://doi.org/10.1109/RBME.2011.2175445

19. Y. Mori, K. Mori, Endoscopy: Computer-aided diagnostic system based on deep learning which supports endoscopists' decision-making on the treatment of colorectal polyps, In: Hashizume, M. (eds) *Multidisciplinary Computational Anatomy*. Springer, Singapore, 2022, 337–342. https://doi.org/10.1007/978-981-16-4325-5_45

20. K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, et al., Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection, in *Proceedings of the 8th ACM multimedia systems conference*, MMSys, (2017), 164–169. https://doi.org/10.1145/3083187.3083212

21. U. K. Lilhore, M. Poongodi, A. Kaur, S. Simaiya, A. D. Algarni, H. Elmannai, et al., Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques, *Comput. Math. Methods Med.*, **2022** (2022), 4688327. https://doi.org/10.1155/2022/4688327

22. W. S. Liew, T. B. Tang, C. H. Lin, C. K. Lu, Automatic colonic polyp detection using integration of modified deep residual convolutional neural network and ensemble learning approaches, *Comput. Meth. Prog. Biomed.*, **206** (2021), 106114. https://doi.org/10.1016/j.cmpb.2021.106114

23. C. M. Lo, Y. W. Yang, J. K. Lin, T. C. Lin, W. S. Chen, S. H. Yang, et al., Modeling the survival of colorectal cancer patients based on colonoscopic features in a feature ensemble vision transformer, *Comput. Med. Imag. Grap.*, **107** (2023), 102242. https://doi.org/10.1016/j.compmedimag.2023.102242

24. H. Ali, M. Sharif, M. Yasmin, M. H. Rehmani, F. Riaz, A survey of feature extraction and fusion of deep learning for detection of abnormalities in video endoscopy of gastrointestinal tract, *Artif. Intell. Rev.*, **53** (2020), 2635–2707. https://doi.org/10.1007/s10462-019-09743-2

25. D. Jha, S. Ali, S. Hicks, V. Thambawita, H. Borgli, P. Smedsrud, et al., A comprehensive analysis of classification methods in gastrointestinal endoscopy imaging, *Med. Image Anal.*, **70** (2021), 102007. https://doi.org/10.1016/j.media.2021.102007

26. S. S. A. Naqvi, S. Nadeem, M. Zaid, M. A. Tahir, Ensemble of texture features for finding abnormalities in the gastrointestinal tract, in *MediaEval Benchmarking Initiative for Multimedia Evaluation*, 2017. Available from: https://api.semanticscholar.org/CorpusID:6396180

27. M. Billah, S. Waheed, Gastrointestinal polyp detection in endoscopic images using an improved feature extraction method, *Biomed. Eng. Lett.*, **8** (2018), 69–75. https://doi.org/10.1007/s13534-017-0048-x

28. A. Rosenfeld, D. G. Graham, S. Jevons, J. Ariza, D. Hagan, A. Wilson, et al., Development and validation of a risk prediction model to diagnose Barrett's oesophagus (MARK-BE): a case-control machine learning approach, *Lancet Digit. Health*, **2** (2020), E37–E48. https://doi.org/10.1016/S2589-7500(19)30216-X

29. K. Kundu, S. A. Fattah, K. A. Wahid, Multiple linear discriminant models for extracting salient characteristic patterns in capsule endoscopy images for multi-disease detection, *IEEE J. Transl. Eng. Health Med.*, **8** (2020), 3300111. https://doi.org/10.1109/JTEHM.2020.2964666

30. T. Agrawal, R. Gupta, S. Sahu, C. E. Wilson, SCL-UMD at the medico task- mediaeval 2017: Transfer learning-based classification of medical images, *MediaEva*, **17** (2017), 13–15.

31. S. Petscharnig, K. Schoffmann, M. Lux, An inception-like CNN architecture for GI disease and anatomical landmark classification, *MediaEva*, **17** (2017), 13–15.

32. H. Gammulle, S. Denman, S. Sridharan, C. Fookes, Two-stream deep feature modelling for automated video endoscopy data analysis, In *Medical Image Computing and Computer Assisted Intervention* – MICCAI 2020: 23rd International Conference, Lima, Peru, October 4-8, 2020, Proceedings, Part III 23. Springer International Publishing, 2020, 742–751. https://doi.org/10.1007/978-3- 030-59716-0_71

33. T. Cogan, M. Cogan, L. Tamil, MAPGI: Accurate identification of anatomical landmarks and diseased tissue in gastrointestinal tract using deep learning, *Comput. Biol. Med.*, **111** (2019), 103351. https://doi.org/10.1016/j. compbiomed.2019.103351

34. S. Jain, A. Seal, A. Ojha, A. Yazidi, J. Bures, I. Tacheci, et al., A deep CNN model for anomaly detection and localization in wireless capsule endoscopy images, *Comput. Biol. Med.*, **137** (2021), 104789. https://doi.org/ 10.1016/j.media.2021.102007

35. S. Mohapatra, G. Kumar Pati, M. Mishra, T. Swarnkar, Gastrointestinal abnormality detection and classification using empirical wavelet transform and deep convolutional neural network from endoscopic images, *Ain Shams Eng. J.*, **14** (2023), 101942. https://doi.org/10.1016/j.asej.2022.101942

36. T. Abraham, J. V. Muralidhar, A. Sathyarajasekaran, K. Ilakiyaselvan, A Deep-Learning Approach for Identifying and Classifying Digestive Diseases, *Symmetry*, **15** (2023), 379. https://doi.org/10.3390/sym15020379

37. C. Gamage, I. Wijesinghe, C. Chitraranjan, I. Perera, GI-Net: Anomalies classification in gastrointestinal tract through endoscopic imagery with deep learning, in *2019 Moratuwa Engineering Research Conference (MERCon)*, 2019. https://doi.org/10.1109/MERCon.2019.8818929

38. J. Yogapriya, V. Chandran, M. G. Sumithra, P. Anitha, P. Jenopaul, C. Suresh Gnana Dhas, Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model, *Comput. Math. Methods Med.*, **2021** (2021), 5940433. https://doi.org/10.1155/2021/ 5940433.

39. M. Khan, K. Muhammad, S. H. Wang, S. Alsubai, A. Binbusayyis, A. Alqahtani, et al., Gastrointestinal diseases recognition: a framework of deep neural network and improved moth-crow optimization with DCCA fusion, *Hum-Cent. Comput. Info. Sci.*, **12** (2022), 1–12. https:// doi.org/10.22967/HCIS.2022.12.025