ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

A Multi-Scale Attention Efficient Network (Mae-Net) For Gastrointestinal Disease Detection and Classification

M. Madhumalini[1,*]

[1]Associate professor, Department of Electronics and Communication Engineering, P. A. College of Engineering and Technology, Pollachi, India. madhumalini2007@gmail.com

ABSTRACT

Gastrointestinal diseases (GIDs) are increasingly prevalent worldwide, necessitating robust methods for their early detection and diagnosis. This study proposes a novel Multi-scale Attention Efficient Network (MAE-Net) for identifying abnormalities in the gastric region using time–frequency analysis. The Kvasir dataset is employed as the primary source of input images. The methodology begins with a preprocessing phase where the input images are resized, converted to grayscale, and enhanced for texture. This is followed by segmentation using approximate spatial fuzzy c-means clustering. Features are extracted using wavelet transform coefficients to analyze frequency and time series data. The decomposed image classes are then input into the MAE-Net model for training and testing, enabling accurate classification and prediction. The performance of the proposed model is evaluated using standard metrics such as accuracy, precision, recall, and F1 score. Comparative analysis with existing methods on similar datasets demonstrates the superiority of the proposed approach in terms of classification performance. This work highlights the potential of MAE-Net in advancing the detection and diagnosis of gastrointestinal abnormalities.

Keywords: Gasto Inestinal Disease (GID), Adaptive Gabor filter, spatial fuzzy c-means clustering, Wavelet transform, Multi-scale Attention Efficient Network (MAE-Net)

1 Introduction

Gastro Intestinal Disease (GID) is one of the main causes of the worldwide disease problem and it is reported as nearly about 8 million people death generally every year [1]. Disease of the gastrointestinal system, containing the esophagus, liver, stomach, small and large intestines, gallbladder, and pancreas, is known as GID. Diseases like Gastro Esophageal Reflux Disease (GERD), cancer, lactose intolerance, irritable bowel syndrome, and hiatal hernia are common in GIDs. Bleeding, bloating, constipation, diarrhea, heartburn, nausea, pain, and vomiting are the common symptoms GID [2]. Most of the digesting and nutritional absorption takes place in the small bowel, also known as the small intestine and it is a long, coiled tube. In this section of the digestive tract the GI is the challenging task to recognize and evaluate gastrointestinal problems [3].Whereas, for patients with limited access, Wireless Capsule Endoscopy (WCE) is a simple way to detect infections such ulcers and polyps in the Gastro Intestinal Tract (GIT) regions without requiring surgery [4].Nowadays, GID is found in a number of places where AI systems are trained to evaluate histological diagnoses, analyse endoscopic and radiographic GI images, and distinguish benign from malignant lesions [5,6]. Additionally, Machine Learning (ML) methods enable computers to learn directly from data, and they are frequently employed in gastrointestinal surgery to analyze large amounts of data [7].

In the framework of Machine Learning, Support Vector Machines (SVM) is widely used in GI image processing. The drawback is it is not suitable for large datasets. In contrast, the data shows that CNN-based supervised learning object detection models are frequently used in Deep Learning (DL) for GI image analysis [8]. However the ResNet-50 model uses a novel deep semantic segmentation model termed deep labv3 as a foundation for 3D-segmentation of the various forms of stomach infections in GI [9].

Additionally, a CNN and a level set method are used in an integrated approach to separate polyps. For surface and form analysis, a modified level set technique called the Local Gradient Weighting embedded Level Set Method (LGWe-LSM) is presented, it destroys high-intensity regions that lead to false positives [10]. And then a 16-layer deep CNN is used to identify polyps in colonoscopy images inside GID image [11].

Moreover, Deep CNN architecture is used that accurately use endoscopic images to detect gastrointestinal disorders in humans. To determine the effectiveness and performance, it is meticulously built using many routes, different image resolutions, and multiple convolutional layers [12]. To overcome these, the classification layer of the AlexNet, GoogleNet, and ResNet topologies uses a classifier based on Long Short-Term Memory (LSTM). To determine the GID, the ImageNet dataset is used to train the LSTM-based CNN models [13]. And also the GI transit time of the pill in the small intestine region and gastric by combining a Long Short-Term Memory (LSTM) with a Convolutional Neural Network (CNN) that is intended to recognize the sequence information of continuous videos [14]. The limitations of Video Capsule Endoscopy (VCE) include incomplete small-bowel transit and low image quality [15].

The contributions of the proposed method are as follows:

- Developed a novel MAE-Net model to enhance the classification of gastrointestinal abnormalities, leveraging advanced attention mechanisms for improved feature extraction and efficiency.
- Designed an effective preprocessing pipeline, including resizing, grayscale conversion, texture enhancement, and segmentation using approximate spatial fuzzy c-means clustering.
- Utilized wavelet transform coefficients to extract frequency and time-series features from gastrointestinal images, enabling detailed and multi-dimensional analysis.
- Demonstrated the efficiency of the suggested method by employing the Kvasir dataset, a widely recognized benchmark for gastrointestinal endoscopy images.

2 Related Work

Cambay *et al* (2024) [16] have proposed a novel Convolutional Neural Network (CNN) called ResNet50 to use endoscopic images and a novel ResNet50-based deep feature engineering model to detect a variety of gastrointestinal disorders.. Gradient-Weighted Class Activation Mapping (Grad-CAM) is used to identify heat maps during feature extraction. The pre-trained ResNet50's final global average pooling layer generates features from these locations. And a kvasir dataset is used to utilize the GID image. However, the value of gradient drops significantly during the propagation in the layers.

Lonseko *et al* (2021) [17] have introduced a deep CNN-based spatial attention mechanism with encoder and decoder layers for classification in endoscopic images. The encoder and decoders of the trained GI image and the potential of attention-guided networks for GI disease is to find the interpretability and classification accuracy. For gastrointestinal disorders an attention mechanism is used to build a reliable computer-aided diagnosis. The drawback of this method is it needs large amount of labelled data to calculate and it takes long time to train the input data.

Khan *et al* (2020) [18] have presented a Mask Region based Convolutional Neural Network (Mask-RCNN) and deep CNN feature for the GI image's ulcer segmentation and gastrointestinal infection classification. The regions are find using the region of interest (ROI) and featured using the pre-trained CNN model ResNet101. To distinguish their GI, the features were enhanced using the grasshopper optimization approach. For the final step the predicted GI disease is fed into a Multi-class Support Vector Machine (MSVM). However, it is not suitable for the real time application.

Xia ***et al* (2021) [19]** have proposed a DL model called Convolutional Neural Network (CNN) and faster Region-based Convolutional Neural Network (RCNN) to find a method that detects stomach lesions automatically to help with diagnosis and inter-physician variances. The GI regions are found using the region based CNN their regions are splits as layers. Then gastric lesions and diagnosis are detected using the RCNN and CNN. However, it takes more time to process the single image.

Haile et al (2022) [20] have invented a neural network model called CNN model by concatenating features of VGGNet and InceptionNet for the classification of multiple gastrointestinal diseases from endoscopy images. Using data from the HyperKvasir database, the CNN model is trained and evaluated. The extracted structures of VGGNet and InceptionNet are given to the SVM to kernel the GD for the classification purpose. Multiple GI disease is classified using CNN technique. However, Training period is computationally costly and demands a large amount of processing power.

3 Proposed Work

MAE-Net is a new DL-based model for medical imaging analysis that uses progressive learning to efficiently extract features from Kvasir dataset for GID prediction. Here, the input

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

GID image is resized for ulilizing the image. Then the grayscale and color texture of the input GID is textured by using Adaptive Gabor filters an appropriate pre-processing method. Subsequently, the essential ROIs of the pre-processed image is segmented individually with the aid of Spatial Fuzzy C-means clustering segmentation method, following which Wavelet Transform is passed out for optimum feature extraction here the time frequency of the horizontal, vertical and diagonals are extracted. The finest feature subset is then consequently selected using the suggested Multi-scale Attention Efficient Net classifier is used for classifying GID.

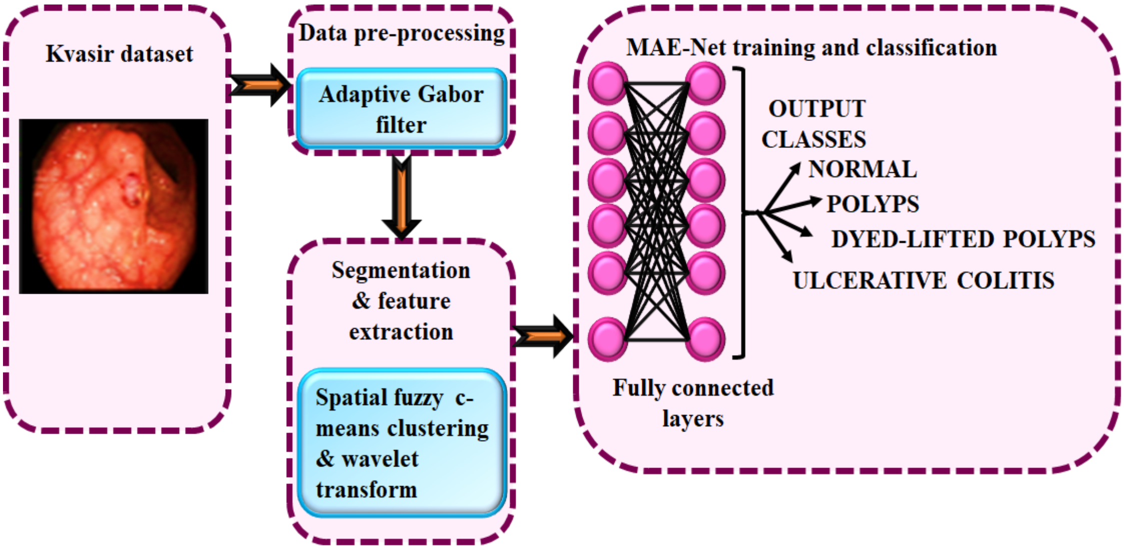

Figure 1: *Block diagram of proposed work*

3.1 Data Pre-processing using Adaptive Gabor Filter (AGF) for GID prediction

Data pre-processing is the initial stage, in this work this filter is used to resize and to texture the GID using AGF. A convolution filter with Gaussian and sinusoidal terms in GID is called the Adaptive Gabor Filter. The best characteristics in GID that enable multi-resolution analysis to extract the optimal GID attributes in the frequency and spatial domains are AGF. Whereas, the sine component provides the directionality, the Gaussian component creates the weights. The 2D Gabor filter described in this paper it is represented by the following equations:

$$g = \exp\left(-\frac{x'^2+\acute{\gamma}^2 y'^2}{2\sigma^2}\right)\exp(i\ (2\pi\frac{x'}{\lambda}+\psi)) \tag{1}$$

$$g_{re} = \exp\left(-\frac{x'^2+\acute{\gamma}^2 y'^2}{2\sigma^2}\right)\cos(\ 2\pi\frac{x'}{\lambda}+\psi) \tag{2}$$

Where,

$$x' = x\cos\theta + y\sin\theta\ ,y' = -x\sin\theta + y\cos\theta \tag{3}$$

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

The real part g_{re} of the Gabor function is utilized for the practical implementation of GID prediction. The wavelength of the sinusoidal is represented by λ in the Gabor function, where $f = \frac{1}{\lambda}$ and f is the sinusoidal frequency. The Gabor-function support's ellipticity is specified by θ, which is the normal orientation to the parallel stripesψ, which is the phase offset, θ which is the Gaussian envelope's standard deviation and γ which is the spatial aspect ratio. The value of ψis often between 0 and π, the range of $\theta$is between 0 and π, and γ typically takes the value of 1. Followed by pre-processing the input image is given to the segmentation process to segment the GID image.

3.2 Segmentation using Spatial Fuzzy C-Means Clustering for GID prediction

Segmentation is the second stage here the pre-processed input GID image is segmented by Spatial Fuzzy C-Means clustering. A soft grouping method known as Spatial Fuzzy C-Means (SFCM), here each GI point represents a member level of a fuzzy logic-based cluster. Fuzzy partitions are used by SFCM, thus several groups with varying membership values (weights) between 0 and 1 preserve a GID data. A member of the cluster will be closer to the center of the cluster if the GID data is closer to the cluster center. All of the objects' membership values must be equal to one. The Spatial Fuzzy C-Means clustering algorithm's fundamental steps are shown as follows:

- First, the GI data to be clustered is inserted as a $n \times m$ matrix. Next, the number of clusters $c(1 \leq c \leq N)$ is specified, along with the matrix partition, power$w$, maximum iteration, least anticipated error $\varepsilon > 0$, initial objective function $P_{o=}0$, and first iteration $t = 1$.
- The second step is to set the initial partition matrix element U to a random number.μ_{ik} Determine each column's number:

$$Q_i = \sum_{k=1}^{c} \mu_{ik} \tag{4}$$

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \tag{5}$$

- The third step is to determine the $k$cluster's center:V_{kj}

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^{w} * X_{ij})}{\sum_{i=1}^{n}(\mu_{ik})^{w}} \tag{6}$$

- Calculate the objective function in the $t$iteration as the fourth step.

$$P_t = \sum_{i=1}^{n}\sum_{i=1}^{c}([\sum_{j=1}^{m}(X_{ij} - V_{kj})^2 \mu_{ik})^{w}) \tag{7}$$

- Determine the matrix partitions which have been changed in fifth step.

$$\mu_{ik} = \frac{[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2]^{-1/(w-1)}}{\sum_{k=1}^{c}[\sum_{j=1}^{m}(X_{ij}-V_{ki})^2]^{-1/(w-1)}} \tag{8}$$

- Check the stop condition in the sixth step.

Stop if$([P_t - P_{t-1}] < \varepsilon)$; else, fourth stop should be repeated if$t = t + 1$.

For every GID data set, Spatial Fuzzy C-Means produces a cluster of center clusters and several degrees of membership. Followed by the segmentation the GID image is given to the feature extraction stage.

3.3 Feature extraction using Wavelet Transform (WT) for GID prediction

Feature extraction is the third stage here the segmented GID image is used as the input to extract the image using wavelet transform. In image analysis, the Wavelet Transform (WT) is a widely used time-frequency representation method. This feature of WT enables the extraction of GID is more important qualities throughout the feature extraction procedure. Therefore, 2D-DWT is taken into consideration in this work for the time-frequency frame representation of the input GI-tract images. Assume that the continuous square-integral function $f(x)$exists. For a real-valued mother wavelet ψ(x)the Continuous Wavelet Transform (CWT) of $f(x)$ is shown as

$$CWT_{\psi}(\propto, t) = \int_{-\infty}^{\infty} f(x)\, \psi_{\propto,t}(x) dx \tag{9}$$

Where,

$$\psi_{\propto,t}(x) = \frac{1}{\sqrt{\alpha}}\ \psi(\frac{x-t}{\alpha}) \tag{10}$$

Where the scaling and translation parameters are denoted by α and t. Following to the feature extraction the GID image is given to the classification their proposed MAE-Net is used.

3.4 Multi-scale Attention Efficient -Net (MAE-Net) for GID prediction

Finally the GID image is given to the propose method called novel network architecture named Multi-scale Attention Efficient-Net (MAE-Net) for GID prediction. In the MAE-Net, the self-attention mechanism is employed in GID. The modules of the suggested MAE-Net are said to accomplish enough feature extraction and synthesis to achieve exact classification. The following information refinement module then applies the attention process, which is primarily made possible by the built-in fused-and MSLIL-attention blocks for GID prediction. Specifically, the earlier is utilized to remove fine-grained info.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

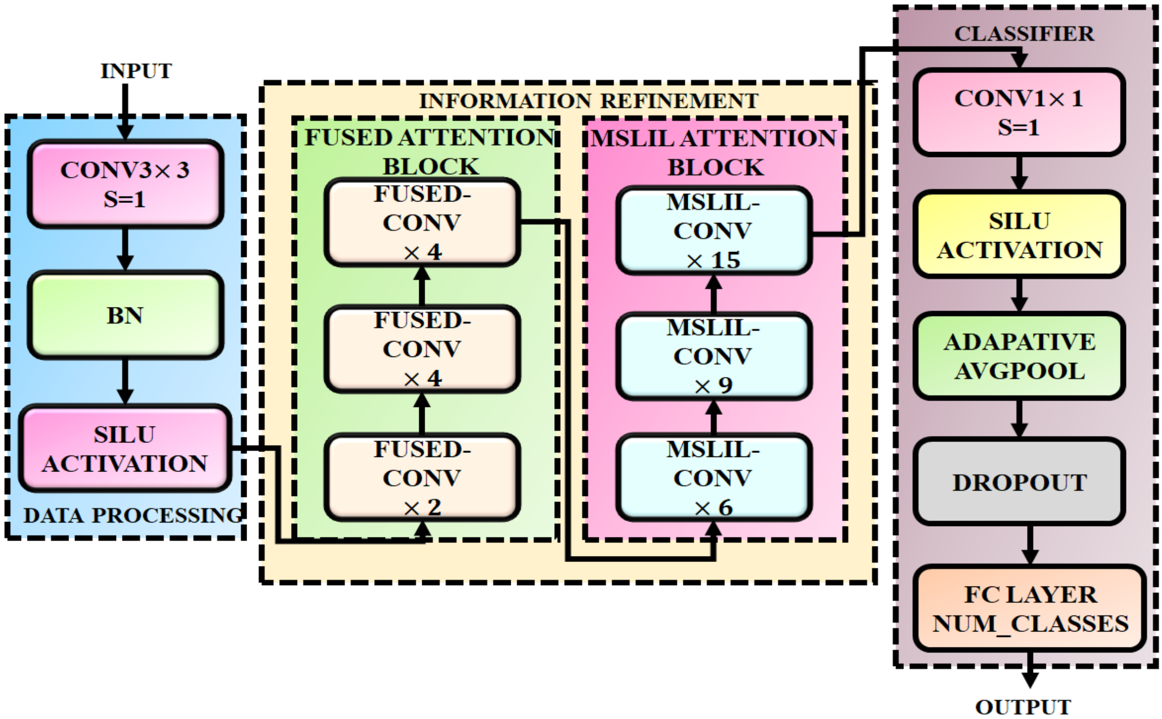

Figure 2: *Outline of the suggested Multi-scale Efficient Network*

3.4.1 Fused-attention block for GID prediction

Figure.3 illuminates the architecture of fused-attention block, which is made up of multiple repeating fused-convolution (Fused-Conv) modules. This module focuses on extracting specific information by implementing the Squeeze-Excitation (SE) attention mechanism for GID. This is because the input GID images of Fused-Conv have a lot of channels but are small in size. Initially, the overall average pooling is used to achieve spatial feature density. Weight coefficients ω then reflect the weight of each channel predicted by an FC layer on the compressed features for GID. The process of determining ω it is explained as follows:

$$s = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} I(i,j) \tag{11}$$

$$\omega = \sigma(W_2\delta(W_1 s)) \tag{12}$$

Where I the initial GID input & s is the compacted one, H and W are number of channels. ω and δrepresent initiation functions in the FC layer, whereas W_1 *and* W_2 are two matrices. Finally, the weight coefficient and original GID input are multiplied channel-wise to produce the output.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

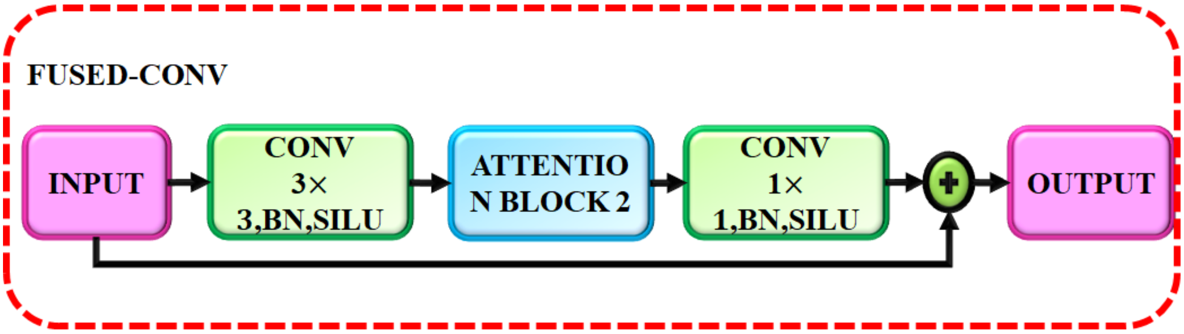

Figure 3: *Architecture of Fused-Attention block*

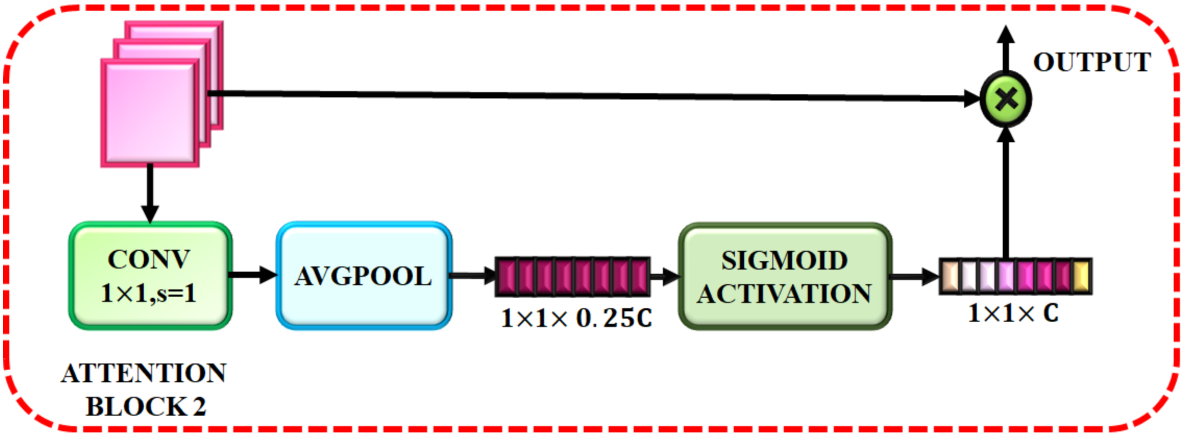

Figure 4: *Squeeze-Excitation (SE) Attention mechanism*

3.4.2 MSLIL-attention block for GID prediction

A new MSLIL-attention block is created in the suggested MAE-Net, it is made up of several MSLIL convolution (MSLIL-Conv) modules, to improve the extraction of semantic info by adding suitable structures for GID. In the MSLIL-Conv module, a multi-branch structure is implemented, with paralleled 2 dilated convolutions and a group convolution, as illustrated in Figure. 5.

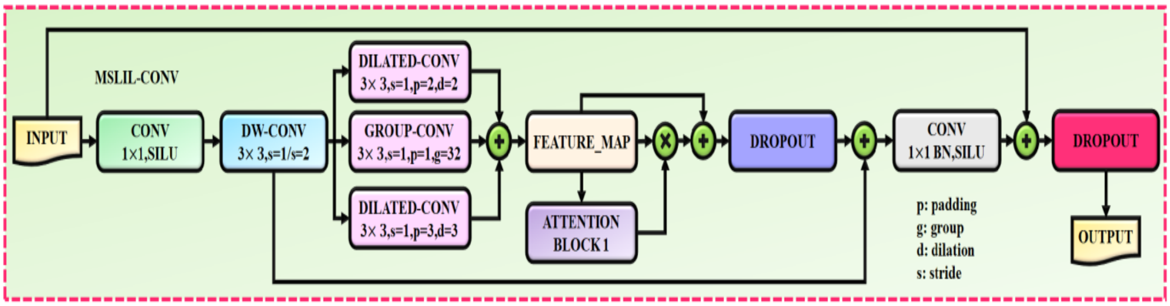

Figure 5: *MSLIL-attention block*

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

By offering rich global information, the dilated convolution that specifically compensate for the missing receptive fields, and the 3 × 3 group convolution for further lessen processing demands in GID. It is possible to merge the outputs GID of the three branches mentioned above and arrive the Efficient Channel Attention (ECA) tool by modifying the hyper-parameters of the convolution operators. Furthermore, to minimize the loss of crucial information, the output GID of the above ECA mechanism is multiplied by the integrated multi-branch output, which retains some of its original feature characteristics.

3.4.3 Efficient Channel Attention mechanism for GID

In the ECA mechanism for GID, the weights of $1D$ convolution are cross-channelly interleaved with one another. These weights are presented in groups, and the number of weights for each GID image group is determined by the size of the convolution kernel k. Only the interaction between y_i and its k neighboring channels is taken into description when calculating the weight of feature y_i, and the result is as follows:

$$\omega_i = \sigma(\textstyle\sum_{j=1}^{k} w^j y_i^j), y_i^j \in \Omega_i^k \tag{13}$$

Where w is weight of y_i &W Ω_i^k is the set of $k$neighboring networks of y_i. Furthermore, the MSLIL-attention blocks for GID prediction have two dropout layers that randomly eliminate 20% of the features in order to lower the computing complexity and avoid overfitting. Followed by the classification the classified output is given to the performance metrics.

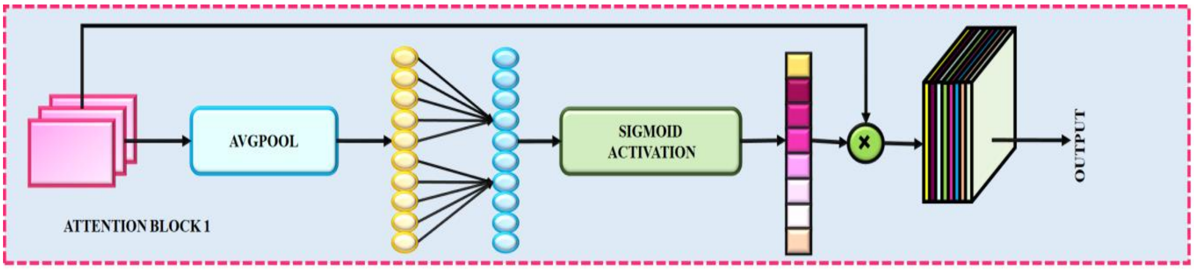

Figure 6: *ECA mechanism*

Performance metrics

Using various performance metrics the suggested outline to training data is evaluated. And these metrics are accuracy, precision, recall, F1 score and AUC. Using the following equation each metrics are defined as

Accuracy: From the predicted data accuracy metric measures the amount of appropriately recognised samples in (GID).

$$Accuracy = \frac{(TN+TP)}{TS} \tag{14}$$

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

Precision: By distributing TP through the amount of TP and False Positives (FP) computation is achieved. The ratio of accurately (TP) to the total number of predicted positive samples (GID) is called precision.

$$Precision = \frac{(TP)}{TP+FP} \tag{15}$$

Recall: The ratio of True Positive (TP) to the overall amount of patients with GID in the kvasir dataset and it is firm by separating TP by the amount of TP and False nNgatives (FN).

$$recall = \frac{(TP)}{(TP+FN)} \tag{16}$$

F1 score: A combined measure of precision and recall is called as F1 score.

$$F1 - score = \frac{(2*\text{Precision}*\text{Recall})}{(Precision + Recall)} \tag{17}$$

At last, the model's performance across different classification thresholds is evaluated using the AUC, which describes the correlation between the true positive and false positive rates for GID prediction.

4 Result and discussion

The suggested detection model categorizes the certain input kvasir dataset for GID prediction in to eight different classes namely dyed-lifted-polyps, dyed-resection-margins, esophagitis, normal-cecum, normal-pylorus, normal-z-line, polyps and ulcerative-colitis. Furthermore, testing in Python applications helps to understand the detection model's capabilities.

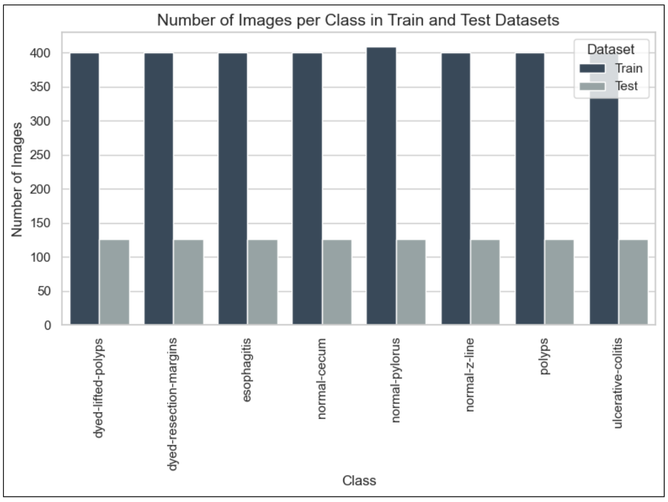

Figure 7: *Number of images per class in test and train datasets*

In figure 7, the number of images per class in test and train dataset is 500 datasets for the specific type of gastrointestinal disease prediction. The eight classes are tested and trained using the techniques.

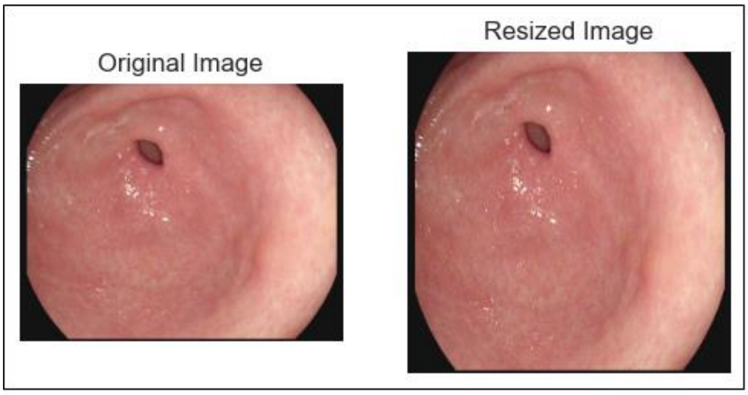

Figure 8: *Resized GID image*

In Figure 8 the original GID image is resized for further pre-processing technique is utilized. When using the resizing technique a clear view of input resized image is captured.

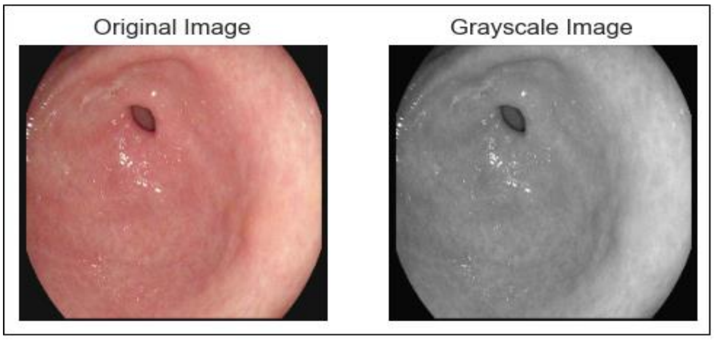

Figure 9: *Gray scale GID image*

Two steps of pre-processing for gastrointestinal images are depicted in the provided image in Figure 9. To reduce the computing complexity and concentrate on the most important elements for analysis, the original endoscopic image is changed to grayscale. Converting images to grayscale preserves important structural details while restructuring image processing.

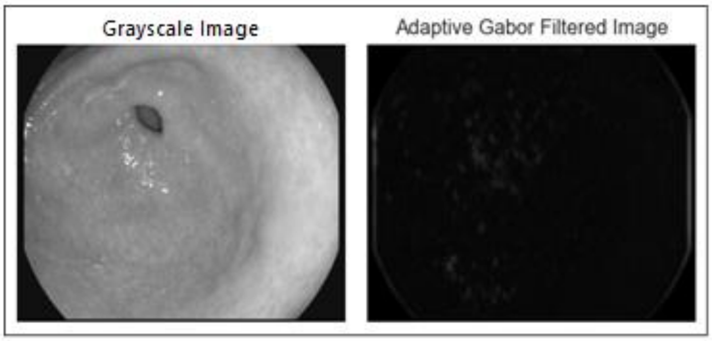

Figure 10: *Adaptive Gabor Filter of GID*

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

After conversion of gray scale image, adaptive gabor filter is utilized to preprocess the GID image to texture enhancement which shows the clear view of the GID image.

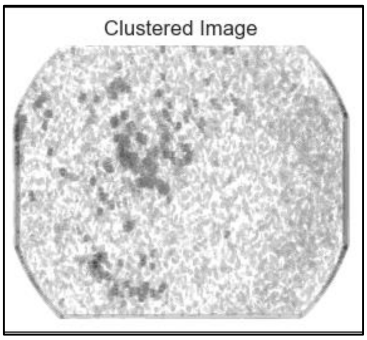

Figure 11: Spatial Fuzzy C-means clustering of GID

Figure 11 illustrates the spatial fuzzy c-means clustering applied to a gastrointestinal disease (GID) image. This method segments the image by grouping similar pixel intensities, emphasizing regions with potential abnormalities. The clustering enhances the focus on key areas for further feature extraction and analysis.

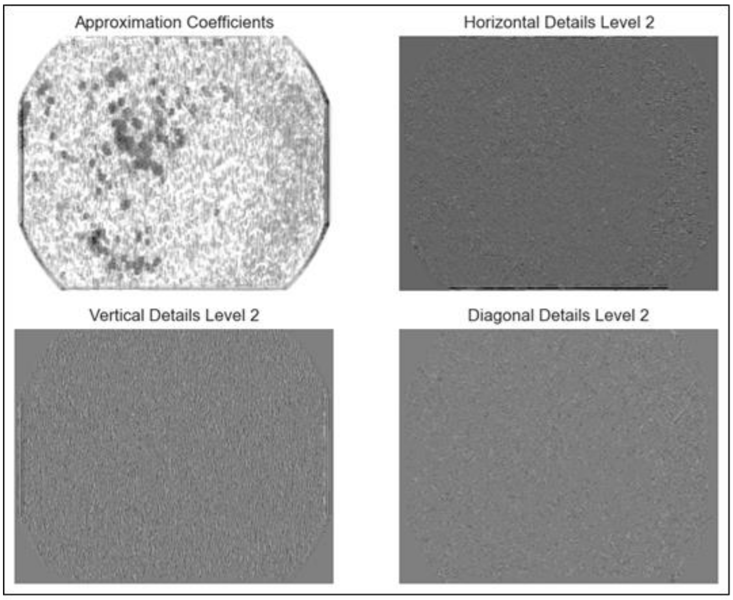

Figure 12: *Wavelet transform for GID*

Figure 12 demonstrates the wavelet transform applied GID image. The transform decomposes the clustered image into multiple components: approximation coefficients and detailed coefficients (horizontal, vertical, and diagonal) at Level 2. This decomposition highlights specific patterns and features, enabling a detailed time–frequency analysis crucial for effective feature extraction and classification.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

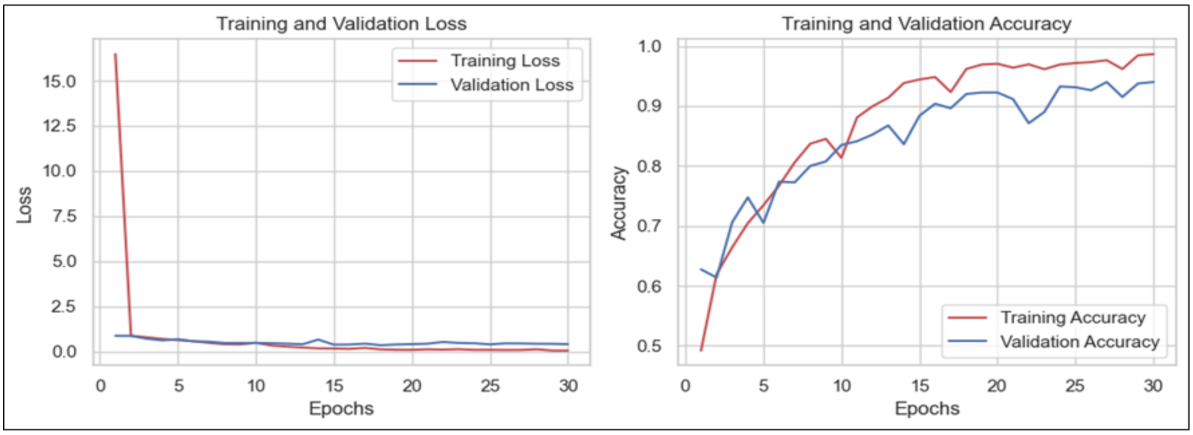

Figure 13: *Training and validation for loss and accuracy*

A loss value of zero is obtained from a perfect prediction, whereas greater loss values are obtained from poor predictions. The ratio of a model's accurate predictions to its total number of predictions is known as model accuracy. The accuracy for this method is 94%.The model's loss is demonstrated, which also shows loss value variations and the accuracy of the model is shown graphically in Figure 13.

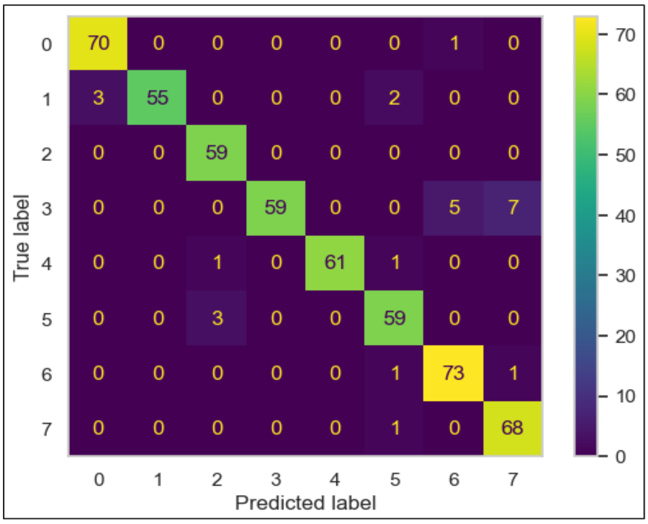

Figure 14: *Confusion matrix for MAE-Net*

Figure 14 illustrates the confusion matrix for MAE-Net. This matrix shows the values for true and predicted labels. Confusion matrices are commonly utilized to evaluate the performance of classification models, particularly those tasked with predicting categorical labels for input instances.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

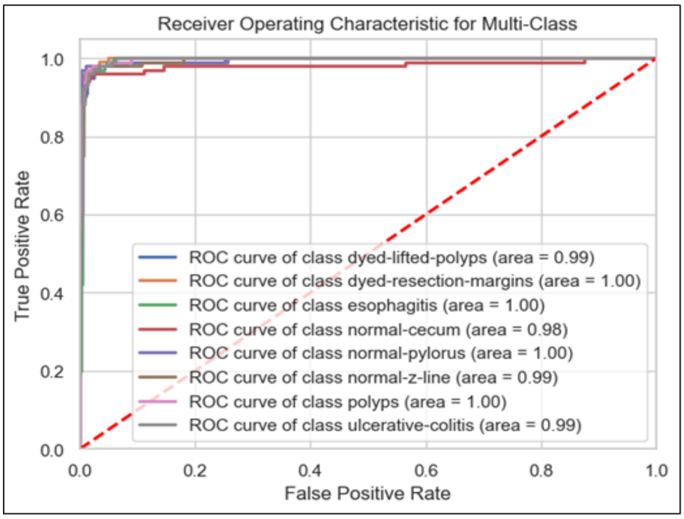

Figure 15: *ROC for Multi class*

According to the results shown in Figure 15, the supplement dataset significantly lowers the AUC value of Efficient Net, the sub-optimal model on datasets, to 0.98. In comparison, the suggested MAE-Net continues to have the highest AUC value of 0.99, demonstrating the unique overview and flexibility of this approach.

Comparison for the proposed method

In contrast, several other studies [17, 21] depend on baseline data, ignoring the training data's time series component. These studies did not sufficiently explain their findings and lacked training data for disease analysis.

Table 1: *Comparison for the proposed method*

Study	Dataset	Accuracy
Lonseko *et al.*[17]	Kvasir Multi-class	93.19 %
Mohapatra *et al.* [21]	Kvasir V2	93.75 %
Proposed approach	Kvasir	94 %

Comparisons for MAE-Net

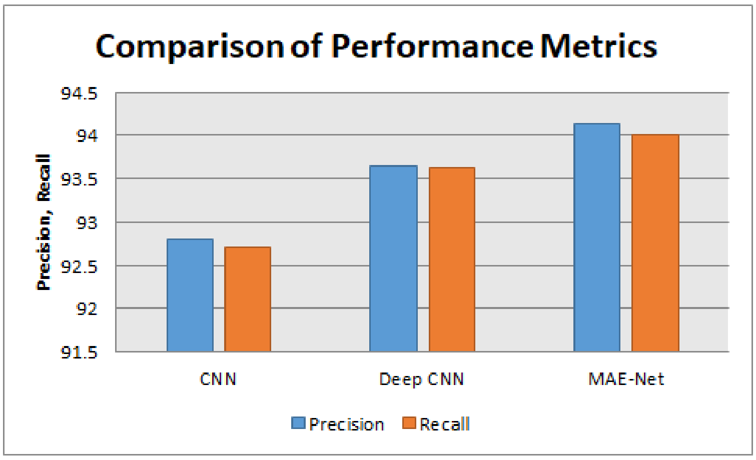

Figure 16: *Comparisons of performance metrics*

For Precision and recall it gives better predicted value when compared to the CNN [17] and Deep CNN [21]. Precision value of 94.14% and recall of 94% for the proposed MAE-Net method.

5 Conclusion

This research introduces a comprehensive and efficient framework for the detection and classification of gastrointestinal diseases using the proposed MAE-Net. The integration of advanced preprocessing techniques, such as resizing, grayscale conversion, and texture enhancement, ensures that the input images are optimized for feature extraction. The use of approximate spatial fuzzy c-means clustering for segmentation further enhances the precision of identifying relevant regions in the images. Additionally, the adoption of wavelet transform coefficients for time–frequency feature extraction enables a detailed and multi-dimensional analysis of the input data, capturing crucial patterns for effective classification. The performance evaluation, conducted on the Kvasir dataset, demonstrates the superior capabilities of the MAE-Net model compared to existing state-of-the-art methods. By using attention mechanism, the model focuses on critical features, resulting in improved accuracy of 94%.

References

1. Vladimir Milivojevic; Tomica Milosavljevic, Year: 2020, “Burden of gastroduodenal diseases from the global perspective”, Current Treatment Options in Gastroenterology, Vol.18, pp.148-157.

2. Kwang-Sig Lee; Eun Sun Kim, Year: 2022, “Explainable artificial intelligence in the early diagnosis of gastrointestinal disease,” *Diagnostics* Vol. 12, no. 11, pp. 2740.

3. Md Faysal Ahamed: Md Nahiduzzaman: Md Rabiul Islam; Mansura Naznine; Mohamed Arselene Ayari; Amith Khandakar; Julfikar Haider, Year: 2024, “Detection of various gastrointestinal tract

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

diseases through a deep learning method with ensemble ELM and explainable AI", Expert Systems with Applications, Vol. 256, pp. 124908.

4. Mohammad Wajih Alam; Seyed Shahim Vedaei, Khan A Wahid, Year: 2020, "A fluorescence-based wireless capsule endoscopy system for detecting colorectal cancer", Cancers, Vol. 12, no. 4, pp. 890.

5. Silvia Pecere; Sebastian Manuel Milluzzo; Gianluca Esposito; Emanuele Dilaghi; Andrea Telese; Leonardo Henry Eusebi, Year: 2021, "Applications of artificial intelligence for the diagnosis of gastrointestinal diseases", Diagnostics, Vol. 11, no. 9, pp. 1575.

6. Muhammad Nouman Noor; Muhammad Nazir; Sajid Ali Khan; Imran Ashraf; Oh-Young Song, Year: 2023, "Localization and classification of gastrointestinal tract disorders using explainable AI from endoscopic images." Applied Sciences 13, no. 15, pp. 9031.

7. Takashi Sakamoto; Tadahiro Goto; Michimasa Fujiogi; Alan Kawarai Lefor, Year: 2022, "Machine learning in gastrointestinal surgery", Vol. 52, no. 7, pp. 995-1007.

8. Malinda Vania; Bayu Adhi Tama; Hasan Maulahela; Sunghoon Lim, Year: 2023, "Recent advances in applying machine learning and deep learning to detect upper gastrointestinal tract lesions", IEEE Access.

9. Javaria Amin; Muhammad Sharif; Eman Gul, and Ramesh Sunder Nayak. Year: 2022, "3D-semantic segmentation and classification of stomach infections using uncertainty aware deep neural networks", Complex & Intelligent Systems, Vol. 8, no. 4, pp. 3041-3057.

10. Debapriya Banik; Kaushiki Roy; Debotosh Bhattacharjee; Mita Nasipuri; Ondrej Krejcar, Year: 2020, "Polyp-Net: A multimodel fusion network for polyp segmentation", IEEE Transactions on Instrumentation and Measurement, Vol. 70, pp.1-12.

11. Tariq Rahim; Syed Ali Hassan; Soo Young Shin, Year: 2021, "A deep convolutional neural network for the detection of polyps in colonoscopy images", Biomedical Signal Processing and Control, Vol. 68, pp. 102654.

12. Imran Iqbal; Khuram Walayat; Mohib Ullah Kakar; Jinwen Ma, Year: 2022, "Automated identification of human gastrointestinal tract abnormalities based on deep convolutional neural network with endoscopic images", Intelligent Systems with Applications, Vol. 16, pp. 200149.

13. Şaban Öztürk; Umut Özkaya, Year: 2020, "Gastrointestinal tract classification using improved LSTM based CNN", Multimedia Tools and Applications, Vol. 79, no. 39, pp. 28825-28840.

14. Seung-Joo Nam; Gwiseong Moon; Jung-Hwan Park; Yoon Kim; Yun Jeong Lim; Hyun-Soo Choi, Year: 2024, "Deep Learning-Based Real-Time Organ Localization and Transit Time Estimation in Wireless Capsule Endoscopy", Biomedicines, Vol. 12, no. 8, pp.1704.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 03 (11), November 2024, pp. 50-66

15. Xianhui Zeng: Liansong Ye,; Jianrong Liu, Xianglei Yuan, Shan Jiang, Minghui Huang, Xiujiang Huang, Chengwei Tang; Bing Hu, Year: 2021 “Value of the diving method for capsule endoscopy in the examination of small-intestinal disease: a prospective randomized controlled trial”, Gastrointestinal Endoscopy, Vol. 94, no. 4, pp.795-802.

16. Veysel Yusuf Cambay; Prabal Datta Barua; Abdul Hafeez Baig; Sengul Dogan; Mehmet Baygin; Turker TuncerP; U. R. Acharya, Year: 2024, “Automated Detection of Gastrointestinal Diseases Using Resnet50*-Based Explainable Deep Feature Engineering Model with Endoscopy Images”, Sensors, Vol. 24, no. 23 pp. 7710.

17. Zenebe Markos Lonseko; Prince Ebenezer Adjei; Wenju Du; Chengsi Luo; Dingcan Hu; Linlin Zhu; Tao Gan; Nini Rao, Year. 2021, “Gastrointestinal disease classification in endoscopic images using attention-guided convolutional neural networks”, Applied Sciences, Vol.11, no. 23, pp.11136.

18. Mehshan Ahmed Khan: Muhammad Attique Khan: Fawad Ahmed: Mamta Mittal: Lalit Mohan Goyal; D. Jude Hemanth; and Suresh Chandra Satapathy, Year. 2021, “Gastrointestinal diseases segmentation and classification based on duo-deep architectures”, Pattern Recognition Letters, Vol.131, pp.193-204.

19. Ji Xia; Tian Xia; Jun Pan; Fei Gao; Shuang Wang; Yang-Yang Qian; Heng Wang; Jie Zhao; Xi Jiang; Wen-Bin Zou; Yuan-Chen Wang; Wei Zhou; Zhao-Shen Li; Zhuan Liao, Year: 2021, “Use of artificial intelligence for detection of gastric lesions by magnetically controlled capsule endoscopy”, Gastrointestinal endoscopy, Vol. 93, no. 1, pp. 133-139.

20. Melaku Bitew Haile; Ayodeji Olalekan Salau; Belay Enyew; Abebech Jenber Belay, Year: 2022, “Detection and classification of gastrointestinal disease using convolutional neural network and SVM”, Cogent Engineering, Vol. 9, no.1, pp. 2084878.

21. Subhashree Mohapatra; Janmenjoy Nayak; Manohar Mishra; Girish Kumar Pati; Bignaraj Naik; Tripti Swarnkar, Year: 2021, “Wavelet transform and deep convolutional neural network-based smart healthcare system for gastrointestinal disease detection”, Interdisciplinary Sciences: Computational Life Sciences, Vol. 13, pp. 212-228.