Diagnosis of Gastrointestinal Diseases Using Modern CNN Techniques

Guduru Varalaxmi
ECE, GRIET
Telangana, India
varalaxmiguduru20@gmail.com

Sahithi Reddy Baddam
ECE, GRIET
Telangana, India
sahireddy17@gmail.com

Emmanuel Suhas Yalamarthi
ECE, GRIET
Telangana, India
suhas.e001@gmail.com

K. Swaraja
ECE, GRIET
Telangana, India
kswaraja@gmail.com

K. Reddy Madhavi
School of Computing, Mohan babu University
Tirupati, India
kreddymadhavi@gmail.com

CN. Sujatha
ECE, Sreenidhi Institute of Science and Technology
Ghatkesar, India
cnsujatha@gmail.com

***Abstract*—Endoscopy is used by the doctors in detecting the gastrointestinal diseases. Endoscopy is a procedure which helps in detection of many gastric diseases like polyps, cancer infections, celiac disease, Crohn's disease, ulcerative colitis, diverticulitis, malabsorption and so on. As in recent days use of endoscopy is increasing drastically because of high increase in gastrointestinal related issues. As this method consumes more time and repeated trials on the patient, it may affect the patient. Therefore, detection of diseases using deep learning helps in detecting the gastric diseases without affecting the patients. By using modern Convolutional Neural Network (CNN) techniques we can acquire accurate results within no time, so that in case of emergencies the treatment can be started immediately without any delays. In this work ResNet50 is implemented which is the subset of Residual Network. RESNET (a neural network), uses the complementary clues of spectral and spatial features to improve the classification. The proposed framework attained an accuracy of 88.05%, sensitivity of 87.05% and specification of 92.33%.**

Keywords*—*Gastrointestinal diseases, ResNet50, Adam

I. INTRODUCTION

In recent times, people are suffering from various cancers which possess a great threat to human health. In this, Gastrointestinal cancer is the most common cancer of the cavity organs, posing a serious threat to human health. So, it is very much important in detecting the gastric cancer in the early stages. In detection of gastrointestinal diseases Endoscopy is used. Endoscopy is performed with a tube that has a tiny camera attached to one end of it, this tube is sent through mouth of the patient and goes until the end of the intestines. The main function of camera is to capture the high-quality footage of the internal organs in video format and gives live output to the doctor performing diagnosis. There are many advantages of using Endoscopy as a primary method, for example Inspection of the internal organs without rupturing the outer tissues of the organ and no need of using expensive surgical methods. But, sometimes endoscopists may make incorrect observations during endoscopy due to fatigue caused by long working hours or inexperience. And this method consumes more time and repeated trials on the patient, may cause inconvenience to the patient. Therefore, detection of diseases using deep learning helps in detecting the gastric diseases without affecting the patients.

Machine learning approaches are utilized to explore the properties of edge, texture as well as color from images of endoscopy which rely on experimentation for detection of disease. The application of supervised learning in CNN's has significantly boosted the capacity of treatment of medical images. To extract features from the videos or images CNN has amazing capacity. Deep learning algorithms have outperformed experts in medical image diagnostics. As a consequence, computer-aided diagnostics for endoscopic images employing algorithms based on deep learning have the potential to outperform skilled doctors in terms of diagnostic accuracy. CNN involves many architectures such as Alexnet, Googlenet, VGG, Res-Net etc. Each architecture has a drawback; how-ever some studies indicate that accuracy seems good when using Res-Net. This is because Res-Net's architecture is built in a way that makes it better able to decrease vanishing and exploding gradients than other architectures. So, in the proposed work, ResNet50 is implemented to diagnosis of gastrointestinal diseases using modern CNN techniques for getting more accuracy in detecting the diseases. From the previous work [3], deep convolutional neural network (CNN) architecture called Single Shot Multi Box Detector is used to detect and classify the gastrointestinal diseases.

The proposed work makes the most significant contribution by developing a computer-aided detection technique for syndromes with lower gastrointestinal, that uses altered method to retrieve features such as shape, color as well as texture through fine-tunning. To select pretrained models, to identify disorders of the lower gastrointestinal tract, numerous trials were carried out. The data is taken from the Kvasir dataset. The dataset is tested by ResNet50 to detect and classify the diseases. First, the dataset is preprocessed. Then the trained ResNet50 is evaluated on the dataset of images to measure its performance in detecting gastrointestinal diseases. The confusion matrix is then computed, and the Accuracy, Specification, and Sensitivity are determined. Then the accuracy and sensitivity of the ResNet50 is compared with the accuracies and sensitivities of previous works.

The proposed work is organized into 5 sections, Literature Survey is discussed in section 2. In section 3, Proposed Methodology which includes the implementation of ResNet50 is analyzed. The simulation results obtained from the Methodology are analyzed and Accuracy, Sensitivity, Specification are calculated in section 4. Finally,

979-8-3503-3401-2/23/$31.00 ©2023 IEEE

Conclusions with Future scope in this field are discussed in section 5.

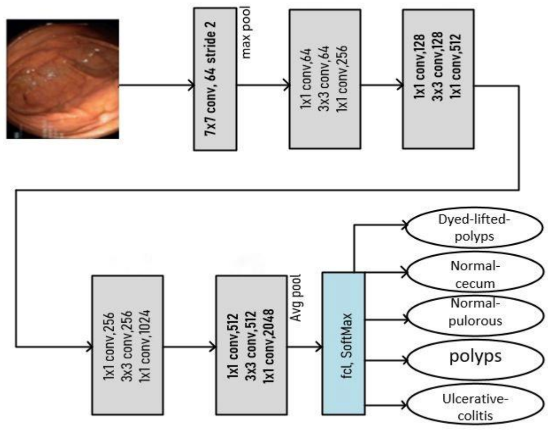

Fig. 1. ResNet-50 structure for detection of gastrointestinal syndrome.

II. LITERATURE SURVEY

Many researchers analyzed about the gastric diseases and detected them by executing their own algorithms. They used different techniques in detecting the gastric diseases. Some of them are taken as reference for the proposed work. Colon polyps can be found using a technique developed by Karkanis et al. [16] that uses wavelet decomposition to extract color characteristics. Many works have used wavelets, context-based features, valley data and edge shape, polyp-based local binary, grey-level co-occurrence matrices (GLCMs), and wavelets to extract features from gastrointestinal images. Taj bakhsh et al. [17] proposed a system that performed better than other methods. It is still challenging to extract manually created elements like structural polyps, camera angle and light reflection. The diagnosis of diagnostic imaging by CNN has shown encouraging outcomes in current history. CNN methods are effective at extracting deep properties. Zhang et al [1].'s Single Shot Multi Box Detector (SSD)-based polyp detection system enhanced detection and classification accuracy by adding the utilized missing features that are extracted from the max pooling layers to the feature maps. Godkhindi and Gowda [18] introduced a CNN method for recognizing polyps from images obtained after a CT colonography. Ribeiro et al. [19] have suggested using CNNs for gastrointestinal epithelium diagnostics to find gastrointestinal cancers in order to diagnose initial colon cancer. In this way, every work has its unique way in the detection of diseases.

In the proposed work, ResNet50 (Fig.1) CNN model is executed to uprise the Accuracy. The Kvasir dataset is taken which consists of 5,000 images and were classified as 5 disorders of the gastrointestinal system using the ResNet-50. The ResNet-50 block diagram is made up of sixteen blocks, each of which contains 177 layers, divided into 49 convolutional layers and an input layer which accepts Red, Green, Blue images (RGB images) with a pixel density of 224 x 224 pixels. Deep features are extracted from input images by the convolutional layer and stored in deep feature vector maps by using a single pooling layer for both average and maximum. The feature vector map's dimensions are reduced by these two layers. Normalization aids in the network's effective learning rate calculation. Convolutional layers are followed by the Rectified Linear Activation function (ReLU), which only transmits positive outcomes and transmits any negative traits to zero. While the completely connected layer obtains 9216 features and generates 4096 features, the second linked layer only generates 1000 features. The five classes, ulcerative colitis, Normal cecum, polyps, and normal pylorus together with dyed-lifted polyps were generated by the soft-max layer through implementing the proposed framework.

III. PROPOSED METHODOLOGY

Endoscopy is a technique that aids in the identification of a variety of gastrointestinal disorders, including cancer infections, polyps, celiac disease, Crohn's disease, diverticulitis, malabsorption, ulcerative colitis and others. In this work, various proposed techniques were analyzed, few of them are ResNet50, Googlenet, Alexnet ResNet34. Among these ResNet50 is chosen, because it explores spectral and spatial features to guide the Res-Net framework. Thus, it helps in improving the classification. We are also using optimization methods and transfer layers this helps in fast processing. Preprocessing enhances images by removing noise and artefacts, whereas image augmentation technology enhances the training process. The most specific and significant features from each image are extracted through convolutional layers. The gastrointestinal images are diagnosed and classified by the fully connected layers.

A. Dataset

The dataset is taken from Kvasir datasets [20] which consists of 5000 images, which are equally divided into 5 diseases: ulcerative colitis, normal pylorus, dyed-lifted tumors, normal cecum, and polyps. The Kvasir dataset includes classes, that include gastroscopic methods in the gastrointestinal system and architectural features, as well as images that have been properly analyzed. The dataset includes numerous images that can be used for transfer learning and deep learning. The RGB color space dataset comprises of images with dimensions ranging from 720x576 to 1920x1072.

B. Preprocessing and Augmentation Techniques:

The mucous membranes that surround interior organs, light reflections, and photographic angles, all these factors cause noise and artefacts. Due to this the performance of CNN will be reduced. Therefore, researchers use optimization processes to improve image quality. Before being fed into the CNN model, gastrointestinal images are preprocessed in our work. In preprocessing, before changing the image sizes to 244x244 pixels for Resnet50 they are scaled for color constancy. Then it was determined for the gastric images what the mean of the RGB channels was. The average filter, which computes and replaces each pixel's average with those of its neighbors, was used to complete the enhancement process. This procedure is carried out again for each pixel in the image.

The most crucial element in CNN approaches is data volume. The model performs better when the training set is larger. Since there is a shortage of diagnostic images, data augmentation methods are utilized, which improves the CNN model for precise categorization. The approaches for data augmentation also aid in balancing the dataset. In this work, training data images were enhanced through rotation, flipping, shifting and zooming.

C. CNN Layers

The primary characteristics of gastrointestinal datasets are color, shape and texture. Along with this, there are other additional advantages also. As every image may not have the syndrome and the disease traits only shown in some images that perhaps the radiologist and experts may miss. Therefore, the manual feature extraction takes a lot of knowledge. So, to avoid this, CNN algorithms are implemented to extract representative features from each disease using convolutional layers. ResNet50 is implemented in the proposed work. Resnet50 has 49 convolutional layers. These layers change the weights while training to handle the deep characteristics and carry them to the subsequent layer through a set of filtration techniques. The max pooling layers and average layers assist to minimize the dimension of the feature maps by using the mean or highest value among pixel groups to symbolize a set of pixels. A total of 9216 features per image were extracted by these convolutional layers and they are represented in extracted features will be sent into the classifying layer.

D. Normalization of the Image

Normalization is used in deep learning neural network training to normalize images in order to accelerate training methods. By gradient decent converging, normalization helps to increase the learning rate. The learning rate becomes more challenging and training will take more time without normalization. The mean of the complete training set was subtracted from each pixel in the proposed work to normalize the images. Data centering and each feature's variance being equal to one resulted from computing and reducing the dataset's variance for each pixel.

E. Dropout Technology

Overfitting happens when a model is unable to generalize and is overly dependent on the training dataset. Overfitting occurs when the train data collection is too small and does not contain sufficient data samples to appropriately represent all potential input data values. Thus, we use the dropout strategy in the present work to lessen overfitting. Thus, ResNet50 is subjected to the dropout approach, which results in the termination of 40% of the neurons throughout each iteration. As a result, each iteration of the networks uses a distinct set of parameters. The training time is doubled with the Dropout technique.

F. ResNet50

It is based on using a dataset that has been trained to solve a particular problem to solve on other dataset's problem. This is the most crucial step in the CNN model. The Resnet50 networks pretrained on the ImageNet dataset are used in the proposed work to apply transfer learning and fine-tuning. In order to apply what has been learnt to a different task, transfer learning first chooses the size of the challenge and the pretrained model. It also helps in avoiding overfitting. In this work, the weights are fine-tuned when the transfer learning is applied to Resnet50. Before even being deployed to the gastrointestinal dataset, the Resnet50 models had been trained on the ImageNet dataset. The completely linked layer was used to replace the last 3 layers of the patterns. In addition to receiving 9,216 neurons, the first linked layer also produced 4,096 neurons, as did the second connected layer. To categorize diseases into distinct classifications, the soft-max layer is used.

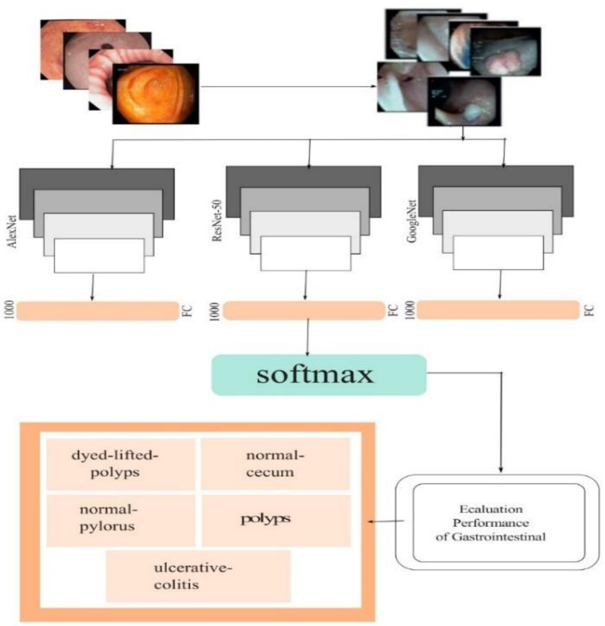

Fig. 2. General architecture of gastric diseases detection using deep learning.

G. Optimizers (Adam)

They are used for adjusting as well as tunning the neural network parameters such as weighting, bias as well as training rate in order to decrease loss. To improve the classification using deep learning, optimizer methods are used that helps in speeding up the performance of the models. The best deep learning optimizer is Adaptive Moment Estimation (Adam). Adam is a cross between RMSProp and momentum. For each parameter, Adam calculates the adaptive learning rate. By storing past decaying average, it preserves mean previous gradients like momentum (A) and squared gradients like declining average (D)and b1 and b2 indicate the decay rate. Eq (1) and Eq (2) describes Adam's approach to tuning parameters, learning rate.

$$A = b1(A - 1) + (1 - b1)G \qquad [1]$$

$$D = b2(D - 1) + (1 - b2)G^2 \qquad [2]$$

IV. .SIMULATION RESULTS

The proposed methodology is implemented using google collab, GPU (Tesla T4), processor intel i5. The programming language used for execution is python. In this work, the Kvasir dataset [21] is utilized to evaluate the gastrointestinal diseases. The dataset used in the proposed work contains 5000 images which are equally divided into five diseases. The dataset is divided into 20% for selection and validation and 80%for training. The implementing elements are as shown in table1.

TABLE I. PREFERENCES IN THE PROPOSED WORK

Optimizer	Adam
Activation layer	ReLU
Dataset	Kvasir

After setting up the preferences, the data is collected from the Kvasir dataset. Then preprocessing of the acquired data is done to improve the contrast and brightness of the image by removing noise and artefacts. Now, the CNN model (ResNet50) is trained on the preprocessed images using a supervised learning approach. The trained CNN model is tested on a Kvasir dataset of images to measure its performance in detecting gastrointestinal diseases. From this the confusion matrix is obtained by which Accuracy, Specificity and Sensitivity are calculated.

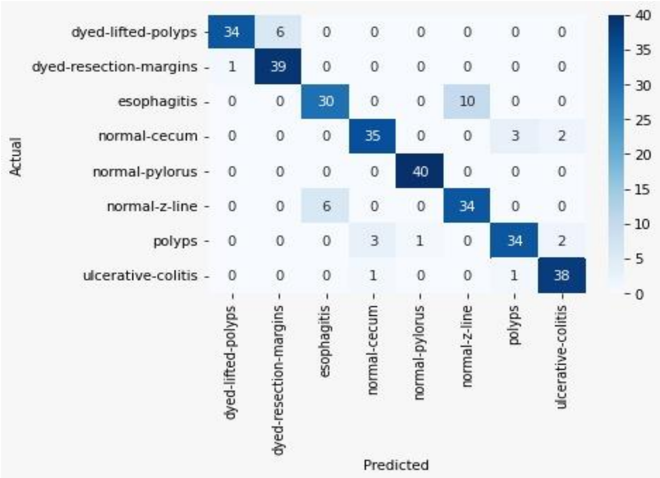

Fig. 3. Confusion matrix

The confusion matrix is depicted in Fig.3. The confusion matrix examines whether the results obtained for each test image is correctly identified as true negative (T1) or true positive (T2), as well as any incorrectly identified test images as false negative (F1) or false positive (F2). Accuracy, Sensitivity, Specificity, and AUC are determined using Eq's [3]-[5].

$$Accuracy = \frac{T2 + T1}{T2 + T1 + F1 + F2} * 100\% \qquad [3]$$

$$Sensitivity = \frac{T2}{T2 + F1} * 100\% \qquad [4]$$

$$Specificity = \frac{T1}{T1 + F2} * 100\% \qquad [5]$$

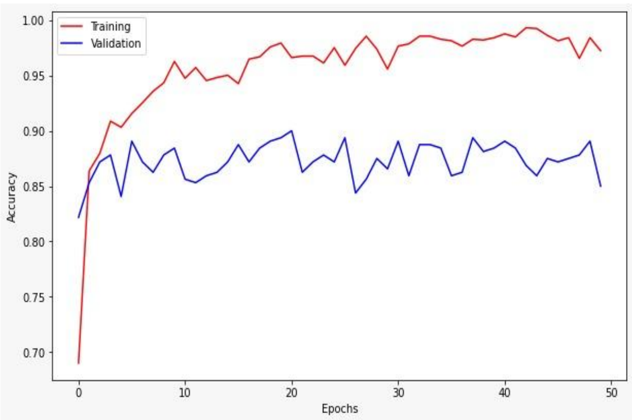

Fig. 4. Model Accuracy

Fig. 4 shows the model accuracy graph. A machine learning model's performance over training and validation datasets is shown visually in the model accuracy graph. It displays the accuracy of the model as a function of the quantity of training iterations or epochs. It is also useful tool for evaluating and comparing different machine learning models. It can help to identify the problems of underfitting and overfitting, as well as the optimal number of training epochs for a given model. The model accuracy graph typically consists of two lines: One for the accuracy of the training, and another for the accuracy of the validation. The model's accuracy on the training data is depicted by the training accuracy line as the model is being trained. The validation accuracy line displays the accuracy of the model on a separate validation dataset, which is used to evaluate the model's performance on unseen data. In the accuracy graph (Fig.4), curve generally starts at lower accuracy level and increases as the model is trained. When validation accuracy starts to decrease while training, accuracy continues to increase then this indicates overfitting. On the other hand, if both the training and validation accuracy are low, then this indicates underfitting. In the accuracy graph (Fig.4), both underfitting and overfitting is avoided as the validation accuracy is following the training accuracy. Thus, the model accuracy graph is helpful in analyzing and improving machine learning models, and is used in the proposed work to make decisions about adjusting the model's architecture or the hyperparameters used during training.

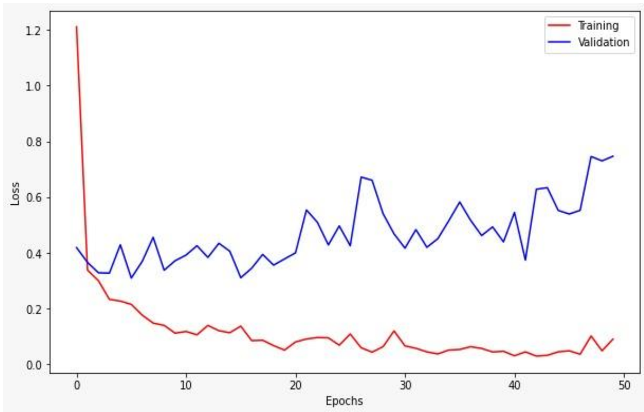

Fig. 5. Model loss

Model loss graph is shown in the Fig.5. The performance of a machine learning model over training and validation datasets is represented visually by a model loss graph. It depicts the loss function of the model as a function of training epochs or iterations. The loss function evaluates how effectively the model can predict the values of the outcomes

Fig.6 shows the output of the proposed model. The diseases are detected and classified by using ResNet50 CNN model. In the classified and detected diseases, B(i), B(ii), B(iii), D(ii), E(i), E(ii), E(iii), E(iv), E(v) are the truly detected diseases. Whereas, the remaining images are falsely detected diseases.

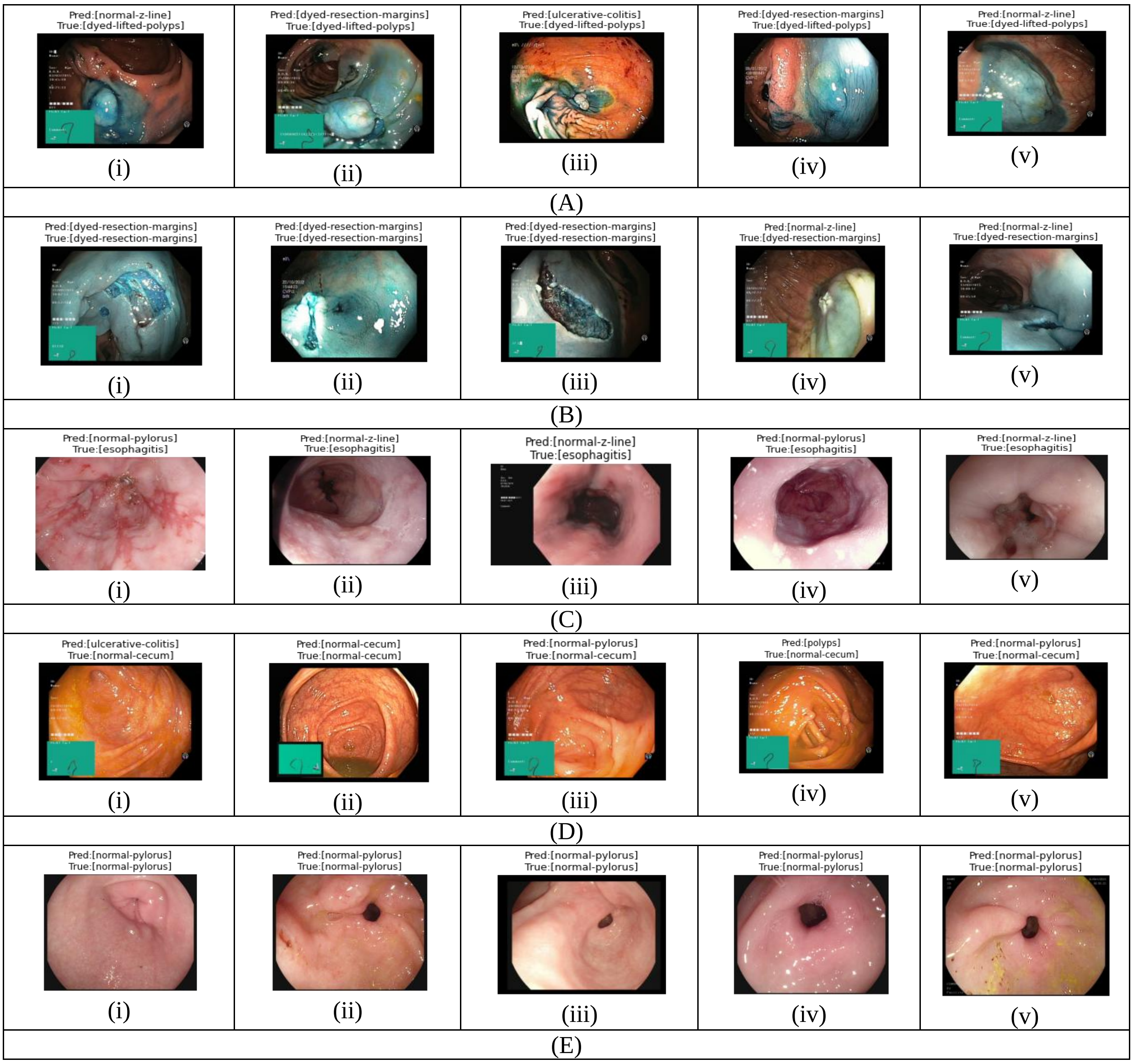

Fig. 6. Predicted and true values of identified five different diseased classes.

TABLE II. PERFORMANCE METRICS OF DETECTED DISEASES IN THE PROPOSED WORK

Diseases	Accuracy	Specificity	Recall
Dyed-lifted polyps	95.31%	100%	84.00%
Dyed-resection margins	95.31%	97.50%	100%
Esophagitis	93.75%	75.00%	75%
Normal-cecum	95.31%	94.59%	97.22%
Normal pyloris	98.44%	100%	100%
Normal z-line	96.88%	84.91%	85%
Polyps	95.31%	94.44%	85%
Ulcerative-Colotis	93.75%	97.50%	95%

The performance metrics (accuracy, recall and specificity) of the proposed work is shown in the table 2. Normal pyloris has more accuracy than other diseases. And the Ulcerative Colotis has least accuracy.

TABLE III. OVERALL PERFORMANCE METRICS

Measurement (ResNet50)	Proposed model
Accuracy	88.05%
Specificity	92.33%
Sensitivity	85.00%
Validation accuracy	87.05%

The Table 3 describes the proposed work's general performance metrics. By substituting the values of T1, T2, F1 and F2 in Eq (3) to Eq (5), Accuracy, Specificity, Sensitivity are calculated.

TABLE IV. COMPARISON OF OUR MODELS WITH PREVIOUS STUDIES

Methods	Accuracy (%)	Specificity (%)	Sensitivity (%)
Zhang et al. [1]	70.40	70.40	70.40
Zhu et al. [2]	85.00	83.00	82.54
T. Ozawa [3]	83.00	-	87.10
Y.Sasaki [15]	85.90	-	87.60
Proposed Model (ResNet50)	88.05%	92.33%	87.05%

In Table 4, it is observed that there is uprise in accuracy, specificity and sensitivity of the proposed work when compared to the previous works as the proposed model is implemented using RESNET 50. RESNET50, a neural network, uses the complementary clues of spectral and spatial features to guide RESNET so that it can improve the classification. Thus, proposed work is more helpful for clinicians in detecting and classifications of gastrointestinal diseases

V. CONCLUSION

In this work, the gastrointestinal diseases are classified and detection with the use of Resnet50. This work helps the clinicians in false detection of diseases. The accuracy, sensitivity and specification are calculated and got more accuracy, sensitivity and specification when compared to previous works.

REFERENCES

[1] X. Zhang, F. Chen, T. Yu et al., "Real-time gastric polyp detection using convolutional neural networks," PLoS One, vol. 14, no. 3, p. e0214133, 2019.

[2] Zhu Y, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, et al. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional.

[3] T. Ozawa, S. Ishihara, M. Fujishiro, Y. Kumagai, S. Shichijo, and T. Tada, "Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks," Therapeutic advances in gastroenterology, vol. 13, p. 1756284820910659, 2020.

[4] Mosleh Hmoud Al-Adhaileh , 1 Ebrahim Mohammed Senan , 2 Fawaz Waselallah Alsaade , 3 Theyazn H. H Aldhyani , 4 Nizar Alsharif,5 Ahmed Abdullah Alqarni , 6 M. Irfan Uddin , 7 Mohammed Y. Alzahrani , 6 Elham D. Alzain , 4 and Mukti E. Jadhav 22 September 2021; Published 18 October 2021.

[5] Ziyi Jin, Tianyuan Gan, Peng Wang, Zuoming Fu, Chongan Zhang, Qinglai Yan, Xueyong Zheng, Xiao Liang & Xuesong Ye BioMedical Engineering OnLine volume 21, Article number: 12 (2022)

[6] Lee JH, Kim YJ, Kim YW, Park S, Choi YI, Kim YJ, et al. Spotting malignancies from gastric endoscopic images using deep learning. Surg Endosc. 2019;33(11):3790–7.

[7] Horiuchi Y, Hirasawa T, Ishizuka N, Tokai Y, Namikawa K, Yoshimizu S, et al. Performance of a computer-aided diagnosis system in diagnosing early gastric cancer using magnifying endoscopy videos with narrow-band imaging (with videos). Gastrointest Endosc. 2020;92(4):2008

[8] Zhu Y, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, et al. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional.

[9] Yoon HJ, Kim S, Kim JH, Keum JS, Oh SI, Jo J, et al. A lesion-based convolutional neural network improves endoscopic detection and depth prediction of early gastric cancer. J ClinMed. 2019;8:9.

[10] Wu L, Zhou W, Wan X, Zhang J, Shen L, Hu S, et al. A deep neural network improves endoscopic detection of early gastric cancer without blind spots. Endoscopy. 2019;51(6):522–31.

[11] Wang H, Ding S, Wu D, Zhang Y, Yang S. Smart connected electronic gastroscope system for gastric cancer screening using multi-column convolutional neural networks. Int J Prod Res. 2019;57(21):6795–806.

[12] D. K. Iakovidis, D. E. Maroulis, and S. A. Karkanis, "An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy," Computers in Biology and Medicine, vol. 36, no. 10, pp. 1084–1103, 2006.

[13] D. K. Iakovidis, D. E. Maroulis, and S. A. Karkanis, "An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy," Computers in Biology and Medicine, vol. 36, no. 10, pp. 1084–1103, 2006.

[14] T. Hirasawa, K. Aoyama, T. Tanimoto et al., "Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images," Gastric Cancer, vol. 21, no. 4, pp. 653–660, 2018

[15] Y. Sasaki, R. Hada, T. Yoshimura, N. Hanabata, T. Mikami, and S. Fukuda, "Computer-aided estimation for the risk of development of gastric cancer by image processing," in Artificial Intelligence in Aeory and Practice III, pp. 197–204, Springer, Berlin, Germany, 2010.

[16] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, and M. Tzivras, "Computer-aided tumor detection in endoscopic video using color wavelet features," IEEE Transactions on information Technology in Biomedicine, vol. 7, no. 3, pp. 141–152, 2003.

[17] N. Tajbakhsh, S. R. Gurudu, and J. Liang, "Automated polyp detection in colonoscopy videos using shape and context information," IEEE Transactions on Medical imaging, vol. 35, no. 2, pp. 630–644, 2015.

[18] A. M. Godkhindi and R. M. Gowda, "Automated detection of polyps in CT colonography images using deep learning algorithms in colon cancer diagnosis," in Proceedings of the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 1722–1728, IEEE, Chennai, India, August 2017.

[19] E. Ribeiro, A. Uhl, and M. Hafner, "Colonic polyp classifi- ¨ cation with convolutional neural networks," in Proceedings of the 2016 IEEE 29 International Symposium on ComputerBased Medical Systems (CBMS), pp. 253–258, IEEE, Dublin, Ireland, June 2016.

[20] https://datasets.simula.no/kvasir/