International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Deep Learning Based Gastrointestinal Disease Detection Using CNN & VGG-16 Neural Networks

Dr. D. Yugandhar[1], K. Arun[2], H.K.V. Santosh Dora[3], N. Binita[4], P. Sumanth[5]

[1]Professor, [2,3,4,5]UG Student

Department of E.C.E, Aditya Institute of Technology & Management, Tekkali

ABSTRACT

In recent years, diagnosing gastrointestinal tract diseases through endoscopic image classification has gained traction. In this regard, the KVASIR V2 dataset comprising of eight classes of GI-tract images such as Normal cecum, Normal pylorus, Normal Z-line, Esophagitis, Polyps, Ulcerative Colitis, Dyed and lifted polyp, and Dyed resection margins are used for classification. In this project, a novel approach is proposed, leveraging a concatenated neural network model. This model integrates features extracted from CNN and VGG-Net on endoscopic images yields promising results, achieving 87% and 96.88% accuracy respectively. Lastly, classification performance is measured through the following measuring indices like accuracy, precision, recall, specificity, and F1 score.

Graphical Abstract

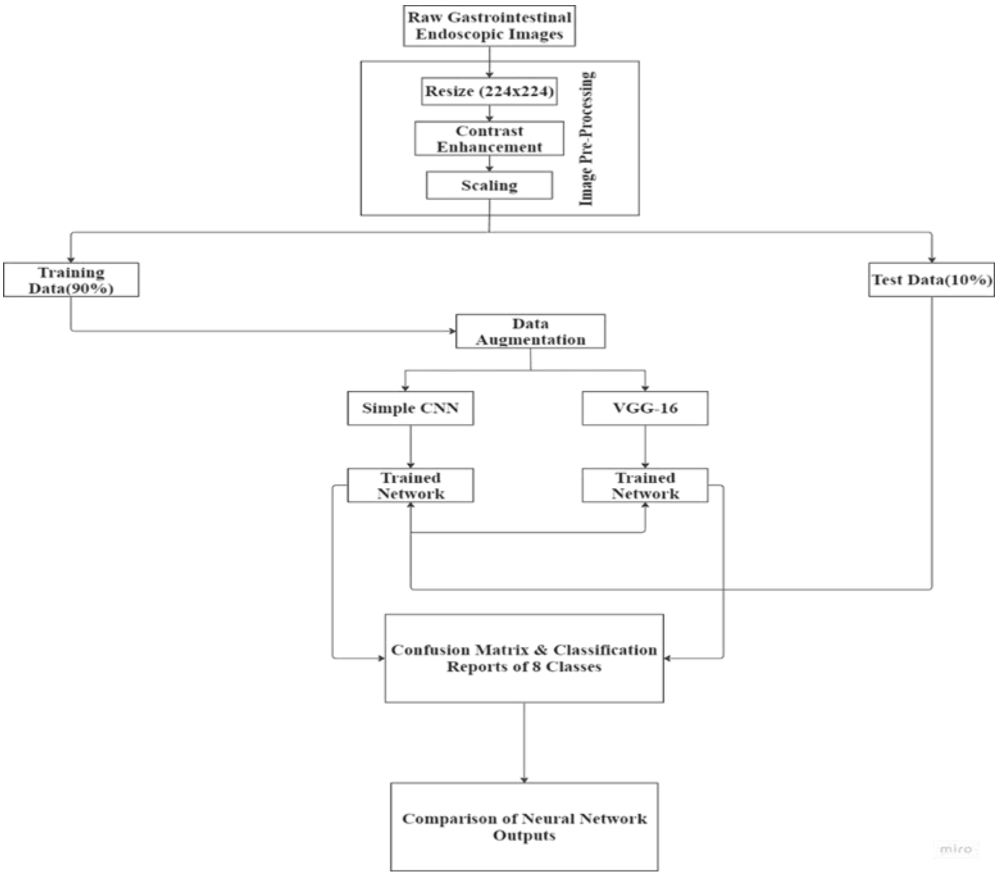

Fig 1. Graphical Abstract

Keywords: Smart healthcare system, Gastrointestinal Tract endoscopic images, Convolutional Neural Network, Visual Geometry Group (VGG).

1.Introduction

1.1 Overview

Gastroenterologists are increasingly over whelmed by the rising number of patients grappling with gastro-intestinal (GI) issues. This surge can be attributed to several factors, including unhealthy dietary habits prevalent among middle and upper-class individuals, demanding work schedules, sedentary lifestyles, heightened stress levels, malnutrition among underprivileged children, and inadequate sanitation in slums and rural areas [1]. A significant aspect of a gastroenterologist's duties involves examining images and videos of the GI tract [2] Fig.1.1, which are crucial for diagnosis and treatment planning. Various imaging techniques such as X-ray, MRI, CT, ultrasound, and PET are commonly used [3], but the most detailed images are obtained through minimally invasive procedures like endoscopy Fig.1.2.

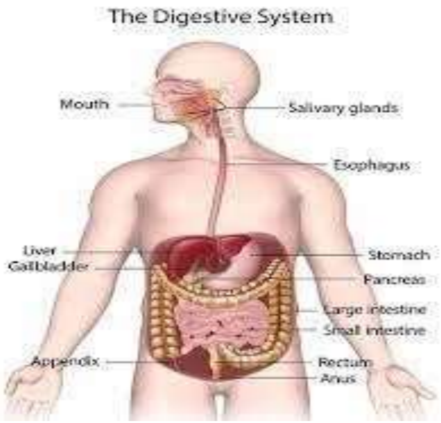

Fig.1.1***: The Gastrointestinal tract***

The timely diagnosis of gastrointestinal diseases is pivotal in mitigating the risk of severe medical complications. Expertise from skilled professionals [5] is indispensable in this process. However, with the escalating number of patients, there's a corresponding surge in data volume necessitating analysis. Conventional machine learning approaches hinge on manually engineered features. Conversely, deep learning methods autonomously extract features and utilize them for disease diagnosis[6],[7]. This autonomy is particularly advantageous in handling extensive volumes of structured or unstructured medical data, which may manifest in textual or image formats[8]. Consequently, for the analysis of such copious datasets, deep learning emerges as the preferred approach.

Fig.1.2: ***The endoscope***

In recent years, convolutional neural networks (CNNs) have become increasingly prominent in the realm of medical image analysis. CNNs excel in learning from image data, autonomously extracting features, and subjecting them to numerous convolutional and pooling layers to identify inherent patterns within the images[9].

1.2 Literature Review

Several models have been introduced in the literature to facilitate the effective diagnosis of GI-tract diseases. These models have undergone rigorous training and testing using specific self-collected endoscopic images from various segments of the GI tract, along with a selection of publicly available datasets. For instance, in a study by researchers [10], a CNN model was trained and tested using 790 endoscopic images of gastric cancer tumors for training and 203 images for testing. These images were sourced from the Endoscopy Center of Zhongshan Hospital in China, with the aim of estimating the invasive depth of cancer. The KVASIR dataset has garnered widespread attention among researchers aiming to devise effective systems for detecting GI-tract issues. Cogan et al. [11] introduced a framework for the automatic and modular preprocessing of GI tract images using deep learning techniques. Additionally, researchers [12] demonstrated the efficacy of data augmentation in enhancing the accuracy of GI-tract disease diagnosis. Hicks et al. [13]

proposed a framework capable of delving into the layers of deep networks to improve image classification and comprehension. Ghatwary et al. [14] presented a deep learning model for detecting esophageal abnormalities, utilizing the KVASIR dataset alongside the MICCAI 2015 dataset [15]. Owais et al. [16] developed an AI-based system for classifying multiple GI-tract diseases, leveraging datasets from Gastrolab [17] and KVASIR. Furthermore, authors [18] proposed an automated model for recognizing GI tract infections, employing traditional feature extraction along with CNN-based feature extraction and genetic algorithm (GA)-based feature selection. Similarly, researchers [19] introduced an efficient long short-term memory (LSTM)-based CNN network for classifying labeled and imbalanced GI-tract disease samples. In recent years, there has been a notable surge in the digitization of medical images, driven by the pursuit of enhancing feature representation and improving abnormality detection capabilities. Amidst this digitalization trend, wavelet transform (WT) has emerged as a significant tool, drawing upon principles of multi-resolution signal processing in computer vision. While WT has found broad application across various medical imaging domains, its utilization in endoscopic imaging of the GI tract has been particularly noteworthy. Numerous studies have successfully employed WT in diverse applications within GI endoscopy, including endoscopic image compression [20], classification of small bowel images [21, 22], GI polyp detection in video endoscopy [23], esophageal cancer detection [24], and identification of ulcers and bleeding [25]. Despite the extensive literature on GI-tract abnormality detection, there remains ample opportunity for further exploration of wavelet analysis within the GI domain. Notably, the potential application of WT in the KVASIR dataset remains largely untapped.

2.Dataset:

According to existing literature, there exist four distinct publicly available datasets relevant to GI-tract research: KVASIR [25], Nerthus, Gastrolab, and HyperKvasir [33]. For this experimental endeavor, the authors have chosen the KVASIR V2 dataset [26], an updated version of KVASIR. The primary rationale behind selecting this dataset is its balanced nature compared to others. Unlike other datasets with varying numbers of images per class, the KVASIR dataset offers balance, with each of its eight classes containing 1000 images. Among these classes, three represent anatomical landmarks, three depict pathological findings, and the remaining two illustrate different endoscopic procedures on the GI tract. The anatomical landmark classes—Z-line, pylorus, and cecum—depict various discernible characteristics of the GI tract observed through endoscopic examination. Pathological classes illustrate different abnormal conditions of the GI tract, encompassing esophagus, pylorus, and ulcerative colitis. Furthermore, there are classes representing endoscopic procedures associated with polyp removal[27], featuring two sets of images: dyed and lifted polyps, and dyed resection margins.

The images within the dataset exhibit varying resolutions, spanning from 720x224 up to 1920x1072 pixels. They are meticulously sorted and categorized into different folders corresponding to their respective classes. To facilitate enhanced representation and interpretation, certain image classes incorporate small green icons within the images, indicating the configuration and position of the endoscopic instrument within the organ. Below are detailed descriptions of individual image classes.

2.1 Anatomical Landmarks

Anatomical landmarks play crucial roles as navigational aids and reference points for interpreting various findings within the gastrointestinal (GI) tract[28]. They are also pivotal areas in diagnosing various diseases. The KVASIR V2 dataset includes three distinct landmarks::

2.1.1 Normal Cecum

The cecum denotes the proximal segment of the large intestine, distinguished prominently by the presence of the appendiceal orifice. The achievement of complete colonoscopy is signified by reaching the cecum, emphasizing the significance of thorough examination and identification. Please refer to Fig.2.1 for an exemplar image of the cecum.

Fig 2.1: ***Cecum***

2.1.2 Normal Pylorus

The pylorus marks the boundary between the opening of the stomach and the duodenum. Circular muscles surrounding the opening control the passage of food from the gastric region. A thorough examination of both sides of the pyloric opening is crucial for detecting any abnormal findings in this particular area. Please refer to Fig.2.2 for an illustrative image of the pylorus.

Fig.2.2**: *Pylorus***

2.1.3 Normal Z-line

The Z-line marks the transition between the esophagus and the stomach, clearly delineated as a line separating the white mucosa of the esophagus from the red mucosa of the stomach. Precise examination of the Z-line is essential for detecting the presence of disease and elucidating pathological conditions in the esophagus. Please refer to Figure 2.3 for an illustrative image of the Z-line.

Fig.2.3: ***Z-line***

2.2 Pathological Findings

Pathological findings signify the presence of abnormal conditions within the GI tract, often presenting as injuries or alterations in the normal mucosa observed during endoscopic procedures. A comprehensive examination of these findings is essential for determining appropriate treatment strategies and assessing patient prognosis. The KVASIR V2 dataset incorporates three classes related to pathological findings:

2.2.1 Esophagitis

Esophagitis refers to inflammation or swelling in the esophagus, often accompanied by breaks in the mucosa. The severity of irritation and inflammation depends on the size of the mucosal break and the extent of involvement. Acid reflux from the stomach or vomiting is a common cause of this condition. An example image of esophagitis is provided in Figure 2.4

Fig.2.4: ***Esophagitis***

2.2.2 Polyps

Polyps are abnormal growths in the intestinal mucosal lining and can assume various shapes. Distinguishable from normal mucosa by their texture and color, most polyps are benign but may pose a risk of cancer development over time. Timely detection and removal of polyps are crucial. Figure 2.5 illustrates an example image of polyps.

Fig. 2.5 ***Polyps***

2.2.3 Ulcerative Colitis

Ulcerative colitis is characterized by chronic inflammation affecting the large intestine. The severity of inflammation ranges from mild to severe, with symptoms varying accordingly. Severe cases may involve bleeding and swelling in the mucosal lining, while milder cases exhibit reddish and inflamed mucosa. An example image of ulcerative colitis is shown in Figure 2.6

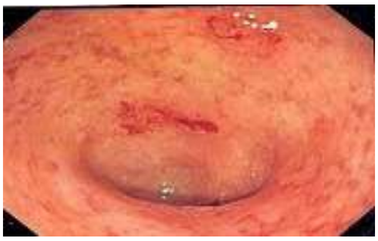

Fig. 2.6 ***Ulcerative Colitis***

2.2.4 Polyp Removal

Polyp removal is advisable during endoscopic examination to mitigate the risk of future cancer development. The KVASIR V2 dataset [29]includes two classes related to polyp removal:

2.2.5 Dyed and Lifted Polyp

Images in this class depict polyps lifted by injecting indigo carmine and saline, with the dyed area distinguishable from normal mucosa. An example image of this class is illustrated in Figure 2.7

Fig. 2.7 ***Dyed-Lifted-Polyps***

2.3 Dyed Resection Margins

Resection margins play a critical role in ensuring complete polyp removal. Incomplete removal may lead to lesion development from residual polyp cells. An example image of dyed resection margins is provided in Figure 2.8

Fig. 2.8 ***Dyed Resection Margins***

3. Proposed Methodology:

The steps of the proposed work are outlined in Graphical Abstract in Figure 1, and each step involved in the classification process is elaborated below:

1. **Data Collection:** The initial step involves collecting the KVASIR V2 dataset, which comprises endoscopic images categorized into eight classes representing various anatomical landmarks, pathological findings, and endoscopic procedures related to the GI tract.

2. **Data Preprocessing:** Following data collection, preprocessing is conducted to enhance the quality and suitability of the images for classification. This may include resizing, contrast enhancement, and scaling to standardize image attributes and improve interpretability.
3. **Model Selection:** Once preprocessed, the next step involves selecting appropriate models for image classification. In this proposed methodology, Convolutional Neural Networks (CNNs) and VGG16 models are chosen for their efficacy in handling image data and extracting meaningful features.
4. **Model Training:** Selected models are trained using the preprocessed dataset to learn patterns and features indicative of the eight classes of GI-tract abnormalities. Training involves iteratively adjusting model parameters to minimize classification errors and optimize performance.
5. **Model Testing:** Trained models are then evaluated using a separate portion of the dataset reserved for testing. This step assesses the models' ability to accurately classify images into the predefined classes and provides insight into their generalization performance.
6. **Performance Evaluation**: Performance metrics such as accuracy, precision, recall, specificity, and F1 score are calculated to quantitatively assess the efficacy of the classification models. These metrics provide valuable insights into the models' ability to correctly classify images across different classes and aid in comparative analysis.

Comparative Analysis: Finally, the results obtained from the proposed methodology are compared with existing literature and other published works to validate the efficacy and novelty of the approach. This step provides context and benchmarking against established methods and helps identify areas for further improvement. The proposed methodology offers a systematic approach to classifying GI-tract abnormalities using advanced deep learning techniques and aims to contribute to the field by providing accurate and efficient diagnostic capabilities.

3.1 Image Pre-processing:

In this study, the KVASIR V2 dataset, publicly available for researchers, is utilized for various technical research purposes. Before applying any classification techniques to the images in the dataset, preprocessing is imperative due to the presence of unwanted artifacts in many images. A challenge encountered with this dataset is the variation in image sizes, ranging from 720×576 up to 1920×1072 pixels. Additionally, many images contain black borders, which can hinder the efficiency of classification networks. To ensure the generalization of the network for application to other datasets and to address these challenges, the following pre-processing steps are performed:

3.1.1 Image Resize, Contrast Enhancement, Scaling:

Image Resize: The raw images are resized to a standardized dimension of 224 × 224 pixels. Resizing facilitates the application of Convolutional Neural Network (CNN) models, including ResNet50, VGG16, and InceptionV3, and the training of the classification network, while also reducing algorithm execution time.

Contrast Enhancement: Contrast enhancement is employed to improve image visibility by increasing the difference in brightness between items and their surroundings. This step is crucial for enhancing image quality, ensuring that important information is not lost during network training.

Scaling: Scaling is applied to stretch or compress images to a desired ratio relative to the original image size. This process helps provide a zoomed-in view of the image and reduces black borders. Given that many images in the KVASIR V2 dataset contain broad black borders, scaling is essential to maximize the image while partially removing these borders.

3.1.2 Data Augmentation:

To enhance the training dataset and improve the performance of the CNN models, data augmentation techniques are employed. With a total of 8000 images available in the KVASIR V2 dataset across eight different classes, 800 images are set aside for testing purposes. Data augmentation techniques such as vertical or horizontal flipping, rotation, image shearing, and image translation are applied.

Rotation: The authors have utilized rotation to generate additional images for each class. Specifically, three different degrees of rotation are applied to each class, resulting in a total of 28,800 images, including the original images. An example of an image with different degrees of rotation, along with the original image, is depicted in Figure 3.1

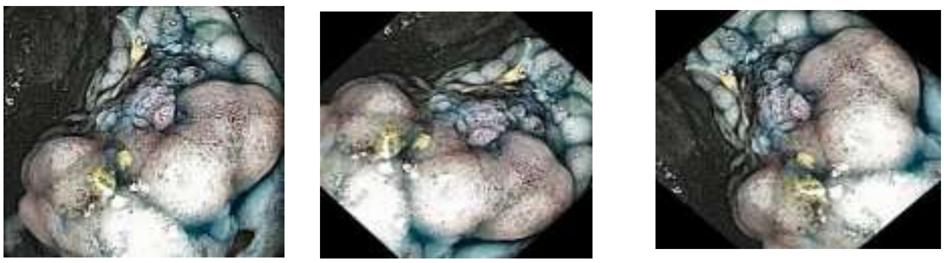

Fig 3.1. ***Three Rotations(45,90,135 respectively) of original image of Dyed Lifted Polyps***

In order to make the most of our few training examples, increase the accuracy of the model, and avoid overfitting, we augmented the data via a number of random transformations.The selected data augmentation techniques are listed below:

Rotation=45°,90°,135°,width_shift_range=0.2,height_shift_range=0.2,horizontal_flip=True,

vertical_flip=True.

4.Evaluation:

Performance evaluation metrics are employed to assess the classification model's effectiveness, including accuracy, recall, precision, F1 score, and specificity:

Accuracy, A= $\frac{\sum TD_k}{\sum TD_k+\sum_{k=1}^{n}\sum_{p=1}^{n} FD_{kp}}\Big|_{k\neq p}$

F1 − Score, F1 = 2 X $\frac{(PXR)}{(P+R)}$

Precision, P= $\frac{1}{n}\sum_{k=1}^{n}\frac{TD_k}{TD_k+\sum_{p=1}^{n} FD_{pk}}$

Recall, R= $\frac{1}{n}\sum_{p=1}^{n}\frac{TD_k}{TD_k+\sum_{k=1}^{n} FD_{pk}\big|_{p\neq k}}$

where TD and FD indicate the truly detected and falsely detected class. The total number of classes is generalized by the symbol ‘n’. These evaluation metrics provide comprehensive insights into the classification model's performance, including its ability to correctly classify different classes, detect true positives, minimize false detections, and achieve an overall balance between precision and recall.

4.1 Result

After training the model with 20 epochs and 32 as a batch size, we achieved an accuracy of 87.00% and 96.00% for CNN[30],VGG-16[32] and INCEPTION V3[33] networks respectively on Kvasir training set of 28,800 images(refer topic 3.1.2) from 8 classes. Progress for training loss and accuracy can be shown in below fig (4.1,4.2,4.3,4.4,4.5).

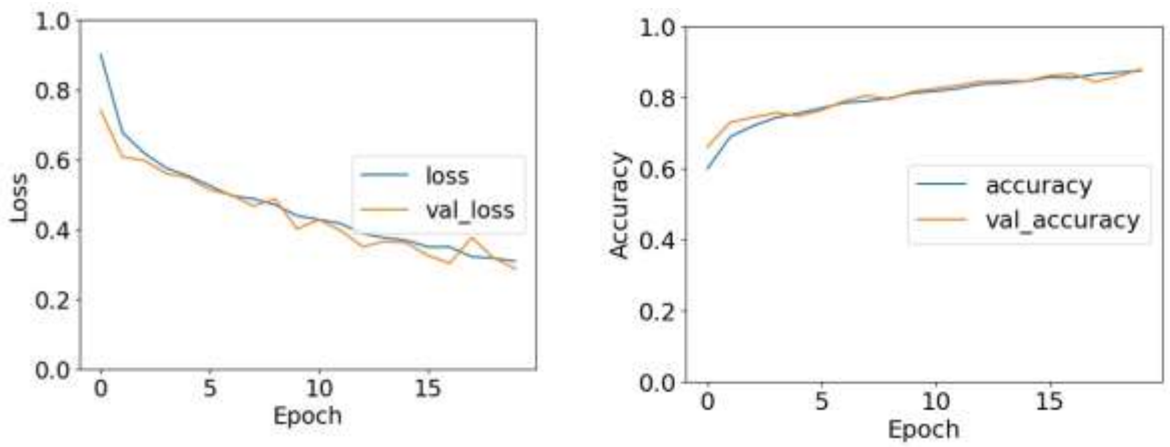

Fig 4.1***: Progress for CNN training loss and accuracy***

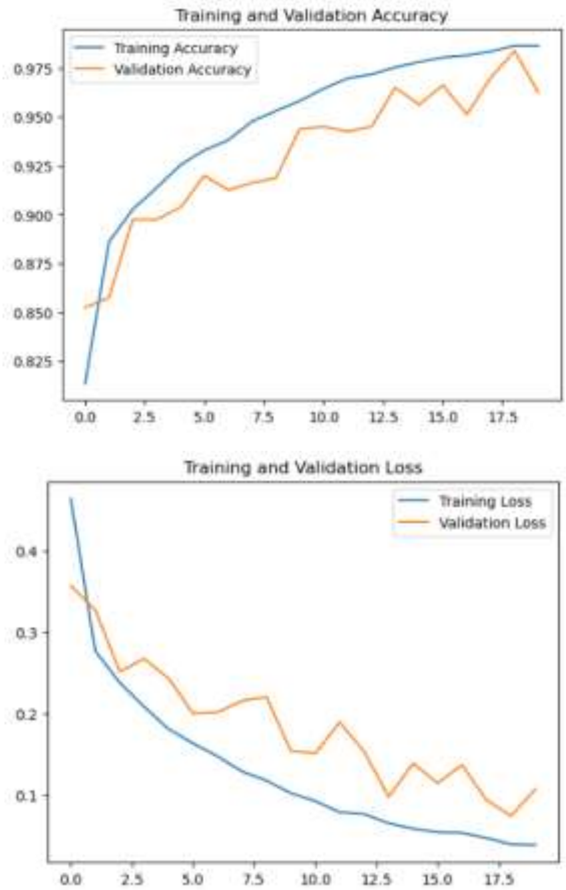

Fig 4.2***: Progress for VGG-16 training loss and accuracy***

4.2 Model Comparison:

The comparison between outputs of neural networks are shown in below bar chart in Fig.4.6. From the below figure, we clearly say that VGG-16 & RESNET-50 exhibited a good performance than CNN & INCEPTION V3. The train and test accuracies of the INCEPTION V3 is very less when compared with the CNN, INCEPTION V3 exhibited some overfitting while training.

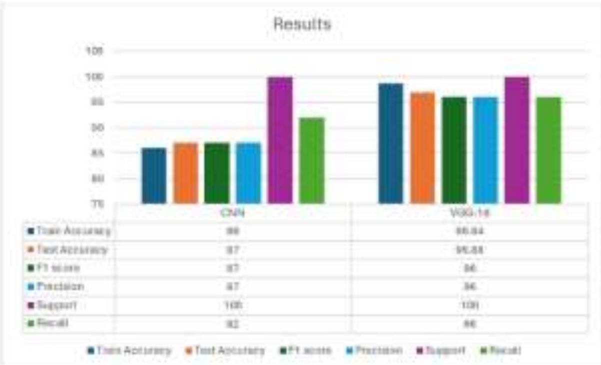

Fig 4.6 ***Comparison of models***

The F1 score, Precision, Recall of INCEPTION V3 are also less than that of other 3 models. The total of 100 images from each class are used to test the models. The output values are plotted as a bar graph for better understanding of Models.

5. Conclusion:

From this project report, we conclude that the gastrointestinal diseases detection using neural networks revealed varying performance among different architectures. VGG16 exhibited the highest accuracy, suggesting their suitability for the task due to their robust feature extraction capabilities. Conversely, Simple CNN showed poor performance, indicating potential limitations in capturing the complexities of the dataset. Images are classified as being or involving an anatomical landmark (pylorus, z-line, cecum), a diseased state (esophagitis, ulcerative colitis, polyps), or a medical procedure (dyed lifted polyps, dyed resection margins). The resulted accuracy after testing each model shows that VGG-16 model was the best model with accuracy of 96.88%.

References

(1). Al Barsh, Y. I., et al. (2020). "MPG Prediction Using Artificial Neural Network." International Journal of Academic Information Systems Research (IJAISR) 4(11): 7-16.

(2). Alajrami, E., et al. (2019). "Blood Donation Prediction using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(10): 1-7.

(3). Alajrami, E., et al. (2020). "Handwritten Signature Verification using Deep Learning." International Journal of Academic Multidisciplinary Research (IJAMR) 3(12): 39-44.

(4). Al-Araj, R. S. A., et al. (2020). "Classification of Animal Species Using Neural Network." International Journal of Academic Engineering Research (IJAER) 4(10): 23-31.

(5). Al-Atrash, Y. E., et al. (2020). "Modeling Cognitive Development of the Balance Scale Task Using ANN." International Journal of Academic Information Systems Research (IJAISR) 4(9): 74-81.

(6). Alghoul, A., et al. (2018). "Email Classification Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 8-14.

(7). Abu Nada, A. M., et al. (2020). "Age and Gender Prediction and Validation through Single User Images Using CNN." International Journal of Academic Engineering Research (IJAER) 4(8): 21-24.

(8). Abu Nada, A. M., et al. (2020). "Arabic Text Summarization Using AraBERT Model Using Extractive Text Summarization Approach." International Journal of Academic Information Systems Research (IJAISR) 4(8): 6-9.

(9). Abu-Saqer, M. M., et al. (2020). "Type of Grapefruit Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 4(1): 1-5.

(10). Afana, M., et al. (2018). "Artificial Neural Network for Forecasting Car Mileage per Gallon in the City." International Journal of Advanced Science and Technology 124: 51-59.

(11). Al-Araj, R. S. A., et al. (2020). "Classification of Animal Species Using Neural Network." International Journal of Academic Engineering Research (IJAER) 4(10): 23-31.

(12). Al-Atrash, Y. E., et al. (2020). "Modeling Cognitive Development of the Balance Scale Task Using ANN." International Journal of Academic Information Systems Research (IJAISR) 4(9): 74-81.

(13). Alghoul, A., et al. (2018). "Email Classification Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 8-14.

(14). Al-Kahlout, M. M., et al. (2020). "Neural Network Approach to Predict Forest Fires using Meteorological Data." International Journal of Academic Engineering Research (IJAER) 4(9): 68-72.

(15). Alkronz, E. S., et al. (2019). "Prediction of Whether Mushroom is Edible or Poisonous Using Back-propagation Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(2): 1-8.

(16). Al-Madhoun, O. S. E.-D., et al. (2020). "Low Birth Weight Prediction Using JNN." International Journal of Academic Health and Medical Research (IJAHMR) 4(11): 8-14.

(17). Al-Massri, R., et al. (2018). "Classification Prediction of SBRCTs Cancers Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 1- 7.

(18). Al-Mobayed, A. A., et al. (2020). "Artificial Neural Network for Predicting Car Performance Using JNN." International Journal of Engineering and Information Systems (IJEAIS) 4(9): 139-145.

(19). Al-Mubayyed, O. M., et al. (2019). "Predicting Overall Car Performance Using Artificial Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(1): 1- 5.

(20). Alshawwa, I. A., et al. (2020). "Analyzing Types of Cherry Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 4(1): 1-5.

(21). Al-Shawwa, M., et al. (2018). "Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(9): 1-6.

(22). Ashqar, B. A., et al. (2019). "Plant Seedlings Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 3(1): 7-14.

(23). Bakr, M. A. H. A., et al. (2020). "Breast Cancer Prediction using JNN." International Journal of Academic Information Systems Research (IJAISR) 4(10): 1-8.

(24). Barhoom, A. M., et al. (2019). "Predicting Titanic Survivors using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(9): 8-12.

(25). Belbeisi, H. Z., et al. (2020). "Effect of Oxygen Consumption of Thylakoid Membranes (Chloroplasts) From Spinach after Inhibition Using JNN." International Journal of Academic Health and Medical Research (IJAHMR) 4(11): 1-7.

(26). Dalffa, M. A., et al. (2019). "Tic-Tac-Toe Learning Using Artificial Neural Networks." International Journal of Engineering and Information Systems (IJEAIS) 3(2): 9-19.

(27). Dawood, K. J., et al. (2020). "Artificial Neural Network for Mushroom Prediction." International Journal of Academic Information Systems Research (IJAISR) 4(10): 9-17.

(28). Dheir, I. M., et al. (2020). "Classifying Nuts Types Using Convolutional Neural Network." International Journal of Academic Information Systems Research (IJAISR) 3(12): 12-18.

(29). El-Khatib, M. J., et al. (2019). "Glass Classification Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 3(2): 25-31.

(30). El-Mahelawi, J. K., et al. (2020). "Tumor Classification Using Artificial Neural Networks." International Journal of Academic Engineering Research (IJAER) 4(11): 8-15.

(31). El-Mashharawi, H. Q., et al. (2020). "Grape Type Classification Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 3(12): 41-45.

(32). Elzamly, A., et al. (2015). "Classification of Software Risks with Discriminant Analysis Techniques in Software planning Development Process." International Journal of Advanced Science and Technology 81: 35-48.

(33). Elzamly, A., et al. (2015). "Predicting Software Analysis Process Risks Using Linear Stepwise Discriminant Analysis: Statistical Methods." Int. J. Adv. Inf. Sci. Technol 38(38): 108-115.