Journal of Information Systems Engineering and Management
2025, 10(19s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Deep Learning Based Early Detection of Gastrointestinal Disorders

[1]Naveen Kumar, [2]Dr. R. Priya, [3]Dr. Sudhansh Sharma,
[1]Ph.D. Research Scholar, Department of Computer Science & Engineering, Annamalai University, Tamilnadu (India)
[2]Professor, Department of Computer Science & Engineering, Annamalai University, Tamilnadu (India)
[3]Assistant Professor, School of Computer and Information Sciences, IGNOU, New Delhi (India)

ARTICLE INFO

Received: 05 Oct 2024

Revised: 05 Dec 2024

Accepted: 22 Dec 2024

ABSTRACT

Each year, around two million people die as a result of gastrointestinal problems worldwide. Medical experts use advanced video endoscopic imaging technology to diagnose and treat gastrointestinal issues such as ulcers, haemorrhaging, and polyps. However, due to the large volume of images produced by medical video endoscopy, healthcare personnel must devote significant time to assessing and interpreting these images. Given the difficulties of manual diagnosis, researchers are investigating computer-aided methods that can accurately and quickly identify all generated images. This research work compares three deep learning networks, GoogleNet, ResNet-50, and AlexNet, using the Kvasir dataset. AlexNet surpassed the other two deep learning networks, GoogleNet and ResNet-50, with a 97% accuracy rate, 96.8% sensitivity, 99.20% specificity, and 99.98% AUC (Area under the Curve).

Keywords: Gastrointestinal (GI) Tract Disease, Endoscopic Gastrointestinal Tract Image, Gastrointestinal Diseases Diagnosis, Deep Learning Network, Convolution Neural Network (CNN), Deep Learning Architecture.

1. INTRODUCTION

Gastrointestinal disorders pose a significant health risk, with nearly two million fatalities occurring worldwide each year as a result of these conditions. To diagnose and manage issues such as ulcers, bleeding, and polyps, medical professionals increasingly rely on advanced medical imaging and cutting-edge video endoscopic technologies. However, the extensive volume of images produced during medical video endoscopy requires physicians to invest considerable time in their evaluation and interpretation. This manual diagnostic process can be quite arduous, prompting researchers to investigate computer-aided systems designed to facilitate the rapid and precise analysis of the images generated. These innovative methods aim to enhance diagnostic accuracy and efficiency in identifying gastrointestinal disorders.

GoogleNet, ResNet-50, and AlexNet are all examples of Convolutional Neural Networks (CNNs), a form of deep learning architecture that is widely used for image categorisation, object detection, and other visual recognition applications. In the goal of computer-assisted gastrointestinal ailment identification, appropriate datasets and precise machine learning models for classification can benefit both healthcare providers and patients. The Kvasir dataset is extensively used by researchers to evaluate their models for early identification of gastrointestinal disorders. This research work compares Three neural networks based on deep learning techniques, GoogleNet, ResNet-50, and AlexNet, using the Kvasir dataset.

This research work evaluates the diagnosis effectiveness of several lower gastrointestinal disorders by comparing three deep learning architectures: GoogleNet, ResNet-50, and AlexNet, using the Kvasir dataset. The Kvasir dataset comprises 5,000 images, uniformly classified into five categories: normal cecum, polyps, ulcerative colitis, and dyed-lifted polyps, indicative of lower gastrointestinal disorders. All images are enhanced and subjected to noise reduction before to being processed using deep learning networks, namely GoogleNet, ResNet-50, and AlexNet.

During the classification phase, transfer learning is used to improve the performance of convolutional neural network (CNN) models on previously learnt tasks. The deep feature vector is activated using the softmax function, resulting in the classification of the input images into five distinct categories. All CNN models delivered amazing results. However, it has been highlighted that AlexNet outperformed the other two deep learning networks - GoogleNet and ResNet-50.

The subsequent sections of the paper are structured as follows. Section 2 pertains to the literature review, Section 3 addresses the datasets and deep learning models employed in the process. Section 4 delineates the adopted methodology. Section 5 presents the experimental results, analysis and findings of this research work, along with a comparison of the proposed systems to those identified in previous research. The Section 6 gives the conclusions.

2. LITERATURE REVIEW

Stomach cancer is a significant contributor to global mortality, ranking as the leading cause of cancer-related deaths. According to the World Health Organization (WHO) [1], approximately 1.8 million individuals succumb to gastrointestinal disorders each year. Additionally, cancer of the digestive tract stands as the 4^{th} most common cause of death worldwide. Intestinal polyps, which are abnormal growths on the linings of the stomach and colon, serve as early indicators of potential gastrointestinal cancer. These polyps typically develop slowly [2], often remaining undetected until they reach a considerable size. However, if identified at an early stage, there are effective strategies for prevention and treatment [3].

Video endoscopy is crucial in decreasing mortality rates and improving the early identification of gastrointestinal polyps [4]. By analyzing mucosal patterns, including variations in color and texture on the mucosal surface of the gastrointestinal tract, endoscopy aids in assessing the severity of ulcerative colitis [5]. Although a video of the gastrointestinal tract can produce numerous images, physicians often lack the time to review them all, and only a limited number of these images may indicate disease symptoms. Consequently, the accuracy of diagnoses is heavily reliant on the physician's expertise, with specialists able to identify polyps in approximately 27% of cases [6]. The potential invisibility of polyps during examinations raises concerns about the risk of future cancer development.

For very accurate polyp diagnosis and early cancer detection, a computer-aided automated method might be beneficial. This is due to the fact that a misdiagnosis may result from the shortcomings of radiologists and other human elements. Artificial intelligence systems have been shown to have the capacity to help people identify illnesses that are invisible to the human eye [7–9]. Many different medical specialities have shown this potential. Artificial intelligence-based techniques, for instance, may be able to identify important elements and extract complex micro-images from endoscopic images. Artificial intelligence methods are capable of distinguishing between neoplastic and non-neoplastic tissues in clinical settings. Additionally, there are ways to assess the risk of stomach cancer by extracting textural characteristics [10, 11]. A comprehensive analysis of the textural features in colonoscopy images has significantly contributed to the classification of colitis [12, 13]. Nevertheless, the precision of these diagnoses is constrained by the challenges associated with capturing the imaging characteristics of the gastrointestinal tract.[14]

In earlier times, the process of diagnosing diseases through endoscopic images relied heavily on trial and error [15, 16]. Recent advancements in machine learning techniques have significantly enhanced this approach, allowing for the extraction of valuable information regarding color, texture, and edge details from these images. The use of convolutional neural networks (CNNs) in supervised learning has resulted in a notable increase in the accuracy of medical image diagnoses [17]. Medical diagnosis accuracy has increased as a result of the use of CNNs, which have begun to overcome these feature engineering restrictions. The incorporation of engineering into the learning process has allowed CNN to exhibit its remarkable ability to extract attributes [18]. When it comes to diagnosing medical image, deep learning algorithms have been shown to perform better than experts [19]. Consequently, computer-aided diagnostics that employ deep learning algorithms for endoscopic images could potentially attain diagnostic accuracy comparable to or even surpassing that of skilled medical practitioners [20].

Gunasekaran H et al. [21] demonstrated how to extract colour patterns that can be used to detect colon polyps by using wavelet decomposition. Studies have used a large amount of machine learning to help medical personnel identify the type of gastrointestinal imaging issue. They have used a variety of approaches, including wavelets, context-based features, edge shape, valley information, polyp-based local binary, and grey-level co-occurrence matrices (GLCMs). [22 – 24]. The approach proposed by Sharif M. et al. [25] has demonstrated superior effectiveness compared to alternative methods. To successfully extract handcrafted features, it is essential to consider several factors, including light reflection, camera positioning, and any structural irregularities present in the object. CNN's techniques are incredibly good at extracting deep characteristics, and in recent years, the network has made some progress in diagnosing medical images.

3. DATA COLLECTION AND NEURAL NETWORK MODELS'

Computer-assisted automatic identification of gastrointestinal disorders is one of the most challenging domains of research. To achieve the objective of computer-assisted identification of gastrointestinal disorders, sufficient datasets and accurate machine learning classification models can benefit both healthcare professionals and patients. This section offers comprehensive descriptions of the datasets and CNN models, including GoogleNet, ResNet-50, and AlexNet. These models are intended to facilitate the prompt and precise identification of decrease gastrointestinal disorders. Researchers utilise the Kvasir dataset to assess their algorithms for the preliminary identification of gastrointestinal disorders.

3.1 Dataset

It is crucial to assess the merits and drawbacks of gastrointestinal problem databases for their accessibility, size, content, and annotations. This comparison will yield a comprehensive grasp of several significant datasets. This research work focuses on the Gastrointestinal Image Dataset (Kvasir), comprising around 8,000 images and video frames obtained from gastrointestinal examinations. It encompasses a range of abnormalities, including oesophagitis, polyps, and various gastrointestinal disorders. Three significant anatomical landmarks and three therapeutically pertinent discoveries are delineated in the collection of images. Furthermore, it comprises two distinct categories of images pertaining to endoscopic polyp ectomy. Medical practitioners, recognised as competent endoscopists, analyse and provide commentary on the dataset. Kvasir plays a pivotal role in the research of computer-aided detection for both singular and many diseases. The dataset is highly beneficial for training machine learning models to address various gastrointestinal illnesses due to its comprehensive annotations and diverse variants. Figure-1 illustrates images from the Kvasir dataset.

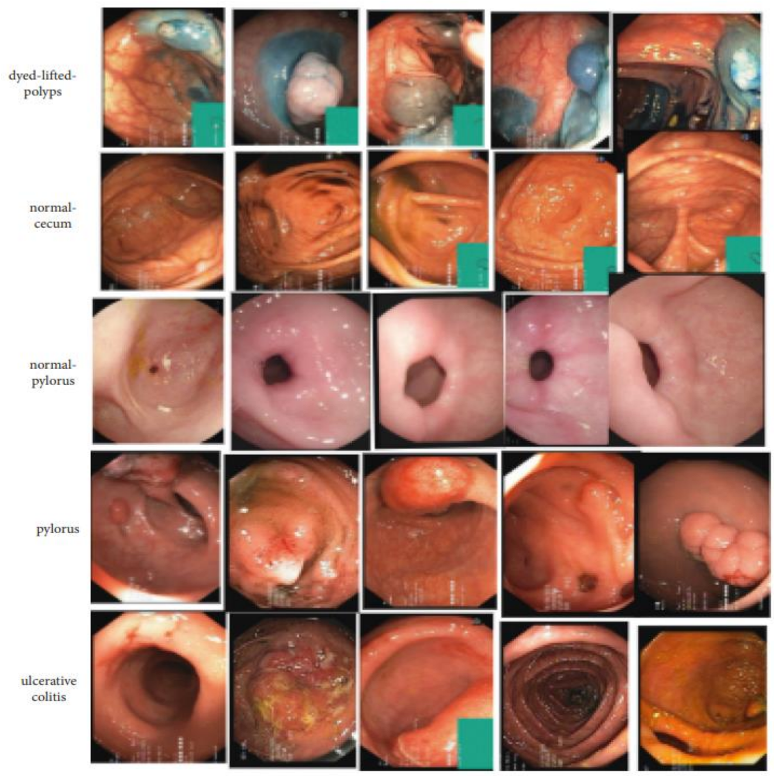

Figure 1: Illustrate images from the Kvasir dataset.

3.2 Deep Learning Models

The following subsection relates to the brief introduction of deep learning models considered under this proposed research work

***3.2.1 ResNet50*:** Figure 2 depicts the ResNet-50 architecture used to classify 5,000 images into five disorders of the lower digestive tract. The ResNet-50 Structure consists of sixteen blocks, each with 177 levels. The layers are divided into two categories: convolutional layers with 49 levels and various filters, and input layers that accept RGB images of 224 × 224 pixels. The convolutional layer collects deep features from input images and saves them as deep feature vector maps. Furthermore, a pooling layer can do both maximum and average pooling. The two layers reduce the dimension of the feature vector map. Batch normalisation is used to help the network reliably determine the ideal learning rate. The Rectified Linear Unit Activation function (ReLU) only allows positive outputs to pass. The Activation layers come before this function. All negative values will be quickly converted to zero. Each interconnected layer generates 4096 characteristics and receives 9216 characteristics. The second convolutional layer creates 1,000 features. During differentiation, the SoftMax layer separates into five classes: ulcerative colitis, normal cecum, normal pylorus, polyps, and coloured elevated polyps.

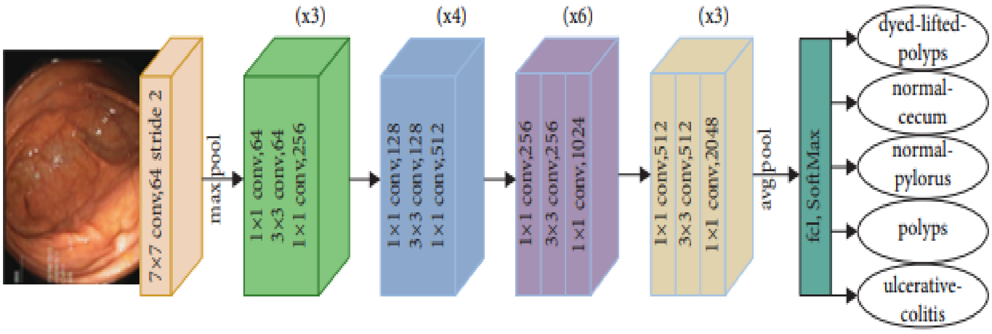

Figure-2: ResNet-50 Structure used to classify 5,000 images divided into five diseases from the lower GI (Gastrointestinal) Tract

3.2.2 GoogleNet: GoogleNet finds usage in many different areas, including medical image categorisation and computer vision. GoogleNet's convolutional neural network was built entirely on Inception design. In virtue of Inception modules, the network can choose from a spectrum of convolutional filter sizes throughout every block. These components are framed vertically in an Inception network, occasionally using stride-two max-pooling layers to halley the grid's resolution. GoogleNet's architectural composition consists of the Initial Layers (standard convolutional layers followed by max pooling to process the input image), the Inception Modules (stacked inception modules that capture multi-scale features by parallel convolutional and pooling operations), the Auxiliary Classifiers (positioned after some Inception modules to aid in training and serve as regularizers), and the Final Layers (global standard pooling layer tracked by a SoftMax layer for classification).

GoogleNet's design has 27 levels, some of which lack explicit parameters. The input layer receives images formatted in an RGB scheme with dimensions of 224 by 224 pixels. This layer contains numerous strata. The initial convolutional layer comprises two 7 x 7 filters. These are surrounded by the largest screens in the system. This flap diminishes the size of the provided image. The subsequent layer is a activation layer with a 3x3 filter, succeeded by a max pooling layer with a 3x3 filter, then another max pooling layer with a 3x3 filter. The output initially proceeds to a block inception module comprising two levels. A maximum pooling layer utilising three-by-three filters follows. Finally, a four-layer block employs an application derived from the Inception module. The subsequent phases comprise a two-layer block origin part, tracked by a max pooling layer utilising 3x3 filters, and concluding with an additional max pooling layer employing 3x3 filters. The standard pooling layer measures 7 pixels by 7 pixels. Stride is a statistic for measuring the extent of filter displacement applied to an image. This mitigates overfitting, a primary purpose of the Dropout approach. During the project, we utilised a Dropout setting of 40%. This indicates that distinct settings are employed for each iteration, and the neurons are deactivated by 40% during every repetition. 9,216 features were received, and 4,096 features were produced by the completely connected layer. The SoftMax layer generates five unique classes. These comprise a normal cecum, a normal pylorus, ulcerative colitis, polyps, and dyed-lifted polyps. Figure 3 below illustrates the GoogleNet structure employed to segregate 5,000 images appear into five illnesses of the lower GI (Gastrointestinal) Tract.

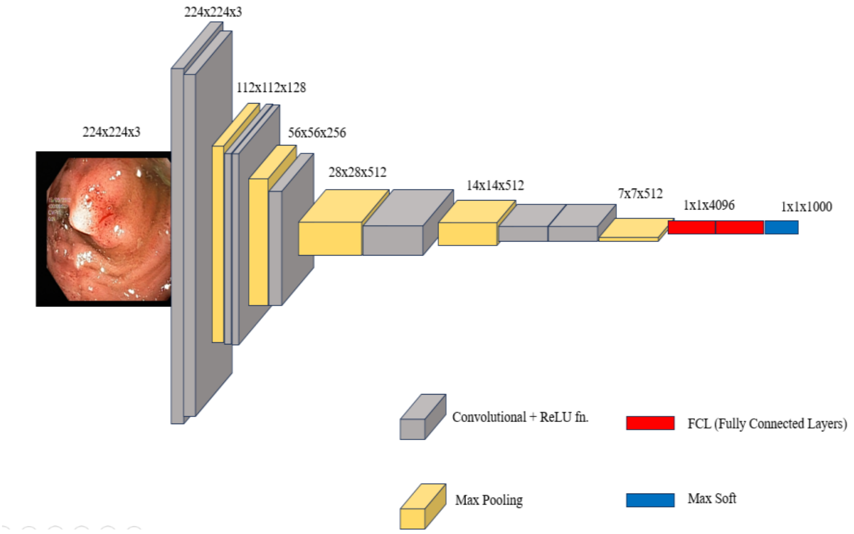

Figure-3: The GoogleNet structure employed to segregate 5,000 images appear into five disorder of the lower GI(Gastrointestinal) Tract

3.2.3 AlexNet: Figure 4 illustrates the AlexNet architecture employed to segregate 5,000 images appear into five disorders of the GI (Gastrointestinal) Tract. The structure consists of five activation layers utilising distinct filters, along with an embedding layer that accepts R (Red), G (Green), B (Blue) images measuring 227 × 227 pixels. This service is delivered by a convolutional layer that pulls profound features from input images. The three layers utilized in max pooling play a crucial role in diminishing the dimensions of the feature vector maps. Max pooling is a novel technique. The two layers of cross-channel normalization are responsible for re-parameterizing the vector weights and determining the suitable learning rate. The seven layers that make up the ReLU (Rectified linear unit) are positioned subsequent to the convolutional layers.. ReLU exclusively produces positive values, rendering negative values as zero. Three interconnected layers collaborate to produce a sequence. The initial layer of connectivity generates 4096 features and receives 9216 unique features. The third layer that is fully connected generates 1000 neurons, which serve as features, whereas The subsequent interconnected layer yields a total of 4096 features. The SoftMax layer is comprised of five distinct categories: polyps, dyed-lifted polyps, normal cecum, normal pylorus, and ulcerative colitis (Figure 4).

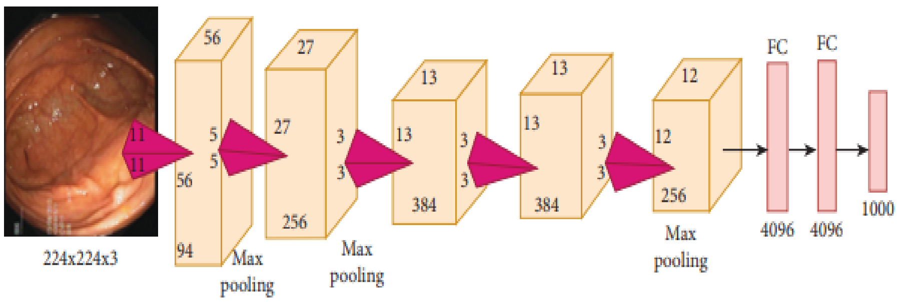

Figure-4: The AlexNet structure used to classify 5,000 images divided into five disorders from the lower GI (Gastrointestinal) Tract

AlexNet usually perform better than other CNN models due to simpler architecture and having less parameters for image processing, making AlexNet faster to train and run. However, AlexNet may perform better with high computational yield where dataset is relatively smaller [39].

4. METHODOLOGY

This section provides detailed descriptions of the methodology and data preprocessing employed to categorise lower digestive tract illnesses utilising CNN models such as GoogleNet, ResNet-50, and AlexNet. Figure 5 below illustrates the detection of gastrointestinal disorders utilising CNN models such as GoogleNet, ResNet-50, and AlexNet.

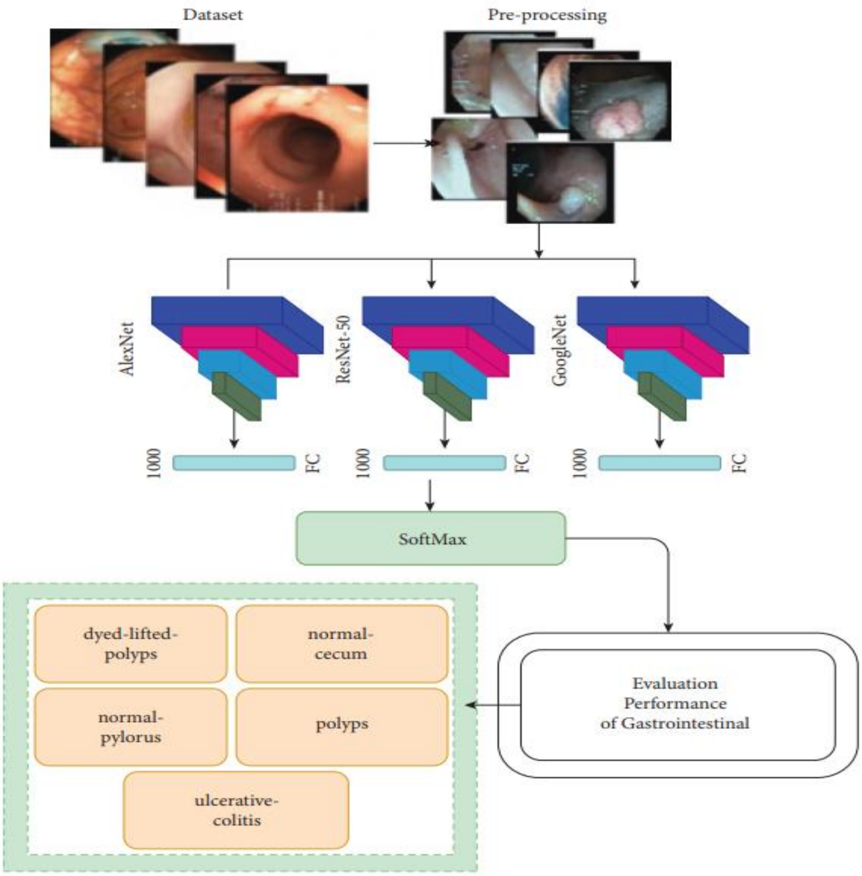

Figure 5: Process of gastrointestinal disease detection through the CNN models such as GoogleNet, ResNet-50, and AlexNet.

4.1 Preprocessing and Augmentation Techniques

Researchers heavily rely on the Kvasir dataset to assess their algorithms for early gastrointestinal disease detection. Dataset preprocessing removes noise and artefacts, therefore enhancing images. Conversely, image augmentation technologies improve the training process. CNN performance is shown to be negatively impacted by light reflections, camera angles, and mucosal membranes enclosing internal organs by means of noise and aberrations, therefore complicating feature extraction [27]. As such, researchers have become rather interested in improving techniques to raise image quality. In this research work, the gastrointestinal images underwent pre-processing prior to being input into the CNN models. The images were resized to dimensions of 244 x 244 pixels for the GoogleNet and ResNet-50 models, and 227 x 227 pixels for the AlexNet model, ensuring consistent colour uniformity. Then for the gastrointestinal images, the three RGB channels' mean values were calculated. The enhancement technique used the average filter, which replaced the value of every pixel with the average of its neighbours. Every pixel in the image underwent this procedure [27, 28]. CNN techniques mostly rely on the volume of data accessible. Increasing the training data helps the model to perform better. By means of compensation for the dearth of medical images, data augmentation techniques improve CNN model accuracy [26, 29]. When there is a disparity in the quantity of images between several classes, data augmentation technology can help to balance the dataset. Using graphics by flipping, zooming, shifting, and rotating [30] this research work improved the training data.

4.2 Convolutional Layers.

The gastrointestinal system dataset contains a wide variety of properties, including colour, shape and consistency. Manual feature extraction requires a lot of information, particularly when extracting images from a movie. This is particularly evident when the illness is only visible in a small number of images, making it difficult for the radiologist and other specialists to detect it. Convolutional layers are used by CNN algorithms to extract features that are specific to each disorder and that characterise the entire illness. There are nine origin layers and several convolutional layers in the GoogleNet network. AlexNet employs five convolutional layers, while ResNet-50 has 49. During training, these layers handle the deep features using a combination of filters and weights. The filters and weights are then transferred to the following phase. The average and maximum pooling levels may also contribute to the feature maps' reduced size. These layers use either the maximum or the average amount of pixels to display a collection of pixels. The task of the convolutional layers is to extract distinct characteristics from every image, totalling 9216 features. These characteristics are then displayed as feature maps prior to being incorporated into the categorisation layers.

4.3 Normalisation of the Image.

Normalisation is an approach used in deep neural network training to standardise images and speed up the training process. Normalisation aids in the selection of an optimal learning rate by allowing gradient descent to converge faster. Learning without the normalisation process is time consuming and difficult. In our investigation, we normalised each pixel's image by removing the mean of the entire training set. The dataset was examined to determine the variance, which was then divided by each pixel. This technique resulted in data centring, which assured that the variance of each characteristic was equal to 1.

4.4 Dropout Technology.

To avoid overfitting, CNNs create millions of parameters, and to lessen the hazards associated with overfitting, CNNs employ a technique known as dropout. During our experiment, we used the Dropout method with GoogleNet, ResNet50, and AlexNet, all of which had a withdrawal rate of 40%. This indicates that 40% of the neurons were halted during each cycle. As a result, each iteration of the network uses a unique set of parameters. In contrast, the Dropout strategy requires twice as much training time as the other way.

4.5 Transfer Learning.

The Convolutional neural networks (CNN) depend critically on transfer learning [31]. We used transfer learning and fine-tuning techniques on the pre-trained GoogleNet, ResNet-50, and AlexNet networks, which were trained on the ImageNet dataset, a vast and well-annotated dataset that has been absolutely vital in training and evaluating state-of- the-art machine learning models, so enabling a fundamental component in the advancement of modern artificial intelligence technology [32]. Training a dataset to solve a particular problem and subsequently using that knowledge to another but similar problem with another dataset is known as transfer learning [33, 28]. Using transfer learning means selecting a pretrained model and determining the task size. Then, the knowledge acquired from the pretrained model helps to enhance the generalisation of another employment. Transference learning lowers overfitting risk. Using GoogleNet, ResNet-50, and AlexNet models, this work applied transfer learning with weight fine-tuning. Trained on ImageNet, the GoogleNet, ResNet-50, and AlexNet models were subsequently used on the gastrointestinal dataset. A completely linked layer took place in place of the last three levels of the designs. The first connected layer produced 4,096 neurons and had 9,216 neurons. The second linked layer also emitted and received 4,096 neurons. Five distinct groups were found by the softmax layer: normal cecum, normal pylorus, polyps, dye-lifted polyps, and ulcerative colitis.

4.6 Adam - Optimizers

Optimizers modify and fine-tune the parameters of a neural network, such as weights, biases and learning rate, in order to minimise loss. An analysis is being conducted on optimizer algorithms in order to enhance the deep learning classifier, hence accelerating the performance of the model. Adam is a leading optimizer in the field of deep learning. Adam is an amalgamation of the RMSProp and momentum optimisation algorithms [34]. Adam computes the adaptive learning rate for every parameter. The algorithm maintains the average previous gradient, referred to as momentum m_t, and stores the squared gradients as a decaying average v_t. The equation shown below outlines Adam's methodology for optimising parameters, including the learning rate and other related factors.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \tag{1}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \tag{2}$$

where m_t refers to the first moment in the gradient, v_t refers to the second moment in the gradient, and β_1 and β_2 indicate the decay rate.

5. EXPERIMENTAL RESULTS

Three tests are carried out to evaluate a dataset connected to the discipline of gastroenterology are presented in this work. There are 5,000 images in the dataset, equally distributed among all the disease kinds. Every three research made use of the same sample size. Training took up eighty percent of the data; the remaining twenty percent was left aside for selection and validation. During the training phase, the CNNs GoogleNet, ResNet-50, and AlexNet had their weights and settings changed in order to assess the dataset of gastrointestinal diseases. Table-1 lists the three different networks' possible training options together with the associated times needed to finish each one.

A dataset evaluation and three distinct Convolutional Neural Network (CNN) models are displayed in Table-2. One can calculate accuracy, sensitivity, specificity, and area under the curve (AUC) using equations (3) through (6). Based on the statistics, Table-2 demonstrates that each network produced favourable outcomes. It is crucial to emphasise that all of the research' findings were quite consistent.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} * 100\%\,,$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\%,$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} * 100\%,$$

$$\text{AUC} = \frac{\text{true positive rate}}{\text{false positive rate}} = \frac{\text{Sensitivity}}{\text{Specificity}} * 100$$

The total number of accurately identified positive samples is represented by TP. The number of accurately categorised negative samples is known as TN. The number of benign samples that are labelled as malignant is known as FP. The total number of malignant samples that were mistakenly classified as benign is represented by the symbol FN.

For each disorder, Table-3 displays the classification results from GoogleNet, ResNet-50, and AlexNet. GoogleNet performed best when it came to classifying ulcerative colitis, with an efficiency of 96%, while ResNet-50 achieved a classification efficiency of 98% for polyps. As indicated in Table 4, a number of measures that have been previously examined in the research literature were employed to assess the CNNs that are now being considered. Compared to earlier research, which revealed accuracy values ranging from 70.40 percent to 90.20 percent, our suggested strategy yielded an accuracy rate of 97%, which is noticeably higher. With a sensitivity of 96.80%, our suggested approach outperformed earlier research studies that showed sensitivity ranging from 70.40 percent to 95.16 percent. The technique we suggested achieved a specificity of 99%, whereas the amount of specificity seen in previous trials ranged from 70.90 percent to 93 percent. The suggested system's area under the curve (AUC) value of 99.99% was higher than any of the previous research results.

Table-1: Deep learning network training parameter configuration options

TRAINING PARAMETERS						
Deep Learning Networks	**Optimizer**	**Batch Size (Min.)**	**Epochs (Max.)**	**Learning Rate**	**Validation Frequency**	**Training Time (Sec.)**
GoogleNet	Adam	20	10	0.0003	3	36501
ResNet50	Adam	20	4	0.0001	5	41605
AlexNet	Adam	128	10	0.0003	50	637

Table-2: Results of Diagnostic outcomes of gastrointestinal disorders implementing Deep Learning approaches

PERFORMANCE EVALUATION PARAMETERS					
Deep Learning Networks	**Accuracy(%)**	**Sensitivity(%)**	**Specificity(%)**	**Validation Accuracy(%)**	**AUC(%)**
GoogleNet	96.70	96.60	99.00	96.70	99.99
ResNet50	99.00	95.00	97.00	98.00	95.00
AlexNet	97.00	96.80	99.20	94.88	99.98

Table-3: **Performance evaluation results for Kvasir - the gastrointestinal disease dataset**

GASTROINTESTINAL DISEASE					
Deep Learning Networks	**Dyed Lifted polyps**	**Normal cecum**	**Normal pylorus**	**Polyps**	**Ulcerative colitis**
GoogleNet	98	93	100	96	96
ResNet50	99	95	97	98	95
AlexNet	99	95	99	88	93

Table-4: **Performance based comparison of our proposed models with the models of other researchers**

PERFORMANCE PARAMETERS				
Deep Learning Networks	**Accuracy(%)**	**Sensitivity(%)**	**Specificity(%)**	**AUC(%)**
Bour et.al [35]	87.10	87.10	93.00	---
Zhang et.al.[36]	70.40	70.40	70.90	90.00
Zhu et.al.[37]	85.00	83.00	82.54	---
Fonollas et.al.[38]	90.20	90.10	90.30	97.00
Proposed Model-1: GoogleNet	96.70	96.60	99.00	99.99
Proposed Model-2: ResNet50	99.00	95.00	97.00	95.00
Proposed Model-3: AlexNet	97.00	96.80	99.20	99.98

6. CONCLUSION

This research work presents a strong and efficient approach for the Kvasir dataset's gastrointestinal disease classification, therefore enabling quick identification of gastrointestinal diseases. According to the conducted research, still there are many personal elements that could lead to the erroneous diagnosis of gastrointestinal diseases. Three deep learning models: ResNet-50, AlexNet, and GoogleNet are presented in this research work. These models can enable doctors concentrate on once neglected but now important areas. While still 80% of the dataset was used for training, the 20% was kept aside for testing and validation. Processing helped the images to be improved by removing any noise and artefacts. The data augmentation method which entails the image duplication during the training phase is used to increase precision. Each image is assigned by the softmax layer to one of the five categories of gastrointestinal illnesses, therefore producing five different classes. All three of them performed really satisfactorily. Nonetheless, when compared to the ResNet-50 and GoogleNet, the AlexNet shows to be performing really well.

Data Availability: The data are available in the following link: https://datasets. simula.no/kvasir/#download.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

REFERENCES

[1] Perrin, A., 2023. Anatomy and physiology of the gastrointestinal tract and associated disease processes. In: Stoma Care Specialist Nursing: A Guide for Clinical Practice. pp. 39-54.
[2] Mahmoudi, H. and Hossainpour, H., 2023. Application and development of fecal microbiota transplantation in the treatment of gastrointestinal and metabolic diseases: A review. Saudi Journal of Gastroenterology, 29(1), pp. 3-11.
[3] Mohapatra, S., Pati, G.K., Mishra, M. and Swarnkar, T., 2023. Gastrointestinal abnormality detection and classification using empirical wavelet transform and deep convolutional neural network from endoscopic images. Ain Shams Engineering Journal, 14(4), p. 101942.
[4] Uboldi, M., Melocchi, A., Moutaharrik, S., Palugan, L., Cerea, M., Foppoli, A., Maroni, A., Gazzaniga, A. and Zema, L., 2022. Administration strategies and smart devices for drug release in specific sites of the upper GI tract. Journal of Controlled Release, 348, pp. 537-552.
[5] Kuang, F.L., De Melo, M.S., Makiya, M., Kumar, S., Brown, T., Wetzler, L., Ware, J.M., Khoury, P., Collins, M.H., Quezado, M. and Pittaluga, S., 2022. Benralizumab completely depletes gastrointestinal tissue eosinophils and improves symptoms in eosinophilic gastrointestinal disease. The Journal of Allergy and Clinical Immunology: In Practice, 10(6), pp. 1598-1605.
[6] Yogapriya, J., Chandran, V., Sumithra, M.G., Anitha, P., Jenopaul, P. and Suresh Gnana Dhas, C., 2021. Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model. Computational and Mathematical Methods in Medicine, 2021.
[7] Naz, J., Sharif, M., Yasmin, M., Raza, M. and Khan, M.A., 2021. Detection and classification of gastrointestinal diseases using machine learning. Current Medical Imaging, 17(4), pp. 479-490.
[8] Raut, V., Gunjan, R., Shete, V.V. and Eknath, U.D., 2022. Gastrointestinal tract disease segmentation and classification in wireless capsule endoscopy using intelligent deep learning model. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, pp. 1-17.
[9] Padmavathi, P. and Harikiran, J., 2022. Wireless capsule endoscopy infected images detection and classification using MobileNetV2-BiLSTM model. International Journal of Image and Graphics, 2350041.
[10] Escobar, J., Sanchez, K., Hinojosa, C., Arguello, H. and Castillo, S., 2021. Accurate deep learning-based gastrointestinal disease classification via transfer learning strategy. In: 2021 XXIII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), pp. 1-5. IEEE.
[11] Laleman, W., Vanderschueren, E., Mehdi, Z.S., Wiest, R., Cardenas, A., Trebicka, J., 2023. Endoscopic procedures in hepatology: Current trends and new developments. Journal of Hepatology.
[12] Andrew, J., DivyaVarshini, M., Barjo, P. and Tigga, I., 2020. Spine magnetic resonance image segmentation using deep learning techniques. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 945-950. IEEE.
[13] Pore, A., Li, Z., Dall'Alba, D., Hernansanz, A., De Momi, E., Menciassi, A., Vander Poorten, E., 2023. Autonomous navigation for robot-assisted intraluminal and endovascular procedures: A systematic review. IEEE Transactions on Robotics.
[14] Zhang, Y.H., Guo, L.J., Yuan, X.L. and Hu, B., 2020. Artificial intelligence-assisted esophageal cancer management: Now and future. World Journal of Gastroenterology, 26(35), pp. 5256.
[15] Dixon, R.V., Skaria, E., Lau, W.M., Manning, P., Birch-Machin, M.A., Moghimi, S.M. and Ng, K.W., 2021. Microneedle-based devices for point-of-care infectious disease diagnostics. Acta Pharmaceutica Sinica B, 11(8), pp. 2344-2361.
[16] Yu, X., Dong, M., Yang, D., Wang, L., Wang, H. and Ma, L., 2023. Deep learning for differentiating benign from malignant tumors on breast-specific gamma image. Technology and Health Care, 31(S1), pp. 61-67.
[17] Pushpa, B., Baskaran, B., Vivekanandan, S. and Gokul, P., 2023. Liver fat analysis using optimized support vector machine with support vector regression. Technology and Health Care, 31(3), pp. 867-886.
[18] Dheir, I.M. and Abu-Naser, S.S., 2022. Classification of anomalies in gastrointestinal tract using deep learning.

[19] Obayya, M., Al-Wesabi, F.N., Maashi, M., Mohamed, A., Hamza, M.A., Drar, S., Yaseen, I., Alsaid, M.I., 2023. Modified salp swarm algorithm with deep learning-based gastrointestinal tract disease classification on endoscopic images. IEEE Access, 11, pp. 25959-25967.

[20] Aliyi, S., Dese, K. and Raj, H., 2023. Detection of gastrointestinal tract disorders using deep learning methods from colonoscopy images and videos. Scientific African, 20, e01628.

[21] Gunasekaran, H., Ramalakshmi, K., Swaminathan, D.K. and Mazzara, M., 2023. GIT-Net: An ensemble deep learning-based GI tract classification of endoscopic images. Bioengineering, 10(7), p. 809.

[22] Ramamurthy, K., George, T.T., Shah, Y. and Sasidhar, P., 2022. A novel multi-feature fusion method for classification of gastrointestinal diseases using endoscopy images. Diagnostics, 12(10), p. 2316.

[23] Khan, M.A., Muhammad, K., Wang, S.H., Alsubai, S., Binbusayyis, A., Alqahtani, A., Majumdar, A. and Thinnukool, O., 2022. Gastrointestinal diseases recognition: A framework of deep neural network and improved moth-crow optimization with DCCA fusion. Human-centric Computing and Information Sciences, 12, p. 25.

[24] Hmoud Al-Adhaileh, M., Mohammed Senan, E., Alsaade, W., Aldhyani, T.H.H., Alsharif, N., Abdullah Alqarni, A., Uddin, M.I., Alzahrani, M.Y., Alzain, E.D. and Jadhav, M.E., 2021. Deep learning algorithms for detection and classification of gastrointestinal diseases. Complexity, 2021, pp. 1-12.

[25] Sharif, M., Attique Khan, M., Rashid, M., Yasmin, M., Afza, F. and Tanik, U.J., 2021. Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. Journal of Experimental & Theoretical Artificial Intelligence, 33(4), pp. 577-599.

[26] Zhang, H., Huang, T., Liu, S., Yin, H., Li, J., Yang, H., Xia, Y., 2020. A learning style classification approach based on deep belief network for large-scale online education. Journal of Cloud Computing, 9, pp. 1-17.

[27] Dakshina, D.S., Della Reasa, V. and Bindhu, A., 2023. Alzheimer disease detection via deep learning-based shuffle network. International Journal of Current Bio-Medical Engineering, 1(1), pp. 9-15.

[28] Reeba Rex, S., Pravin Rose, T. and Amudaria, S., 2024. Real-time remote monitoring via horse head optimization deep learning network. International Journal of Data Science and Artificial Intelligence, 2(2), pp. 42-47.

[29] Anlin Sahaya Infant, Tinu M., Ahilan, A. and Muthukumaran, N., 2024. Detection of brain tumour via reversing hexagonal feature pattern for classifying double-modal brain images. IETE Journal of Research, Published online, February.

[30] Nithya, A., Appathurai, A., Venkatadri, N., Ramji, D.R. and Palagan, C.A., 2020. Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. Measurement, 149, p. 106952.

[31] Awan, R., Koohbanani, N.A., Shaban, M., Lisowska, A. and Rajpoot, N., 2018. Context-aware learning using transferable features for classification of breast cancer histology images. In: Proceedings of the International Conference on Image Analysis and Recognition, pp. 788-795, Povoa de Varzim, Portugal, June.

[32] Baccouche, A., Garcia-Zapirain, B., Castillo Olea, C. and Elmaghraby, A., 2020. Ensemble deep learning models for heart disease classification: A case study from Mexico. Information, 11(4), p. 207.

[33] Senan, E.M. and Jadhav, M.E., 2021. Analysis of dermoscopy images by using ABCD rule for early detection of skin cancer. Global Transitions Proceedings, 2(1).

[34] Ali, M.N.Y., Sarowar, M.G., Rahman, M.L., Chaki, J., Dey, N. and Tavares, J.M.R.S., 2019. Adam deep learning with SOM for human sentiment classification. International Journal of Ambient Computing and Intelligence, 10(3), pp. 92-116.

[35] Bour, A., Castillo-Olea, C., Garcia-Zapirain, B. and Zahia, S., 2019. Automatic colon polyp classification using convolutional neural network: A case study at Basque country. In: Proceedings of the 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 1-5, Ajman, UAE, December.

[36] Zhang, R., Zheng, Y., Mak, T.W.C. et al., 2016. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. IEEE Journal of Biomedical and Health Informatics, 21(1), pp. 41-47.

[37] Zhu, R., Zhang, R. and Xue, D., 2015. Lesion detection of endoscopy images based on convolutional neural network features. In: Proceedings of the 2015 8th International Congress on Image and Signal Processing (CISP).

[38] Fonolla, R., Van Der Sommen, F., Schreuder, R.M., Schoon, E.J. and De With, P.H., 2019. Multi-modal classification of polyp malignancy using CNN features with balanced class augmentation. In: Proceedings of

the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 74-78, Venice, Italy, April.

[39] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Neural Information Processing Systems (NeurIPS), 1097-1105.