ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

CONTRASTIVE LEARNING: A NOVEL APPROACH FOR GASTROINTESTINAL DISEASE CLASSIFICATION

Muhammad Amir Shehzad[1]
(amirshehzad3974@gmail.com),
Javaria Amin[1],
Muhammad Imran Khan[1] ,
Sajid Iqbal[2]
Adeel Ahmed[3]
[1] Department of Computer Science, University of Wah, Wah Cantt, Pakistan
[2] Department of Computer Science and IT, The University of Lahore, Pakistan
[3] Faculty of Computer Science and Information Technology, The Superior University Lahore, 54000, Pakistan.

Abstract

A wide range of serious health issues, such as pain, discomfort, poor digestion, and in certain situations, chronic difficulties, are brought on by structural or functional abnormalities of the digestive system, which are referred to as gastrointestinal (GI) diseases. The digestive system problems cause severe disruptions to people's quality of life. Early detection improves patient care and lowers medical costs, and it is strongly related to successful treatment. Through the examination of endoscopic images, artificial intelligence (AI), and intense learning, have shown outstanding potential in detecting gastrointestinal disorders. However, problems like incorrectly categorising conditions that are identical in appearance and limited model generalizability still present serious difficulties. Therefore, Further research is necessary to develop robust solutions that address these challenges. This work proposed a GI classification model using a self-supervised contrastive learning technique using the ResNet-18 as a CNN backbone. The proposed model performs exceptionally well in classifying GI disease on the publicly available Kvasir dataset. The proposed model efficiently classifies eight categories of GI disease with an excellent overall 100% accuracy. The placement of ROC curves for all classes close to the top-left corner of the graph and AUC ≈ 1 for all classes reflects the strong classification ability of the proposed model. The suggested model is a state-of-the-art method for precise and effective diagnosis, outperforming current AI models in the classification of gastrointestinal diseases.

1. Introduction

Gastrointestinal (GI) diseases are some of those global diseases that need immediate attention as they are causing serious health issues to a wide majority. These diseases include colorectal cancer, Crohn's disease, other types of gastritis, and ulcerative colitis [1]. Their destructive impacts unfold as a noticeable decrease in patients' Standard of living and an enormous increase in healthcare costs, as continuous monitoring, treatment, and surgery, in most cases, require a huge sum of money [2]. Additionally, GI diseases, including colorectal cancer, have increased exponentially along with the increase in other forms of cancers. This increased likelihood of cancerous patients has increased the frequency of GI diseases, which demand early and rapid detection and classification of inhalation injury to improve patients' handling [3]. Endoscopic examinations were now adequate for examining GI diseases but require specialized interpretation with potential diagnostic variability [4]. These examination processes have shown effective results in consistent and correct classification of many GI diseases at various levels of regularity and objectivity that were impossible to achieve through traditional methods [5].

However, Gastrointestinal (GI) diseases require an early and accurate diagnosis for in-time and effective control. An early identification of these diseases helps prevent complications and improve treatment results. Thus, minimizing the invasiveness of treatment. On the other hand, effective though traditional diagnostic procedures such as endoscopy are invasive and time-consuming: highly skilled clinical dependency, diagnostic variations, and plenty of time

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

to treat. This variability is problematic, especially when working in a high-workload environment or the availability of skilled staff in a limited capacity [4]. However, the implementation of artificial intelligence can bring some ease in the diagnosis of stomach and intestinal diseases. Different deep learning AI models like CNNs mainly address the challenges of preparing training data and recognizing fine details in outer images. These techniques improve the accuracy, consistency, and objectivity of the medical data and require no or minimal human involvement. In fact, AI models have been employed to Find different classes of GI disease from endoscopic images, resulting in a noninvasive, faster, and more accurate diagnostic [6]. Some of the many advantages of AI models are reduced diagnostic errors up to 60 times, early detection, early application of treatment, reduced spending, and improved patient results [7].

Though AI models are greatly helpful in healthcare innovation, Artificial intelligence techniques, including ML and DL, have mainly evolved to become increasingly important for diagnosing Gastrointestinal (GI) abnormalities. Learning algorithms such as CNN-based approaches have been used in endoscopic, histopathological, and capsule endoscopy images to facilitate high accuracy in identifying GI diseases [8]. With the help of these models, the examination and resultantly diagnostic of large datasets become more efficient and reliable [9]. Deep learning architectures like DNNs are used to automate the autopsy of medical images and eliminate human errors and diagnostic variability. A Herbst theory of the preservation of constitutional value in decisions concerning the denial of parole statehood can be developed. Additionally, ML models are greatly helpful for clinicians to prioritize tasks based on severity, quicker decision-making, and effective utilization of time and resources [10]. Moreover, with the help of AI models, supervised learning has also further improved in specific supervised and transfer learning techniques. Specifically, it transfers learning that enables us to reuse pre-tested/trained models on new smaller datasets, particularly when we have limited annotated medical images. This type of advancement in the field of healthcare results in early, accurate and effective detection of GI diseases, even with limited data. ML-based diagnostic models can be effectively integrated into all clinical practices for faster, more accurate, and cost-effective results. They are significantly more helpful in the underserved regions with no or low specialist access [11]. ML improve healthcare efficiency and patient outcomes by providing timely and accurate diagnosis. In future, due to the continuous efforts to include multimodal data, reduce bias in predictions and increase model interpretability by using Explainable AI (XAI), ML will be able to provide more accurate and detailed predictions and serve as a critical factor in Gastrointestinal (GI) disease diagnosis using endoscopic images [12].

Diverse learning algorithms, notably convolutional networks, provide outstanding results in Gastrointestinal (GI) disease diagnosis using endoscopic images and help physicians provide timely and accurate medical facilities. But ML algorithms are facing several challenges also, and these challenges still limit the abilities of ML learning algorithm's practical applications in gastrointestinal (GI) disease diagnosis [13].

It is a common problem with some Convolutional Neural Network (CNN) models that they can provide outstanding results on specific data. However, if we use this same CNN model with a new dataset, the prediction accuracy and efficiency decrease due to less generalization. If the small dataset is used for training purposes, it worsens the overfitting problem. To overcome these issues and provide better performance on patient data belonging to different demographics and conditions, The Convolutional Neural Network (CNN) models must be more generalized and should avoid overfitting issues, which require larger and more varied data [14]. It is a complex task to implement Machine Learning (ML) models in clinical

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

environments that require real-time and fast-paced diagnosis due to problems like computational delays and time limitations [15]. Clear and actionable decisions from AI models are the basic requirement of clinicians. A major drawback to using Convolutional Neural Networks (CNNs) is less interpretability, reducing the practical implementation of these learning models [12]. Many publicly available datasets are either too small or poorly annotated. Non-availability of large, versatile, well-annotated and accurate datasets creates a great hurdle for machine learning (ML) algorithms to develop robust machine learning-based diagnostic systems [16]. Transfer learning provides a strong way to deal with situations like limited data availability and computational resources by using previously trained models already trained on huge datasets. However, transfer learning has some limitations, such as it cannot deal with huge variations in datasets. Therefore, very diverse datasets decrease the performance of an AI model [9].

Addressing these challenges is a critical issue to deal with while implementing any machine learning (ML) model for gastrointestinal disease diagnosis. Our work provides an efficient solution for these challenges by proposing a Neural Network model with Convolutional layers (CNN) along with contrastive learning techniques for classifying GI diseases. Combining contrastive learning techniques with a Neural network model (ResNet18) already trained on a vast dataset can enhance the ability to detect, classify, and monitor gastrointestinal diseases. The model will be able to work with limited datasets, capture subtle differences, provide better generalization and increase the model accuracy and efficiency in the categorization of different GI disease conditions. This proposed model can efficiently and accurately diagnose and classify diseases based on subtle visual patterns, detecting early signs of diseases like Crohn's disease, colon cancer, and gastritis and segmenting affected areas in the GI tract for further analysis. Labelled data like endoscopic images for gastrointestinal disease detection and classification is very difficult to obtain and sometimes it is very expensive to collect because it needs expert doctors for labelling. Contrastive learning overcomes this issue because it does not always need labelled data [17, 18].

To main contribution of the proposed model is as follows:

1. To classify gastrointestinal disease, develop a state-of-the-art model that achieves high performance by leveraging a contrastive learning technique with a ResNet-18 as CNN backbone.
2. Reducing misclassification rates through efficient feature representation for classifying difficult and overlapping classes.
3. The study emphasises an effective training approach, as demonstrated by reduced training and learning loss across several epochs.

The article's structure is as follows: Section II discusses related work, Section III elaborates on suggested method steps, and Section IV describes the results and discussion. Lastly, Section V has the conclusion.

2. Related work

Different AI models like CNNs can be used to detect GI diseases accurately and efficiently using endoscopic and histopathological medical images [19]. An efficient technique for classifying gastrointestinal diseases is recommended by applying a GA to improve the sharpness and brightness of WCE images using enhanced contrast adjustment with an optimized brightness control method. On these enhanced WCE images, a transfer-based learning method is also used to Refine a pre-trained model. Optimization of the features, Fusion, and training of numerous models can enhance performance in the future [11]. In contrast with a flat Neural network with a convolutional layers (CNN) model, the hierarchical deep convolutional neural network has better classification performance. The higher layers of

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

convolutional neural networks extract more abstract features than the lower layers, which only capture the low-level features. Using this strategy in the hierarchical deep convolutional neural network can reduce the computational cost, and in the future, optimising this model can provide outstanding results for the multi-category classification of GI disorders [20]. Due to their inherent mechanism, full utilisation of spatial information is impossible in CNN models. To improve the classification accuracy of GI disorders, a deep convolutional neural network with encoder-decoder layers and a spatial attention method is suggested. Data augmentation technique is used to deal with data imbalance problems [9]. CNN models that are already trained on large datasets can be used with small datasets by implementing augmented techniques to increase the performance of GI disease detection [11]. Combining convolutional network (CNN) models and supervised contrastive learning can optimize the model's performance in identifying gastrointestinal conditions. Classes having limited data samples as compared to other classes can create an imbalanced dataset problem. The cost-sensitive learning method is useful to tackle the imbalance problem. Explainable artificial intelligence (XAI) techniques provide valuable insights into interpretability for an enhanced understanding of the decision-making process of the CNN model. Further developments in Incremental learning techniques in CNN models can increase the accuracy in detecting critical conditions and improve diagnostic performance [18].

FLATer is a transfer-based, highly accurate, fast, and lightweight model that leverages the capabilities of both Vision transformers (ViT) and convolutional networks (CNNs). To focus on local features and global attention, FLATer has a spatial attention block, along with a residual layer and vision transformer module. FLATER provides outstanding classification results. It has 99.7% accuracy for multi-class classification with three categories (ulcerative colitis, polyps, and esophagitis) and 96.4% in binary classification (standard vs. pathological). Even when the model is trained from scratch, the dual approach can propagate robustness and high precision at negligible computational cost; due to this, it is broadly applicable, with less reliance on large datasets [21] . Transformer-based models are very efficient in classifying medical images for gastrointestinal disease detection with great accuracy and with an outstanding speed of 16.4k images per second. This accuracy, efficiency, and speed made this model a good choice for real-time applications in clinical practice. Combining convolutional neural networks (CNNs) with transformer models can achieve improved computation efficiency and better model precision [8, 22, 23].

To classify GI diseases using a combination of pre-trained CNN model and Vision Transformer (ViT) model on wireless capsule endoscopy (WCE) images can show a performance of 95.63%. The ability of Vision Transformer (ViT) model to focus on fine features in endoscopic images made this model better than the standard CNN models as its predictions are very precise and also decreased the computational costs of medical image classification [8]. For the early identification of gastric cancer, a deep learning-based CADx system that can distinguish between normal and abnormal endoscopic pictures is presented. This approach improves diagnostic accuracy and tackles data-gathering issues by combining DCGAN with CIFAR-based data variation [24]. Using AI models as random forests to identify gastrointestinal (GI) disease some predictors use gastrointestinal bleeding as vital signs and lab values as key predictors. SHAP values as explainable AI (XAI) is also used to support and explain the predictions and transparent diagnostic aids for gastrointestinal disease [25]. Deep learning models, VGG-19 and ResNet-50, with the Kvasir dataset, were used to classify gastrointestinal (GI) diseases to identify the gastrointestinal polyps early to avoid unnecessary tumour removal. VGG-19 and ResNet-50 classify the endoscopic images into seven gastrointestinal disorder categories with great accuracy by focusing on critical areas

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

only. If the dataset is expanded with diverse expert annotations performance and accuracy of the model also can be increased [26].

Classify gastrointestinal (GI) diseases by optimizing ResNet-50 and ResNet-152 models for pattern extraction and recognition of dedicated regions of interest; Mask Recurrent CNN, called (R-CNN) results in a high accuracy rate of 96.43%. The extracted features are used to optimize the feature selection by an improved Ant-Colony optimization (ACO) algorithm used with a serial approach [27].

Limitations like small datasets, heterogenous equipment, population and clinical settings compromise the quality of decision-making of AI models for identifying and categorizing gastrointestinal (GI) diseases. On the other hand, for accurate and transparent identification and categorization of GI diseases, An extensive collection of endoscopic images and the use of explainable AI (XAI) with a Convolutional Neural network hold significant importance [4]. For AI models, data imbalance is a huge problem for identifying and categorising GI diseases. The availability of very few samples of rare conditions in gastrointestinal disease as compared to other diseases causes biases in AI models which can reduce the efficiency and accuracy rate of the AI model [28]. Classification of the disease as early staged or late-staged is a very difficult task for AI models to predict accurately. To achieve an acceptable accuracy, a huge dataset of endoscopic images and real-time endoscopy videos of gastrointestinal disease needed to be processed by the AI model for feature extraction and model training which required expensive and advanced hardware support [29].

The strength and adaptability of machine learning models for endoscopic gastrointestinal diagnosis can be improved by using more extensive and diverse datasets comprising heterogeneous patient demographics, imaging, and disease types [30]. By using the heatmaps method, explicitly pointing out the targeted region of the images participating most in the process by which the AI model makes decisions, introduces the concept of explainability [31]. Gastrointestinal (GI) disease endoscopic image datasets are still very limited for accurate detection and classification for rare types of Gastrointestinal (GI) disease. With the use of Transfer learning and GANs, issues like data scarcity for rare gastrointestinal conditions can be minimized effectively and increase the model performance for classifying rare types of Gastrointestinal (GI) disease [32]. This study examines potential remedies for data-driven decision-making for precise GI disease diagnosis while highlighting the difficulties caused by class imbalance. Techniques like cost-sensitive learning can address data imbalance, oversampling minority classes, and synthetic augmentation so that the models can better classify rare diseases [28]. To overcome the data imbalance, oversampling minority classes, and synthetic augmentation problems, Cost-sensitive learning techniques and multi-tiered models to classify the type and the stage of GI diseases show high accuracy rates [12, 28].

3. Research Framework

The proposed methodology processes endoscopic images to classify gastrointestinal diseases in two stages. In the first stage, a self-supervised learning technique called Contrastive learning is used to learn image representations [33]. At this stage, the pre-trained CNN backbone (ResNet-18) is trained/modified to map the endoscopic images of GI disease in multidimensional embedded space as positive pairs (Similar images) labelled as “1” and negative pairs (dissimilar images) labelled as “0”, in such a way that positive pairs are positioned close together, while negative pairs are placed at the maximum distance in embedding space [34]. At this stage, a contrastive loss function minimises the distance between similar pairs and maximises the distance at a minimum margin of m between dissimilar pairs in embedding space [35]. During the second step, optimize the classifications

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

using a supervised learning technique by adding a fully connected layer to associate the embeddings into predefined classes of Gastrointestinal disease as shown in **Figure 1**. ResNet-18 extracts features of the input images that are mapped into a meaningful multidimensional embedding space, while the learning goal ensures that the space is structured in a way that supports accurate classification of gastrointestinal diseases.

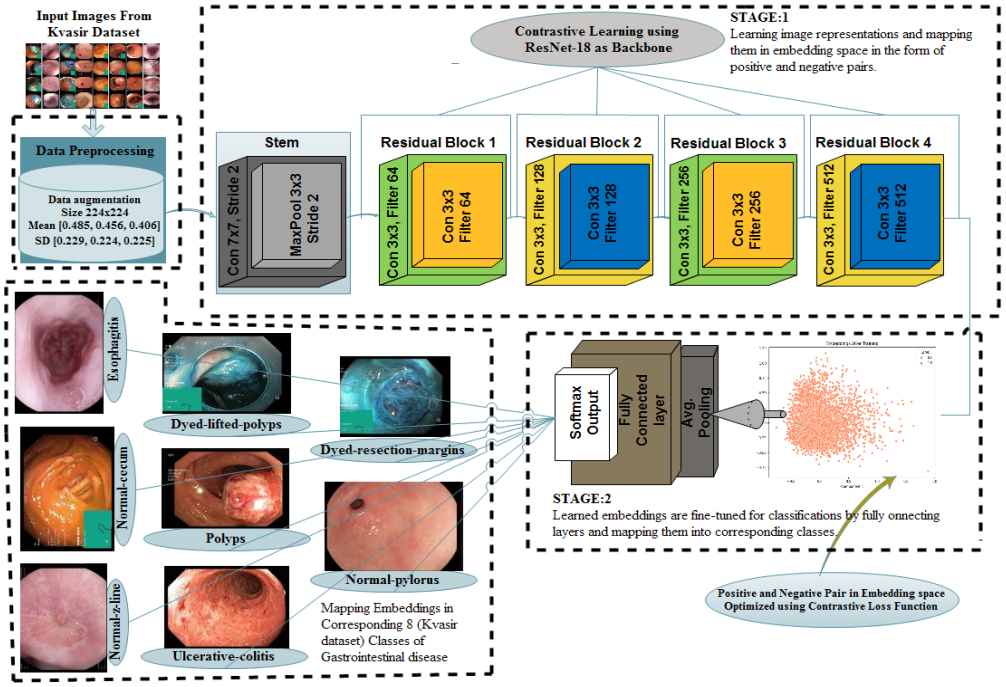

Figure 1 Proposed Model Architecture Using Contrastive Learning Technique with RestNet-18 Model as CNN Backbone

3.2. Contrastive Learning for Feature Representation in Positive and Negative Pairs

The contrastive learning technique plays a significant role in learning image representations by structuring the multidimensional embedding space of endoscopic images for GI disease classification. Utilizing a tagged dataset, positive pairs (Images belonging to the same category) designated as label "1" and negative pairs (images from distinct classes) are tagged with label "0", and are generated during preprocessing; as shown in Figure 3, some samples of positive and negative pairs of images from different classes. The ResNet-18 (CNN backbone) maps these pairs into multidimensional embedding space. In **Figure 2,** some positive and negative pairs generated by ResNet-18 using contrastive learning are presented. The contrastive loss function minimizes the distance between similar pairs and maximizes the distance at a minimum margin of *m* between dissimilar pairs in embedding space, which is the primary goal of the contrastive learning technique. **Figure 3** provides an overview of the Contrastive learning architecture based on ResNet-18.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

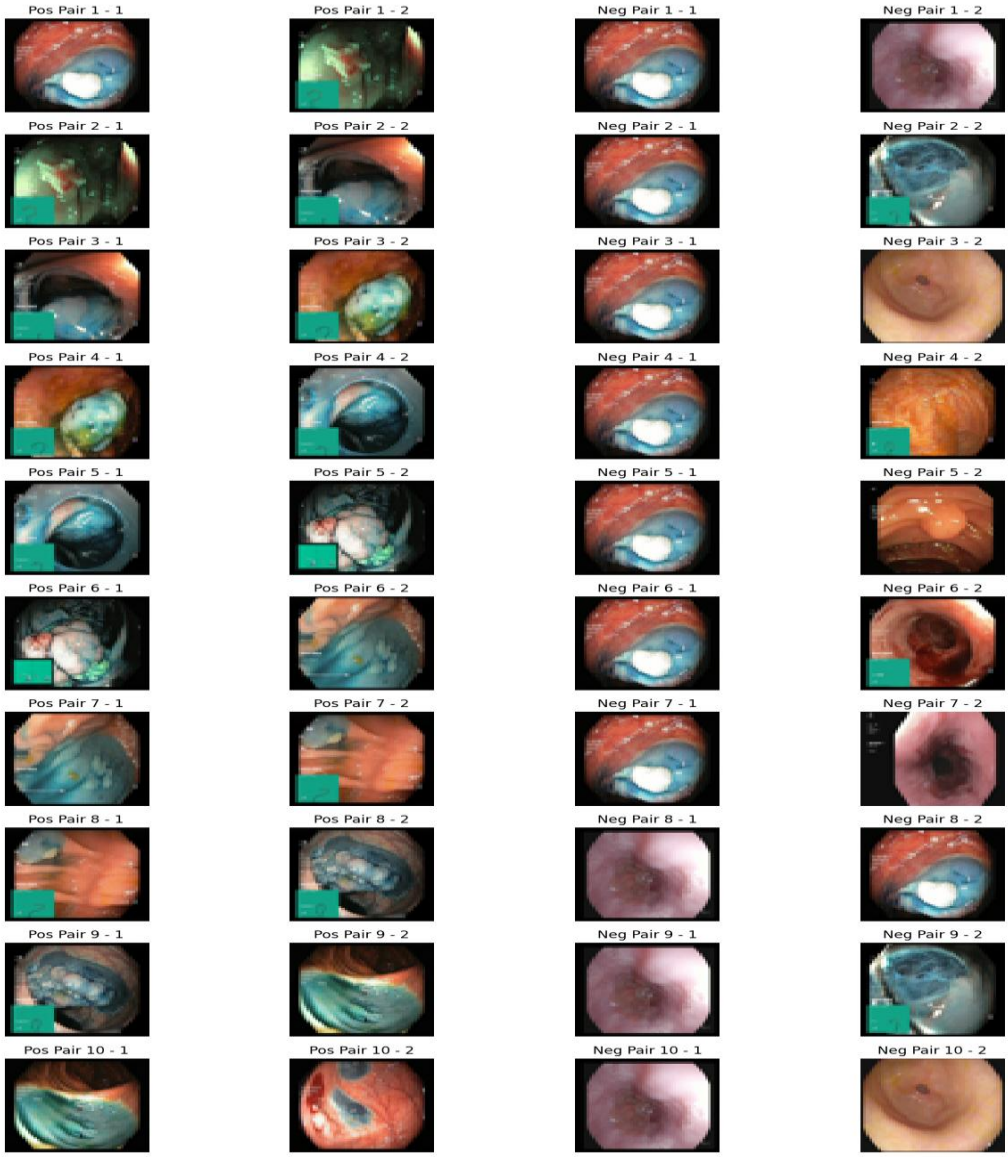

Figure 2 A Sample of 10 Positive Pairs and 10 Negative Pairs Generated Using Contrastive Learning Technique

The contrastive learning technique enables the model to learn meaningful representations from the endoscopic images of gastrointestinal diseases. Using the contrastive learning technique at this stage improves the quality of image representations, which helps the further classification of images at the next stage. By learning robust features at the pretraining stage, the contrastive learning technique provides a well-initialized model for the next fine-tuning stage, ultimately enhancing classification accuracy.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

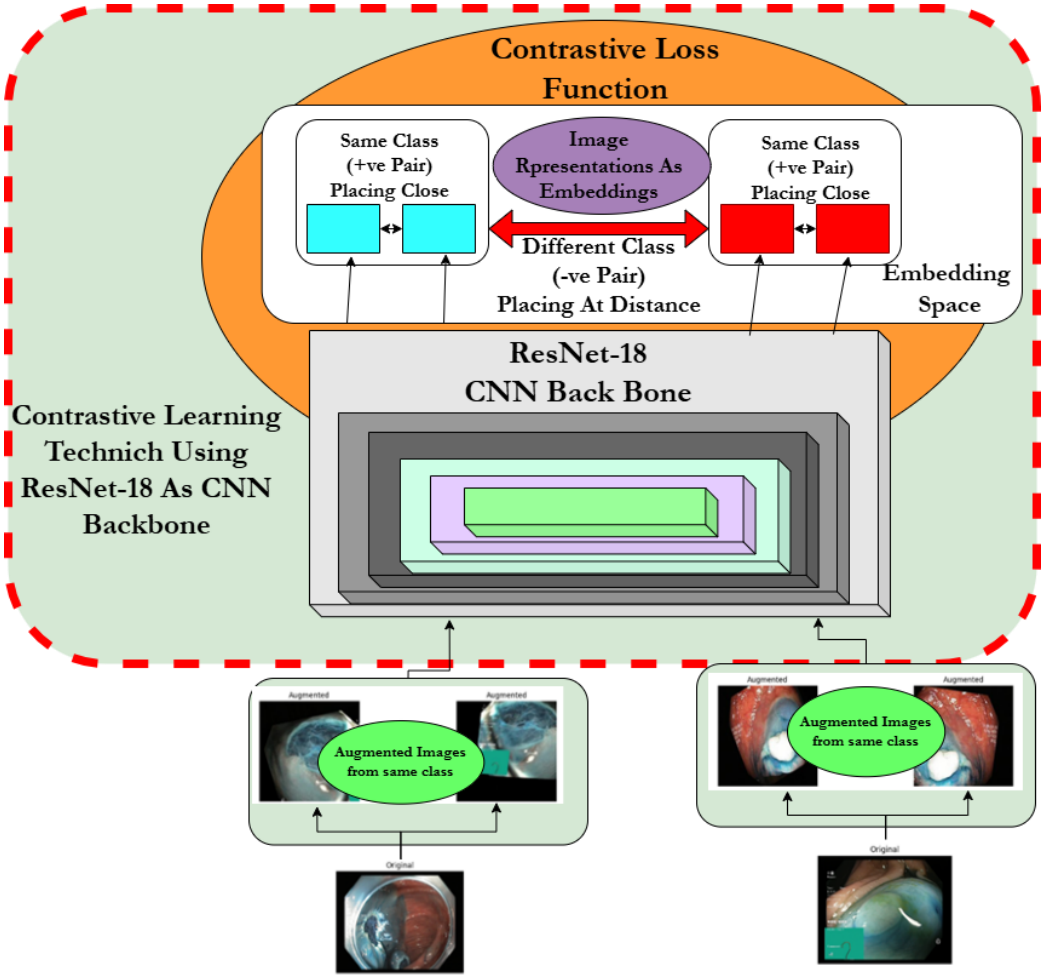

Figure 3 Implementation of Contrastive Learning Technique based on RestNet-18 as Backbone CNN

3.2. Trained Resnet-18 As CNN Backbone To Embed Endoscopic Image Representations in Multidimensional Embedding Space

In the proposed model, the role of ResNet-18, serving as the CNN backbone, is to capture useful features of endoscopic images and generate embeddings. ResNet-18 is fine-tuned to process endoscopic images to effectively capture hierarchical features that play a pivotal role in characterizing gastrointestinal diseases. ResNet-18 is a pre-trained CNN architecture that contains 18 layers that are trained on a huge database named ImageNet. The database consists of millions of images, which makes it very efficient for image classifications. ResNet-18 takes images of size 224x224 and can efficiently classify images into up to 1000 categories. The resNet-18 is a very efficient CNN model in extracting and classifying the features of the endoscopic images due to its residual learning framework. While the deeper layers of the ResNet-18 model extract more intricate, high-level characteristics, the early levels catch basic patterns like textures and edges. Residual neural networks (ResNet), apply identity mapping, meaning that a certain layer's input can be directly passed to some other layers [34]. The input X is directly forwarded to some other layer. Skipping one or more connections is actually called identity mapping. The output of the residual block can be written by treating f(x) as the output of a subsequent layer in the network and x as the input to one layer, as shown in **equation 1**:

$Z = X + f(x)$ ---------- (1)

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

To address the size mismatch of input and output in terms of dimensionality due to the convolutional operations. The updated **equation 1** can be written as **equation 2**:

$Z = X * Ws + f(x, Wi)$ ---------- (2)

Here, the Ws represents the implementation of convolutional configuration for matching the size, and Wi is the CNN layer's parameter. Figure 3 further explains the concepts

As shown in **Figure 4,** The first few layers of the ResNet-18 are called Stem; this comprises a first convolution layer with a stride of two and a 7x7 kernel size, followed by a 3x3 max-pooling layer also of stride 2. Residual Blocks 1, 2, 3 and 4 contain 2 convolutional layers of size 3x3, with 64, 128, 256 and 512 filters, respectively and a skip connection. Residual Block 3 contains 2 convolutional layers of size 3x3 with 256 filters. A stride of size 2 for downsampling is used for each layer in all residual blocks, reducing the feature map's spatial dimension [34].

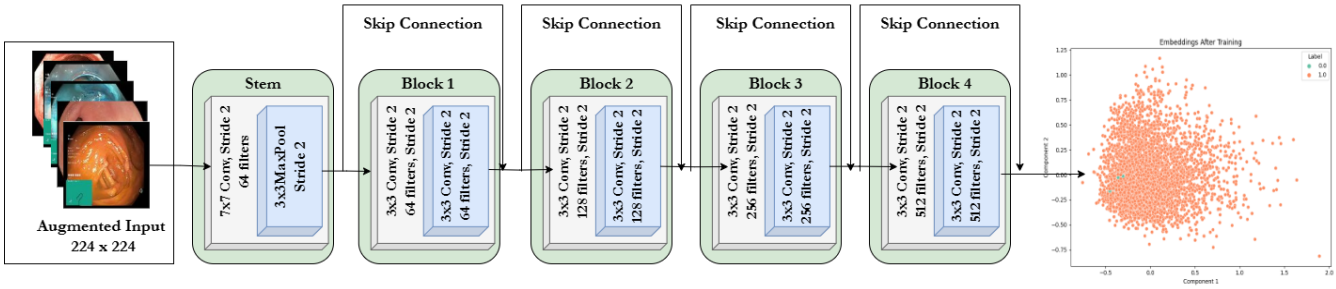

Figure 4 The Layers Configuration of The Resnet-18 Model to Learn and Map Image Representation on Embedding Space

ResNet-18 maps input images into a high-dimensional embedded vector so that similar images (positive pairs) are brought closer together and unlike Images (negative pairs) are separated by a greater distance using a contrastive loss function. As shown in **Figure 5**, it encodes the critical features of an image in a compact form.

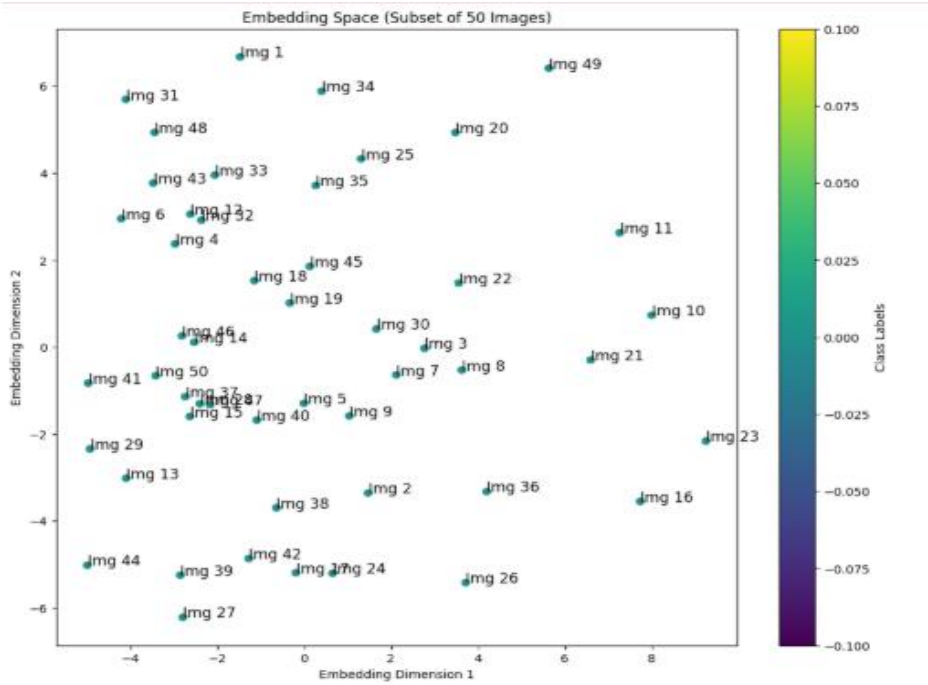

Figure 5 A sample of embedding space is shown, and 50 images are placed in the form of positive and negative pairs

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

3.3. Learning Goals

The learning goal of the ResNet-18 model is to map image representations into a structured multidimensional embedding space, which pushes different images farther away and brings identical ones closer together. This is attained using the contrastive loss function during training. This structured embedding space enables the model to distinguish between different classes of GI disease effectively and enhances the final classification task.

3.4. Contrastive Loss Function

The contrastive loss function plays a significant role in structuring the embeddings generated by the ResNet-18 backbone in the training process of our model for the classification of Gastrointestinal disease. Contrastive loss function ensures the placement of endoscopic images in the embedding space. By maintaining clear boundaries between different disease classes, optimizing the contrastive loss facilitates the improvement in the class separation. Grouping similar diseases together and dissimilar ones separated in embedding spaces improves the classification task, as shown in Figures **6 and 7**. The contrastive loss function optimize the embedding space in such a way that:

- Positive pairs (images depicting the same disease) placed closer to each other
- Negative pairs (images from different disease classes) placed far enough by maintaining the minimum distance defined by the margin m.

For a single pair of images, the loss function is defined as In **Equation 3**:

$$L = y \cdot [\{D\big(f(x_1), f(x_2)\big)^2\} + (1-y)] \cdot [\max\{0, m - D\big(f(x_1), f(x_2)\big)\}^2] \quad \text{.....(3)}$$

- For images x_1 and x_2, ResNet-18 backbone generates $f(x_1)$ and $f(x_2)$ embeddings
- The Euclidean distance between the two embeddings is represented by $D\big(f(x_1), f(x_2)\big)^2$
- To indicate whether a pair is positive or negative y is used as a binary label indicating $= 1\ for\ positive\ and\ y = 0\ for\ negative\ pairs$.
- To ensure adequate separation between positive and negative pairs, a margin m is used

To meet the learning goal contrastive loss function optimises the embeddings:

- Distance $D\big(f(x_1), f(x_2)\big)^2$ is minimised to bring embedding closers in the feature space for $y = 1\ for\ positive\ pairs.$
- Distance $D\big(f(x_1), f(x_2)\big)^2$ is maximised to separate embedding up to the distance margin m in the feature space for $= o\ for\ negative\ pairs.$.

In the gastrointestinal disease detection context, the total contrastive loss is averaged over all pairs for *N* pairs and can be elaborated in **equation 4**:

$$L_{total} = \frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot D\big(f(x_{1i}), f(x_{2i})\big)^2 + (1-y_i) \cdot \max\Big(0, m - D\big(f(x_{1i}), f(x_{2i})\big)\Big)^2\right] \ldots (4)$$

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

Figure 6 The Embedding Space Before Optimization

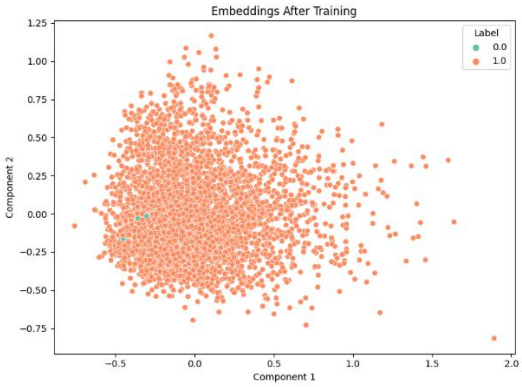

Figure 7 The Embedding Space After Optimization Using a Contrastive Loss Function

3.5. Supervised Fine-Tuning for Disease Classification Using a Fully Connected Layer with Cross-Entropy Loss

At the fine-tuning stage, feature representations learned in the previous stage are refined to optimize classification performance. Labelled data is used to train the classifier to map the embeddings to the desired output classes, which is why it is called a supervised fine-tuning process. The goal of this phase is to accurately distinguish between N disease classes' alignment of the learned embeddings with specific classification tasks.

A fully connected layer, as shown in Figure 8, acts as a classifier because it inputs high-dimensional embeddings and produces a probability distribution over the *N* classes of gastrointestinal disease. The average pooling stage serves as a bridge between the convolutional feature extraction layers and the classification layers with dense connections, transforming multidimensional feature maps into a one-dimensional feature vector. The output of the fully connected layer corresponds to a particular class. Since the function SoftMax uses logits (raw output) from the fully connected layer and a distribution of probabilities over the N classes of GI disease, it is important for final classification at this stage. The class with the greatest overall probability is chosen as the expected class for the input.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

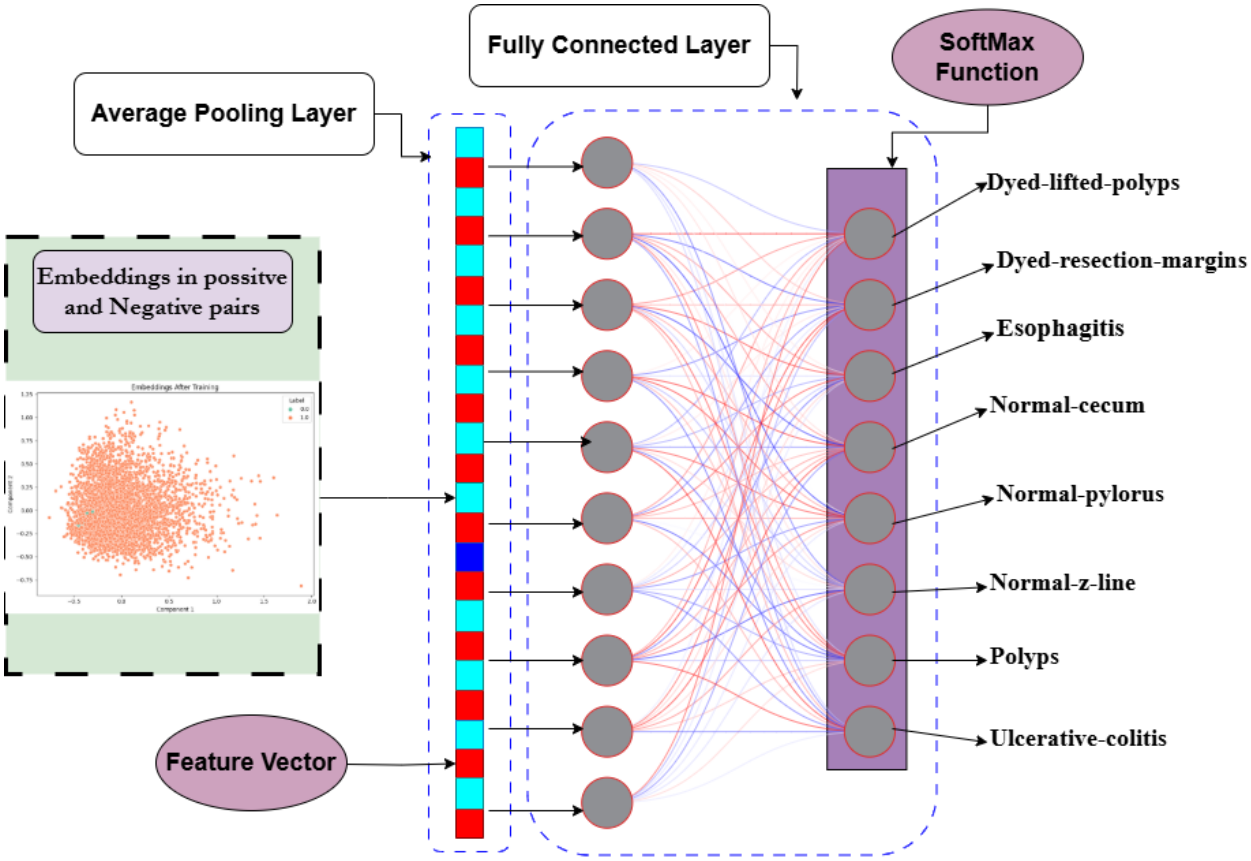

Figure 8 Fully connected layer for classification of embeddings in *N* classes of GI Disease as Per Kvasir Dataset

Cross-entropy loss function (L_{ce}) optimise the fine-tuning process by measuring the discrepancy between the expected probability distribution (p) and the actual class labels (y). High probabilities are assigned to the correct classes, while cross-entropy loss penalises incorrect classifications to optimize the model performance. Cross-entropy loss formula can be defined in **Equation 5**:

$$L_{ce} = -\sum_{i=1}^{n} y_i \log(p_i) \quad \text{..... 5}$$

Here n refers to the count of classes, y_i represents a binary indicator for the true class, and $p_i$represents the predicted probability for the corresponding class. During the fine-tuning stage, In reaction to the loss, the system learns to improve its predictions by adjusting the weights [36]. Specifically:

- The cross-entropy loss increases, If the model assigns a high probability to an incorrect class, signaling that there is a large discrepancy between the model's prediction and the true label.
- On the other hand, the decrease in cross-entropy loss indicates a correct classification when the model's predicted probability aligns closely with the true class.

4. Experiment and results

Extensive experiments were carried out utilizing the well-known Kvasir dataset to evaluate the effectiveness and reliability of the suggested methodology for categorizing GI diseases. The accuracy and effectiveness of classification algorithms in gastrointestinal research are commonly assessed using this dataset, which is renowned for its variety and high-quality endoscopic pictures. The dataset contains 4000 annotated endoscopic images, grouped into 8

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

distinct categories, each containing 500 images [37]. Various data augmentation techniques, as given in **Table 1,** were applied to enhance the diversity of the dataset and improve the robustness of the model, such as randomly rotating images ±30 degrees, horizontally and vertically shifting images by up to 20% of the total dimensions, to generate distortion along a specified axis shear intensity was applied, Random zooming of ±20%, Images were horizontally flipped to create mirrored variations. The nearest fill mode was employed to fill missing pixels with the closest neighbour values.

Table 1 The augmentation parameters applied to the Kvasir Dataset

Augmentation Technique	Value %
Image rotation	30
W-Shift	20
H-Shift	20
H-Flip	True
Shearing transformation	20
Magnification	20
Image filling approach	Nearest

Table 2 demonstrates the actual and after-augmentation statistics. An 80/20 split is used to separate the data into training and testing sets, with 80% of the data going toward training and 20% toward testing.

Table 2 Dataset Details Before and After Augmentation

Class Name	**Number of Images**		**Dataset Splitting 80-20 Ratio**	
	Before Augmentation	**After Augmentation**	**Training**	**Testing**
DL Polyps	500/Class	8000/Class	6448	1552
Polyps			6402	1598
U Colitis			6392	1608
N Cecum			6430	1570
N Pylorus			6375	1625
N Z Line			6388	1612
DR Margins			6418	1582
Esophagitis			6347	1653
Total	**4000**	**64000**	**51200**	**12800**
Training Images	**51200**			
Testing Images	**12800**			

In **Figure 9**, some sample images from each class are shown before and after augmentation. The dataset is resized as 224 x 224 to match the input requirement of ResNet-18. The pixel intensity of the images is adjusted on the values based on the mean value and standard

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

deviation values, [0.485, 0.456, 0.406], [0.229, 0.224, 0.225], respectively, to match the normalization parameters for the pre-trained ResNet-18.

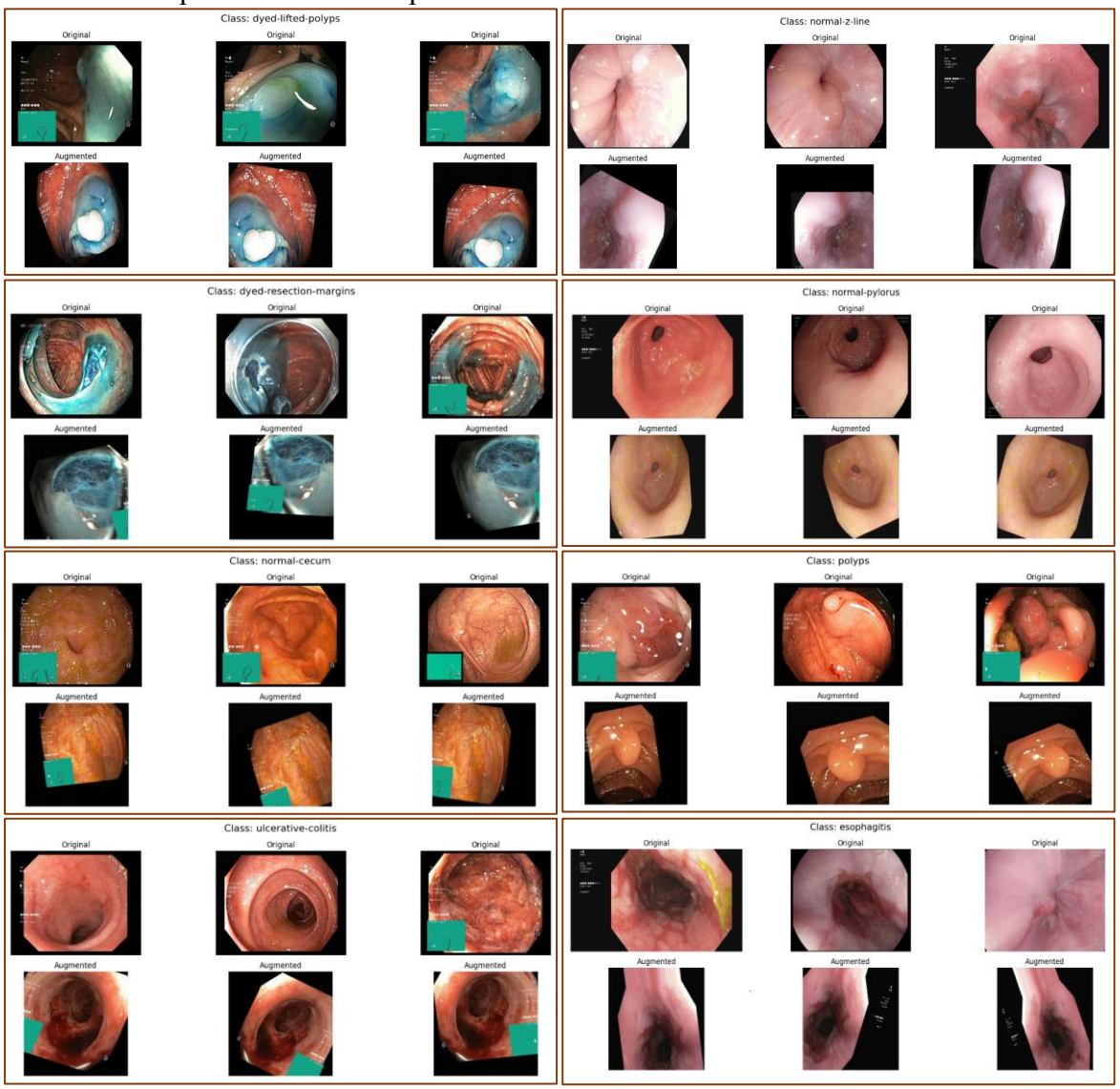

Figure 9 Sample Images from Kvasir Dataset of Each Class Before and After Augmentation

The proposed model is implemented in a Kaggle notebook environment, which provides GPU100 access for accelerated model training and testing. A machine equipped with a Core i7 (8th generation) processor and running Windows 10 operating system is used to perform experiments.

A detailed evaluation using various matrices of the model performance is carried out during the experiment, which ensures a comprehensive analysis. **Figure 10** presents the model's classification results in the form of a confusion matrix. **Figure 11** shows the training loss over the Epochs graph. The ROC curve for each class is presented in **Figure 12**. The classification report in **Table 4** includes precision, recall, f-1 score and support; this table also presents the overall accuracy, global average, and Class-weighted average of the model **Table 05** presents a comparison of the performance of the proposed model with other AI models.

Table 3 Hyperparameters for the proposed model

Hyperparameters	Selected Values

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

Batch Size	32
Training Rate (lr)	1e-4
Loss Function	Cross Entropy Loss
Optimizer	Adam
Epochs (Contrastive Learning)	10
Epochs (Fine-Tuning)	5

The hyperparameters listed in Table 3 were meticulously tuned to ensure the model's maximum learning efficiency during training and validation for classifying gastrointestinal diseases. In order to balance computation time and memory use, a batch size of 32 was selected, guaranteeing stable and effective training. A good trade-off between avoiding overshooting the ideal solution and achieving rapid convergence was achieved by setting the learning rate at 1e-4. The cross-entropy loss was chosen because it is perfect for multi-class prediction problems and effectively penalizes inaccurate classifications. The Adam optimizer was selected for its adaptive learning rate capabilities, offering robust performance in training deep neural networks. Lastly, the model was trained in two stages for feature learning and embedding generation using 10 epochs followed by an additional 5 epochs for image classification, a value that allowed sufficient time for learn the intricate patterns in the data while preventing overfitting.

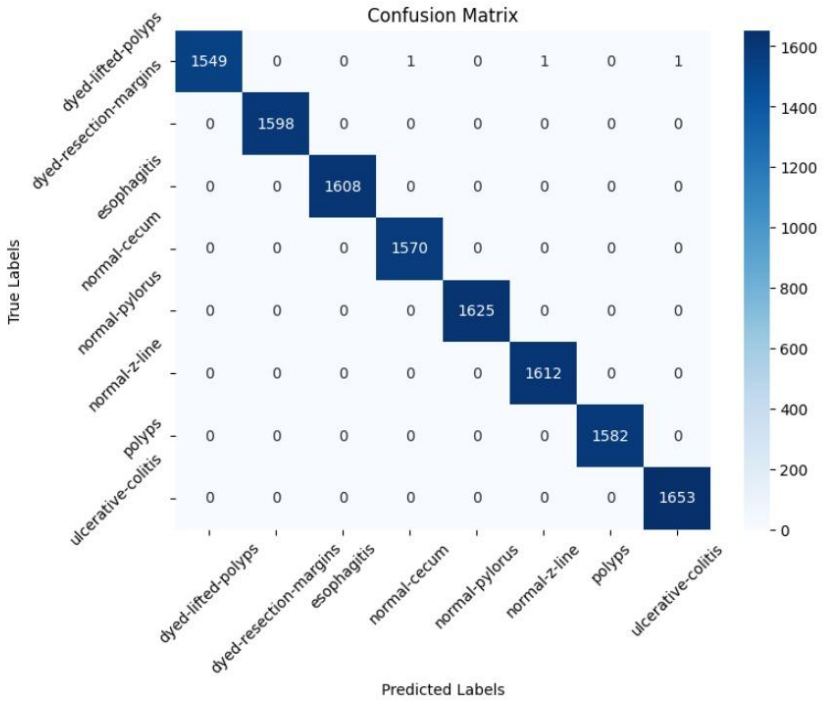

Figure 10 Confusion Matrix Representing the Classification Performance om Kvasir Dataset for GI Disease Classification

Figure 10 displays the model's classification performance across multiple classes using a confusion matrix. Actual labels are represented in the rows and predicted labels are represented in the columns. The off-diagonal elements show the examples that were incorrectly classified, whereas the diagonal elements show the number of samples that were

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

correctly identified for each class. Out of 1552 samples of Dyed-lifted-polyps, the model classifies 1549 samples correctly and misclassifies 3 classes because a few categories, such as "Normal-cecum," "normal-z-line," and "ulcerative-colitis," share characteristics. For the classes of Ulcerative Colitis, Polyps, Normal-z-line, Normal-cecum, Normal-pylorus, Normal-esophagitis, and Dyed-resection-margins, all test samples were correctly identified with 100% accuracy without any misclassification. Most of the classifications are concentrated along the diagonal in the confusion matrix, which reflects the strong performance of the proposed model that has effectively learned to distinguish between the majority of the classes.

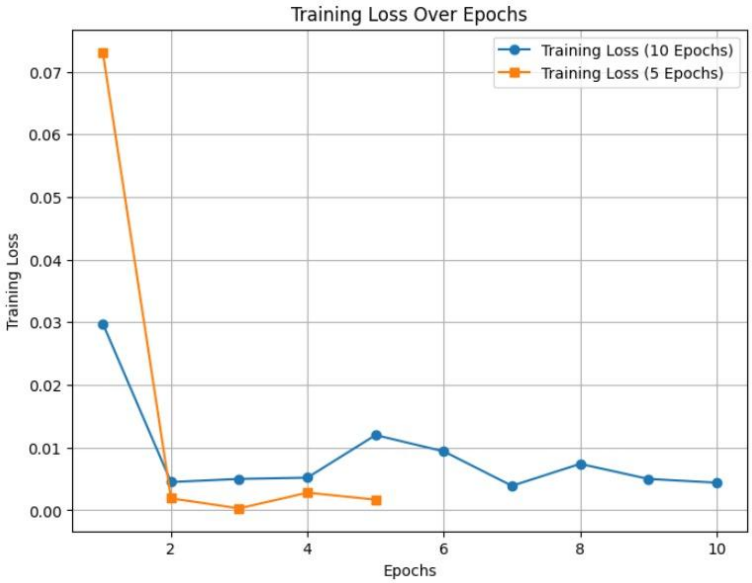

Figure 11 Training Over Epochs for Embedding Generation and Classification Using ResNet-18 CNN Model

Using a contrastive learning method with the ResNet-18 CNN backbone, the graph in **Figure 11** shows the training loss over epochs for classifying gastrointestinal disorders. The graph contrasts the training loss for two scenarios: In multidimensional embedding space, 10 epochs are applied to embed image features near one another for positive pairs and far from one another for negative pairs. Up to five epochs are also used for fine-tuning to map learnt embeddings in N classes.

In both situations, the training loss quickly drops in the first few epochs, suggesting that the model is learning effectively. Overfitting can be seen in shorter training cycles since the training loss for five epochs rapidly stabilizes and approaches zero loss after the second epoch. Better generalization over longer training durations is suggested by the training loss for 10 epochs, which exhibits minor variations but continuously achieves low loss values beyond the initial epochs. By successfully limiting both training and learning loss, the above graph illustrates effective learning and training in the classification of gastrointestinal Disease, underscoring the model's capacity to attain stability and accuracy in a short training period.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

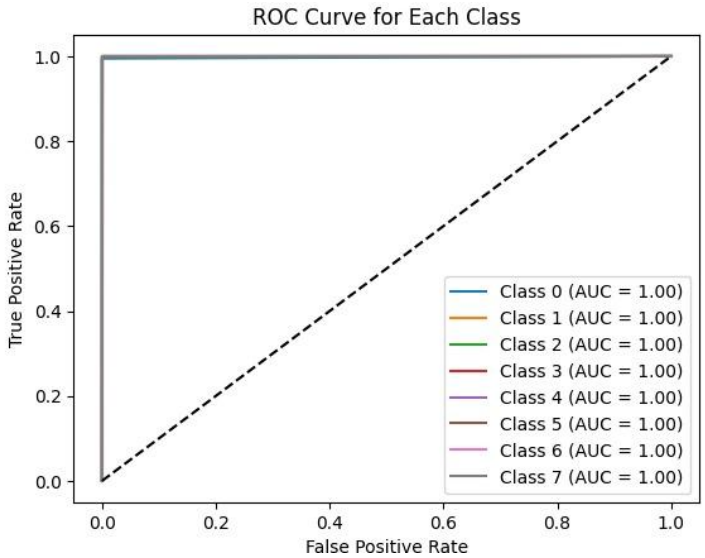

Figure 12 ROC Curve and AUC Values for Each Class of GI disease

Classifying each class effectively, the ROC curve in **Figure 12** shows the performance of the proposed model. For various threshold values, the ROC curve of the proposed model illustrates the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR). An AUC (Area Under the Curve) score of 1.00 signifies flawless classification with no errors or misclassifications across all classes. Close to the top-left edge of the graph the ROC curves for all classes, proposed model reflecting strong classification performance. The proposed model performs exceptionally well for all classes, as clearly shown by the AUC and ROC metrics. The high discriminative ability of the proposed model is confirmed by the perfect AUC score of 1.00, which makes the model highly reliable for the classification tasks of Gastrointestinal diseases.

Table 4 Classification Summary of 8 Classes of GI Disease as Per the Kvasir Dataset

Class	Recall	F1-Score	Precision	Accuracy %	Support
DL Polyps	0.99	1.0	1	100	1552
Polyps	1.0				1598
U Colitis					1608
N Cecum					1570
N Pylorus					1625
N Z Line					1612
DR Margins					1582
Esophagitis					1653
Macro Average	1.0				**12800**
Weighted Average	1.0				
Total					

For the classification of Gastrointestinal diseases, the performance of the proposed model based on Precision, Recall, F1-score, Accuracy, Macro Average, Weighted Average and support for each class is summarized in **Table 04.** The model performs perfect classification for Esophagitis, Dyed-resection-margins, Normal-cecum, Normal-pylorus, Polyps, Normal-z-line and Ulcerative-colitis, achieving 1.00 Precision, Recall and F1- score. This clearly indicates that there is no false negative or false positive for 12800 test samples. For Dyed-lifted-polyps, the proposed model achieved 1 precision, 0.99 recall, and 1 F1-score, which shows that 99% of instances of this class are identified perfectly. From 1552 test images, only 3 cases are marked as a false negative. Achieving an overall accuracy of 100% in the classification of Gastrointestinal disease across 12800 test samples of multiple classes, our

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

proposed model demonstrates the robustness and strong reliability for the classification of GI disease. The unweighted average across all classes is represented by macro average values, 1.00 for precision, 1 for recall, and 1 for F1-score), which clearly identifies that the proposed model treats all classes with consistent performance. The proposed model also performed efficiently for imbalanced class data, as shown by the weighted average values (1.00 for precision, 1 for recall, and 1 for F1-score, considering the number of samples in each class.

Table 5 compares the proposed model with state-of-the-art, latest models for Gastrointestinal disease classification using the Kvasir dataset. For 8-class Gastrointestinal disease classification, 85.7% accuracy is achieved by the InceptionResNetV2-based architecture [38]. ResNet-50 achieved an accuracy of 96.81% for classifying gastrointestinal diseases on the 7 classes of the Kvasir dataset. On the other hand, on the same dataset, VGG-19 achieved 94.21 accuracy [26]. A combination of fine-tuned ResrNet-50 with Mask R-CNN for region detection and ResNet-152 for feature extraction, along with an improved Ant Colony Optimization algorithm for feature selection and serial feature fusion approach, achieved 96.43% accuracy for classifying the gastrointestinal disease classification using the Kvasir dataset [27]. Utilizing a spatial attention mechanism to enhance the ConvMixer model "called Spatial-Attention ConvMixer (SAC) architecture" achieved an overall accuracy of 93.37 for Gastrointestinal classification using the Kvasir dataset [39]. For GI disease classification, a combination of transfer learning technique with EfficientNetB0 achieved a very good 98% accuracy [40]. RestNet50 trained on the augmented version of the Kvasir dataset achieved an overall 92.87 accuracy for classifying Gastrointestinal disease [41]. Compared to the above-stated AI model using different techniques and CNN models, The proposed model's overall performance for classifying gastrointestinal diseases is excellent. Most of the classes achieved 100% accuracy. These perfect results clearly highlight the proposed model's efficiency, robustness, and effectiveness in real-time clinical practices.

Table 5 Comparison With Other AI Models for Gastrointestinal Disease Classification

Ref#	Year	Dataset	Accuracy %	Recall	F1- score	Precision
[38]	2024	Kvasir	85.7	x	x	x
[26]	2023		96.81(7 Classes)	95.28	94.85	95
[27]	2023		96.43	x	x	x
[39]	2024		93.37	93.37	93.42	93.66
[40]	2023		98	98	x	98
[41]	2024		92.87	x	92.87	x
Proposed Model			**100**	**100**	**100**	**100**

5. Conclusion

The proposed model, leveraging the Contrastive learning technique with ResNet-18 as Backbone CNN, exhibited exceptional performance in classifying Gastrointestinal disease on the Kvasir dataset. The proposed model efficiently classifies eight categories of Gastrointestinal disease with an excellent overall accuracy of 100%. The main reason for some minor misclassifications is that the visual and structural features of various classes overlap with each other, which probably leads to confusion during the categorization process. The confusion matrix, which demonstrates that the suggested model can accurately classify most classes, supports these findings even more. At the same time, the number of

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.01 (2025)

misclassifications between some classes is also relatively low. Low AUC values ≈ 1.00 confirmed the excellent discriminative ability of the model for different classes of Gastrointestinal disease and the high reliability of the proposed model for the classification task. 100% Precision, recall, and F1 score across classes indicate the proposed model's robustness. Despite overlapping features, the suggested model performs well and receives perfect scores for all classes. This study demonstrates the perfect combination of ResNet-18 with a contrastive learning technique for classifying endoscopic images with 100% accuracy. The proposed model draws attention to the potential use of these models in clinical settings to aid in accurately diagnosing gastrointestinal diseases. Future work may improve the model's classification capabilities for diverse datasets to address data scarcity and provide better explainability and interpretability to make the models more transparent by providing better insight into the decision-making process for real-time applications.

References

[1] Y. Wang, Y. Huang, R. C. Chase, T. Li, D. Ramai, S. Li, *et al.*, "Global burden of digestive diseases: a systematic analysis of the global burden of diseases study, 1990 to 2019," *Gastroenterology,* vol. 165, pp. 773-783. e15, 2023.

[2] P. Danpanichkul, K. Suparan, P. Tothanarungroj, D. Dejvajara, K. Rakwong, Y. Pang, *et al.*, "Epidemiology of gastrointestinal cancers: a systematic analysis from the Global Burden of Disease Study 2021," *Gut,* 2024.

[3] R. Wang, Z. Li, S. Liu, and D. Zhang, "Global, regional, and national burden of 10 digestive diseases in 204 countries and territories from 1990 to 2019," *Frontiers in Public Health,* vol. 11, p. 1061453, 2023.

[4] Y. Okagawa, S. Abe, M. Yamada, I. Oda, and Y. Saito, "Artificial Intelligence in Endoscopy," *Digestive Diseases and Sciences,* vol. 67, pp. 1553-1572, 2022/05/01 2022.

[5] C. Labaki, E. N. Uche-Anya, and T. M. Berzin, "Artificial Intelligence in Gastrointestinal Endoscopy," *Gastroenterology Clinics,* vol. 53, pp. 773-786, 2024.

[6] V. Chinnasamy, V. Sakulsaengprapha, and S. C. Mathews, "The Benefit of Artificial Intelligence-Based Diagnosis in Gastroenterology and Hepatology Is Highly Variable: A Diagnostic Need and Burden Analysis," *Gastroenterology,* vol. 165, pp. 788-790.e3, 2023.

[7] O. Parkash, A. T. S. Siddiqui, U. Jiwani, F. Rind, Z. A. Padhani, A. Rizvi, *et al.*, "Diagnostic accuracy of artificial intelligence for detecting gastrointestinal luminal pathologies: A systematic review and meta-analysis," *Frontiers in Medicine,* vol. 9, 2022-November-04 2022.

[8] A. S. Hosain, M. Islam, M. H. K. Mehedi, I. E. Kabir, and Z. T. Khan, "Gastrointestinal disorder detection with a transformer based approach," in *2022 IEEE 13th annual information technology, electronics and mobile communication conference (IEMCON)*, 2022, pp. 0280-0285.

[9] Z. M. Lonseko, P. E. Adjei, W. Du, C. Luo, D. Hu, L. Zhu, *et al.*, "Gastrointestinal Disease Classification in Endoscopic Images Using Attention-Guided Convolutional Neural Networks," *Applied Sciences,* vol. 11, p. 11136, 2021.

[10] J. Escobar, K. Sanchez, C. Hinojosa, H. Arguello, and S. Castillo, "Accurate Deep Learning-based Gastrointestinal Disease Classification via Transfer Learning Strategy," in *2021 XXIII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, 2021, pp. 1-5.

[11] M. Nouman Noor, M. Nazir, S. A. Khan, O.-Y. Song, and I. Ashraf, "Efficient Gastrointestinal Disease Classification Using Pretrained Deep Convolutional Neural Network," *Electronics,* vol. 12, p. 1557, 2023.

[12] Q. Jiang, Y. Yu, Y. Ren, S. Li, and X. He, "A review of deep learning methods for gastrointestinal diseases classification applied in computer-aided diagnosis system," *Medical & Biological Engineering & Computing,* 2024/09/30 2024.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

[13] V. Thambawita, D. Jha, H. L. Hammer, H. D. Johansen, D. Johansen, P. Halvorsen, *et al.*, "An extensive study on cross-dataset bias and evaluation metrics interpretation for machine learning applied to gastrointestinal tract abnormality classification," *ACM Transactions on Computing for Healthcare,* vol. 1, pp. 1-29, 2020.

[14] S. Lee and I. Lee, "Comprehensive assessment of machine learning methods for diagnosing gastrointestinal diseases through whole metagenome sequencing data," *Gut Microbes,* vol. 16, p. 2375679, 2024/12/31 2024.

[15] S. Parasa, T. Berzin, C. Leggett, S. Gross, A. Repici, O. F. Ahmad, *et al.*, "Consensus statements on the current landscape of artificial intelligence applications in endoscopy, addressing roadblocks, and advancing artificial intelligence in gastroenterology," *Gastrointestinal Endoscopy*.

[16] H. Gunasekaran, K. Ramalakshmi, D. K. Swaminathan, A. J, and M. Mazzara, "GIT-Net: An Ensemble Deep Learning-Based GI Tract Classification of Endoscopic Images," *Bioengineering,* vol. 10, p. 809, 2023.

[17] W. Du, N. Rao, J. Yong, Y. Wang, D. Hu, T. Gan, *et al.*, "Improving the classification performance of esophageal disease on small dataset by semi-supervised efficient contrastive learning," *Journal of Medical Systems,* vol. 46, pp. 1-13, 2022.

[18] D. Nath and G. Shahariar, "Gastrointestinal disease classification through explainable and cost-sensitive deep neural networks with supervised contrastive learning," *arXiv preprint arXiv:2307.07603,* 2023.

[19] J. Naz, M. I. Sharif, M. I. Sharif, S. Kadry, H. T. Rauf, and A. E. Ragab, "A comparative analysis of optimization algorithms for gastrointestinal abnormalities recognition and classification based on ensemble XcepNet23 and ResNet18 features," *Biomedicines,* vol. 11, p. 1723, 2023.

[20] R. Sali, S. Adewole, L. Ehsan, L. A. Denson, P. Kelly, B. C. Amadi, *et al.*, "Hierarchical deep convolutional neural networks for multi-category diagnosis of gastrointestinal disorders on histopathological images," in *2020 IEEE International Conference on Healthcare Informatics (ICHI)*, 2020, pp. 1-6.

[21] S. Wu, R. Zhang, J. Yan, C. Li, Q. Liu, L. Wang, *et al.*, "High-Speed and Accurate Diagnosis of Gastrointestinal Disease: Learning on Endoscopy Images Using Lightweight Transformer with Local Feature Attention," *Bioengineering,* vol. 10, p. 1416, 2023.

[22] Y. El Khalfaoui, B. Alibouch, and A. F. El Ouafdi, "Gastro-CNN-VIT: Vision Transformer and Deep CNNs for Detecting GI Diseases in WCE Images," Cham, 2024, pp. 131-137.

[23] A. Subedi, S. Regmi, N. Regmi, B. Bhusal, U. Bagci, and D. Jha, "Classification of Endoscopy and Video Capsule Images Using CNN-Transformer Model," in *MICCAI Workshop on Cancer Prevention through Early Detection*, 2024, pp. 26-36.

[24] Y.-j. Kim, H. C. Cho, and H.-c. Cho, "Deep learning-based computer-aided diagnosis system for gastroscopy image classification using synthetic data," *Applied Sciences,* vol. 11, p. 760, 2021.

[25] K.-S. Lee and E. S. Kim, "Explainable artificial intelligence in the early diagnosis of gastrointestinal disease," *Diagnostics,* vol. 12, p. 2740, 2022.

[26] X. A. Mary, A. Raj, C. S. Evangeline, T. M. Neebha, V. B. Kumaravelu, and P. Manimegalai, "Multi-class Classification of Gastrointestinal Diseases using Deep Learning Techniques," *The Open Biomedical Engineering Journal,* vol. 17, 2023.

[27] M. Alhajlah, M. N. Noor, M. Nazir, A. Mahmood, I. Ashraf, and T. Karamat, "Gastrointestinal diseases classification using deep transfer learning and features optimization," *Comput. Mater. Contin,* vol. 75, pp. 2227-2245, 2023.

[28] S. Jiang, T. Wang, and K.-H. Zhang, "Data-driven decision-making for precision diagnosis of digestive diseases," *BioMedical Engineering OnLine,* vol. 22, p. 87, 2023/09/01 2023.

[29] S. Dilmaghani and N. Coelho-Prabhu, "Role of Artificial Intelligence in Colonoscopy: A Literature Review of the Past, Present, and Future Directions," *Techniques and Innovations in Gastrointestinal Endoscopy,* vol. 25, pp. 399-412, 2023.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

[30] J. R. Lewis, S. Pathan, P. Kumar, and C. C. Dias, "AI in Endoscopic Gastrointestinal Diagnosis: A Systematic Review of Deep Learning and Machine Learning Techniques," *IEEE Access,* pp. 1-1, 2024.

[31] S. Mukherjee, S. Vagha, and P. Gadkari, "Navigating the Future: A Comprehensive Review of Artificial Intelligence Applications in Gastrointestinal Cancer," *Cureus,* vol. 16, p. e54467, 2024/2/19 2024.

[32] C. Le Berre, W. J. Sandborn, S. Aridhi, M.-D. Devignes, L. Fournier, M. Smaïl-Tabbone*, et al.*, "Application of Artificial Intelligence to Gastroenterology and Hepatology," *Gastroenterology,* vol. 158, pp. 76-94.e2, 2020.

[33] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*, 2020, pp. 1597-1607.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[35] F. Wang and H. Liu, "Understanding the behaviour of contrastive loss," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 2495-2504.

[36] A. Mao, M. Mohri, and Y. Zhong, "Cross-entropy loss functions: Theoretical analysis and applications," in *International conference on Machine learning*, 2023, pp. 23803-23828.

[37] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen*, et al.*, "Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection," in *Proceedings of the 8th ACM on Multimedia Systems Conference*, 2017, pp. 164-169.

[38] A. Ali, A. Iqbal, S. Khan, N. Ahmad, and S. Shah, "A two-phase transfer learning framework for gastrointestinal diseases classification," *PeerJ Computer Science,* vol. 10, p. e2587, 2024.

[39] A. A. Demirbaş, H. Üzen, and H. Fırat, "Spatial-attention ConvMixer architecture for classification and detection of gastrointestinal diseases using the Kvasir dataset," *Health Information Science and Systems,* vol. 12, p. 32, 2024.

[40] J. Thomas Abraham, A. Muralidhar, K. Sathyarajasekaran, and N. Ilakiyaselvan, "A deep-learning approach for identifying and classifying digestive diseases," *Symmetry,* vol. 15, p. 379, 2023.

[41] V. Y. Cambay, P. D. Barua, A. Hafeez Baig, S. Dogan, M. Baygin, T. Tuncer*, et al.*, "Automated Detection of Gastrointestinal Diseases Using Resnet50*-Based Explainable Deep Feature Engineering Model with Endoscopy Images," *Sensors,* vol. 24, p. 7710, 2024.