

Article

Localization and Classification of Gastrointestinal Tract Disorders Using Explainable AI from Endoscopic Images

Muhammad Nouman Noor ¹, Muhammad Nazir ^{1,*} , Sajid Ali Khan ², Imran Ashraf ³ and Oh-Young Song ^{4,*} 

¹ Department of Computer Science, HITEC University, Taxila 47080, Pakistan

² Department of Software Engineering, Foundation University Islamabad, Islamabad 44000, Pakistan

³ Department of Computer Engineering, HITEC University, Taxila 47080, Pakistan; imran.ashraf@hitecuni.edu.pk

⁴ Software Department, Sejong University, Seoul 05006, Republic of Korea

* Correspondence: muhammad.nazir@hitecuni.edu.pk (M.N.); oysong@sejong.edu (O.-Y.S.)

Abstract: Globally, gastrointestinal (GI) tract diseases are on the rise. If left untreated, people may die from these diseases. Early discovery and categorization of these diseases can reduce the severity of the disease and save lives. Automated procedures are necessary, since manual detection and categorization are laborious, time-consuming, and prone to mistakes. In this work, we present an automated system for the localization and classification of GI diseases from endoscopic images with the help of an encoder–decoder-based model, XceptionNet, and explainable artificial intelligence (AI). Data augmentation is performed at the preprocessing stage, followed by segmentation using an encoder–decoder-based model. Later, contours are drawn around the diseased area based on segmented regions. Finally, classification is performed on segmented images by well-known classifiers, and results are generated for various train-to-test ratios for performance analysis. For segmentation, the proposed model achieved 82.08% dice, 90.30% mIOU, 94.35% precision, and 85.97% recall rate. The best performing classifier achieved 98.32% accuracy, 96.13% recall, and 99.68% precision using the softmax classifier. Comparison with the state-of-the-art techniques shows that the proposed model performed well on all the reported performance metrics. We explain this improvement in performance by utilizing heat maps with and without the proposed technique.

Keywords: gastrointestinal diseases; deep learning; endoscopic images; heat maps



Citation: Nouman Noor, M.; Nazir, M.; Khan, S.A.; Ashraf, I.; Song, O.-Y. Localization and Classification of Gastrointestinal Tract Disorders Using Explainable AI from Endoscopic Images. *Appl. Sci.* **2023**, *13*, 9031. <https://doi.org/10.3390/app13159031>

Academic Editors: Zhonghua Sun and Jan Egger

Received: 28 April 2023

Revised: 3 August 2023

Accepted: 4 August 2023

Published: 7 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

GI tract diseases are disorders related to the digestive system. Diagnoses of these diseases are highly dependent on medical imaging. The processing of large visual data is difficult for medical professionals and radiologists; this renders it subject to incorrect medical evaluation [1]. The most common diseases that occur in the digestive system are ulcerative colitis, ulcers, esophagitis, and polyps, which can transform into colorectal cancer. These diseases are the key causes of mortality around the globe [2].

As per the survey conducted on colorectal cancer for the year 2019, 26% of men, as well as 11% of women, around the globe are diagnosed with this cancer [3]. In 2021, more than 0.3 million cases of colorectal cancer were diagnosed in the US, and the death toll rose to 44% [4]. Roughly 0.7 million new instances of diseases are reported each year worldwide [5]. Alongside GI malignant growth [6,7], ulcer advancement in the GI tract is additionally a significant illness. The authors of [8] announced that the most noteworthy yearly predominance of ulcers was 141 per 1000 people in Spain, and the least was around 57 in Sweden.

During a routine endoscopic checkup, many lesions are missed due to the factors like the presence of stool and because of the organ's multifaceted topology. Although the bowel is cleansed for improvement in the detection of cancer or its predecessor lesions, still, the ratio of missed polyps is immoderate, from 21.4–26.8% [9]. Moreover, the interclass

similarity between lesions also plays a challenging role in identifying them. A recent procedure called wireless capsule endoscopy (WCE) [10] empowers specialists to see inside the intestinal tract, a region that is undeniably challenging to reach with regular endoscopy. In WCE, the patient swallows a camera-containing capsule that catches many images as it travels through the GI tract. These images are stitched together to form a video, which is then, at that point, examined by the specialists (expert gastroenterologists) to track down deformations. This manual strategy requires 2–3 h overall, so analysts are currently creating different computerized techniques [11,12]. Therefore, an automated system is required that not only classifies the diseases, but also highlights the diseased area.

Several techniques for identifying colorectal cancer and other diseases using endoscopic images have been proposed in the literature by computer vision (CV) and machine learning (ML) researchers. In [13], the authors proposed a technique with which they established a feature matrix, and then classified these features using a support vector machine (SVM) and decision trees. The maximum achieved accuracy using this method was 93.64%. Similarly, in another paper [7], the authors employed the information of local as well as global features and fused them. Deep discriminant features were obtained using an adaptive aggregation feature module. The best achieved accuracy using the stated techniques was 96.37%. The challenge has persisted up to this point due to the similarities between many symptoms, such as color, shape, and texture of lesions. Moreover, many researchers have focused on single disease detection and binary classification [14–16]. Furthermore, the localization of diseases is also a major challenge that requires addressing. Hence, in this paper, we propose a technique to address these challenges and perform multiclass classification after the localization of diseases. The major diseases analyzed in our research are polyps, esophagitis, ulcers, and ulcerative colitis, as well as a normal class.

The following summarizes the paper's primary contributions:

- Development of an encoder–decoder-based model for segmentation and localization of diseases.
- Development of an explainable AI-based model that is utilized for the classification of endoscopic images with contours into four main diseases.
- Development of an efficient and robust framework having better accuracy, precision, and recall rate.

The remainder of this paper is structured as follows: related works are presented in Section 2. The methodological specifics are then provided in Section 3. The experimental data are included in Section 4. In Section 5 discussion and analysis of experimental results are presented and Finally, Section 6 concludes the paper.

2. Literature Review

In the past few years, the detection of diseases using medical imaging has been a hot area of research, especially in the domain of the gastrointestinal tract. The segmentation of polyps, in particular, has been the major focus because of the availability of ground truths. Furthermore, the classification of gastrointestinal diseases has also been an active area of research. The performance of machine learning algorithms reported in the literature has been quite impressive [17,18], but deep learning algorithms surpass the ML approaches and achieve better results [19].

For the detection of GI tract diseases, numerous studies are available in the literature that use the ML method. For example, in [17], the authors developed a ML model based on the longitudinal training cohort of over 20 thousand patients undergoing treatment for peptic ulcers between the years 2007 and 2016. Their greatest accuracies were 82.6% and 83.3% using logistic regression and ridge regression, respectively. Sen Wang et al. [18] established ML architecture for ulcer diagnosis and performed experimentation on a private developed dataset of WCE videos, 1504 to be exact. The effectiveness of this technique was evaluated using the ROC curve and the AUC, and achieved a 0.9235 peak value. In a different work [13], Jinn-Yi Yeh et al. used color characteristics and a WCE image collection to identify bleeding and ulcers. They used texture information in addition to combining

all the picture attributes into a single matrix. Several classifiers, including SVM, neural networks, as well as decision trees, were presented in this matrix of characteristics. Various performance metrics were included for examination, and the accuracy ranged from 92.86% to 93.64%.

It has been observed that deep learning (DL) models generally performed better in detecting GI tract diseases. The authors of [20] developed the VGGNet model based on CNN to detect GI ulcers, with a dataset of 854 images, and achieved 86.6% accuracy. However, these tests took place using conventional endoscopy images. In [21], the authors developed a CNN-based DL model; the dataset consisted of 5360 images containing ulcers and erosions, and contained merely 450 normal class images. The method achieved 90.8% detection accuracy. Sekuboyina and co-authors, in [22], proposed models based on CNN to detect dissimilar forms of diseases in WCE images, like ulcers, and more. They developed multiple subsections of images and applied the DL model. This experiment attained 71% sensitivity and 72% specificity.

Apart from the classification techniques, researchers also proposed segmentation techniques for the detection of the predecessor disease of colorectal cancer. A fully convolutional network (FCN) was proposed in [23], which is trained from start to finish as well as pixel by pixel, and yields the segmentation of polyps. There are no extra postprocessing procedures needed for the suggested model, which is the major contribution of this research. In another paper [24], the authors discussed and enhanced the FCN network and named it the U-Net architecture. The U-Net model achieved good results for localization. Furthermore, many researchers have tried to modify and enhance the U-Net architecture to achieve better segmentation and localization results [25–27], but in medical images, these are not evaluated or do not provide better results. By maximizing the characteristics gleaned from two pre-trained models, the authors of [28] established a framework for gastrointestinal illness categorization and achieved 96.43% accuracy. In another framework [29], MobileNet-V2 is used for the multiclass classification of gastrointestinal illnesses, and a contrast enhancement approach was suggested.

Based on the literature, it can be said that sufficient related work has been performed in the field of GI tract disease detection and classification. The presented results show reasonable performance in terms of accuracy. However, performance can be improved. Accuracy is an important performance metric; however, for multiclass classification problems, accuracy is less significant as compared to other performance metrics, especially when there is an imbalance in the dataset. For instance, we would like to emphasize that precision and recall rate are important performance measures for life-critical applications. Most of the presented works have reasonable accuracy, but they suffer from lower precision and recall rate, and require improvement.

Review of the existing work also highlights that most of the work on GI tract diseases has been conducted on datasets that are not publicly available. This makes it hard to generalize the results and compare the performance. Furthermore, researchers mostly focused on single disease detection and binary classification [14–16]. The focus of our work is to conduct experiments on publicly available datasets and target the multiclass classification of GI tract diseases like polyps, ulcers, ulcerative colitis, and esophagitis. Also, the suggested strategy has significantly improved the performance across practically all indicators.

3. Methodology

Various diseases can attack the human GI tract, like colorectal cancer, and their predecessor diseases, like polyps, as well as other diseases, such as ulcers, esophagitis, and ulcerative colitis, to name a few. To diagnose such diseases, traditional endoscopic images or WCE images are needed and play a vital role. Artificial intelligence-based methods like DL have proved to be helpful for the diagnosis of such diseases. Therefore, in this paper, we have developed a DL-based model for segmentation as well as a multiclass classification of GI tract diseases. The core aim of our research is to put forward a DL model based on

the segmented images. This approach is used for the detection of multiple GI tract diseases, and hence, is used for reducing the doctors' time to manually diagnose or use multiple applications separately for each malady.

In our proposed methodology, we undertake the five steps shown in Figure 1. As a first step, we acquired the publicly available datasets, namely, Kvasir-Seg, Kvasir V-2, and Hyper-Kvasir datasets. After that, the dataset was increased by applying data augmentation using multiple transformations. Subsequently, segmentation was performed using U-Net, an encoder–decoder-based model, with Resnet-34 as a backbone, and then, contours were drawn around the diseased area. In the second-last step, heat maps were generated to compare and analyze the model's performance on segmented and non-segmented images. In the last step, images with contours around the diseased area were used as an input of the Xception model for feature extraction, and multiple classifiers are applied for classification.

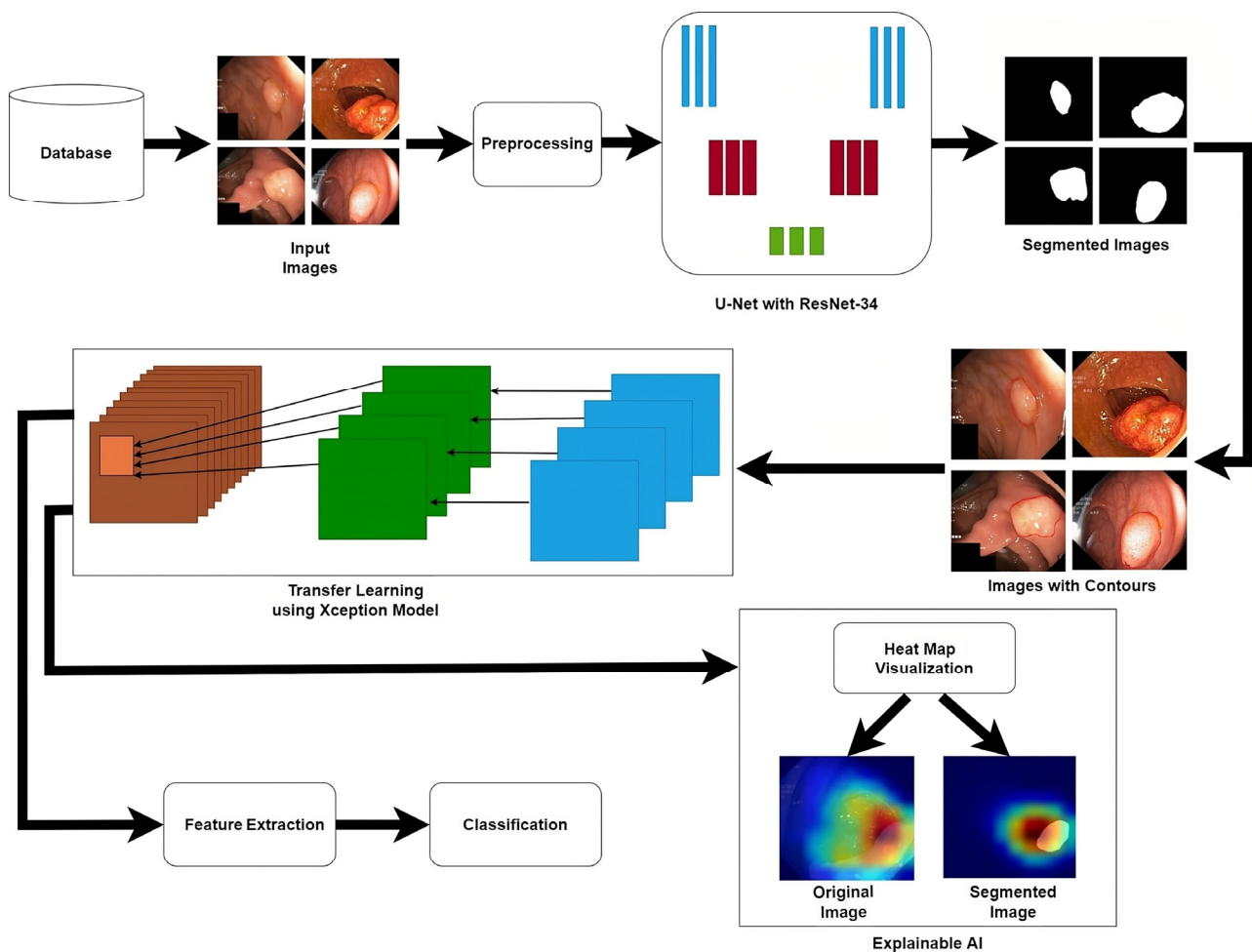


Figure 1. Proposed methodology of localization and classification of GI tract disorders.

3.1. Dataset Collection and Preparation

The Kvasir-Seg [30] dataset was utilized for segmentation, and the Kvasir-V2 [31] and Hyper-Kvasir [32] datasets were utilized for classification. Our dataset contains four diseases, i.e., ulcers, polyps, esophagitis, and ulcerative colitis, as well as a normal class, with 1000 instances for each malady, other than the ulcer malady, which has only 854 instances. As a result, our dataset consists of 4854 images divided into five classes: ulcerative colitis, polyps, ulcers, esophagitis, and normal. For segmentation, the Kvasir-Seg dataset is used, which contains the 1000 images of the polyp class with their ground truths.

Initially, the segmentation results were collected based on Kvasir-Seg dataset, and then this method was applied to all other diseased images, and classification was performed.

3.2. Preprocessing

DL models require more data to train on as compared with ML models, otherwise they start overfitting the data and lacking generalization. Hence, augmentation was performed after the dataset was initially collected to enhance its size. Moreover, data augmentation is a very powerful technique used to reduce the validation error along with the training error [33]. The main transformations that are applied during the data augmentation are rotation, width shifting, height shifting, horizontal/vertical flip, and zoom-in/out. The total dataset size after applying data augmentation increased to 30,000 images, with 6000 images for each class. The images generated after applying data augmentation are shown in Figure 2.

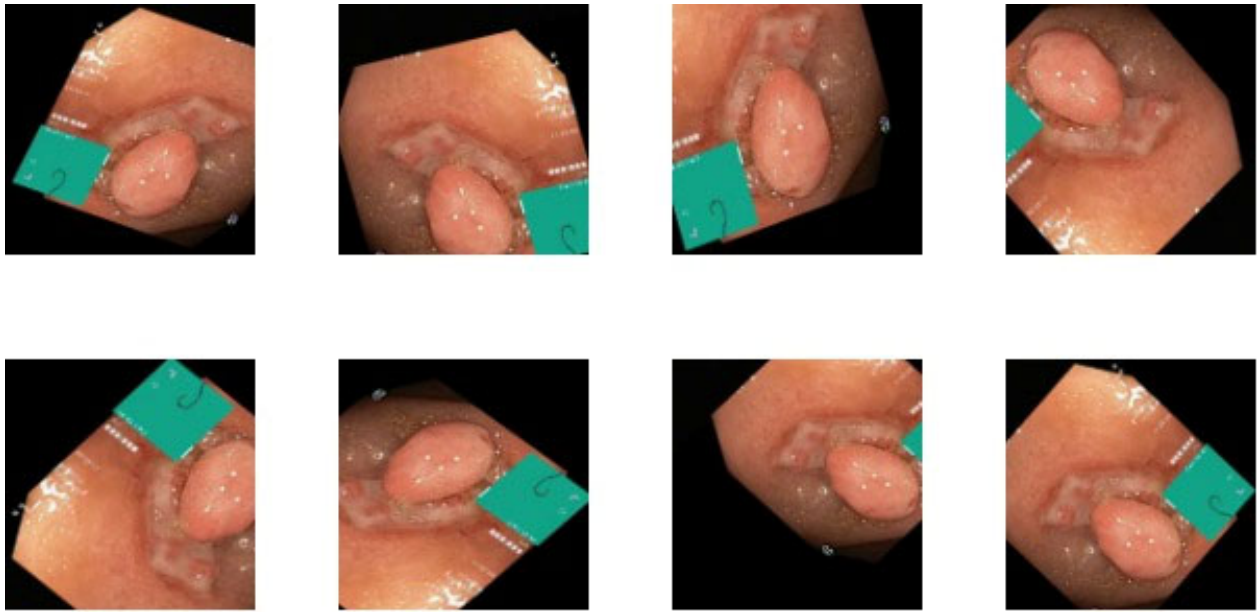


Figure 2. Polyp images after applying data augmentation.

3.3. Segmentation

Segmentation of the diseased region was performed using the U-Net model. U-Net is a CNN-based segmentation model that was proposed in 2015 for biomedical images [23]. It has one encoder module and another decoder module. Figure 3 depicts the U-Net model's architecture. In the encoder module, two convolutional (3×3) layers are applied repeatedly with one stride. The Relu layer and a 2×2 Maxpooling layer with two and four strides follow each convolutional layer. A dropout layer is applied following the first convolutional layer. The bottom layers consist of 3×3 convolutional layers. The decoding part up-samples the dimensions of the image to its original by applying two convolutional (3×3) layers. The first layer is stacked by Relu; the dropout layer and the next convolutional layer are stacked by the Relu layer only. The top layer, which is also the last layer, is a convolutional (1×1) layer. The first encoder part is used for the extraction of features, and is similar to the VGG-16 model [34]. The up-sampling operation combines both low-resolution as well as high-resolution information, which is the provision of object-based recognition, as well as accurate positioning and segmentation, which is useful for medical image segmentation [34]. As the foundational model for U-Net, we utilized the ResNet-34 model, which was observed to outperform other segmentation models [35]. The U-Net model outputs the black-and-white image mask, which was then used to draw contours around the diseased area of an image.

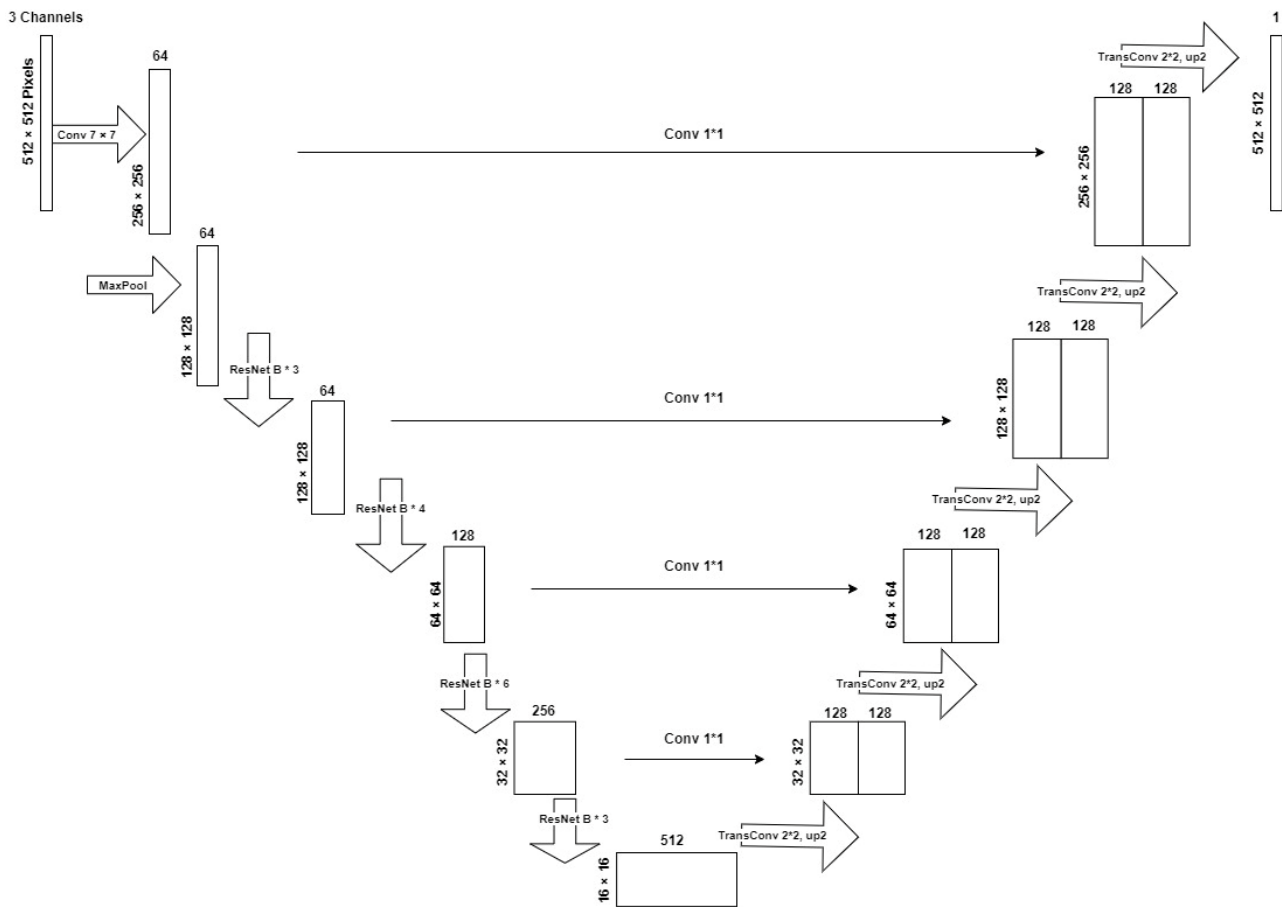


Figure 3. U-Net architecture diagram with ResNet-34 as the backbone model [36].

During model training, the Adam optimizer was applied. Due to its outstanding outcomes and adaptable learning gain, the Adam optimizer is frequently used by researchers for CNNs [37], and root mean squared error (RMSE) is used as a loss function. The model was trained for a total of 250 epochs with a batch size of 50.

The Adam optimizer is used to control the gradient descent rate in such a way that there is minimum fluctuation near to global optima, and it takes large steps near to local optima to avoid it and reach global minima efficiently. Adam combines the features of two gradient descent techniques, namely, momentum and root mean squared propagation (RMSP). Mathematical equations of momentum and RMSP are expressed as follows:

$$R_t = \sigma_1 * R_t + (1 - \sigma_1) * \left(\frac{\delta S}{\delta \nabla_t} \right) \quad (1)$$

$$\varphi_t = \sigma_2 * \varphi_t + (1 - \sigma_2) * \left(\frac{\delta S}{\delta \nabla_t} \right)^2 \quad (2)$$

where R_t is the gradient aggregate at t , δS is a derivative of a loss function, $\delta \nabla_t$ is a derivative of weights at t , σ is an average parameter that is moving, and φ_t is the sum of the square of past gradients. Initially, both R_t and φ_t are set to zero, and it is observed that both tend to be biased towards zero as σ_1 and σ_2 are set to one. The Adam optimizer solved this problem by calculating bias-corrected $\widehat{R}_t$ as well as $\widehat{\varphi}_t$. Mathematical equations of these biased corrected values are expressed as follows:

$$\widehat{R}_t = \frac{R_t}{1 - \sigma_1^t} \quad (3)$$

$$\hat{\varphi}_t = \frac{\varphi_t}{1 - \sigma_2^t} \quad (4)$$

After each iteration, new positions of weights by substituting the updated values are given as follows:

$$\forall_t = \forall(t-1) - \exists * \left(\frac{\hat{R}_t}{\sqrt{(\hat{\varphi}_t)}} + \mu \right) \quad (5)$$

where $\forall_t$ is a weight at time t , $\exists$ is a learning rate, and μ is a constant.

Mean squared error (MSE) is called an average of squares of errors. It is the square of the difference between the actual attribute and estimator. Mathematically, the equation of mean squared error is expressed as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\mathfrak{L}_i - \hat{\mathfrak{L}}_i) \quad (6)$$

where $\mathfrak{L}_i$ is the original valuation, and $\hat{\mathfrak{L}}_i$ is the anticipated valuation of the model.

3.4. Heat Maps

Explainable artificial intelligence (XAI) in medical imaging is a set of techniques and approaches to enable medical experts to understand the diseased judgment process of artificial intelligent models. The gradient-weighted class activation map (Grad-CAM) is a tool created in 2017 that produces an explanation for each type of CNN model [38,39]. The heat map of the anticipated labels is the Grad-CAM result.

Heat maps of images were generated before segmentation and after segmentation for the analysis of the diseased area in an image. The magnitude with which the model highlights the area is called activation, and we exhibit this on the Jet color map. Violet color highlights the lowest-magnitude area, and red represents the high-magnitude area. The process of heat map generation is shown in Figure 4.

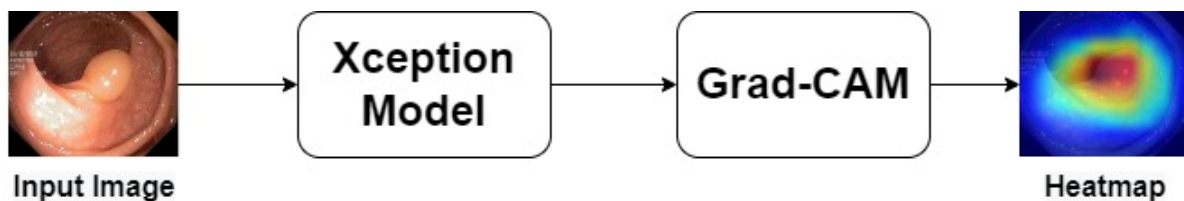


Figure 4. Heat map generation process for an endoscopic image.

Grad-CAM works by checking the last convolutional layer before and after the examination of gradient information that is flowing to that layer. In our case, we applied the transfer learning concept and used the pre-trained Xception model, as it provides the best heat maps, and is therefore used for classification as well. The results of the image and its heat map before and after segmentation are shown in Figure 5. It is apparent from Figure 5 that after segmentation, the model is more focused, and looks exactly at the diseased region as the high-magnitude area; therefore, we used images with contours drawn around the diseased area for classification.

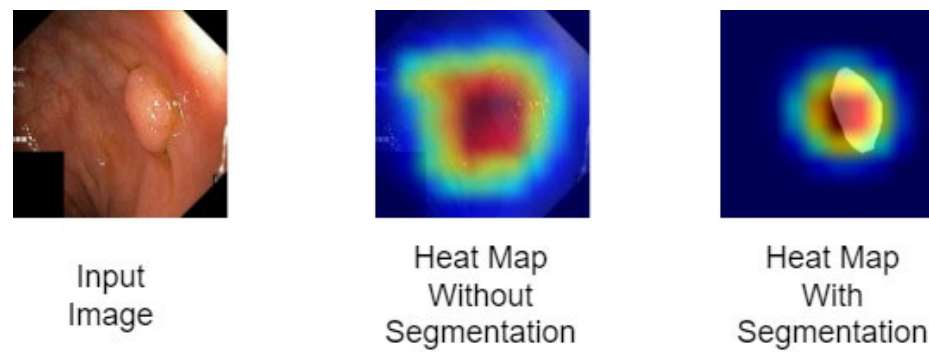


Figure 5. Heat map visualization of input image before and after segmentation using the Xception-Net model.

3.5. Features Extraction and Classification

As a final step, the Xception model was fine-tuned, and multiple classifiers were applied to predict the true labels. In our proposed model, the transfer learning approach is utilized as it performs better than training completely from beginning [40,41]. The Xception model, which was pre-trained on the ImageNet dataset, was used and fine-tuned on our dataset by applying a dropout layer with 0.4 probability. The input of the Xception model is the images with contours, and the output is the features. These features are used for classification by applying multiple classifiers, like softmax, linear SVM, quadratic SVM, and Bayesian.

The Xception model is based on CNN with depth-wise separable convolutional layers. This model has 36 convolutional layers that are arranged into 14 modules. In simple terms, the Xception model is a depth-wise separable CNN with a residual connection. The architecture of Xception is shown in Figure 6. The authors of [42] proved through experimentation that Xception outperforms other CNN models like VGG-16, ResNet-152, and Inception V3 on the ImageNet dataset.

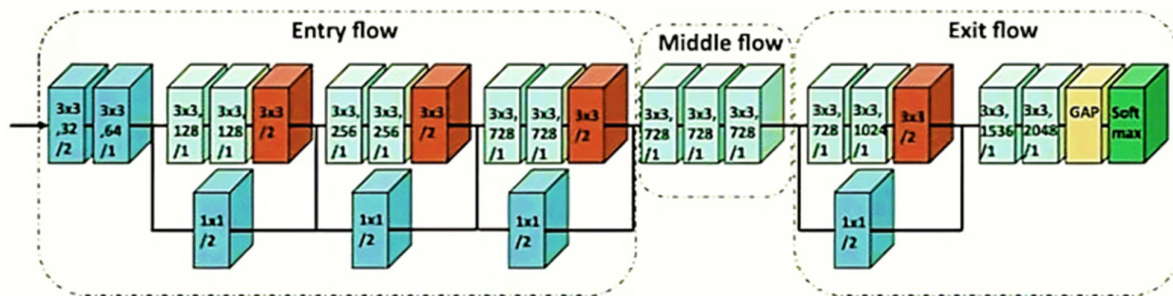


Figure 6. XceptionNet architecture [42].

For experimentation, Python was used, and other settings are shown here in order to reproduce the results. During model training, the Adam optimizer was applied. Because of its outstanding outcomes and adaptable learning gain, the Adam optimizer is frequently used by researchers for CNNs [37]. Categorical cross-entropy (CCE) was also employed as a loss function. During the training of the DL model, the loss function determines the difference between the original class and the anticipated class. It also adjusts the weights of the CNN to produce a better-fitting model [43]. The set batch size was 50 and the model was trained on 250 epochs.

CCE loss is an excellent measure for calculating loss by computing how distinguished two discrete probabilities are from each other. The mathematical equation of this loss is as follows:

$$CCE = - \sum_{n=1}^{Outcome\ Size} S_i \cdot \log \hat{S}_i \quad (7)$$

where S_i is the original valuation, and $\hat{S}_i$ is the anticipated valuation of the model.

4. Results

The results of our proposed model are compiled separately for both segmentation as well as classification in Sections 4.1 and 4.2, respectively. Evaluation matrices used for evaluating the results are dice, mIOU, precision, recall, and accuracy. Python 3.10, Matplotlib 3.6.2, PyTorch 1.12.0, and Keras 2.11.0 are the primary tools and libraries used for experimentation. The Adam optimizer and CCE are used, and the entire framework is developed on a GPU with a 4 GB NVIDIA Tesla graphics card and 32 GB of RAM. The model was trained on 250 epochs with a fixed batch size of 50.

4.1. Segmentation Results

Colorectal cancer and its predecessor disease segmentation results can be evaluated using different measures. It is highly dependent on the rate of detection as well as on the fraction between complete pixels and diseased pixels. To check the effectiveness of segmentation using U-Net with ResNet-34 as a backbone model, we performed a set of experiments on the Kvasir-Seg dataset. Performance measures used to check the efficiency of segmentation are dice, mIOU, precision, and recall.

Dice, which is also known as the overlap measure, is the most frequently used measure for evaluating and testing the effectiveness of medical image segmentation [44]. This overlap region between the predicted segmented image and the ground truth is doubled, and the result is divided by the total number of pixels in both images. mIOU, known as the mean intersection over union, is usually used to check for medical segmentation. IOU is calculated as the anticipated segmentation overlap over the ground truth divided by the total number of pixels. Mean IOU is calculated by taking the IOU of each label and averaging them. A precision measure is defined as the quality of being accurate. It measures the quality of our predictions. Recall is a measure used to calculate the positive points in the ground truth that are predicted positively by a model. Mathematical equations of these performance measures are provided as Equations (8)–(11).

$$Dice = \frac{2 * M}{2 * M + N + O} \quad (8)$$

$$mIOU = \frac{M}{M + N + O} \quad (9)$$

$$Precision = \frac{M}{M + N} \quad (10)$$

$$Recall = \frac{M}{M + O} \quad (11)$$

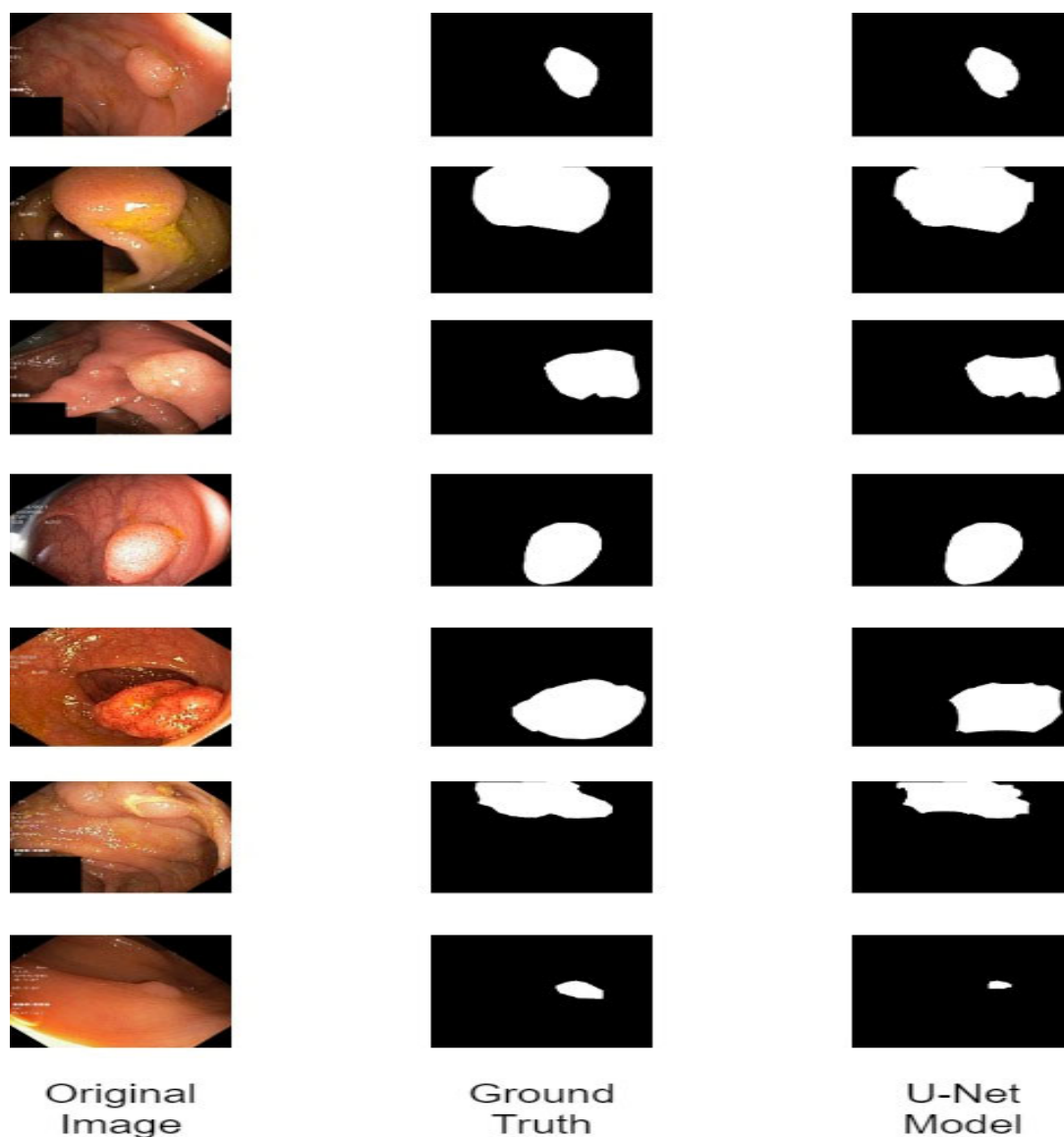
where M is defined as true-positive, N is defined as false-positive, O is defined as false-negative, and P is defined as true-negative.

For the segmentation results, the K-fold cross-validation technique was applied with K fixed to 10, as it is evident from research that when K is 10, the model performs better [40]. After we applied U-Net with ResNet-34 as a backbone model on our dataset, the model achieved a 0.9030 mIOU score, 0.8208 dice score, 0.9435 precision, and 0.8597 recall score. Table 1 compares the quantitative findings based on the Kvasir-Seg dataset using several segmentation methods.

Table 1. Quantitative findings based on the Kvasir-Seg dataset.

Sr. No.	Method	Dice	mIOU	Precision	Recall
1.	ResUNet [45]	0.5144	0.4364	0.7292	0.5041
2.	ColonSegNet [46]	0.7980	0.6980	0.8432	0.8193
3.	HarDNet-MSEG [47]	0.8102	0.7459	0.8652	0.8485
4.	PraNet [48]	0.8142	0.8796	0.9126	0.8453
5.	UNet with ResNet-34	0.8208	0.9030	0.9435	0.8597

Qualitative results of segmentation and localization of polyps based on the Kvasir-Seg dataset are shown in Figure 7. By looking at the ground truth, it can be noticed that the segmentation results generated by UNet with the ResNet-34 model as a background are up to the mark. Furthermore, the results show that the model detected the large diseased area and produced high-quality masks at similar locality but with a slightly different shape. This same segmentation model is applied to all other classes, like ulcer, polyp, ulcerative colitis, and esophagitis, for drawing contours around the diseased area and passing these images for classification.

**Figure 7.** Qualitative findings based on the Kvasir-SEG dataset after applying U-Net with ResNet-34 as the backbone.

4.2. Classification Results

We evaluated the performance of classification on a fine-tuned Xception model using different performance measures, namely, precision, recall, and accuracy. In medical applications, we are more concerned that recall should be high so that no disease case should be treated as normal. Precision and recall are already discussed in the segmentation, and their equations are also shown in (10) and (11); therefore, only accuracy is discussed in this section. Accuracy points out the number of true predictions from total predictions. The equation of accuracy is shown below:

$$Accuracy = \frac{M + P}{M + N + O + P} \quad (12)$$

where M is defined as true-positive, N is defined as false-positive, O is defined as false-negative, and P is defined as true-negative.

Classification results were collected by distributing the dataset into various train-to-test ratios, namely, 80/20, 70/30, and 60/40, and using 10-fold cross-validation. Initially, results are collected based on input images with no contours using 10-fold cross-validation to compare the performance. It is evident from Table 2 that the softmax classifier outperforms other classifiers, with 89.62% precision, 78.25% recall, and 81.06% accuracy. However, quadratic SVM performance cannot be overlooked, as it is near to that of softmax.

Table 2. Performance matrices based on images without contours using 10-fold cross-validation.

	Precision	Recall	Accuracy
Softmax	89.62%	78.25%	81.06%
Linear SVM	77.16%	71.33%	75.94%
Quadratic SVM	88.97%	78.17%	80.68%
Bayesian	77.16%	71.33%	75.94%

On an 80/20 ratio, the best achieved results using the softmax classifier are 87.67% precision, 80.13% recall, and 85.27% accuracy. Results achieved by applying multiple classifiers using our proposed model are shown in Table 3. It is evident from Table 3 that quadratic SVM performance is also satisfactory and near to that of softmax. Moreover, the testing accuracy graph based on the model trained for the 80/20 train-to-test ratio is also shown in Figure 8.

Table 3. Performance matrices based on the 80/20 train-to-test ratio.

	Precision	Recall	Accuracy
Softmax	87.67%	80.13%	85.27%
Linear SVM	81.92%	77.42%	79.19%
Quadratic SVM	86.22%	80.10%	85.02%
Bayesian	80.31%	74.22%	78.94%

The confusion matrix generated based on the 80/20 ratio using the softmax classifier is shown in Figure 9. Looking at the results, it is clear that in the ulcer class, four of the cases are shown as normal, and in the esophagitis class, three of the cases are treated as normal, which is an attentive sign. Moreover, six of the normal cases are treated as disease cases.

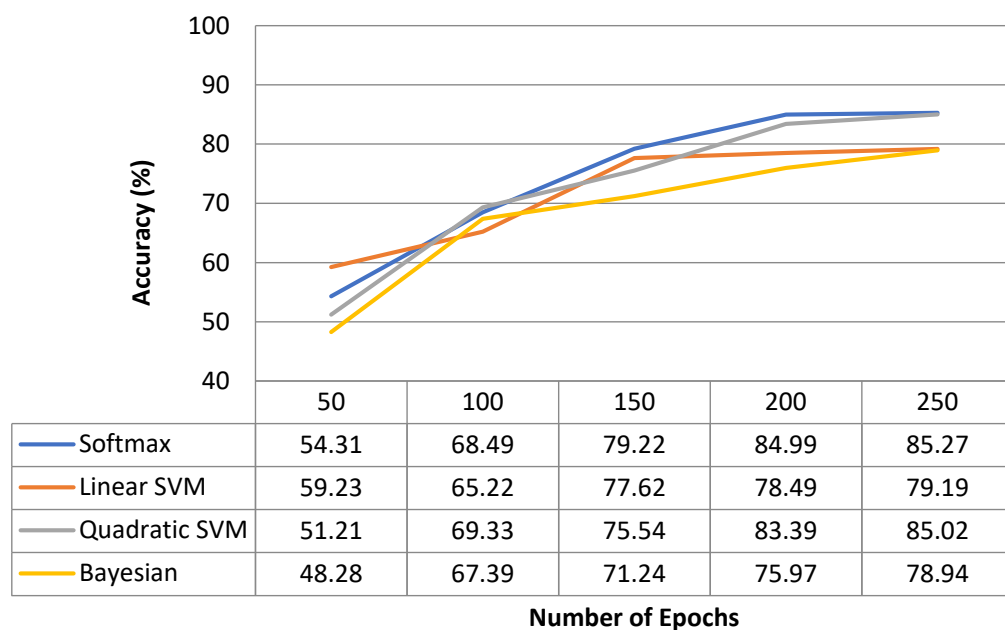


Figure 8. Testing accuracy graph based on the 80/20 train-to-test ratio.

		Predicted Class				
		Polyp	Ulcer	Ulcerative-Colitis	Esophagitis	Normal
Actual Class	Polyp	84	2	6	8	0
	Ulcer	5	82	8	1	4
	Ulcerative-Colitis	6	9	80	5	0
	Esophagitis	3	4	10	80	3
	Normal	1	3	0	2	94

Figure 9. Confusion matrix based on the 80/20 train-to-test ratio using softmax.

On a 70/30 ratio, the best results achieved using the softmax classifier are 96.94% precision, 93.22% recall, and 94.68% accuracy. Results achieved by applying multiple classifiers using our proposed model are shown in Table 4. It is evident from Table 4 that after softmax, Bayesian performance is better than the other classifiers. Moreover, the testing accuracy graph on the model trained with the 70/30 train-to-test ratio is also shown in Figure 10.

Table 4. Performance matrices based on the 70/30 train-to-test ratio.

	Precision	Recall	Accuracy
Softmax	96.94%	93.22%	94.68%
Linear SVM	90.35%	87.01%	87.19%
Quadratic SVM	91.77%	82.01%	86.22%
Bayesian	91.64%	88.35%	88.34%

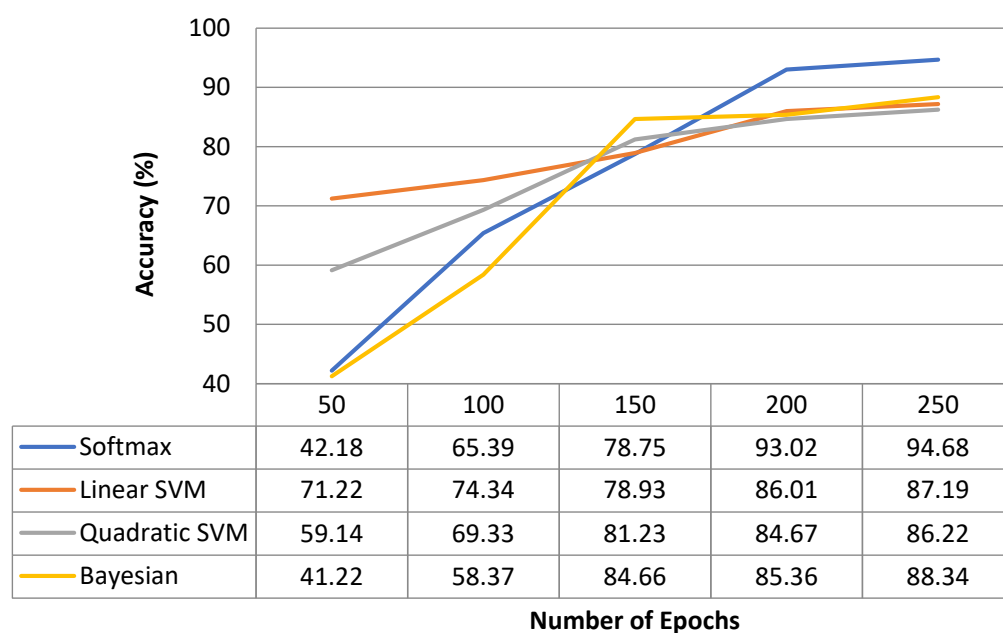


Figure 10. Testing accuracy graph based on the 70/30 train-to-test ratio.

The confusion matrix generated based on the 70/30 ratio using the softmax classifier is shown in Figure 11. Looking at the results, it is clear that while using this ratio, the model performs much better, and only one disease case is treated as normal, which is in the esophagitis class. Moreover, only four of the normal cases are treated as disease cases.

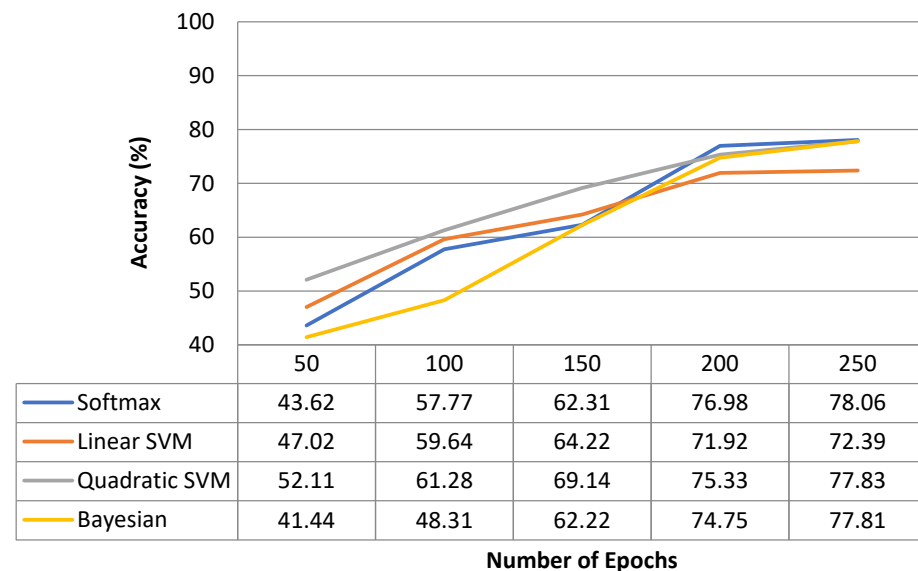
		Predicted Class				
Actual Class		Polyp	Ulcer	Ulcerative-Colitis	Esophagitis	Normal
	Polyp	97	1	2	0	0
	Ulcer	0	98	1	1	0
	Ulcerative-Colitis	1	5	91	3	0
	Esophagitis	0	1	8	90	1
	Normal	1	1	2	0	96

Figure 11. Confusion matrix based on the 70/30 train-to-test ratio using softmax.

On a 60/40 ratio, the best results obtained using the softmax classifier are 82.56% precision, 73.69% recall, and 78.06% accuracy. Results achieved by applying multiple classifiers using our proposed model are shown in Table 5. It is evident from Table 5 that quadratic SVM performance is also satisfactory and near to that of softmax. Moreover, the testing accuracy graph on the model trained with the 60/40 train-to-test ratio is shown in Figure 12.

Table 5. Performance matrices based on the 60/40 train-to-test ratio.

	Precision	Recall	Accuracy
Softmax	82.56%	73.69%	78.06%
Linear SVM	74.22%	67.42%	72.39%
Quadratic SVM	82.18%	72.81%	77.83%
Bayesian	78.16%	71.23%	77.81%

**Figure 12.** Testing accuracy graph based on the 60/40 train-to-test ratio.

The confusion matrix generated based on the 60/40 ratio using the softmax classifier is shown in Figure 13. Looking at the results, it is clear that while using this ratio, model performance worsens, as three polyp cases, six ulcer cases, six ulcerative colitis cases, and ten cases of esophagitis were predicted as non-diseased. Moreover, it is also a great concern that 20 of the normal cases were treated as disease cases. We believe that the behavior of the model worsened as training data were reduced; hence, the model was not properly tuned.

		Predicted Class				
Actual Class		Polyp	Ulcer	Ulcerative-Colitis	Esophagitis	Normal
	Polyp	77	7	9	5	3
	Ulcer	5	81	5	3	6
	Ulcerative-Colitis	0	11	75	8	6
	Esophagitis	1	4	0	85	10
	Normal	4	8	2	6	80

Figure 13. Confusion matrix based on the 60/40 train-to-test ratio using softmax.

For 10-fold cross-validation, the best results achieved using the softmax classifier are 99.68% precision, 96.13% recall, and 98.32% accuracy. Results achieved by applying multiple classifiers using our proposed model are shown in Table 6. It is evident from

Table 6 that the quadratic SVM as well as Bayesian performances are satisfactory, and cannot be ignored. Moreover, the testing accuracy graph based on the model trained with 10-fold cross-validation is depicted in Figure 14.

Table 6. Performance matrices based on 10-fold cross-validation.

	Precision	Recall	Accuracy
Softmax	99.68%	96.13%	98.32%
Linear SVM	91.72%	89.29%	90.07%
Quadratic SVM	99.24%	95.04%	97.64%
Bayesian	97.63%	94.46%	97.28%

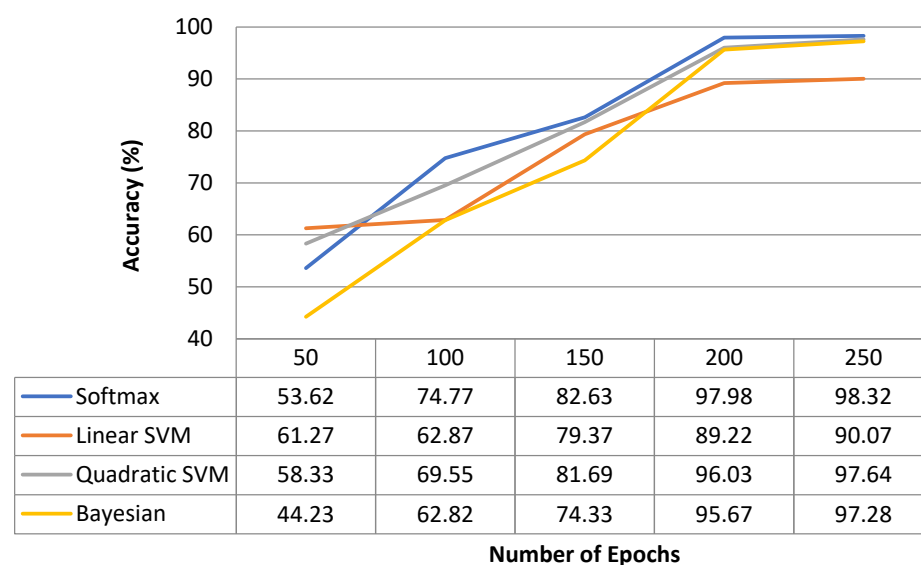


Figure 14. Testing accuracy graph based on 10-fold cross-validation.

The confusion matrix produced by 10-fold cross-validation with the softmax classifier is shown in Figure 15. Looking at the results, it is clear that the model performs significantly better, and no disease case is treated as normal, which is the prime focus in medical applications. Hence, we achieved our desired performance using the proposed methodology. Moreover, only one of the normal cases is treated as a disease case. After analyzing the results, it is understandable that better results are achieved with 10-fold cross-validation, which shows the reliability of our model; hence, we select it as our proposed model.

		Predicted Class				
Actual Class		Polyp	Ulcer	Ulcerative-Colitis	Esophagitis	Normal
	Polyp	100	0	0	0	0
	Ulcer	0	100	0	0	0
	Ulcerative-Colitis	0	3	96	1	0
	Esophagitis	0	1	2	97	0
	Normal	0	0	0	1	99

Figure 15. Confusion matrix based on 10-fold cross-validation using softmax.

5. Discussion

This section focuses on the analysis of the proposed methodology and its effectiveness, along with its limitations. Better segmentation and heat maps contribute towards improved classification accuracy, precision, and recall. The dataset is split into various train-to-test ratios in order to ensure that no bias exists, and that samples are actual representatives of the dataset. If we analyze the results in Tables 3–6, it is clear that when the training data are reduced to 60%, the accuracy is reduced drastically, and the model treated 25 disease cases as normal, which means that the model is not generalized well when the training data are reduced. Moreover, better results on 10-fold cross-validation indicate balance between bias and variance of model. It is also evident from the results that there is a significant improvement in precision and recall rate, along with accuracy, which is also an indication of robustness. Upon analysis of the confusion matrix presented in Figure 15, a clear indication of better performance in the case of diseased data is observed, as no disease case is treated as normal. The proposed framework includes numerous significant steps, and major classification results were improved by using the images with contours, which indicates the significance of this step. The performance improvement between original and contour images can be observed by looking at the results presented in Tables 2 and 6. There is drastic improvement in accuracy of up to 17.26% for images with contours. This step highlights the boundary region of disease in an image, which in turn improves the classification outcomes. Moreover, heat maps also reveal that when the segmentation is performed, the model is more focused on a diseased area. Overall, the performance of the model in terms of false-positive rate (no diseased instance is classified as normal) with the 10-fold cross-validation technique demonstrates the robustness of our proposed methodology.

The proposed methodology outperformed the cutting-edge methods, thus having major contributions; however, there are certain limitations that need to be addressed for future study. For instance, this study does not take into account the contrast and brightness issues of the endoscopic images. Moreover, neither the influence of training several models nor the optimization of features was taken into account in this study, which might lead to better results.

Finally, we also present a comparison with cutting-edge methods. Table 7 shows that the suggested model outperforms the state-of-the-art methods in terms of accuracy.

Table 7. Proposed model comparison with other approaches.

Methods	Accuracy
Logistic and ridge regression [17]	83.3%
CNN-based framework [18]	85.69%
Various classifiers are applied to multiple handcrafted extracted features [13]	93.64%
Modified VGGNet model on preprocessed images [20]	86.6%
Divided images into numerous regions and then applied the modified DenseNet Model [22]	94.03%
Maximizing the characteristics gleaned from two pre-trained models [28]	96.43%
A contrast enhancement approach is suggested with MobileNet-V2 [29]	96.40%
An attention image-based classification is performed and best features selected [3]	98.07%
Proposed Model	98.32%

6. Conclusions

The manual detection and classification of GI diseases is a challenging task; therefore, an automated system is needed for improved results. In this work, we proposed a DL-based architecture to accurately segment and classify GI diseases. The main idea is to

perform localization using an encoder–decoder-based segmentation technique and draw contours around the diseased area of an image. Furthermore, heat maps are generated using Grad-CAM for unsegmented and segmented images to visualize the high-magnitude region within an image. The images with contours are then used for classification using a deep learning-based model. Segmentation performance is evaluated using various performance metrics like dice, mIOU, accuracy, precision, and recall. For segmentation, our proposed model achieved 82.08% dice, 90.30% mIOU, 94.35% precision, and 85.97% recall. For classification, we reported effectiveness in terms of accuracy, precision, and recall rate. The proposed model achieved 98.32% accuracy, 96.13% recall, and 99.68% precision using the softmax classifier. Our findings show that the presented model did not treat any disease case as normal, which is crucial when human life is involved. Although the proposed model achieved better results as compared to the existing state-of-the-art techniques, several interesting questions need to be researched in the future. For instance, the effects of contrast enhancement and illumination variation were not considered in this research. These preprocessing steps will be the focus of future work, as these highlight the region of interest, and may result in improved performance. Furthermore, we plan to assess the performance of the proposed method in diverse domains, such as those mentioned in references [49–53].

Author Contributions: All authors contributed equally to the writing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the Ministry of Trade, Industry, and Energy (MOTIE) and the Korea Institute for Advancement of Technology (KIAT) through the International Cooperative RD program (Project No. P0016038); the MSIT (Ministry of Science and ICT), Republic of Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-RS-2022-00156354) supervised by the IITP (Institute for Information Communications Technology Planning and Evaluation); Institute of Information & communications Technology Planning & Evaluation (IITP) under the metaverse support program to nurture the best talents (IITP-2023-RS-2023-00254529) grant funded by the Korea government(MSIT); and the faculty research fund of Sejong University in 2022.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Research is conducted on publically available dataset.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ling, T.; Wu, L.; Fu, Y.; Xu, Q.; An, P.; Zhang, J.; Hu, S.; Chen, Y.; He, X.; Wang, J.; et al. A deep learning-based system for identifying differentiation status and delineating the margins of early gastric cancer in magnifying narrow-band imaging endoscopy. *Endoscopy* **2021**, *53*, 469–477. [CrossRef]
2. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249.
3. Noor, M.N.; Nazir, M.; Ashraf, I.; Almujaal, N.A.; Aslam, M.; Fizzah Jilani, S. GastroNet: A robust attention-based deep learning and cosine similarity feature selection framework for gastrointestinal disease classification from endoscopic images. *CAAI Trans. Intell. Technol.* **2023**, 1–14. [CrossRef]
4. Available online: <https://www.cancer.net/cancer-types/colorectal-cancer/statistics> (accessed on 20 April 2023).
5. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2015. *CA Cancer J. Clin.* **2015**, *65*, 5–29. [CrossRef]
6. Korkmaz, M.F. Artificial Neural Network by Using HOG Features HOG_LDA_ANN. In Proceedings of the 15th IEEE International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia, 14–16 September 2017; pp. 327–332.
7. Li, S.; Cao, J.; Yao, J.; Zhu, J.; He, X.; Jiang, Q. Adaptive aggregation with self-attention network for gastrointestinal image classification. *IET Image Process* **2022**, *16*, 2384–2397. [CrossRef]
8. Azhari, H.; King, J.; Underwood, F.; Coward, S.; Shah, S.; Ho, G.; Chan, C.; Ng, S.; Kaplan, G. The global incidence of peptic ulcer disease at the turn of the 21st century: A study of the organization for economic co-operation and development (oecd). *Am. J. Gastroenterol.* **2018**, *113*, S682–S684. [CrossRef]
9. Kim, N.H.; Jung, Y.S.; Jeong, W.S.; Yang, H.J.; Park, S.K.; Choi, K.; Park, D.I. Miss rate of colorectal neoplastic polyps and risk factors for missed polyps in consecutive colonoscopies. *Intest. Res.* **2017**, *15*, 411. [CrossRef]

10. Iddan, G.; Meron, G.; Glukhovsky, A.; Swain, P. Wireless capsule endoscopy. *Nature* **2000**, *405*, 417. [[CrossRef](#)]
11. Khan, M.A.; Sarfraz, M.S.; Alhaisoni, M.; Albeshier, A.A.; Wang, S. StomachNet: Optimal deep learning features fusion for stomach abnormalities classification. *IEEE Access* **2020**, *8*, 197969–197981. [[CrossRef](#)]
12. Khan, M.A.; Sharif, M.; Akram, T.; Yasmin, M.; Nayak, R.S. Stomach deformities recognition using rank-based deep features selection. *J. Med. Syst.* **2019**, *43*, 329. [[CrossRef](#)]
13. Yeh, J.Y.; Wu, T.H.; Tsai, W.J. Bleeding and ulcer detection using wireless capsule endoscopy images. *J. Softw. Eng. Appl.* **2014**, *7*, 422.
14. Dewi, A.K.; Novianty, A.; Purboyo, T.W. Stomach disorder detection through the Iris Image using Backpropagation Neural Network. In Proceedings of the 2016 International Conference on Informatics and Computing (ICIC), Mataram, Indonesia, 28–29 October 2016; pp. 192–197.
15. Korkmaz, S.A.; Akcicek, A.; Binol, H.; Korkmaz, M.F. Recognition of the stomach cancer images with probabilistic HOG feature vector histograms by using HOG features. In Proceedings of the 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia, 14–16 September 2017; pp. 339–342.
16. De Groen, P.C. Using artificial intelligence to improve adequacy of inspection in gastrointestinal endoscopy. *Tech. Innov. Gastrointest. Endosc.* **2020**, *22*, 71–79. [[CrossRef](#)]
17. Wong, G.L.-H.; Ma, A.J.; Deng, H.; Ching, J.Y.-L.; Wong, V.W.-S.; Tse, Y.-K.; Yip, T.C.-F.; Lau, L.H.-S.; Liu, H.H.-W.; Leung, C.-M.; et al. Machine learning model to predict recurrent ulcer bleeding in patients with history of idiopathic gastroduodenal ulcer bleeding. *APT—Aliment. Pharmacol. Ther.* **2019**, *49*, 912–918.
18. Wang, S.; Xing, Y.; Zhang, L.; Gao, H.; Zhang, H. Second glance framework (secG): Enhanced ulcer detection with deep learning on a large wireless capsule endoscopy dataset. In Proceedings of the Fourth International Workshop on Pattern Recognition, Nanjing, China, 31 July 2019; pp. 170–176.
19. Majid, A.; Khan, M.A.; Yasmin, M.; Rehman, A.; Yousafzai, A.; Tariq, U. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. *Microsc. Res. Tech.* **2020**, *83*, 562–576. [[CrossRef](#)]
20. Sun, J.Y.; Lee, S.W.; Kang, M.C.; Kim, S.W.; Kim, S.Y.; Ko, S.J. A novel gastric ulcer differentiation system using convolutional neural networks. In Proceedings of the 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), Karlstad, Sweden, 18–21 June 2018; pp. 351–356.
21. Aoki, T.; Yamada, A.; Aoyama, K.; Saito, H.; Tsuboi, A.; Nakada, A.; Niikura, R.; Fujishiro, M.; Oka, S.; Ishihara, S.; et al. Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest. Endosc.* **2019**, *89*, 357–363.
22. Sekuboyina, A.K.; Devarakonda, S.T.; Seelamantula, C.S. A convolutional neural network approach for abnormality detection in wireless capsule endoscopy. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017; pp. 1057–1060.
23. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
24. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
25. Milletari, F.; Navab, N.; Ahmadi, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571.
26. Zhang, Z.; Liu, Q.; Wang, Y. Road extraction by deep residual unet. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 749–753. [[CrossRef](#)]
27. Guo, Y.B.; Matuszewski, B. Giana polyp segmentation with fully convolutional dilation neural networks. In Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Prague, Czech Republic, 25–27 February 2019; pp. 632–641.
28. Alhajlah, M.; Noor, M.N.; Nazir, M.; Mahmood, A.; Ashraf, I.; Karamat, T. Gastrointestinal Diseases Classification Using Deep Transfer Learning and Features Optimization. *Comput. Mater. Contin.* **2023**, *75*, 2227–2245. [[CrossRef](#)]
29. Nouman, N.M.; Nazir, M.; Khan, S.A.; Song, O.-Y.; Ashraf, I. Efficient Gastrointestinal Disease Classification Using Pretrained Deep Convolutional Neural Network. *Electronics* **2023**, *12*, 1557. [[CrossRef](#)]
30. Jha, D.; Smedsrud, P.H.; Riegler, M.; Halvorsen, P.; Lange, T.D.; Johansen, D.; Johansen, H.D. Kvasir-SEG: A Segmented Polyp Dataset. In Proceedings of the MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, 5–8 January 2020; Proceedings, Part II 26. Springer International Publishing: Cham, Switzerland, 2020; Volume 11962, pp. 451–462.
31. Pogorelov, K.; Randel, K.R.; Griwodz, C.; Eskeland, S.L.; de Lange, T.; Johansen, D.; Spampinato, C.; Dang-Nguyen, D.T.; Lux, M.; Schmidt, P.T.; et al. KVASIR: A multi-class image dataset for computer aided gastrointestinal disease detection. In Proceedings of the 8th ACM on Multimedia Systems Conference, Taipei, Taiwan, 20–23 June 2017; pp. 164–169.
32. Borgli, H.; Thambawita, V.; Smedsrud, P.H.; Hicks, S.; Jha, D.; Eskeland, S.L.; Randel, K.R.; Pogorelov, K.; Lux, M.; Nguyen, D.T.D.; et al. Hyper-Kvasir: A Comprehensive Multi-Class Image and Video Dataset for Gastrointestinal Endoscopy. *Sci. Data* **2020**, *7*, 283. [[CrossRef](#)]
33. Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big. Data* **2019**, *6*, 60. [[CrossRef](#)]
34. Ding, Y.; Chen, F.; Zhao, Y.; Wu, Z.; Zhang, C.; Wu, D. A Stacked Multi-Connection Simple Reducing Net for Brain Tumor Segmentation. *IEEE Access* **2020**, *7*, 104011–104024. [[CrossRef](#)]

35. Kaiming, H.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
36. Livne, M.; Rieger, J.; Aydin, O.U.; Taha, A.A.; Akay, E.M.; Kossen, T.; Sobesky, J.; Kelleher, J.D.; Hildebrand, K.; Frey, D.; et al. A U-Net Deep Learning Framework for High Performance Vessel Segmentation in Patients with Cerebrovascular Disease. *Front. Neurosci.* **2019**, *13*, 97. [\[CrossRef\]](#)
37. Bae, K.; Heechang, R.; Hayong, S. Does Adam optimizer keep close to the optimal point? *arXiv* **2019**, arXiv:1911.00289.
38. Kusakunniran, W.; Karnjanapreechakorn, S.; Siriapisith, T.; Borwarnginn, P.; Sutassananon, K.; Tongdee, T.; Saiviroonporn, P. COVID-19 detection and heatmap generation in chest x-ray images. *J. Med. Imaging* **2021**, *8*, 014001. [\[CrossRef\]](#)
39. van der Velden, B.H.M.; Kuijff, J.H.; Gilhuijs, K.G.A.; Viergever, M.A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med. Image Anal.* **2022**, *79*, 102470. [\[CrossRef\]](#)
40. Noor, M.N.; Khan, T.A.; Haneef, F.; Ramay, M.I. Machine Learning Model to Predict Automated Testing Adoption. *Int. J. Softw. Innov.* **2022**, *10*, 1–15. [\[CrossRef\]](#)
41. Noor, M.N.; Nazir, M.; Rehman, S.; Tariq, J. Sketch-Recognition using Pre-Trained Model. In Proceedings of the National Conference on Engineering and Computing Technology, Islamabad, Pakistan, 12–13 June 2021; Volume 8.
42. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
43. Ho, Y.; Wookey, S. The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE Access* **2019**, *8*, 4806–4813. [\[CrossRef\]](#)
44. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 29. [\[CrossRef\]](#)
45. Jha, D.; Smedsrud, P.H.; Riegler, M.A.; Johansen, D.; De Lange, T.; Halvorsen, P.; Johansen, H.D. ResUNet++: An Advanced Architecture for Medical Image Segmentation. In Proceedings of the 2019 IEEE International Symposium on Multimedia (ISM), San Diego, CA, USA, 9–11 December 2019; pp. 225–2255. [\[CrossRef\]](#)
46. Jha, D.; Ali, S.; Tomar, N.K.; Johansen, H.D.; Johansen, D.; Rittscher, J.; Riegler, M.A.; Halvorsen, P. Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. *IEEE Access* **2021**, *9*, 40496–40510. [\[CrossRef\]](#)
47. Huang, C.-H.; Wu, H.-Y.; Lin, Y.-L. Hardnet-mseg: A simple encoder-decoder polyp segmentation neural network that achieves over 0.9 mean dice and 86 fps. *arXiv* **2021**, arXiv:2101.07172.
48. Fan, D.P.; Ji, G.P.; Zhou, T.; Chen, G.; Fu, H.; Shen, J.; Shao, L. Pranet: Parallel reverse attention network for polyp segmentation. In Proceedings of the International Conference on Medical image Computing and Computer-Assisted Intervention, Lima, Peru, 4–8 October 2020.
49. Habib, M.; Ramzan, M.; Khan, S.A. A Deep Learning and Handcrafted Based Computationally Intelligent Technique for Effective COVID-19 Detection from X-ray/CT-scan Imaging. *J. Grid Comput.* **2022**, *20*, 23. [\[CrossRef\]](#)
50. Ramzan, M.; Habib, M.; Khan, S.A. Secure and efficient privacy protection system for medical records. *Sustain. Comput. Inform. Syst.* **2022**, *35*, 100717. [\[CrossRef\]](#)
51. Masmoudi, Y.; Ramzan, M.; Khan, S.A.; Habib, M. Optimal feature extraction and ulcer classification from WCE image data using deep learning. *Soft Comput.* **2022**, *26*, 7979–7992. [\[CrossRef\]](#)
52. Riaz, A.; Riaz, N.; Mahmood, A.; Khan, S.A.; Mahmood, I.; Almutiry, O.; Dhahri, H. ExpressionHash: Securing telecare medical information systems using biotransforming. *Comput. Mater. Contin.* **2021**, *67*, 2747–2764. [\[CrossRef\]](#)
53. Hussain, A.; Alawairdhi, M.; Alazemi, F.; Khan, S.A.; Ramzan, M. A Hybrid Approach for the Lung (s) Nodule Detection Using the Deformable Model and Distance Transform. *Intell. Autom. Soft Comput.* **2020**, *26*, 857–871. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.