Al-Azhar Engineering 16th International Conference

Vol. 19, No. 72, July 2024, 333 - 350

DIAGNOSIS OF GASTROINTESTINAL CANCER METASTASIS WITH DEEP LEARNING TECHNIQUE

K. Fahmy[1,2*], M. Zorkany[2], Abd El-Hady Ammar[1,2]

[1] Communication and Electronics Department, Faculty of Engineering, Al-Azhar University, Nasr City, 11884, Cairo, Egypt

[2] Department of Electronics, National Telecommunications Institute, Nasr City, 11884, Cairo, Egypt

***Correspondence:** khadega.al_sayed@yahoo.com

Citation:

K. Fahmy, M. Zorkany and A. Ammar, "Diagnosis of gastrointestinal cancer metastasis with deep learning technique", Journal of Al-Azhar University Engineering Sector, vol. 19, pp. 333 - 350, 2024.

Received: 05 March 2024

Revised: 18 May 2024

Accepted : 30 May 2024

Dol:10.21608/auej.2024.282660.1646

ABSTRACT

Gastrointestinal cancer is a leading cause of cancer-related deaths globally, prompting significant research into using artificial intelligence (AI) for detection. Researchers have been exploring AI applications in this field since the 1960s, leveraging its ability to handle repetitive tasks and complex computations. In Phase I of this study, various AI models, including basic CNN and more advanced ones like vgg16, Alex, and DenseNet121, were employed to diagnose gastrointestinal cancer using datasets comprising images of benign tumors and malignancies from patients. However, accuracy rates with conventional techniques were found to be insufficient. Thus, Phase II focused on refining the DenseNet121 model, leading to improved accuracy, sensitivity, and specificity. The modified model demonstrated enhanced diagnostic performance, albeit with slightly longer processing times, compared to existing approaches.

KEYWORDS: Artificial Intelligence (AI), Deep Learning, Convolutional Neural Networks (CNN),Diagnosis, Gastrointestinally cancer.

تشخيص ورم خبيث في سرطان الجهاز الهضمي باستخدام تقنية التعلم العميق

خديجه السيد فهمي[1,2*], محمد الزرقاني[2], عبدالهادي عمار[1,2]

[1]قسم الأتصالات والألكترونيات، كلية الهندسة، جامعة الأزهر، مدينة نصر، 11884 القاهرة، مصر

[2]قسم الألكترونيات، المعهد القومي للأتصالات، مدينة نصر، 11884 القاهرة، مصر

*البريد الالكتروني للباحث الرئيسي: khadega.al_sayed@yahoo.com

الملخص

يعد سرطان الجهاز الهضمي سببًا رئيسيًا للوفيات المرتبطة بالسرطان على مستوى العالم، مما دفع إلى إجراء أبحاث مهمة حول استخدام الذكاء الاصطناعي (AI) للكشف عنه. لقد ظل الباحثون يستكشفون تطبيقات الذكاء الاصطناعي في هذا المجال منذ الستينيات، مستفيدين من قدرته على التعامل مع المهام المتكررة والحسابات المعقدة. في المرحلة الأولى من هذه الدراسة، تم استخدام نماذج مختلفة للذكاء الاصطناعي، بما في ذلك CNN الأساسية والنماذج الأكثر تقدمًا مثل vgg16 وAlex وDenseNet121، لتشخيص سرطان الجهاز الهضمي باستخدام مجموعات بيانات تشتمل على صور للأورام الحميدة والأورام الخبيثة من المرضى. ومع ذلك، فقد وجد أن معدلات الدقة مع التقنيات التقليدية غير كافية. وهكذا، ركزت المرحلة الثانية على تحسين نموذج DenseNet121، مما أدى إلى تحسين الدقة والحساسية والخصوصية. أظهر النموذج المعدل أداءً تشخيصيًا مُحسّنًا، وإن كان مع أوقات معالجة أطول قليلاً، مقارنةً بالطرق الحالية.

الكلمات المفتاحية : الذكاء الاصطناعي (AI)، التعلم العميق، الشبكات العصبية التلافيفية (CNN)، التشخيص، سرطان الرئة.

1. INTRODUCTION

Digestive system cancer encompasses cancers affecting organs like the stomach, esophagus, gallbladder, pancreas, intestines, rectum, and anus, each presenting unique symptoms such as difficulties in swallowing, bowel movements, bleeding, etc. Diagnosis typically involves endoscopy and tissue analysis. Treatment and prognosis depend on the type, location, and spread of the tumor.

With the increasing integration of CAD systems and AI into clinical practice, physicians will need to grasp the basics of AI's functionality and potential, which may appear as intriguing and enigmatic to the uninitiated as they are fascinating [1]. In recent years, there has been a notable rise in the adoption of artificial intelligence within the medical sector [2, 3].

Substantial research efforts are underway to explore various AI applications aimed at improving diagnostic accuracy in clinical settings. Within gastrointestinal endoscopy, AI is being leveraged for numerous purposes, notably in the detection and classification of dysplastic and neoplastic lesions [4, 5].

Even for proficient physicians, accurately interpreting disease entities or lesions can pose significant challenges. The growing demand for computer-aided diagnosis stems from artificial intelligence's exceptional diagnostic capabilities within particular areas. While research driven by preclinical and engineering efforts in AI for gastrointestinal endoscopy is prevalent, recent publications also highlight real-world clinical investigations [6].

Doctors often struggle to grasp the technical aspects of artificial intelligence (AI) and the diverse CAD and machine learning (ML) techniques encompassed by this umbrella term. It's crucial for endoscopists to comprehend the core technological and clinical implications of AI, as it will significantly impact medicine, particularly gastrointestinal endoscopy. Deep learning methods and image processing are being employed to detect gastrointestinal cancer, aiming to enhance accuracy in identifying tumors and determining their shape, size, and location.

Swift detection not only saves time but also enables early treatment initiation. This project focuses on classifying tumors as malignant or benign through preprocessing (noise removal), post-processing (segmentation), and classification algorithms. Malignant tumors, characterized by uncontrollably proliferating cells that can invade nearby tissues, contrast with benign tumors, which do not spread or pose cancer risks. The primary objective is to explore various approaches to gastrointestinal cancer diagnosis, utilizing 3D imaging to identify tumors.

While CT scans are typically used for cancer identification, the sheer volume can challenge prompt and effective diagnosis by medical professionals or radiologists. However, CAD leveraging technological advancements streamlines this process. Various imaging modalities, such as CT, PET,

X-rays, and MRI, have been utilized in deep learning research, with endoscopy emerging as a promising tool despite its historical absence in gastrointestinal cancer treatment.

While endoscopic images offer insights into certain cancer features, the suggested research underscores their potential effectiveness. CNN, a popular DL technique for image classification, comes in multiple versions like AlexNet, VGG, DenseNet, and ResNet, each with distinct layer structures and internal configurations. Our joint effort between physicians and engineers can be encapsulated as follows:

1. Employing a range of standalone CNN models for diagnosing gastrointestinal cancer.
2. Developing a customized CNN model tailored specifically for gastrointestinal cancer diagnosis to improve upon existing models.
3. Conducting performance validation of our proposed model structure by comparing it to several other CNN designs.

The subsequent sections are structured to address these contributions: Section II presents the pertinent literature. Section III delineates the suggested criteria and their evaluation for the chosen CNN types within the proposed Multi-model architecture. Research findings, along with recommendations for further exploration, are detailed in Section IV. Section V delves into the assessment and findings.

2. RELATED WORK

Theoretical physicist predicts that computers equipped with artificial intelligence (AI) will surpass human capabilities within the next century, emphasizing the importance of aligning their objectives with those of humans. The potential applications of artificial intelligence in medicine are garnering increasing attention as a result.

Current developments in deep learning-based multi-omics data processing and their use in illness prediction are thoroughly assessed in this study. The existing obstacles in this subject are listed, and the importance of developing deep learning techniques and enhancing their use to overcome these obstacles is discussed [7].

The article reviews systematic reviews, meta-analyses, randomized controlled trials, and original research on the performance of AI in diagnosing both malignant and benign esophageal and gastric diseases, while also discussing key characteristics of AI [8].

Large, controlled trials in real-time settings are needed to evaluate AI's role in clinical practice. This review highlights key AI applications in GI diseases, noting advantages, limitations, and future development considerations [9].

Key aspects of various neural network architectures used to evaluate gastrointestinal conditions, focusing on tasks like lesion detection and characterization (distinguishing between benign and malignant lesions in the esophagus, stomach, and colon), are examined in this review. An overview of recent achievements and future prospects in deep learning methods for analyzing radiology, endoscopy, and histologic whole-slide images of the gastrointestinal tract is provided [10].

It was said in their paper that AI is utilized in various endoscopic procedures such as esophagogastroduodenoscopy, capsule endoscopy, and colonoscopy. Endoscopic physicians are aided by AI in improving diagnosis rates, reducing missed diagnoses, enhancing endoscopy quality, assessing disease severity, and increasing efficiency. However, challenges remain due to the diversity, susceptibility, and imaging specificity of gastrointestinal endoscopic images. More large-scale, high-quality, multicenter prospective studies are needed to explore AI's clinical applicability, along with considerations of ethical issues [11].

An assessment and synopsis of the utility of radiomics in prognosticating treatment responses in gastrointestinal tumor patients were presented [12]. Various machine-learning methods for real-time polyp detection during live colonoscopies were outlined and evaluated [13].

Frame-by-frame detection and pixel-by-pixel localization techniques integrating deep learning with manual approaches were proposed. The feasibility of analyzing multimedia content in clinical environments, with the potential for workflow improvements and increased detection rates for medical practitioners, was indicated by findings from the study.

A review examining the efficacy of AI in detecting both benign and malignant esophageal and gastric disorders was conducted. Original research articles, meta-analyses, and systematic reviews were encompassed in their analysis, alongside discussions on the basics of AI [14].

The trade-offs associated with systems employing multiple binary classifiers versus a single multiclass classification, specifically focusing on multi-class image classification, were highlighted in their study [15]. Several modern neural network architectures, including DenseNet, Inception v3, Inception ResNet v2, Xception, NASNet, and MobileNet, were examined to assess these classifiers. A comprehensive comparison of various aspects of these architectures' performance, including classification speed and accuracy measurements during both training and testing phases, was conducted by employing both classification approaches.

Current challenges in developing computer-based digital assistants were addressed in their article. Examples of proposed tools leveraging various technologies were presented, existing difficulties were identified, and recommendations for the future creation and evaluation of CAD systems were provided [16].

The persistent global issue of inconsistencies in endoscopy reports was highlighted by them. It was pointed out that artificial intelligence (AI) techniques, particularly neural networks, have shown promise in automatically identifying lesions within the gastrointestinal tract mucous membrane, offering the potential to mitigate human variability in disease diagnosis. However, deep neural networks are often perceived as "black boxes," as their decision-making processes are not always transparent to end users. Hicks et al. aim to address this opacity and leverage their insights to develop a device capable of generating standardized endoscopic reports automatically [17].

A comprehensive AI-based framework designed to extract both temporal and spatial data from endoscopic videos simultaneously was introduced by them, aiming to enhance classification performance across various gastrointestinal disorders. Their approach involved coupling two separate residual networks and a long memory model in cascade mode to extract the required data. Experiments were conducted using a pooled dataset consisting of one of the largest endoscopic video collections, comprising 52,471 frames [18].

A deep convolutional neural network is detailed in their paper. This network is engineered to analyze images captured by digital endoscopes, automatically detecting disorders from gastrointestinal (GI) test footage [19].

In their study, a two-input network, GFD Faster R-CNN, was proposed to tackle the challenges associated with identifying esophageal abnormalities. Initially, a fractal Gabor image was generated from the original endoscopic picture by applying multiple Gabor filter responses, capturing various orientations and scales to enhance the fractal texture data within the image. Then, DenseNet was incorporated as the underlying network to independently extract features from both the original endoscopic image and the generated GF image. The number of trained parameters was effectively reduced by DenseNet while maintaining network accuracy and facilitating maximum information flow. During the ROI pooling stage in Faster R-CNN, the features extracted from the GF and endoscopic images were merged via binary fusion, resulting in a comprehensive feature representation that ultimately enhanced detection performance [20].

Comprehensive evaluations of five distinct machine learning models utilizing global features integrated into deep neural networks were conducted. These models were designed to identify 16 major categories of gastroenteritis, encompassing pathological results, anatomical landmarks, polyp eliminations, and normal findings from standard gastrointestinal inspection equipment photos [21].

In this work, the GooleNet model was refined to better target medical pathological images, guaranteeing diagnosis accuracy while also greatly reducing computing load. The improved model lowers the chance of misdiagnosis by improving the accuracy of stomach cancer diagnoses [22].

3. SUGGESTED CNN MULTI-MODEL- BASIC STRUCTURE

The main image classification algorithms utilized in automatic gastrointestinal cancer detection systems encompass those tailored for X-ray, CT, positron emission tomography, MRI, and endoscopic images. Deep learning techniques, which are rooted in neural networks and form a fundamental aspect of artificial intelligence, are considered superior to other image classification algorithms. Typically, the conventional CNN architecture comprises three primary levels: convolution, pooling, and fully connected layers. CNNs are categorized into various types based on their internal organization, with the most prominent ones in image classification being DenseNet, DenseNet 121, VGG-16, VGG-19, AlexNet, and ResNet 101.

Recent studies have employed various CNN architectures to detect gastrointestinal tract organ cancers. However, while certain CNN types excel with specific image types, they may struggle with others of similar quality. For instance, a CNN model optimized for X-ray images may not perform as effectively with CT scans, and vice versa. Consequently, much of the research in this domain tends to utilize a particular CNN type tailored to a specific medical imaging modality. Consequently, there's a paucity of studies exploring gastrointestinal cancer detection across diverse medical imaging modalities.

Although certain CNN types exhibit strong performance with specific medical image types, these favorable outcomes are often constrained to particular datasets. CNN performance tends to deteriorate when confronted with different datasets. Hence, the prevailing literature in this domain frequently showcases positive results achieved by a specific CNN type applied to a singular class of medical images and a specific dataset. Despite the uniform CNN structure, these disparities stem from the varied structural configurations inherent in each CNN model. This underscores that each CNN variant offers advantages and can proficiently extract features from certain images, but optimal performance is contingent upon compatibility with specific image sets and datasets.

Consequently, by leveraging endoscopic images from various models, each model contributes to the detection of digestive system cancer across all its organs. Within this framework, the optimal model, as determined in the testing phase, is utilized to ascertain the presence of gastrointestinal cancer and its type. This approach enhances the efficiency of the integrated model and facilitates performance evaluations, as illustrated in Figures 1, and 2.

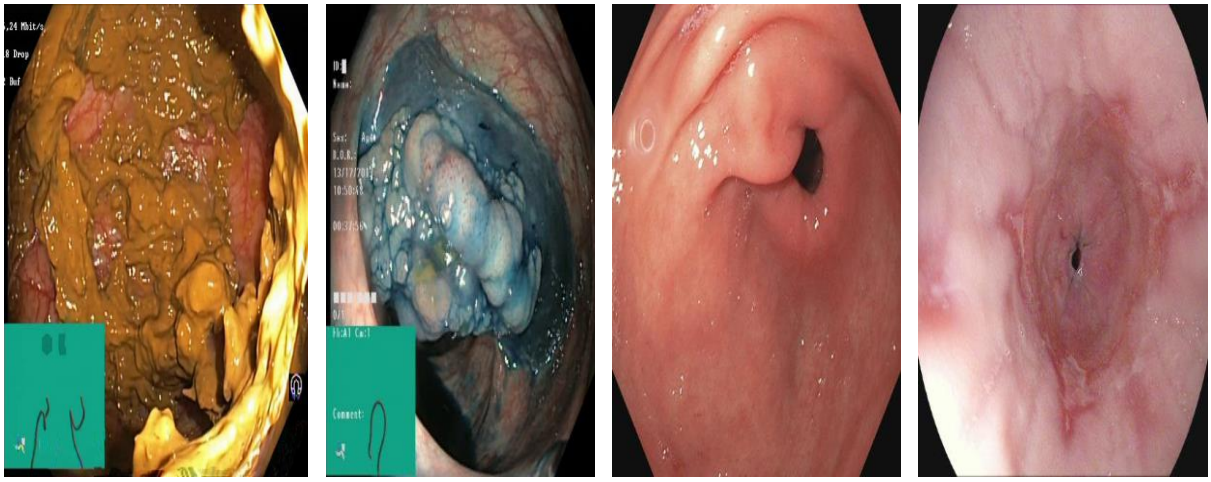


Fig. 1: Pictures of endoscopy for various diagnoses of gastrointestinal tumors.

The proposed automated detection model for gastrointestinal cancer will adopt a multi-model CNN architecture, integrating diverse CNN models with endoscopic images to identify the type of gastrointestinal cancer. The primary objective is to reduce error rates and enhance validation accuracy by leveraging multiple models or algorithms for diagnosing gastrointestinal cancer using endoscopic images. The integration of multi-model systems will entail variations in circumstances.

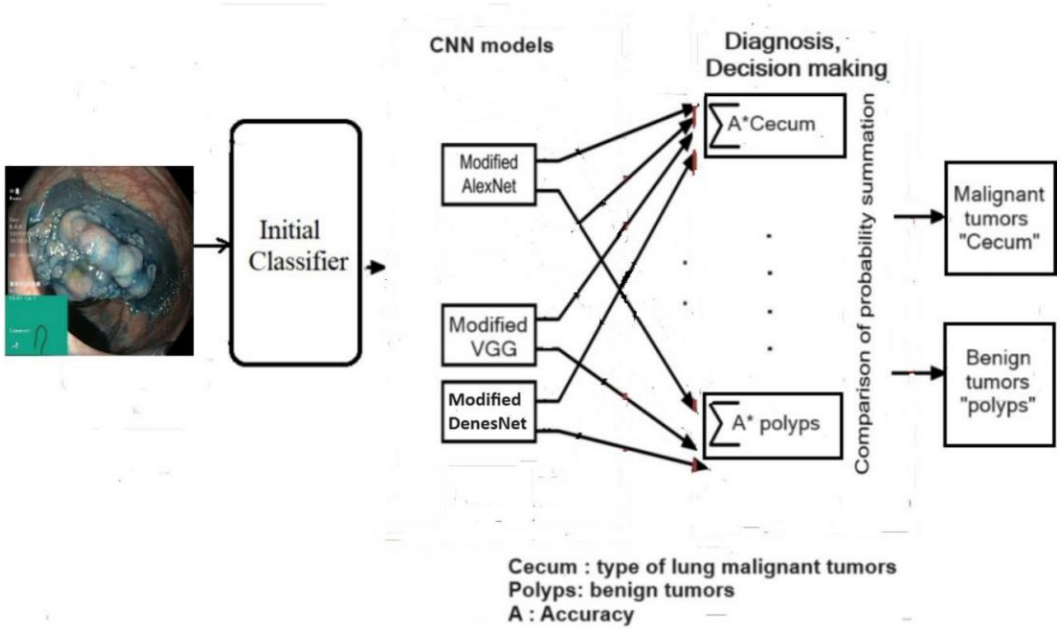


Fig. 2: Diagnosis and decision-making diffusion.

3.1. Data from Multimodal Imaging

Multimodal learning relies on diverse sources of relevant data, with image classification models like CNN benefiting from various imaging datasets owing to multi-modal image data classification technology. In our proposal, we will utilize medical endoscopic images, leveraging their precision in diagnosis compared to X-rays used for detecting infections, tumors, diseases, and bone fractures. Our primary data sources encompass common dataset types listed in Table 1. By

employing various endoscopic datasets, the model's learning accuracy can be enhanced, enabling the extraction of more features.

For our work, we utilize the Gastrointestinal Cancer Dataset, which includes 12 classes extracted from recent Medical Dataset research papers and a Kaggle dataset. This dataset comprises 9717 photos, each in jpg format and sized (224,224). It is a recent and sizable dataset, sourced from a large study detailed in reference [23,24,25], forming the cornerstone of our research. The dataset consists of samples representing both normal and abnormal instances, including various forms of gastrointestinal cancer, facilitating the development of machine learning techniques for diagnosis. It is divided into training, testing, and validation sets, with 70%, 10%, and 20% of the data respectively. Notably, the dataset comprises 14% normal instances and 86% abnormal instances.

Table 1: Dataset summary that were used.

NO	DATA SET	USED
	Image Type	**No. of Image**
1	bbps-0-1	(646 files)
2	bbps-2- 3	(1148 files)
3	Cecum	(1009 files)
4	dyed-lifted-polyps	(1002 files)
5	dyed-resection-margins	(989 files)
6	esophagitis-a	(403 files)
7	Polyps	(1028 files)
8	Pylorus	(999 files)
9	retroflex-rectum	(391 files)
10	Retroflex-stomach	(764 files)
11	ulcerative-colitis-grade-1-2	(406 files)
12	z-line	(100 files)

3.2. CNN Structure Based on Multi Models

In this study, several steps were undertaken to develop the proposed model for diagnosing gastrointestinal cancer using endoscopic images.

- We utilized a basic CNN model, known as the simple initial classifier, to distinguish between different types of endoscopic images.
- We selected the top CNN deep learning technique types for image classification.
- Subsequently, we employed standard datasets outlined in Table 1, focusing on models that exhibited superior performance in identifying gastrointestinal cancer disease.
- Some of these models underwent modifications to improve diagnostic outcomes, as elaborated later.
- The final diagnosis of this condition was determined using the decision-making diffusion methodology, which involved aggregating the outputs of both the original and modified CNN models.

The CNN models that were chosen and utilized to construct our suggested model are VGG, AlexNet, and ResNet 121. In order to get better outcomes with particular image or dataset categories, some of these algorithms have also undergone modifications. Additionally, we changed a few CNN varieties, and customized CNN and modified AlexNet, DenseNet, and VGG in the ways listed below:

3.2.1. Customized CNN

Multiple convolutional, pooling, and fully connected neural network layers that are intended to learn in an automatic and flexible manner feature hierarchies make up a customized CNN. The customized CNN model is displayed in Fig. 4. It employs four convolution layers for feature extraction, three max-pooling layers for feature reduction, and three fully connected neural network layers to map the extracted features into the final SoftMax output layer for classification. All convolution layers employed tiny kernel filters with a 3x3 size. When compared to large kernel size filters, using this tiny size filter will yield higher accuracy in various image classification applications. Additionally, we include five dropout layers: three in the fully connected layers section after the first and second layers, and three after the third, fourth, and fifth convolution layers. The activation function is Relu utilized in the fully interconnected and convolution layers, whilst SoftMax is used in the output layer. A straightforward regularization technique called dropout is utilized to stop the deep learning neural networks from over fitting. The percentage of connections, which stands for the CNN's parameters or weights, is randomly removed from the neural network throughout each training cycle. In actuality, each CNN update, or iteration, appears to be a distinct "view" within the component layer. CNN's convolution and fully linked layers both support dropout.

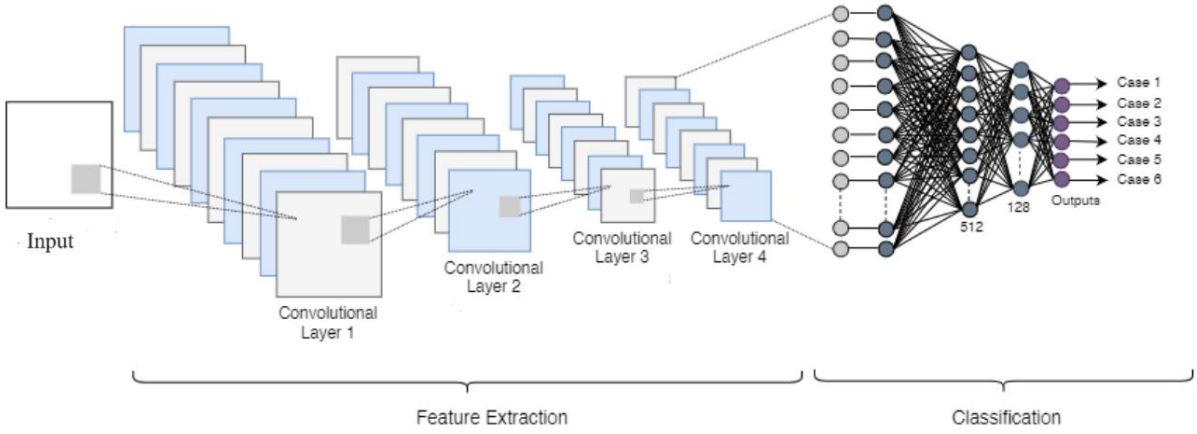
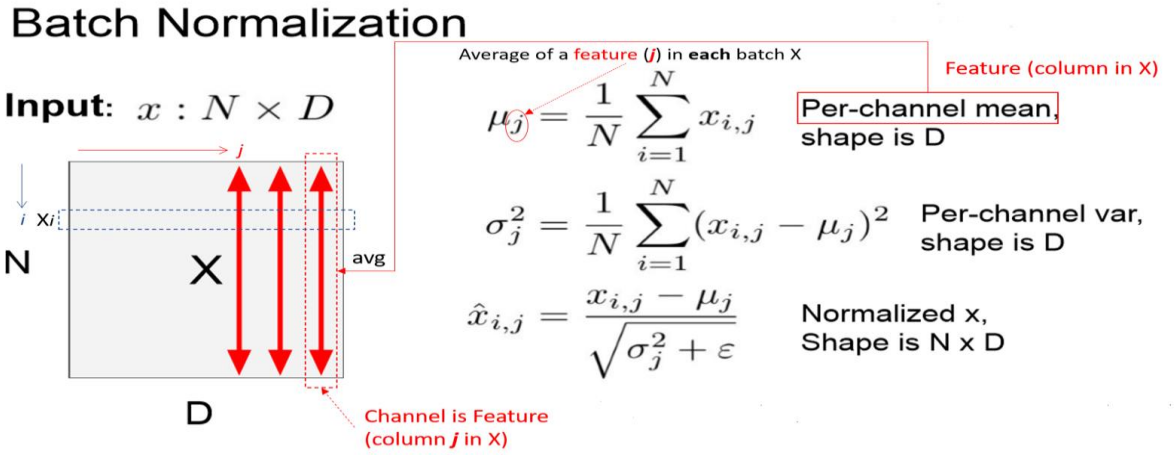


Fig. 3: Customized CNN structure.

3.2.2. Modified AlexNet

Introduced in 2012, AlexNet uses a deep CNN design that made significant progress in computer vision tasks including image categorization. There are three fully linked layers and five convolutional layers totaling eight layers in this structure. Subsequent layers employ smaller receptive fields to catch increasingly sophisticated and abstract elements, while the initial convolutional layer has a large receptive field to record low-level features like edges and textures. (ReLU) activation functions were first effectively used by lexNet, the first deep network, and are now widely used in deep learning. Additionally, dropout regularization was employed to keep training from overfitting. The triumph of AlexNet on the ImageNet dataset, encompassing more than a million photos, showcased deep neural networks' capacity for image identification assignments and cleared the path for more progressions in the domain of computer vision.

To improve performance, this model incorporates several AlexNet model adjustments, such as batch normalization and dropout. Three dropout layers having a rate of 0.5 were inserted between three fully connected layers, 64 layers of batch normalization were added to each layer, and a layer of batch normalization was placed at the model's input. Additionally, rather than using the "SGD" optimizer in this model, we utilized the "Adam" optimizer (lr=0.0001, momentum=0.9). Batch normalization is utilized to train the networks in this modified version of the AlexNet CNN model. The following is the pseudo code for the batch normalization transform steps:



By using this method to change the AlexNet model, the learning process is stabilized and training time and iterations are decreased.

3.2.3. Modified VGG16

Feature extraction is necessary to recognize patterns that differentiate between a healthy GI tract and a cancerous GI tract. Pre-trained models are more accurate and efficient than newly developed CNN architectures built from scratch, especially when it comes to classification. Since the pre-trained VGG16 model is thought to be an advanced version of the Alex Net neural network, we looked into VGG16 for feature extraction in our investigation. The Image Net dataset served as pre-training for the VGG16 transfer learning algorithms, which were used to extract features.

To VGG 16, we introduced batch normalization, step-by-step dropouts, and the use of an Adam optimizer. As with the prior method, applying this series of normalization can increase the stability of a CNN network. In addition, the dropout layers were included to stop DL neural networks from over fitting. This will be confirmed in the results section.

3.2.4. Modified DenseNet 121

Because each layer in the DenseNet structure receives feature maps from the anterior levels, the architecture may be made thinner, and more compact, with fewer channels, so the error signal can more quickly spread to the earlier layers. According to recent studies, DenseNet outperforms traditional CNN techniques in picture categorization. Thus, the DenseNet model—one of the models chosen for gastrointestinal cancer diagnosis based on multimodal images—will be the foundation of the suggested multi-model technique. CNN is the basis for the DenseNet-121 model [22], which represents the classification process. The primary components of the suggested model were the four implanted layers, as seen in Fig. 4 displays the DenseNet 121, and an input image is used to launch the CNN model. The feature map is produced by applying the filters to the input image. Equation (1) explains how the first layer creates feature maps like $x0$, …, $xl-1$ from the subsequent levels.

$$X_l = H_l([X_0, X_1, \dots, X_{l-1}]) \qquad (1)$$

Where the feature maps that have been concatenated and created with layers ranging from 0,..., $l-1$ are known as [$X0$,$X1$, …, $Xl-1$]. When the feature maps are fed, an activation function such as ReLU is employed to raise the non-linearity at the layer used for pooling. After the first layer is assessed using Eq. (2), the Hl function produces the k level-based mapping features.

$$H_l = K_0 + K \times (L - 1) \qquad (2)$$

$K0$ is the total number of channels in the input layer, as determined by the aforementioned Eq. (2). The hyper parameter with a higher rate of growth in the network is called K. Every layer adds feature maps, each of which has its own state. Uses for the dot product include calculate the weights produced by the neuron, and the volume is taken into consideration while computing the input images.

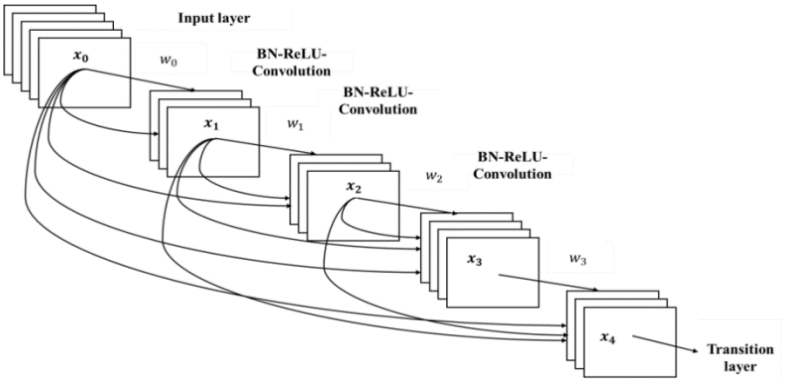


Fig. 4: DenseNet 121 model [26].

We added batch normalization, dropouts at every step, and the usage of an Adam optimizer to denseNet 121. Applying this batch normalization can improve a CNN network's stability, just like the previous technique did.

3.3. Diagnosis and Decision Making

Gastrointestinal cancer diagnosis —whether it is positive or negative and if negative, what kind—is the cornerstone of the suggested paradigm. It verifies the accuracy of each unique standard and customized model. Additionally, it takes into account each standard model's probability results for gastrointestinal cancer, either positive or negative, as shown in Fig. 1.

Let **Pi** be the probability of a positive gastrointestinal cancer outcome, and **Nai** is one of the different probabilities of a negative gastrointestinal cancer result. Assume that **Ai** represents the validation accuracy for each standard model.

Where (i) stands for each of the eight models that were previously specified, (m) stands for each unique model number (for example, the eighth model in our architecture that has been evaluated), and (a) is Type-negative gastrointestinal cancer. In the parts that follow. The following equations illustrate how the validation accuracy for each prior model has been multiplied by the corresponding positive and negative results probability in order to arrive at a judgment or diagnosis:

$$\frac{1}{m}\sum_{i}^{m} A_i * P_i \quad (3)$$

$$\frac{1}{m}\sum_{i}^{m} A_i * N_{ai} \quad (4)$$

The results of equation 3 were compared to the results of other equations to determine which is superior for diagnosis and decision-making. gastrointestinal cancer (+ve) will be the ultimate option if the output of equation 3 is higher than the other output, and vice versa with the negative decision, gastrointestinal cancer (-ve).

3.4. Performance Assessment

We looked at the most well-known performance metrics used in deep learning to assess the effectiveness of the various tested models. The following formula was used to determine those indicators: loss function, accuracy, recall, and precision.

Commonly employed in CNN as the gradient descent function, the loss function for cross-entropy is as follows:

$$\text{Loss} = -\frac{1}{n}\sum_{n=1}^{\infty}[(td\ \ln od) + (1 - td)\ln(1 - od)] \quad (5)$$

Where **od** is the calculated value, **td** is the necessary output value, and **n** is the output size.

The common measures for the right model are the following: precision, recall, and accuracy (i.e., recognition rate).

Forecasting as follows in this categorization domain:

$$Accuracy = \frac{TP+TN}{P+N} \quad (6)$$

$$Recall = \frac{TP}{P} \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

Where P stands for positive images (gastrointestinal cancer), N for negative images (non-gastrointestinal cancer), TP for true positive, TN for true negative, and FP for false positive rate.

4. OUTCOME OF TRAINING AND TESTING

In this paper, we used various CNN algorithms on various gastrointestinal cancer datasets. The goal of the output models was to identify the patient from his endoscopic pictures. To create the suggested integrated model, we created several models and employed various datasets. In addition to multimodal imaging, our model took into account the multi-model CNN component of the gastrointestinal cancer issue. The following will be our main performance focus.

4.1. Custom CNN

As previously indicated in the suggested model, various numbers of convolution and pooling layers were utilized in the construction of this tailored detection model for gastrointestinal cancer. This model's obtained accuracy is 57%. The model's training and validation losses were s: Loss: 1.0577 - Accuracy: 0.5774 –Test loss: 1.057 - test Accuracy: 0.577.

This model's accuracy was increased to 62% by adjusting the number of convolution, pooling, and dropout to 0.5 instead of 0.2 while increasing the batch size from 32 to 64. Modified Model 1 assessments: Loss: 1.0691 - Accuracy: 0.6290 –Test loss: 1.069 - test accuracy: 0.629.

4.2. Modified AlexNet

The accuracy rate climbed to 84% while utilizing the pre-installed AlexNet technology model, which is explained. We observed that the findings differ significantly from the prior model (Custom CNN). For instance, the simulation results for this model, which take into consider the curves used for training and validation in Fig. 5, are displayed in the following images. Where model 2 evaluations: Test accuracy: 0.841; test loss: 0.354; accuracy: 0.8419. Loss: 0.3549.

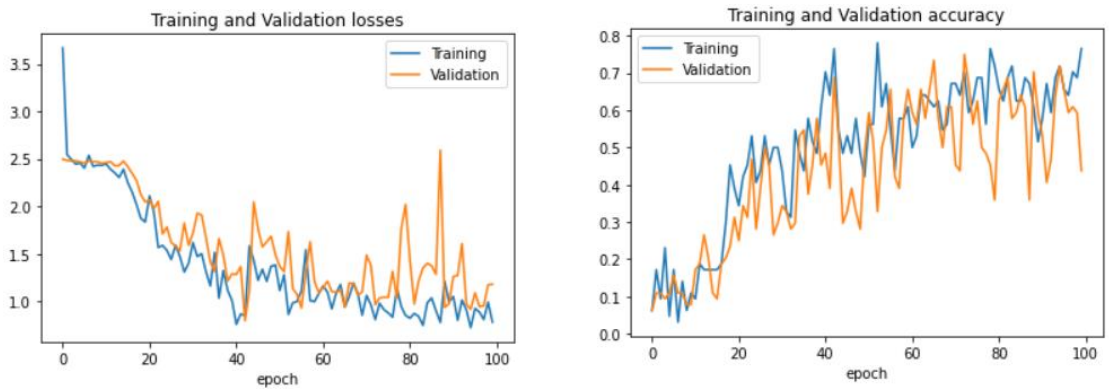


Fig. 5: AlexNet Model Losses and Accuracy

The modification we mentioned previously to this model did not significantly improve accuracy, as it was observed that the accuracy increased by only 1% to become 87%. The results are shown in the following figure 6. Where Test accuracy: 0.859; test loss: 0.325; accuracy: 0.8591. Loss: 0.3251.

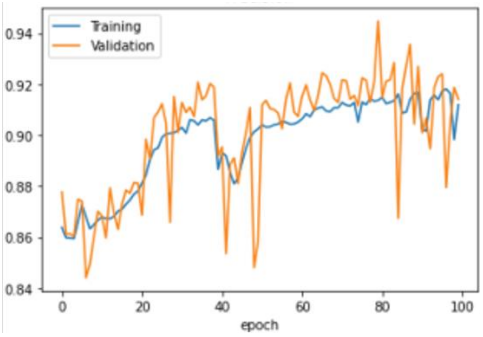


Fig.6: AlexNet modified Model Accuracy

4.3. Modified VGG16

From the simulation results performed based on different models, the lowest performance was achieved for the models in VGG the training accuracy does not exceed 15%, while the obtained test accuracy is close to 11%. The curves show VGG the instability of the model. Highlight the figure 7. These results are for sample data sets. Evaluations of model: Test accuracy: 0.113; Test loss: 2.406; Accuracy: 0.1140 Loss: 2.4069.

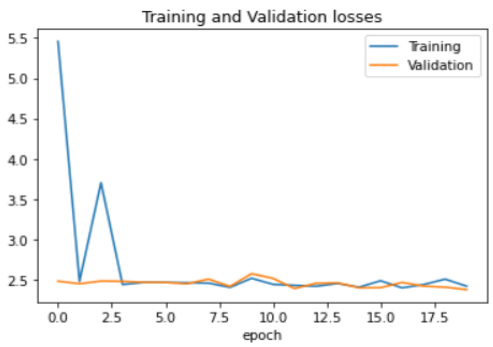


Fig. 7: VGG16 Model Losses.

VGG 16 was used to measure the accuracy of gastrointestinal cancer in gastrointestinal endoscopy images based on the standard model (i.e., without modifications). But it obtains a verification accuracy of about 11%, as we mentioned previously. In order to increase this accuracy, we applied some modifications as mentioned before. We thought that these modifications would lead to improving accuracy, but unfortunately it turned out that the VGG models are not at all suitable for the type of data used, as the outcomes demonstrated that the accuracy does not exceed 15% based on the same data sets. The results of this part are shown in Figure 8.

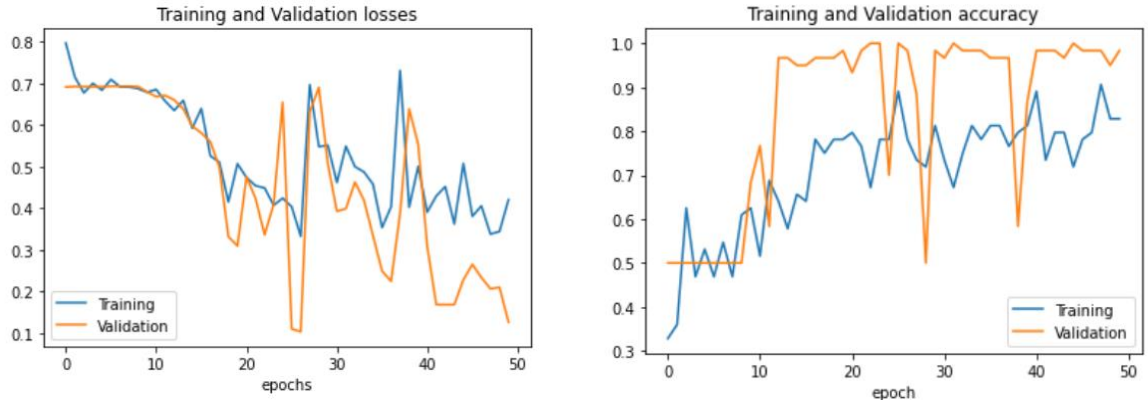
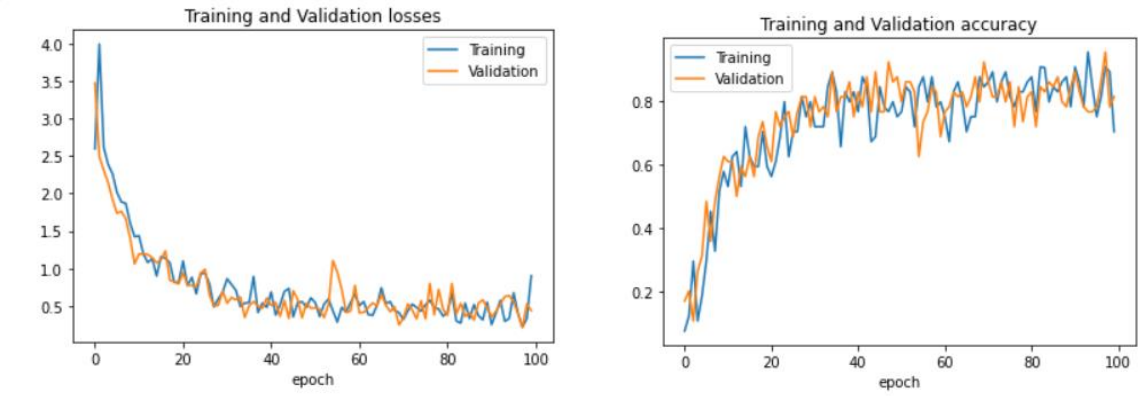


Fig. 8: VGG16 Model modifed Losses and Accuracy.

4.4. Modified DenseNet 121

When the DenseNet approach is applied, curve fluctuations exhibit more stability compared to those of the VGG and AlexNet models. It is observed that applying DenseNet to a larger number of datasets improves learning model accuracy, and that the accuracy is also influenced by the dataset. DenseNet technology used to datasets yields good performance. A few samples from DenseNet models using gastrointestinal endoscopic images are displayed in Fig. 9. As we were able to acquire a training accuracy of approximately 89% as well as a validation accuracy of approximately 89%, precision of 0.866499, and recall of 0.882051 for the dataset based on the simulation results.

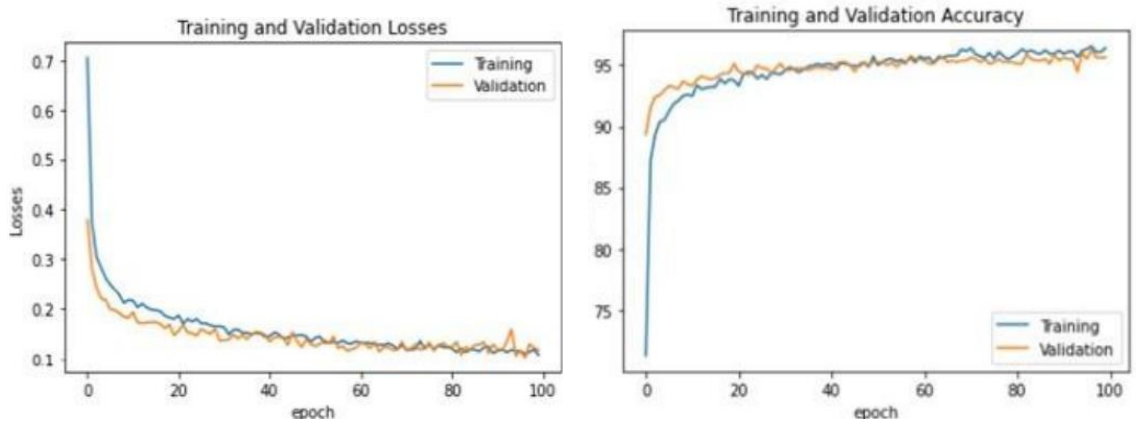


Out[17]:

	0	1	accuracy
precision	0.797357	0.866499	0.891346
recall	0.773504	0.882051	0.891346

Fig. 9: DenseNet 121Losses, Accuracy, precision and recall.

As we previously indicated, we make certain adjustments to the DenseNet 121 model in order to improve accuracy. When compared to earlier models, these improvements result in an improvement of up to 93%, fig 10.Which is quite good. Evaluations of model: Test accuracy: 0.929; test loss: 0.13; accuracy: 0.9298 Loss: 0.04069, precision of 0.855037, and recall of 0.892308.



Out[27]:

	0	1	accuracy
precision	0.806452	0.855037	0.938141
recall	0.747863	0.892308	0.938141

Fig. 10: DenseNet 121 modofied Losses, Accuracy precision and recall.

Table 2: Training and Testing Result.

Model's Name	Training Result Loss	Training Result Accuracy	Testing Result Loss	Testing Result Accuracy
Custom CNN	1.0577	0.5774	1.057	0.577
Modified Custom CNN	1.0691	0.6290	1.069	0.629
AlexNet	0.3549	0.8419	0.354	0.841
Modified AlexNet	0.3251	0.8591	0.325	0.859
VGG16	2.4069	0.1140	2.406	0.113
Modified DenseNet 121	0.04069	0.9298	0.13	0.929

SUMMARY AND CONCLUSIONS

1) This study addresses the challenge of diagnosing gastrointestinal cancer through a multi-model approach based on gastrointestinal endoscopic imaging, combining training and utilizing multiple endoscopic imaging datasets.
2) The proposed multi-model incorporates widely-used CNN varieties such as AlexNet, VGG, and DenseNet, renowned for their performance in image classification tasks.
3) Initially, accuracy was observed to be higher with endoscopic images compared to X-ray or CT scan data.
4) The suggested model integrates a multi-model framework comprising various CNN structures and multi-modal datasets, presenting an automated approach for detecting gastrointestinal cancer.
5) The model Modified DenseNet 121 significantly reduces error rates and increases the validation accuracy, achieving a 92% diagnosis rate for gastrointestinal cancer using endoscopic images.
6) The proposed multi-model approach outperforms individual models employing different CNN structures, as evidenced by testing and validation results.
7) Comparative analysis with other models indicates the superiority of our method in accurately identifying gastrointestinal cancer using gastrointestinal endoscopic imaging.
8) Testing results, based on extensive and diverse datasets, suggest that the suggested multi-model CNN structure holds promise for identifying gastrointestinal cancer infections.
9) Leveraging endoscopic images, this model serves as an automated method for diagnosing gastrointestinal cancer.

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of the National Telecommunication Institute(NTI) in Cairo, Egypt.

CONFLICT OF INTEREST

The authors have no financial interest to declare in relation to the content of this article.

REFERENCES

[1] Visaggi, P., de Bortoli, N., Barberio, B., Savarino, V., Oleas, R., Rosi, E. M., ... & Savarino, E. (2022). Artificial intelligence in the diagnosis of upper gastrointestinal diseases. Journal of Clinical Gastroenterology, 56(1), 23-35.

[2] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639), 115-118.

[3] Baker, J. A., Rosen, E. L., Lo, J. Y., Gimenez, E. I., Walsh, R., & Soo, M. S. (2003). Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion. American Journal of Roentgenology, 181(4), 1083-1088.

[4] Kominami, Y., Yoshida, S., Tanaka, S., Sanomura, Y., Hirakawa, T., Raytchev, B., ... & Chayama, K. (2016). Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. Gastrointestinal endoscopy, 83(3), 643-649.

[5] Urban, G., Tripathi, P., Alkayali, T., Mittal, M., Jalali, F., Karnes, W., & Baldi, P. (2018). Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. Gastroenterology, 155(4), 1069-1078.

[6] Wang, P., Berzin, T. M., Brown, J. R. G., Bharadwaj, S., Becq, A., Xiao, X., ... & Liu, X. (2019). Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut, 68(10), 1813-1819.

[8] Visaggi, P., de Bortoli, N., Barberio, B., Savarino, V., Oleas, R., Rosi, E. M., ... & Savarino, E. (2022). Artificial intelligence in the diagnosis of upper gastrointestinal diseases. Journal of Clinical Gastroenterology, 56(1), 23-35.

[9] Pecere, S., Milluzzo, S. M., Esposito, G., Dilaghi, E., Telese, A., & Eusebi, L. H. (2021). Applications of artificial intelligence for the diagnosis of gastrointestinal diseases. Diagnostics, 11(9), 1575.

[10] Berbís, M. A., Aneiros-Fernández, J., Olivares, F. J. M., Nava, E., & Luna, A. (2021). Role of artificial intelligence in multidisciplinary imaging diagnosis of gastrointestinal diseases. World Journal of Gastroenterology, 27(27), 4395.

[11] Wu, J., Chen, J., & Cai, J. (2021). Application of artificial intelligence in gastrointestinal endoscopy. Journal of Clinical Gastroenterology, 55(2), 110-120.

[12] Wesdorp, N. J., Hellingman, T., Jansma, E. P., van Waesberghe, J. H. T., Boellaard, R., Punt, C. J., ... & Kazemier, G. (2021). Advanced analytics and artificial intelligence in gastrointestinal cancer: a systematic review of radiomics predicting response to treatment. European journal of nuclear medicine and molecular imaging, 48, 1785-1794.

[13] Pogorelov, K., Ostroukhova, O., Jeppsson, M., Espeland, H., Griwodz, C., de Lange, T., ... & Halvorsen, P. (2018, June). Deep learning and hand-crafted feature based approaches for polyp detection in medical videos. In 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS) (pp. 381-386). IEEE.

[14] Visaggi, P., de Bortoli, N., Barberio, B., Savarino, V., Oleas, R., Rosi, E. M., ... & Savarino, E. (2022). Artificial intelligence in the diagnosis of upper gastrointestinal diseases. Journal of Clinical Gastroenterology, 56(1), 23-35.
[15] Berstad, T. J. D., Riegler, M., Espeland, H., de Lange, T., Smedsrud, P. H., Pogorelov, K., ... & Halvorsen, P. (2018, December). Tradeoffs using binary and multiclass neural network classification for medical multidisease detection. In 2018 IEEE International Symposium on Multimedia (ISM) (pp. 1-8). IEEE.
[16] De Lange, T., Halvorsen, P., & Riegler, M. (2018). Methodology to develop machine learning algorithms to improve performance in gastrointestinal endoscopy. World journal of gastroenterology, 24(45), 5057.
[17] Hicks, S., Smedsrud, P. H., Riegler, M. A., de Lange, T., Petlund, A., Eskeland, S. L., ... & Halvorsen, P. (2019). 383 deep learning for automatic generation of endoscopy reports. Gastrointestinal Endoscopy, 89(6), AB77.
[18] Owais, M., Arsalan, M., Choi, J., Mahmood, T., & Park, K. R. (2019). Artificial intelligence-based classification of multiple gastrointestinal diseases using endoscopy videos for clinical diagnosis. Journal of clinical medicine, 8(7), 986.
[19] Harzig, P., Einfalt, M., & Lienhart, R. (2019, October). Automatic disease detection and report generation for gastrointestinal tract examination. In Proceedings of the 27th ACM international conference on multimedia (pp. 2573-2577).
[20] Ghatwary, N., Zolgharni, M., & Ye, X. (2019). Gfd faster r-cnn: Gabor fractal densenet faster r-cnn for automatic detection of esophageal abnormalities in endoscopic images. In Machine Learning in Medical Imaging: 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 10 (pp. 89-97). Springer International Publishing.
[21] Thambawita, V., Jha, D., Hammer, H. L., Johansen, H. D., Johansen, D., Halvorsen, P., & Riegler, M. A. (2020). 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, 1(3), 1-29.
[22] Xie, K., & Peng, J. (2023). Deep learning-based gastric cancer diagnosis and clinical management. Journal of Radiation Research and Applied Sciences, 16(3), 100602.
[23] https://www.kaggle.com/datasets/balasubramaniamv/lung-cancer-dataset
[24] https://github.com/simula/hyper-kvasir/tree/master?tab=readme-ov-file
[25] Borgli, H., Thambawita, V., Smedsrud, P. H., Hicks, S., Jha, D., Eskeland, S. L., ... & de Lange, T. (2020). HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. Scientific data, 7(1), 283.
[26] Hiremath, G., Mathew, J. A., & Boraiah, N. K. (2023). Hybrid Statistical and Texture Features with DenseNet 121 for Breast Cancer Classification. International Journal of Intelligent Engineering & Systems, 16(2).