International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Automated Classification of Gastrointestinal Abnormalities using Convolutional Neural Networks

[1]Soham Navale, [1]Aditya Anjanikar, [1]Akheel Mohammed, [1]Sairam V. A. , [2]Shilpa Gite, [2]Smita Mahajan , Nandhini K.[1]

Submitted: 23/11/2023 **Revised**: 29/12/2023 **Accepted**: 09/01/2024

Abstract: With the use of a tiny, ingestible capsule fitted witha camera, a cutting-edge medical imaging technology called capsule endoscopy, high-resolution images of the digestive systemcan be obtained. By allowing for non-invasive digestion system viewing, this technique has revolutionized the diagnosis and monitoring of gastrointestinal illnesses. This paper is focused onthe creation of sophisticated picture classification algorithms to increase the clinical value of capsule endoscopy. Five distinct CNN models were used in the study: DenseNet121, EfficientNetB4, EfficientNetV2B3, ResNet101, and InceptionV3. Out of all the models, the DenseNet121 model performed the best, showing higher accuracy, AUC, precision, and recall. Its accuracy was 94.93%, recall was 93.87%, precision was 96.c9%, and AUC was 99.79%. =

Keywords: *Deep learning, Convolutional neural networks (CNNs), Capsule endoscopy, Pre-trained Modes, Gastrointestinal abnormalities.*

1. Introduction

The advancement of wireless medical technologies has drawn a lot of interest. Implantable and ingestible medical devices with wireless telemetry functionalities are regarded as one of the most significant approaches to biological health monitoring due to their non-invasive, painless, and incredibly successful qualities [1]. One such non-invasive diagnostic technique that greatly reduces patient discomfort is wireless capsule endoscopy (WCE), which transmits biological information and takes pictures of the gastrointestinal (GI) tract without the need for sedatives. By enabling the physician to directly observe gastrointestinal lesions and diagnose illnesses, it significantly resolves the issue that traditional endoscopy is unable to precisely locate the lesion in the gastrointestinal system. Light-emitting diodes (LEDs), cameras, wireless transmitters with antennae, and batteries are the standard components of WCE. According to studies, the capsules typically remain in the body for ten hours after being taken [2].For this reason, battery power longevity is especially crucial. Because the battery's volume and power supply are correlated, other parts' designs must be as small as feasible. Furthermore,a great deal of gastrointestinal tract images are taken by the capsule and transmitted to the portable recorder via the antenna. Hence, for a capsule endoscopy system, the design ofimplanted antennas is crucial.

[1]*Artificial Intelligence and Machine Learning Department, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India.*
[2] *Symbiosis Centre for Applied AI, Symbiosis International (Deemed University), Pune, India.*
Corresponding Author-Nandhini K.
2020sohamnavale@gmail.com,
aanjanikar@gmail.com,
akheel.sb@gmail.com
sairamadithya2002@gmail.com
shilpagite15@gmail.com,
nandhinik2@gmail.com

The non-invasive WCE was first introduced in 2000. It is shaped as a pill structure device that the patient swallows to allow it to pass through the digestive system and use the peristaltic motions of the stomach to offer an internal visual ofthe gastrointestinal tract [3]. It is regarded as the recommended instrument for the identification of small bowel problems because of its capacity to cover the entire gastrointestinal tract,including any restricted access area, and its remarkable interiorimaging results when compared with other methods for abnormalities visualization. Even though, the main disadvantage of this procedure is the time-consuming (40–60 minutes), laborious, and prone to human error review process carried out by medical personnel, especially considering the resulting eight-hour recorded video [4]. This has spiked interest in developers for developing methods for the automatic recognition, classification, and lesions diagnosis to help gastroenterologists in the speedy and effective examination ofvideo capsule endoscopies (VCE).

Computer aided pathology diagnosis is arguably the pinnacle of vision-based anomaly recognition and sickness diagnosis [4]. However, digital pathology has limitations much like any other technology. For instance, full scan imaging usually results in a large number of files that are difficult for computer algorithms to evaluate and need to be digitally saved. Image classification, segmentation, and detection appear to be extremely challenging challenges for computer vision systems[5]. Examining deep learning's extensive classification potential appears to be a wise course of action for overcoming the challenges associated with digital pathology [5]. Many deep neural network architectures have been trained on wide- scale image collections, such those from the ImageNet project, to tackle difficult problems like object categorization. The results achieved were nothing but extraordinary. It has becomecommon to get accuracy values in the mid- to high-90s when deep networks are evaluated to identify unknown samples [5]. Millions of pictures were used to train deep networks. The application of deep learning in digital pathology has a long wayto go, despite all of the advances. To successfully train any type of multi-layer neural network, the primary challenge appears to be the scarcity of big labelled data or datasets, which might notbe accessible for several years [6]. Thus, the only option is to begin building and training deep networks using the datasets that are now accessible.

Especially in the field of gastrointestinal medicine, capsule endoscopy has had a considerable impact. Among its contributions are:

Improved diagnostic capacity: The entire small intestine may be seen with a capsule endoscopy, which is difficult to achieve with conventional endoscopy and imaging methods [7]. As a result, it is now easier to diagnose a number of gastrointestinal disorders, including Crohn's disease, small bowel tumors, celiac disease, and mysterious gastrointestinal haemorrhage.

Early Detection: Capsule endoscopy has made it possible tofind abnormalities, such as precancerous lesions and early- stage cancers, early by delivering high-quality images of the digestive tract [8]. Better treatment outcomes and early intervention may result from this.

Monitoring Disease Progression: Chronic gastrointestinal disorders like Crohn's disease can be effectively tracked via capsule endoscopy. Doctors can monitor the progression of acondition, spot problems, and gauge how well a medication is working overtime [9].

Minimally Invasive: Capsule endoscopy, in contrast to standard endoscopy, is minimally invasive and does not call for sedation or the insertion of a protracted, flexible tube intothe body. As a result, it is a safer and more pleasant choice for patients [10].

Patient compliance: Because patients just need to swallowa little capsule, capsule endoscopy is typically well-tolerated by patients. This could motivate more patients to get crucial diagnostic tests, resulting in earlier identification andtreatment of digestive problems [11].

Research and Education: The data and images produced bycapsule endoscopy have advanced our knowledge of theanatomy and pathophysiology of the small intestine and contributed to research in gastroenterology. Improvements in medical education and training have followed from this [12].

Technical Developments: The advancements in capsule endoscopy technology have made it possible for more breakthroughs in medical imaging and diagnostics. It has sparked advancements in image processing algorithms, data transport, and camera technology [13].

All of the above supports the importance of the capsule endoscopy technology but It's crucial to REMEMBER that capsule endoscopy has some drawbacks as well, such as the inability to conduct therapeutic interventions or obtain biopsies [14]. When abnormalities are found, more steps maybe required for confirmation and treatment. The capacity to diagnose and treat gastrointestinal illnesses has been greatly aided by the use of capsule endoscopy, which is still a key tool in the diagnostic and monitoring toolbox of gastroenterologists.

2. Related Works

From [15], It is determined that the diagnosis and classification processes use the Efficient Attention Mechanism Network (EAM-Net). With certain limitations like system needs to be validated on a larger dataset and in clinical settings. And the dataset used contains 15120 colonoscopy pictures from 768 ulcerative colitis patients and768 non-ulcerative colitis patients.

From [16], It is determined that the diagnosis and classification processes use the DenseNet Model. With certain limitations like sample size and data used is from onlya single dataset. And the dataset used contains 1000 GI tractimages; 500 images of polyps and 500 images of without polyps

From [17], It is determined that the diagnosis and classification processes use the DenseNet. With certain limitations like DenseNet can be computationally expensiveto train and deploy. And the dataset used contains 2,000 WCE images, which were collected from a single hospital.

From [18], It is determined that the diagnosis and classification processes use the conventional methods of image processing, deep learning, and machine learning. Withcertain limitations like limited size, potential for bias in deeplearning. And the dataset used is MICCAI WCE Challenge dataset, the Kvasir-Capsule dataset, and the M2CAI WCE Polyp Detection Challenge dataset.

From [19], It is determined that the diagnosis and classification processes use the SVM classifier which is trained directly on the raw WCE video images, without the need for any manual. With certain limitations like system needs to be validated on a larger dataset and in clinical settings. And the dataset used contains 1000 upper gastrointestinal endoscopy images from 100 patients.

From [20], It is determined that the diagnosis and classification processes use the Multi-Layered Perceptron Neural Network (MLNN). With certain limitations like system needs a substantial volume of training data. And the dataset used is CUI WA Stomach Diseases dataset.

From [21], It is determined that the diagnosis and classification processes use the SVM classifier which is trained directly on the raw WCE video images, without the need for any manual feature extraction steps. With certain limitations like system requires a large amount of training data. And the dataset used is Ten WCE films with a total frame count of fifty thousand. Images of the stomach, colon, small intestine, and esophagus are included.

From [22], It is determined that the diagnosis and classification processes use the support vector machine (SVM) classifier. With certain limitations like system requires a large dataset, might not be able to detect bleeding that is very small. And the dataset used is Private.

From [23], It is determined that the diagnosis and classification processes use the convolutional neural network(CNN). With certain limitations like the system requires a large dataset, might not be able to detect bleeding that is very small. And the dataset contains 21,320 CE (Capsule Endoscopy) images.

From [24], It is determined that the diagnosis and classification processes use the convolutional neural network(CNN). With certain limitations like algorithm can be noise-sensitive. could result in both false negatives and false positives. And the dataset contains 2417 SBCE images.

From [25], It is determined that the diagnosis and classification processes use two dense layers, ReLu activation, average pooling layers, and a multilayer convolutional neural network (CNN). With certain limitations like the method may not be robust to variations inthe quality and appearance of WCE images. And the datasetcontains 1319 WCE images, which were collected from a single hospital.

From [26], It is determined that the diagnosis and classification processes use pre-trained AlexNet CNN, custom CNN. With certain limitations like the framework may not be robust to variations in the quality and appearanceof CE images. And the dataset contains two publicly available CE image datasets: the capsule endoscopy.org dataset and the KID dataset with 1311 and 1186 images present in each dataset respectively.

From [27], It is determined that the diagnosis and classification processes use a hybrid ELM-based CAD system that blends a modified rotation-invariant local binary pattern (RLBP) descriptor with a histogram of oriented gradients (HOG) descriptor, ELM classifier. With certain limitations like the system's performance could be enhanced by employing a multi-classifier strategy. And the dataset contains 1,000 WCE images, which were collected from a single hospital.

From [28], It is determined that the diagnosis and classification processes use a modified YOLOv3 CNN-based target detection method. With certain limitations like the system needs a substantial volume of data for training, the framework may not be robust to variations in the quality andappearance of CE images. And the dataset contains 1,500 WCE images.

From [29], It is determined that the diagnosis and classification processes use a k-nearest neighbors (k-NN) model. With certain limitations like it requires the construction of a bespoke optical sensor to measure the spectral features of the GI tract. And the dataset contains 300 non-bleeding and 100 bleeding pictures from ten different patients.

From [30], It is determined that the diagnosis and classification processes use Fuzzy logic and PCA model. With certain limitations like model is evaluated on a small dataset of 10 patients. And the dataset contains 100 bleedingand 300 non-bleeding images from 10 patients.

From [31], It is determined that the diagnosis and classification processes use Deep convolutional neural network (CNN) and geometric features. With certain limitations like system won't be able to adapt to new illness kinds that aren't represented in the training dataset. And the dataset contains 5500 WCE images from 110 patients with gastrointestinal tract diseases, such as ulcers, tumors, and inflammation.

From [32], It is determined that the diagnosis and classification processes use pre-trained ResNet-101 CNN model. With certain limitations like system needs a large amount of training data. And the dataset contains 1,000 CE images.

From [33], It is determined that the diagnosis and classification processes use pre-trained ResNet-50 CNN witha custom CNN model. With certain limitations like the modelmay not be robust to variations in the quality of data. And thedataset contains 1,000 endoscopy images.

From [34], It is determined that the diagnosis and classification processes use pre-trained ResNet-50 CNN. With certain limitations like the model should be trained with huge data. And the dataset contains 1,000 Capsule endoscopyimages.

From [35], It is determined that the diagnosis and classification processes use pre-trained ResNet-50 CNN with a custom CNN. With certain limitations like the model may not be robust to variations in the quality of data. And the dataset contains 1,000 Capsule endoscopy images.

From [36], It is determined that the diagnosis and classification processes use combination of U-Net and EfficientNetB5. With certain limitations like with a rather small dataset of colonoscopy images, the model was trained. Second, if the photos differ greatly from the training set, the model might not be able to generalize adequately to them. Third, noise and visual artifacts may cause the model to become sensitive. And the dataset are combined which are KVASIR and CVC-ClinicDB. 1000 colonoscopy pictures with ground truth labelling for polyp segmentation make up the KVASIR dataset. 612 colonoscopy pictures with groundtruth labelling for polyp classification make up the CVC- ClinicDB dataset.

From [37], It is determined that the diagnosis and classification processes use Mask R-CNN. With certain limitations like comparatively limited collection of colonoscopy pictures was used to train the model. Second, pictures that deviate greatly from the training set of images may cause the model to perform poorly in terms of generalization. Third, there's a chance the model is susceptible to picture artifacts and noise. And the dataset used is the Kvasir-SEG dataset, which consists of 1,600 colonoscopy images with pixel-level annotations of polyps. Other researchers in the medical imaging domain demonstrated different aspects such segmentation, explainablity, optimization in [38-41] to signify latest techniques.

3. Materials and Methodology

A. Outline of the Methodology

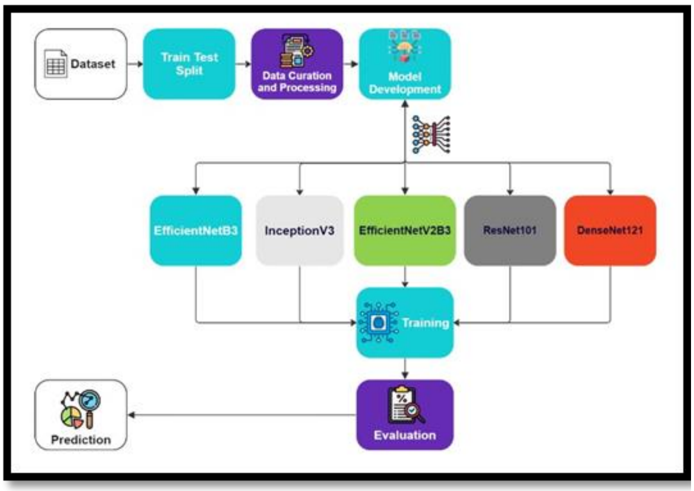


Fig 1: Flowchart of Classification

Utilizing deep learning techniques, the image

B. Data Processing

component organizes collected frames into pertinent anatomical regions or clinical findings. Convolutional neural networks(CNNs) and transfer learning are used in the model to identify images with high accuracy, assisting in the early diagnosis and localization of diseases. The efficiency and accuracy of capsuleendoscopy analysis are greatly increased by our suggestedframework, allowing medical personnel to make prompt and well-informed decisions regarding patient care. The findings show how AI-driven image analysis has the potential to revolutionize the area of gastroenterology and open the door for more efficient and individualized medical interventions.

From **Figure 1** which describes flow and working of the

classification

project. To describe the process below points are need to be considered.

C. Dataset Description

A sizable VCE (Video Capsule Endoscopy) dataset called Kvasir-Capsule was gathered from hospital examinations in Norway. From the 117 videos that make up Kvasir-Capsule, 4,741,504 picture frames can be extracted. In order to identify anomalies from 14 different classes of findings. We have labelled and medically verified 47,238 frames. To fully realizethe promise of VCE technology, improved algorithms can be developed with the help of the Kvasir-Capsule dataset. Below **Table 1** describes the data which is collected in various classesThe data was obtained from the given link https://osf.io/dv2ag/..

Table 1: Class distribution of the dataset

Classes	Images
ampulla of vater	10
angiectasia	866
blood-fresh	446
blood-hematin	12
erosion	506
erythema	159
foreign body	776
ileocecal valve	4189
lymphangiectasia	592
normal clean mucosa	34338
polyps	55
pylorus	1529
reduced mucosal view	2906
ulcer	854

D. Data Processing

In order to gain better performance for prediction, the data needed to be pre-processed and ideal for training and testing purposes. But here, the data was divided into 14 classes wheretrain and test data were combined by default. Therefore, the data needed to be splitted manually and randomly so that training and test could be done easily. So, the data was splittedmanually and randomly on 80% and 20% resp. by combining every class together which gave 37783 training and 9455testing images.

After splitting the data, for better performance image generation method is used for randomly generating training and testing images by shifting image angles, zooming the images and randomly shifting images horizontally and vertically. This gives more data for training purposes which can lead to better model performance. For this various parameter were used suchas **rescale, rotation_range, zoom_range, width_shift_range,height_shift_range, image size and batch size.**

E. Model Development

In this a simple CNN (Convolutional Neural Network) model is created using various layers like Input layer, Base model (Functional), GlobalAveragePooling2D and multiple dense layers (here three). This gives a total 18,726,765 paramsfrom which 18,601,558 are Trainable and 125,207 are Non- trainable.

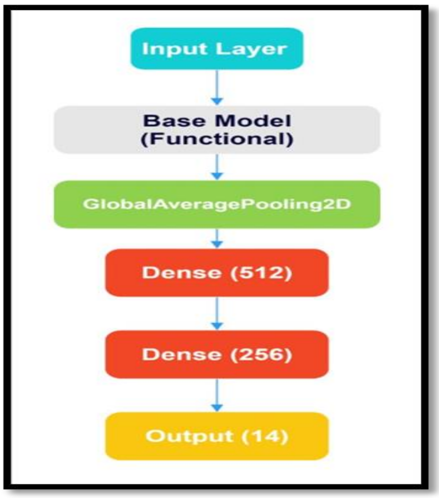


Fig 2: Flowchart of **CNN Model**

In this, the model is compiled with various parameters as **loss** which is set to “categorical_crossentropy”, **optimizer** which is set to “Adam” with “learning rate = 0.001” and **metrics** is set to a list including values like “accuracy”, “AUC”, “'Precision”, “Recall”. Model is then trained with given parameters like **epochs** which are 40 with each epoch contains 50 steps, then **ReduceLROnPlateau** for monitoring loss per epoch, then **ModelCheckPoint** is used to save best results which monitors Validation Accuracy and finally the results are stored in a csv file using a **CSVLogger.** After training the model for 40 epochs the model is evaluated and the ideal performance values are selected. Here **Loss, Accuracy, AUC (Area Under Curve), Precision, Recall** are evaluated.

F. Prediction

Based on the model performance the ideal model is selected and saved which then later used for Prediction or classification of new dataset or image.

G. Models Used

Following are the different models used for classification purposes:

1. InceptionV3

In 2014, Google researchers unveiled the Inception model, a convolutional neural network (CNN) architecture. It has been used to win multiple competitions in the field and is well- known for its excellent accuracy on image identification tasks. The Inception model's foundation is the notion that processing images in numerous parallel layers enables it to extract more information from the data. This improves its ability to identify things in photos, regardless of how blurry or distorted the images might appear.

The following are some of InceptionV3's salient features:

• **Convolution factorization**: A factorization technique is used by InceptionV3 to lower the computational cost of convolutions. To do this, bigger convolutional filters are divided into smaller, more effective filters.

• **Multiple parallel routes**: To extract characteristics from images, InceptionV3 makes use of multiple parallel paths. This can help the network perform better on a range of tasks by enabling it to learn a more varied collection of features.

• **Global average pooling**: To lower the dimensionality of the convolutional layer output, InceptionV3 employs global average pooling. As a result, the network performs better during generalization and avoids overfitting.

2. EfficientNetB4:

Google AI unveiled EfficientNetB4, a convolutional neural network architecture that uses less computing power than earlier models to deliver state-of-the-art performance in object detection and picture classification tasks in 2019. It belongs to the family of neural networks called EfficientNet, which aims to be more economical in terms of computation and accuracy.

Essential Elements of EffectiveNetB4:

- **Compound Scaling:** To expand the network's size in a more balanced manner, EfficientNetB4 employs a method known as compound scaling. This entails changing the number of channels and filter sizes in addition to simultaneously expanding the network's depth, width, and resolution.

- **Squeeze-and-Excitation (SE) Layers:** To enhance the network's capacity for feature extraction, EfficientNetB4 makes use of SE layers. The network's SE layers constantly modify the weights of each channel, enabling it to suppress less significant features and concentrate on more crucial ones.

- **Depth-wise Separable Convolutions:** EfficientNetB4 lowers the computational cost of convolutions by employing depth-wise separable convolutions. A conventional convolution is factored into two distinct operations: depth-wise convolution and pointwise convolution, thanks to depth-wise separable convolutions.

3. Resnet 101:

ResNet101, a convolutional neural network (CNN) architecture, was first revealed in 2015. It has been used to win multiple competitions in the field and is well-known for its excellent accuracy on image identification tasks. The foundation of ResNet101 is the concept of residual connections, which allows the model to learn more complex picture representations without running into the vanishing gradient issue that can arise with deep CNNs.

Important ResNet101 Features:

• **Residual Connections**: ResNet101 solves the vanishing gradient problem through the use of residual connections. Gradients can move more directly from the input to the output since residual connections bypass some of the network's layers. This aids in keeping the gradients from getting too tiny and disappearing completely.

• **Batch Normalization**: To enhance the training procedure, ResNet101 employs batch normalization. The activations of each layer are normalized via batch normalization, which lessens internal covariate shift and enhances the generalization performance of the network.

• **Bottleneck Design**: The convolutional blocks of ResNet101 are designed with a bottleneck. One kind of convolutional block called a bottleneck block reduces the dimensionality of the input using 1x1 convolutions before applying 3x3 convolutions.

4. **EfficientNetV2B3**:

Convolutional neural network (CNN) architecture EfficientNetV2B3 is a member of the EfficientNet model family. This high-performance model aims to be precise and effective at the same time. Based on the EfficientNetV2 architecture, EfficientNetV2B3 achieves great accuracy at low computing cost by combining squeeze-and-excitation layers, compound scaling, and depth-wise separable convolutions.

Important attributes of EfficientNetV2B3:

- **Compound Scaling:** To expand the network's size in a more balanced manner, EfficientNetV2B3 employs a method known as compound scaling. This entails changing the number of channels and filter sizes in addition to simultaneously expanding the network's depth, width, and resolution.

- **Squeeze-and-Excitation (SE) Layers:** To enhance the network's capacity for feature extraction, EfficientNetV2B3 makes use of SE layers. The network's SE layers constantly modify the weights of each channel, enabling it to suppress less significant features and concentrate on more crucial ones.

- **Depth-wise Separable Convolutions:** EfficientNetV2B3 lowers the computational cost of convolutions by employing depth-wise separable convolutions. A conventional convolution is factored into two distinct operations: depth-wise convolution and pointwise convolution, thanks to depth-wise separable convolutions.

5. **DenseNet121**:

Huang et al. unveiled DenseNet121, a convolutional neural network (CNN) architecture, in 2017. It has been used to win multiple competitions in the field and is well-known for its excellent accuracy on image identification tasks. Dense connections, which enable information sharing between all tiers in the network, are the foundation of DenseNet121. In comparison to conventional CNN architectures, this enables the model to learn more intricate characteristics and attain greater accuracy.

Important DenseNet121 Features:

- **Dense connection:** It is a technique used by DenseNet121 to link every layer in the network to every other layer. As a result, the network is able to extract more intricate features from the input data, potentially improving its performance on a range of image categorization tasks.

- **Growth Rate:** DenseNet121 regulates the number of connections between layers via a growth rate parameter. Increased growth rate can result in a denser network and better performance, but it also adds to the computational burden and number of parameters.

- **Global Average Pooling:** DenseNet121 reduces the dimensionality of the convolutional layers' output by utilizing global average pooling. As a result, the network performs better during generalization and avoids overfitting.

4. Results Analysis and Discussion

Below **Table 2** shows the performance of various models used for classification purposes. The performance difference between these models is not significant apart from ResNet101, but clearly the Densnet121 shows better results than other models hence, it should be used for classification or prediction purposes.

Table 2- Comparison of the Models

Models	**Loss**	**Accuracy**	**AUC**	**Precision**	**Recall**
Efficient NET B4	0.2221	93.63%	99.58%	95.26%	92.20%
InceptionV3	0.2665	92.22%	94.49%	94.98%	90.47%
Resnet101	0.5962	82.59%	98.23%	88.61%	77.02%
EfficientNetV2	0.3091	91.21%	99.30%	93.41%	89.75%
DenseNet121	**0.1735**	**94.93%**	**99.79%**	**96.39%**	**93.87%**

From the Models from **Table 2**. The trained and tested models are then fine-tuned for better performance. And below chart analysis represents the model’s performance respectively. Following are the chart analysis of **InceptionV3 with** performance across epochs using various parameters:

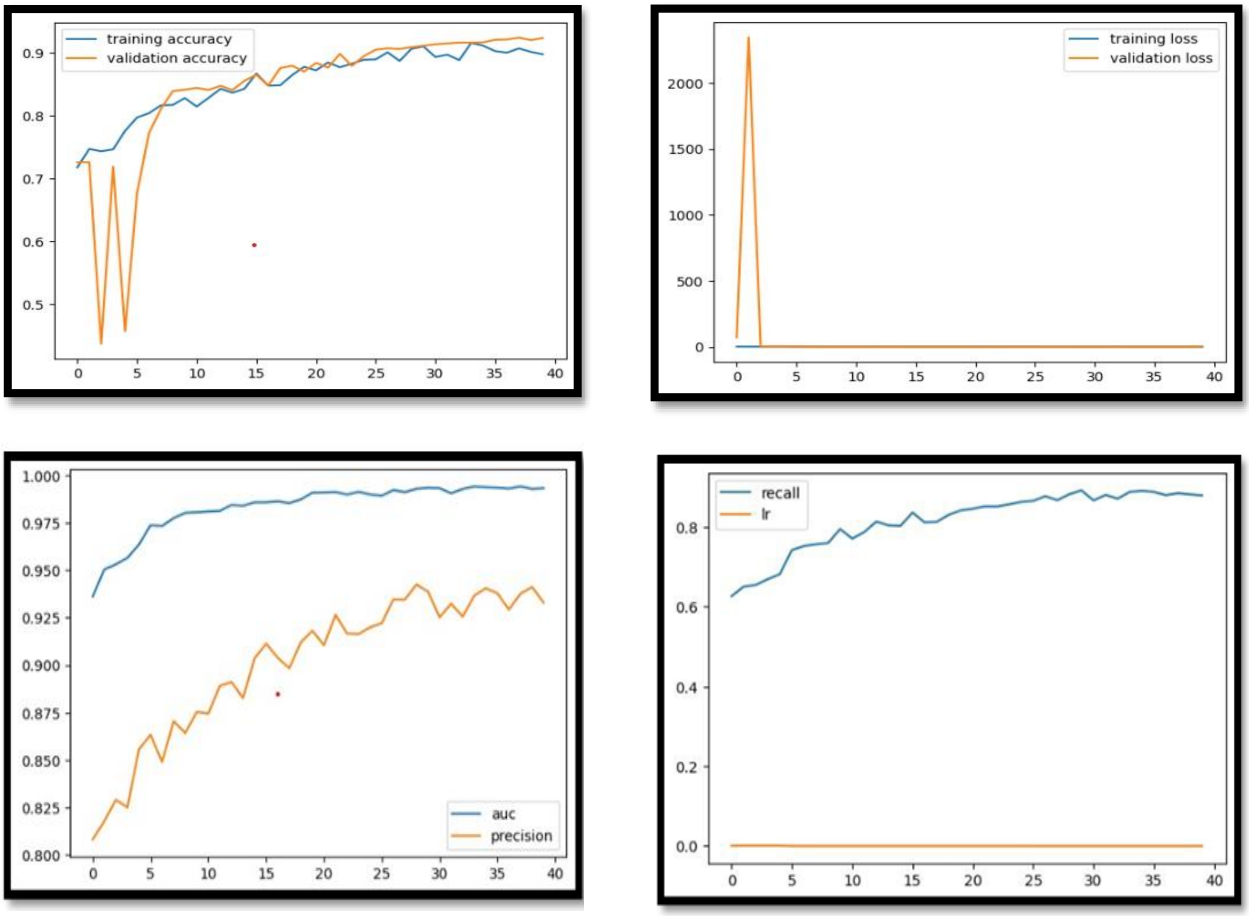
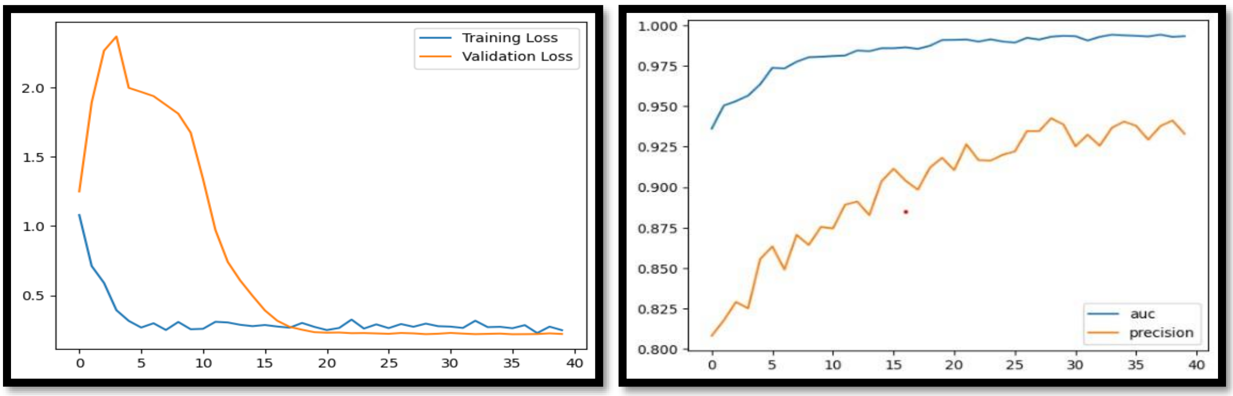


Fig 3- Chart Analysis of InceptionV3

Figure 3 represents different charts for **InceptionV3** such as **3.a** represents **"Training and Validation for Accuracy"**, **3.b** represents **"Training and Validation for Loss"**, **3.c** represents **"Area Under Curve and Precision"** and **3.d** represents **"Recall and Learning rate"**. All of the above charts are analysed for performance across the epochs on which the models are trained and tested. Following are the chart analysis of **EfficientNetB4** with performance across epochs using various parameters:



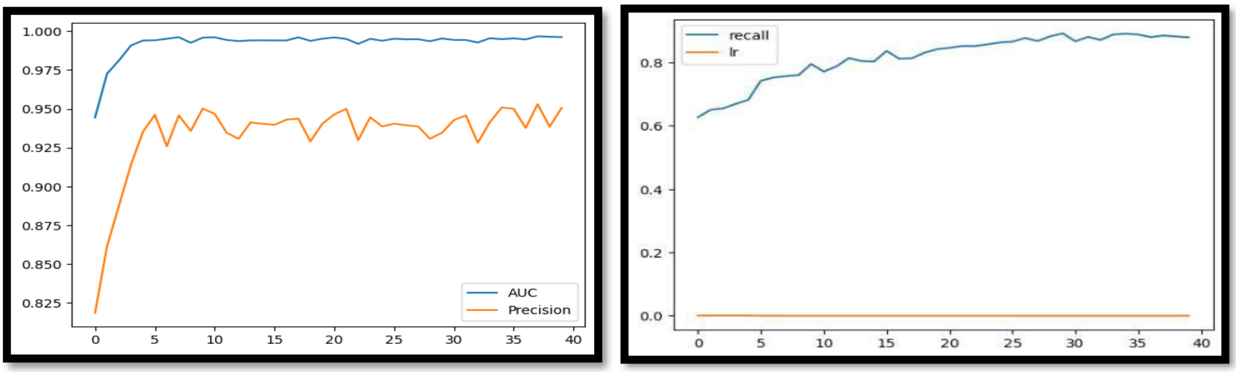


Fig 4- Chart Analysis of EfficientNetB4

Figure 4 represents different charts for EfficientNetB4 such as 4.a represents "Training and Validation for Accuracy", 4.b represents "Training and Validation for Loss", 4.c represents "Area Under Curve and Precision" and 4.d represents "Recall and Learning rate". All of the above charts are analysed for performance across the epochs on which the models are trained and tested. Following are the chart analysis of Resnet101 with performance across epochs using various parameters:

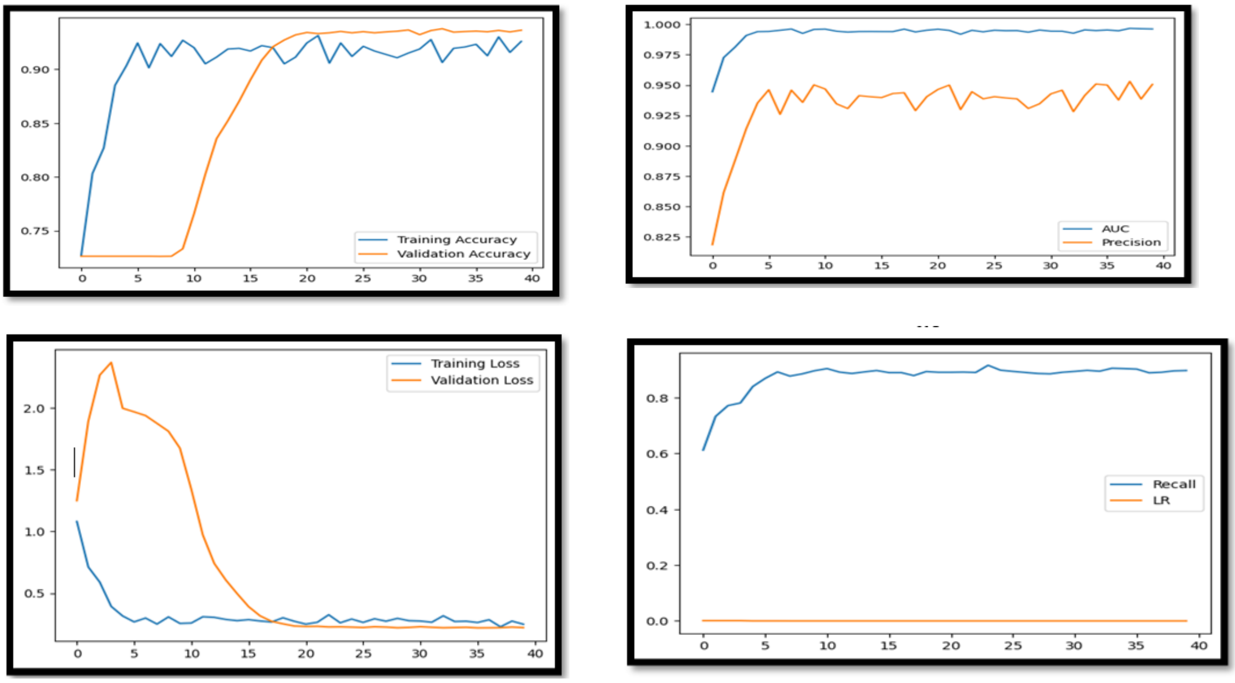


Fig 3- Chart Analysis of EfficientNetB5

Figure 5 represents different charts for EfficientNetB5 such as 5.a represents "Training and Validation for Accuracy", 5.b represents "Training and Validation for Loss", 5.c represents "Area Under Curve and Precision" and 5.d represents "Recall and Learning rate". All of the above charts are analysed for performance across the epochs on which the models are trained and tested. Following are the chart analysis of EfficientNetV2B3 with performance across epochs with various parameters.

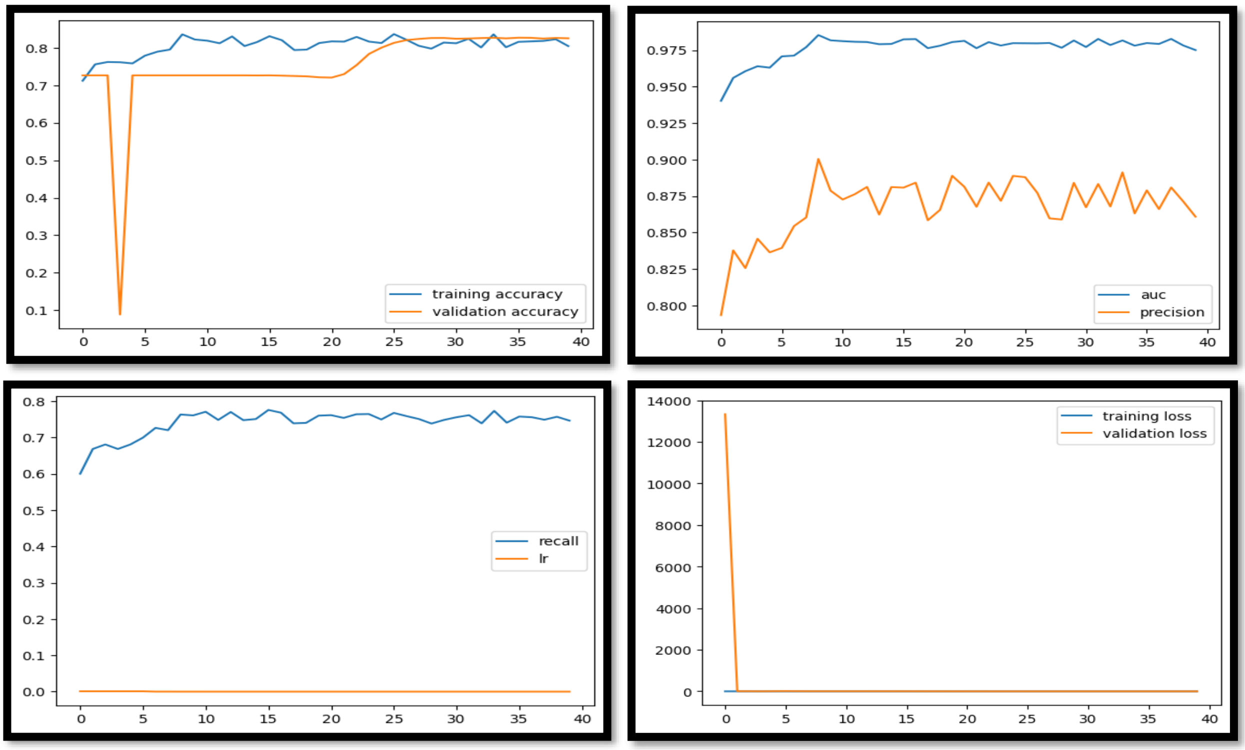


Fig 6- Chart Analysis of EfficientNetV2B3

Figure 6 represents different charts for **EfficientNetV2B3** such as **6.a** represents **"Training and Validation for Accuracy"**, **6.b** represents **"Training and Validation for Loss"**, **6.c** represents **"Area Under Curve and Precision"** and **6.d** represents **"Recall and Learning rate"**. All of the above charts are analysed for performance across the epochs on which the models are trained and tested. Following are the chart analysis of **Densenet121** with performance across the epochs using various parameters.

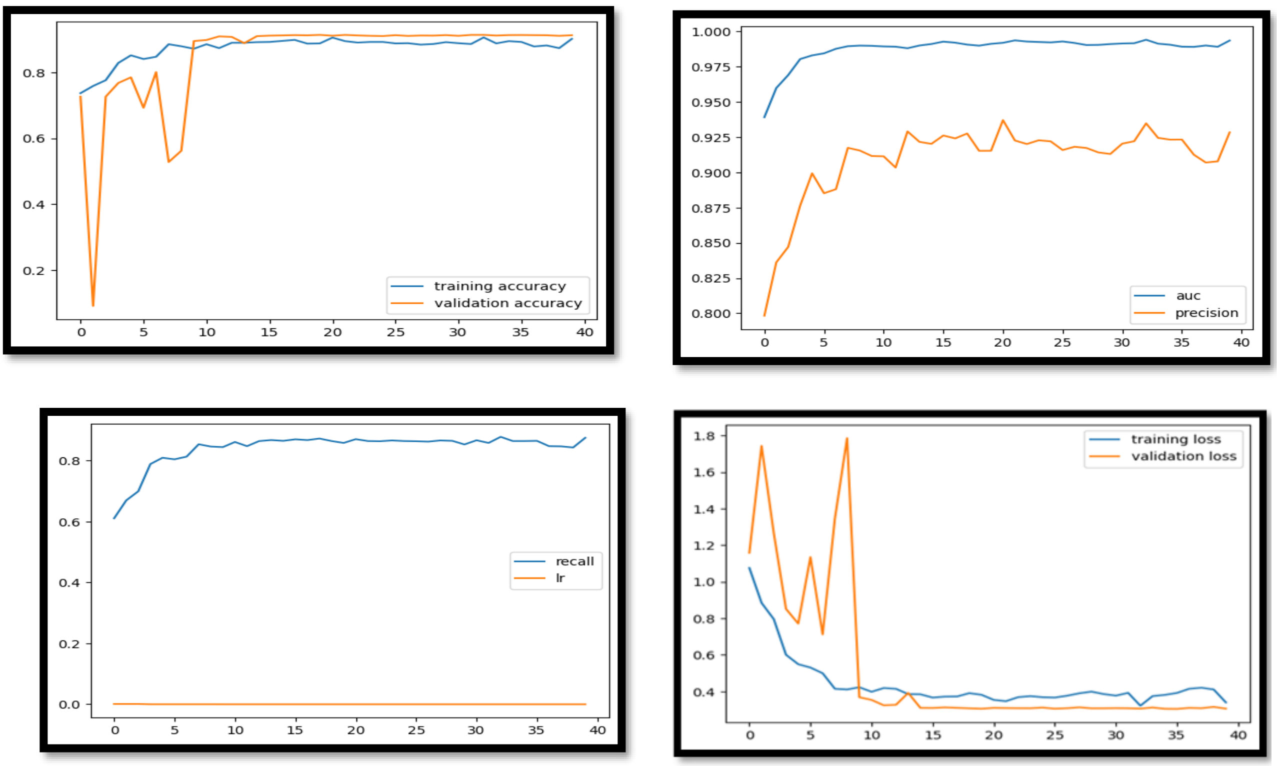


Fig 7: Chart Analysis of the Densenet121

Figure 7 represents different charts for **Densenet121** such as **7.a** represents **"Training and Validation for Accuracy"**, **7.b** represents **"Training and Validation for Loss"**, **7.c** represents **"Area Under Curve and Precision"** and **7.d** represents **"Recall and Learning rate"**. All of the above charts are analysed for performance across the epochs on which the models are trained and tested.

Above Chart Analysis of different models shows their performance according to the epochs on which they are trained with various parameters. These models are fine-tuned or developed to gain better accuracy or performance. These performed models are used for classification of diseases or abnormalities present in the gastronomical tract.

5. Conclusions

Using the Kvasir-Capsule dataset, the study provided a thorough investigation of the deep learning-based classification of capsule endoscopy (CE) images. The dataset includes a sizable number of labelled CE Images which are divided into various classes that show a range of gastrointestinal disorders. Five distinct CNN models were used in the study: DenseNet121, EfficientNetB4, EfficientNetV2B3, ResNet101, and InceptionV3. Every model was trained and assessed with a range of parameters, such as metrics, optimizer, loss function, and epochs. Out of all the models, the DenseNet121 model performed the best, showing higher accuracy, AUC, precision, and recall. Its accuracy was 94.93%, recall was 93.87%, precision was 96.c9%, and AUC was 99.79%. These astounding findings demonstrate the potential of deep learning for the classification of CE images, opening exciting new possibilities for the diagnosis and treatment of gastrointestinal disorders. This research project opens avenues for future work in several directions. First, to develop capsules that can also deliver therapy to the digestive tract. Second, to develop AI algorithms that can automatically analyze CE images and identify abnormalities in real-time. Third, to develop wireless power transmission technology that would allow capsules to be powered continuously. This would extend the lifespan of capsules and make them more practical for long-term monitoring of patients with gastrointestinal disorders. And many more developments can be done in the similar fields.

6. Acknowledgement

We would like to thank our institute Symbiosis Institute of Science and Technology, Pune for providing us the computational facilities with latest CPU's, memory and TPU's to carry out our project and we would like to thank our mentor Dr.Shilpa Gite and Dr.Nandhini for providing us valuable guidance and assistance in carrying out this project. The authors show no conflict of interest.

References

[1] Kiourti, A., Psathas, K.A. and Nikita, K.S., 2014. Implantable and ingestible medical devices with wireless telemetry functionalities: A review of current status and challenges. Bioelectromagnetics, 35(1), pp.a-15.

[2] Wang, G.B., Xuan, X.W., Jiang, D.L., Li, K. and Wang, W., 2022. A miniaturized implantable antenna sensor for wireless capsule endoscopy system. AEU-International Journal of Electronics and Communications, 143, p.a54022.

[3] Iddan, G., Meron, G., Glukhovsky, A. and Swain, P., 2000. Wireless capsule endoscopy. Nature, 405(6785), pp.d17- 417.

[4] Fonseca, F., Nunes, B., Salgado, M. and Cunha, A., 2022.Abnormality classification in small datasets of capsule endoscopy images. Procedia Computer Science, 196, pp.d69-476.

[5] B. Kieffer, M. Babaie, S. Kalra and H. R. Tizhoosh, "Convolutional neural networks for histopathology image classification: Training vs. Using pre-trained networks," 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, Canada, 2017, pp. 1-6, doi: 10.a109/IPTA.b017.8310149.

[6] Ching, T., Himmelstein, D.S., Beaulieu-Jones, B.K., Kalinin, A.A., Do, B.T., Way, G.P., Ferrero, E., Agapow, P.M., Zietz, M., Hoffman, M.M. and Xie, W., 2018. Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), p.b0170387.

[7] Costamagna, G., Shah, S.K., Riccioni, M.E., Foschia, F., Mutignani, M., Perri, V., Vecchioli, A., Brizi, M.G., Picciocchi, A. and Marano, P., 2002. A prospective trial comparing small bowel radiographs and video capsule endoscopy for suspected small bowel disease. Gastroenterology, 123(4), pp.999-1005.

[8] Namikawa, K., Hirasawa, T., Yoshio, T., Fujisaki, J., Ozawa, T., Ishihara, S., Aoki, T., Yamada, A., Koike, K., Suzuki, H. and Tada, T., 2020. Utilizing artificial intelligence in endoscopy: a clinician's guide. Expert review of gastroenterology & hepatology, 14(8), pp.689- 706.

[9] Zonderman, J. and Vender, R.S., 2009. UnderstandingCrohn disease and ulcerative colitis. Univ. Press of Mississippi.

[10] Haycock, A., Cohen, J., Saunders, B.P., Cotton, P.B. and Williams, C.B., 2014. Cotton and Williams' practical gastrointestinal endoscopy: the

fundamentals. John Wiley& Sons.

[11] De'Angelis, G.L., Fornaroli, F., De'Angelis, N., Magiteri, B. and Bizzarri, B., 2007. Wireless capsule endoscopy for pediatric small-bowel diseases. Official journal of the American College of Gastroenterology| ACG, 102(8), pp.a749-1757.

[12] Wahab, H., Mehmood, I., Ugail, H., Sangaiah, A.K. and Muhammad, K., 2023. Machine learning based small bowel video capsule endoscopy analysis: Challenges and opportunities. Future Generation Computer Systems.

[13] M. R. Neuman et al., "Advances in Medical Devices and Medical Electronics," in Proceedings of the IEEE, vol. 100, no. Special Centennial Issue, pp. 1537-1550, 13 May 2012,doi: 10.a109/JPROC.b012.b190684.

[14] Di Nardo, G., Esposito, G., Ziparo, C., Micheli, F., Masoni, L., Villa, M.P., Parisi, P., Manca, M.B., Baccini, F. and Corleto, V.D., 2020. Enteroscopy in children and adults with inflammatory bowel disease. World Journal of Gastroenterology, 26(39), p.5944.

[15] Luo, X., Zhang, J., Li, Z. and Yang, R., 2022. Diagnosis of ulcerative colitis from endoscopic images based on deep learning. Biomedical Signal Processing and Control, 73, p.a03443.

[16] Souaidi, M., Lafraxo, S., Kerkaou, Z., El Ansari, M. and Koutti, L., 2023. A Multiscale Polyp Detection Approach for GI Tract Images Based on Improved DenseNet and Single-Shot Multibox Detector. Diagnostics, 13(4), p.733.

[17] Zhou, T., Ye, X., Lu, H., Zheng, X., Qiu, S. and Liu, Y., 2022. Dense convolutional network and its application in medical image analysis. BioMed Research International, 2022.

[18] Rahim, T., Usman, M.A. and Shin, S.Y., 2020. A survey on contemporary computer-aided tumor, polyp, and ulcer detection methods in wireless capsule endoscopy imaging.Computerized Medical Imaging and Graphics, 85, p.a01767.

[19] Wang, S., He, Q., Zhang, P., Chen, X. and Zuo, S., 2022. Toward Automatic Detection of Gastric Lesion for Upper Gastrointestinal Endoscopy with Neural Network. Journal of Medical Robotics Research, 7(01), p.b141003.

[20] Khan, M.A., Majid, A., Hussain, N., Alhaisoni, M., Zhang, Y.D., Kadry, S. and Nam, Y., 2021. Multiclass stomach diseases classification using deep learning features optimization.

[21] Mackiewicz, M., Berens, J. and Fisher, M., 2008. Wireless capsule endoscopy color video segmentation. IEEE Transactions on Medical Imaging, 27(12), pp.a769-1781.

[22] Vajravelu, A., Selvan, K.S., Jamil, M.M.B.A., Jude, A. and Diez, I.D.L.T., 2023. Machine learning techniques to detect bleeding frame and area in wireless capsule endoscopy video. Journal of Intelligent & Fuzzy Systems,44(1), pp.c53-364.

[23] Afonso, J., Mascarenhas, M., Ribeiro, T., Cardoso, H., Andrade, P., Ferreira, J.P., Saraiva, M.M. and Macedo, G., 2022. Deep Learning for Automatic Identification andCharacterization of the Bleeding Potential of Enteric Protruding Lesions in Capsule Endoscopy. Gastro Hep Advances, 1(5), pp.835-843.

[24] Houdeville, C., Souchaud, M., Leenhardt, R., Beaumont, H., Benamouzig, R., McAlindon, M., Grimbert, S.,Lamarque, D., Makins, R., Saurin, J.C. and Histace, A., 2021. A multisystem-compatible deep learning-based algorithm for detection and characterization of angiectasis in small-bowel capsule endoscopy. A proof-of-concept study. Digestive and Liver Disease, 53(12), pp.a627-1631.

[25] Vani, V. and Prashanth, K.M., 2022. Ulcer detection in Wireless Capsule Endoscopy images using deep CNN. Journal of King Saud University-Computer and Information Sciences, 34(6), pp.c319-3331.

[26] Ghosh, T. and Chakareski, J., 2021. Deep transfer learningfor automated intestinal bleeding detection in capsule endoscopy imaging. Journal of Digital Imaging, 34, pp.d04-417.

[27] Ellahyani, A., Jaafari, I.E., Charfi, S. and Ansari, M.E., 2021. Detection of abnormalities in wireless capsule endoscopy based on extreme learning machine. Signal, Image and Video Processing, 15, pp.877-884.

[28] Xiao, Z. and Feng, L.N., 2020. A study on wireless capsule endoscopy for small intestinal lesions detection based on deep learning target detection. IEEE Access, 8, pp.a59017-159026.

[29] Sohag, M.H.A., 2020. Detection of intestinal bleeding in wireless capsule endoscopy using machine learningtechniques (Doctoral dissertation, University of Saskatchewan).

[30] Al Mamun, A., Em, P.P., Ghosh, T., Hossain, M.M., Hasan, M.G. and Sadeque, M.G., 2021. Bleeding recognition technique in wireless capsule endoscopy images using fuzzy logic and principal component analysis. International Journal of Electrical and Computer Engineering (IJECE), 11(3), pp.b688-

2695.

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

[32] Laiz, P., Vitrià, J., Gilabert, P., Wenzek, H., Malagelada, C., Watson, A.J. and Seguí, S., 2023. Anatomical landmarks localization for capsule endoscopy studies. Computerized Medical Imaging and Graphics, 108, p.a02243.

[33] Sharma, A., Kumar, R. and Garg, P., 2023. Deep learning- based prediction model for diagnosing gastrointestinal diseases using endoscopy images. International Journal of Medical Informatics, 177, p.a05142.

[34] Laiz, P., Vitrià, J., Gilabert, P., Wenzek, H., Malagelada, C., Watson, A.J. and Seguí, S., 2023. Anatomical landmarks localization for capsule endoscopy studies. Computerized Medical Imaging and Graphics, 108, p.a02243.

[35] Sharma, A., Kumar, R. and Garg, P., 2023. Deep learning- based prediction model for diagnosing gastrointestinal diseases using endoscopy images. International Journal of Medical Informatics, 177, p.a05142.

[36] Sadagopan, R., Ravi, S., Adithya, S.V. and Vivekanandhan, S., 2023. PolyEffNetV1: A CNN based colorectal polyp detection in colonoscopy images. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, *237*(3), pp.406-418.

[37] Bekkanti, A., Ganesan, S. and Satyala, N., 2023.DETECTION AND SEGMENTATION OF COLORECTAL POLYPS IN GASTROINTESTINAL COLONOSCOPY IMAGES USING MASK R-CNNFRAMEWORK. Journal of Data Acquisition and Processing, 38(2), p.1442.

[38] P. Shukla, R. Aluvalu, S. Gite and U. Maheswari, Computer Vision: Applications of Visual AI and Image Processing. Vol. 15. 2023: Walter de Gruyter GmbH & Co KG.

[39] Deshpande, N. M., Gite, S., Pradhan, B., Assiri, M. E. (2022). Explainable Artificial Intelligence–A New Step towards the Trust in Medical Diagnosis with AI Frameworks: A Review. CMES-Computer Modeling in Engineering & Sciences, 133(3), 843–872.

[40] Nilkanth Mukund Deshpande, Shilpa Gite, Biswajeet Pradhan, Ketan Kotecha, Abdullah Alamri. Improved Otsu and Kapur approach for white blood cells segmentation based on LebTLBO optimization for the detection of Leukemia[J]. Mathematical Biosciences and Engineering, 2022, 19(2): 1970-2001. doi: 10.3934/mbe.2022093

[41] Gite, S., Mishra, A. & Kotecha, K. Enhanced lung image segmentation using deep learning. Neural Comput & Applic 35, 22839–22853 (2023). https://doi.org/10.1007/s00521-021-06719-8

[42] Mishra, A.; Dharahas, G.; Gite, S.; Kotecha, K.; Koundal, D.; Zaguia, A.; Kaur, M.; Lee, H.-N. ECG Data Analysis with Denoising Approach and Customized CNNs. *Sensors* 2022, *22*, 1928. https://doi.org/10.3390/s2205.

[43] Tripathi, R.C., Gupta, P., Anand, R., Jayashankar, R.J., Mohanty, A., Michael, G., Dhabliya, D. Application of information technology law in India on IoT/IoE with image processing (2023) Handbook of Research on Thrust Technologies? Effect on Image Processing, pp. 135-150.