An Advanced Diagnostic ColoRectalCADx Utilizes CNN And Unsupervised Visual Explanations to Discover Malignancies

[1]Akella S Narasimha Raju, [2]Kayalvizhi Jayavel, [3]T. Rajalakshmi,
[1 2]Department of Networking and Communications, School of Computing
[3]Department of Electronics and Communication Engineering, School of Electrical and Electronics Engineering
SRM Institute of Science and Technology, Kattankulathur, 603203 Chennai
Corresponding author(s). Email (s): ar9488@srmist.edu.in,
Contributing authors: kayalvij@srmist.edu.in, rajalakt@srmist.edu.in,

Abstract:

Colorectal cancer (CRC) is one of the most lethal kinds of cancer, so early detection is critical. Three datasets, CNN transfer learning with discrete wavelet transform (DWT), discrete cosine transform (DCT), and support vector machines (SVM) were used to find CRC. In these instances, a quick and precise visual diagnosis of polyps is need in the current scenario. The proposed process involves four distinct phases. First and foremost, convolutional neural networks (CNN) are developed to test the efficacy of the model. Further, a transfer learning approach was incorporated using SVM and LSTM. Using the K-means technique, a visual explanation is finally presented. This system works with the balanced Hyper Kvasir and Mixed Dataset, which are made up of CVC Clinic DB, Kvasir2, and Hyper Kvasir. The system is called "ColoRectalCADx." The convolutional neural network (CNN) models ResNet50V2, DenseNet-201, VGG-16, and RDV-22. The system achieved the highest accuracy with CNN DesnseNet-201 in Hyper Kvasir (98.92% training, 98.91% testing, 93.62% SVM training, and 95.87% SVM tests). CNN DesnstNet-201 also achieved a highest accuracy with the mixed dataset (98.91% training, 96.13% testing, 95.41% SVM training, and 94.86% SVM testing). The process involved three phases namely individual CNN, combination of CNN with SVM and combination of CNN, LSTM and SVM. After three phases of the system, across both datasets, the CNN+SVM+LSTM combination was proved to be the most effective. Finally, the unsupervised K-means learning algorithm extracts the location of any cancerous polyps and upon classification using SVM classifier resulted with an accuracy of 80%. The K-means algorithm, which uses segmented images as input, accurately predicts the sites of tumors in colorectal cancer patients.

Keywords: Colorectal Cancer (CRC), Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Long Short-Term Memory (LSTM), K-means clustering.

I. Introduction

Taking care of one's own physical and mental health should be an absolute top priority for every human being on our globe. Many different diseases and mental problems may affect humans. Several distinct illnesses have the potential to damage one or more organ systems in the body. These illnesses have their origins in the fact that these people inhabit a wide range of settings, each of which plays a role in the spread of the disease. [1]. These environments each play a role in the transmission of the disease. One's lifestyle, including dietary habits and negative actions, plays a prominent part in the genesis of medical problems. Cancer has the highest mortality rate, which causes the most potential damage when compared to any sickness on Earth. This fatal disease may invade organs' blood cells and consume their tissues, causing extensive harm. [2]. The many manifestations of cancer are often categorized by the organs they first occur in. The fatality rate attributable to colorectal cancer ranks third overall. This condition begins in the rectum and subsequently spreads to the rest of the colon. Over sixty percent of colorectal cancer cases worldwide are attributed to industrialized nations. An estimated 608,000 persons die each year because of CRC-related conditions. In India, the annual incidence rate (AIR) for colon cancer is 4.4 per 100,000 males, whereas the AIR for rectal cancer is 4.1 per 100,000 males. The incidence of colon cancer in women is 3.9 per 100,000 individuals. [3].

More than 1.8 million new cases of colorectal cancer and more than 880,000 deaths were reported by the World Health Organization in 2018. (WHO) [4]. As the disease progresses it gets increasingly difficult to treat, so early discovery is crucial for effective treatment and a better prognosis. Several deficiencies in existing diagnostic approaches lead to diagnostic delays and poor patient outcomes; these deficiencies must be addressed. Early identification of colorectal cancer may have a considerable effect on survival rates, in addition to enhancing diagnostic accuracy. According to the American Cancer Society [5], people with locally contained colorectal cancer have a 5-year survival rate of more than 90 percent, but those with metastatic illness that has spread to other organs have a 5-year survival rate of just 14 percent [6]. Consequently, there may be a direct correlation between the efficacy of therapy and the efficiency of early diagnosis. If a problem is detected early, less intrusive treatment alternatives may be available. In the case of cancer, for example, early detection might mean the difference between a curative operation requiring extensive surgery and a less invasive treatment requiring just a tiny incision. Regular screening and early diagnosis are required to discover cancer at an earlier, more curable stage, when there may be less intrusive treatment alternatives.

The large, five-foot-and-two-inches-long large intestine is essential for digesting. The ascending, transverse, descending, and sigmoid colons all contribute to digestion, although in various ways. Some polyps may grow into cancer, even though CRC often begins as benign polyps. [7]. The growth of possibly cancer-causing polyps, also known as malignant polyps, occurs at an exponential rate. There may be a connection between this malignancy and the lifestyle choices of the patient, including their diet, mental health, and sleeping patterns. Four phases comprise the CRC: in situ, regional, national, and international. [8, 9]. Malignant polyps may first develop if cancerous cells make it through the mucosal lining of the colon, reach a critical mass where they try to break through the rectal wall and spread further. Blood in the faeces or a bleeding rectal area are both typical symptoms of the disease. Colorectal cancer and other gastrointestinal diseases may be detected by a number of different colon examinations. [10]. Due to the thorough, inch-by-inch examination of the whole colon by a gastroenterologist, colonoscopy is regarded as the gold standard for colon cancer screening. [11]. As the colonoscope, this technique makes use of a narrow, illuminated tube fitted with a high-definition front camera. All five meters of the colon may be captured on film or still images using this equipment, which is inserted into the rectum. Data from all four sections of the colon is sent to a personal computer (PC) at regular intervals for analysis. [12, 13]. The radiologist can tell how far the cancer has spread by examining the video and still images.

Colonoscopy data, including video and still images, is utilized in machine learning (ML) and deep learning (DL) algorithms. Through the use of computational methods, diagnosis may now occur much sooner and with much greater speed. Modern cutting-edge developments like automation and robotics owe a great deal to AI. [14]. AI specialists with expertise in machine learning (ML) and deep learning have contributed to technological advances (DL). The advancements in biomedical engineering and computing aid in the identification of colorectal cancer (CRC). The early identification of cancer requires machine learning methods. For the DL of this experiment, convolutional neural networks (CNN) were used [15]. Functional models in the CADx system, such as CNNs, are crucial to this research. Multiple iterations of CADx creation are needed for accurate diagnosis of malignant neoplasms that are likely to spread quickly. The ColoRectalCADx system is utilised to detect colon cancer in this investigation. The purpose is to define "ColoRectalCADx" and describe how it works.

Organization of the paper

The remaining sections of the research are as follows: In Section 2, a comprehensive literature review is presented, while in Section 3, materials and methods are outlined and explained. Results are presented and discussed in Section 4, and the research is concluded in Section 5.

II. Literature Survey

Even though gastroenterologists are competent at identifying worrisome areas in the colon, there has been a significant increase in the use of Computer-Aided Detection (CADx) systems to aid in the localization of Colorectal Polyps from colonoscopy images. The development of CADx systems that allow real-time detection of colorectal polyps has led to the emergence of deep learning approaches in medical imaging and the availability of enormous datasets, both public and private. Several studies have recommended the use of end-to-end Convolutional Neural Networks to overcome the difficulties of identifying colorectal polyps (CNNs). These convolutional neural networks (CNNs) are trained to discern between healthy intestinal tissue and polyps using annotated medical images. These technologies assist gastroenterologists in establishing more accurate diagnosis by finding and identifying potentially troublesome areas of the colon. Ultimately, developments in CADx systems have the potential to enhance the precision and efficiency of colorectal polyp identification, reduce the workload of gastroenterologists, and increase their productivity. Deep learning algorithms used in medical imaging have the potential to improve both diagnosis and therapy in the future. According to Banik et al. [16], a reliable computer-aided detection (CAD) system is required since manual colorectal cancer screening is known to be error-prone. A revolutionary CAD system using deep learning and polyp identification techniques is offered. PolypNet, a CNN-DT-WpCNN pooling approach for recognising polyps, is presented. This model generates a segmented mask with an abundance of brilliant detail outside of the polyp zone. PolypNet is evaluated using the CVC Clinic DB dataset, and the authors claim a dice score of 0.839. It should be emphasised, however, that this evaluation is limited to a single dataset, and a more comprehensive evaluation on many datasets may be necessary to properly comprehend the system's performance. Overall, the strategy reported by the authors has tremendous potential for enhancing the accuracy of polyp detection in colorectal cancer screening.

Muthu Subhash Kavitha et al. [17], provide a deep learning technique that uses convolutional neural networks (CNNs) to detect colorectal cancer in colonoscopy pictures. The authors use end-to-end learning, hybrid learning, and explainable AI techniques to identify and localise colorectal polyps. Although the article does a decent job of describing the theoretical basis, it lacks experimental evidence to support the authors' statements. Nevertheless, the proposed technique has the potential to improve colorectal cancer detection, which is a crucial step in reducing the disease's incidence and mortality rates. The availability of colonoscopy imaging data increases the possibility for developing cutting-edge deep learning models to aid in the early diagnosis and treatment of colorectal cancer. Cao et al. [18] propose combining Raman spectra

with deep learning to enhance the diagnostic precision of colorectal cancer. The authors classified Raman spectroscopy data from 26 people with colorectal cancer using a 1D-ResNet CNN. The experimental results confirmed the efficacy of the proposed approach by demonstrating that it properly detected 98.5% of instances of colorectal cancer. This model was only examined using colorectal Raman spectroscopy data, hence its effectiveness on other types of image datasets is unknown. In addition, the collection of Raman spectroscopy data is tedious and time-consuming, which may limit its use in clinical practise. However, the findings of this study open the way for the future development of diagnostic systems based on deep learning for the early diagnosis of colorectal cancer.

Siva Nagaraju et al. [19] proposed an AI-based technique for the detection of colorectal cancer. They use advanced deep learning techniques to improve machine learning for cancer detection. Before categorising a collection of colonoscopy pictures, the authors employ a GAN cycle to enhance and rectify image imbalance. We employ CNN MobleNet V2 and MNIST histopathological data to classify images. The proposed method for identifying colorectal cancer is 99.9 percent accurate. Detection of colorectal cancer is done utilizing a single CNS and visual enhancement using GAN. The classification of colonoscopy images is the core emphasis of Ozturik et al.'s [20] research, which focuses mostly on CNN models. Combining CNN-based classification with LSTM on a dataset comprising between 2,000 and 6,000 images is the focus of this study. All of AlexNet, GoogleNet, and ResNet's pooling layers are turned to LSTM layers and then utilised to classify data. Despite the availability of several colonoscopy datasets, none were used as references in this study. In this experiment, three distinct convolutional neural network (CNN) models are compared and shown to be helpful for diagnosing colorectal cancer.

Yan Wang et al. [21] reported research on transfer learning in CNNs and its application to colonoscopy image datasets. The proposed deep learning architecture classified colonoscopy images using a trained network that transfers learning to CNN layers. A variety of CNN models, including ResNet50, ResNet101, VGG16, and VGG19, were used to evaluate the performance of the suggested approach. However, no attempts were made to analyse or include transfer learning, and different CNN algorithms were not used with other colonoscopy datasets. It is important to note that colorectal diseases are not discussed in detail here. Meryem Souaidi [22] demonstrated a novel device known as the MP-FSSD for locating deep polyps in wireless capsule endoscopy (WCE) and colonoscopy images. The study used VGG-16 backbones for polyp identification but made no mention of the usage of alternative CNN models. The performance of polyp detection was tested using WCE and CVC clinic DB datasets. However, the focus of this research is confined to the detection of polyps; no other colorectal illnesses or anomalies are investigated. Guo et al. [23] proposed a unique approach to solve challenges in segmenting colonoscopy polyps, such as imprecise borders and similarities to surrounding tissue. To evaluate the accuracy of the malignant polyp, a transformer-based detection network called Exploration (UnX) was employed to identify the locations of uncertainty inside the polyps. This study examined a novel UnX approach on five datasets; hence, the results are delayed. Although dependable, the system's intricacy may discourage typical users. The study does not address the issue of the colonoscopy imaging system's unequal colour distribution.

Hence in the current scenario there is a need for early detection and diagnosis of colorectal cancer. The aim and objectives of the proposed study is as follows:

i. The procedure for recognising colorectal cancer consist of four stages.
ii. The first step in diagnosing cancer is to classify images obtained from a colonoscopy using three CNN architecture. The system provides a single CNN integrated model and generates the model that might be the most successful by utilising datasets.
iii. Individual and integrated CNN models exhibit transfer learning, simplifying the implementation of SVM.
iv. Continuing ahead, LSTM transfer learning shall be used. The best model for diagnosing cancer may be found using this classification.
v. Visual explanation of K-means clustering for identifying malignant cancer polyps.

III. Materials and Methods

The purpose of this research is to develop a system that, through a combination of automated analysis and human review of suspicious lesions, can reliably classify colorectal ailment as cancer. This problem has a dedicated CADx solution in the form of the ColoRectalCADx system. The proposed ColoRectalCADx system's data flow is shown in Figure 1.

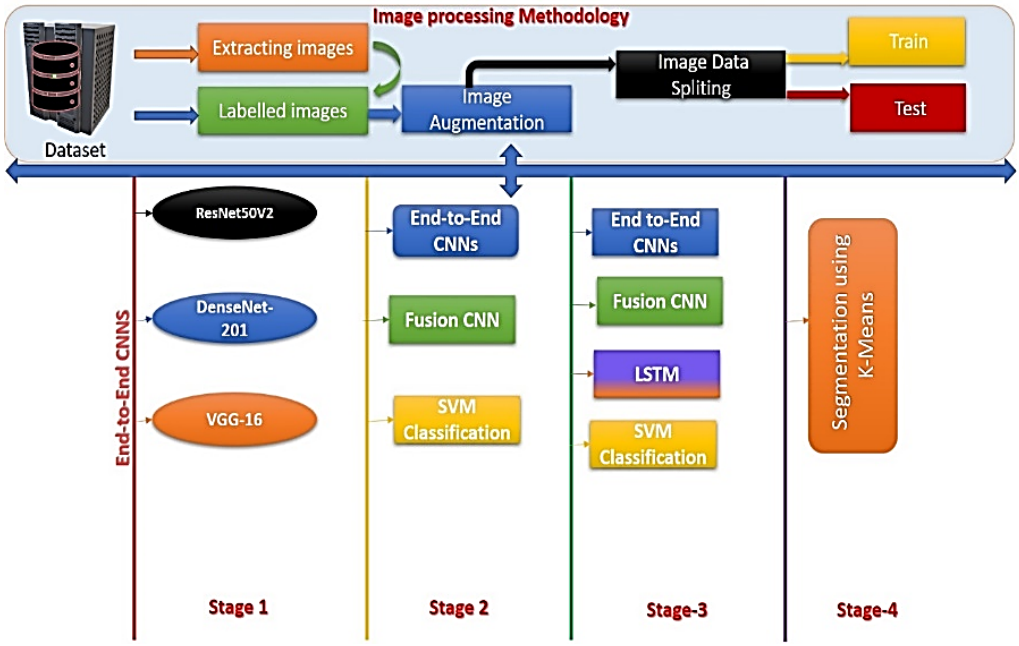


Figure 1: ColoRectalCADx' s Flow diagram

From the local server archives, the CVC Clinic DB, Kvasir2, and Hyper Kvasir datasets are retrieved and placed in their respective directories. It is clear from the titles that these documents are part of classes 2, 8, and 23. Multiple versions of the modified datasets exist. Then, these datasets are made available in two distinct varieties: the 9-class balanced dataset and the 10-class mixed dataset. Folders containing colonoscopy images are organized and labelled for storage. These folders contain the dataset's images, each of which has a different pixel dimension. Images retrieved from CNN experiments often need to be scaled down to a size of 224 by 224 pixels. Images may be resized by a factor of 0.4, rotated by 15 degrees, and flipped horizontally using these improved methods. Image augmentation improves the rendered image. Also, the input for the 10 and 9 class datasets are split 70:30 across the training and testing sets.

- Proposed intelligent and automated CADx system, ColoRectalCADx, uses 4 steps to find and identify polyps that could be signs of colorectal cancer.
- In the first step, the features of the three CNNs are separated and put into groups.
- Individual and integrated CNNs are also liable. SVM helps to classify CNNs. SVMs are useful for multiclass and high-dimensional situations. The kernel function moves the feature space of a dataset to a new domain. It is used for CNN's enormous DL feature dimension.
- Extracting temporal and spatial information is the focus of the third stage, which combines the transfer learning of several standalone and integrated CNNs with LSTM and SVM multi-class classification.
- Through every level, the ColoRectalCADx model is validated for accuracy and other performance parameters and compared to determine the best one.
- The fourth step uses K-means clustering to identify the precise cancerous polyp.

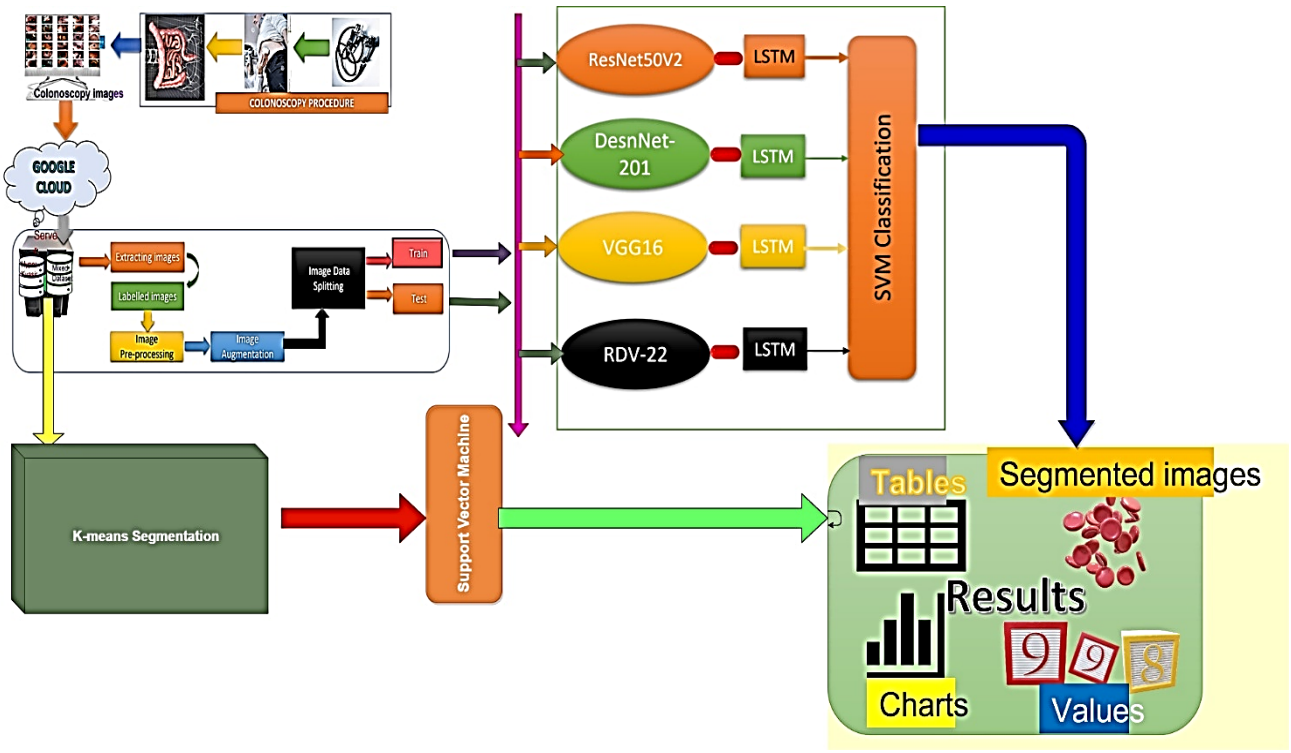
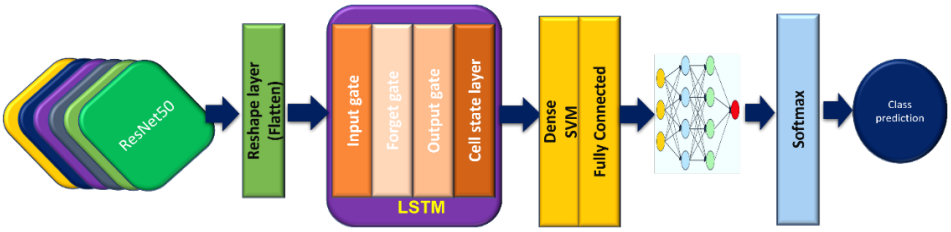
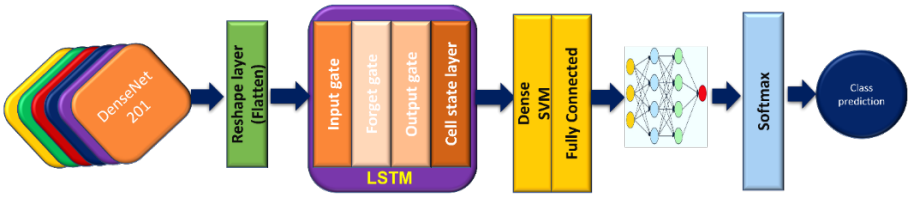


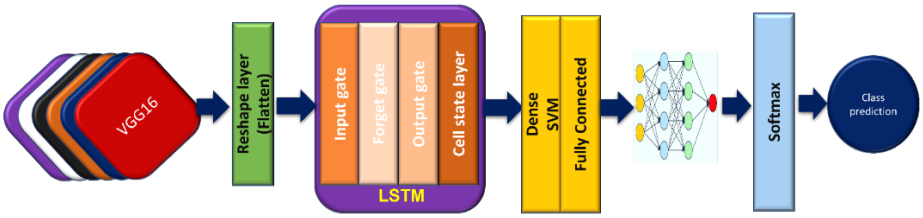
Figure 2: The Block Diagram of the ColoRectalCADx system



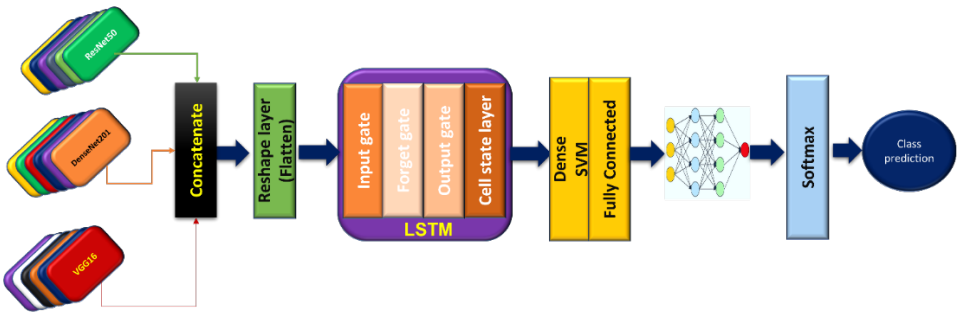
(a)



(b)



(c)



(d)

Figure 3: The Architecture for CNN+LSTM+SVM combinations for (a) ResNet-50V2V2+LSTM+SVM (b) DenseNet-201+LSTM+SVM (c) VGG-16+LSTM+SVM and (d) RDV-22+LSTM+SVM

Figure 2 and 3 shows the block diagram and architectures of the ColoRectalCADx system.

Colonoscopy Procedure

Gastrointestinal disorders may be diagnosed by a colonoscopy. The colon, which is 5 feet and 2 inches in length, is examined with the use of a colonoscope. This compact equipment is equipped with an incredible camera. This device is implanted in the colon and it records video and images at regular intervals. The gastroenterologist is in position to conduct this procedure, while the radiologist is responsible for analysing the results [24, 25]. The resulting images shall be used for cancer investigation. Once the radiologist's post-colonoscopy images have been saved to computer system, they may be categorised according to the different parts of the colon using labelled images, also known as datasets. After that, the images are stored on both the high-end local servers and on the Google cloud, with the necessary account information associated [26]. Google Drive is now storing the images. In order to discover the findings, extensive research operations are carried out with the support of Google's co-laboratory, which is abbreviated as "Co-Lab" [27].

Datasets

The coloscopy medical motion image collections are available to everyone with an internet access. Some sources provide free downloads of images from the CVC Clinic DB, Kvasir2, and Hyper Kvasir datasets. Labels range from two in the smallest dataset to twenty-three in the largest. The given system includes a new mixed dataset consisting of all 24 class labels, which was generated by merging the CVC Clinic DB, the Kvasir2 images, and the Hyper Kvasir images.

- The CVC Clinic DB dataset included 2 classes such as labelled polyps (818) and non-polyp (822) with a total of 1640 images [28].
- The Kvasir2 dataset includes eight labelled classes, labelled as Dyed-Lifted Polyps (1000), Dyed-Resection Margins (1000), Esophagitis (1000), Normal-cecum (1000), Normal-cecum (1000), Normal-z-line (1000), Polyps, and Ulcerative Colitis (1000) with a total of 8000 images [29].
- Lower gastrointestinal tract (GI) and upper gastrointestinal tract (GI) datasets are included in Hyper kvasir, with a total of 23 subclasses within these two major categories. They are named as barrettes (42), barrettes-short-segment (53), bbps-0-1(653), bbps-2-3 (1148), cecum (1009), dyed-lifted-polyps (1003), dyed-resection margins (990), esophagitis-a (404), esophagitis-b-d (260), haemorrhoids (10), ileum (9), impacted-stool (132), polyps (1028), pylorus (1000), retroflex-rectum (391), retroflex-stomach (765), ulcerative-colitis-grade-0-1(35), ulcerative-colitis-grade-1(201), ulcerative-colitis-grade-1-2 (11), ulcerative-colitis-grade-2 (443), ulcerative-colitis-grade-2-3 (28), ulcerative-colitis-grade-3 (133), z-line (933), totalling 10672 images [30].
- The Mixed dataset is a novel dataset that combines the aforementioned datasets, and these classes of datasets are further classified and labelled into 24 classes. These 24 classes are named as barrettes (42), barrettes-short-segment (53), bbps-0-1(653), bbps-2-3 (1148), cecum (2009), dyed-lifted-polyps (2003), dyed-resection margins (1990), esophagitis-a (1404), esophagitis-b-d (260), haemorrhoids (10), ileum (9), impacted-stool (132), Non-Polyps (818) polyps (2150), pylorus (2000), retroflex-rectum (391), retroflex-stomach (765), ulcerative-colitis-grade-0-1(1035), ulcerative-colitis-grade-1(201), ulcerative-colitis-grade-1-2 (11), ulcerative-colitis-grade-2 (443), ulcerative-colitis-grade-2-3 (28), ulcerative-colitis-grade-3 (133), z-line (1933), with a total of 19,621 images.

The proposed mixed dataset has 19,621 images from 24 unique classifications. 14 of these classes are particularly unbalanced due to difficulties with the classification process, which may impede the efficacy of image classification models. The Hyper Kvasir dataset comprises 10,672 classes across 23 classes, while the varied sample contains just 14 classes in common. The performance of this model's classification process worsens because of this high bias toward these 14 classes. In investigations conducted on the 23 and 24 class mixed and Hyper Kvasir datasets, a substantial level of misclassification was seen, resulting in a decrease in accuracy. To concentrate on the most stable classes and images, the authors of this study intended to exclude 14 unstable classes. Therefore, the authors of this research were able to focus on finding and analysing

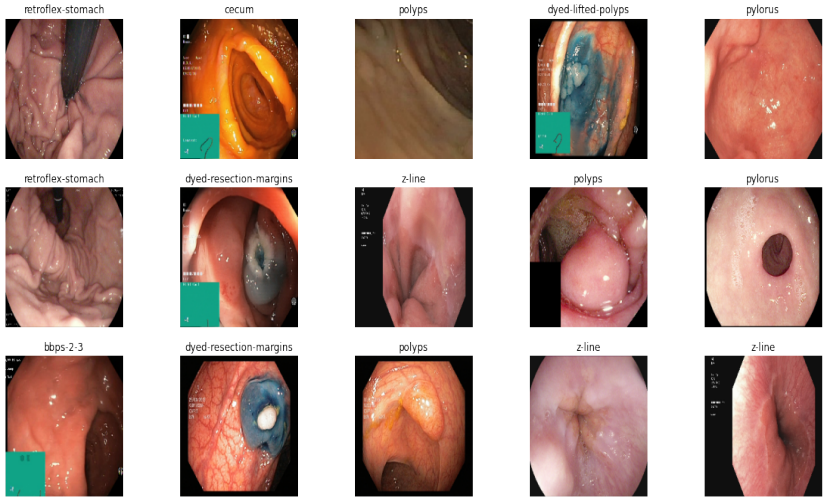
the best quality data. By focusing on the most consistent classes and images, the accuracy and performance of the system were improved. This strategy improved the system's capacity to meet the requirements of its users.

To investigate this issue, the authors of this article ran tests using the 10 most consistent classes from the mixed dataset. The 9 classes of the Hyper Kvasir dataset were the most balanced. The purpose was to overcome the dataset's constraints and increase image classification performance and accuracy. After establishing balanced classes, the mixed-dataset images collection for this classification experiment has 14287 images with 10 top balanced classes and 8529 images with 9 top balanced classes. Table 1 shows the distribution of datasets between the mixed and Hyper Kvasir datasets.

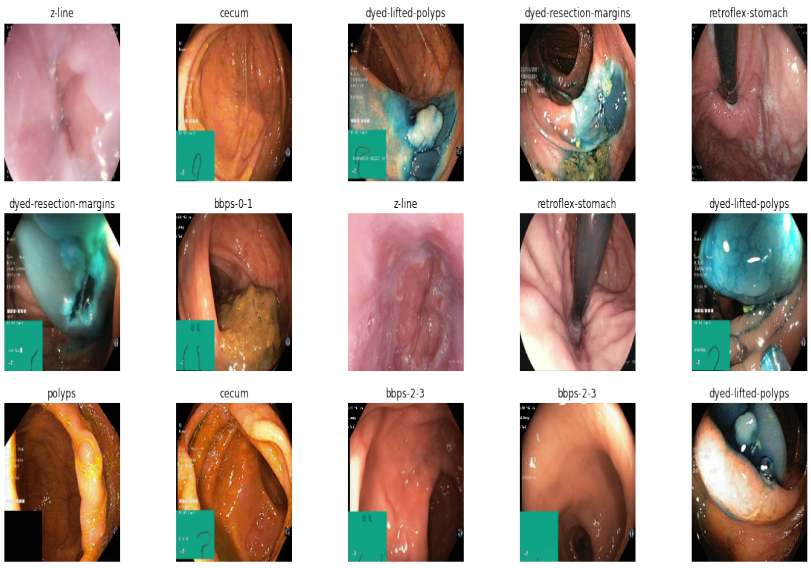
Table.1 Balanced Mixed dataset and Hyper Kvasir dataset class labels

Balanced Classes	Balanced Mixed Dataset	Images	Classes	Balanced Hyper Kvasir	Images
0	**bbps-0-1**	653	**0**	**bbps-0-1**	653
1	**bbps-2-3**	1148	**1**	**bbps-2-3**	1148
2	**cecum**	2009	**2**	**cecum**	1009
3	**dyed-lifted-polyps**	2003	**3**	**dyed-lifted-polyps**	1003
4	**dyed-resection-margins**	1990	**4**	**dyed-resection-margins**	990
5	**Non-Polyps**	818	**5**	**polyps**	1028
6	**polyps**	818	**6**	**pylorus**	1000
7	**pylorus**	2150	**7**	**retroflex-stomach**	765
8	**retroflex-stomach**	765	**8**	**z-line**	933
9	**z-line**	1933			
	Total images in dataset	14287		**Total images in dataset**	8529

The mixed and Hyper Kvasir dataset of the typical sample images is represented in Figure 4



(a)



(b)

Figure 4. Sample Datasets: (a) Mixed (b) Hyper Kvasir dataset

Image Pre-Processing

The images used in this experiment are divided into groups, and each group is provided with a unique amount of data. There are 10 primary classes that have 14,287 images, while the other 9 classes contain 8,529 images.

To experiment efficiency, each image has been scaled to a resolution of 224 by 224 pixels, but this results in varying pixel sizes for each image. In addition to being extensively supported by convolutional models, this resolution is also commonly used in image classification tasks. By scaling the images to the same dimensions throughout the dataset, model training may be accelerated and improved in accuracy. A pre-processing technique known as "image augmentation" is used to further enhance the image quality. These alterations to the proportions, zooming (by a factor of 0.2), rotation (by a factor of 15 degrees), and horizontal flip of the images enhance their visual appeal [31].

The dataset is then divided 70:30 between a training set and a test set after the completion of the augmentation procedure. Keeping a larger portion of the data for training the model while leaving enough for testing and assessment is a usual practise in machine learning, therefore this ratio was chosen appropriately. The images in the training set are used to train convolutional neural networks how to recognise patterns and make accurate predictions, whilst the images in the test set are used to evaluate the models' performance on unique data. Using this splitting method, overfitting of the model to the training data may be avoided and the model's performance on unique data can be assessed more precisely. Overall, this technique gives more accurate and reliable results when evaluating the usefulness of convolutional neural networks for image classification, and it allows for a more complete examination of the dataset.

Convolutional Neural Networks (CNN)

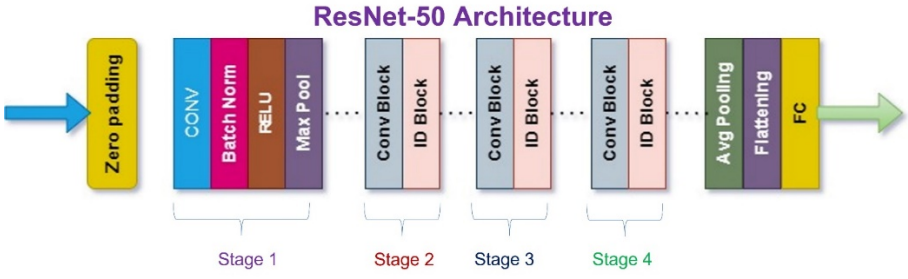
CNN models have been demonstrated to be useful in addressing classification challenges in healthcare and deep learning applications. When attempting to classify images, CNNs seem to be quite favourable. The ColoRectalCADx system and CNNs play an important part in the present investigating protocol for the diagnosis of colorectal cancer. A considerable number of images from both the mixed data set and Hyper Kvasir are classified using CNN [32]. In order to determine how well an image may be classified, convolutional neural networks (CNNs) take in the input images and run them through a number of operations, such as convolution, pooling, activation, dropout, and finally a full connection to the underlying neural networks [33]. Three alternative convolutional neural network models (CNNs)—ResNet-50V2, DenseNet-201, and VGG-16—are tested on the ColoRectalCADx system [33, 34]. Three top-tier CNN models were evaluated, and the one with the highest accuracy was chosen as the optimal model for detecting colorectal cancer. The most accurate method for spotting colorectal cancer is now being tested using a combination of the three best individual models and one integrated model. The RDV-22 integrated model is the result of combining ResNet50, DenseNet-201, and VGG16. We evaluate the performance of this integrated model in comparison to three other models [35].

Integrated models combine individual models. In ColoRectalCADx, four models are analysed, and the most relevant are merged into a colorectal cancer model. Figure 5 shows the three CNN models' (ResNet50, DenseNet-201, and VGG16) architectures Including ResNet50 [36] is a 50-layer deep neural network consisting of convolutional layers, pooling layers, and fully connected layers. It also features residual connections, allowing for the training of more complicated networks and the learning of residual functions.

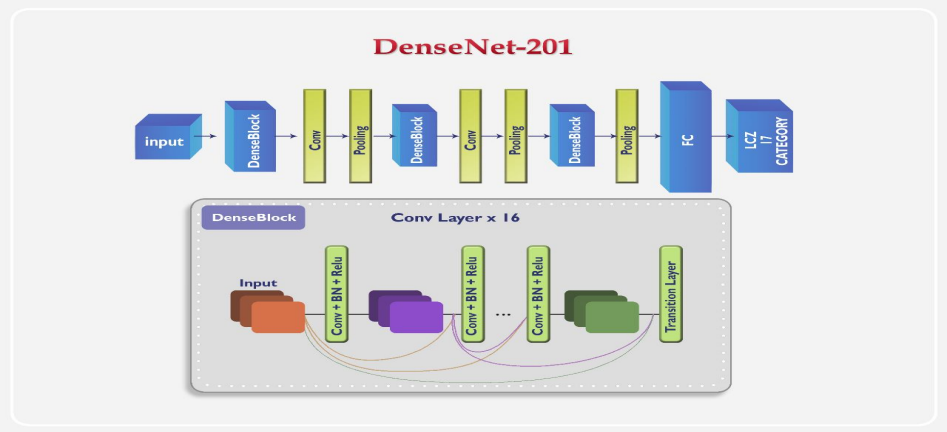
Its architecture, which is based on DenseNet-201 [37] [38] , is famous for the many connections between its various layers. Feature maps are sent from one layer to the next in DenseNet-201, with each layer getting feature maps from all layers below it. By enhancing gradient flow and reusing features, dense connections are produced that enhance accuracy with fewer parameters. The building comprises of dense blocks, each of which consists of a complicated network of linked layers. To reduce the spatial dimensions of the feature maps, they are separated into dense blocks and connected by transition layers including a convolutional layer, a pooling layer, and a batch normalisation layer. The network outputs the predicted class probabilities at the final layer, which consists of a global average pooling layer followed by a fully connected layer with a SoftMax activation function. There are a total of 16 layers in the VGG16 [39] architecture, 13 of which are convolutional and 3 of which are fully connected. Small convolutional filters (3x3) with stride = 1 and equivalent padding form the foundation of the architecture of VGG16, which is subsequently followed by max pooling layers (2x2) with stride = 2. Consistently repeating this pattern of convolutional and pooling layers produces a deep network with several trainable parameters. VGG16 employs completely connected layers with a total of 4096, 4096, and 1000 units after the convolutional layers. The predicted class probabilities for each input image are then generated by the last fully connected layer, which employs a SoftMax activation function.

To mitigate the issue of vanishing gradients and increase speed, the design also includes skip links that avoid some levels. The image datasets used by CNN are balanced between mixed and Hyper Kvasir. These images' features are extracted using convolution. Image features also send image data to the maximum pooling layer. Following that, with a

fully connected and constructed neural network, ReLU activation values are supplied to the neural network. SoftMax is the last layer used to classify multi-class images. It is possible to determine the state of the imaging polyp [40].



(a)



(b)

(c)

Figure 5: CNN Architecture (a) ResNet-50V2 (b) DenseNet-201 (c) VGG-16

When it comes to the classification of medical colonoscopy images, both the independent model and the integrated model have their merits. This deep learning approach is extremely helpful in the identification of colorectal cancer. This includes the main subjective understanding essential for swift and precise cancer detection. Cutting down on the amount of the connections between the layers closest to the input and the layers closest to the output has been shown to improve the depth, accuracy, and learning efficiency of CNNs. Table 2 lists all of the suggested CNNs' parameters, including those for integration and those for each individual CNN, as well as the total number of training parameters that are effective. [35, 41].

Table.2 Number of Parameters of CNNs

CNN Architecture Models	Introduced year	Total params	Trainable params	Non-trainable params	Layers
ResNet-50V2 [36]	2016	2,59,33,975	23,69,175	2,35,64,800	50
DenseNet-201 [37] [38]	2018	1,94,29,463	11,07,479	1,83,21,984	201
VGG-16 [39]	2014	1,53,14,391	5,99,703	1,47,14,688	16
RDV-22	2022	60,675,604	4,074,132	56,601,472	267

ColoRectalCADx system in CNN Individual mode with ResNet50 and wrap up with the integrated model RDV-22. Experiments ranging from using the most basic convolutional neural networks (CNNs) to the most complex CNNs possible are being conducted in these models. All the experimental studies of individual and integrated CNNs have focused on feature mining. Filtering results might be limited to just one image dataset if CNN maps are used. CNN layers are swapped using transfer learning. The term "transfer learning" refers to the process of applying parts of one machine learning model to a different issue [42]. Because of this, the model may be adjusted to solve a specific issue. Transfer learning occurs when a model is moved from one dataset to another [43]. For CNN object recognition, only the last convolutional layer makes a prediction.

Long Short-Term Memory (LSTM)

The computational architecture of deep learning includes LSTM neural networks. In this DL, a recurrent neural network processes the time series inputs. The LSTM may be used to solve time series prediction problems in applications such as machine translation, speech recognition, etc. The acronym LSTM describes a subclass of recurrent neural network (RNN) architecture known as Long Short-Term Memory. There are primarily four levels:

- Input gate layer: This layer makes the decision on what parts of the input will become part of the cell's state.
- Forget gate layer: In this stage, it is determined whether bits of history from the prior cell state may safely be forgotten.
- Output gate layer: This layer is responsible for determining the output of some elements of the cell's state.
- Cell state layer: This layer is responsible for maintaining the network's memory and for moving information forward in time.

In addition, a Peephole connection layer, the fifth possible layer, may also be present. To make decisions about the input, forget, and output gates, the LSTM uses this layer to sneak a look at the state of the cells.

In summary, an LSTM cell takes in data as input, processes it along with what was previously stored in memory, and then produces an output and a new memory state. To do classification or regression, the output is often fed into a dense layer, where the memory state is passed on to the next cell in the sequence [44].

Support Vector Machines (SVM)

The SVM is one of the cornerstone ideas in machine learning, and a common supervised learning approach. Classification, regression, and the identification of outliers are just some of the applications where this method shines [45, 46]. This technique produces the hyperplane used in data classification. The objective is to discover a hyperplane that maximally constrains the space of possible media vectors inside the given data. The SVM returns a hyperplane with the highest possible margin. It produces hyperplanes for improved classification. Data classification, whether it be into two (binary) or more groupings, is a major topic (multi-class) [47, 48]. The ColoRectalCADx system employs four unique CNN investigations in total. The CNN's convolutional layers are used to extract significant visual features from the Hyper Kvasir and Mixed datasets. Typically, these layers consist of a collection of filters that are applied to the input images and convolved with the pixels to generate feature maps. Depending on the individual assignment, the size of the filters and the number of feature maps are hyperparameters that may be modified. The feature maps produced by the convolutional layers are used to represent the features that have been learnt from the input images. Later layers of the CNN learn more complex features that capture higher-level notions like forms and object pieces, while the early levels learn more fundamental features like edges and textures.

The Long Short-Term Memory (LSTM) layer is a recurrent neural network (RNN) that processes data in sequences. The LSTM layer processes the CNN's output and learns the sequential correlations between features for use in image classification. Information is stored in the LSTM layer's cells over time, and the cells' inputs and outputs are regulated by gates. A series of feature maps generated by the CNN serves as the input to the LSTM layer. The LSTM layer iteratively analyses feature maps, where each map represents a distinct region of the input images. To capture the semantic meaning of the input sequence, the LSTM layer updates the current cell with data from prior cells. The SVM layer is a technique for machine learning used to classify data into distinct groups based on its learnt properties. In the context of image classification, the SVM layer receives as input the output of the LSTM layer and generates a probability distribution for each class. The SVM method uncovers a hyperplane that maximally divides the feature space into distinct classes. To predict the class label of incoming input data, a hyperplane is selected to minimise the classification error.As a final layer, the SoftMax output normalises the SVM layer's output into a probability distribution across the various classes. To assign a probability to each class, the SoftMax function uses the results from the SVM layer as input. The result is a legitimate probability distribution since the total of the probabilities for all classes equals 1.

Overall, this approach uses the strengths of both CNNs and LSTMs for input Hyper Kvasir and Mixed datasets image classification. The CNN extracts relevant features from the input images, while the LSTM processes these features to learn the sequential relationships in the data. The SVM then uses the learned features to classify the input image into different categories, and the output SoftMax layer provides a probability distribution over the different classes. The ColoRectalCADx CNN models are needed to apply the transfer learning idea to SVM for use in this cancer detection investigation. The linear activation function is used as the output layer's final activation function owing to the "kernel regularizer" option in SVM, which follows the l2 norm. The "SoftMax" activation feature of SVM is required for multi-class classification. For multi-class classification, the "squared hinge" loss is used. Therefore, the adjustments are implemented in the CNN's final layers, the linear SVM is represented, and the overall CNN's individual and integrated accuracy is maximised [49, 50].

K-Means Clustering of Colonoscopy Image Datasets:

Using K-means clustering-based segmentation, polyps, which may be precursors to colorectal cancer, may be detected and separated from colonoscopy images. Here's a description of how everything unfolds:

- *Acquiring an Image*: Obtaining a quality colonoscopy image is the first step. Subsequently, the image is pre-processed to remove any noise or artefacts that might possibly disrupt the study.

- *Segmenting Images*: Then, K-means clustering is utilised to build individual images segments. The approach works by clustering pixels with similar values together. Clusters in colonoscopy images may be identified by searching for certain colour, texture, or intensity variations.
- *Identifying a Specific Area of Interest*: During image segmentation, polyp-containing regions of interest (ROIs) may be chosen. To do this, threshold settings may be modified based on the image's luminance or texture.
- *Extracting Features*: The extracted features are then based on the discovered ROIs. The size, shape, and texture of the polyp, as well as its precise location in the colon, are all considerations that may be made.
- *Classification*: The discovered ROIs are then classified as polyps or non-polyps based on the extracted features. Machine learning algorithms can do this.

Segmenting colonoscopy images using K-means clustering with varying K values can be crucial for a number of reasons. Initially, the selection of K value can impact the precision and quality of segmentation. Using a range of K values, the optimal K value for a given dataset can be determined. This can ensure that the segmentation correctly identifies the location and extent of polyps in the images. In addition, K-means clustering can aid in enhancing the efficiency and velocity of image segmentation. By grouping pixels with similar features, the algorithm can reduce the overall difficulty of the segmentation task. This can reduce the required computational resources and accelerate the segmentation procedure. In conclusion, K-means clustering can provide a standardised and automated approach to segmentation, which can reduce inter-observer variability in the identification of polyps. By employing a consistent and reproducible segmentation algorithm, the polyp detection process can be made more reliable and accurate. Using K-means clustering to segment colonoscopy images with different K values can enhance the precision, efficiency, and reproducibility of polyp detection in medical imaging applications. Because of this lack of readily available labelled data, clustering algorithms like K-Means must rely on unsupervised learning approaches [51, 52]. Data similarity is measured to determine the existence of distinct classes or clusters within the information provided. Similarities within a given set of data points are stronger than those between groupings. It is usual practise to use the term "k" to denote the total number of groups in k-means clustering, a popular clustering method.

The working of K-means clustering is as follows:

A. In k-means clustering, the total number of clusters is denoted by the number k.
B. Randomly divide the information into the k-clusters.
C. Locate the centres of the clusters.
D. Find out how far off each data point is from the centres of the distinct clusters.
E. Based on how far away it is from that cluster, each data point should be moved to the cluster that is geographically closest to it.
F. Perform the computation once again to identify the new cluster centre.
G. Repeat steps D, E, and F until the data points no longer cause the clusters to move or until the allotted number of repeats is complete.

As k-value grows, the image becomes more distinct, revealing the K-means algorithm's ability to categorise an increasing number of colour classes or clusters. K-means clustering performs well when our data collection is modest in size. Objects inside images may be properly segmented, and improved results are also produced; however, when used with big datasets (a larger number of images), the time required to analyse all of the samples increases exponentially. ColoRectalCADx uses the K-means algorithm to analyse its datasets, which comprise the CVC clinic DB, Kvasir2, and Hyper Kvasir as a mixed dataset. Some of the well-known datasets include 14287 images from 10 different classes. Since K-means clustering is used to find cancerous polyps inside images and provide a precise cancer diagnosis, it has recently been used to these images. The method creates segmented images indicating the location of polyps in the colon. From these images, the size, shape, and likely malignancy of the polyps may be determined. Improved patient outcomes and a lower disease burden are the results of using this method of polyp detection and separation in the diagnosis and treatment of colorectal cancer. This technique may help clinicians find worrisome polyps or tumours and provide sharp diagnostic images that may be used to direct treatment when used to find colorectal cancer. Figure 6. shows the block diagram of K-means clustering part of the ColoRectalCADx system.

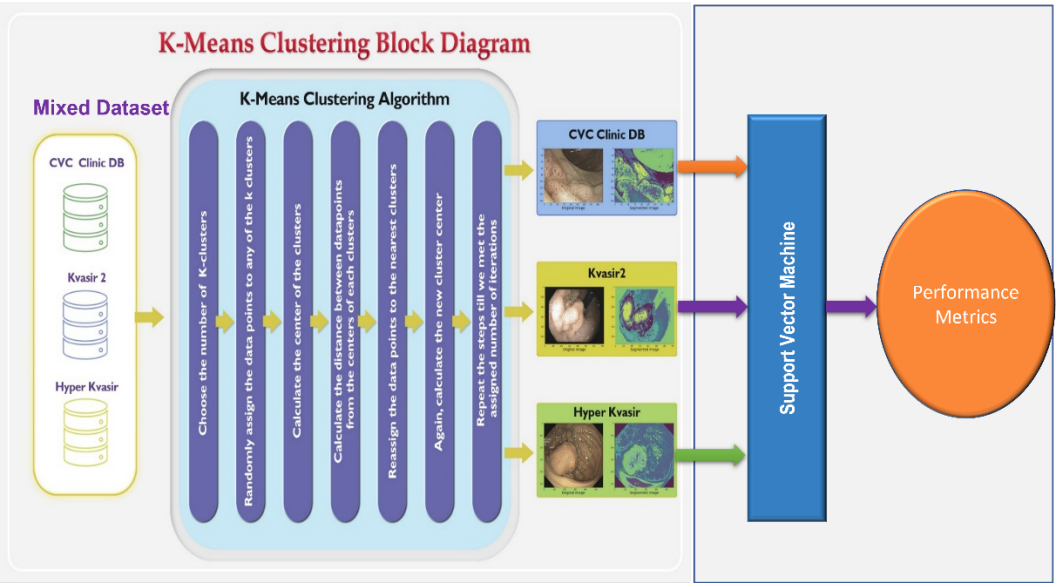


Figure 6: The K-means Clustering Block Diagram

IV. Results

For our ColoRectalCADx system, the datasets serve as the foundation. Alternatively, the number nine represents Hyper Kvasir with equally balanced classes, whereas the number ten represents mixed datasets. Previously, these records had 24 and 23 classes, respectively. To attain the highest possible degree of precision in the identification of colorectal cancer, most classes are unbalanced; thus, these classes are eliminated and balanced classes are created. Colonoscopy images from medical facilities are archived and made available for use in CNN training using any and all datasets. Additional information on the dataset is provided in Table 3.

Table 3: Train and Test split of Three Datasets

Datasets	Training set	Test sets	Total images
Mixed Datasets	10594	3693	14287
Hyper Kvasir	5970	2559	8,529

The four CNNs used in the experiments were ResNet-50V2, DenseNet-201, VGG-16, and the Integrated CNN RDV-22, as well as LSTM and SVM; the findings and hyper parameters were found as follows for the two datasets, Hyper Kvasir and Mixed. The results of the experiment were adjusted by means of hyperparameters. The specific hyperparameters of the ColoRectalCADx system are listed in Table 4. Utilizing hyperparameters, the experiment's work was modified. The entire ColoRectalCADx hyperparameters are shown in Table 4.

Table 4: Hyperparameters for ColoRectalCADx system

CNN hyperparameters	
ResNet50:	50 convolutional layers, filter sizes of 3x3, max pooling with size 2x2, ReLU activation function, dropout rate of 0.5, learning rate of 0.001, batch size of 16, epochs of 30, optimizer of Adam, momentum of 0.9
DenseNet201	201 convolutional layers, filter sizes of 3x3, max pooling with size 2x2, ReLU activation function, dropout rate of 0.5, learning rate of 0.001, batch size of 16, epochs of 30, optimizer of Adam, momentum of 0.9
VGG16	13 convolutional layers, filter sizes of 3x3, max pooling with size 2x2, ReLU activation function, dropout rate of 0.5, learning rate of 0.001, batch size of 16, epochs of 30, optimizer of Adam, momentum of 0.9
LSTM hyperparameters	
Number of LSTM units	64
Number of LSTM layers	1
Dropout rate	0.4
Learning rate	0.001
Batch size	16
SVM hyperparameters	

Type of kernel function	Linear function
Kernel parameters	Degree of polynomial kernel (2), gamma for radial basis function kernel (0.1)
Regularization parameter	C=10
Class weights	Balanced
Learning rate	0.001

ColoRectalCADx was used throughout the study at various points. At these points, the combined investigative background of the three individuals was most useful. By comparing the three CNNs' overall, individual, and consecutive performances, the best model was found. After each of these procedures, the results were added to the dataset, which already included the results of the individual CNN experiments used in the preceding procedure. On the dataset, we first ran individual CNN + SVM experiments, and then separate CNN + SVM + LSTM experiments. Following extensive testing, this article concludes which CNNs are the best for certain tasks. Results from many different CNNs were compared using the mixed and Hyper Kvasir datasets, and the best model for detecting colorectal cancer was selected. Experiments using ResNet50, DenseNet-201, and VGG16, three of the most widely used CNN models, are compared to find the best-recognized CNN. In this case, we also think of evaluating the integrated CNN RDV-22 inside the ColoRectalCADx system, which is designed to detect cancer quickly and accurately.

Stage1: Individual CNN experimentation

Table 5: Comparison of accuracy of individual CNN

Dataset	**Individual CNNs**	**Accuracy in %**
Mixed Dataset	**ResNet-50V2v2**	88
	DenseNet-201	88
	VGG-16	83
Dataset	**Integrated CNNs**	**Accuracy in %**
Hyper Kvasir	**ResNet-50V2v2**	*94*
	DenseNet-201	*95*
	VGG-16	*92*

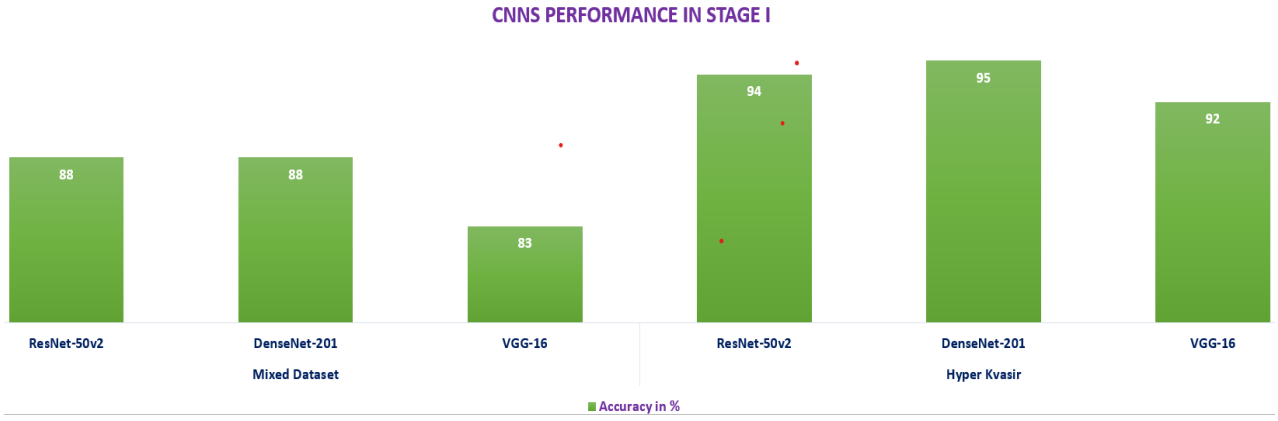


Figure 7: Comparison individual CNN for Mixed and Hyper Kvasir datasets

The mixed dataset using CNN ResNet-50V2V2 and DenseNet-201 produced the highest accuracy (88%) of any CNN studied, whereas VGG-16 showed the lowest drive accuracy (83%). All the research results and comparisons showed that DenseNET-201, using its own CNN, had the highest accuracy (95% on the Hyper Kvasir dataset), whereas VGG-16 had the lowest drive accuracy (92%). Table 5 and Figure 7 show the in-depth comparison findings for the CNNs.

Stage2: Experimentation: Individual CNN+SVM experimentation

Table 6: Comparison of accuracy of individual CNN+SVM

Dataset	CNNs	Accuracy in %(Training)	Accuracy in %(Testing)	SVM Accuracy in %(Training)	SVM Accuracy in %(Testing)	AUC in %
Mixed Dataset	ResNet-50V2v2	92.15	88.44	84.5	84.62	98.74
	DenseNet-201	96.74	88.21	88.33	87.29	98.33
	VGG-16	89.35	82.94	84.16	81.86	98.47
Dataset	**CNNs**	**Accuracy in %(Training)**	**Accuracy in %(Testing)**	**SVM Accuracy in %(Training)**	**SVM Accuracy in %(Testing)**	**AUC in %**
Hyper Kvasir	ResNet-50V2v2	90.07	94.25	84.05	89.8	99.6
	DenseNet-201	96.38	95.23	88.21	91.48	99.7
	VGG-16	90.92	91.63	85.54	88.47	99.5

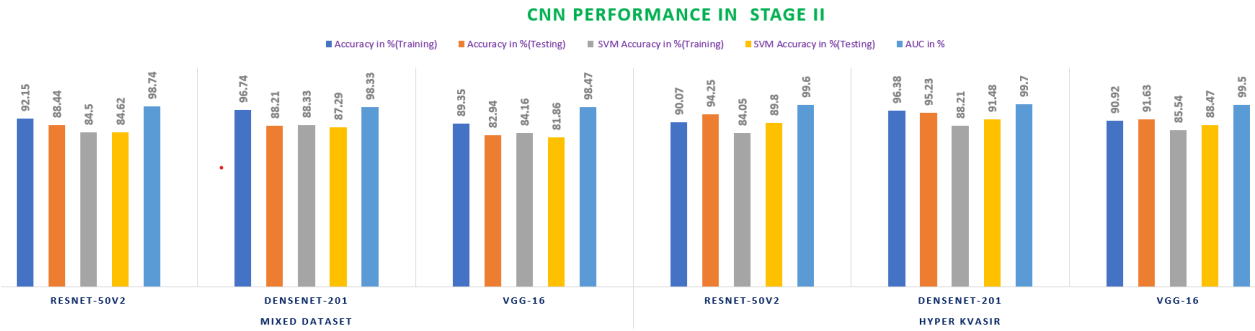


Figure 8: comparison individual CNN for mixed and Hyper Kvasir datasets

The mixed dataset with individual CNN+SVM (DenseNET-201) achieved the highest training accuracy (96.74%), testing accuracy (88.21%), SVM training accuracy (88.33%), and SVM testing accuracy (87.29%) when compared to all other CNNs. In contrast, VGG-16 showed the lowest training accuracy (89.35%), the highest testing accuracy (82.94%), the lowest SVM training accuracy (84.16%), and the highest overall accuracy. VGG-16 is ranked well among the finest CNNs, but it underperforms when compared to ResNet-50V2V2 and DenseNet-201. The Hyper Kvasir dataset with individual CNN+SVM (DenseNet-201) achieved the best training accuracy of 96.74%, the best testing accuracy of 88.21%, the best SVM training accuracy of 88.33%, and the best SVM testing accuracy of 87.29% with respect to any CNN, while VGG-16 displayed the lowest training accuracy of 89.35%, the lowest testing accuracy of 82.94%, and the lowest SVM training accuracy of 84.16 VGG-16 is ranked well among the finest CNNs; however, it underperforms when compared to ResNet-50V2V2 and DenseNet-201.

To measure how well a dataset performs overall, statisticians calculate its area under the curve (AUC), which takes into account both the true-positive (TP) and false-positive (FP) rates for each dataset type. According to these specifics, a performance of less than 50% during dataset creation is considered subpar. DenseNet-201's AUC accuracy was maximised on the suggested mixed dataset, reaching 98.33% for each individual CNN. The DenseNet-201 CNN has the maximum accuracy (99.7%) in the projected Hyper Kvasir dataset. Table 6 and Figure 8 depicts the comparison of the proposed model with the existing literature.

Stage 3: Experimentation: Individual CNN+SVM+LSTM Experimentation

Table 7: Comparison of accuracy of individual CNN+SVM+LSTM

		LSTM			
Dataset	**Integrated CNNs**	**Accuracy in %(Training)**	**Accuracy in %(Testing)**	**SVM Accuracy in %(Training)**	**SVM Accuracy in %(Testing)**
Mixed Dataset	ResNet-50V2v2	96.38	94.51	90.62	90.53
	DenseNet-201	98.91	96.13	95.41	94.86
	VGG-16	93.26	87.08	89.79	90.1
Dataset	**CNNs**	**Accuracy in %(Training)**	**Accuracy in %(Testing)**	**SVM Accuracy in %(Training)**	**SVM Accuracy in %(Testing)**

Hyper Kvasir	ResNet-50V2v2	94.79	97.89	89.93	93.21
	DenseNet-201	98.92	98.91	93.62	95.87
	VGG-16	94.68	95.06	90.95	93.52

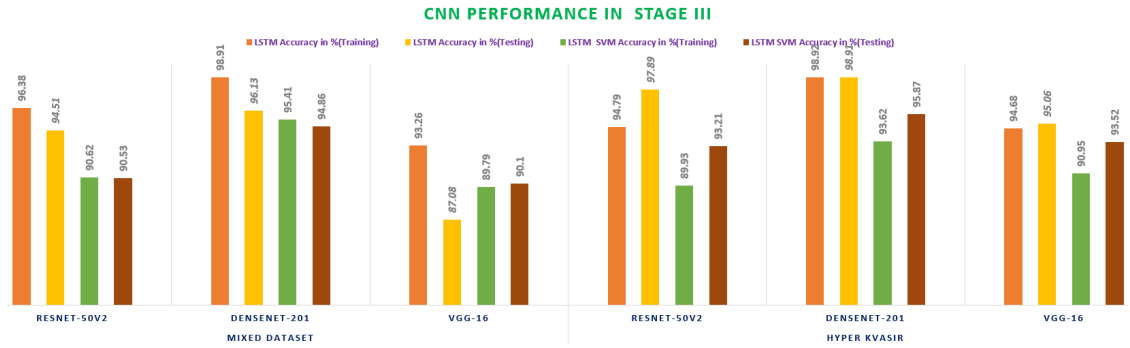


Figure 9: Comparison individual CNN+SVM+LSTM for mixed and Hyper Kvasir datasets

When compared to other CNNs, the DenseNET-201 mixed dataset with individual CNN+SVM+LSTM had the highest training accuracy (98.91%), testing accuracy (96.13%), SVM training accuracy (95.41%), and SVM testing accuracy (94.86%). When compared to other CNNs, VGG-16 had the lowest accuracy rates in training (93.26 %), testing (87.08 percent), SVM training (89.79 percent), and SVM testing (90.1 %). VGG-16 is ranked well among the finest CNNs; however, it underperforms when compared to ResNet-50V2V2 and DenseNet-201. The Hyper Kvasir dataset with individual CNN+SVM+LSTM DenseNet-201 achieved the best training accuracy of 98.92%, the best testing accuracy of 98.91%, the best SVM training accuracy of 93.62%, and the best SVM testing accuracy of 95.87% with respect to any CNN, while VGG-16 displayed the lowest training accuracy of 94.68% the lowest testing accuracy of 95.06%. The lowest SVM training accuracy of 90.95% VGG-16 is ranked well among the finest CNNs; however, it underperforms when compared to ResNet-50V2V2 and DenseNet-201. Table 7 and Figure 9 shows the comparison chart of the proposed model with the existing model.

The integrated CNN +LSTM+SVM

Stages 1 ,2 & 3: Integrated CNN+LSTM+SVM experimentation

In the RDV-22 experiment, the first, second, and third stages of this integrated CNN were studied. In the second stage of the process, the results are acquired by merging the integrated CNN with the SVM. In the third and final stage of this process, the LSTM and SVM models are combined with CNN input. This results in the ultimate outcome. Following tables display the findings of a comprehensive research.

Table 8: Comparison of accuracy of integrated CNN+LSTM+SVM of balanced Mixed dataset and balanced Hyper Kvasir

Dataset	Integrate CNN	Loss	Accuracy	Validation loss	Validation accuracy	SVM Accuracy	LSTM Accuracy	AUC
Mixed Dataset	RDV-22	23.21%	94.72%	43.04%	86.62%	83.14%	96.76%	98.45%
Hyper Kvasir	RDV-22	36.92%	88.73%	36.03%	88.59%	84.97%	92.83%	98.82%

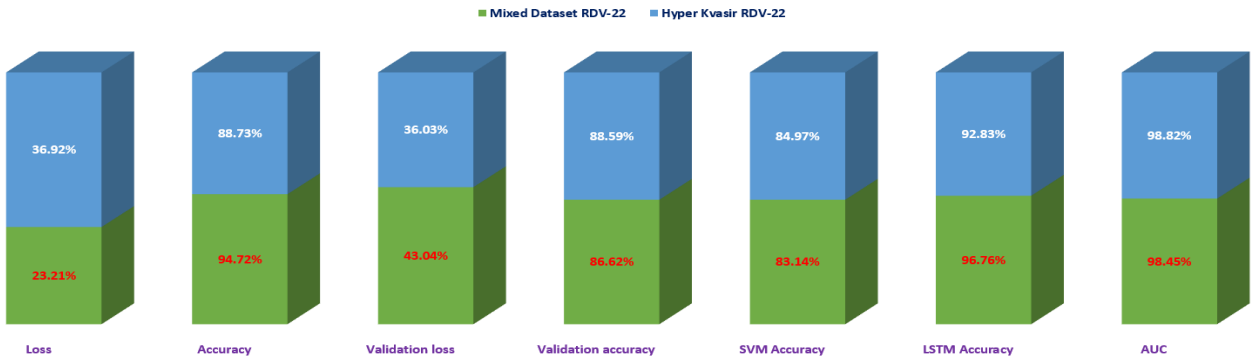


Figure 10: Comparison integrated CNN+LSTM+SVM for mixed and Hyper Kvasir datasets

Table 9: Comparison of the precision, accuracy, F1 score, and recall metrics CNN+LSTM+SVM of balanced Mixed dataset and balanced Hyper Kvasir

Dataset	Individual CNN	Accuracy	F1Score	Recall	Precision	SVM Accuracy	LSTM Accuracy
Mixed Dataset	RDV-22	86.61%	86.52%	86.61%	88.73%	81.86%	90.96%
Hyper Kvasir	RDV-22	88.58%	87.13%	88.58%	90.45%	89.80%	93.68%

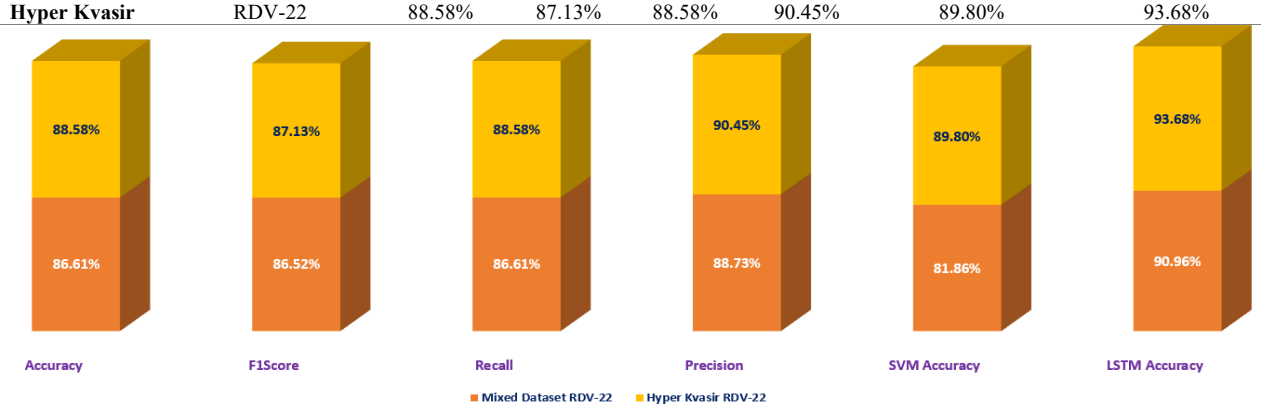


Figure 11: Comparison performance metrics individual CNN+SVM+LSTM for mixed and Hyper Kvasir datasets

The CNN RDV-22 mixed dataset has the best training accuracy (94.72%), the highest testing accuracy (86.61%), the highest SVM training accuracy (83.14%), and the highest SVM testing accuracy (81.86%). When using the combined CNN+SVM+LSTM dataset, RDV-22 achieved the highest training accuracy (96.76%) and testing accuracy (90.96%). The Hyper Kvasir with integrated CNN RDV-22 achieves the maximum accuracy of 88.73 percent in training, 88.58 percent in testing, 84.97% in SVM training, and 89.80% in SVM testing. With a training accuracy of 98.92% and a testing accuracy of 93.68%, respectively, the Hyper Kvasir with individual CNN+SVM+LSTM, RDV-22, topped all other methods. Tables 8 and 9 show all the findings that the integrated CNN RDV-22 has obtained, while figures 10 and 11 show the corresponding visualization.

Performance measures of best CNN models

To understand more about the Hyper Kvasir and Mixed dataset, experiments are undertaken with training and testing. The outcomes of the training and testing detailed in the previous section are provided here and evaluated in depth. After that, each class of the datasets performs the prescribed test, which are known as precision, recall, and F1 score for each of the test support images. For each form of two datasets ColoRectalCADx system undertakes extensive experiment employing a large variety of individual CNNs and integrated CNN. The outcomes of these datasets and CNNs demonstrate that the individual CNN DenseNet-201 and the fusion CNN RDV-22 present the greatest performance compared to the other CNN models. Looking at Table 10 and Figure 112, the various colon component types and their individual performance data can be compared.

Table 10: Performance Metrics of CNNs in ColoRectalCADx in two datasets

Classes		Performance metric for CNN-DenseNet-201 using Hyper Kvasir dataset			
		Precision	**Recall**	**F1 score**	**support**
0	**bbps-0-1**	0.98	0.98	0.98	194
1	**bbps-2-3**	0.99	0.99	0.99	345

2	**cecum**	0.99	0.99	0.99	303
3	**dyed-lifted-polyps**	0.83	0.86	0.85	301
4	**dyed-resection-margins**	0.86	0.83	0.85	297
5	**polyps**	0.97	0.97	0.97	309
6	**pylorus**	0.97	0.97	0.97	300
7	**retroflex-stomach**	0.99	1	0.99	230
8	**z-line**	0.99	0.99	0.99	280

Classes		**Performance metric for CNN-DenseNet-201 using Mixed dataset**			
		Precision	**Recall**	**F1 score**	**support**
0	**bbps-0-1**	0.98	0.98	0.98	198
1	**bbps-2-3**	0.99	0.99	0.99	345
2	**cecum**	0.85	0.99	0.92	603
3	**dyed-lifted-polyps**	0.59	0.92	0.72	601
4	**dyed-resection-margins**	0.86	0.48	0.62	597
5	**Non-Polyps**	0.98	1	0.99	257
6	**polyps**	0.99	0.77	0.87	868
7	**pylorus**	0.97	1	0.99	600
8	**retroflex-stomach**	0.99	0.99	0.99	230
9	**z-line**	0.99	0.99	0.99	580

Classes		**Performance metric for Integrated CNN-RDV-22 using Hyper Kvasir dataset**			
		Precision	**Recall**	**F1 score**	**support**
0	**bbps-0-1**	0.99	0.97	0.98	194
1	**bbps-2-3**	0.98	0.99	0.99	345
2	**cecum**	0.95	0.98	0.97	303
3	**dyed-lifted-polyps**	0.81	0.24	0.37	301
4	**dyed-resection-margins**	0.56	0.93	0.69	297
5	**polyps**	0.95	0.98	0.96	309
6	**pylorus**	0.98	0.96	0.97	300
7	**retroflex-stomach**	0.99	0.98	0.99	230
8	**z-line**	0.97	0.99	0.98	280

Classes		**Performance metric for Integrated CNN-RDV-22 using Mixed dataset**			
		Precision	**Recall**	**F1 score**	**support**
0	**bbps-0-1**	0.97	0.96	0.96	198
1	**bbps-2-3**	0.99	0.99	0.99	345
2	**cecum**	0.87	1	0.93	603
3	**dyed-lifted-polyps**	0.56	0.86	0.67	601
4	**dyed-resection-margins**	0.78	0.49	0.6	597
5	**Non-Polyps**	0.95	0.99	0.97	257
6	**polyps**	0.98	0.74	0.85	868
7	**pylorus**	0.96	0.99	0.98	600
8	**retroflex-stomach**	0.98	0.99	0.98	230
9	**z-line**	0.99	0.98	0.99	580

Hyper Kvasir dataset analysis showed that CNN DenseNet-201 had the highest performance in individual classes, with more than 90% precision, recall, and F1 score with testing support images 194,345,303,309,300,230,280, for the good classification of bbps-0-1, bbps-2-3, cecum, polyps, pylorus, retroflex-stomach, and z-line classes. Of the nine classes, seven have a classification accuracy of 90% or higher, while only two have a classification precision of 80% or higher but 90%. For evaluating support images 194,345,303,309,300,230,280, the Integrated CNN RDV-22 performs best in each class on the Hyper Kvasir dataset, with over 90% precision, recall, and F1 score. The bbps-0-1, bbps-2-3, cecum, polyps, pylorus, retroflex-stomach, and z-line classes are all well classified in this. Seven out of the nine classes have classification accuracy over 90%, while the other three fall between 50% and 90%.

CNN DenseNet-201 outperformed all other models on a mixed dataset, achieving over 90% precision, recall, and F1 score on a total of 198,345,257,868,600,230,580 test support images for good classification of bbps-0-1, bbps-2-3, non-Polyps, polyps, pylorus, retroflex-stomach, and z-line classes. Seven of the ten classes have a classification accuracy of 90% or above, while just three achieve an accuracy of over 50% but below 90%.

On a mixed dataset, CNN RDV-22 surpasses all other models, achieving over 90% precision, recall, and F1 score for accurate classification of the bbps-0-1, bbps-2-3, non-Polyps, polyps, pylorus, retroflex-stomach, and z-line classes.

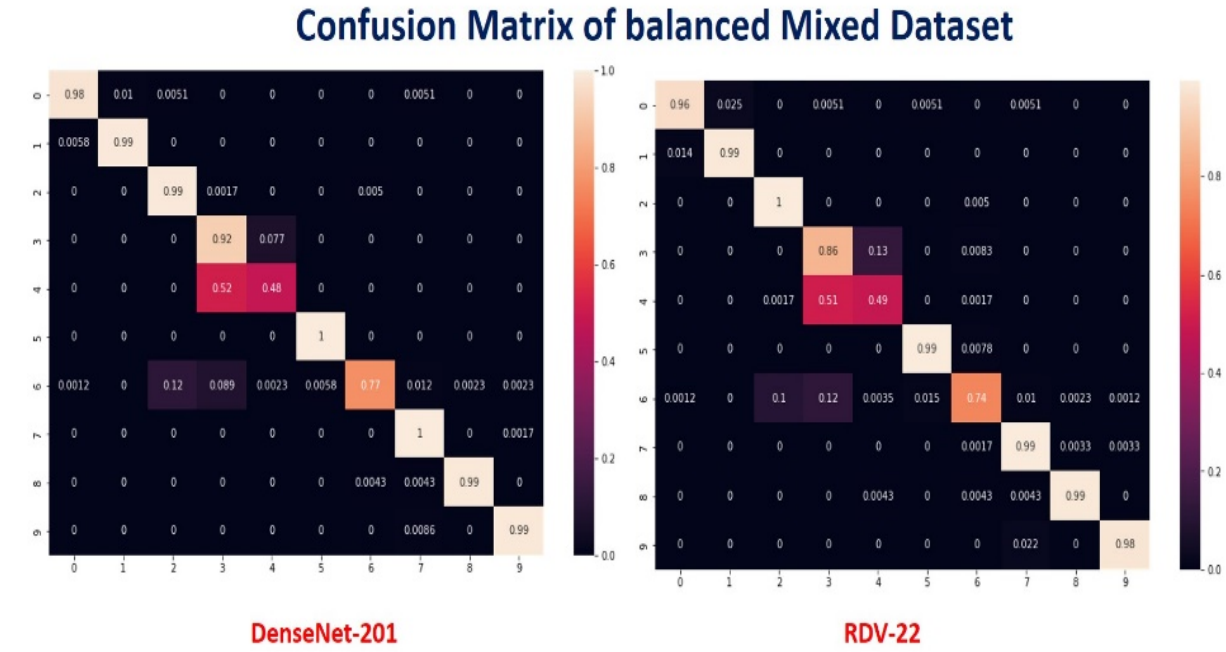
Only two of the 10 classes have classification accuracy between 80% and 90%, whereas seven of the ten have accuracy over 90%. Figure 11 depicts the performance metrics of the ColoRectalCADx system for CNN DenseNet-201 and CNN RDV-22 on Hyper Kvasir and mixed datasets, respectively.

Figure 12: High performed CNN classes: (a) Hyper Kvasir Dataset for DenseNet-201 (b) Mixed Dataset for DenseNet-201 (c) Hyper Kvasir Dataset for RDV-22 (d) Mixed Dataset for RDV-22

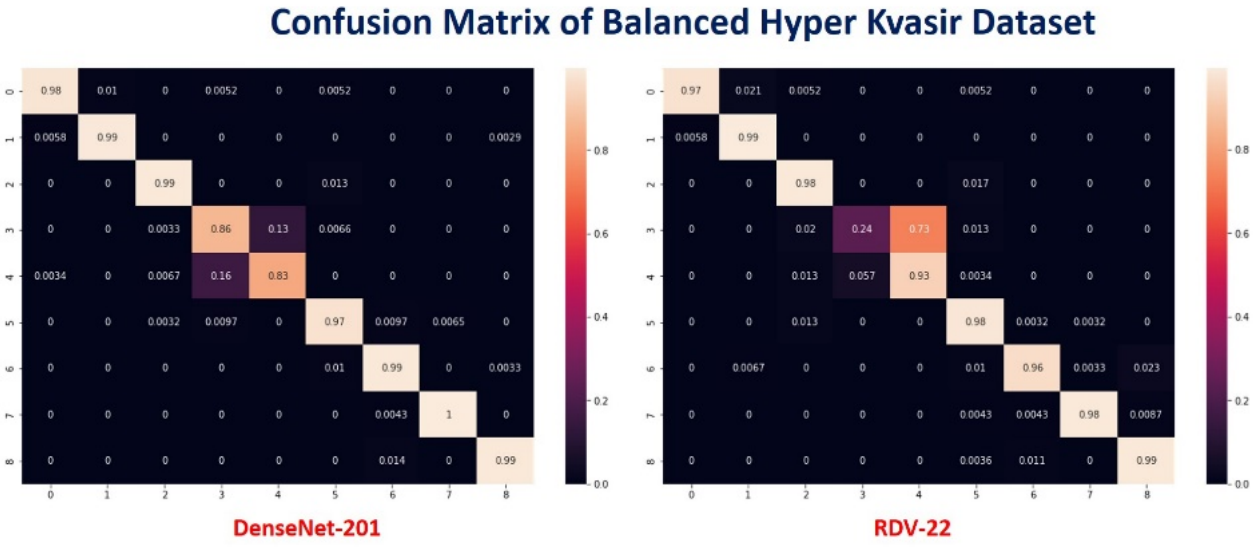
Confusion Matrices and ROC curves

Classification findings from evenly distributed mixed datasets show that individual DenseNet-201 CNNs achieve the best accuracy, while the integrated RDV-22 CNN achieves the highest accuracy. The RDV-22 combined CNN obtains the best accuracy, while the DenseNet-201 individual CNN wins the highest accuracy, based on the classification of the well-balanced Hyper Kvasir datasets. In this scenario, four distinct sorts of results shall be presented: True Positives (TP),

True Negatives (TN), False Positives (FP), and False Negatives (FN). Using the classifications in the dataset, a confusion matrix is constructed. Confusion matrices may be used to evaluate the precision of assessment models. While each row represents a genuine category, each column is an approximation. It evaluates the observed data against the predicted values. In realistic representations, the diagonal should thus contain a disproportionately large number of components. Because of normalising the confusion matrix, the number 1 is now considered the highest diagonal value. Our model indicates that diagonal values for all classes are close to 1. The most accurate confusion matrices are shown in Figure 13.



(a)

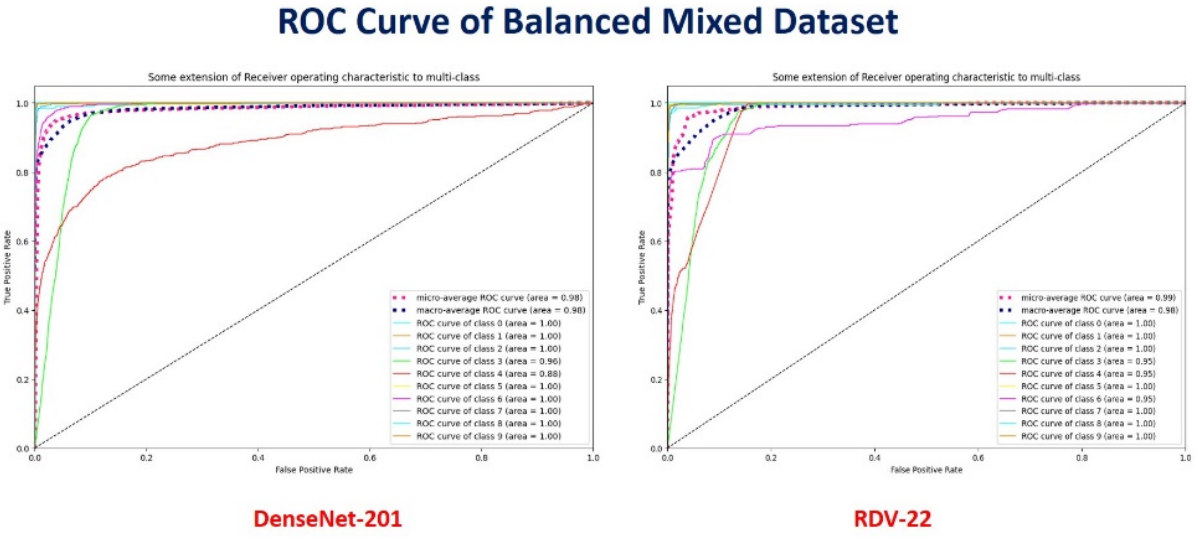


(b)

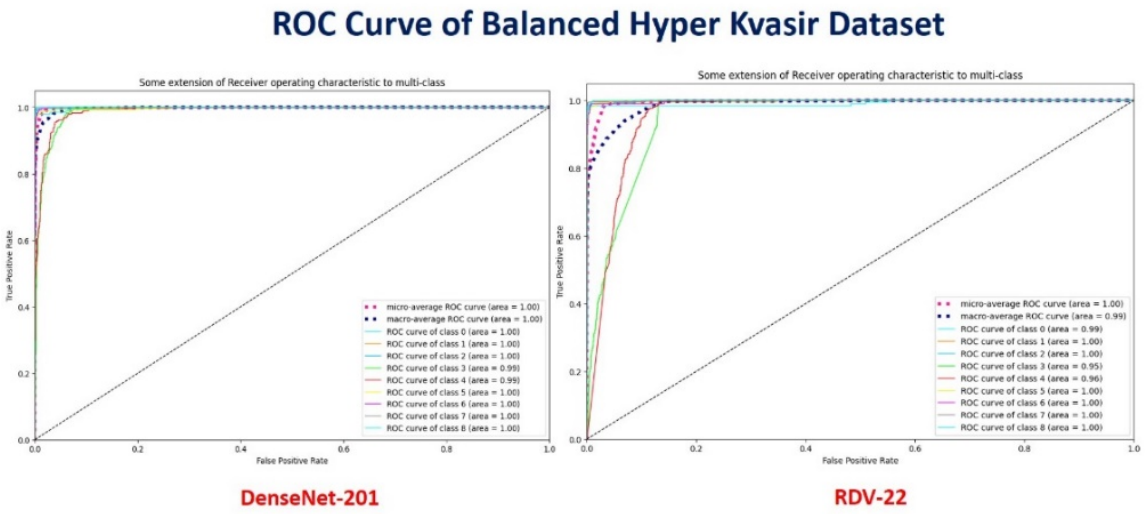
Figure 13: The confusion matrices: (a) Balanced Mixed Datasets (b) Balanced Hyper Kvasir datasets.

Different medical motion colonoscopy data sets are used to evaluate the CNN approaches' recognition capabilities. DenseNet-201 for individual CNNs and RDV-22 for integrated CNNs achieve the highest ratio findings and ROC curves for medical motion image recognition on the balanced mixed dataset. For the balanced hyper-Kvasir dataset, the best results were achieved by DenseNet-201 for individual CNNs and by RDV-22 for integrated CNNs. Figure. 14 shows ROC curves. These recognition rate curves may be produced at varying degrees of system accuracy. The accuracy category is shown using CNN and ROC visualizations, and the top accuracy category is identified and displayed graphically. Rates of TP and

FP are compared graphically. The DenseNet-201 and RDV-22 both achieve between 98% and 100% accuracy on all 10 classes in this evenly balanced Mixed dataset. Across all nine classes in the DenseNet-201 and RDV-22's balanced Hyper Kvasir dataset, both reach 99-100% accuracy.



(a)



(b)

Figure 14: The ROC curves: (a) Balanced Mixed Datasets (b) Balanced Hyper Kvasir datasets.

In its default configuration, the ColoRectalCADx system employs four CNN models: three individual CNN models and one integrated CNN model. To achieve optimum colorectal cancer detection, both the CNN DenseNet-201 and the CNN RDV-22 may undertake a total of four investigations. Key inputs are publicly accessible datasets, adjusted versions of these datasets known as balanced mixed datasets and balanced Hyper Kvasir, and individual datasets that are integrated with these publicly accessible datasets. Optimizing the CNN model using the whole ColoRectalCADx system, and the results of processing various datasets previously worked with, is shown in Table 13.

Table 12: ColoRectalCADx system's top performance across all datasets evaluated.

Dataset	Model Type	Model	Training Accuracy in %	Testing Accuracy in %
CVC Clinic DB	Individual CNN	CNN+DWT+SVM(DenseNet-201)	95.37	99
	Integrated CNN	CNN+DWT+SVM(DaRD-22)	93.86	96
Kvasir2	Individual CNN	CNN+DWT+SVM(DenseNet-201)	80.53	88

	Integrated CNN	CNN+DWT+SVM(DaRD-22)	77.01	82
Hyper Kvasir	Individual CNN	CNN+DWT+SVM(DenseNet-201)	78.17	84
	Integrated CNN	CNN+DWT+SVM(RDEV2-22)	62.38	66
	Individual CNN	CNN+SVM+LSTM(DenseNet-201)	98.92	98.91
Hyper Kvasir Balanced	Integrated CNN	CNN+ SVM+LSTM(RDV-22)	92.83	93.68
Mixed Dataset	Individual CNN	CNN+LSTM(DenseNet-201)	87.1	84.7
	Integrated CNN	CNN+LSTM(ADaDR-22)	84.61	82.17
Mixed Dataset Balanced	Individual CNN	CNN+SVM+LSTM(DenseNet-201)	98.91	96.13
	Integrated CNN	CNN+ SVM+LSTM(RDV-22)	96.76	90.96

ColoRectalCADx supports the import of both independent datasets and mixed datasets, including the combination of CVC Clinic DB, Kvasir2, and Hyper Kvasir. This system further accepts input from distinct datasets. This system provides the best individual CNN performance generated by DenseNet-201 at the current time. During training and testing, its accuracy was 87.1 % and 84.7 %, respectively. The accuracy of the combined training and testing for CNN ADaDR-22 was 84.61 %, while the accuracy of the exams itself was 82.17%. The last phase of the system is the integration of CNN and LSTM, which produces an accurate output. Before commencing tests with mixed-dataset inputs, the system could work with individual dataset inputs and producing results for each dataset independently. Before commencing experiments using mixed-dataset inputs, this was completed. Using the CVC Clinic DB dataset for each CNN, DenseNet-201 had the highest accuracy rates throughout both training and testing. Training accuracy was 95.37% while testing accuracy was 99%. Additionally, the system is compatible with the integrated CNN DarD-22, which has the greatest training and testing accuracy rates of 93.86 and 96%, respectively. DenseNet-201 was the individual CNN with the best accuracy in both training and testing, with a training accuracy of 80.53% and a test accuracy of 88.88% when using the Kvasir dataset. In addition, the system is compatible with the integrated CNN DarD-22, resulting in the greatest training (77.01%) and testing (82%) accuracy rates, respectively. DenseNet-201 attained maximum training and test accuracies of 78.17% and 84%, respectively, using the Hyper Kvasir dataset for the individual CNN. In addition, the system is compatible with the integrated CNN DarD-22, which claims the greatest training accuracy rates of 62.38 and 66%, respectively. For each of the three data sets, the system's findings were combined with those from CNN, DWT, and SVM. Figure 15 displays all information on the combined datasets and the separate datasets, as well as the best results obtained.

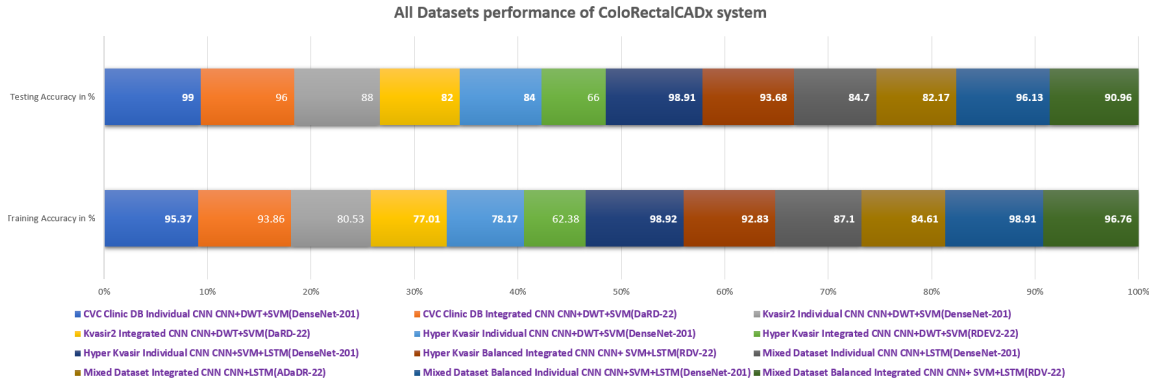


Figure 15: All datasets best performances of ColoRectalCADx

Stage 4 Experimentation: Recognition of Colorectal polyps with K-means clustering.

Using the K-means clustering method, the data points are separated into a given number of groups based on their similarities. This method may be used to identify groupings of pixels within a medical image that have similar features, such as intensity or texture. Then, leveraging these clusters, regions of interest, such as polyps or tumours, might well be pinpointed. The ColoRectalCADx system was created to aid in the detection of cancerous polyps and the diagnosis of colorectal cancer. Identifying polyps that contain cancerous cells is the primary focus of this study. The input for this specific use of K-means clustering is the image itself. In contrast, in this scenario, Python is used to implement the k-means algorithm in a Google Co-Lab environment once the images have been downloaded from the Google Cloud. Due to the co-labs to the Google Cloud, the images from the colonoscopy dataset may be promptly applied to the K-means input.

With a total of 14,287 images, cluster analysis was performed using the mixed datasets of CVC Clinic DB, Kvasir2, and Hyper Kvasir. Different K values are selected and shown as image clusters so that polyps may be easily identified. This allows for the isolation and targeted treatment of linked polyps [53]. These images, which go along with the data sets and display the classification of the data, are accurate depictions of the real analysis. Table 13 displays the k-

means clustering of the datasets according to the image classes, and Figure 16 provides a visual representation of the datasets.

Table 13: Number of clusters for each class in different datasets

Mixed Dataset	
Class	**K Clusters**
bbps-0-1	**9**
bbps-2-3	**15**
cecum	**9**
dyed-lifted-polyp	**7**
dyed-resection-margins	**9**
Non-Polyps	**5**
polyps	**9**
pylorus	**7**
retroflex-stomach	**11**
z-line	**7**

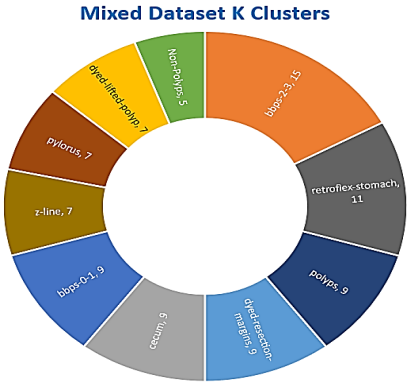
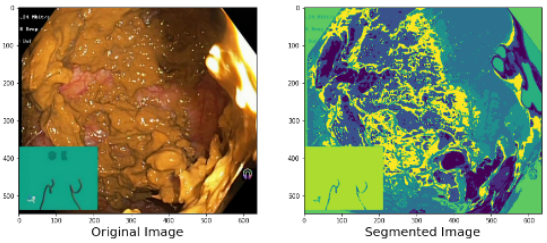


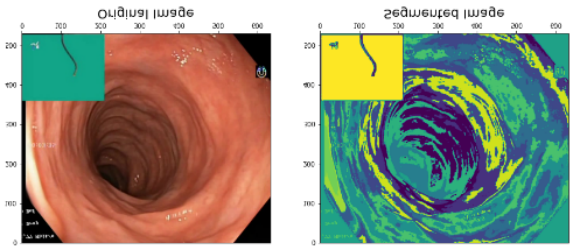
Figure 16: K -Clusters representation of each class in datasets

Mixed dataset K-means segmented clusters.

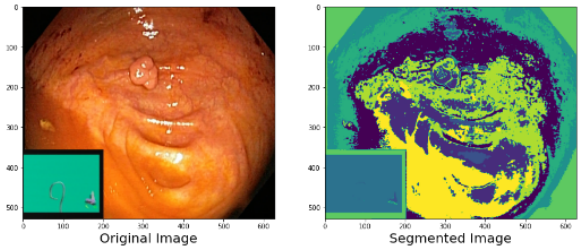
There are ten equally balanced types of medical images in the mixed dataset. The K-means technique was used to create clusters of class images for segmentation and representation. The subsequent subsection provides a thorough analysis and description of the images belonging to each class [54]. Clustered examples of K-means image segmentation for the mixed dataset are shown in Figure 16.



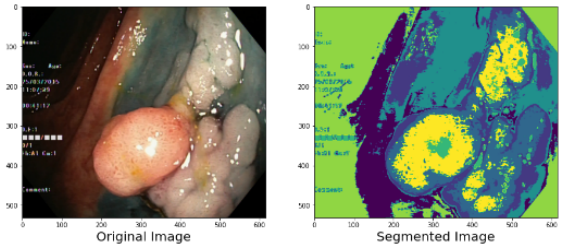
(a)



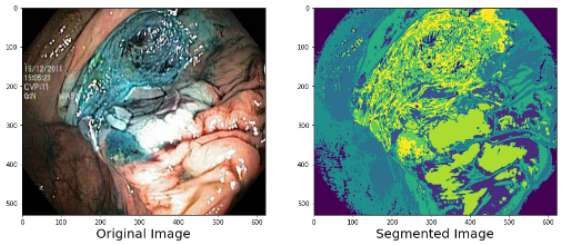
(b)


Original Image
Segmented Image

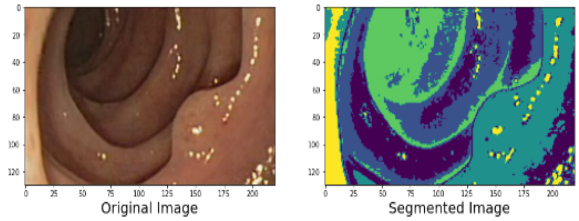
(c)


Original Image
Segmented Image

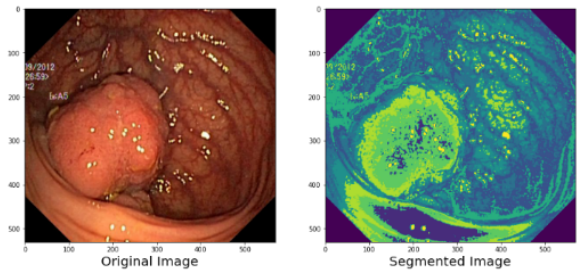
(d)


Original Image
Segmented Image

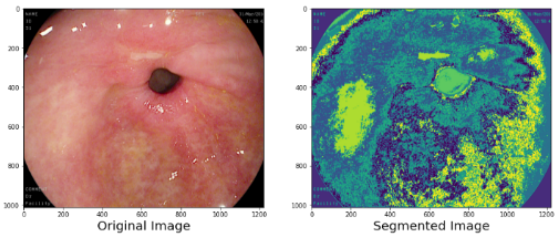
(e)


Original Image
Segmented Image

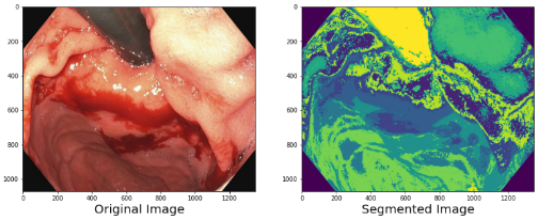
(f)


Original Image
Segmented Image

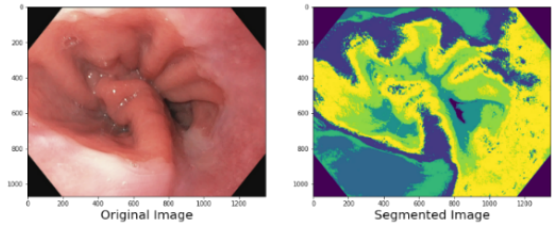
(g)



(h)



(i)



(j)

Figure 17: Balanced Mixed datasets K- values (a) bbps-0-1, K=9 (b) bbps-2-3, K=15 (c) Cecum, K=9 (d) dyed-lifted-polyp, K=7 (e) dyed-resection-margins, K=9 (f) Non-Polyps, K=7 (g) Polyps, K=9 (h)(Pylorus, K=7 (i) retroflex-stomach K=11 (j) z-line, K=7

Figure 17(a) is a nine-valued k composition of the colon component class bbps-0-1. Nine image clusters were created using K-means segmentation. These nine distinct clusters represent possible sites for malignant tumours; the segmented area shown in yellow is most at risk, as evidenced by the presence of polyps that raise concerns about cancer. The colon component class bbps-2-3 is depicted in Figure 17(b). Using K-means clustering with a k value of 15, the image is divided into 15 clusters. The Dark Blue cluster is the most cancer-suspicious of the 15 segmented cluster regions in the displayed image, but it may or may not be malignant. Figure 17(c) represents the colon component class Cecum; in this image, K-means clustering was used using a k=9 value, which resulted in the image being segmented into 9 area clusters. The image that was acquired reveals that out of 9 clusters, the light green cluster demonstrates which cluster is most likely to be impacted. This is simply a suspected carcinogenic zone, but it is possible that this area does not have cancer. Figure 17(d) shows a K-means clustering of the image of the colon component class dyed-lifted-polyp using a K=7 value, yielding seven different clusters. The image depicts seven distinct clusters, with the yellow cluster being the most suspect as a possible site of cancer polyps. The risk of developing colon cancer is greatest in this area.

Using K-means with K = 9, the image in Figure 17(e) is segmented into 9 cluster areas representing the colon component class dyed-resection-margins. The image has been partitioned into nine clusters, the two most affected being the lime green and yellow areas. These two sections indicate high-risk zones for colon cancer because they are likely to contain malignant polyps. Figure 17(f) shows the results of using K-means to segment the image into classes, with K = 7 being used to segment the non-Polyps class. Seven clusters are produced, one of which is dark blue and contains an abnormal polyp that should not be malignant. Figure 17(g) depicts a colon component class containing polyps; using a K-means segmentation algorithm with k=9, the image can divide this class into nine distinct regions of segments. From nine cluster segments, the most probable and hazardous cancer adenoma is identified in the green colour cluster. Figure 17(h) consists of the colon component class termed pylorus; the image is divided into seven distinct cluster segments using the K-means clustering algorithm with K = 7. The cluster with light green and yellow colours is the most potentially malignant location out of the seven segmented colour regions, but it may or may not be diseased.

A representation of the retroflex-stomach colon component class may be seen in Figure 17(i). By using K-means with a K value of 11, the image is divided up into 11 separate clusters. Since the cancerous polyp, also referred to as the malignant polyp, is found inside the dark green clusters, these clusters are the most susceptible to damage out of the eleven clusters.

The image is first segmented before being divided into seven regions for the z-line class illustrated in Figure 17(j). This is accomplished by using the K-means clustering technique with a K value of 7. Although the light green cluster has a higher incidence of cancer than the other six clusters, each of the seven clusters has its own unique significance. The K means clusters and the values are shown in Figure 17. Images of malignant polyps are shown in a balanced mixed dataset as the study of these tumours progresses. K-means aided in identifying the position of the polyps in relation to the colon area, enabling a more precise diagnosis and identification of the cancer. The classes represented in the classified datasets are dependent on the various portions of the colon, and the region of the colon that is afflicted by cancer exactly indicates the location of the malignant tumour. Using an image processing approach based on machine learning, the optimal number of malignant polyps is identified, and colorectal cancer is diagnosed with exceptional intelligence.

K-means clustering is utilised to extract features from colonoscopy images, which are then fed into a machine learning system such as Support Vector Machine (SVM). The resulting SVM model is trained on the labelled images before being evaluated on an independent unlabelled image dataset. When comparing its predictions on the real labels of the 5019 testing images to the training set of 9,268 images, the linear support vector machine (SVM) employed here achieved an 80% rate of accuracy. In Table 13 and Figure 18, we can see the results of a collection of mixed-class performance measures.

Table 13: Performance metrics of K-means clustering with SVM.

Mixed dataset	Precision	Recall	F1 Score	Support
CVC Clinic DB	0.60	0.61	0.60	516
Kvasir2	0.78	0.78	0.78	1800
Hyper Kvasir	0.80	1.00	0.89	2703

CVC Clinic DB, Kvasir2, and Hyper Kvasir images have been incorporated in the mixed dataset, which can be fed to the SVM based on the K-means clustering features of the mixed dataset. The accuracy of the CVC clinic's DB image classification reports is 60%, the recall rate is 61%, and the F1 score is 60% with support of 516 testing images. The kvasir2 image classification reports yielded the following results: 78% precision, 78% recall, and 78% F1 score with support of 1800 images. The accuracy of the Hyper Kvasir image classification reports is 80%, the recall rate is 100%, and the F1 score is 89% with support of 2703 testing images. A total of 5019 test images are supported. The following shown in Figure 17 are graphs depicting the performance metrics of the ColoRectalCADx system for K-means clustering images of mixed datasets are fed to the SVM and obtained the results.

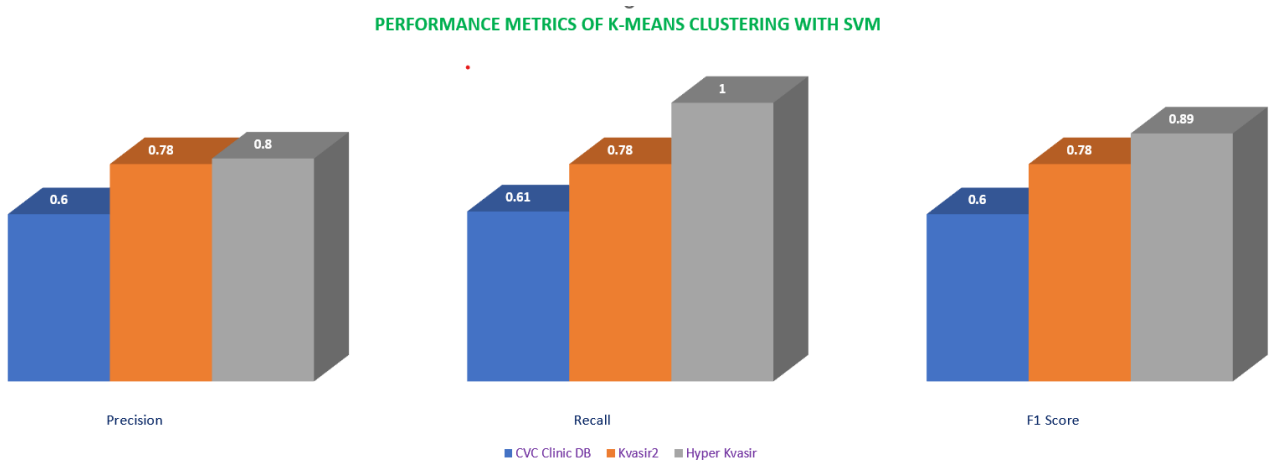


Figure 18: Performance metric comparison between K-means clustering with SVM classification.

V. Discussion

The results of the comprehensive ColoRectalCADx system were analysed and contrasted with those of the three-step GastroCADx procedure slated for 2021. The findings of the analysis done on all three data sets are shown in Table 14
In the GastroCADx research, the system was evaluated against four separate CNN models, while in the ColoRectalCADx study, it was evaluated against seven distinct CNN models. GastroCADx found DenseNet-201 to be the best model, whereas ColoRectalCADx also found DenseNet-201 to be the best model. In contrast, some systems are better described using approaches that compare well with one another while using differing technical models. There was no uniformity in how they tackled the issue.

Table 13: Comparative analysis of prior state-of-the-art procedures from 2021 to 2022

Dataset used in experiment	Author	Method	Model used	Accuracy in %
CVC Clinic DB	Liew et al. (2021) [55]	**Ensemble classifier**	ResNet50+Adaboost, AlexNet, GoogLeNet, and VGG-19	***97.91***

CVC Clinic DB, Aichi Medical Dataset	Pallabi Sharma et al. (2022) [56]	**Ensemble classifier**	ResNet101, GoogleNet and XceptionNet	***98.3***
CVC Clinic DB, ETIS-Larib	Nisha J.S et al. (2022) [57]	**DP-CNN**	Dual Path CNN	***99.6***
CVC Clinic DB, WCE dataset	Maryem Souaidi et al. (2022) [22]	**MP-FSSD**	VGG16 with feature Fusion Module	***91.56***
Kvasir2	Omneya Attallah et al. (2021) [58]	**GastroCADx**	AlexNet, DarkNet19, ResNet50 and DenseNet-201, DWT and DCT functions, SVM	***97.3***
Kvasir2,	Pallabi Sharma et al. (2022) [56]	**Ensemble classifier**	ResNet101, GoogleNet and XceptionNet	***97***
Hyper Kvasir	Omneya Attallah et al. (2021) [58]	**GastroCADx**	AlexNet, DarkNet19, ResNet50 and DenseNet-201, DWT and DCT functions, SVM	***99.7***
Hyper Kvasir Balanced	Proposed model	**ColoRectalCADx**	ResNet-50V2, DenseNet-201, VGG16, LSTM and SVM	***98.91***
Mixed Dataset Balanced	Proposed model	**ColoRectalCADx**	ResNet-50V2, DensNet-201, VGG16, LSTM and SVM	***96.13***

In 2021 and 2022, researchers began comparing and contrasting ColoRectalCADx with several system models, including ensemble classifiers, DP-CNN, and MP-FSSD. When examining all three datasets, the previous version of the ColoRectalCADx system obtained an accuracy of 84%, 88%, and 100%, respectively. The accuracy of a dataset including several entries is 84.7%. DenseNet-201 had the greatest performance in detecting colorectal cancer, with a sensitivity of 98.91% and a maximum sensitivity of 96.13%, respectively, for the balanced datasets of mixed and Hyper Kvasir. While experimenting with the ColoRectalCADx system, it is necessary to consider limitations. Complex models such as ResNet-50V2V2, DenseNet-201, VGG-16, LSTM, and SVM may overfit training data, resulting in high accuracy on the training set but poor generalisation to new data. Cross-validation and regularisation are indispensable model training strategies for mitigating this risk. The findings may not be applicable to other situations. ResNet-50V2V2, DenseNet-201, VGG-16, LSTM, and SVM have demanding computational requirements, necessitating a substantial amount of memory and processing capacity. This could limit their applicability in environments with limited resources or on devices with limited capabilities.

Conclusion and Future Work

On a worldwide scale, colorectal cancer is the third most common and deadly form of the disease. To identify colorectal cancer, colonoscopy screening is crucial. Proposed study involves publicly accessible colonoscopy datasets including the CVC Clinic DB, Kvasir2, and Hyper Kvasir. Each of the three datasets has 2.8.23 classes, with varying numbers of clinical motion colonoscopy images in each. Originally, there were 24 classes and 19621 images in this new dataset, but after modifications have been made to include more evenly distributed samples, just 10 classes and 14,287 images existed. The original Hyper Kvasir dataset included 23 classes, however it was updated to accommodate the more evenly distributed classes. Too far, there have been 9 classes updated, and a total of 8529 images have been included in the dataset. These two standardised datasets are being investigated to the accuracy in cancer prediction The employed CNNs were broken down into its component parts, with ResNet50V2, DenseNet-201, and VGG16 making up the whole. Additionally, these models were integrated into a single RDV-22 model. All four CNNs are put through its paces in order to find the best effective and practical colorectal cancer detection model. The proposed system achieved the highest level of accuracy with CNN DesnseNet-201 in Hyper Kvasir (98.92% accuracy during training, 98.91% accuracy during testing, 93.62% accuracy during SVM training, and 95.87% accuracy during SVM testing). The mixed dataset provided CNN DesntNet-201 with the highest degree of success in terms of accuracy (98.91% training, 96.13% testing, 95.41% SVM training, and 94.86% SVM testing). The CNN+SVM+LSTM configuration was found to be the most effective after three system iterations, with positive results on both datasets. With the goal of identifying cancerous polyps in imaging studies. After convolutional neural networks (CNNs), support vector machines (SVMs), and long short-term memory (LSTMs), K-means clustering visualisation is used to display the desired category from the input balanced mixed dataset and achieved an accuracy of 80% using SVM. The original frame polyps suspected of being cancerous were isolated using the segmentation-based clustering method of K-means clustering. Consistent experiments are conducted on multiple collections of colonoscopy motion videos and other types of images for the future aspect of colorectal cancer recognition. Using deep learning models trained on large datasets of colonoscopy images to enhance the images can improve accuracy and efficiency. This technique has the potential to detect lesions in colonoscopy imaging, as well as MRI and CT scans.

Acknowledgements

The authors of this article are grateful to Chennai's SRM Institute of Science and Technology. Our work has been delegated to artificial intelligence, machine learning, and deep learning laboratories with advanced GPU systems. This initiative receives no funding from outside sources.

Declarations

Conflicts of Interest

There are no conflicts of interest between authors.

Funding

There is no external funding for this project.

Data Availability

Collecting data from publicly accessible colonoscopy datasets, CVC clinic DB, Kvasir 2, and Hyper Kvasir for the project. Even though this work with internal human organs is available to the public via Internet sources, it has not been deemed ethical by official authorities.

Data are publicly available from the following websites:

1. CVC clinic DB Dataset was obtained from
https://www.kaggle.com/datasets/balraj98/cvcclinicdb.
2. Kvasir2 Dataset was obtained from
https://datasets.simula.no/kvasir/.
3. Hyper Kvasir Dataset was obtained from
https://datasets.simula.no/hyper-kvasir/.

In this research paper, the combining of three datasets is presented as a new dataset known as a mixed dataset.

References

[1] Neil Murphy, Victor Moreno, David J. Hughes, Ludmila Vodicka, Pavel Vodicka, Elom K. Aglago, Marc J. Gunter, Mazda Jenab, "Lifestyle and dietary environmental factors in colorectal cancer susceptibility," *Molecular Aspects of Medicine,* vol. 69, pp. 2-9.

[2] "Americal cancer Society," ACS. [Online]. [Accessed 15 November 2022].

[3] "Colorectal Cancer Facts & Figures 2020-2022," American Cancer Society, 2022.

[4] "World Health Orgnization(WHO," 2018. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer. [Accessed 1 December 2022].

[5] "American Cancer Society," 2021. [Online]. Available: https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/survival-rates.html. [Accessed 2 December 2022].

[6] "Cancer Facts & Figures 2021," American Cancer Society, 2021.

[7] " Colorectal Cancer: Risk Factors and Prevention," ASCO, May 2022. [Online]. Available: https://www.cancer.net/cancer-types/colorectal-cancer/risk-factors-and-prevention. [Accessed 20 November 2022].

[8] Baopu Li, Max Q.-H. Meng, "Tumor Recognition in Wireless Capsule Endoscopy Images Using Textural Features and SVM-Based Feature Selection," *IEEE Transactions on Information Technology in Biomedicine ,* vol. 16, no. 3, pp. 323 - 329, 24 January 2012.

[9] Pittman, Erika Hissong & Meredith E., "Colorectal carcinoma screening: Established methods and emerging technology," *Critical Reviews in Clinical Laboratory Sciences,,* pp. 22-36, January 2020.

[10] Iyad A Issa,alak Noureddine, "Colorectal cancer screening: An updated review of the available options," *world journal of Gastroenterology,* p. 5086–5096., 28 July 2017 .

[11] M. Bernard Levin, M. M. Durado Brooks, P. Robert A. Smith and M. Amy Stone, "Emerging Technologies in Screening for Colorectal Cancer: CT Colonography, Immunochemical Fecal Occult Blood Tests, and Stool Screening Using Molecular Markers,,," *Emerging Technologies in screening for colorectal cancer,* pp. 53:44-55, 2003.

[12] Bilal Taha, Naoufel Werghi, Jorge Dias,, "Automatic polyp detection in endoscopy videos: A survey," in *Proceedings of the lASTED International Conference Biomedical Engineering (BioMed 2017)*, Innsbruck, Aus t r i a, February 20 - 21, 201 7 published:06 April 2017.

[13] Hanna Borgli # 1 2, Vajira Thambawita, Pia H Smedsrud , Steven Hicks,S, Jha D, Eskeland SL, Randel KR, Pogorelov K, Lux M, Nguyen DTD, Johansen D, Griwodz C, Stensland HK, Garcia-Ceja E, Schmidt PT, Hammer HL, Riegler MA, Halvorsen P, de Lange T, "HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy," *Scientific Data,* vol. PMID: 32859981, p. 7(1):283., 28 August 2020.

[14] Hai Tang; Zhihui Hu, "Research on Medical Image Classification Based on Machine Learning," *IEEE Access,* vol. 8, pp. 93145 - 93154, 11 May 2020.

[15] Muthu Subash Kavitha , Prakash Gangadaran , Aurelia Jackson , Balu Alagar Venmathi Maran , Takio Kurita and Byeong-Cheol Ahn, "Deep Neural Network Models for Colon Cancer Screening," *Cancers,* 29 July 2022.

[16] D. Banik, K. Roy, D. Bhattacharjee, M. Nasipuri and O. Krejcar, "Polyp-Net: A Multimodel Fusion Network for Polyp Segmentation," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-12, 10 August 2020.

[17] Kavitha MS, Gangadaran P, Jackson A, Venmathi Maran BA, Kurita T, Ahn BC, "Deep Neural Network Models for Colon Cancer Screening. Cancers (Basel)," *Cancers,* vol. 14, no. 15, p. 3707, 29 July 2022.

[18] Zheng Cao , Xiang Pan, Hongyun Yu, Shiyuan Hua, Da Wang, Danny Z. Chen ,Min Zhou , And Jian Wu5, "A Deep Learning Approach for Detecting Colorectal Cancer via Raman Spectra," *BMEF,* vol. 2022, 02 May 2022.

[19] M. Siva Naga Raju B. Srinivasa Rao, "Colorectal Cancer Disease Classification and Segmentation Using A Novel Deep Learning Approach," *International Journal of Intelligent Engineering and Systems,* vol. 4, 29 April 2022.

[20] Saban ozturik, Umut ozakaya, "Residual LSTM layered CNN for classification of gastrointestinal tract diseases," *Journal of Biomedical Informatics,* vol. 113, January 2021.

[21] Yan Wang , Zixuan Feng, Liping Song,Xiangbin Liu, and Shuai Liu, "Multiclassification of Endoscopic Colonoscopy Images Based on Deep Transfer Learning," *Computational and Mathematical Methods in Medicine,* vol. 2021, 3 July 2021.

[22] Meryem Souaidi ,Mohamed El Ansari, "A New Automated Polyp Detection Network MP-FSSD in WCE and Colonoscopy Images Based Fusion Single Shot Multibox Detector and Transfer Learning," *IEEE Access,* vol. 10, pp. 47124 - 47140, 29 April 2022.

[23] Qingqing Guo; Xianyong Fang; Linbo Wang; Enming Zhang, "Polyp Segmentation of Colonoscopy Images by Exploring the Uncertain Areas," *IEEE Access,* vol. 10, pp. 2169-3536, 17 May 2022.

[24] Fenlon HM, Nunes DP, Schroy PC 3rd, Barish MA, Clarke PD, Ferrucci JT, " comparison of virtual and conventional colonoscopy for the detection of colorectal polyps.," *The New England Journal of Medicine,* vol. 341, no. 20, p. 1496–503, 11 November 1999.

[25] Jason A. Dominitz,, Douglas J. Robertson, "Understanding the Results of a Randomized Trial of Screening Colonoscopy," *The New England Journal of Medicne,* vol. 387, pp. 1609-1611, 27 October 2022.

[26] Fausto Milletar,Johann Frei, Gerome Vivar,Seyed-Ahmad Ahmadi, "Cloud Deployment of High-Resolution Medical Image Analysis With TOMAAT," *IEEE Journal of Biomedical and Health Informatics,* Vols. Volume: 23,, no. Issue: 3, pp. 969 - 977, 05 December 2018.

[27] B.Prashanth, Mruthyunjaya Mendu, Ravikumar Thallapalli, "Cloud based Machine learning with advanced predictive Analytics using Google Colaboratory," *Materials Today: Proceedings,* 3 March 2021.

[28] "https://www.kaggle.com/balraj98/cvcclinicdb," 2015. [Online]. [Accessed 25 May 2021].

[29] "https://datasets.simula.no/kvasir/," 2016. [Online]. [Accessed 3 July 2021].

[30] "https://datasets.simula.no/hyper-kvasir/," 2020. [Online]. [Accessed 3 July 2021].

[31] Oza, P.; Sharma, P.; Patel, S.; Adedoyin, F.; Bruno, "A. Image Augmentation Techniques for Mammogram Analysis.," *Journal of Imaging,* vol. 8, no. 5, 20 May 2022.

[32] Abol Basher, Byeong C. Kim,Kun Ho Lee,Ho Yub Jung, "Automatic Localization and Discrete Volume Measurements of Hippocampi From MRI Data Using a Convolutional Neural Network," *IEEE Access,* vol. Volume: 8, pp. 91725 - 91739, 14 May 2020.

[33] AminKabir Anaraki,MoosaAyati,FoadKazemi, "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *Biocybernetics and Biomedical Engineering,* vol. 39, no. 1, pp. Pages 63-74, January–March 2019.

[34] Haya Alaskar , Abir Hussain , Nourah Al-Aseem , Panos Liatsis , Dhiya Al-Jumeily , "Application of Convolutional Neural Networks for Automated Ulcer Detection in Wireless Capsule Endoscopy Images," *sensors,* p. 19(6):1265, 13 March 2019.

[35] Hemin ali qadir, Johannes solhusvik,Jacob bergsland, Lars aabakken,and Ilangko balasingham, "A Framework With a Fully Convolutional Neural Network for Semi-Automatic Colon Polyp Annotation," *IEEE Access,* vol. Volume: 7, pp. 169537 - 169547, 20 November 2019.

[36] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition," in *Computer Vision and Pattern Recognition (cs.CV),*, 10 Dec 2015.

[37] Gao Huang,Zhuang Liu,Laurens van der Maaten, "Densely Connected Convolutional Networks," *Computer Vision and Pattern Recognition (cs.CV),* 28 January 2018.

[38] Noha Ghatwary, Xujiong Ye,Massoud Zolgharni, "Esophageal Abnormality Detection Using DenseNet Based Faster R-CNN With Gabor Features," *IEEE Access,* vol. 7, pp. 84374 - 84385, 27 June 2019.

[39] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *Computer Vision and Pattern Recognition (cs.CV),arXiv:1409.1556v6*, 10 Apr 2015.

[40] Tao yang, Ning liang, Jing li, Yi yang, Yuedan li, Quan, Huang,, "Intelligent Imaging Technology in Diagnosis of Colorectal Cancer Using Deep Learning," *IEEE Access,* vol. 7, pp. 178839-178847, 06 December 2019.

[41] Qiaoliang Li, Guangyao Yang, Zhewei Chen, Bin Huang, Liangliang Chen, Depeng Xu, Xueying Zhou, Shi Zhong, Huisheng Zhang and Tianfu Wang,, "Colorectal Polyp Segmentation Using A Fully Convolutional Neural Network," in *10th International Congress on Image and signal Processing , Biomedical Engineering and infromatics(CISP-BMWI2017)*, Shanghai, China, 14-16 Oct. 2017 Published: 27 February 2018.

[42] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, Quoc V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *Computer Vision and Pattern Recognition (cs.CV); Machine Learning (cs.LG); Machine Learning (stat.ML),arXiv:1707.07012v4*, 11 Apr 2018.

[43] Taruna Agrawal, Rahul Gupta, Saurabh Sahu, C. Espy-Wilson, "SCL-UMD at the Medico Task-MediaEval 2017: Transfer Learning based Classification of Medical Images," in *Medico challenge,,MediaEval*, Dublin, Ireland, 13th september,2017.

[44] Öztürk, Ş., Özkaya, U, "Gastrointestinal tract classification using improved LSTM based CNN," *Multimedia Tools and Applications,* vol. 79, p. 28825–28840, 06 August 2020.

[45] HaifengWu,QingHuang,DaqingWang,LifuGao, "A CNN-SVM combined model for pattern recognition of knee motion using mechanomyography signals," *Journal of Electromyography and Kinesiology,* vol. 42, pp. 136-142, 26 July 2018.

[46] Wentao Wu, Daning Li, Jiaoyang Du,1 Xiangyu Gao, Wen Gu,Fanfan Zhao, Xiaojie Feng, and Hong Yan, "An Intelligent Diagnosis Method of Brain MRI Tumor Segmentation Using Deep Convolutional Neural Network and SVM Algorithm," *Computational and Mathematical methods in medicine volume 2020,* vol. Volume 2020, 14 July 2020.

[47] MuhammedTalo,Ulas BaranBaloglu,ÖzalYıldırım,URajendra Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," *Cognitive Systems Research,* vol. 54, pp. 176-188, 5 May 2019.

[48] Wener Borges Sampaio 1, Edgar Moraes Diniz, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Marcelo Gattass, "Detection of masses in mammogram images using CNN, geostatistic functions and SVM," *computers in Biology and Medicine,* vol. 21703605, pp. 41(8):653-64, 17 August 2011.

[49] Takumi Okamoto, Tetsushi Koide, Shigeto Yoshida Hiroshi Mieno, Hiroshi Toishi Takayuki, Sugawara Masayuki Tsuji, Masayuki Odagawa Nobuo Tamba, Toru Tamaki, Bisser Raytchev ,Kazufumi Kaneda, Shinji Tanaka, "Implementation of Computer-Aided Diagnosis Sysstem on Custamizable DSP core for Colorectal Endoscopic images with CNN Features and SVM," in *Proceedings TENCON 2018-2018 IEEE Region 10 Conference*, Jeju, Korea, 28-31 Oct. 2018, Published:25 February 2019.

[50] Le Hoang Thai,Tran Son Hai,Nguyen Thanh Thuy, "Image Classification using Support Vector Machine and Artificial Neural Network," *I.J. Information Technology and Computer Science,* vol. 5, pp. 32-38, 2012.

[51] W. Kim, A. Kanezaki and M. Tanaka, "Unsupervised Learning of Image Segmentation Based on Differentiable Feature Clustering," *IEEE Transactions on Image Processing,* vol. 29, pp. 8055-8068, 28 July 2020.

[52] Takayasu Moriya, Holger R. Roth, Shota Nakamura, Hirohisa Oda, Kai Nagara, Masahiro Oda, Kensaku Mori, "Unsupervised pathology image segmentation using representation learning with spherical k-means," *Medical Imaging, Digital Pathology;,* vol. 10581, 2018.

[53] Li Liu, Liang Kuang, Yunfeng Ji, "Multimodal MRI Brain Tumor Image Segmentation Using Sparse Subspace Clustering Algorithm," *Computational and Mathematical Methods in Medicine,* vol. 2020, p. 13, 4 July 2020.

[54] NameirakpamDhanachandra, KhumanthemManglem, Yambem JinaChanu, "Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm," *Procedia Computer Science,* vol. 54, pp. 764-771, 2015.

[55] Liew, W.S., Tang, T.B., Lu, CK, "Computer-Aided Diagnostic Tool for Classification of Colonic Polyp Assessment.," in *International Conference on Artificial Intelligence for Smart Community.*, 2022.

[56] Pallabi Sharma, Bunil Kumar Balabantaray, Kangkana Bora, Saurav Mallik , Kunio Kasugai and Zhongming Zhao, "An Ensemble-Based Deep Convolutional Neural Network for Computer-Aided Polyps Identification From Colonoscopy," *Frontiers in Genetics,* 26 April 2022.

[57] Nisha J.S., Varun P. Gopi , Palanisamy P., "Automated colorectal polyp detection based on image enhancement and dual-path CNN architecture," *Biomedical Signal Processing and Control,* vol. 73, March 2022.

[58] Omneya Attallah , Maha Sharkas, "GASTRO-CADx: a three stages framework for diagnosing gastrointestinal diseases," *PeerJ Computer Science,* vol. PMID: 33817058, p. 7:e423, 10 March 2021.