

Reptile Search Algorithm with Deep Convolutional Neural Network for Cloud Assisted Colorectal Cancer Detection and Classification

Rajeev Kumar Tiwari¹, Dr. S. Murugappan²

^{1, 2}Dept. of Computer and Information Science, Annamalai University, Annamalainagar – 608 002, Tamil Nadu, India

Abstract

Cloud-based automatic colorectal cancer (CC) detection involves the usage of cloud computing technology and system to help in the earlier and accurate diagnosis of CC in medical images and patient information. This cloud-based technology aims to improve the efficiency and reliability of CC screening, monitoring, and diagnoses. Automatic CC detection refers to the use of computer-based technology and systems to aid in the earlier and accurate detection of CC in patient data and medical images. This automated system aims to increase the reliability and efficiency of CC monitoring, screening, and diagnosis. Deep learning (DL) methods, especially convolutional neural networks (CNNs), exhibit promising results in automatic CC diagnosis. They can be trained on wide-ranging datasets of medical images to learn patterns and features related to precancerous and cancerous lesion. This study develops a new Reptile Search Algorithm with Deep Learning for Colorectal Cancer Detection and Classification (RSADL-CCDC) technique. The main aim of the RSADL-CCDC method focuses on the automatic classification and recognition of the CC in the cloud environment. Once the medical images are stored in the cloud server, the detection process is carried out. In the presented RSADL-CCDC approach, the initial stage of preprocessing is performed by bilateral filtering (BF) approach. For feature extraction, the RSADL-CCDC technique applies ShuffleNetv2 model. Besides, the recognition and classification of CC take place using convolutional autoencoder (CAE) model. Finally, the hyperparameter tuning of the CAE technique takes place by utilizing RSA. The experimental validation of the RSADL-CCDC system is performed on benchmark medical database. Extensive results stated the enhanced performance of the RSADL-CCDC technique on CC recognition over other models with respect to various actions.

Keywords: Colorectal cancer; Cloud environment; Computer diagnosis; Medical imaging; Reptile search algorithm; Deep learning.

1. Introduction

Cloud computing has become a game-changing technology in the healthcare field, offering wide array of benefits that have revolutionized the way healthcare organization manage data, deliver services, and collaborate with stakeholders [1]. With respect to healthcare, cloud computing refers to the delivery of computing services, including storage, processing, and data access, over the internet, which allows healthcare providers and organizations to remotely access and leverage these resources. One of the most important benefits of cloud computing in healthcare is enhanced interoperability and data accessibility. Healthcare generates abundance of data, including medical images, electronic health records (EHRs), research data, and patient histories [2]. Storing and managing this data on local servers can be inefficient, costly, and prone to data silos. Cloud-based solutions standardize and centralize data, making it easily accessible to authorized users across various healthcare sectors. This improved interoperability supports continuous data sharing amongst healthcare providers, improving care coordination, enhancing patient outcomes, and ultimately reducing duplication of tests [3]. Furthermore, cloud platforms enable real-time and secure access to patient data from anyplace, enabling healthcare experts to make informed decisions at the point of care. Also, patients can benefit from cloud-based personal health records (PHRs), gaining access to their medical information and taking a more active role in their healthcare management. Fig. 1 shows the architecture of cloud computing in the healthcare sector.

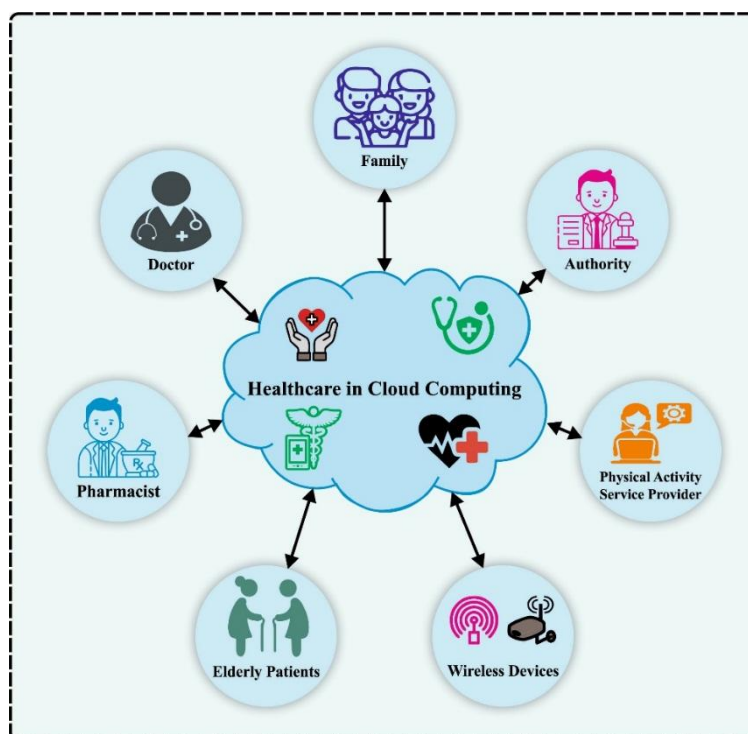


Fig. 1. Cloud Computing in Healthcare

Colorectal cancer (CRC) is a major common cancer that stands at third position globally. Even though, the treatment techniques develop faster, earlier identification performs a vital part in reducing mortalities [4]. In addition to that, it is highly recognized that adenomas involve a 50% harmful alteration capability and almost one-fourth of them could be lost in the conventional colonoscopy. Due to this reason, an effective colonoscopy is very important to discover CRC and its potential signs [4]. The CRC risk factors are family history, sex, age, pre-existing conditions like Lynch syndrome, inflammatory bowel disease, etc. and other unnecessary lifestyle factors such as alcohol, physical inactivity, obesity, a diet high in red meat, low-fibre diet and smoking. Moreover, CRC is said to be a serious health issue because it is symptomless until the later stages when the

cancer is improved. In the initial stage, if CRC is determined as adenomatous polyps, it is a mainly curable disease, which can benefit from curative surgery [5]. Presently, histopathological analysis plays an essential role in evaluating cancer potential of a lesion.

With high resolution of diagnosis, screening and treatment approaches for CRC patients, the existing research has proven that Artificial Intelligence (AI) plays an important role in clinical practice [5]. In recent days, the researchers proposed an AI technique to decrease the neglected adenomas rates and then the risk of increasing cancer by enhancing CRC screening results [6]. Characterization systems and Computer-aided detection have gained more attention as well as interest. The AI helps with optical diagnosis and colorectal polyp detection in colonoscopy which may aid endoscopists in making correct and on-time diagnoses. AI is one of the important fields in computer science [7]. It is committed to developing smart machines, which are proficient in performing tasks that usually need human-level intelligence. There are numerous AI applications around us, so it is very difficult to recognize and estimate their effect on current society [8]. Recently, the impact of Deep Learning (DL) and Support Vector Machine (SVM) model are played a vital role in healthcare and medicine structures. In the medical domain, AI applications can be a two major parts namely physical and virtual. DL and Machine learning (ML), which is a subcategory of ML establish the effective part of AI [9]. Further, ML techniques have been categorized like supervised, unsupervised learning and reinforcement learning (RL). Most of the important DL systems and Convolutional Neural Networks (CNNs) symbolize a certain category of multilayer artificial neural network (ANN), which can be very effective for image classification [10].

This research designs a new Reptile Search Algorithm with Deep Learning for Colorectal Cancer Detection and Classification (RSADL-CCDC) technique in the cloud environment. The main aim of the RSADL-CCDC method focuses on the automatic identification and classification of the CC in the cloud platform. In the presented RSADL-CCDC approach, the medical images are stored in the cloud server and the diagnostic process take place in it. For feature extraction, the RSADL-CCDC technique applies ShuffleNetv2 model. Besides, the recognition and classification of CC take place using convolutional autoencoder (CAE) model. Finally, the hyperparameter tuning of the CAE system occurred by employing of RSA. The experimental validation of the RSADL-CCDC approach can be performed on benchmark medical dataset.

2. Related Works

In [11], the two-determination DL network with self-attention mechanism (DRSANet) combines details and then context for CRC binary detection and localisation in Computer-Aided Diagnosis (CAD) and Whole Slide Images (WSIs) is proposed. Two input systems was mainly developed to learn context and details at the same time, and then self-attention appliance was employed to learn dissimilar positions in the images to enhance performance. In [12], a structure is established on several DL techniques are projected. The images are mainly fed into the SqueezeNet, MobileNet as well as ShuffleNet methods. The features are decreased by employing Principal Component Analysis (PCA) and then Fast Walsh–Hadamard Transform (FWHT) models. In addition, Discrete Wavelet Transform (DWT) is mainly utilized to combine the FWHT's removed feature acquired from a 3 DL methods. In addition to that, these DL techniques PCA features are connected. At last, results are fed to the four separate ML methods.

Hamida et al. [13] considered an implementation of DL framework to identify as well as emphasize colon tumour areas in lightly marked histopathological content information. Primarily, advanced CNNs comprising the vgg, ResNet, AlexNet, DenseNet, and Beginning methods are revised and associated. This approach employs the utilization of transmission learning methods. In [14], an highly-efficiency WSI inquiry method for locating cancer areas precisely by a patch-related CNN method is projected. The research uses Monte Carlo adjustable selection for a quick recognition of cancers at slide stage and a conditional random field (CRF) method for mixing space association for well identification precision. Three datasets from The Cancer Genome Atlas (TCGA) are employed to assess.

In [15], a colon cancer recognition system employing a transmission-learning framework to remove top-level features spontaneously from colon surgery images for automatic analysis of victims and diagnosis is projected. In the research, the image features are removed from a pretrained CNN and employed to develop the Bayesian enhanced Care Vector Device classification. Also, VGG-16, Alexnet, and InceptionV3 pretrained NN could be employed. Dif and Elberichi [16] target to invent a novel energetic collective DL approach. Initially, it produces a group of methods based on transfer learning tactics from DNN. Then, applicable subclass of methods is nominated by using the PSO technique and mixed of averaging or voting approaches. The projected method was verified on a histopathological database for CRC identification depends upon seven kinds of CNN. In [17], a ranking-based DNN for cancer grading in HI is used. By employing DNN, HI are charted in a hidden area. It was created based on ranking loss, and triplet loss, as well as developed to optimize the inter-group space between tumour positions in the hidden area regarding the violence of tumour, prominent to the precise arrangement or grade of pathology images. A few colorectal pathology images have been used for assessment.

3. The Proposed Model

We have established an automated CC detection and classification employing the RSADL-CCDC technique in the cloud environment. The main objective of the RSADL-CCDC system focuses on the automatic recognition and classification of the CC. The cloud server executes the proposed model for the detection and classification of CC. In the presented RSADL-CCDC approach, four stages of operations are involved namely BF-based preprocessing, ShuffleNetv2-based feature extraction, CAE-based classification, and RSA-based hyperparameter tuning. Fig. 2 depicts the entire flow of RSADL-CCDC approach.

3.1. BF based Pre-processing

Initially, the BF algorithm is applied to eliminate the noise. BF is a digital image processing approach to denoise or enhance images while retaining important details and edges [18]. It is especially helpful to smooth images without blurring sharp transitions and boundaries between regions or objects within the image. The term "bilateral" represents the fact that the filter considers spatial and intensity data while implementing the smoothing process. The bi-lateral filter evaluates the weighted average of the pixel value within the neighborhood, where the weight can be defined by the spatial and intensity kernels. Pixels that are intensively and spatially closer to the target pixel have high weight, while those that are different have low weight. This implies that the filter would retain fine details and edges meanwhile pixels with considerable intensity differences will less contribute to the averaging.

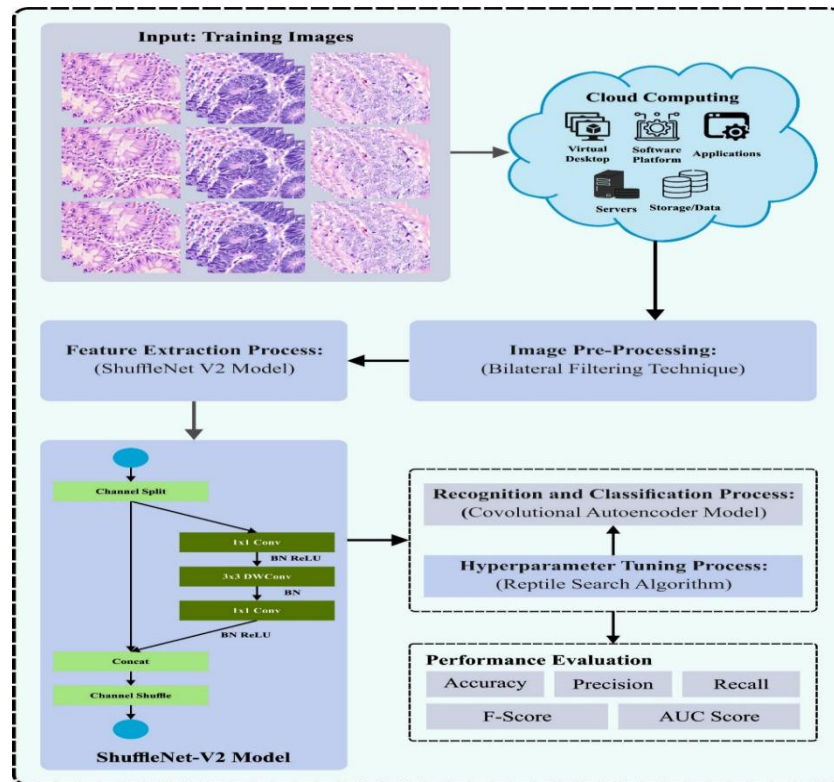


Fig. 2. Overall flow of RSADL-CCDC algorithm

3.2. Feature Extraction

The ShuffleNet_V2 architecture is utilized for producing a set of feature vectors. The major function of the Shufflenet_V2 architecture is the residual block (unit), which comprises a 2 branches [19]. At first, it carries out a channel division at input and splits the input feature maps into a 2 subdivisions; the former has 3 convolutional functions and the next branch doesn't carry out any task, the input and output channel groups of all the branches remain unchanged. Next, the feature map is divided into two branches, the initial branch with 3 convolutional functions and the next branch has 1 depthwise convolution and 1 pointwise convolution. The residual block combines the output feature map of both subdivisions with merging at the output and carrying out channel combined to feature map. Various branches are extracted at random for rearranging into a new feature maps such that the group convolution may combine the input feature in various sets, enhancing the data flow among channels and ensuring to be the input and output channel groups were reconnected. The ShuffleNet_V2 architecture has essentially comprised the MaxPool, Conv5, FC, Conv1, Stage2, Stage3 layer, and Stage4 layers. The Stage2 layer, Stage3 layer, and Stage4 layer includes the superposition of residual units. Particularly, the Stage2 and Stage4 layers are superimposed with overall of 4 residual blocks, as well as the Stage3 layer could be superimposed with overall of 8 residual blocks. The stepsize of initial residual block in all the Stages is 2, the primary objective is to downsample, and the stepsize of other residual blocks is 1. The network with various complexities was intended by shifting the amount of output channel groups in network architecture. According to ShuffleNet_V21, the amount of output channel groups in the Conv1, Max-Pool, Stage2, Stage3, Stage4, Conv5, and FC layers are 24, 24, 116, 232, 464, 1024, and 1000.

3.3. Image Classification

For image classification, the CAE model is exploited. By merging the encoder and decoder with the classifier, CAE can be used for classification. The CAE and classifier are trained end-to-end to minimize the classifier's classification error and the CAE's reconstructed error [20]. This technique may result in higher performance than directly training the classifier on the raw input dataset. The study aims to improve the classifier's

performance and train the classifier and the AE simultaneously. Fig. 3 depicts the infrastructure of CAE. Typically, the loss function of CAE utilized for classification comprises classification and reconstruction loss. The reconstruction loss is used to measure the distinction among the regenerated image and the input image produced by the decoder. The MSE and binary cross-entropy (BCE) loss are the reconstructed loss function. In this work, MSE Loss was used to measure the reconstruction loss of AE.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - Y_i')^2 \quad (1)$$

In Eq. (1), the number of samples or observations in the database is n , the actual value of target parameter for the i^{th} sample is Y_i , and Y_i' are the predictive value of the targeted parameter for the i^{th} sample. A convolutional layer assists in extracting the feature map through the filter on the input. w_i represents the filter, and b_i represents the bias.

$$f_i = \text{ReLU} \left(\sum X * w_i + b_i \right) \quad (2)$$

The data is processed by the activation function after the convolution layer. Now, ReLU is utilized as an activation function. ReLU output zero for negative input value and a similar value for the positive input values. Now, x and $f(x)$ can be the input and the output values.

$$f(x) = \max(0, x) \quad (3)$$

The pooling layer reduces the feature maps' size and reduces the computation difficulty. Various pooling approaches are available; however, the Max-pooling method was preferred in this work. Max-pooling moves a window across the feature map and outputs the maximal value of all the windows.

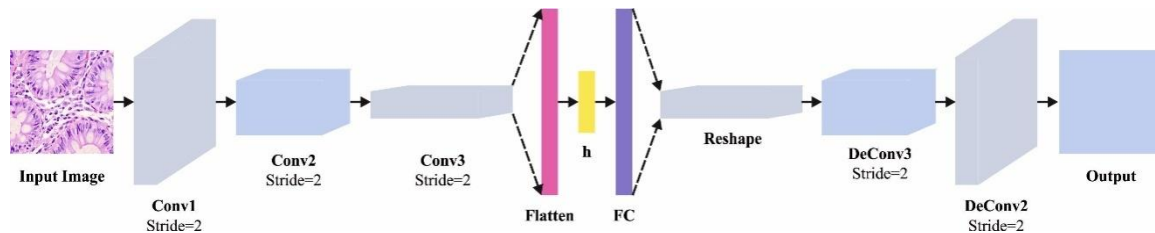


Fig. 3. Architecture of CAE

The bottleneck layer is used to create compressed data representation. Deconvolution is the reverse process of convolution, where the deconvolution kernel is denoted as w' , b is a bias, and i indicates the number of channels.

$$D_i = \text{ReLU} \left(\sum X * w'_i + b_i \right) \quad (4)$$

The encoder layer is used to construct summaries for the hidden layer, the up-sampling model is used in the decoder layer to recreate the original size image using those summaries. Once this classifier and AE are trained, then the feature was learned to compress the data and make accurate class predictions. Thus, the AE feature is expected to improve performance of the classifier performance.

3.4. Hyperparameter Tuning

Finally, the hyperparameter selection of the CAE model is implemented by the use of RSA. It is a new metaheuristic optimization approach that seeks to emulate the natural habitat of crocodile [21]. This algorithm stimulates the hunting strategy of crocodiles that mainly prefer regions with rich food and water sources and capable of hunting inside and outside of the water. The steps for RSA are discussed in the following:

Stage 1: RSA parameter initialization

It is crucial to initialize the algorithmic and control parameters before running the RSA algorithm. The control parameter comprises T , the maximum number of iterations, N , the number of crocodiles (using count of candidate solutions); α , and β , which controls the exploitation and exploration capabilities. During the search process, this parameter is used to balance exploitation and exploration.

Stage2: Population initialization of RSA

By using Eq. (5), a random set of solutions can be initialized:

$$\chi_{ij} = \text{rand} * (\text{UB} - \text{LB}) + \text{LB}, i = 1, 2, \dots, N, \text{ and } j = 1, 2, \dots, n. \quad (5)$$

Now, χ_{ij} represents the j^{th} location of the i^{th} solutions, n denotes the dimensional size of the problem, the random integer within $[0,1]$ is represented by rand , the lower and upper limits of the search space are LB and UB . Therefore, N solution set is produced and stored in the matrix form:

$$X = \begin{bmatrix} x_{1,1} & \dots & x_{1,j} & x_{1,n-1} & x_{1,n} \\ x_{2,1} & \dots & x_{2,j} & x_{2,n-1} & x_{2,n} \\ \vdots & \dots & \vdots & \vdots & \vdots \\ x_{N-1,1} & \dots & x_{N-1,j} & x_{N-1,n-1} & x_{N-1,n} \\ x_{N,1} & \dots & x_{N,j} & x_{N,n-1} & x_{N,n} \end{bmatrix} \quad (6)$$

Stage 3: Fitness function assessment

Where X is the fitness values of solution χ_{ij} in the population, calculated as $f(\chi_{ij})$.

Stage 4: Exploration stage

RSA uses two different strategies namely belly walking and high walking to determine best solution by exploring novel areas in the search range. The updating position of mathematical formula can be given below:

$$x_{i,j}(t+1) = \text{Best}_j(t) * -\eta_{i,j}(t) * \beta - R_{i,j}(t) * \text{randif}(t \leq T/4), \quad (7)$$

and

$$x_{i,j}(t+1) = \text{Best}_j(t) * x_{r_1,j} * \text{ES}(t) * \text{randif}(T/4 < t \leq 2T/4) \quad (8)$$

with $x_{i,j}$ representing the search space of i^{th} solution at j^{th} location. The value of $\text{Best}_j(t)$ corresponds to j^{th} location in the optimum solution attained at t^{th} iteration, $t+1$ indicates the novel iteration, and t shows the prior iteration. The hunting operators of j^{th} location at i^{th} solution, $\eta_{i,j}(t)$. $x_{r_1,j}$ denotes the search space at j^{th} location in the i^{th} solution, where r_1 refers to a value within $[1, N]$. The belly walking strategy has controlled with $T/4 < t \leq 2T/4$, but, a high walking strategy was controlled by $t \leq T/4$. The values of $\eta_{i,j}$, $M(x_i)$, $P_{i,j}$, $R_{i,j}$ and $\text{ES}(t)$ are computed by the following expression:

$$\eta_{i,j} = \text{Best}_j(t) * P_{i,j}, \quad (9)$$

$$M(x_i) = \frac{1}{n} \sum_{i=1}^n x_{(i,j)}, \quad (10)$$

$$P_{i,j} = \alpha + \frac{x_{i,j} - M(x_i)}{Best_j(t) * (UB_j - LB_j) + \epsilon}, \quad (11)$$

$$ES(t) = 2 * r_3 * (1 - 1/T), \quad (12)$$

and

$$R_{i,j} = \frac{Best_j(t) - x_{r_2',j}}{Best_j(t) + \epsilon}. \quad (13)$$

Where, the percentage difference between the search space at j^{th} location of the best solution ($Best(t)$) and the search space at the location of the existing solution (x) is represented by $P_{i,j}$, the parameter α controls the exploration capability of RSA, with value of $\alpha = 0.1$. Furthermore, e is a random integer range within $[0,2]$, and $M(X)$ indicates the average value of each search space of the existing solutions. The $R_{i,j}$ parameter lessen the decision variable area of j^{th} location in the i^{th} solution. The evolutionary sense probability, $ES(t)$, is arbitrarily allocated a value reducing from 2 to -2 , and is computed by Eq. (9). The parameter r_2 has a random integer within $[1, N]$ and r_3 is a random number that lies in $(-1, 0, \text{ and } 1)$.

Stage 5: Exploitation stage

Using hunting cooperation and hunting coordination strategies, this stage exploits present search area, to find optimum solution as follows:

$$x_{i,j}(t+1) = Best_j(t) * P_{i,j} * randif(2T/4 \leq t \leq 3T/4), \quad (14)$$

and

$$x_{i,j}(t+1) = Best_j(t) - \eta_{i,j} * \epsilon - R_{i,j} * randif(3T/4 \leq t \leq T). \quad (15)$$

The hunting cooperation is utilized at the time interval $3T/4 \leq t \leq T$, while the hunting coordination is used at the time interval $2T/4 \leq t \leq 3T/4$.

Stage 6: Stopping condition

This procedure is reiterated, until the maximum iteration is attained.

The RSA method derives a FF to accomplish high effectiveness of classifier. It defines a positive integer to characterize the superior values of the solution candidate. Here, the failure of classifier error rate can be assumed as FF.

$$\begin{aligned} \text{fitness}(x_i) &= \text{ClassifierErrorRate}(x_i) \\ &= \frac{\text{number of misclassified samples}}{\text{Total number of samples}} * 100 \end{aligned} \quad (16)$$

4. Results and Discussion

The CC detection and classification performance of the RSADL-CCDC method could be validated on the Kaggle datasets [22], comprising 10000 instances with two classes as defined in Table 1.

Table 1 Details on database

Class Names	Description	No. of Instances
Col-Ad	Colon Adenocarcinoma	5000
Col-Be	Colon Benign Tissue	5000
Total Number of Instances		10000

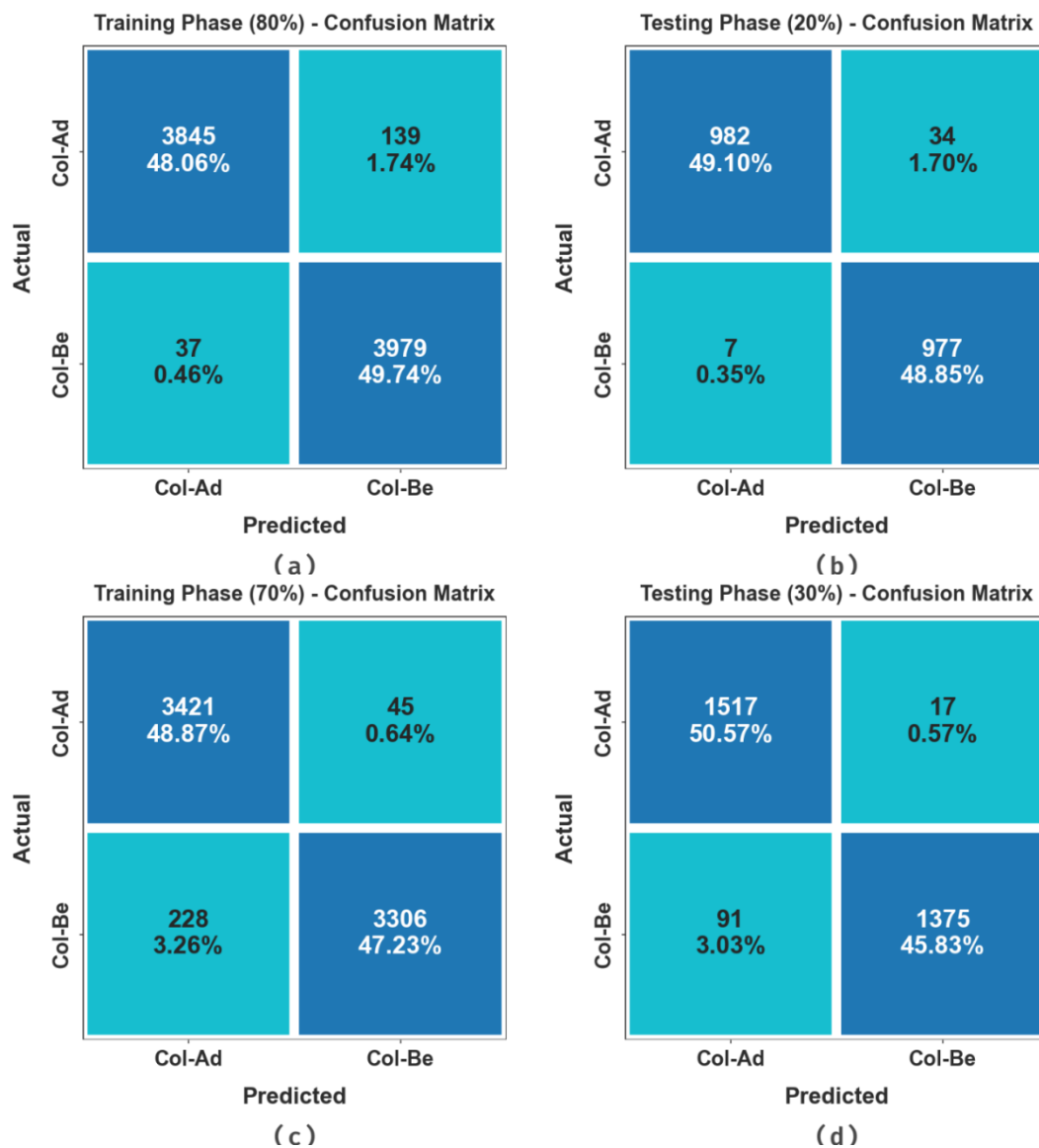
**Fig. 4.** Confusion matrices of (a-c) TR phase of 80% and 70% and (b-d) TS phase of 20% and 30%

Fig. 4 shows the confusion matrices formed by the RSADL-CCDC technique at 80:20 and 70:30 of TR phase/TS phase. The simulated values exhibit the effectual recognition of the Col-Ad and Col-Be samples with all 2 classes.

The CC detection results of the RSADL-CCDC technique with 80:20 of TR Phase/TS Phase are reported in Table 2 and Fig. 5. The simulated values pointed out the RSADL-CCDC technique properly categorizes the samples. With 80% of TR Phase, the RSADL-CCDC system offers average accu_y of 97.79%, prec_n of 97.84%, reca_l of 97.79%, F_{score} of 97.80%, and $\text{AUC}_{\text{score}}$ of 97.79%. Along with that, based on 20% of TS Phase, the RSADL-CCDC model offers average accu_y of 97.97%, prec_n of 97.96%, reca_l of 97.97%, F_{score} of 97.95%, and $\text{AUC}_{\text{score}}$ of 97.97% correspondingly.

Table 2 CC detection outcome of RSADL-CCDC algorithm on 80:20 of TR phase/TS phase

Class Labels	Accu_y	Prec_n	Reca_l	F_{score}	$\text{AUC}_{\text{score}}$
TR Phase (80%)					
Col-Ad	96.51	99.05	96.51	97.76	97.79
Col-Be	99.08	96.62	99.08	97.84	97.79
Average	97.79	97.84	97.79	97.80	97.79
TS Phase (20%)					
Col-Ad	96.65	99.29	96.65	97.96	97.97
Col-Be	99.29	96.64	99.29	97.94	97.97
Average	97.97	97.96	97.97	97.95	97.97

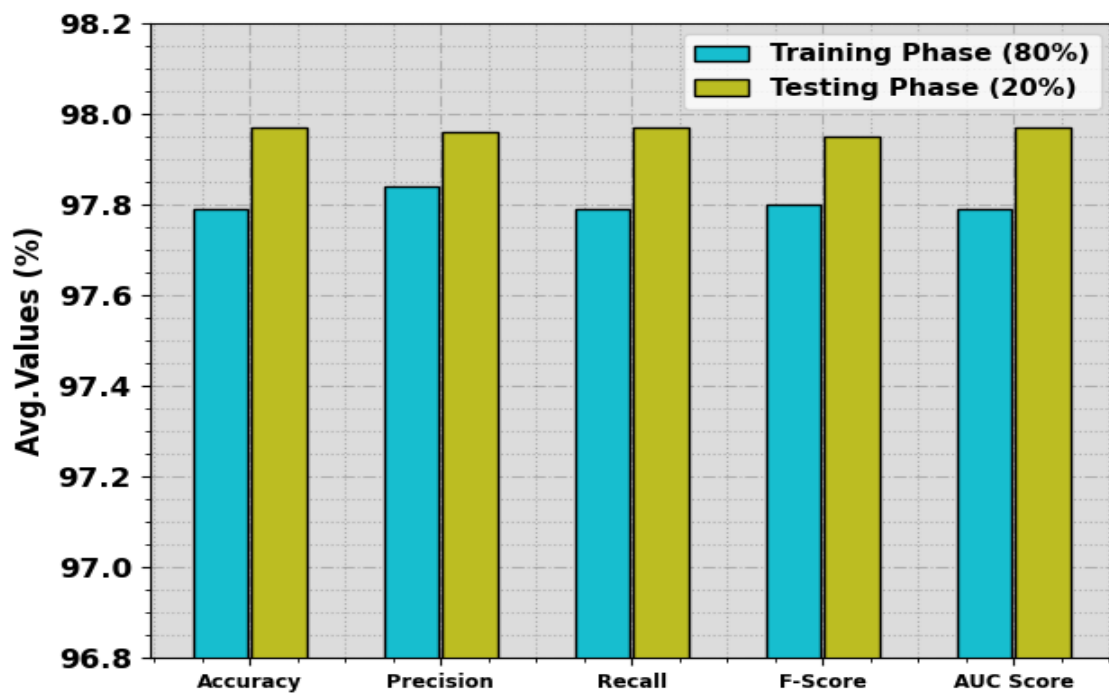


Fig. 5. Average of RSADL-CCDC algorithm on 80:20 of TR phase/TS phase

The CC detection results of the RSADL-CCDC method with 70:30 of TR Phase/TS Phase are described in Table 3 and Fig. 6. The simulated values reported that the RSADL-CCDC system appropriately categorizes the samples. With 70% of TR Phase, the RSADL-CCDC methodology offers average $accu_y$ of 96.13%, $prec_n$ of 96.20%, $reca_l$ of 96.13%, F_{score} of 96.10%, and AUC_{score} of 96.13%. Besides, with 30% of TS Phase, the RSADL-CCDC methodology gives average $accu_y$ of 96.34%, $prec_n$ of 96.56%, $reca_l$ of 96.34%, F_{score} of 96.39%, and AUC_{score} of 96.34% respectively.

Table 3 CC detection outcome of RSADL-CCDC algorithm on 70:30 of TR phase/TS phase

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	AUC_{score}
TR Phase (70%)					
Col-Ad	98.70	93.75	98.70	96.16	96.13
Col-Be	93.55	98.66	93.55	96.03	96.13
Average	96.13	96.20	96.13	96.10	96.13
TS Phase (30%)					
Col-Ad	98.89	94.34	98.89	96.56	96.34
Col-Be	93.79	98.78	93.79	96.22	96.34
Average	96.34	96.56	96.34	96.39	96.34

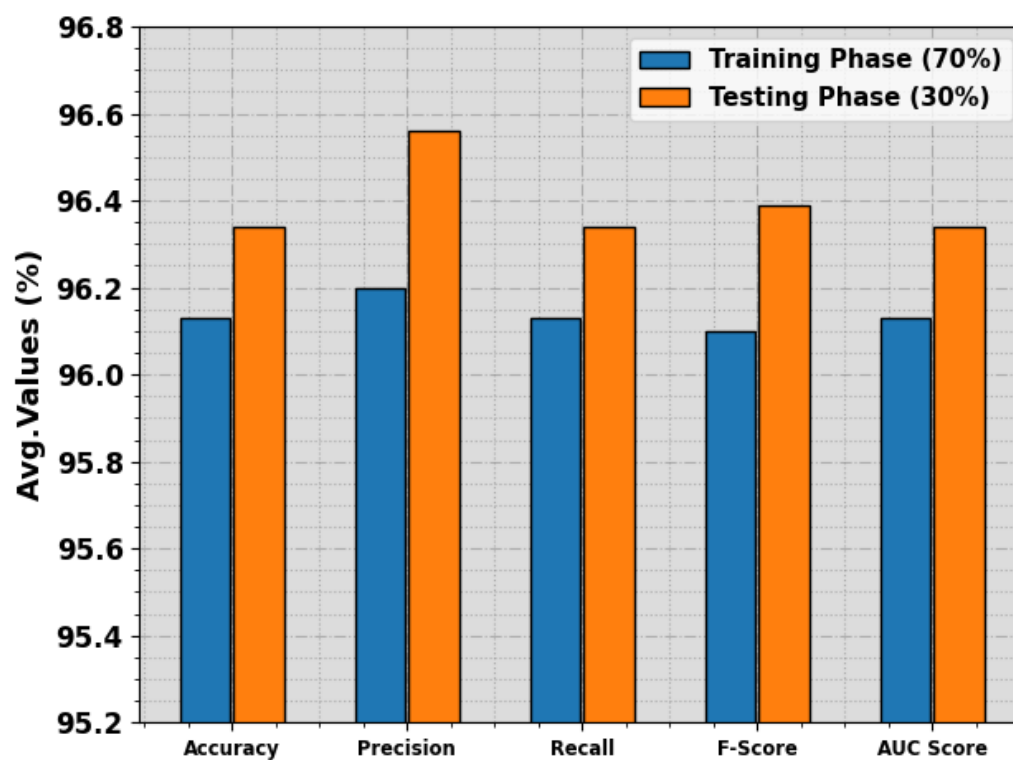


Fig. 6. Average of RSADL-CCDC algorithm at 70:30 of TR phase/TS phase

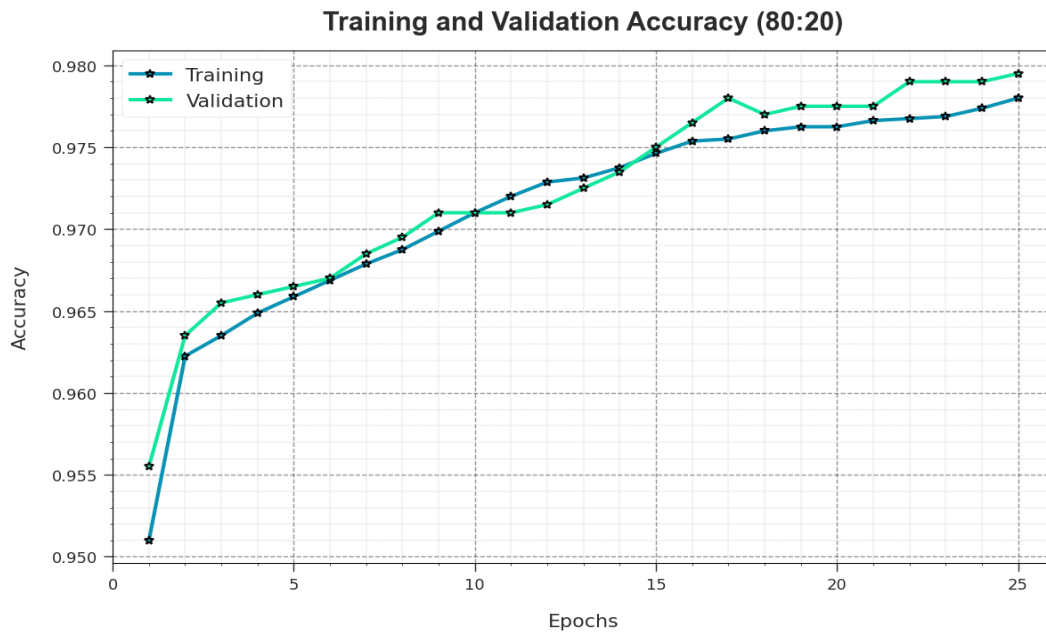


Fig. 7. Accu_y curve of RSADL-CCDC algorithm on 80:20 of TR phase/TS phase

To calculate the performance of the RSADL-CCDC technique with 80:20 of TR Phase/TS Phase, TR and TS accu_y curves are determined, as illustrated in Fig. 7. The TR and TS accu_y curves exhibit the performance of the RSADL-CCDC method over numerous epochs. The figure offers important details about the learning tasks and generalization capabilities of the RSADL-CCDC system. With an improvement in epoch count, it is observed that the TR and TS accu_y curves acquire enhanced. It is noticed that the RSADL-CCDC algorithm attains improved testing accuracy that can potentially recognize the patterns in the TR and TS data.

Fig. 8 shows the overall TR and TS loss values of the RSADL-CCDC system with 80:20 of TR Phase/TS Phase over epochs. The TR loss reveals the model loss is reduced over epochs. Mainly, the loss values become decreased as the model adapts the weight for diminishing the predicted error on the TR and TS data. The loss curves exhibit the extent to where the model is fitting the training data. It is evidenced that the TR and TS loss is gradually reduced and described that the RSADL-CCDC methodology successfully learns the patterns represented in the TR and TS data. It is also remarked that the RSADL-CCDC approach changes the parameters to lessen the difference among the actual and predicted training label.

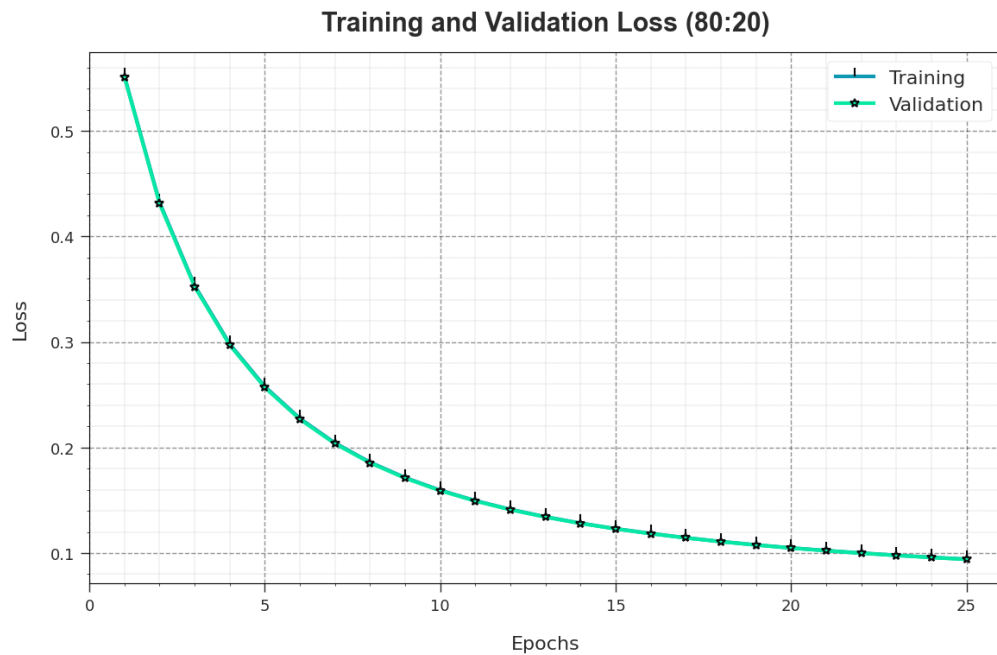


Fig. 8. Loss curve of RSADL-CCDC method with 80:20 of TR phase/TS phase

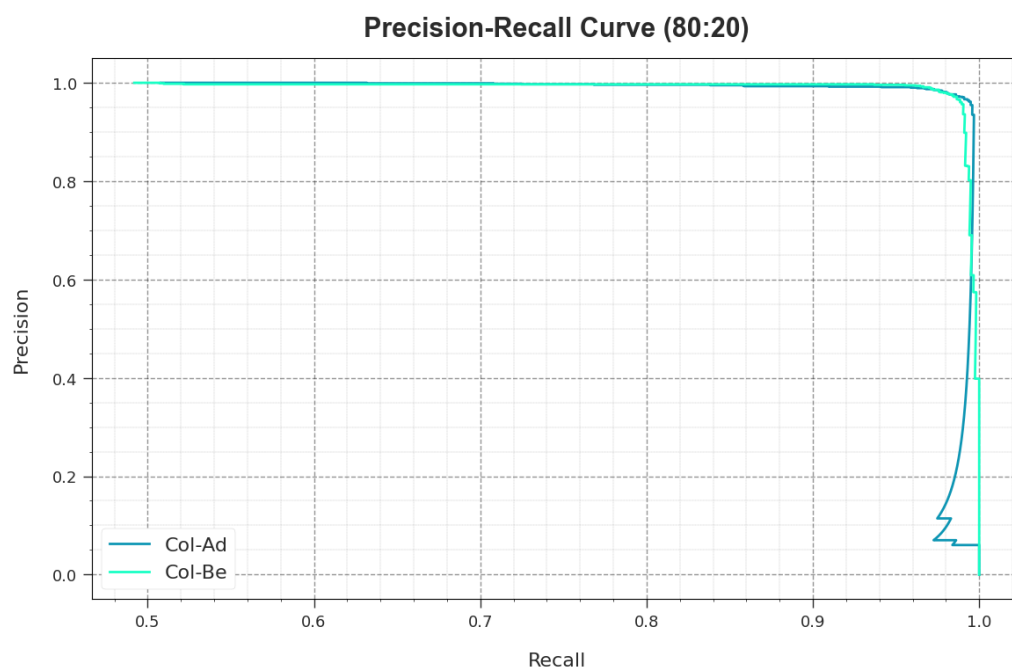


Fig. 9. PR curve of RSADL-CCDC algorithm with 80:20 of TR phase/TS phase

The PR curve of the RSADL-CCDC technique with 80:20 of TR Phase/TS Phase is exhibited by plotting precision against recall as shown in Fig. 9. The simulated values confirm that the RSADL-CCDC approach gets improved PR values with each 2 class. The figure describes that the model learns to recognize different class

labels. The RSADL-CCDC methodology achieves improved outcomes in the recognition of positive samples with lower false positives.

The ROC analysis provided by the RSADL-CCDC method with 80:20 of TR Phase/TS Phase is exhibited in Fig. 10, which has the ability the differentiation of the class labels. The figure specifies valuable insights into the trade-off among the TPR and FPR rates over various classification thresholds and changing numbers of epochs. It offers accurately predicted performance of the RSADL-CCDC system on the classification of separate 2 classes.

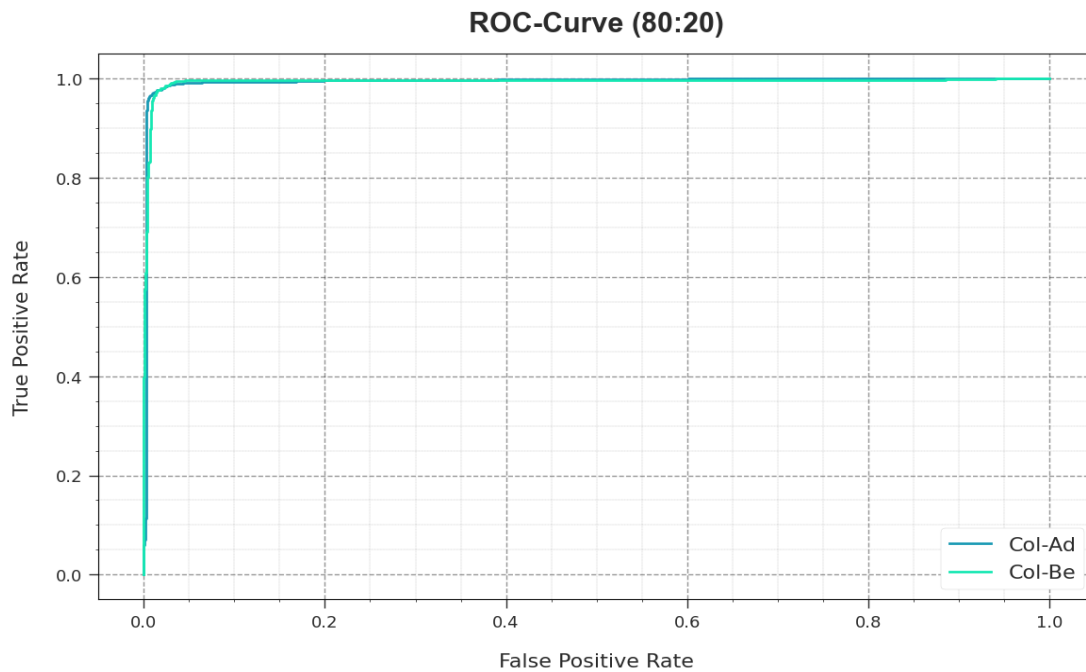


Fig. 10. Loss curve of RSADL-CCDC systemwith 80:20 of TR phase/TS phase

In Table 4 and Fig. 11, the comparative analysis of the RSADL-CCDC technique is confirmed. The simulated values show that the mSRC model leads to poorer performance. At the same time, the ResNet-50, DenseNet169-SVM, and VGG-16 models have shown slightly improved performance. Meanwhile, the CNN and DL methodologies have gained considerable performance. Nevertheless, the RSADL-CCDC system illustrates maximum performance with accu_y of 97.97%, prec_n of 97.96%, reca_l of 97.97%, and F_{score} of 97.95%. These simulated values confirmed the enriched performance of the RSADL-CCDC method over other models.

Table 4 Comparative analysis of RSADL-CCDC model with existing systems

Methods	Accu_y	Prec_n	Reca_l	F_{score}
mSRC Algorithm	88.21	85.21	91.78	86.78
RESNET-50	93.64	96.12	97.49	96.94
CNN method	97.11	97.07	97.44	97.61
DL technique	96.32	96.86	96.44	97.08
DenseNet169 and SVM	92.08	95.82	95.67	96.15

VGG-16 Model	91.09	95.08	96.53	97.00
RSADL-CCDC	97.97	97.96	97.97	97.95

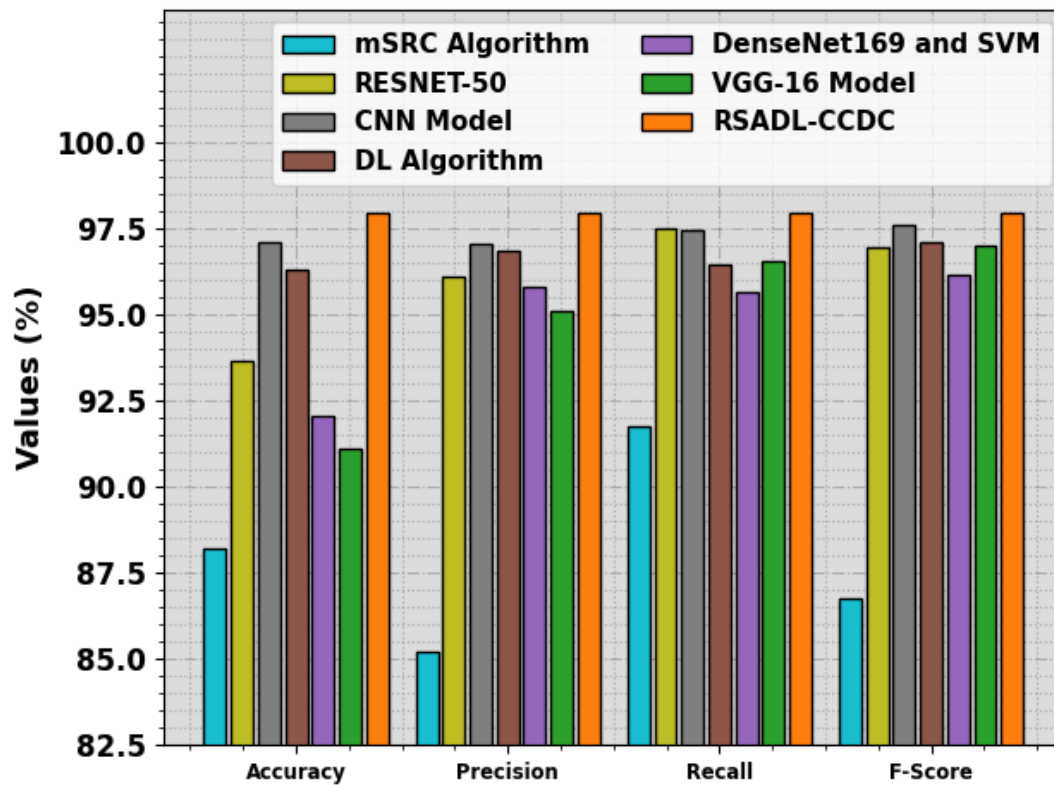


Fig. 11. Comparative outcome of RSADL-CCDC approach with existing methods

5. Conclusion

In this study, we have developed an automatic cloud assisted CC detection and classification by employing the RSADL-CCDC technique. The main aim of the RSADL-CCDC technique focuses on the automated recognition and classification of the CC in the cloud environment. In the presented RSADL-CCDC approach, four stages of operations are involved namely BF-based preprocessing, ShuffleNetv2 based feature extraction, CAE based classification, and RSA based hyperparameter tuning. In this work, the RSADL-CCDC technique applies ShuffleNetv2 model for feature extraction and CAE is employed for CC classification. Furthermore, the recognition and classification of CC take place using the CAE model and the hyperparameter tuning method is carried out by the use of RSA. The experimental validation of the RSADL-CCDC technique can be performed on benchmark medical dataset. Wide-ranging outcomes stated the enhanced performance of the RSADL-CCDC system on CC recognition over other models with respect to various assessment. In future, the performance of the RSADL-CCDC method is tuned by ensemble voting classifier.

6. References

- [1] Raju, A.S.N., Jayavel, K. and Rajalakshmi, T., 2022. An IoT-Enabled Healthcare System: Auto-predictive Colorectal Cancer with Colonoscopy Images Combined with the Convolutional Neural Network. In High Performance Computing and Networking: Select Proceedings of CHSN 2021 (pp. 283-294). Singapore: Springer Singapore.

- [2] Xu, B. and Zhou, F., 2022. The roles of cloud-based systems on the cancer-related studies: a systematic literature review. *IEEE Access*.
- [3] Kaur, H., Jameel, R., Alam, M.A., Alankar, B. and Chang, V., 2023. Securing and managing healthcare data generated by intelligent blockchain systems on cloud networks through DNA cryptography. *Journal of Enterprise Information Management*.
- [4] Javed, S., Mahmood, A., Fraz, M.M., Koohbanani, N.A., Benes, K., Tsang, Y.W., Hewitt, K., Epstein, D., Snead, D. and Rajpoot, N., 2020. Cellular community detection for tissue phenotyping in colorectal cancer histology images. *Medical image analysis*, 63, p.101696.
- [5] Masud, M., Sikder, N., Nahid, A.A., Bairagi, A.K. and AlZain, M.A., 2021. A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework. *Sensors*, 21(3), p.748.
- [6] Lorenzovici, N., Dulf, E.H., Mocan, T. and Mocan, L., 2021. Artificial Intelligence in Colorectal Cancer Diagnosis Using Clinical Data: Non-Invasive Approach. *Diagnostics*, 11(3), p.514.
- [7] Sarwinda, D., Paradisa, R.H., Bustamam, A. and Anggia, P., 2021. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Computer Science*, 179, pp.423-431.
- [8] Zhou, C., Jin, Y., Chen, Y., Huang, S., Huang, R., Wang, Y., Zhao, Y., Chen, Y., Guo, L. and Liao, J., 2021. Histopathology classification and localization of colorectal cancer using global labels by weakly supervised deep learning. *Computerized Medical Imaging and Graphics*, 88, p.101861.
- [9] Tsai, M.J. and Tao, Y.H., 2021. Deep learning techniques for the classification of colorectal cancer tissue. *Electronics*, 10(14), p.1662.
- [10] Alqudah, A.M. and Alqudah, A., 2022. Improving machine learning recognition of colorectal cancer using 3D GLCM applied to different color spaces. *Multimedia Tools and Applications*, 81(8), pp.10839-10860.
- [11] Xu, Y., Jiang, L., Huang, S., Liu, Z. and Zhang, J., 2023. Dual resolution deep learning network with self-attention mechanism for classification and localisation of colorectal cancer in histopathological images. *Journal of clinical pathology*, 76(8), pp.524-530.
- [12] Attallah, O., Aslan, M.F. and Sabanci, K., 2022. A framework for lung and colon cancer diagnosis via lightweight deep learning models and transformation methods. *Diagnostics*, 12(12), p.2926.
- [13] Hamida, A.B., Devanne, M., Weber, J., Truntzer, C., Derangère, V., Ghiringhelli, F., Forestier, G. and Wemmert, C., 2021. Deep learning for colon cancer histopathological images analysis. *Computers in Biology and Medicine*, 136, p.104730.
- [14] Ke, J., Shen, Y., Guo, Y., Wright, J.D., Jing, N. and Liang, X., 2020, August. A high-throughput tumor location system with deep learning for colorectal cancer histopathology image. In *International Conference on Artificial Intelligence in Medicine* (pp. 260-269). Cham: Springer International Publishing.
- [15] Babu, T., Singh, T., Gupta, D. and Hameed, S., 2021. Colon cancer prediction on histological images using deep learning features and Bayesian optimized SVM. *Journal of Intelligent & Fuzzy Systems*, 41(5), pp.5275-5286.
- [16] Dif, N. and Elberrichi, Z., 2020. A new deep learning model selection method for colorectal cancer classification. *International Journal of Swarm Intelligence Research (IJSIR)*, 11(3), pp.72-88.
- [17] Le Vuong, T.T., Kim, K., Song, B. and Kwak, J.T., 2021. Ranking loss: a ranking-based deep neural network for colorectal cancer grading in pathology images. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII* 24 (pp. 540-549). Springer International Publishing.
- [18] Asokan, A. and Anitha, J., 2020. Adaptive Cuckoo Search based optimal bilateral filtering for denoising of satellite images. *ISA transactions*, 100, pp.308-321.

- [19] Meng, L., Cui, X., Liu, R., Zheng, Z., Shao, H., Liu, J., Peng, Y. and Zheng, L., 2023. Research on Metallurgical Saw Blade Surface Defect Detection Algorithm Based on SC-YOLOv5. *Processes*, 11(9), p.2564.
- [20] Arslan, N.N., Ozdemir, D. and Temurtas, H., 2023. ECG heartbeats classification with dilated convolutional autoencoder. *Signal, Image and Video Processing*, pp.1-10.
- [21] Krishanthi, G., Jayetileke, H., Wu, J., Liu, C. and Wang, Y.G., 2023. Enhancing Feature Selection Optimization for COVID-19 Microarray Data. *COVID*, 3(9), pp.1336-1355.
- [22] <https://www.kaggle.com/datasets/andrewmvd/lung-and-colon-cancer-histopathological-images>
- [23] Abdullah, S. and Ragab, M., 2023. Tunicate swarm algorithm with deep convolutional neural network-driven colorectal cancer classification from histopathological imaging data. *Electronic Research Archive*, 31(5), pp.2793-2812.
- [24] Sakr, A.S., Soliman, N.F., Al-Gaashani, M.S., Plawiak, P., Ateya, A.A. and Hammad, M., 2022. An efficient deep learning approach for colon cancer detection. *Applied Sciences*, 12(17), p.8450.
- [25] AlGhamdi, R., Asar, T.O., Assiri, F.Y., Mansouri, R.A. and Ragab, M., 2023. Al-Biruni Earth Radius Optimization with Transfer Learning Based Histopathological Image Analysis for Lung and Colon Cancer Detection. *Cancers*, 15(13), p.3300.