



## OPEN Advances in colorectal cancer diagnosis using optimal deep feature fusion approach on biomedical images

Sultan Refa Alotaibi<sup>1</sup>, Manal Abdullah Alohal<sup>2</sup>✉, Mashael Maashi<sup>3</sup>, Hamed Alqahtani<sup>4</sup>, Moneerah Alotaibi<sup>1</sup> & Ahmed Mahmud<sup>5</sup>

Colorectal cancer (CRC) is the second popular cancer in females and third in males, with an increased number of cases. Pathology diagnoses complemented with predictive and prognostic biomarker information is the first step for personalized treatment. Histopathological image (HI) analysis is the benchmark for pathologists to rank colorectal cancer of various kinds. However, pathologists' diagnoses are highly subjective and susceptible to inaccurate diagnoses. The improved diagnosis load in the pathology laboratory, incorporated with the reported intra- and inter-variability in the biomarker assessment, has prompted the quest for consistent machine-based techniques to be integrated into routine practice. In the healthcare field, artificial intelligence (AI) has achieved extraordinary achievements in healthcare applications. Lately, computer-aided diagnosis (CAD) based on HI has progressed rapidly with the increase of machine learning (ML) and deep learning (DL) based models. This study introduces a novel Colorectal Cancer Diagnosis using the Optimal Deep Feature Fusion Approach on Biomedical Images (CCD-ODFFBI) method. The primary objective of the CCD-ODFFBI technique is to examine the biomedical images to identify colorectal cancer (CRC). In the CCD-ODFFBI technique, the median filtering (MF) approach is initially utilized for noise elimination. The CCD-ODFFBI technique utilizes a fusion of three DL models, MobileNet, SqueezeNet, and SE-ResNet, for feature extraction. Moreover, the DL models' hyperparameter selection is performed using the Osprey optimization algorithm (OOA). Finally, the deep belief network (DBN) model is employed to classify CRC. A series of simulations is accomplished to highlight the significant results of the CCD-ODFFBI method under the Warwick-QU dataset. The comparison of the CCD-ODFFBI method showed a superior accuracy value of 99.39% over existing techniques.

**Keywords** Colorectal Cancer, Biomedical images, Osprey optimization Algorithm, Feature Fusion, Computer-aided diagnosis

Colorectal cancer (CRC) is ranked third for incidence (6.1%) and second in fatality rate (9.2%) around the world. With regard to new cases and deaths, the global CRC burden will continue to be increased by 60% in 2030<sup>1</sup>. The prompt and accurate diagnosis of CRC is crucial to optimize the efficiency of treatment and survivorship. The existing CRC diagnoses need visual monitoring with the help of skilled pathologists. Diagnosis is implemented by digital whole-slide images (WSIs) of the hematoxylin and eosin (H&E)-stained samples attained from frozen or formalin-fixed paraffin-embedded (FFPE) tissues<sup>2</sup>. In general, it is made through pathologists that automatically analyze the HIs of CRC tissues that remain the standard for tumour diagnosis and staging. On the other hand, the training, time assessment/pressure condition, and experience of pathologists may result in diagnosis judgment<sup>3</sup>. Thus, the automated classification of CRC for reasonable evaluation has considerable pathological significance. The pathologist's state and the specific disease entities may vary according to the

<sup>1</sup>Department of Computer Science, College of Science and Humanities Dawadmi, Shaqra University, Shaqra, Saudi Arabia. <sup>2</sup>Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia. <sup>3</sup>Department of Software Engineering, College of Computer and Information Sciences, King Saud University, PO Box 103786, Riyadh 11543, Saudi Arabia. <sup>4</sup>Department of Information Systems, College of Computer Science, Center of Artificial Intelligence, Unit of Cybersecurity, King Khalid University, Abha, Saudi Arabia. <sup>5</sup>Research Center, Future University in Egypt, New Cairo 11835, Egypt. ✉email: maalohaly@pnu.edu.sa

pathologist's experience<sup>4</sup>. Moreover, the circling determination of cancer region and tumour content on H&E-stained samples to detect lesion regions for the downstream genomic analysis is a vital pre-analytical step to guarantee precise determination and improve the tumour content of genetic variants<sup>5</sup>. Early diagnosis of cancer has contributed significantly to diagnoses and increased the survival rate. Medical imaging techniques assist in earlier diagnosis and detection of cancer. As a result, medical imaging has been utilized to diagnose, detect, and classify tumours<sup>6</sup>.

The medical imageries are manually interpreted, but it is a time-consuming and very tedious process and may have errors because humans suffer from distractions, fatigue, and so on. This resulted in the implementation of a computer-aided diagnosis (CAD) scheme in the early 80s that assisted clinicians and doctors in interpreting medical images<sup>7</sup>. Machine Learning (ML) approaches and medical images are utilized in CAD systems. In ML methods, the feature extraction of images is the key step. Several researchers have identified various feature extraction methods for different medical imaging modalities and multiple kinds of cancer<sup>8</sup>. Deep learning (DL) is a state-of-the-art technique considered an advancement of ML because it uses numerous layers of NN to learn and progressively extract abstract features to decrease human interference in detecting image classes<sup>9</sup>. Lately, Convolutional neural networks (CNN) have shown promising solutions in image classification in the DL field, where an NN might have hundreds or dozens of layers to learn images with various characteristics. A convolution layer has a small kernel for generating advanced features and passes over an activation function as an output<sup>10</sup>. The major benefit of using CNN over a classical NN is that the model parameter is reduced for accurate outcomes.

This study introduces a novel Colorectal Cancer Diagnosis using the Optimal Deep Feature Fusion Approach on Biomedical Images (CCD-ODFFBI) method. The primary objective of the CCD-ODFFBI technique is to examine the biomedical images to identify colorectal cancer (CRC). In the CCD-ODFFBI technique, the median filtering (MF) approach is initially utilized for noise elimination. The CCD-ODFFBI technique utilizes a fusion of three DL models, MobileNet, SqueezeNet, and SE-ResNet, for feature extraction. Moreover, the DL models' hyperparameter selection is performed using the Osprey optimization algorithm (OOA). Finally, the deep belief network (DBN) model is employed to classify CRC. A series of simulations highlights the significant results of the CCD-ODFFBI method under the Warwick-QU dataset.

## Literature review

Hicham et al.<sup>11</sup> developed a CAD system using CT colonography data to prevent CRC by categorizing scans as polyp or polyp-free. After the preprocessing stage, a DL approach is introduced with two variations: a 3D CNN-BN and a 3D CNN-BN & Dropout. 3D CT scans of the abdomen classification based on the polyps' presence or absence using CNN is primordial to improve the possibility of earlier diagnosis. In<sup>12</sup>, a unique diagnostic tool, such as complementary Artificial Intelligence algorithms and Vision-based Surface Tactile Sensor (VS-TS), was presented. This model used statistical analysis (accuracy, sensitivity, and reliability) and Support Vector Machine (SVM) models. Using the VS-TS, the tumour types are classified by using the SVM model and employing the t-distributed Stochastic Neighbor Embedding approach to effectively detect the difficulty of polyp phantom classified according to the output image. In<sup>13</sup>, proposed a CNN-based DL approach. First, guided image filter and dynamic histogram equalization techniques filter and enhance the CRC images. Afterwards, a Single Shot MultiBox Detector (SSD) is utilized to recognize and categorize colorectal polyps from CRC images effectively. Lastly, a fully connected (FC) layer with dropout is used for the polyp classification. Alqushaibi et al.<sup>14</sup> introduce an effective method for image synthesis and CRC segmentation, integrating an Attention U-Net and Pix2Pix Generative Adversarial Network (Pix2Pix-GAN) method guided by Sine Cosine Algorithm (SCA) for hyperparameter tuning within the GAN model. Using SCA has been instrumental in enhancing the fine balance between discriminator and generator dynamics. Mohamed et al.<sup>15</sup> designed a strong CRC cancer diagnosis technique based on the feature selection approach. Initially, the images' feature is extracted according to the CNN model. Squeezenet, Resnet50, AlexNet, and GoogleNet are utilized for the CNN. Next, the metaheuristic approach is used to reduce the feature count. This study exploits the grasshopper optimization algorithm to select the optimum features from the dataset. Ragab et al.<sup>16</sup> developed an intelligent DL-based CC detection and classification (IDL-CCDC) approach. The proposed method includes a fuzzy filtering method for the noise reduction method. Furthermore, a water wave optimization (WWO) based EfficientNet architecture is applied for feature extraction. Additionally, chaotic glowworm swarm optimization (CGSO) based variational autoencoder (VAE) is employed for the CC segmentation into malignant or benign. The proposed model assists in increasing the overall classifier performance. Alzubaidi et al.<sup>17</sup> present an effective pre-trained model for diagnosing and classifying CRC. First, the benchmark data of Warwick-QU is gathered to estimate effectiveness. Next, preprocessing is performed using noise removal and contrast enhancement. Lastly, the extracted feature outcomes are categorized using Mask Recurrent-CNN (Mask RCNN).

In<sup>18</sup>, a new architecture based on UNet for cancer segmentation is called Fovea-UNet. A pooling operation known as the Fovea Pooling (FP) model is developed to aggregate detailed and non-local context data based on the importance of pixel-level features. Furthermore, the lightweight backbone network with GhostNet is applied to lower the computation cost. Karaman et al.<sup>19</sup> aim to improve real-time colorectal cancer (CRC) polyp detection systems by incorporating the artificial bee colony (ABC) method with the You Only Look Once (YOLO) object detection algorithm. Nur-A-Alam et al.<sup>20</sup> present an ensemble ML approach for detecting colorectal cancer from colonoscopy images, utilizing preprocessing, feature fusion, and an ensemble classifier to improve accuracy in cancer or polyp detection. Obayya et al.<sup>21</sup> present the Biomedical Image Analysis for Colon and Lung Cancer Detection using Tuna Swarm Algorithm with Deep Learning (BICLCD-TSADL) model for colon and lung cancer detection, utilizing Gabor filtering (GF) for preprocessing, GhostNet for feature extraction, AFAO for hyperparameter tuning, and Tuna Swarm Algorithm (TSA) with Echo State Network (ESN) for classification. Pacal and Karaboga<sup>22</sup> improve YOLOv4 for real-time polyp detection by utilizing CSPNet, Mish activation,

and DIoU loss. The study also optimizes YOLOv3/YOLOv4 with ResNet, VGG, DarkNet53, and Transformers, applying data augmentation, ensemble learning (EL), and NVIDIA TensorRT. Raju and Venkatesh<sup>23</sup> introduce an EnsemDeepCADx system by integrating CNNs with TL via BiLSTM and SVM. The model also utilizes pre-trained models, namely AlexNet, DarkNet-19, DenseNet-201, and ResNet-50, and features fusion from various image datasets with multiple CNN ensembles and SVM to optimize accuracy. Abd El-Aziz et al.<sup>24</sup> propose a refined DL methodology for multi-classifying lung and colon cancers, incorporating ResNet-101V2, NASNetMobile, and EfficientNet-B0. Xia, Yun, and Liu<sup>25</sup> propose two essential modules, the Multi-Scale Feature Fusion Block (MSFFB) and the Reducing Difference Block (RDB), to improve feature interactions and long-distance dependencies. Furthermore, Polarized Self-Attention (PSA) and Balancing Attention Module (BAM) refine local regions and improve foreground-background details. Pacal<sup>26</sup> proposes an improved Swin Transformer with a Hybrid Shifted Windows Multi-Head Self-Attention (HSW-MSA) module and a Residual-based MLP (ResMLP) to improve accuracy, reduce memory usage, and speed up training. Karaman et al.<sup>27</sup> present a DL approach using YOLOv5 for polyp detection optimized with the ABC model to enhance activation functions and hyper-parameters. Pacal<sup>28</sup> introduced the Multi-Axis Vision Transformer (MaxViT), optimized for Pap smear data with a lightweight structure for better accuracy and speed. The study improves performance by replacing MBConv blocks with ConvNeXtV2 and MLP blocks with GRN-based MLPs.

Dwivedi, Srivastava, and Pradhan<sup>29</sup> propose a novel nested feature fusion method by utilizing pre-trained EfficientNet models for early detection and classification of colorectal carcinoma. Pacal et al.<sup>30</sup> propose DL methods for reliable polyp detection, improving YOLOv3 and YOLOv4 by integrating CSPNet for real-time performance. The study also uses advanced data augmentation and transfer learning (TL) and replaces activation functions with SiLU for improved detection, along with Ciou as the loss function. Singh et al.<sup>31</sup> propose an ensemble classifier by integrating Random Forest (RF), SVM, and Logistic Regression (LR), using majority voting for predictions. Deep features from lung and colon images are extracted with VGG16 and LBP and then incorporated for classification. Sangeetha et al.<sup>32</sup> present a Multimodal Fusion Deep Neural Network (MFDNN) methodology to integrate medical imaging, genomics, and clinical data for improved lung cancer diagnosis. It also discusses the ethical considerations, validation, and regulatory needs for deploying AI in clinical settings. Poalelungi et al.<sup>33</sup> emphasize the importance of collaboration between physicians and tech experts to leverage AI's potential fully. Sureshkumar et al.<sup>34</sup> present a hybrid CAD model integrating CNN and pruned ensembled extreme learning machine (HCPELM) for breast cancer detection, utilizing ReLU activation and feature extraction with convolutional and fully connected layers. ELM handles classification, and TL mitigates parameters for easier detection. Srivastava, Chauhan, and Pradhan<sup>35</sup> apply EL with Differential Evolution optimization and Condorcet's Jury Theorem for lung and colon cancer detection, improving classification performance and reducing computational efforts. Gowthamy and Ramesh<sup>36</sup> integrate pre-trained DL approaches, namely ResNet-50, InceptionV3, and DenseNet, with the Kernel Extreme Learning Machine (KELM) model for accurate lung cancer diagnosis using histopathology images. Feature fusion improves classification, while the Mutation Boosted Dwarf Mongoose Optimization Algorithm (MB-DMOA) optimizes model parameters for better accuracy and faster convergence. Ho et al.<sup>2</sup> develop and validate a unique AI DL method by incorporating Faster-RCNN for glandular segmentation and a classical ML classifier to assist pathologists in screening colorectal specimens for malignancies, enhancing cancer detection with high sensitivity. Raju et al.<sup>37</sup> developed TumorDiagX, a framework integrating DL and computer vision (CV) for precise cancer detection. The framework computes multiple CNNs, integrates diverse networks to improve detection, and utilizes U-Net for image segmentation to enhance the detection of malignant lesions.

Existing studies mainly depend on limited datasets, which affect the generalization of models to diverse populations and real-world scenarios. Many approaches must also address the challenge of varying image quality and artefacts, which can affect performance. Additionally, more research must be done on incorporating real-time diagnostic capabilities and confirming model interpretability for clinical adoption. These gaps emphasize the requirement for more robust, scalable solutions with improved adaptability and transparency in healthcare applications. Furthermore, there is a requirement for improved generalization of techniques across diverse patient populations and medical conditions to confirm widespread applicability and accuracy.

## Materials and methods

This work introduces a novel CCD-ODFFBI technique. The major aim of the technique is to examine biomedical images for the identification of CRC. It encompasses different processes involved, such as noise removal, feature fusion, hyperparameter selection, and DBN-based CRC classification. Figure 1 establishes the entire flow of the CCD-ODFFBI method.

### Noise removal process

Initially, the CCD-ODFFBI technique utilizes the MF approach for the noise elimination process<sup>38</sup>. This model is chosen for noise elimination due to its simplicity, efficiency, and ability to preserve edges while removing noise. Unlike linear filters, MF is a nonlinear technique that replaces each pixel with its neighbours' median, making it specifically effectual in reducing impulsive noise (e.g., salt-and-pepper noise) without blurring sharp edges. This is crucial in medical image processing, where retaining structural details is vital for precise diagnosis. Moreover, MF is computationally efficient and easy to implement, making it appropriate for real-time applications. Unlike other noise reduction methods like Gaussian or Wiener filtering, MF performs better in dealing with extreme noise values while maintaining essential image features. Figure 2 illustrates the structure of the MF model.

Median filter (MF) is a popular image preprocessing method that eliminates noise while retaining edges. It moves a window of a predetermined size over all the image pixels, replacing the central pixel values with the median pixel value within the window. Unlike the mean filter, which could blur edges, MF is highly efficient in removing salt-and-pepper noise, a well-known type of impulsive noise where the pixel value is considerably

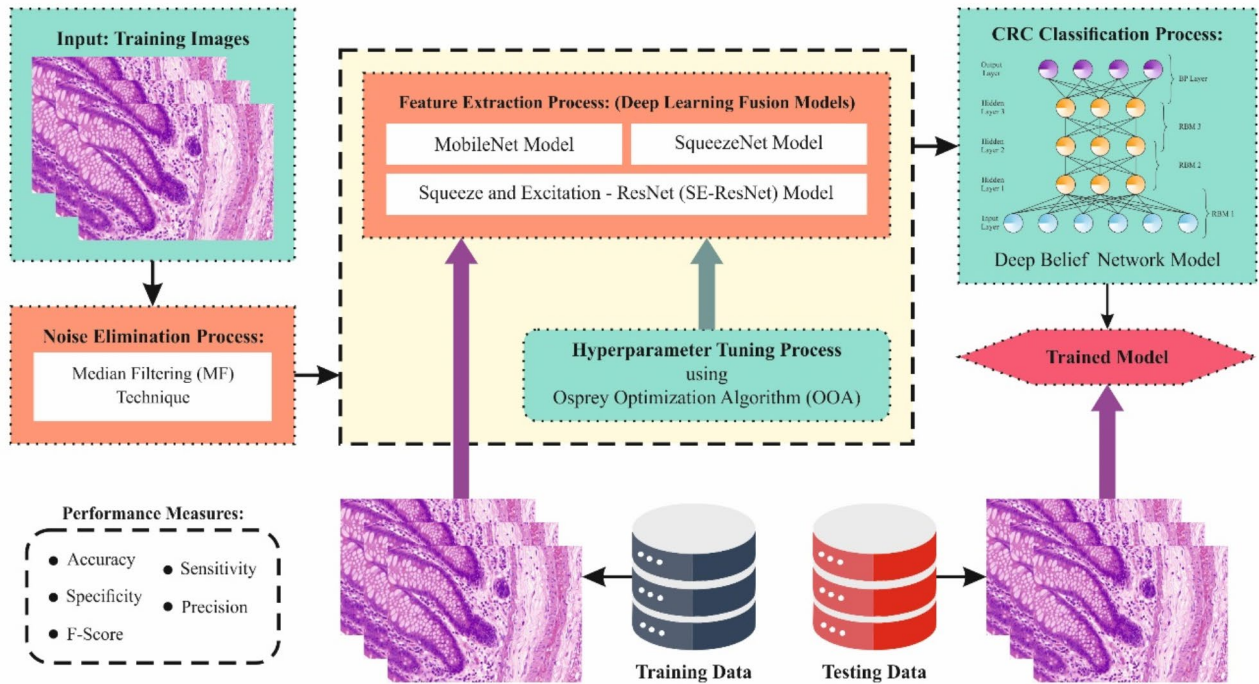


Fig. 1. Working flow of the CCD-ODFFBI method.

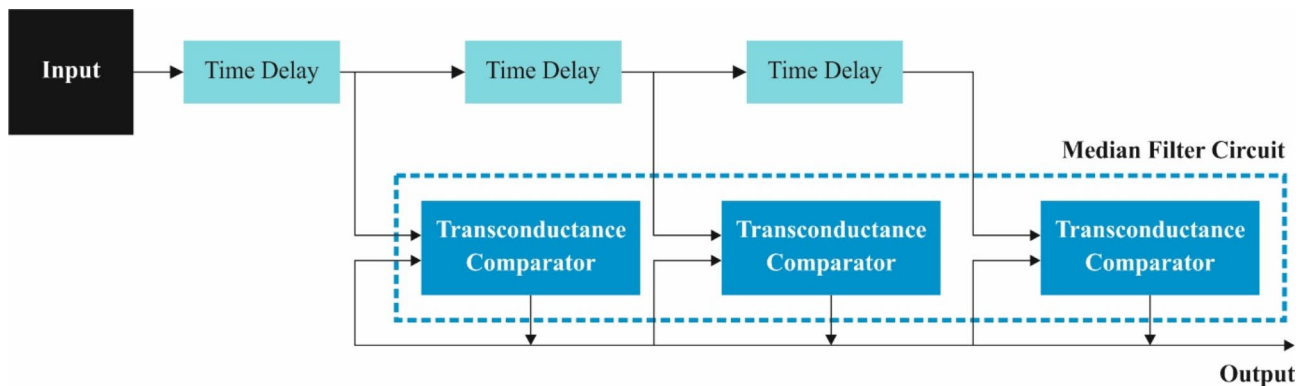


Fig. 2. MF workflow.

distinct from the surroundings. The nonlinear nature of MF enables it to smoothen noise while retaining quick changes between dissimilar regions in an image, which makes it especially helpful in scenarios where edge preservation is crucial.

**Feature extraction process**

For feature extraction, the CCD-ODFFBI technique employs a fusion of three DL methods: MobileNet, SqueezeNet, and SE-ResNet. These techniques are chosen due to the complementary merits of every model in capturing both fine-grained and large-scale features. MobileNet is lightweight and optimized for efficiency, making it appropriate for mobile and edge computing environments. At the same time, SqueezeNet outperforms in mitigating the number of parameters without sacrificing accuracy, ideal for applications requiring low computational resources. With its attention mechanism, SE-ResNet improves feature representation by emphasizing crucial features and suppressing irrelevant ones, resulting in improved performance in complex image recognition tasks. By integrating these models, the system benefits from a balance of efficiency, accuracy, and adaptability, making it more robust than depending on any single model. This fusion approach presents the potential to attain greater performance while minimizing the computational burden, which is critical in real-time applications.

### MobileNet model

MobileNet is a DL method proposed to be employed in low-cost hardware devices<sup>39</sup>. Object classification, identification, and segmentation are executed utilizing the MobileNet method. The MobileNet technique was recognized as MobileNet-V1, and MobileNet-V2 models were developed from the MobileNet-V1. When equated to the MobileNet-V2 method with the previous form, this novel technique provides the main contribution to the issues of linearity among the layers. If a linear bottleneck arises among the layers, the issues are set in this form. Its input size is 224x224 pixels, and its structure contains an in-depth (DW) separable filter. The performance of the model upsurges as it inspects DW, and the input feature is separated into dual layers. All the layers are subdivided into the subsequent ones by uniting them with an output feature till the procedure is done. MobileNet-V2 technique utilizes ReLU among layers. Therefore, it permits the nonlinear output from the preceding layer to be linearly conveyed as an input to the subsequent layers in the future. The method ensures its training procedure till an easy step is made. In this method, the convolution layer spreads filters over input imageries and generates activation mappings. The activation mapping covers the feature and moves to the subsequent layers. Also, the pooling layer is employed in the MobileNet-V2 method. The matrices attained over these layers are transformed into small sizes. The mobileNet-V2 method was used as pre-trained, and the SVM technique was utilized in the stage of classification. Figure 3 depicts the working flow of the MobileNet model.

### SqueezeNet model

SqueezeNet is a thorough learning method of input size 224x224 pixels, including pooling, convolutional, fire, and ReLU layers. The SqueezeNet doesn't contain full connection (FC) and dense layers. On the other hand, the Fire layer executes the function of these parallel layers. The benefit is that it executes surveys effectively by decreasing the number of parameters, thus declining the dimension of model ability. The squeezeNet method formed further effective outcomes, reducing the model cost. While the data regarding the layers is given in the MobileNet-V2 method, the Fire (F2, F3, and F9) layers look like a novel layer containing dual parts, i.e. Expansion and Compression. This method utilizes a 1x1 convolution filter to input imagery in the Compression part. Meanwhile, the expansion part utilizes 1x1 and 3x3 convolution filters for the input imagery. The Expansion and Compression retain similar sizes of feature maps. In the Compression part, an input image depth is decreased and enlarged. Next, the depth is enlarged in the Expansion part. Figure 4 demonstrates the SqueezeNet architecture.

### SE-ResNet Model

The SE attention module is a sign of channel attention<sup>40</sup>. It mainly concentrates on the issue of interdependence among networks. The convolution (Conv) function initially combines every input channel and then totals the outcomes of Conv for every channel. This permits the spatial feature to be merged with the channel's feature, resulting in a main diverse feature set. The SE element extracts this confusion and lets the DL technique absorb the channel feature straight. SE takes out the interdependency among feature channels, attaining the significance of every network. Every channel feature is biased, emphasizing significant features and overwhelming secondary ones. The SE module is effortlessly accessed to other network structures. This module contains three essential processes: excitation, compression, and measurement.  $W$ ,  $H$ , and  $C$  denote width, height, and channels, correspondingly. Figure 5 shows the architecture of the SE-ResNet model.

Squeeze uses global pooling to reduce the spatial features of every network into a particular overall feature, efficiently incorporating the data from every network feature. Then, an FC layer is united to evaluate the importance of every channel depending upon the compacted global features attained. The weight value for every channel is defined by the SE module, which is multiplied by the matrix equivalent to the separate network in the new feature mapping. ResNet has a standard application in feature extraction through various areas. The ResNet module uses the shallow feature to get additional vital features. The residual element is employed as the foremost feature extractor structure in feature detection and identification tasks. After conducting numerous experiments, it was definite that the SE attention module and ResNet50 model should be employed in this research work:

### Hyperparameter selection

Meanwhile, the DL models' hyperparameter selection is performed using OOA. The OOA comprises two stages: the primary stage contains osprey finding the place of the fish and catching it (global exploration), and

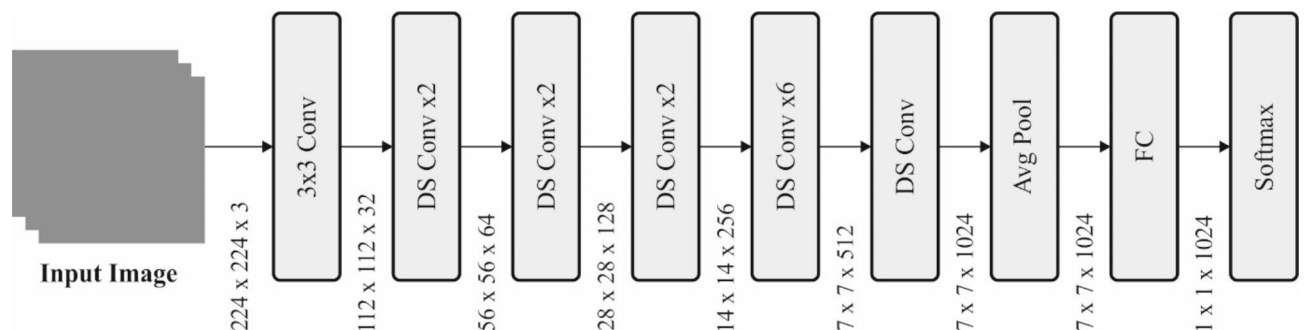
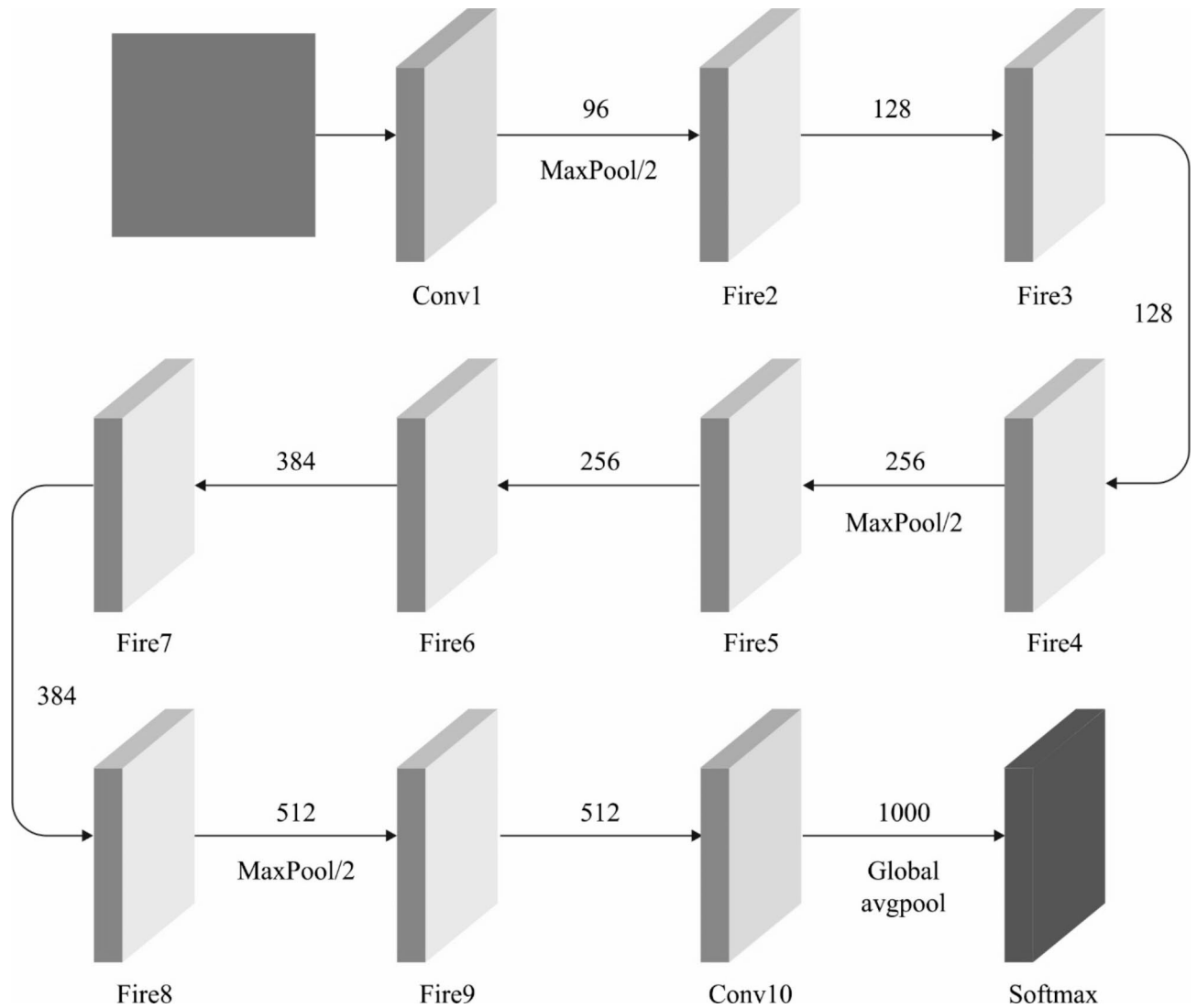


Fig. 3. Workflow of the MobileNet technique.



**Fig. 4.** Workflow of the SqueezeNet method.

the secondary stage contains bringing the fish to a safe place to eat (local exploitation)<sup>41</sup>. This method was selected due to its robust exploration and exploitation capabilities, inspired by the natural hunting behaviour of osprey birds. OOA effectually balances global search and local refinement, confirming that it averts premature convergence while effectually narrowing down the optimal hyperparameter space. This makes it specifically appropriate for intrinsic DL methods that require fine-tuning to attain high accuracy. Additionally, the capability of the OOA model to handle large search spaces and multidimensional optimization problems enables it to work well with deep neural networks, where hyperparameter interactions are often complex. OOA gives a more adaptive and computationally efficient solution than other optimization techniques, such as grid or random search, resulting in enhanced performance and faster convergence. Its flexibility also allows for easy adaptation to different architectures, making it a versatile tool in hyperparameter optimization. Figure 6 represents the OOA structure.

#### 1) Population initialization.

The OOA is stimulated by hunting osprey behaviour by employing search and predation approaches to determine optimum performance for engineering problems. During this OOA, every osprey signifies the probable solution with its location under the searching space equivalent to the variable rates of the problem. Every osprey is defined by a vector, with all the elements equivalent to a problem variable rate. This method examines the complete solution space to acquire the optimum performance. The osprey population is a mathematical model demonstrated as a matrix (Eq. (1)). Primarily, the osprey positions are randomly initialized utilizing Eq. (2) for distributing them throughout the search space, so improving the search range.

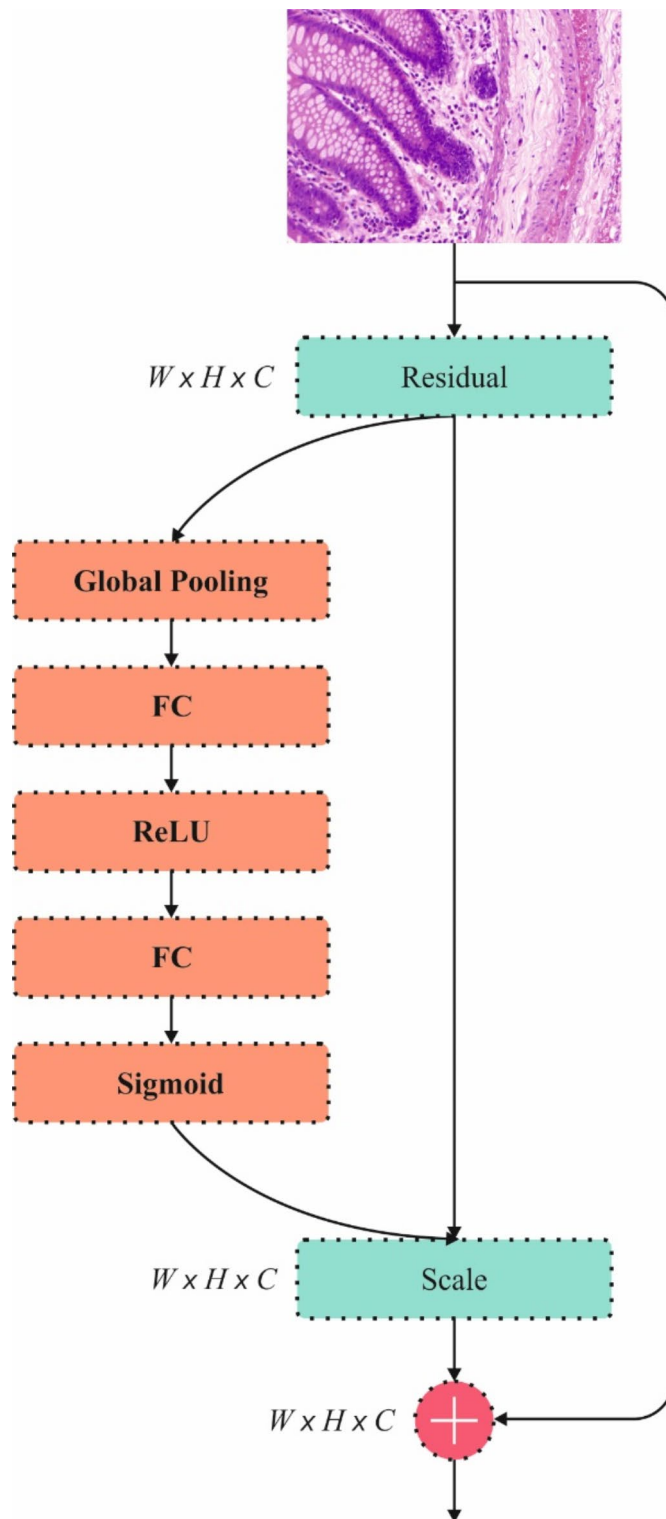


Fig. 5. Architecture of SE-ResNet Model.

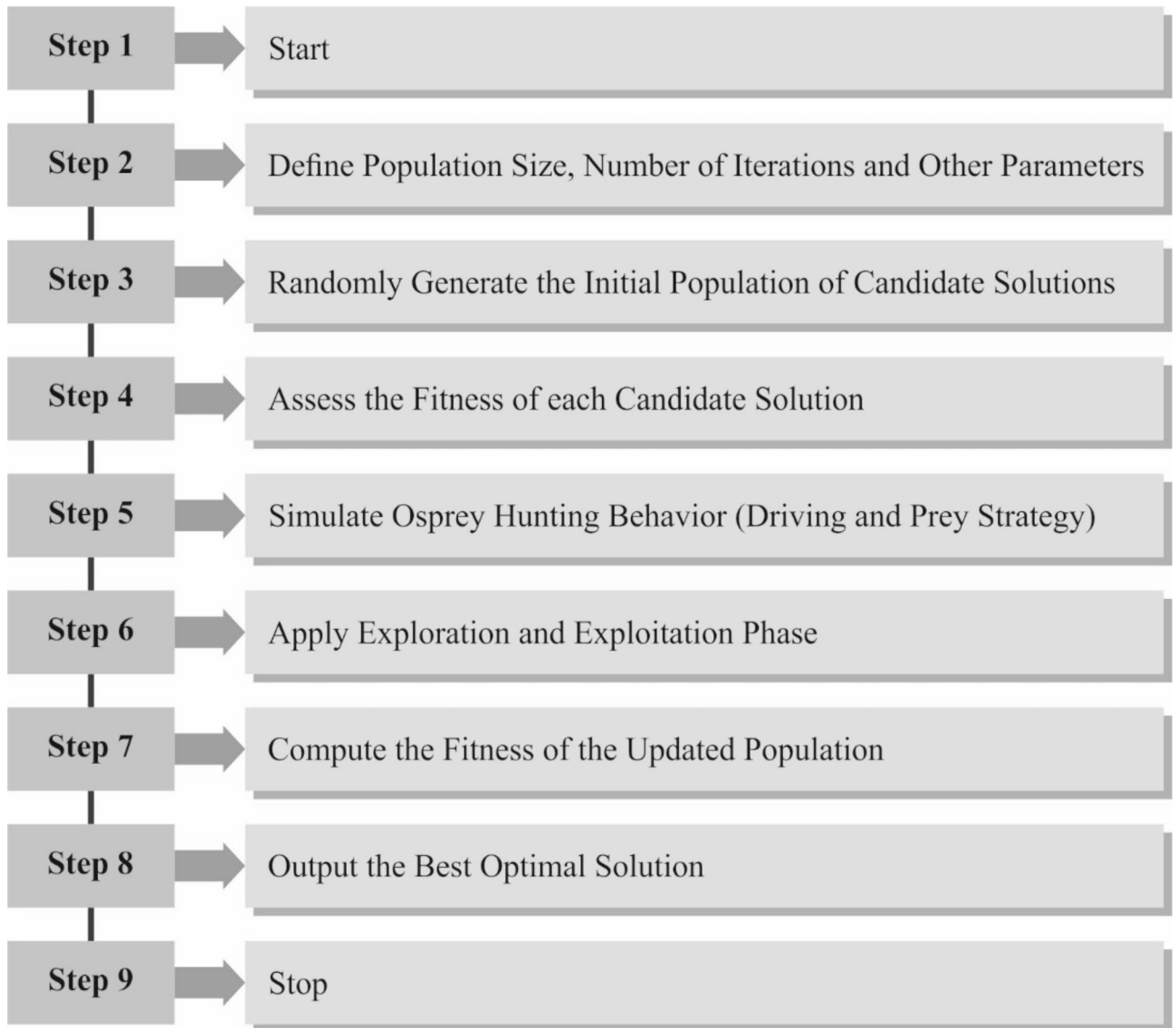


Fig. 6. Steps involved in the OOA methodology.

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \\ \vdots \\ X_N \end{bmatrix}_{N \times 1} = \begin{bmatrix} x_{1,1} & \dots & X_{1,m} & \dots & X_{1,M} \\ \vdots & \ddots & \vdots & \bullet & \vdots \\ x_{n,1} & \dots & x_{n,m} & \dots & x_{n,M} \\ \vdots & \bullet & \vdots & \ddots & \vdots \\ x_{N,1} & \dots & x_{N,m} & \dots & x_{N,M} \end{bmatrix} \tag{1}$$

$$x_{n,m} = lb_m + r_{n,m} \cdot (ub_m - lb) \tag{2}$$

Whereas  $n = 1, 2, \dots, N, m = 1, 2, \dots, M$ ,  $X$  refers to the population matrix of osprey places,  $X_n$  stands for the  $n^{th}$  osprey (candidate performances),  $x_{n,m}$  denotes the  $m^{th}$  dimensional (problem variable),  $N_o$  signifies the osprey counts,  $M$  refers to the size of the variable,  $r_{n,m} \in [0,1]$  represents the random number,  $lb_m$ , and  $ub_m$  stands for the lower and upper boundaries of  $m^{th}$  problem variable, correspondingly.

The fitness value of all the osprey's items has been computed depending on the matching objective function (OF) values. Equation (3) defines the fitness values of every osprey, which are utilized to assess the quality of all the solutions.

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_n \\ \vdots \\ F_N \end{bmatrix}_{N \times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_n) \\ \vdots \\ F(X_N) \end{bmatrix}_{N \times 1} \quad (3)$$

In which  $F$  defines the vector OF values, and  $F_n$  represents the vector of the OF values for the  $n^{th}$  osprey.

### 2) Global exploration.

Once the fish were caught, the osprey launched an attack to catch them. During this OOA, hunting is demonstrated in the primary phase of population renewal. By simulating, the osprey positions in the population are changed dramatically, improving the exploratory capability of the method for identifying the optimum area and escape from the local optimum. During this OOA design, the set of positions for all the ospreys is expressed in Eq. (4).

$$FP_n = \{X_v | v \in \{1, 2, \dots, N\} \wedge F_k < F_n\} \cup \{X_{best}\} \quad (4)$$

$FP_n$  refers to the set of positions, but fish is locked for  $n^{th}$  osprey, and  $X_{best}$  represents the osprey with the best position. The osprey randomly finds the fish's position and initiates an attack. The effort of the osprey nearby fish is simulated and the osprey position is measured using Eqs. (5) and (6), correspondingly. Once the novel position enhances the OF value, the osprey position is upgraded based on Eq. (7).

$$x_{n,m}^{P1} = x_{n,m} + r_{n,m} \cdot (SF_{n,m} - I_{n,m} \cdot x_{n,m}) \quad (5)$$

$$x_{n,m}^{P1} = \begin{cases} x_{n,m}^{P1}, & lb_m \leq x_{n,m}^{P1} \leq ub_m \\ lb_m, & x_{n,m}^{P1} < lb_m \\ ub_m, & x_{n,m}^{P1} > ub_m \end{cases} \quad (6)$$

$$X_n = \begin{cases} x_n^{P1}, & F_n^{P1} < F_n \\ X_n, & \text{else} \end{cases} \quad (7)$$

Whereas  $x_n^{P1}$  denotes the novel place of  $n^{th}$  osprey at the initial stage,  $x_{n,m}^{P1}$  implies the novel place of  $n^{th}$  osprey at the initial stage, in its  $m^{th}$  dimensional,  $F_n^{P1}$  stands for the OF in its  $m^{th}$  dimensional,  $SF_n$  represents the fish chosen by  $n^{th}$  osprey,  $SF_{n,m}$  is its  $m^{th}$  dimensional,  $r_{n,m} \in [0,1]$  denotes the random number, and  $I_{n,m} \in \{1,2\}$  refers to the random number.

### 3) Localized exploitation.

During the OOA design, a novel random location was calculated for all the individuals from the population utilizing Eqs. (8) and (9), which signifies the position of an appropriate predatory fish. Once the OF value is enhanced at this novel place, the preceding location of Osprey has been upgraded based on Eq. (10).

$$x_{n,m}^{P2} = x_{n,m} + \frac{lb_m + r \cdot (ub_m - lb_m)}{I_k} \quad (8)$$

In this case,  $I_k$  defines the iteration counter of the algorithm, and  $I_k = 1, 2, \dots, T$  and  $T$  demonstrate the entire iteration count.

$$x_{n,m}^{P2} = \begin{cases} x_{n,m}^{P2}, & lb_m \leq x_{n,m}^{P2} \leq ub_m \\ lb_m, & x_{n,m}^{P2} < lb_m \\ ub_m, & x_{n,m}^{P2} > ub_m \end{cases} \quad (9)$$

$$X_n = \begin{cases} x_n^{P2}, & F_n^{P2} < F_n \\ X_n, & \text{else} \end{cases} \quad (10)$$

Whereas  $x_n^{P2}$  illustrates the novel place of  $n^{th}$  osprey,  $x_{n,m}^{P2}$  represents the novel position of  $n^{th}$  osprey from the  $m^{th}$  dimensional,  $F_n^{P2}$  signifies the OF,  $SF_n$  denotes the fish chosen by the  $n^{th}$  osprey,  $SF_{n,m}$  defines its  $m^{th}$  dimensional, and  $r_{n,m} \in [0,1]$  is a random number.

The OOA derives an FF to gain superior results of the classifier. It defines a positive integer to describe the enhanced efficiency of the solution candidate. Here, the reduction of the classifier error rate is assumed as the FF.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{No. of misclassified samples}{Total No. of samples} \times 100 \quad (11)$$

## CRC classification using DBN

Lastly, the classification of CRC is implemented by using the DBN model<sup>42</sup>. This technique is chosen due to its capability to learn hierarchical representations of data through diverse layers of abstraction. DBNs are a DL method that outperforms extracting complex patterns from massive datasets, which is specifically useful for medical image analysis like CRC detection. Using unsupervised pretraining and fine-tuning, DBNs can effectually handle high-dimensional data and improve generalization, making them ideal for tasks with limited labelled data. Compared to conventional ML models, DBNs have shown superior performance in capturing

nonlinear relationships and detecting subtle features critical in CRC detection. Furthermore, DBNs can be trained to automatically learn crucial features from raw data, reducing the requirement for manual feature engineering and allowing for a more robust and scalable model. Their capacity to model intrinsic decision boundaries makes DBNs highly effectual for classifying CRC cases, giving more accurate and reliable predictions. Figure 7 portrays the DBN architecture.

RBM is a special kind of Markov Random Filed (MRF). And is a two-layer stochastic network. Its two layers are the visible layer (VL) and the hidden layer (HL). The neurons in the VLs and HLs are fully connected and bidirectional. In a binary RBM, based on the VL parameters  $v = (v_1, v_2, \dots, v_n)$  and HL parameters  $h = (h_1, h_2, \dots, h_m)$ , the joint distribution of the RBM's VL and HL is represented as

$$p(v, h) = \frac{1}{Z} \exp(-E(v, h)) \quad (12)$$

Where

$$z = \sum_{v, h} \exp(-E(v, h))$$

the normalization factor is called the partition function, and  $E(v, h)$  is the energy function.

The unit's connection weight and bias define the probability distribution over the binary state vector  $v$  of the VL unit through the energy function.

$$E(v, h; \theta) = - \sum_{i=1}^V \sum_{j=1}^H w_{ij} v_i h_j - \sum_{i=1}^V b_i v_i - \sum_{j=1}^H a_j h_j, \quad (13)$$

Where  $\theta = (w, b, a)$  and  $w_{ij}$  signifies the symmetric interaction between the  $i^{th}$  and  $j^{th}$  VL and HL units, and  $b_i$  and  $a_j$  represent their biased terms.  $V$  and  $H$  are the number of VL and HL units. The model assigns a probability to the VL vector  $v$  is

$$p(v; \theta) = \frac{\sum_h \exp(-E(v, h))}{\sum_u \sum_h \exp(-E(u, h))}. \quad (14)$$

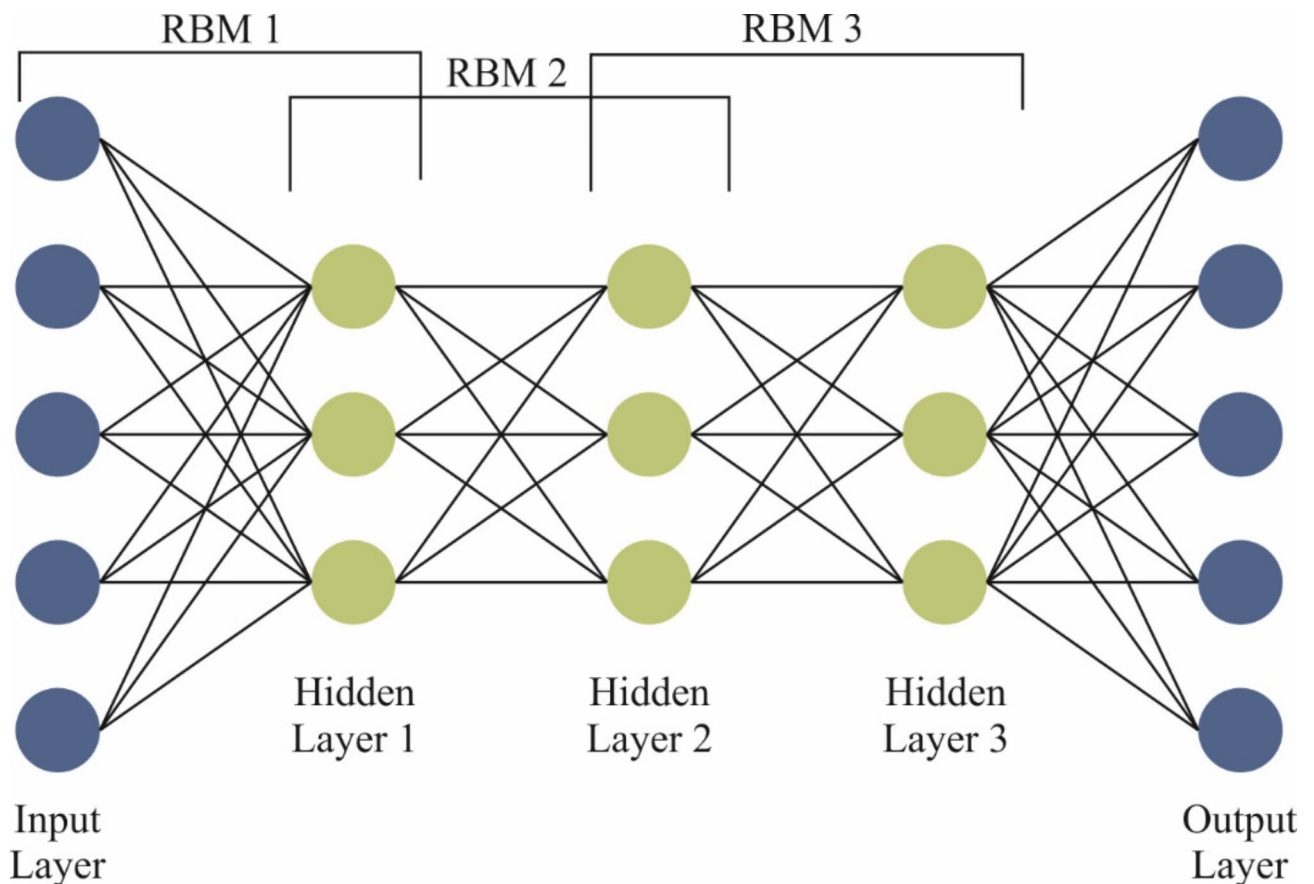


Fig. 7. DBN architecture.

The factorial  $p(v|h)$  and  $p(h|v)$  are conditional distributions because there are no visible-visible or hidden-hidden links, as follows:

$$p(h_j = 1|v; \theta) = \sigma \left( \sum_{i=1}^V w_{ij}v_i + a_j \right),$$

$$p(v_j = 1|h; \theta) = \sigma \sum_{j=1}^H w_{ij}h_j + b_j, \tag{15}$$

where  $\sigma(x) = (1 + \exp(-x))^{-1}$ . The RBM is trained to model the joint distribution of data and class labels; the VL vectors are connected to the binary vector.

$$E(v, 1, h; \theta) = - \sum_{i=1}^V \sum_{j=1}^H w_{ij}h_jv_i - \sum_{y=1}^L \sum_{j=1}^H w_{yj}h_jl_y$$

$$- \sum_{j=1}^L a_jh_j - \sum_{y=1}^L c_yl_y - \sum_{i=1}^V b_iv_i, \tag{16}$$

$$p(l_y = 1|h; \theta) = \text{softmax} \left( \sum_{j=1}^H w_{yj}h_j + c_y \right). \tag{17}$$

Furthermore,  $p(1|v)$  is computed exactly by

$$p(1|v) = \frac{\sum_h e^{-E(v,1,h)}}{\sum_1 e^{-E(v,1,h)}}. \tag{18}$$

The value of  $p(1|v)$  is efficiently calculated using the conditional independence of the HL units, which makes the marginalization time of the HL units linear in the number of HL units.

RBM has two training algorithms: gradient method and Contrastive Divergence (CD). Where the gradient method treats  $\log p(v, \theta)$  as a likelihood function, which does not change monotonicity. The parameters are corrected along the gradient  $\partial \log p(v, \theta) / \partial \theta$  to achieve higher learning efficiency. The details are as follows:

$$\theta(n+1) = \theta(n) + \alpha^* \left( - \frac{\partial \log p(v, \theta)}{\partial \theta} \right), \theta \in \{w, a, b\},$$

$$- \frac{\partial \log p(v, w_{ij})}{\partial w_{ij}} = E_v [p(h_i|v) * v_j] - v_j^{(i)} * f(w_i * v^{(i)} + b_i),$$

$$- \frac{\partial \log p(v, b_i)}{\partial b_i} = E_v [p(h_i|v) * v_j] - f(w_i * v^{(i)}),$$

$$- \frac{\partial \log p(v, a_j)}{\partial a_j} = E_v [p(h_i|v) * v_j] - v_j^{(i)}. \tag{19}$$

The update rule for the visible-hidden weights is based on the gradient of the joint probability function of the data and labels.

$$\Delta w_{ij} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model} \tag{20}$$

The expectation  $\langle v_i h_j \rangle_{data}$  is the frequency of simultaneous occurrences of the VL unit  $v_i$  and the HL unit  $h_j$  in the training set. To accurately compute the  $\langle \cdot \rangle_{model}$ , it takes an exponential amount of time, so the CD approximation of the gradient is used. As shown in the following equation:

$$\Delta w_{ij} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_1, \tag{21}$$

Where  $\langle \cdot \rangle_1$  denotes the expected value of the distribution of samples for running the Gibbs sampler, which is initialized with data and run for a whole step.

The energy of joint configuration for Gaussian-Bernoulli RBM is

$$E(v, h; \theta) = \sum_{i=1}^{\Sigma} \frac{(v_i - b_i)^2}{2} - \sum_{i=1}^V \sum_{j=1}^H w_{ij}v_ih_j - \sum_{j=1}^H a_jh_j. \tag{22}$$

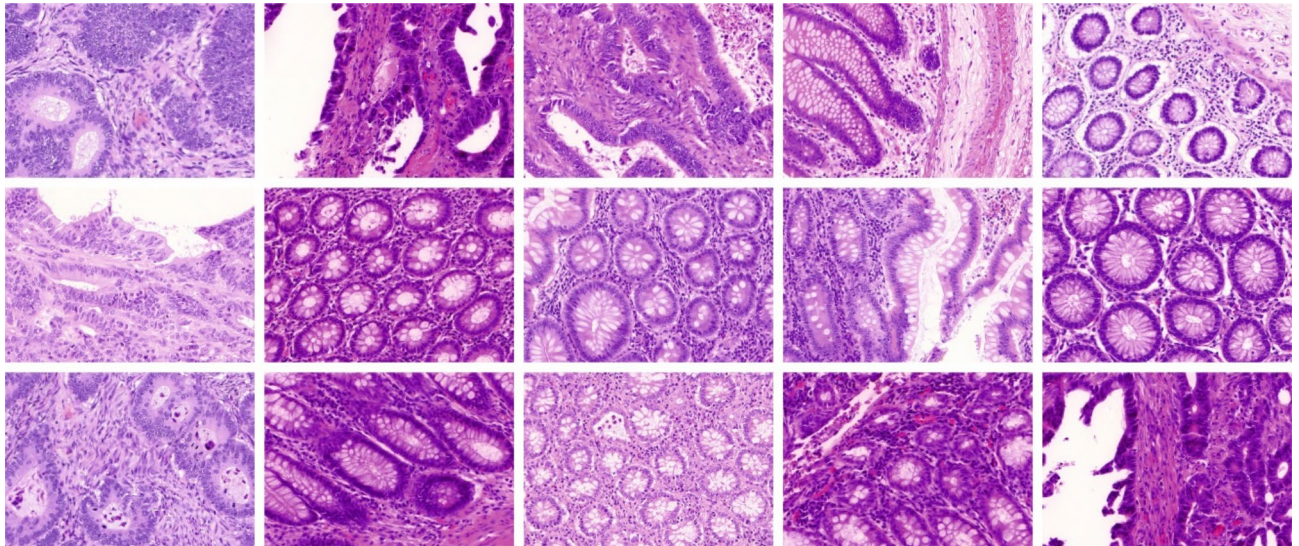
The conditional distribution  $p(v; h, \theta)$  is a factorial distribution value because there is no visible-visible connection,

$$p(v_i; h, \theta) = N \left( b_i + \sum_{j=1}^H w_{ij}h_j, 1 \right) \tag{23}$$

Here,  $N(\mu, V)$  refers to a Gaussian with mean  $\mu$  and variance  $V$ . Besides, the Gaussian-Bernoulli RBM has the exact inference and learning rules as the binary RBM, except that the learning rate needs to be smaller.

Class	No. of Images
Benign	74
Malignant	91
Total Images	165

**Table 1.** Details of the dataset.



**Fig. 8.** Sample images.

A single hidden layer RBM needs to capture data features more precisely. After training the RBM, the learned features are fed into a second RBM as an input dataset. This layer-by-layer learning system is used to build DBN. DBN is a DNN with multiple RBMs and a BPNN. As with other DNNs, the key of DBN is to initialize an FFNN with unsupervised pretraining using an unlabeled dataset, then fine-tune the FFNN using labelled data. The initial RBM is trained using the CD method during the pretraining phase. The learning states of the HL units in the first RBM are used as input data for the VL units of the second RBM. The weights of all RBMs are trained layer-by-layer in the same way until the last RBM. The features of the previous RBM are those learned by the whole training system. The FFNN will use the highest RBM weights as its initial weights when the unsupervised pretraining of the RBM is finished. Using a back-propagation algorithm, the FFNN is subsequently trained or fine-tuned on the labelled training data.

### Performance validation

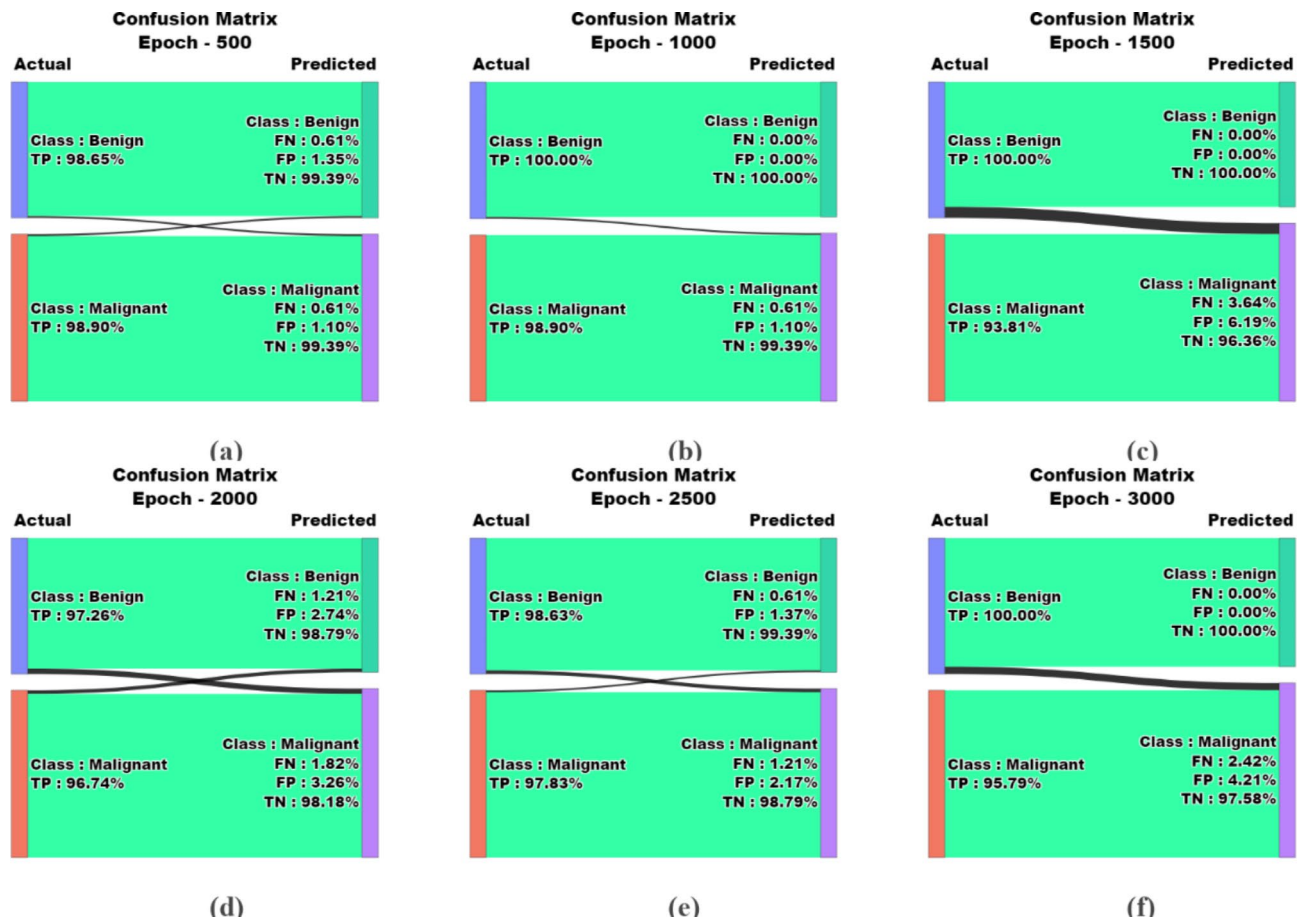
The simulation analysis of the CCD-ODFFBI method is examined using the Warwick-QU dataset<sup>43</sup>. The dataset contains 165 sample images with two classes illustrated in Table 1. Figure 8 exhibits the sample images.

Figure 9 shows the confusion matrices produced by the CCD-ODFFBI method in different epochs. The outcomes show that the CCD-ODFFBI method effectively detects benign and malignant samples under different classes.

Table 2 illustrates the overall classification outcomes of the CCD-ODFFBI technique under various epochs. The outcomes imply that the CCD-ODFFBI method has properly detected the benign and malignant samples. With 500 epochs, the CCD-ODFFBI technique gains an average  $accu_y$  of 98.79%,  $prec_n$  of 98.77%,  $sens_y$  of 98.77%,  $spec_y$  of 98.77%, and  $F_{score}$  of 98.77%. In addition, with 1000 epochs, the CCD-ODFFBI technique gains an average  $accu_y$  of 99.39%,  $prec_n$  of 99.46%,  $sens_y$  of 99.32%,  $spec_y$  of 99.32%, and  $F_{score}$  of 99.39%. Moreover, with 1500 epochs, the CCD-ODFFBI method obtains an average  $accu_y$  of 96.36%,  $prec_n$  of 96.91%,  $sens_y$  of 95.95%,  $spec_y$  of 95.95%, and  $F_{score}$  of 96.29%.

Figure 10 shows the CRC detection outcomes of the CCD-ODFFBI method under epochs 2000–3000. The result indicates that the CCD-ODFFBI approach has accurately detected the benign and malignant samples. With 2000 epochs, the CCD-ODFFBI approach obtains an average  $accu_y$  of 96.97%,  $prec_n$  of 97.00%,  $sens_y$  of 96.87%,  $spec_y$  of 96.87%, and  $F_{score}$  of 96.93%. Moreover, with 2500 epochs, the CCD-ODFFBI method obtains an average  $accu_y$  of 98.18%,  $prec_n$  of 98.23%,  $sens_y$  of 98.10%,  $spec_y$  of 98.10%, and  $F_{score}$  of 98.16%. Furthermore, with 3000 epochs, the CCD-ODFFBI method obtains an average  $accu_y$  of 97.30%,  $prec_n$  of 97.89%,  $sens_y$  of 97.30%,  $spec_y$  of 97.30%, and  $F_{score}$  of 97.54%.

In Fig. 11, the training and validation accuracy outcomes of the CCD-ODFFBI technique are described. The accuracy value is calculated within the range of 0–1000 epochs. The figure shows that the training and validation accuracy value shows a growing tendency, which indicates the capability of the CCD-ODFFBI approach with



**Fig. 9.** Confusion matrices of CCD-ODFFBI technique (a-f) Epochs 500–3000.

enriched performance over dissimilar iterations. Furthermore, the training and validation accuracy remain closer over the epochs, which exhibits enhanced performance and indicates minimal overfitting of the CCD-ODFFBI approach, which guarantees consistent prediction on hidden samples.

In Fig. 12, the training and validation loss graph of the CCD-ODFFBI method is demonstrated. The loss value is calculated within the range of 0–1000 epochs. It is signified that the training and validation accuracy values demonstrated a declining tendency, which notified the capability of the CCD-ODFFBI method in balancing a tradeoff between data fitting and generalization. The continual reduction in loss values additionally ensures the superior performance of the CCD-ODFFBI method and tunes the predictive outcomes over time.

In Fig. 13, the PR inspection of the CCD-ODFFBI method under 1000 epochs offers an interpretation of its performance by plotting Precision against Recall for different classes. The figure indicates that the CCD-ODFFBI method continuously obtains superior PR values across various classes, representing its capability to maintain a considerable portion of true positive predictions amongst all the positive predictions (precision) while capturing a large proportion of actual positives (recall). The continuous rise in PR outcomes among all classes depicts the efficiency of the CCD-ODFFBI method in the classifier model.

Figure 14 shows the ROC curve of the CCD-ODFFBI method under 1000 epochs. The outcomes indicate that the CCD-ODFFBI approach obtains superior ROC outcomes over all the classes, representing the substantial ability to discriminate them. This consistent trend of high ROC values over different classes represents the promising solution of the CCD-ODFFBI method on prediction class, which highlights the robust nature of the classification model.

Table 3; Fig. 15 illustrates a widespread comparison study of the CCD-ODFFBI technique under distinct aspects<sup>44–46</sup>. The ResNet-50 model with a 60–40 data split has lower  $accu_y$  at 78.92% but high  $spec_y$  at 93.99%. The ResNet-50 model with an 80–20 split shows improved  $accu_y$  at 89.89% and  $sens_y$  at 94.66%. VGG16, AlexNet, and Inception-v3 also portray robust performance with  $accu_y$  values ranging from 96.84 to 98.06%. The MDCC-Net and SMADTL-CCDC models achieve improved results with  $accu_y$  above 99%. The CCD-ODFFBI model attains the highest performance, with an  $accu_y$  of 99.39%,  $sens_y$  of 99.32%, and  $spec_y$  of 99.32%.

Figure 16 shows the  $sens_y$  analysis of the CCD-ODFFBI approach with existing techniques. The outcomes show that the ResNet-50 (60–40) technique has shown ineffective performance with  $sens_y$  of 61.16%. Meanwhile, the DL-CP, DL-SC, ResNet-50 (80–20), and SMADTL-CCDC approaches have demonstrated moderately closer outcomes with  $sens_y$  of 70.47%, 84.61%, 94.66%, and 98.18%. Moreover, the VGG16,

Class	Accu <sub>y</sub>	Prec <sub>n</sub>	Sens <sub>y</sub>	Spec <sub>y</sub>	F <sub>Score</sub>
Epoch – 500					
Benign	98.79	98.65	98.65	98.90	98.65
Malignant	98.79	98.90	98.90	98.65	98.90
Average	98.79	98.77	98.77	98.77	98.77
Epoch – 1000					
Benign	99.39	100.00	98.65	100.00	99.32
Malignant	99.39	98.91	100.00	98.65	99.45
Average	99.39	99.46	99.32	99.32	99.39
Epoch – 1500					
Benign	96.36	100.00	91.89	100.00	95.77
Malignant	96.36	93.81	100.00	91.89	96.81
Average	96.36	96.91	95.95	95.95	96.29
Epoch – 2000					
Benign	96.97	97.26	95.95	97.80	96.60
Malignant	96.97	96.74	97.80	95.95	97.27
Average	96.97	97.00	96.87	96.87	96.93
Epoch – 2500					
Benign	98.18	98.63	97.30	98.90	97.96
Malignant	98.18	97.83	98.90	97.30	98.36
Average	98.18	98.23	98.10	98.10	98.16
Epoch – 3000					
Benign	94.59	100.00	94.59	100.00	97.22
Malignant	100.00	95.79	100.00	94.59	97.85
Average	97.30	97.89	97.30	97.30	97.54

**Table 2.** CRC detection outcomes of CCD-ODFFBI technique under various epochs.

AlxNet, Inception-v3, and MDCC-Net models have portrayed slightly improved *sens<sub>y</sub>* values of 96.70%, 97.47%, 97.97%, and 98.55%. However, the CCD-ODFFBI method outperforms the other techniques with an increased *sens<sub>y</sub>* of 99.32%.

Figure 17 shows the *spec<sub>y</sub>* analysis of the CCD-ODFFBI approach with existing techniques. The outcomes show that the DL-CP approach has demonstrated ineffective performance with *spec<sub>y</sub>* of 71.37%. The RestNet-50 (80–20) and DL-SC methods have shown slightly enhanced outcomes with *spec<sub>y</sub>* of 84.34% and 82.11%. Meanwhile, the RestNet-50 (60–40) and SMADTL-CCDC approaches have illustrated moderately closer outcomes with *spec<sub>y</sub>* of 93.99% and 98.26%. Moreover, the VGG16, AlxNet, Inception-v3, and MDCC-Net models have portrayed slightly improved *spec<sub>y</sub>* values of 96.68%, 97.45%, 98.08%, and 98.59%. However, the CCD-ODFFBI method outperforms the other techniques with a high *spec<sub>y</sub>* of 99.32%.

Table 4; Fig. 18 demonstrate the ablation study of the proposed model. The CCD-ODFFBI model attained an *accu<sub>y</sub>* of 99.39%, *sens<sub>y</sub>* of 99.32%, and *spec<sub>y</sub>* of 99.32%. The MobileNet model had an *accu<sub>y</sub>* of 98.86%, *sens<sub>y</sub>* of 98.74%, and *spec<sub>y</sub>* of 98.70%. The SqueezeNet model showed an *accu<sub>y</sub>* of 98.20%, *sens<sub>y</sub>* of 98.07%, and *spec<sub>y</sub>* of 97.94%. The SE-ResNet model achieved an *accu<sub>y</sub>* of 97.56%, *sens<sub>y</sub>* of 97.46%, and *spec<sub>y</sub>* of 97.33%. Lastly, the OOA model demonstrated an *accu<sub>y</sub>* of 96.91%, *sens<sub>y</sub>* of 96.79%, and *spec<sub>y</sub>* of 96.80%.

## Conclusion

In this study, a novel CCD-ODFFBI technique is introduced. The CCD-ODFFBI technique aimed to examine the biomedical images for the identification of CRC. It utilized different processes such as noise removal, feature fusion, hyperparameter selection, and DBN-based CRC classification. Initially, the CCD-ODFFBI technique utilized the MF approach for the noise elimination process. Three DL models, namely MobileNet, SqueezeNet, and SE-ResNet, are employed for feature extraction. Meanwhile, the DL models' hyperparameter selection was performed using OOA. Furthermore, the classification of CRC was accomplished by utilizing the DBN model. A series of simulations highlighted the significant results of the CCD-ODFFBI method under the Warwick-QU dataset. The comparison of the CCD-ODFFBI method showed a superior accuracy value of 99.39% over existing techniques. The CCD-ODFFBI method's limitations include reliance on a single dataset, which may limit the model's generalization to diverse populations or imaging conditions. Furthermore, image quality discrepancies and artefacts could affect the model's performance. The study also lacks a comprehensive evaluation across diverse real-world scenarios, such as different stages of cancer or diverse histological types. Future work should focus on integrating larger, more varied datasets to enhance generalization. Moreover, integrating real-time diagnostic capabilities and addressing interpretability could improve the clinical application of the model.

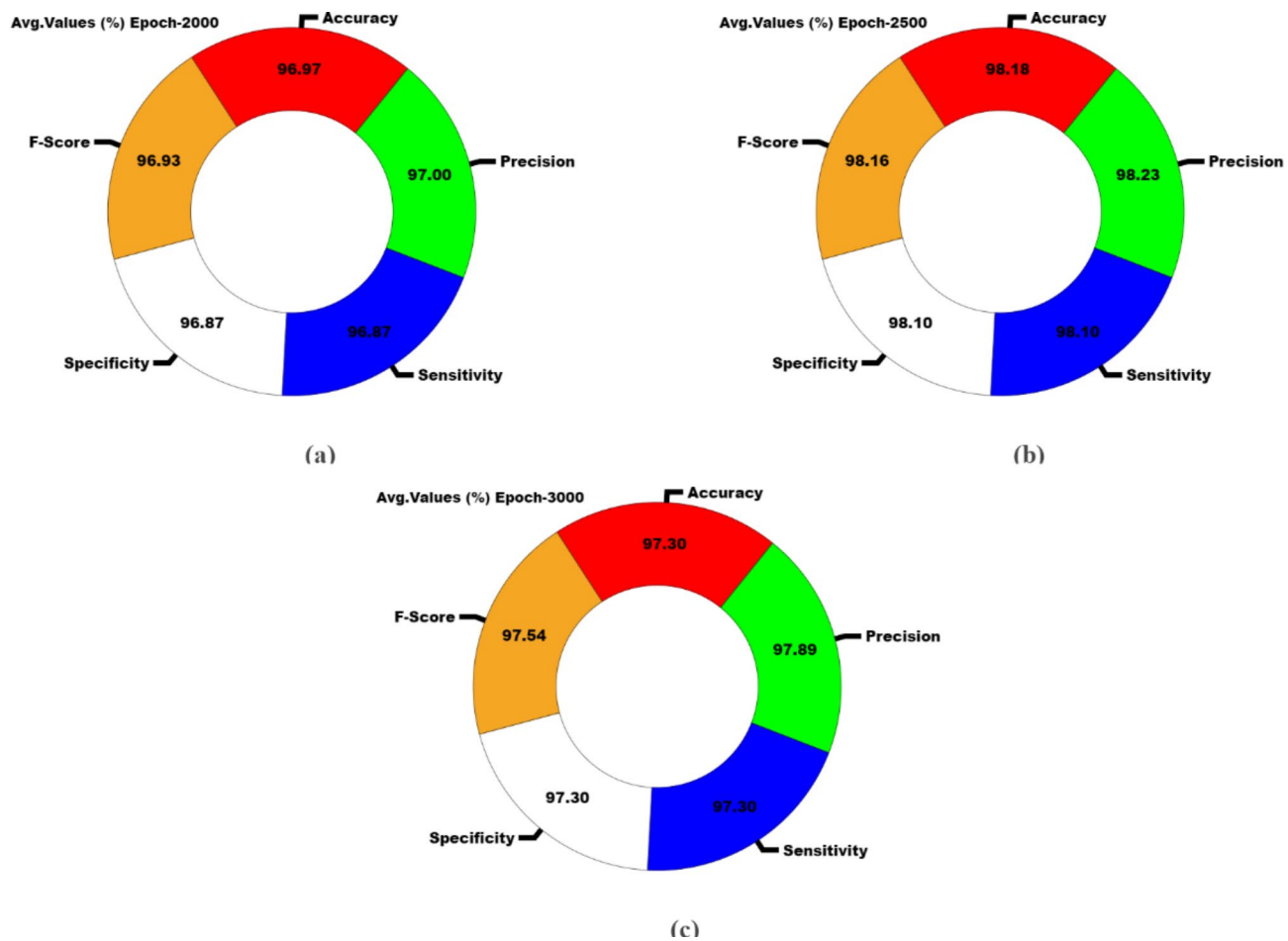
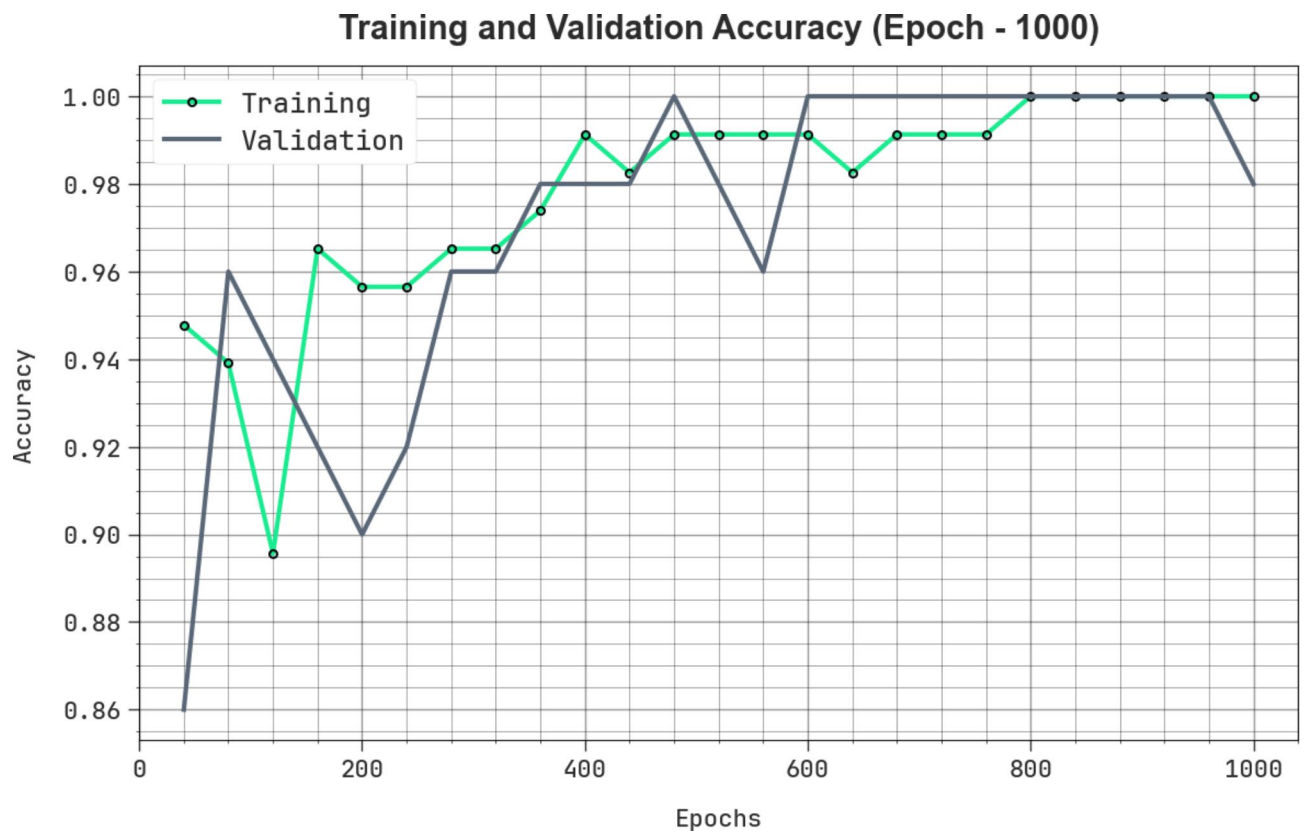


Fig. 10. Average outcome of CCD-ODFFBI technique (a-c) Epochs 2000–3000.



**Fig. 11.**  $Accu_y$  curve of CCD-ODFFBI technique under 1000 epochs

### Training and Validation Loss (Epoch - 1000)

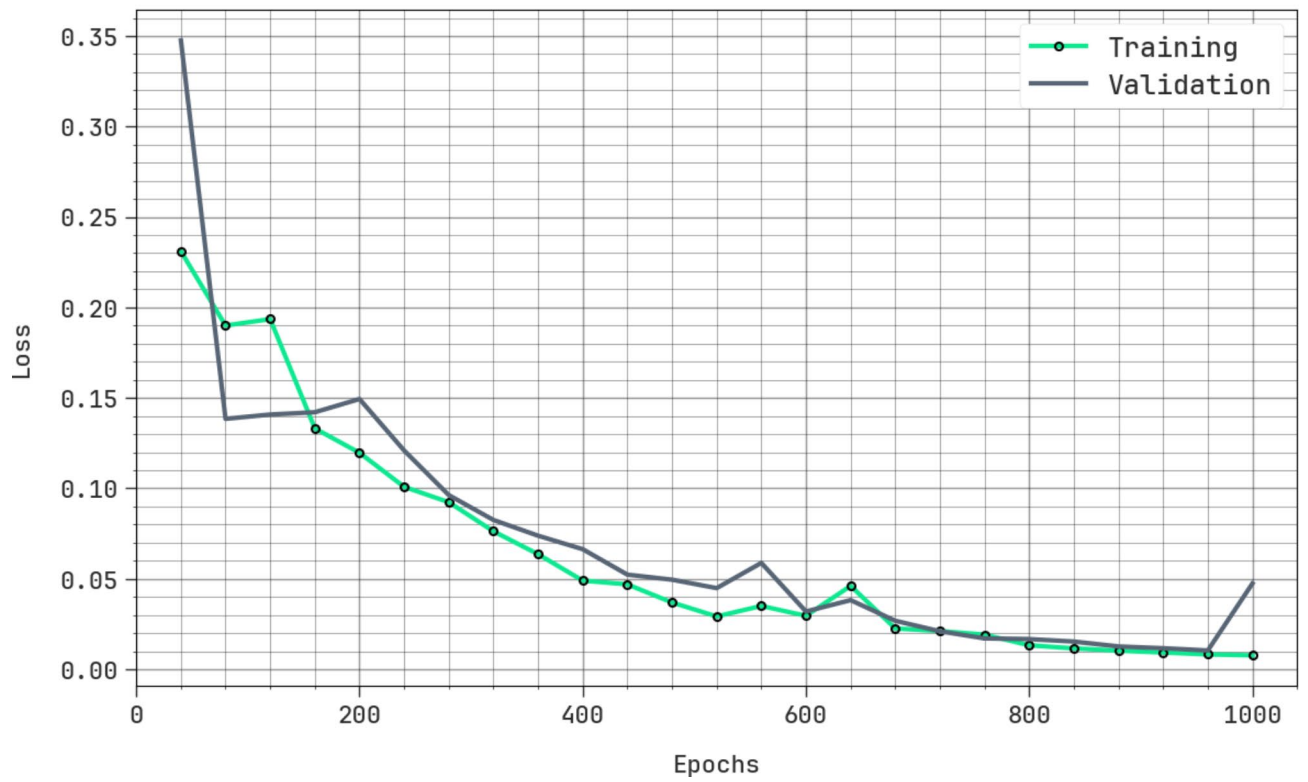


Fig. 12. Loss curve of CCD-ODFFBI technique under 1000 epochs.

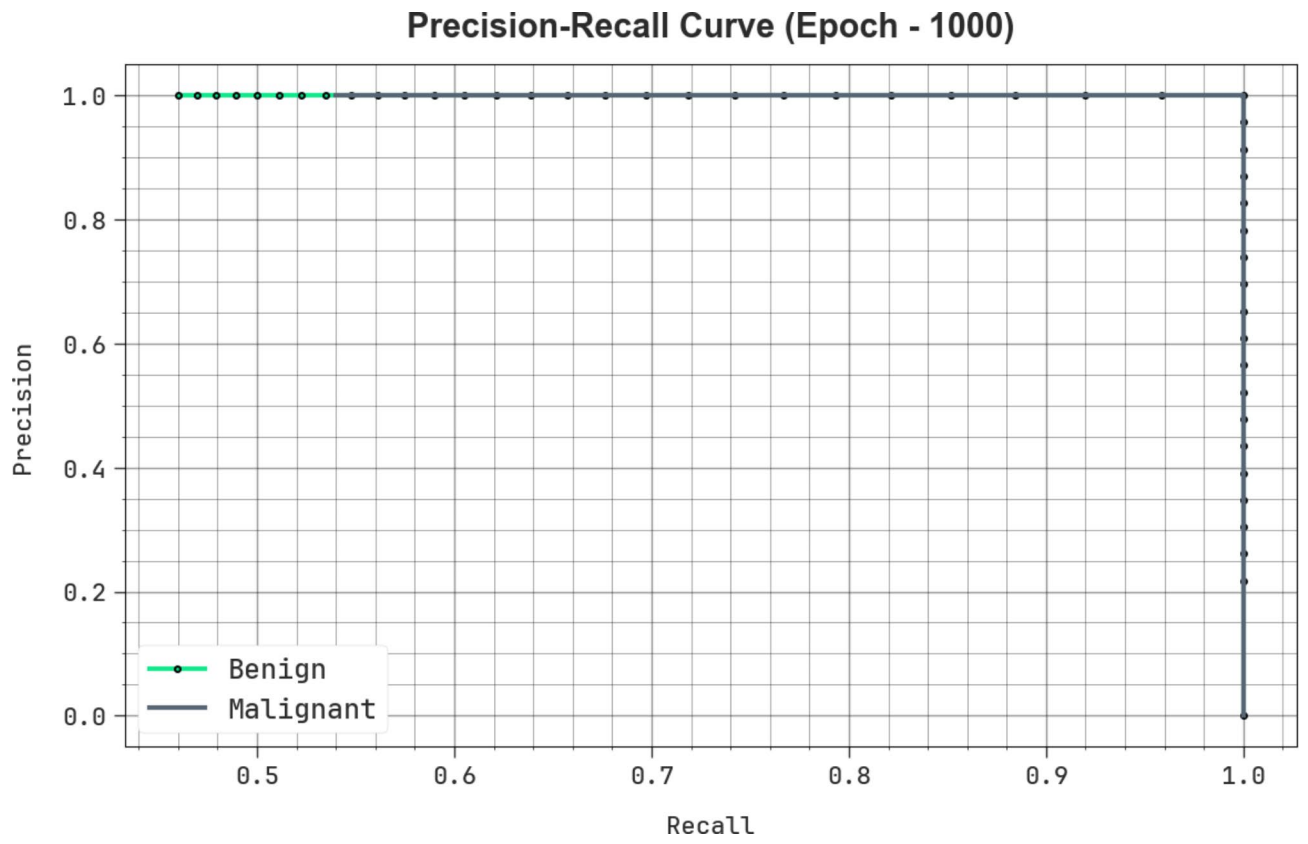
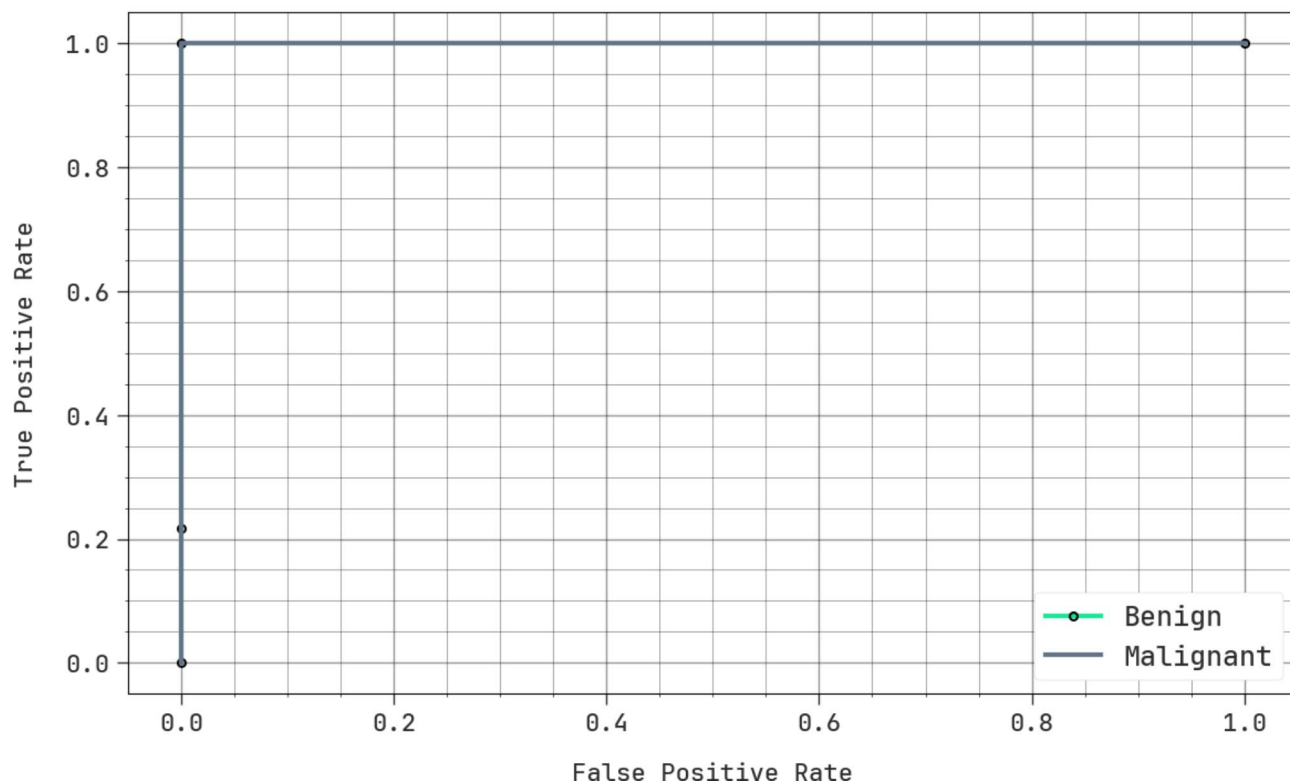


Fig. 13. PR curve of CCD-ODFFBI technique under 1000 epochs.

### ROC-Curve (Epoch - 1000)



**Fig. 14.** ROC curve of CCD-ODFFBI technique under 1000 epochs.

Methods	Accu <sub>y</sub>	Sens <sub>y</sub>	Spec <sub>y</sub>
RestNet-50 (60 – 40)	78.92	61.16	93.99
RestNet-50 (80 – 20)	89.89	94.66	84.34
DL-SC Model	82.69	84.61	82.11
DL-CP Model	71.41	70.47	71.37
SMADTL-CCDC	99.17	98.18	98.26
VGG16	96.84	96.70	96.68
AlxNet	97.52	97.47	97.45
Inception-v3	98.06	97.97	98.08
MDCC-Net	98.67	98.55	98.59
CCD-ODFFBI	99.39	99.32	99.32

**Table 3.** Comparative outcome of the CCD-ODFFBI method with existing techniques<sup>44–46</sup>.

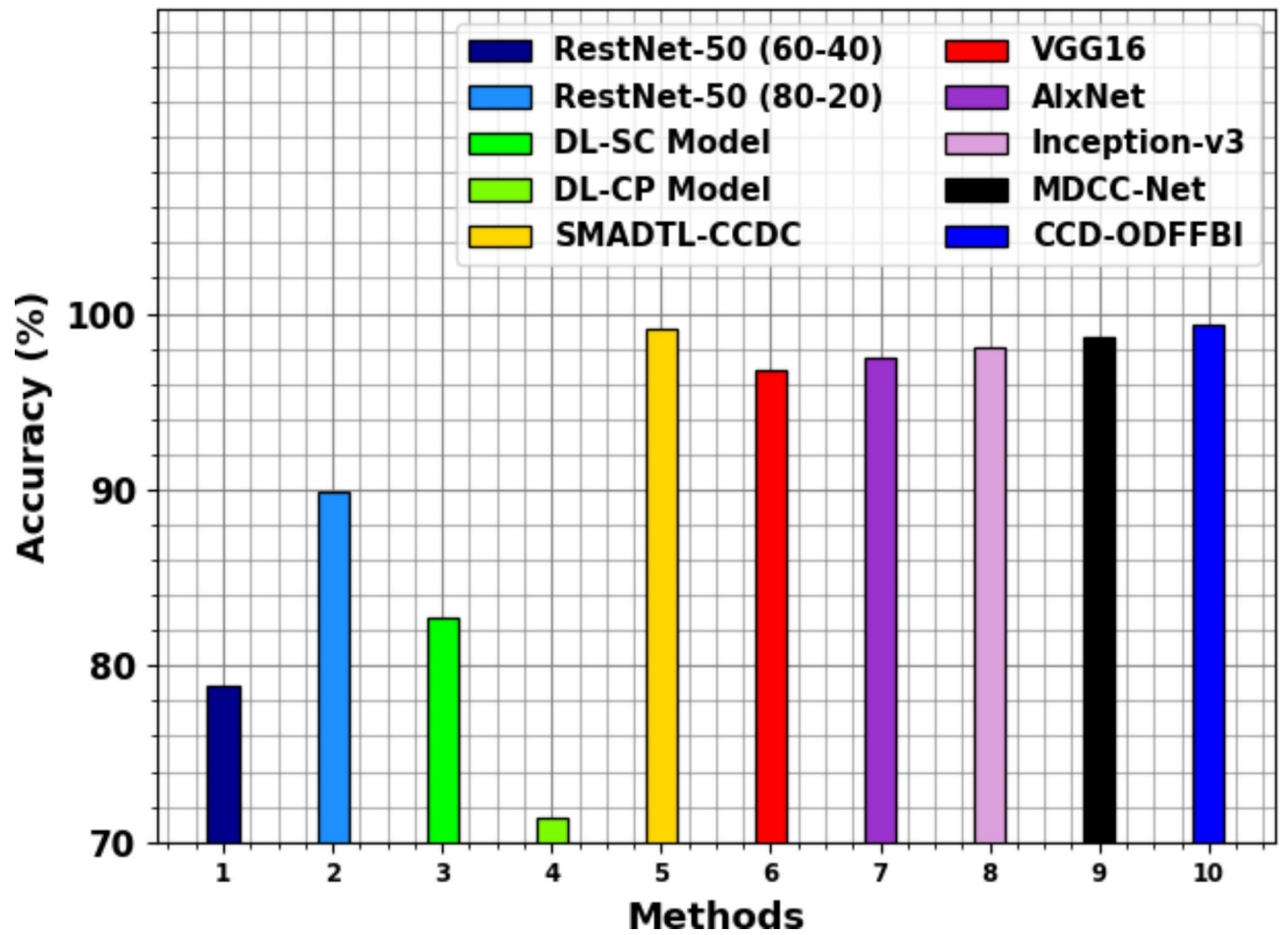


Fig. 15.  $Accu_y$  outcome of CCD-ODFFBI method with existing techniques

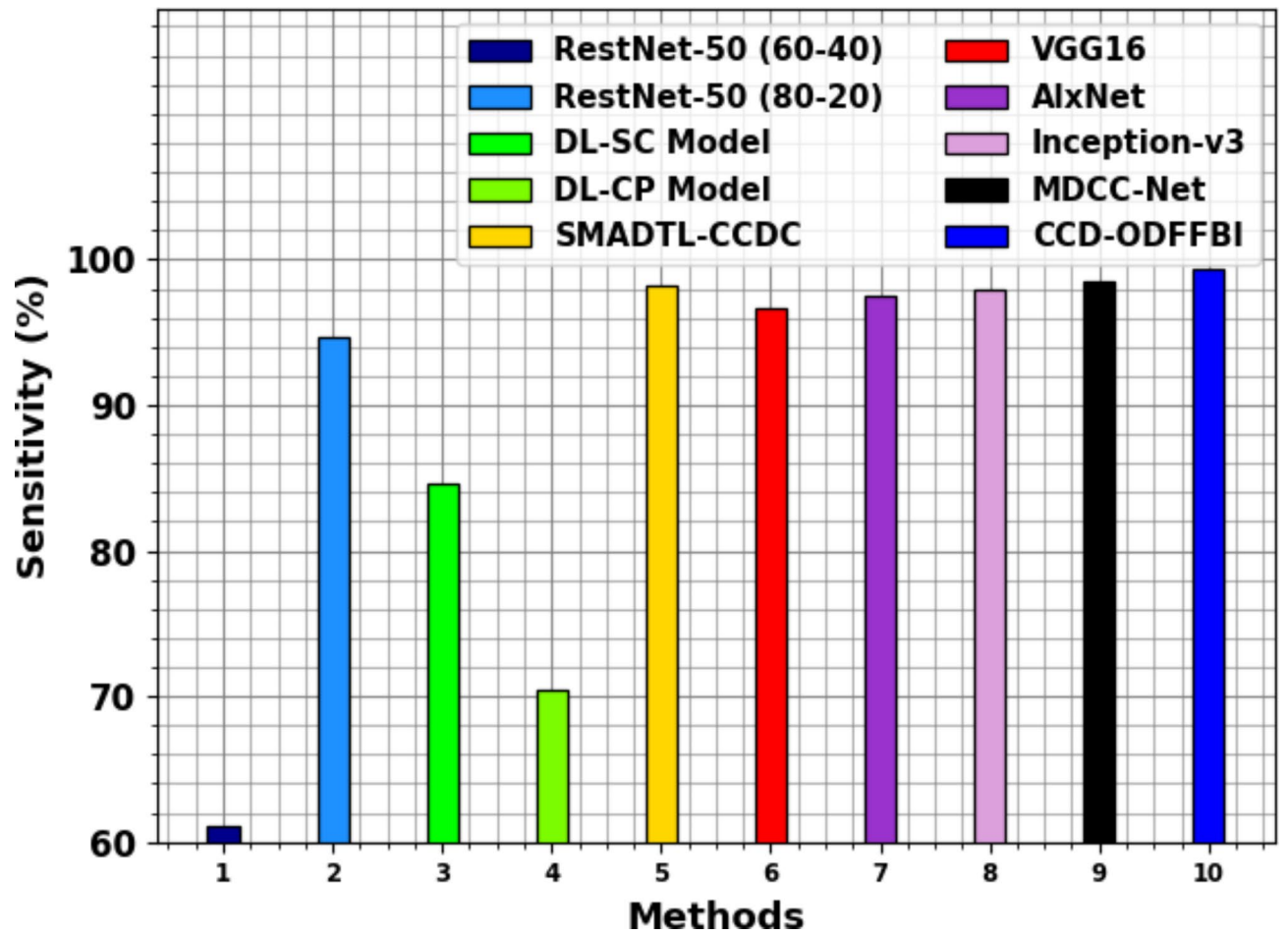


Fig. 16.  $Sens_y$  outcome of CCD-ODFFBI method with existing techniques.

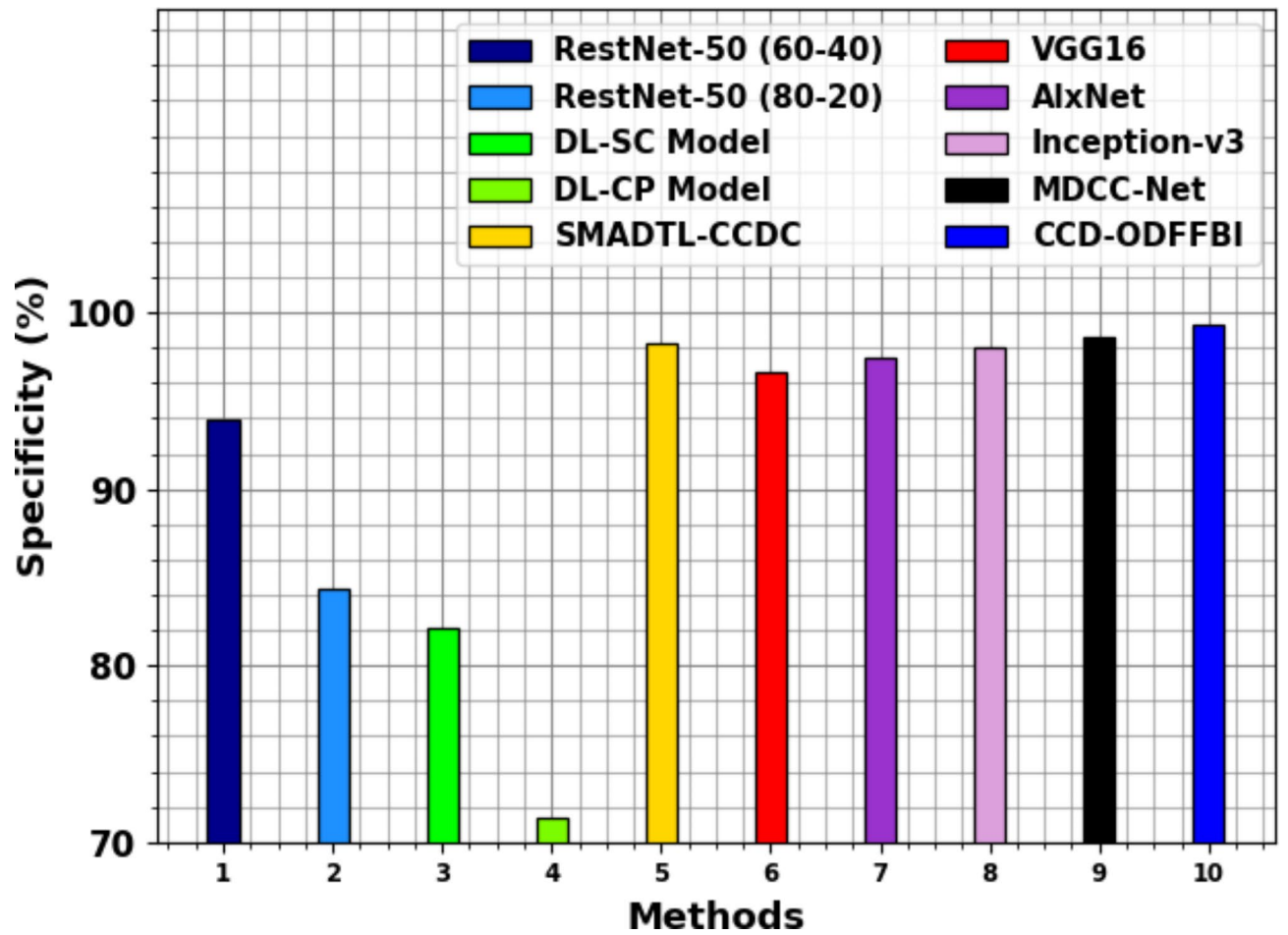


Fig. 17.  $Spec_y$  outcome of CCD-ODFFBI method with existing techniques

Methods	Accu <sub>y</sub>	Sens <sub>y</sub>	Spec <sub>y</sub>
CCD-ODFFBI	99.39	99.32	99.32
MobileNet Model	98.86	98.74	98.70
SqueezeNet Model	98.20	98.07	97.94
SE-ResNet Model	97.56	97.46	97.33
OOA	96.91	96.79	96.80

Table 4. Result analysis of the ablation study of CCD-ODFFBI method.

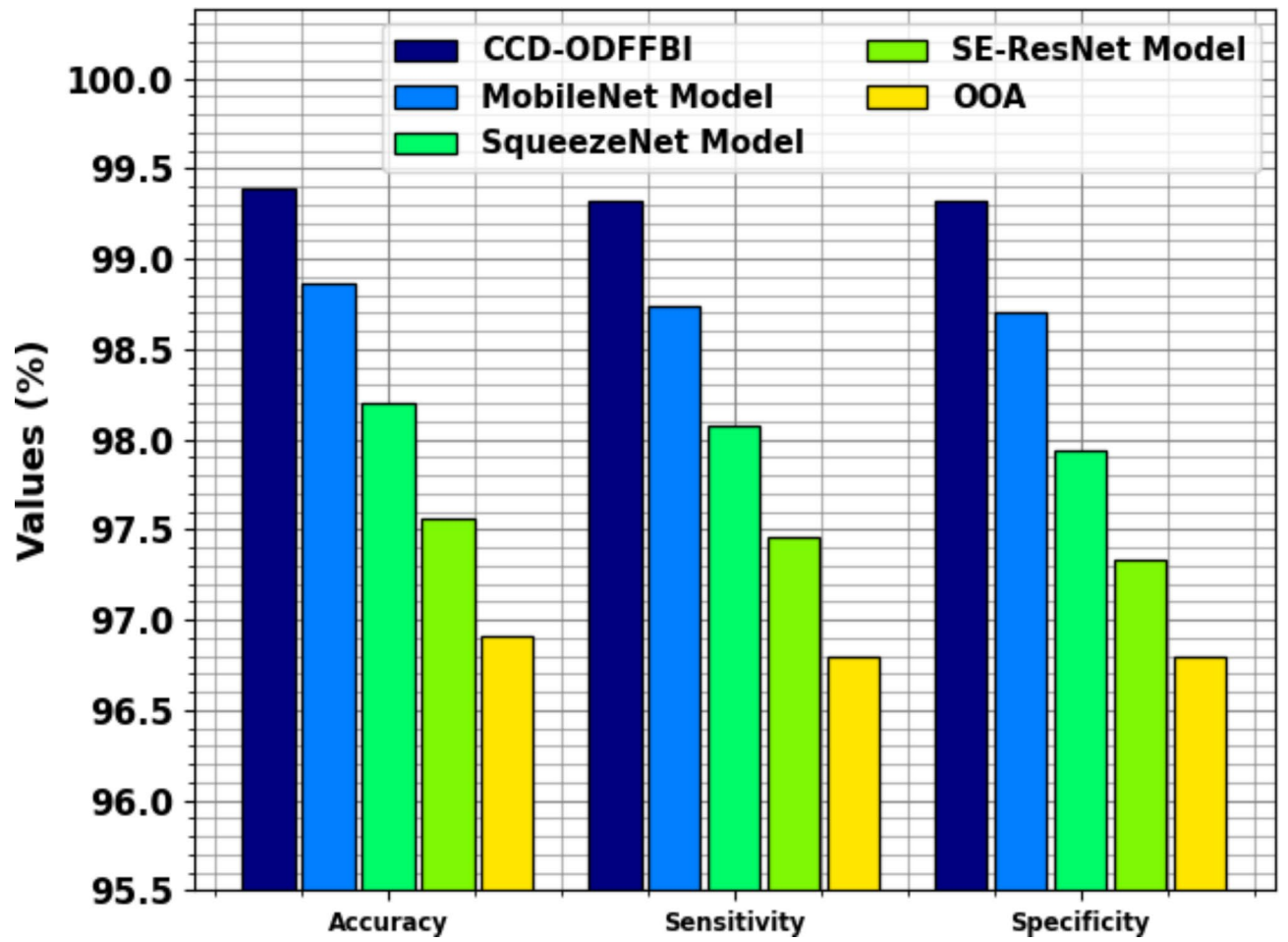


Fig. 18. Result analysis of the ablation study of CCD-ODFFBI method.

## Data availability

The datasets used and analyzed during the current study available from the corresponding author on reasonable request.

Received: 5 July 2024; Accepted: 16 December 2024

Published online: 04 February 2025

## References

- Skrede, O. J. et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* **395** (10221), 350–360 (2020).
- Ho, C. et al. A promising deep learning-assistive algorithm for histopathological screening of colorectal cancer. *Scientific reports*, *12*(1), p.2222. (2022).
- Tsai, M. J. & Tao, Y. H. Deep learning techniques for the classification of colorectal cancer tissue. *Electronics*, *10*(14), p.1662. (2021).
- Kather, J. N. et al. Multi-class texture analysis in colorectal cancer histology. *Sci. Rep.* **6** (1), 1–11 (2016).
- Dif, N. & Elberichi, Z. A new deep learning model selection method for colorectal cancer classification. *Int. J. Swarm Intell. Res. (IJSIR)*. **11** (3), 72–88 (2020).
- Sarwinda, D., Paradisa, R. H., Bustamam, A. & Anggia, P. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Comput. Sci.* **179**, 423–431 (2021).
- Mulenga, M. et al. Feature extension of gut microbiome data for deep neural network-based colorectal cancer classification. *IEEE Access*. **9**, 23565–23578 (2021).
- Bychkov, D. et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific reports*, *8*(1), p.3395. (2018).
- Rodríguez, J. V., Martínez, J. R. & Salazar, F. F. J. Colorectal Cancer prediction using machine learning and neutrosophic MCDM methodology: a Case Study. *Full Length Article*, **21**(2), (2023). pp.118–18.
- Zhang, D. et al. An efficient ECG denoising method based on empirical mode decomposition, sample entropy, and improved threshold function, *Wireless Communications and Mobile Computing*, vol. no. 2, pp. 1–11, 2020. (2020).
- Hicham, K. et al. May. 3D CNN-BN: A Breakthrough in Colorectal Cancer Detection with Deep Learning Technique. In *2024 4th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)* (pp. 1–6). IEEE. (2024).
- Kara, O. C., Venkatayogi, N., Ikoma, N. & Alambeigi, F. A reliable and sensitive framework for simultaneous type and stage detection of colorectal cancer polyps. *Ann. Biomed. Eng.* **51** (7), 1499–1512 (2023).
- Tanwar, S. et al. Detection and classification of colorectal polyp using deep learning. *BioMed Research International*, *2022*(1), p.2805607. (2022).
- Alqushaibi, A. et al. Enhanced Colon Cancer Segmentation and Image Synthesis through Advanced Generative Adversarial Networks based-Sine Cosine Algorithm. *IEEE Access*. (2024).
- Mohamed, A. A. A., Hançerlioğullari, A., Rahebi, J., Ray, M. K. & Roy, S. Colon disease diagnosis with convolutional neural network and grasshopper optimization algorithm. *Diagnostics*, *13*(10), p.1728. (2023).
- Ragab, M. et al. Optimized Deep Learning Model for Colorectal Cancer detection and classification model. *CMC-Comput Mat. Contin.* **71** (3), 5751–5764 (2022).
- Alzubaidi, L. H., Priyanka, K., Suhas, S. K., Kumar, B. K. & Malathy, V. ResNet110 and Mask Recurrent Convolutional Neural Network Based Detection and Classification of Colorectal Cancer. In *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)* (pp. 1–4). IEEE. (2024).
- Liu, Y. et al. Fovea-UNet: detection and segmentation of lymph node metastases in colorectal cancer with deep learning. *BioMedical Engineering OnLine*, *22*(1), p.74. (2023).
- Karaman, A. et al. Hyper-parameter optimization of deep learning architectures using artificial bee colony (ABC) algorithm for high performance real-time automatic colorectal cancer (CRC) polyp detection. *Appl. Intell.* **53** (12), 15603–15620 (2023).
- Nur-A-Alam, M., Uddin, K. M. M., Manu, M. M. R., Rahman, M. M. & Nasir, M. K. An automatic system to detect colorectal polyp using hybrid fused method from colonoscopy images. *Intelligent Systems with Applications*, *22*, p.200342. (2024).
- Obayya, M. et al. Biomedical image analysis for colon and lung cancer detection using tuna swarm algorithm with deep learning model. *IEEE Access*. (2023).
- Pacal, I. & Karaboga, D. A robust real-time deep learning based automatic polyp detection system. *Computers in Biology and Medicine*, *134*, p.104519. (2021).
- Raju, A. S. N. & Venkatesh, K. EnsemDeepCADx: Empowering Colorectal Cancer Diagnosis with Mixed-Dataset Features and Ensemble Fusion CNNs on Evidence-Based CKHK-22 Dataset. *Bioengineering*, *10*(6), p.738. (2023).
- Abd El-Aziz, A. A., Mahmood, M. A. & Abd El-Ghany, S. Advanced Deep Learning Fusion Model for Early Multi-Classification of Lung and Colon Cancer Using Histopathological Images. *Diagnostics*, *14*(20), p.2274. (2024).
- Xia, Y., Yun, H. & Liu, Y. MFENet: Multi-scale feature enhancement and Fusion Network for polyp segmentation. *Computers in Biology and Medicine*, *157*, p.106735. (2023).
- Pacal, I. A novel swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *Int. J. Mach. Learn. Cybernet.*, pp.1–19. (2024).
- Karaman, A. et al. Robust real-time polyp detection system design based on YOLO algorithms by optimizing activation functions and hyper-parameters with artificial bee colony (ABC). *Expert systems with applications*, *221*, p.119741. (2023).
- Pacal, I. MaxCerVixT: a novel lightweight vision transformer-based Approach for precise cervical cancer detection. *Knowl. Based Syst.* **289**, 111482 (2024).
- Dwivedi, A. K., Srivastava, G. & Pradhan, N. NFF: A Novel Nested Feature Fusion Method for Efficient and Early Detection of Colorectal Carcinoma. In *Proceedings of Fourth International Conference on Computer and Communication Technologies: IC3T 2022* (pp. 297–309). Singapore: Springer Nature Singapore. (2023).
- Pacal, I. et al. An efficient real-time colonic polyp detection with YOLO algorithms trained by using negative samples and large datasets. *Computers in biology and medicine*, *141*, p.105031. (2022).
- Singh, O. & Singh, K. K. An approach to classify lung and colon cancer of histopathology images using deep feature extraction and an ensemble method. *Int. J. Inform. Technol.* **15** (8), 4149–4160 (2023).
- Sangeetha, S. K. B. et al. An enhanced multimodal fusion deep learning neural network for lung cancer classification. *Systems and Soft Computing*, *6*, p.200068. (2024).
- Poalelungi, D. G. et al. Advancing patient care: how artificial intelligence is transforming healthcare. *Journal of personalized medicine*, *13*(8), p.1214. (2023).
- Sureshkumar, V. et al. Breast cancer detection and analytics using hybrid CNN and extreme learning machine. *Journal of Personalized Medicine*, *14*(8), p.792. (2024).
- Srivastava, G., Chauhan, A. & Pradhan, N. Cjt-deo: Condorcet's jury theorem and differential evolution optimization based ensemble of deep neural networks for pulmonary and colorectal cancer classification. *Applied Soft Computing*, *132*, p.109872. (2023).
- Gowthamy, J. & Ramesh, S. A novel hybrid model for lung and colon cancer detection using pre-trained deep learning and KELM. *Expert Systems with Applications*, *252*, p.124114. (2024).

37. Raju, A. S. N., Rajababu, M., Acharya, A. & Suneel, S. Enhancing Colorectal Cancer Diagnosis With Feature Fusion and Convolutional Neural Networks. *Journal of Sensors*, 2024. (2024).
38. Shah, A. et al. Comparative analysis of median filter and its variants for removal of impulse noise from gray scale images. *J. King Saud University-Computer Inform. Sci.* **34** (3), 505–519 (2022).
39. Toğaçar, M., Ergen, B. & Cömert, Z. COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Computers in biology and medicine*, *121*, p.103805. (2020).
40. Gu, X., Tian, Y., Li, C., Wei, Y. & Li, D. Improved SE-ResNet Acoustic–Vibration Fusion for Rolling Bearing Composite Fault Diagnosis. *Applied Sciences*, *14*(5), p.2182. (2024).
41. Chang, Y. & Bao, G. Enhancing Rolling Bearing Fault Diagnosis in Motors using the OCSA-VMD-CNN-BiLSTM Model: A Novel Approach for Fast and Accurate Identification. *IEEE Access*. (2024).
42. Jiao, D. et al. Wavelength detection of serial WDM ultra-short fiber Bragg grating sensor networks based on a CCD interrogator using deep belief networks and sparrow search algorithm. *Opt. Express*. **32** (13), 22263–22279 (2024).
43. [www.warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/download](http://www.warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/download)
44. Ragab, M., Mahmoud, M. M., Asseri, A. H., Choudhry, H. & Yacoub, H. A. Optimal deep transfer learning based colorectal Cancer detection and classification model. *Computers Mater. Continua*, **74**(2) (2023).
45. Qiu, S., Lu, H., Shu, J., Liang, T. & Zhou, T. Colorectal Cancer Segmentation Algorithm based on deep features from enhanced CT images. *Computers Mater. Continua*, **80**(2). (2024).
46. Li, R. et al. Few-shot learning based histopathological image classification of colorectal cancer. *Intelligent Medicine*. (2024).

## Acknowledgements

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/267/45. Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R330), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Researchers Supporting Project number (RSPD2024R787), King Saud University, Riyadh, Saudi Arabia. This study is partially funded by the Future University in Egypt (FUE).

## Author contributions

Conceptualization: Sultan Refa Alotaibi Data curation and Formal analysis: Mashael Maashi, Ahmed Mahmud Investigation and Methodology: Sultan Refa Alotaibi, Manal Abdullah Alohali Project administration and Resources: Supervision; Manal Abdullah Alohali Validation and Visualization: Hamed Alqahtani, Moneerah Alotaibi Writing—original draft, Sultan Refa Alotaibi Writing—review and editing, Manal Abdullah Alohali, Ahmed Mahmud All authors have read and agreed to the published version of the manuscript.

## Declarations

### Competing interests

The authors declare no competing interests.

### Ethics approval

This article does not contain any studies with human participants performed by any of the authors.

### Additional information

**Correspondence** and requests for materials should be addressed to M.A.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025