International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 2, April 2024, pp. 2167~2180
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i2.pp2167-2180

A deep learning framework for accurate diagnosis of colorectal cancer using histological images

Maria M. Attia[1], Nihal Fayez F. Areed[1], Hanan M. Amer[1], Mervat El-Seddek[2]
[1]Department of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University, Mansoura, Egypt
[2]Department of Communications and Electronics Engineering, MISR Higher Institute for Engineering and Technology, Mansoura, Egypt

Article Info

Article history:

Received Sep 18, 2023
Revised Nov 26, 2023
Accepted Dec 5, 2023

Keywords:

Colorectal cancer
Convolutional neural network
Data preprocessing
Deep learning
Image classification

ABSTRACT

Colorectal cancer (CRC) is one of the most prevalent malignancies worldwide, with high mortality and incidence rates. Early detection of the disease may increase the probability of survival, making it critical to develop effective procedures for precise treatment. In the past few years, there has been an increased use of deep learning techniques in image classification that aid in the detection of various types of cancer. In this study, convolutional neural network (CNN) models were used to classify colorectal cancer into benign and malignant. After applying various data preprocessing techniques to the image dataset, we evaluated our prototypes using three distinct subsets of testing data, representing 20%, 30%, and 40% of the total dataset. Additionally, four pre-trained CNN models (ResNet-18, ResNet-50, GoogLeNet, and MobileNetV2) were trained, and the network architectural techniques were compared by applying the Adam optimizer. Finally, we assessed the performance of algorithms in terms of accuracy, sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic (ROC) curve (AUC). In this research, deep learning approaches demonstrated high efficacy in accurately diagnosing colorectal cancer. This indicates that these techniques have an important and significant value for advancing medical research.

Corresponding Author:

Mervat El-Seddek
Department of Communications and Electronics Engineering, MISR Higher Institute for Engineering and Technology
Mansoura-35516, Egypt
Email: mervat.elseddek@ieee.org

1. INTRODUCTION

Colorectal cancer (CRC) ranks third among the most frequently diagnosed illnesses in both men and women [1]. According to projections published by the American Cancer Society, the United States is expected to witness 46,050 newly diagnosed cases of rectal cancer and 106,970 newly diagnosed cases of colon cancer in 2023. The incidence rate of colon or rectal cancer has been decreasing in the United States since the mid-1980s. This decline is believed to be due to increased awareness of the importance of screening and individual risk factors resulting from lifestyle choices. Between 2011 and 2019, the incidence rates decreased by approximately 1%, as per the statistics. However, since the mid-1990s, rates for individuals under the age of 50 have been observed to increase by 1%–2% per year [2].

Once a polyp transitions to cancer, it has the potential to infiltrate the colon or rectum wall, where it may infect the blood or lymph vessels responsible for removing and transporting waste and fluid from cells. Cancer cells frequently spread to adjacent lymph nodes, which are bean-shaped structures that play a role in fighting infections. These cells can also travel through blood vessels to various other tissues and organs, including the liver or lungs, or they can be discharged directly into the peritoneum, the membrane that lines the abdomen [3], [4]. This process of cancer cells spreading to distant areas from the original tumor site is known as metastasis. Figure 1 illustrates the stages of colorectal cancer development, ranging from stage 0 to stage IIIC, and the spread of cancer to other organs.

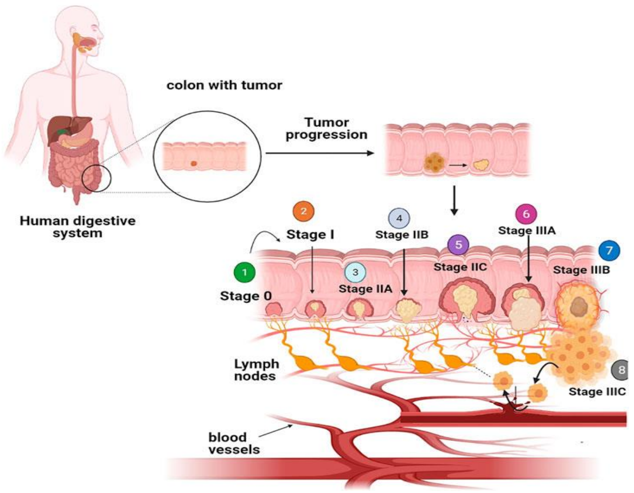

Figure 1. Colorectal cancer growth stages [5]

Colorectal cancer has a high prevalence globally, and early detection and diagnosis of CRC can significantly improve patients' survival rates. However, traditional CRC scanning and detection methods are often invasive, exhaustive, time-consuming, and more susceptible to errors. To address these limitations, researchers are investigating the integration of deep learning (DL), a subset of machine learning (ML), and artificial intelligence (AI) as a potential solution for developing more reliable diagnostic tools for CRC [6]. Convolutional neural networks (CNNs) have been extensively employed in successful deep learning applications and specifically utilized in this study, involving models such as ResNet-18, ResNet-50, GoogLeNet, and MobileNetV2. These CNNs have demonstrated successful applications in various deep-learning tasks, particularly in predicting and diagnosing colorectal cancer at an early stage. In this study, these CNN models have shown promising results and outperformed other approaches [7], [8].

The primary aim of this study is to predict and diagnose colorectal cancer at an early stage by implementing a comprehensive structure that employs CNN, which includes all the critical steps from image acquisition through disease detection. The proposed structure includes all necessary steps for accurately classifying and determining the presence of colorectal cancer. The paper presents several important contributions, including the following:

a. Accurate and effective prediction for colorectal cancer based on CNNs for more accurate colorectal cancer diagnosis, which can improve outcomes.
b. To preserve the consistency of the images, an image processing technique has been established. Data preprocessing and data augmentation techniques are also employed to expand and balance the data set.
c. ResNet-18, ResNet-50, GoogLeNet, and MobileNetV2 are four pre-trained CNN models that are applied for image classification.
d. Additionally, the Adam optimizer was employed to improve efficiency and performance. This optimization technique led to faster convergence and improved results.
e. The performance is assessed using several evaluation metrics, including accuracy, sensitivity, specificity, precision, F1-score, and area under the receiver operating characteristic (ROC) curve (AUC). Comparison and evaluation based on key performance measures are essential for evaluating models.
f. These methods are expected to advance the understanding of models for classification, increase the ability for the precise and efficient detection and diagnosis of colorectal cancer, and promote further study into the application of related methods to medical image processing, establishing opportunities for medical diagnosis and treatment.

A full article usually follows a standard structure: section 2 the dataset as well as the techniques for classification used in this study are demonstrated, along with the pre-processing methods, data augmentation, deep learning algorithms, and the applied optimizer. Additionally, it contains the criteria used to assess the performance of the model. Section 3 the empirical results of the suggested strategies are introduced, along with a comparison of the outcomes. Section 4 the conclusion and future work are provided.

Using various approaches and datasets, numerous studies have been conducted concerning the diagnosis of colorectal disease. The widespread use of computer vision tasks and supervised learning approaches demonstrates how strongly digital technology is currently employed in classifying medical images. In this study, Lusted [9] discusses how computers and electronic devices can be used in healthcare and medicine, as well as how digital technology may be used for the investigation and diagnosis of medical diseases.

Al-Shawesh and Chen [10] introduced a technique to enhance the accuracy of classification, which included the use of data sets from the national center for tumor diseases (NCT). To classify the CRC histopathology images, they employed the ResNet-50 model. The results demonstrate that the ResNet-50 structure, with a validation dataset accuracy of 97.7%, is the best CNN architecture for categorizing CRC.

Yoon *et al*. [11] employed CNN to propose a technique for classifying tumors from colorectal histology images. The performance of the model was assessed by implementing two experiments. The second experiment involved using a modified model, known as visual geometry group-E (VGG-E), derived from the visual geometry group (VGG), and produced the highest accuracy and specificity (93.48% and 92.76%).

Kather *et al*. [12] conducted an analysis of a multi-class issue in 5000 histological images involving tumor epithelium along with simple stroma using various textual descriptors. Four methods were suggested for classification, including the utilization of the k-nearest neighbor algorithm (k-NN), a support vector machines (SVM) decision function for classifying all categories, the construction of decision tree models using the random under-sampling boosting (RUSBoost) method, and training the classifiers with a 10-fold cross-validation without a specific stratification approach. The study showed that the SVM method achieved the highest accuracy of 87.4% across eight classes.

Kassani *et al*. [13] evaluated several deep learning-based models for the automatic segmentation of tumors in colorectal tissue samples. The suggested method emphasizes the value of implementing transfer learning and convolutional neural network modules in a segmentation architecture's encoder section for histopathology image processing. The results show that the accuracy is 87.07%±1.56.

Wang *et al*. [14] suggest a novel based on the bilinear convolutional neural network technique (BCNN) that initially separates the images into hematoxylin and eosin stain components in order to classify colorectal cancer histopathological images. BCNN has been applied to deconstructed images for combing and enhancing feature representation efficiency. The proposed BCNN-based algorithm is more accurate than the conventional CNN, according to experimental findings, which showed an accuracy of 92.6%±1.2.

Sarwinda *et al*. [15] conducted a study that employed deep learning techniques with residual network (ResNet) variants for image classification in the identification of colorectal cancer. The study employed ResNet-18 and ResNet-50, achieving an accuracy range of 73% to 88%. The findings of this study demonstrate the efficacy of DL in precisely diagnosing images of CRC.

Du *et al*. [16] conducted a study where they found that learning the fundamental features of CNNs outperformed the creation of custom features. This approach also allowed for the automatic differentiation of the epithelial and stromal parts of the breast. Moreover, using a network architecture layer technique, they achieved an accuracy of 84% in distinguishing colorectal cancers from malignant tissue. A transfer learning strategy was also employed with GoogLeNet, which resulted in an accuracy of 90.2%. This suggests that GoogLeNet can potentially be used to categorize the tumor-stroma ratio (TSR). Additionally, Kather *et al*. [17] reached the best accuracy rate of 98.7% by substituting the classification layer with VGG19.

Song *et al*. [18] developed a diagnostic model for colorectal adenomas with the implementation of deep CNNs. The model was constructed using the DeepLab v2 architecture combined with ResNet-34. The accuracy of the deep learning model was 90.4%, with an area under the curve of 0.92.

Hsu *et al*. [19] employed CNN to detect and classify colorectal polyps in grayscale, red-green-blue (RGB), and narrow-band images. Grayscale images estimated the highest accuracy of 95.1%, surpassing RGB (94.1%) and narrow-band (82.8%) images. These findings suggest that grayscale images have the potential to achieve even higher accuracy in polyp detection, especially in scenarios where lightweight processing is preferred.

In this study, Iizuka *et al*. [20] devised a highly efficient system to detect gastric and colonic epithelial cancer by combining convolutional and recurrent neural networks. The research involved collecting histological images of colorectal cancer from patients, followed by the extraction of effective features. These features were then trained using a convolutional network and a recurrent neural network. The developed network analyzed whole slide images comprehensively and accurately categorized adenoma, adenocarcinoma, and non-neoplastic tumors. The efficacy of the system was evaluated through experiments, achieving high predictive accuracy for colon tumors with an area under the curve (AUC) value of 0.97.

2. METHOD

This study presents a proposed approach that utilizes images of colorectal glands for the purpose of detecting and diagnosing colorectal cancer. The methodology involves several essential stages, which include data collection, data preprocessing, feature extraction, classification, and performance evaluation. These stages are visually represented in Figure 2. In the subsequent sections, a comprehensive discussion of each step will be conducted.

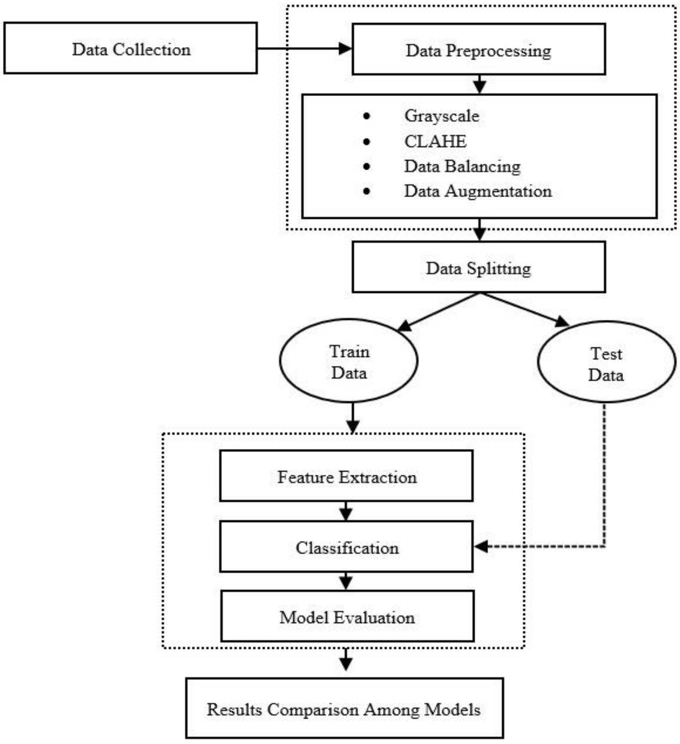

Figure 2. Diagram illustrating the proposed method

2.1. Data description

Colorectal gland images were obtained from a dataset provided by Warwick-QU via www.warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/download. The dataset contains 165 images, 74 of which indicate benign tumors and 91 of which indicate malignant tumors. The dataset was stained with H&E and scanned with a Zeiss MIRAX MIDI Scanner at 20x magnification. The image data resolution range of 567×430 pixels to 775×522 pixels, and the image data weight range of 716 kilobytes, 1.187 kilobytes. The images had an isotropic 0.62 µm pixel resolution. The depicted model in Figure 3 showcases images of colon glands extracted from a subset of the Warwick-QU dataset. In Figure 3(a), a benign gland is represented, while Figure 3(b) illustrates a malignant gland.

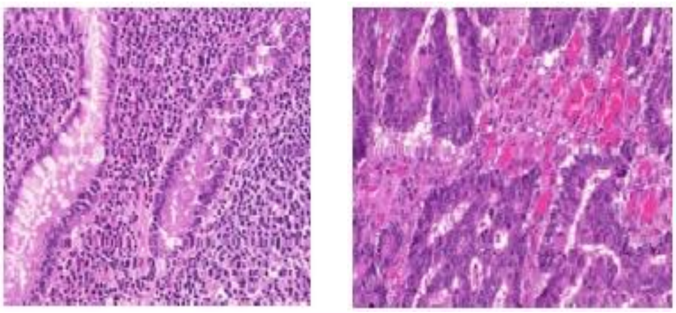

Figure 3. The Warwick-QU dataset sample (a) benign gland and (b) malignant gland [21]

2.2. Data preprocessing

Data preprocessing refers to the transformation and preparation of the initial data before its input into a deep learning model. Normalization, feature scaling, dimensionality reduction, and other preparation techniques are only a few examples of the preprocessing operations in deep learning. Deep learning models' effectiveness and precision can be improved with the right preprocessing steps. The objective is to improve the quality of the data, its relevance, and its suitability for the specified employment. To expedite processing and ensure that all images would work with trained CNN models, each image was initially downsized to 224×224×3. The application of preprocessing methods to images is indicated in Figure 4. In this Figure, Figure 4(a) depicts the initial input images, while Figure 4(b) showcases the grayscale versions of these images. Additionally, Figure 4(c) showcases the outcome after applying contrast-limited adaptive histogram equalization.

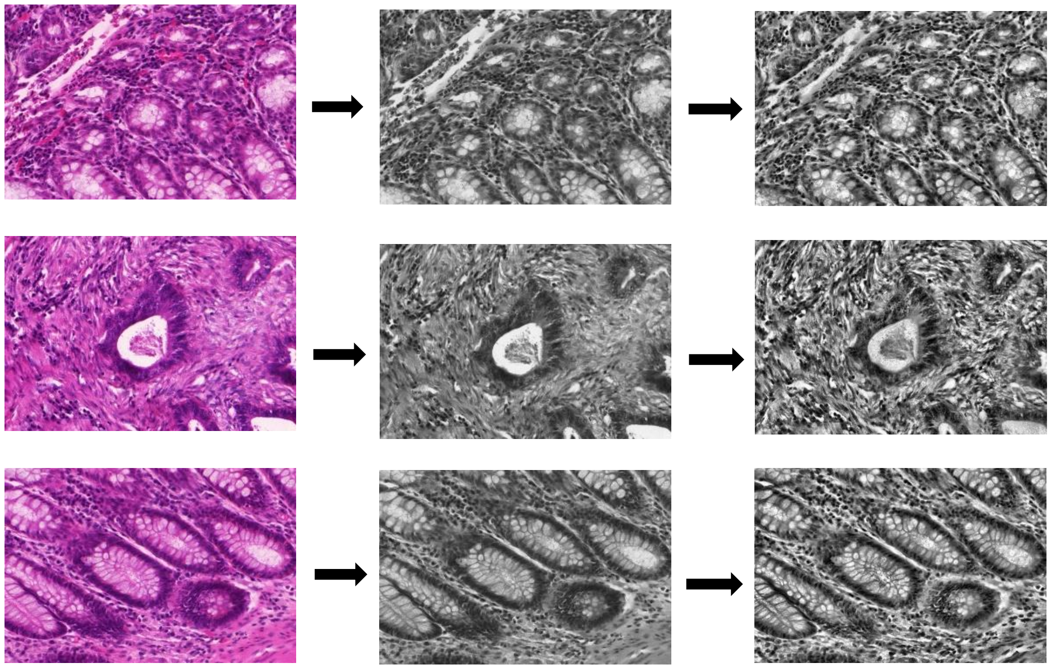

Figure 4. Illustrates the pre-processing steps (a) the input dataset, (b) images in grayscale, and (c) applying the contrast-limited adaptive histogram equalization

2.2.1. Employ grayscale images

Grayscale images offer a range of advantages, such as decreased memory usage and computational complexity, making them popular in medical image processing. Grayscale images are widely employed in digital photography, computer-generated imaging, and colorimetry. Each pixel in a grayscale image is assigned a value that exclusively represents the amount of light it emits, ranging from black for the weakest intensity to white for the strongest. These images are monochromatic, composed of various shades of gray. To generate grayscale images, the light intensity produced by each pixel is measured, utilizing a weighted combination of frequencies from any part of the electromagnetic spectrum, such as infrared or ultraviolet. According to the studied characteristics of human vision, a colorimetric grayscale image is defined as having a specific grayscale color space that correlates the obtained numeric sample values to the achromatic channel of a standard color space [22].

2.2.2. Contrast-limited adaptive histogram equalization

The contrast-limited adaptive histogram equalization (CLAHE) method, an extension of the histogram equalization (HE) approach, was initially developed for medical images. It has been successfully employed to enhance low-contrast images [23], [24]. The CLAHE approach aims to improve the visibility of curves and edges in various parts of an image [25]. The number of sub-images and the clip limit significantly influence the outcomes obtained from the CLAHE technique. Three distinct sets of regions, namely the corner region (CR), the border region (BR), and the inner region (IR), are created by dividing the input image into sub-images (tiles) [26]. For each tile, a contrast transform function is generated using the "Clip Limit" as the contrast factor [27]. The study used the CLAHE enhancement technique with a clip limit of 0.02. By employing the CLAHE method, solid images can be produced, and the contrast across all areas of the image can be made consistent. The expression as in (1) for standard CLAHE with uniform distribution:

$$g = [\, g_{max} - g_{min}] * P(f) + g_{min} \tag{1}$$

where g is the calculated pixel value, g_{max} is maximum pixel value, g_{min} is minimum pixel value. For exponential distribution to enhance results expressed as shown in (2).

$$g = g_{min} - \left(\frac{1}{\alpha}\right) * \ln[1 - P(f)\,] \tag{2}$$

where $P(f) = CPD$ (cumulative probability distribution), α is the clip parameter.

2.2.3. Data balancing

Dealing with imbalanced training datasets is recognized as a significant challenge within the realm of machine learning [28]. Both down-sampling and over-sampling are popular methods used to address this issue. When there is an imbalance in the distribution of images among different classes, over-sampling the training dataset becomes a technique of major importance. On the other hand, in the down-sampling approach, the number of images for each specific class is reduced to match the lowest count observed across all classes. In contrast, over-sampling involves increasing the number of images within each class. The over-sampling strategy was ultimately chosen for use in this study. The difference between over-sampling and down-sampling is shown in Figure 5.

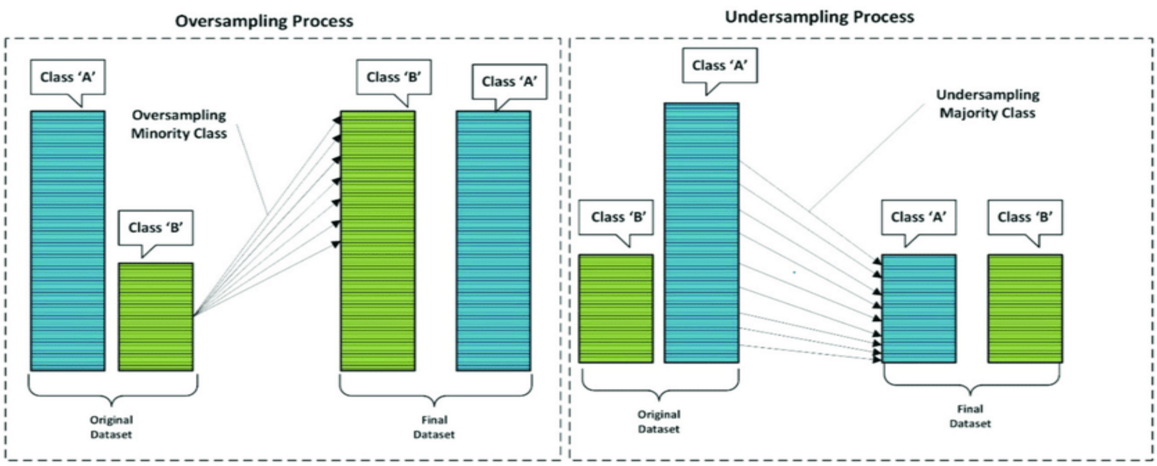

Figure 5. Illustrates the distinctions between down-sampling and over-sampling [29]

2.2.4. Data augmentation

In the field of deep learning, a variety of techniques referred to as data augmentation are employed to increase the number of datasets. By expanding the quantity and diversity of the datasets, data augmentation helps enhance the accuracy and generalization of deep learning models by generating additional samples through various transformations [30]. In addition, data augmentation aids in the prevention of overfitting during the training process, enhancing the robustness and stability of the models [31], [32]. When using data augmentation techniques, new data samples are generated through various transformations. Several techniques are used to augment the data, including random reflections, shearing, translations, scaling, and rotating the images by an angle of up to 20 degrees.

2.3. Feature extraction with deep learning

In deep learning, feature extraction involves the automated process of learning and extracting significant features from raw data using deep neural networks [33]. This procedure is essential for enabling deep learning models to carry out a variety of tasks, including object recognition, categorizing images, and processing natural language. Deep learning models, including CNNs, deep belief networks (DBNs), restricted Boltzmann networks (RBMs), and recurrent neural networks (RNNs), have been developed to automatically learn and extract features from input data [34]. These models, with their multiple layers, are designed to facilitate the extraction of relevant features. Four pre-trained CNN models are used in this study for extracting features. An overview of the chosen models is presented in the subsequent sections.

2.3.1. Convolutional neural network architecture

CNN, or ConvNet, which is a type of deep neural network, has been effectively utilized in a wide range of applications. CNNs are primarily designed for processing and analyzing grid-like data, such as images and videos. Unlike standard neural networks, CNNs employ specialized convolution and pooling operations to recognize and extract features from the input data. Therefore, they are well-suited for tasks like image identification and classification. CNNs can also be applied to various artificial intelligence and computer vision applications. They can automatically learn and recognize patterns within visual data [35]. The CNN architecture is illustrated in Figure 6.

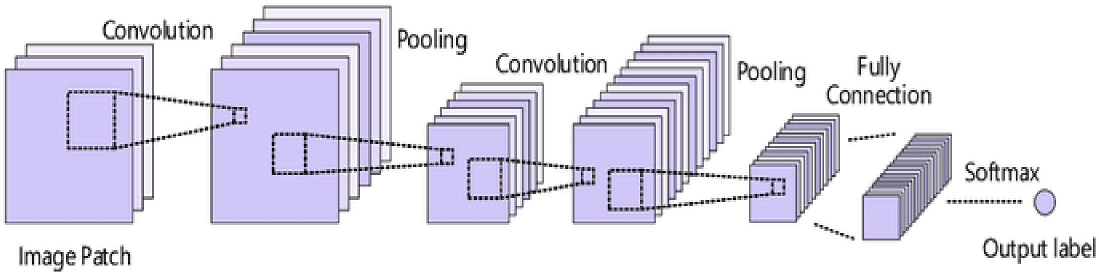

Figure 6. The structure of CNN [36]

a. The input layer of an artificial neural network is where the initial data is introduced for processing by subsequent layers of neurons. This initial data comprises features such as images, their sizes, and input channels. This layer serves as the starting point of the network and aids in extracting significant features from the raw data, which facilitates processing by subsequent layers.
b. Convolutional layers are utilized for feature extraction from input images, including edges and corners. These layers perform convolution operations between the image input, an image-sized filter for convolution of the input data, and scanning the entire image with a fixed pixel stride. The feature map obtained is then forwarded to additional layers in order to determine additional features of the input image.
c. In artificial neural networks, the rectified linear unit (ReLU) is the most prevalent activation function. As shown in (3), the nonlinear function ReLU is displayed. It enables the elimination of negative values from the CNN architecture.

$$F(x) = max\ (0,1) \tag{3}$$

d. In convolutional neural networks, pooling layers are a type of layer that is frequently used to downsample feature maps by summing the presence of features in patches of the feature maps. They enhance the effectiveness of the network by reducing the dimensionality of the data by integrating the outputs of neuron clusters at the previous layer into one neuron in the subsequent layer. Max, average, and stochastic pooling are a few of the several types of pooling.
e. Fully connected layers are frequently added at the end of a neural network and are typically employed as classifiers to determine classification. These layers, which stand for the final output of the network, can also be layered to enhance efficiency.

2.3.2. Four pre-trained convolutional neural network models

Four deep neural networks built using CNN models are examined in this study. Deep neural networks have been developed to distinguish different tissue samples by utilizing characteristics identified in CNN models. An approach for improving the accuracy and efficacy of diagnosing colorectal cancer in applications employing medical image processing is provided by the proposed methodology. An overview of the key features of the models employed in this investigation is shown in Table 1.

Table 1. Features of the models employed in this study

Network	Parameters (Millions)	Layers (Depth)	Total layers	Input size
ResNet-18	11.7	18	71	224x224x3
ResNet-50	25.6	50	177	224x224x3
GoogLeNet	7	22	144	224x224x3
MobileNetV2	3.5	53	154	224x224x3

Residual learning, sometimes referred to as residual networks or ResNets, is a deep learning method for enhancing neural network performance. ResNet proves to be a highly efficient and successful artificial neural network architecture that applies the application of residual learning to construct ultra-deep networks. The "deep residual learning" neural network architecture was introduced by He *et al*. [37]. ResNet-18 and ResNet-50 were implemented in this study. Residual learning operates by including a shortcut connection that passes through one or more neural network layers. The vanishing gradients issue, which can arise in deep neural networks, is addressed by the shortcut connection, which enables the model to transport information without loss from earlier layers to later layers. The architecture for residual learning is illustrated in Figure 7. The ResNet block function can be represented as shown in (4).

$$output = F(x) + x \tag{4}$$

where x represents the input that is being passed through the residual block.

GoogleNet is an Inception-based deep convolutional neural network consisting of 22 layers. The Inception architecture is designed to incorporate optimal local sparse structure spatially repeated from the beginning to the end of a vision network. To enhance performance and efficiency in various scenarios, GoogleNet introduces three Inception structures. In most cases, these structures utilize 1×1 convolution to compute reductions before the more resource-intensive 3×3 and 5×5 convolutions are applied. By using GoogleNet as an inspiration, advancements in identification capability are achieved through higher-powered hardware, greater datasets, and larger models, as well as through new algorithms, ideas, and improved network architectures. Figure 8 shows how GoogleNet is structured.

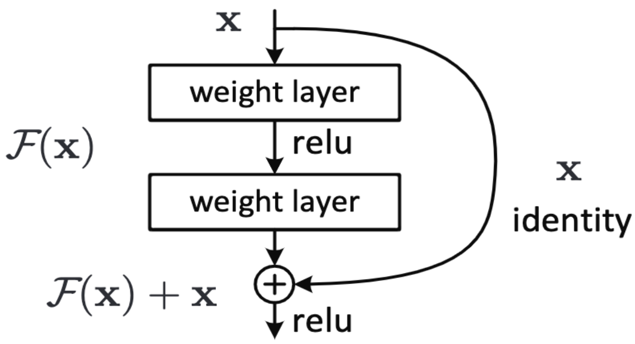

Figure 7. Residual block

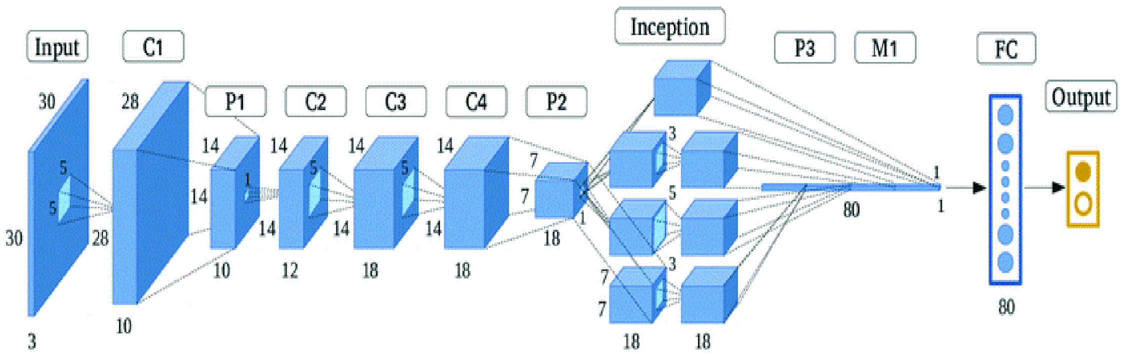

Figure 8. GoogLeNet architecture [38]

A convolutional neural network structure indicated by MobileNetV2 employs an inverted residual structure. MobileNetV2, a pre-trained convolutional neural network, consists of 53 layers and was trained on a vast dataset of over a million images from ImageNet [39]. The model has been trained effectively to categorize images into 100 distinct object classes. The basic structure of MobileNetV2 is comprised of 32 filters and 19 bottleneck layers, which have been organized into fully convolutional layers. There are two distinct blocks in it, each with three levels. Both blocks have a 32-filter with 1x1 convolutional layers at the start and end, although the second layer is a convolutional layer that functions in depth order. At every level of the structure, the ReLU is utilized. The stride values vary among the two blocks, with the first block having a stride size of 1 and the second block having a stride length of 2, indicating a difference in stride values between the two blocks. Overall, MobileNetV2 is a simple and effective model. Figure 9 provides an illustration of the construction of the model.

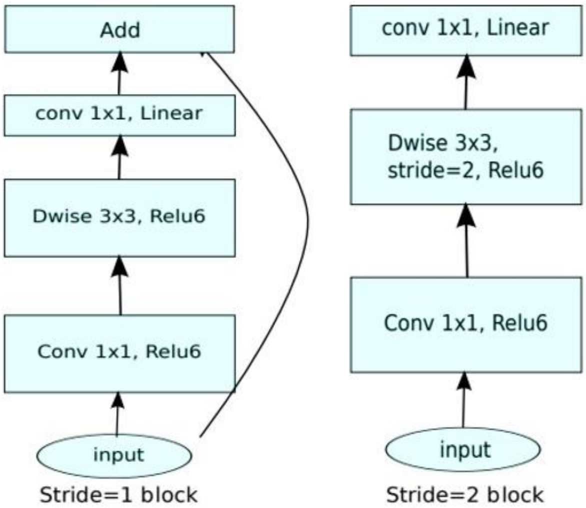

Figure 9. Architecture of MobileNetv2 [40]

2.4. Classification

Several classification techniques have been employed to differentiate between benign and malignant instances of colorectal cancer. These methods aim to simplify the identification of colorectal cancer. Along with training and evaluating the image dataset using a CNN model, several additional metrics, including accuracy, sensitivity, specificity, precision, F1-score, and area under the ROC curve (AUC), have been employed for evaluating the effectiveness of these classification methods.

2.5. Network optimizer

Deep learning optimizer techniques are employed to modify the weights and learning rate of a neural network with the objective of reducing the loss function during training. Deep learning employs a variety of optimizers. Some of the commonly used optimizers in deep learning are stochastic gradient descent with momentum (SGDM), adaptive moment estimation (Adam), adaptive gradient optimizer (AdaGrad), and root mean square propagation (RMSProp). The main objective of these optimizers is to find the ideal weights that enable the neural network to successfully convert input data into the desired result. There are various types of optimization algorithms available for deep learning models, and the choice of optimizer should be based on the specific requirements of the model and the dataset being used. Each optimizer has its own set of strengths and weaknesses. In this study, Adam optimizer employed [41].

The Adam optimization algorithm employs a first-order gradient strategy to optimize a stochastic function. It is a suitable method to be simply implemented with many parameters and data. It provides high computational efficiency and requires little in the form of hardware resources [42]. It is also appropriate to handle non-stationary goals and issues with noisy or sparse gradients. A popular optimization strategy in deep learning is the Adam technique, which combines both RMSProp and stochastic gradient descent. Adam contributes to balancing the learning rate for each weight in the model networks, which improves the stability of the optimization process and lowers the possibility of overfitting. Overall, Adam is an efficient optimization strategy that has been widely employed in the field of deep learning due to its successful performance. The Adam optimizer is expressed mathematically, as shown in (5) and (6).

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\, g_t \quad (5)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\, {g_t}^2 \quad (6)$$

2.6. Performance evaluation

The confusion matrix produced for different models is used to generate components including true positive (TP), false positive (FP), true negative (TN), and false negative (FN). These elements are crucial for evaluating the models [43]. Additionally, the confusion matrix components are employed to calculate important metrics like accuracy (ACC), precision positive predictive value (PPV), sensitivity (Recall), specificity (SPC), and F1-score.

– Accuracy is the ratio of correctly predicted values to all predicted values, according to (7).

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

– Recall is the ratio of positive samples predicted out of the total number of samples that are positive, as given in (8).

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

– The number of samples predicted to be negative out of all samples that are negative is known as the specificity (SPC), as given in (9).

$$SPC = \frac{TN}{TN+FP} \quad (9)$$

– As given in (10), precision refers to the ratio of correctly classified samples from the actual positive samples in the detected class.

$$PPV = \frac{TP}{TP+FP} \quad (10)$$

– F1-score, which can be calculated from (11), represents overall performance as determined by the model's recall and precision.

$$F1 - score = \frac{2*precision*Recall}{Precision+Recall} \quad (11)$$

3. RESULTS AND DISCUSSION

3.1. The analysis

The analysis was performed using a computer running Windows 11 (64-bit), an Intel (R) Core (TM) i5-1135G7 CPU, and the MATLAB 2021a programming language. Three different models were employed to train and test the data. In the first model, 80% of the data was allocated for training and 20% for testing. The second model used 70% of the data for training and 30% for testing. Finally, the third model distributed the data, with 60% used for training and 40% for testing. Four distinct CNN models, specifically ResNet-18, ResNet-50, GoogLeNet, and MobileNetV2, were trained using the Adam optimizer to improve the productivity and effectiveness of the learning process. These techniques made it possible to assess how well the models were able to classify and estimate the outcomes based on test data. Table 2 includes the variables used to train CNN models.

Table 2. MATLAB training options for deep learning neural networks

Configurations	Values
Input size	224x224x3
Max epochs	8
Mini batch size	10
Shuffle	Every epoch
Validation frequency	5
Learning rate	0.00001
Optimizer	Adam

3.2. The results of the applied models

Table 3 presents the distinct outcomes achieved by employing four pre-trained CNN models across three different models. The first model demonstrated exceptional performance, with ResNet-18 and ResNet-50 achieving 100% accuracy, specificity, and sensitivity. GoogLeNet also performed well, achieving an accuracy of 97%, a specificity of 100%, and a sensitivity of 94.4%. MobileNetV2 achieved an accuracy of 97%, a specificity of 93.3%, and a sensitivity of 100% in this first model. In the second model, the highest accuracy, specificity, and sensitivity were achieved by ResNet-18, ResNet-50, and GoogLeNet, all reaching 100%. On the other hand, MobileNetV2 achieved an accuracy of 89.8%, a specificity of 77.3%, and a sensitivity of 100% in this model. Moving to the third model, ResNet-18 attained an accuracy of 97%, a specificity of 96.7%, and a sensitivity of 97.2%. ResNet-50 achieved a high accuracy of 97%, a specificity of 100%, and a sensitivity of 94.4%. GoogLeNet achieved an accuracy of 90.9%, a specificity of 83.3%, and a sensitivity of 97.2%. The lowest accuracy was obtained from MobileNetV2, with an accuracy of 89.4%, a specificity of 76.7%, and a sensitivity of 100%. Overall, the best results were obtained from both the first model, which employed 80% of the training data and 20% of the testing data, and the second model, which utilized 70% of the training data and 30% of the testing data.

Numerous studies have been conducted to provide projections for colorectal cancer employing deep learning, machine learning, and other techniques depending on various imaging data for cancer cells. Table 4 displays several approaches utilized by researchers, highlighting the various datasets used in addition to the suggested approach. While earlier studies have demonstrated remarkable accuracy in predicting colorectal cancer, the proposed approach has also produced promising results, with accuracy of up to 100% attained in specific cases.

Table 3. Performance results of the suggested methods for multiple testing datasets

Training data : Testing data	Network	SEN. %	SPE. %	ACC. %	F1-score. %	PRE. %	AUC	Elapsed time per epoch (second)
80%:20%	ResNet-18	100%	100%	100%	100%	100%	1.0000	21.28
	ResNet-50	100%	100%	100%	100%	100%	1.0000	51.07
	GoogLeNet	94.4%	100%	97%	97.12%	100%	1.0000	21.56
	MobileNetV2	100%	93.3%	97%	97.3%	94.7%	1.0000	36.05
70%:30%	ResNet-18	100%	100%	100%	100%	100%	1.0000	20.96
	ResNet-50	100%	100%	100%	100%	100%	1.0000	58.36
	GoogLeNet	100%	100%	100%	100%	100%	1.0000	18.14
	MobileNetV2	100%	77.3%	89.8%	91.5%	84.4%	1.0000	27.82
60%:40%	ResNet-18	97.2%	96.7%	97%	97.2%	97.2%	0.9991	16.67
	ResNet-50	94.4%	100%	97%	97.12%	100%	0.9991	44.18
	GoogLeNet	97.2%	83.3%	90.9%	92.1%	87.5%	0.9870	14.34
	MobileNetV2	100%	76.7%	89.4%	91.13%	83.7%	0.9954	29.82

Table 4. Comparison between the proposed method and previous results

References	Method	Accuracy %
Al. Shawesh and Chen [10]	Categorized the CRC histopathology image model employing ResNet-50.	97.7%
Yoon *et al*. [11]	Implementing the modified VGG E model-based CNN architecture.	93.48%
Kather *et al*. [12]	Applying decision trees, linear SVM, radial-basis SVM, and one-nearest neighbor for multi-class texture analysis in colorectal cancer histology.	87.4%
Kassani *et al*. [13]	The use of DenseNet and LinkNet architectures was demonstrated for automatically segmenting tumors in colorectal tissue samples.	87.07%±1.56
Wang *et al*. [14]	Applying BCNN for the categorization of colorectal cancer histopathological images.	92.6%±1.2
Sarwinda *et al.* [15]	Deep learning approach using the ResNet architecture for the detection of colorectal cancer through image classification.	73%-88%
Du *et al*. [16]	Classify tumor epithelium and stroma through transfer learning with the GoogLeNet architectur.	90.2%
Kather *et al*. [17]	Evaluating colorectal cancer histology slides to predict survival applying several CNN models.	98.7%
Song *et al*. [18]	Developed a model using the DeepLab v2 architecture with ResNet-34 for identifing colorectal adenomas.	90.4%
Hsu *et al*. [19]	Proposed a system that employs grayscale images and deep learning, specifically CNN model, to detect and classify colorectal polyp images.	82.8%-95.1%
Tanwar *et al*. [44]	Using deep learning for the detection and categorization of colorectal polyps.	92%
Wulczyn *et al*. [45]	A DL procedure was implemented to predict patient survival in colorectal cancer.	87%-95.5%
Proposed method	Early prediction and diagnosis of colorectal cancer by implementing four pre-trained CNN models.	89.4%-100%

The surgical sector has witnessed tremendous and ongoing technical advancements in recent years. The adoption of the internet of things (IoT) concept in surgical practice stands out as one of the most revolutionary developments [46]. A subset of IoT identified as the internet of surgical things (IoST) involves the integration of software, sensors, and smart surgical instruments to enhance the effectiveness, safety, and results of surgical procedures. IoST may be beneficial in primary disease diagnosis, particularly in colorectal cancer. DL could serve as a prompt and effective confirmation test following the initial conventional clinical inquiry while also assisting trainee doctors in gaining expertise in CRC diagnosis throughout their internship years. Moreover, global health systems may benefit from the incorporation of DL into CRC evaluation [47]. Further investigation has been conducted into the potential clinical practice implementation of deep learning algorithms for the categorization and diagnosis of colorectal cancer histopathology images. The advancement made possible by deep learning algorithms has the potential to improve colorectal cancer detection's accuracy and efficacy [48].

4. CONCLUSION AND FUTURE WORK

Over the past few years, deep learning and machine learning techniques have made a significant impact on various fields, including image processing and the medical industry. These techniques have proven to be valuable in the screening of different types of cancer and in highlighting areas of concern during cancer diagnosis by specialists. The study aims to speed up and simplify the procedure for accurate identification of colorectal cancer, along with making it more effective and real-time. To achieve this, multiple deep learning models, particularly CNNs were employed with the Adam optimizer. According to the results of this study, the four pre-trained CNN models can effectively identify colorectal cancer, achieving a precision range of 83.7% to 100%, sensitivity values ranging from 94.4% to 100%, and an accuracy range of 89.4% to 100%. In future work, the suggested approach can undo training and testing with a larger dataset while also being assessed on different datasets for the purpose of classification and prediction. Collaborating with medical researchers specializing in colorectal cancer within hospitals or clinics would greatly contribute to the advancement and practical application of this research within the medical field.

REFERENCES

[1] A. S. Narasimha Raju, K. Jayavel, and T. Rajalakshmi, "Intelligent recognition of colorectal cancer combining application of computer-assisted diagnosis with deep learning approaches," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, pp. 738–747, 2022, doi: 10.11591/ijece.v12i1.pp738-747.

[2] R. L. Siegel, N. S. Wagle, A. Cercek, R. A. Smith, and A. Jemal, "Colorectal cancer statistics, 2023," *CA: A Cancer Journal for Clinicians*, vol. 73, no. 3, pp. 233–254, May 2023, doi: 10.3322/caac.21772.

[3] D. V. F. Tauriello, A. Calon, E. Lonardo, and E. Batlle, "Determinants of metastatic competency in colorectal cancer," *Molecular Oncology*, vol. 11, no. 1, pp. 97–119, Jan. 2017, doi: 10.1002/1878-0261.12018.

[4] P. Quere *et al.*, "Epidemiology, management, and survival of peritoneal carcinomatosis from colorectal cancer," *Diseases of the Colon & Rectum*, vol. 58, no. 8, pp. 743–752, Aug. 2015, doi: 10.1097/DCR.0000000000000412.

[5] B. Brar *et al.*, "Nanotechnology in colorectal cancer for precision diagnosis and therapy," *Frontiers in Nanotechnology*, vol. 3, Sep. 2021, doi: 10.3389/fnano.2021.699266.

[6] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.

[7] S. Suhirman, S. Saifullah, A. T. Hidayat, M. A. Kusuma, and R. Drezewski, "Real-time mask-wearing detection in video streams using deep convolutional neural networks for face recognition," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 1, pp. 1005-1014, Feb. 2024, doi: 10.11591/ijece.v14i1.pp1005-1014.

[8] D. Alboaneen *et al.*, "Predicting colorectal cancer using machine and deep learning algorithms: Challenges and opportunities," *Big Data and Cognitive Computing*, vol. 7, no. 2, p. 74, Apr. 2023, doi: 10.3390/bdcc7020074.

[9] L. B. Lusted, "Medical electronics," *New England Journal of Medicine*, vol. 252, no. 14, pp. 580–585, Apr. 1955, doi: 10.1056/NEJM195504072521405.

[10] R. A. Al-Shawesh and Y. X. Chen, "Enhancing histopathological colorectal cancer image classification by using convolutional neural network," *medRxiv*, p. 2021.03.17.21253390, 2021, doi: 10.1101/2021.03.17.21253390.

[11] H. Yoon *et al.*, "Tumor identification in colorectal histology images using a convolutional neural network," *Journal of Digital Imaging*, vol. 32, no. 1, pp. 131–140, Feb. 2019, doi: 10.1007/s10278-018-0112-9.

[12] J. N. Kather *et al.*, "Multi-class texture analysis in colorectal cancer histology," *Scientific Reports*, vol. 6, no. 1, Art. no. 27988, Jun. 2016, doi: 10.1038/srep27988.

[13] S. Hosseinzadeh Kassani, P. Hosseinzadeh Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Deep transfer learning based model for colorectal cancer histopathology segmentation: A comparative study of deep pre-trained models," *International Journal of Medical Informatics*, vol. 159, p. 104669, Mar. 2022, doi: 10.1016/j.ijmedinf.2021.104669.

[14] C. Wang, J. Shi, Q. Zhang, and S. Ying, "Histopathological image classification with bilinear convolutional neural networks," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, IEEE, Jul. 2017, pp. 4050–4053. doi: 10.1109/EMBC.2017.8037745.

[15] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer," *Procedia Computer Science*, vol. 179, pp. 423–431, 2021, doi: 10.1016/j.procs.2021.01.025.

[16] Y. Du *et al.*, "Classification of tumor epithelium and stroma by exploiting image features learned by deep convolutional neural

networks," *Annals of Biomedical Engineering*, vol. 46, no. 12, pp. 1988–1999, 2018, doi: 10.1007/s10439-018-2095-6.

[17] J. N. Kather *et al.*, "Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study," *PLOS Medicine*, vol. 16, no. 1, p. e1002730, Jan. 2019, doi: 10.1371/journal.pmed.1002730.

[18] Z. Song *et al.*, "Automatic deep learning-based colorectal adenoma detection system and its similarities with pathologists," *BMJ Open*, vol. 10, no. 9, Art. no. e036423, Sep. 2020, doi: 10.1136/bmjopen-2019-036423.

[19] C.-M. Hsu, C.-C. Hsu, Z.-M. Hsu, F.-Y. Shih, M.-L. Chang, and T.-H. Chen, "Colorectal polyp image detection and classification through grayscale images and deep learning," *Sensors*, vol. 21, no. 18, p. 5995, Sep. 2021, doi: 10.3390/s21185995.

[20] O. Iizuka, F. Kanavati, K. Kato, M. Rambeau, K. Arihiro, and M. Tsuneki, "Deep learning models for histopathological classification of gastric and colonic epithelial tumours," *Scientific Reports*, vol. 10, no. 1, Art. no. 1504, Jan. 2020, doi: 10.1038/s41598-020-58467-9.

[21] K. Sirinukunwattana, S. E. A. Raza, Y. W. Tsang, D. R. J. Snead, I. A. Cree, and N. M. Rajpoot, "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1196–1206, 2016, doi: 10.1109/TMI.2016.2525803.

[22] A. Güneş, H. Kalkan, and E. Durmuş, "Optimizing the color-to-grayscale conversion for image classification," *Signal, Image and Video Processing*, vol. 10, no. 5, pp. 853–860, 2016, doi: 10.1007/s11760-015-0828-7.

[23] G. Yadav, S. Maheshwari, and A. Agarwal, "Contrast limited adaptive histogram equalization based enhancement for real time video system," Proceedings of the 2014 International Conference on Advances in Computing, Communications and *Informatics, ICACCI 2014*, pp. 2392–2397, 2014, doi: 10.1109/ICACCI.2014.6968381.

[24] M. S. Hitam, E. A. Awalludin, W. N. Jawahir Hj Wan Yussof, and Z. Bachok, "Mixture contrast limited adaptive histogram equalization for underwater image enhancement," *International Conference on Computer Applications Technology, ICCAT 2013*, 2013, doi: 10.1109/ICCAT.2013.6522017.

[25] P. Singh, R. Mukundan, and R. De Ryke, "Feature enhancement in medical ultrasound videos using contrast-limited adaptive histogram equalization," *Journal of Digital Imaging*, vol. 33, no. 1, pp. 273–285, 2020, doi: 10.1007/s10278-019-00211-5.

[26] A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color retinal image enhancement using CLAHE," *Proceedings - International Conference on ICT for Smart Society 2013: "Think Ecosystem Act Convergence," ICISS 2013*, pp. 215–217, 2013, doi: 10.1109/ICTSS.2013.6588092.

[27] N. Salem, H. Malik, and A. Shams, "Medical image enhancement based on histogram algorithms," *Procedia Computer Science*, vol. 163, pp. 300–311, 2019, doi: 10.1016/j.procs.2019.12.112.

[28] R. Mohammed, J. Rawashdeh, and M. Abdullah, "Machine learning with oversampling and undersampling techniques: Overview study and experimental results," *2020 11th International Conference on Information and Communication Systems, ICICS 2020*, pp. 243–248, 2020, doi: 10.1109/ICICS49469.2020.239556.

[29] V. Kumar *et al.*, "Addressing binary classification over class imbalanced clinical datasets using computationally intelligent techniques," *Healthcare (Switzerland)*, vol. 10, no. 7, 2022, doi: 10.3390/healthcare10071293.

[30] L. T. Menon, I. A. Laurensi, M. C. Penna, L. E. S. Oliveira, and A. S. Britto, "Data augmentation and transfer learning applied to charcoal image classification," *International Conference on Systems, Signals, and Image Processing*, vol. 2019-June, pp. 69–74, 2019, doi: 10.1109/IWSSIP.2019.8787271.

[31] L. G. Falconi, M. Perez, W. G. Aguilar, and A. Conci, "Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database," *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 2, pp. 154–165, 2020, doi: 10.25046/aj050220.

[32] S. Akter, F. M. J. M. Shamrat, S. Chakraborty, A. Karim, and S. Azam, "Covid-19 detection using deep learning algorithm on chest X-ray images," *Biology*, vol. 10, no. 11, 2021, doi: 10.3390/biology10111174.

[33] M. Ishaque and L. Hudec, "Feature extraction using deep learning for intrusion detection system," *2nd International Conference on Computer Applications and Information Security, ICCAIS 2019*, 2019, doi: 10.1109/CAIS.2019.8769473.

[34] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[35] Y. Rikiya, N. Mizuho, D. R. K. Gian, and T. Kaori, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, pp. 1–19, 2018.

[36] X. Mei *et al.*, "Spectral-spatial attention networks for hyperspectral image classification," *Remote Sensing*, vol. 11, no. 8, 2019, doi: 10.3390/rs11080920.

[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[38] Z. Guo, Q. Chen, G. Wu, Y. Xu, R. Shibasaki, and X. Shao, "Village building identification based on ensemble convolutional neural networks," *Sensors (Switzerland)*, vol. 17, no. 11, 2017, doi: 10.3390/s17112487.

[39] F. M. J. M. Shamrat *et al.*, "Alzheimernet: an effective deep learning based proposition for Alzheimer's disease stages classification from functional brain changes in magnetic resonance images," *IEEE Access*, vol. 11, pp. 16376–16395, 2023, doi: 10.1109/ACCESS.2023.3244952.

[40] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[41] A. Kumar, S. Sarkar, and C. Pradhan, "Malaria disease detection using CNN technique with SGD, RMSprop and ADAM optimizers," *Studies in Big Data*, vol. 68, pp. 211–230, 2020, doi: 10.1007/978-3-030-33966-1_11.

[42] M. Reyad, A. M. Sarhan, and M. Arafa, "A modified Adam algorithm for deep neural network optimization," *Neural Computing and Applications*, vol. 35, no. 23, pp. 17095–17112, 2023, doi: 10.1007/s00521-023-08568-z.

[43] D. A. Arafa, H. E. D. Moustafa, H. A. Ali, A. M. T. Ali-Eldin, and S. F. Saraya, "A deep learning framework for early diagnosis of Alzheimer's disease on MRI images," *Multimedia Tools and Applications*, 2023, doi: 10.1007/s11042-023-15738-7.

[44] S. Tanwar, S. Vijayalakshmi, M. Sabharwal, M. Kaur, A. A. Alzubi, and H. N. Lee, "Detection and classification of colorectal polyp using deep learning," *BioMed Research International*, vol. 2022, 2022, doi: 10.1155/2022/2805607.

[45] E. Wulczyn *et al.*, "Interpretable survival prediction for colorectal cancer using deep learning," *NPJ Digital Medicine*, vol. 4, no. 1, 2021, doi: 10.1038/s41746-021-00427-2.

[46] F. Mulita, G. I. Verras, C. N. Anagnostopoulos, and K. Kotis, "A smarter health through the internet of surgical things," *Sensors*, vol. 22, no. 12, 2022, doi: 10.3390/s22124577.

[47] D. Bousis *et al.*, "The role of deep learning in diagnosing colorectal cancer," *Przeglad Gastroenterologiczny*, vol. 18, no. 3, pp. 266–273, 2023, doi: 10.5114/PG.2023.129494.

[48] D. D. Chlorogiannis *et al.*, "Tissue classification and diagnosis of colorectal cancer histopathology images using deep learning algorithms. Is the time ripe for clinical practice implementation?," *Gastroenterology Review*, 2023, doi: 10.5114/pg.2023.130337.

BIOGRAPHIES OF AUTHORS

Maria M. Attia received the B.Sc. degree in communication engineering from the Department of Communications and Electronics Engineering, MISR Higher Institute for Engineering and Technology (MET), Mansoura, Egypt, and is currently pursuing the M.Sc. degree in communication engineering from the Department of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University. Major research interests include image processing, medical imaging, neural networks, and machine learning. She can be contacted at email: maria.elnagar@yahoo.com.

Nihal Fayez F. Areed received the B.Sc. degree in electronics and communications from the Faculty of Engineering, Mansoura University, in May 2000, and the M.Sc. and Ph.D. degrees in electrical communications from the Faculty of Engineering, Mansoura University, in 2003 and 2008, respectively. She was appointed as a Professor in September 2018. Microwave and antenna research, optical fibers, and plasmonic models are among her research interests. She is an IEEE Communication Society and OSA member. In 2001, she received the Egyptian Society of Engineers' Scientific Excellence Award. In July 2018, she received the State Incentive Awar from the Academy of Scientific Research. She can be contacted at email: nfayez@zewailcity.edu.eg.

Hanan M. Amer is an assistant professor at Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University. She received her Ph.D. degree from Mansoura University, Egypt, in 2018. Her field of experience includes signal and image processing, medical image analysis, and deep learning applications in the medical field. She can be contacted at email: hanan.amer@yahoo.com.

Mervat El-Seddek IEEE senior member, associate professor at the Department of Communications and Electronics Engineering at the Mansoura Higher Institute of Engineering and Technology. She received the B.Sc. degree in electronics and communications from the Electronic and Communication Department, Faculty of Engineering, Mansoura University, in 1999, and the M.Sc. and Ph.D. degrees in electrical communications from the Faculty of Engineering, Mansoura University, in 2009 and 2015, respectively. Image processing, medical imaging, and machine learning are among the primary research areas. She can be contacted at email: mervat.elseddek@ieee.org.