Received: 27 March 2021 | Revised: 31 May 2021 | Accepted: 21 June 2021

DOI: 10.1111/exsy.12764

ORIGINAL ARTICLE

Expert Systems WILEY

Optimal deep learning based fusion model for biomedical image classification

Romany F. Mansour[1] | Nada M. Alfar[2] | Sayed Abdel-Khalek[2,3] | Maha Abdelhaq[4] | Rashid A. Saeed[5] | Raed Alsaqour[5]

[1]Department of Mathematics, Faculty of Science, New Valley University, El-Kharga, Egypt

[2]Department of Mathematics, College of Science, Taif University, Taif, Saudi Arabia

[3]Department of Mathematics, Faculty of Science, Sohag University, Sohag, Egypt

[4]Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia

[5]Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia

Correspondence
Romany F. Mansour, Department of Mathematics, Faculty of Science, New Valley University, El-Kharga, Egypt.
Email: romanyf@sci.nvu.edu.eg

Abstract

Automated examination of biomedical signals plays a vital role to diagnose diseases and offers useful data to several applications in the areas of physiology, sports medicine, and human–computer interface. The latest advancements in Artificial Intelligence (AI) have the ability to manage and analyse enormous biomedical datasets resulting in clinical decision making and real time applications. At the same time, Colorectal cancer (CRC) is the third most deadly disease affecting people over the globe. The utilization of AI techniques for the earlier identification of CRC has gained significant interest among the research communities. Therefore, this paper presents a novel AI based fusion model for CRC disease diagnosis and classification, named AIFM-CRC. The presented AIFM-CRC model primarily undergoes Gaussian filtering based noise removal and contrast enhancement as a preprocessing stage. In addition, a fusion based feature extraction process takes place where the SIFT based handcrafted features and Inception v4 based deep features are fused together. Besides, whale optimization algorithm tuned deep support vector machine model is employed as a classification technique to determine the existence of CRC. In order to highlight the proficient results analysis of the AIFM-CRC model, a comprehensive simulation analysis takes place. The resultant experimental values pointed out the betterment of the AIFM-CRC model by accomplishing a maximum accuracy of 96.18%.

KEYWORDS
artificial intelligence, biomedical signals, colorectal cancer, deep learning, fusion model, image processing

1 | INTRODUCTION

Colorectal cancer (CRC) has become the deadliest disease which affects both genders globally and it remains a second major reason for cancer related mortality (Colorectal Cancer Statistics, 2019). In fortunate, the occurrence and death rate of CRC gets gradually reduced in the last 30 years in the United States due to the efficient diagnosis method. In 1993, landmark National Polyp research illustrated that the occurrence of this tumour can be decreased with polypectomy and colonoscopy of adenomatous polyps (Winawer, 1993), that is approved in several investigations. Therefore, several main social rules have suggested colonoscopy as preferable colon cancer (CC) diagnosis approach. Adenoma detection rate (ADR) refers the ratio of diagnosis colonoscopy having minimum of one adenoma—has been a main quality measure in the US for estimating an endoscopists capability to detect adenomas. Advanced ADRs are related to low post colonoscopy CC and deaths (Burt et al., 2013; Corley et al., 2014). Numerous additional methods and gadgets have been investigated for enhancing an endoscopists capability to find adenomas. To raise the effectiveness and reduce entire health care cost, several techniques are needed to design to precisely screening and neglect tiny non-

neoplastic polyps and extract precancerous polyps. This technique obtains important function is computer aided diagnoses (CAD), that is computer enabled image investigation which includes both increasing polyp detection and histopathologic variations with no changes to the colonoscope or original process. Moreover, in contrast to alternative methods (like virtual chromoendoscopy and narrow band imaging), CAD is mostly operator-independent (Chao et al., 2019).

Recently, artificial intelligence (AI) has resulted in the design of deep neural network (DNN) and machine learning (ML) techniques, particularly in the field of computer vision (LeCun et al., 2015). A convolutional neural network (CNN) is the type of DNN which is extremely efficient in carrying out an image and video analysis. A CNN based CAD method for colonoscopy can support endoscopists in the identification polyps and carry out visual diagnoses (Alagappan et al., 2018). These AI based methods have high significance to increase the ADR and decrease the cost of polypectomy for hyperplastic polyps (Mori et al., 2017). For obtaining maximum efficiency, the polyp detection method must contain higher sensitivity for the identification of polyps with false positive rate, whereas it maintains faster computation and finds appropriate for real-world colonoscopy. By the utilization of AI for colonoscopy, automated identification and classification of colorectal polyps have gained more interest over the variation of polyps in either hyperplastic or adenomatous. The previous goal is to identify polyps, regardless of the pathology (either neoplastic nor hyperplastic polyps). Subsequently, it assists visibly to categorize the polyp's detection into pathological classes. In the earlier studies, polyp identification is more investigated by white light imaging, whereas polyp classifier needs recent imaging techniques like magnifying narrow band imaging (Kudo et al., 2019).

1.1 | Conventional and AI based histologic analysis

The large slice of all tumours is chosen to estimate the succeeding histologic features such as ineffectively distinguished clusters, lymphatic invasions, venous invasions, cancer depths, and budding. Venous invasion is ensured via H&E by Victoria blue staining highlighting the elastic fibre (Takamatsu et al., 2019). The deepness of submucosal invasion is determined based on the Japanese Society of the Colon and Rectum criteria. The badly distinguished clustering is determined as tumour cluster consisting of more than five cells and insufficient to glandular structure and considered as positive when the size is approximately 0.24 mm^2 by 40$\times$ objective lens. Cancer budding is estimated at H&E slide and determined based on past researches as a tumour cell or clustering consist of 1–4 cells in the invasive frontal region. The AI assisted histologic estimation of the colorectal region is growing at a faster rate. The AI has started to utilize for predicting the prognoses of T1 CRC persons by traditional histologic estimation information and further medical data. The classical histologic estimation is yet needed, but, and the issue of interobserver argument still exists. In recent times, AI based image investigation of tissue microarrays is proposed and employed to forecast the medical result of CRC persons. Accordingly, ML might helpful to predict the lymph node metastasis (LNM) of T1 CRC by image investigation.

1.2 | Paper contributions

This paper presents a novel AI based fusion model for CRC disease diagnosis and classification, named AIFM-CRC. The presented AIFM-CRC model primarily undergoes Gaussian filtering (GF) based noise removal and contrast enhancement as a preprocessing stage. Moreover, a fusion based feature extraction process is carried out where the SIFT based handcrafted features and Inception v4 based deep features are fused together. Furthermore, whale optimization algorithm (WOA) tuned deep support vector machine (DSVM) model is applied as a classification technique to compute the existence of CRC. For assessing the proficient results analysis of the AIFM-CRC model, a complete simulation analysis is performed.

This paper is structured as follows. Section 2 briefs the background information and existing works related to the study. Then, section 3 discusses the proposed model and section 4 validates the performance of the proposed model. Finally, section 5 highlights the concluding remarks of the study.

2 | BACKGROUND INFORMATION AND LITERATURE REVIEW

This section is two folded to discuss the basic concepts of AI and state of art works (Abdulsahib et al., 2021; Al-Waisy et al., 2021; Awan et al., 2021; Jasim Hussein et al., 2021; Mohammed et al., 2021). Firstly, the recent technological developments made in AI are identified and discussed. Secondly, the existing CRC diagnosis and classification models available in the literature are reviewed.

2.1 | Technological developments in artificial intelligence

Data driven techniques depending upon ML received main attention for advanced growths in AI. Assuming the image classifiers as a demonstrative instance, to create ML methods, labelled trained information (i.e., image marked up by respective class names) are gathered and utilized for

model optimization via estimating efficiency from trained information. The previous techniques typically separated the model into feature extraction and classification, where feature extraction is a vital feature of the image and next categorizes an image based on feature extraction. The ML techniques perform learning processes however, the extracted feature portion frequently needs human activities for engineering (Lowe, 2004). In recent times, DL, that includes designing two (i.e., classifier portions and feature extraction) combined with DNN and studying the whole method by ML technique, have gained major interest because of its effective enhancement in several applications. The DL surpasses human engineering via learning robust features straightaway from trained information. In case of image and video domains, CNN is one of the conventional types of DNN (LeCun & Bengio, 1995).

Over-fitting arises if learning machine is excessively adjusted to the training data. Two general approaches to avoid over-fitting are to gather additional training data and constraints the complexity of the model (Abu-Mostafa et al., 2012). By the use of additional training data, the learning method is able to gather additional generalized data, and restricting the difficulties of the method from excessively spending the capacity for fitting the training data. Preferably, the collection of additional information is difficult, expensive, and assets intense. Contrastingly, restricting the complexity is highly effective however it might occasionally result in "underfitting," an occurrence in which a machine does not operate better over trained information. Selecting the proper approach is therefore based on the ML problem, typically computed as (i.e., cross) validation to split the dataset into training and testing parts. By the use of AI technologies, particularly DL methods (i.e., NN) are typically very difficult and requires more data, it should capture generalization and over-fitting to account in the learning rate where the aim is to carry out CAD in colonoscopies on original persons (i.e., invisible in training) with exclusive endoscopic features.

2.2 | Existing works on biomedical imaging based disease diagnosis

In current growths in IT, data classification techniques denote a significant study domain and have become helpful methods to assist medical diagnosis principles (Ting et al., 2020). The ML is a kind of data mining technique and is employed to investigate significant data hidden in the huge quantity of information saved in medicinal records. Several distinct types of ML techniques are utilized for designing prediction methods to classify or predict CRC tumours. For instance, Tseng et al. (2014) utilized SVM and extreme learning machines (ELM) to forecast the recurrence proneness of cervical tumours. Tseng et al. (2017) used ELM, SVM, Random forest (RF), and multivariate adaptive regression splines (MARS) to recognize factors of risk and diagnosis ovarian tumour repetition. Ting et al. (2018) employed MARS, RF, ELM, and SVM techniques to identify repetition in persons diagnose by CC. Chang and Chen (2019) developed a classifier method utilizing XGBoost as classification for forecasting second major tumours in females with breast tumours. The XGBoost, SVM, RF, ELM, and MARS techniques have attained better enhancements for building efficient prediction methods of tumours. Kopetz et al. (2015) proposed ColoPrint predictive technique for enhancing prognosis accuracy autonomous of microsatellite state. Gao et al. Gao et al. (2016) proposed tumour hallmark-based genetics signature sets (i.e., CSS sets) for prognosis prediction and facilitate the detection of persons with phase II CC.

Automated identification of CRC by ML techniques has received attention among several scientists. Numerous investigations have utilized conventional ML techniques to categorize CRC based microbiome instances (Wirbel et al., 2019). This technique contains numerous restrictions like lower accuracy and requires manual FS. The feature engineering is widely utilized to enhance the classifier efficiency of ML technique on series data. The study in Topçuoğlu et al. (2020) noticed that the efficiency of DNN techniques to classify microbiome information is restricted if the instance size is small. Though additional studies have displayed that utilizing DNN methods can enhance the classifier efficiency of CRC depending upon microbiome instances. Especially, data augmentation finds useful to resolve the curse of dimensionality that is predominant in series information (Simidjievski et al., 2019). Researches which integrate feature engineering and data augmentation needs to be developed for CRC identification (Mulenga et al., 2021).

3 | THE PROPOSED CRC DIAGNOSIS MODEL

The working principle involved in the presented AIFM-CRC model is demonstrated in Figure 1. The figure showcases that the input image is primarily preprocessed in two stages namely GF based noise removal and contrast enhancement. Followed by, the fusion of feature extraction process takes place through SIFT and Inception v4 model. Finally, WOA-DSVM method is applied for determining the class label of the input biomedical images. The WOA is applied to tune the parameters of the DSVM model.

3.1 | Data collection

The experimental validation of the AIFM-CRC model takes place using colorectal gland images collected from the Warwick-QU dataset (www.warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/download). The dataset holds a set of 165 images including 74 benign and 91 malignant tumour images respectively. It is obtained with the Zeiss MIRAX MIDI Scanner by employing an image data weight range of 716 kilobytes, 1.187

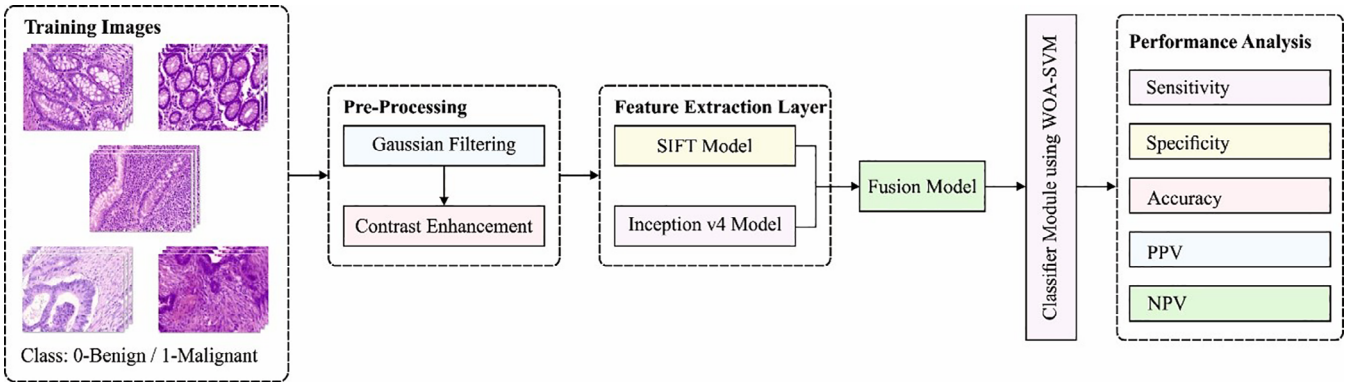

FIGURE 1 Working process of AIFM-CRC model

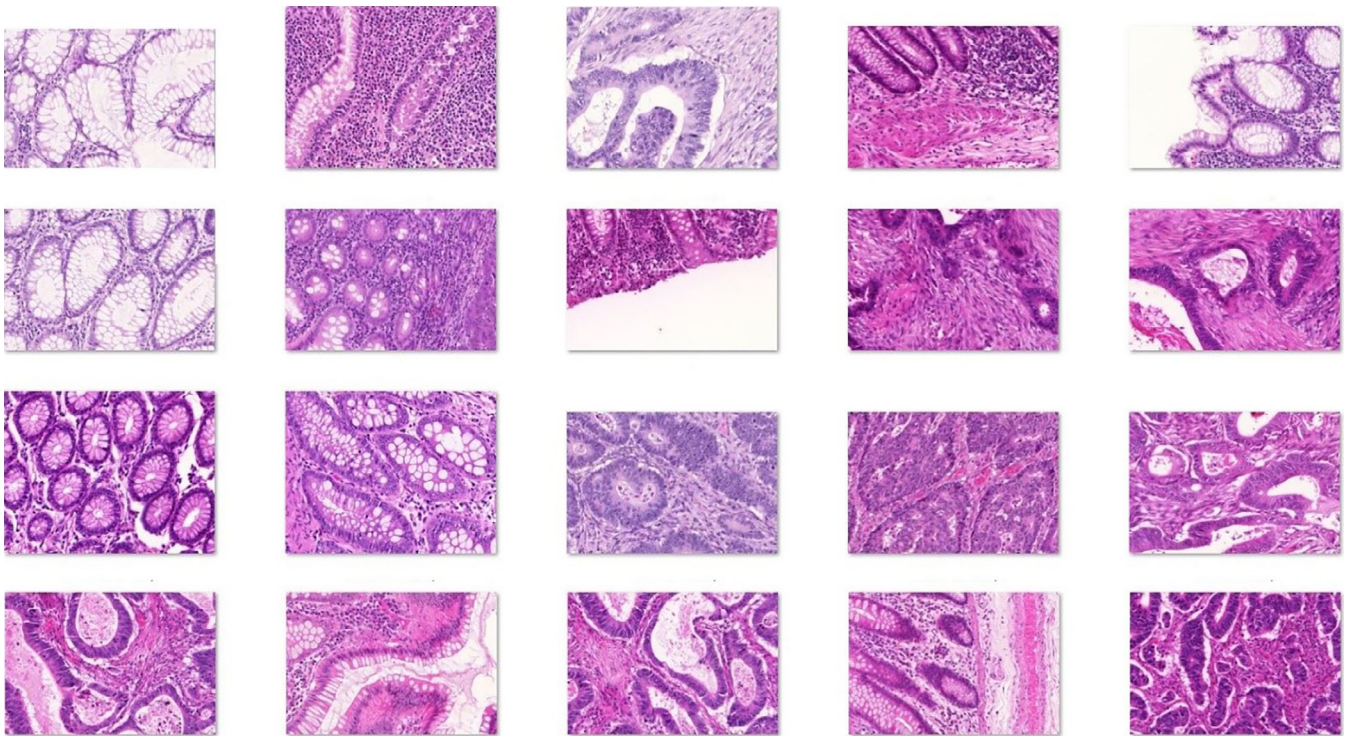

FIGURE 2 Sample images

kilobytes, and an image data resolution range of 567 × 430 pixels to 775 × 522 pixels with every individual pixel has a distance of 0.6 μm from the original distance. Figure 2 shows the sample images.

3.2 | Image preprocessing

At the beginning stage, the proposed AIFM-CRC model performs preprocessing to remove the noise using GF technique and improve the contrast using CLAHE model.

3.2.1 | Noise removal process

The design of two-dimensional GF is broadly utilized for noise removal and smoothing. The convolution operator is Gaussian operator and the concept is attained with convolution. The Gaussian operators in one dimension can be represented by:

$$G_{ID}(x) = \frac{1}{\sqrt{2\pi}o} e^{-\left(\frac{x^2}{2o^2}\right)}. \quad (1)$$

The optimum smoothing filter of image is located in the frequency as well as spatial fields, thus, fulfilling the unpredictability relationship as Nandan et al. (2018):

$$\Delta x \Delta \omega \geq \frac{1}{2}. \quad (2)$$

The Gaussian operator in two dimensions (i.e., circularly symmetric) can be represented by:

$$G_{2D}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)}, \quad (3)$$

Let σ (Sigma) represent the SD of Gaussian function. It has a huge value and the image smoothing results are high. Here, (x,y) represent the Cartesian coordinates of image that display the window dimensions. These filters consist of multiplication and addition processes among image and kernel, in which the image is denoted as matrix via value from 0 to 255 (i.e., 8 bits). The kernel in standardized square matrix value lies in the range of 0 and 1. The kernel is denoted as number of bits. In convolution procedure, the result in every bit of the kernel and component of image is later divided by power 2.

3.2.2 | Contrast enhancement process

CLAHE is very effective in bio medicinal image investigation and is mainly utilized to enhance the contrast level of image. It needs two input parameters such as clip limit and dimensional of sub windows. It recognizes grid size of window and value is attained from a topmost leftward corner part of the image in which calculation starts from earlier index of windows. Next, the size of region w^2 and clip limits are estimated from general values of an image. In case of every grid point, the H[I] nearby images are defined. Next, the histogram where existing over the level of estimated clip limits are clipped and cumulative distribution function (CDF) is calculated. Then, for all pixels, the 4 nearby adjusting grid points are attained. By employing the intensity rate, pixel index, and map functions of 4 grid points, the CDF measurement is evaluated. Next, the contextual pixel N and recently established pixel are interpolated. A similar procedure is accompanied by transferring the window to image. However, this procedure is ended and it yet achieves the last index of window (Shankar et al., 2021). This method does not store histogram that exceeds clipping limits, but it could rearrange in an equivalent way for every histogram bins. Thus, the presented technique helps to enhance the segmentation outcome of the applied test image.

3.3 | Fusion based feature extraction process

The preprocessed image is fed into the feature extractor where the fusion of handcrafted and deep features was taken place. The handcrafted features are derived by the SIFT technique and the deep features are generated by the Inception v4 model. The fusion process takes place in such a way that the classification performance can be increased.

3.3.1 | SIFT based handcrafted features

This technique is to detect and extract local feature descriptors that are unvarying to image scaling, rotation, and brightness. The SIFT model has numerous benefits as given here. It is non-variance to orientation, uniform scaling, and partial unvarying to brightness variations. It has optimum error toleration by some matching and better performance interms of speed. It is comfortable to integrate and produce helpful data. The detection phase of SIFT features is separated into four stages:

1. Scale space extrema identification: In the initial phase, the image $I(x,y)$ is convoluted by the GF at distinct scales in Equation (3) as given by:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (4)$$

where, $L(x,y,\sigma)$ denotes convolution of image $I(x,y)$ by GFs $G(x,y,\sigma)$ on scale σ. Variances among w Gaussian image on scale is $k\sigma$ and σ made in Equation (5) can be demonstrated as:

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \tag{5}$$

The variance on these 2 scales is known as a DoG (Difference of Gaussian).

2. Keypoint localization: In this phase, the key point also called Interest points are recognized as local maximum or minimum of the DoG image over scale. Every pixel in the DoG image is related to eight adjacent pixels by a similar scale. It is required to perform accurate key point' localization with removal point via predefined values as given by.

$$\frac{D(\widehat{x}) = D + \frac{1}{2}\partial D^T}{\partial x \widehat{x}} \tag{6}$$

where $\widehat{x}$ is estimated by setting up the derivation $D(x,y,\sigma)$ to 0.

3. Orientation assignment: To attain invariant to orientation $\theta\,(x,y)$ and gradient magnitude, the $m(x,y)$ is recalculated by Equation (7) and (8):

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \tag{7}$$

$$\theta(x,y) = arctan\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \tag{8}$$

4. Key point descriptors creation: When a key point orientation is chosen, the feature descriptor is calculated as a class of orientation histogram at 4×4 pixel adjacencies. Histogram comprises eight bins, thus SIFT feature vector is defined by $4 \times 4 \times 8 = 128$ components and standardized to distance unit. It is the characteristic of SIFT feature descriptor.

3.3.2 | Inception v4 based deep features

This technique varies from convolutional procedure in the investigation of max-pooling layer and manner of grids get trained. The network consisting of five layers of weight like F3, M2, Cl, input, and output layers. Now, θ represent entire trained feature (i.e., weights measure), $\theta = \{\theta_l\}$ and $l = 1,2,3,4$, in which θ_l denotes variable set among $(l-l)$ th and lth layer. In CRC, all pixel instances are assumed as a two-dimension image that has an equivalent height of 1. Henceforth, the size of input layer of n_1 and $(n_1, 1)$ indicate band count. The primary hidden conv layer Cl extract $n_1 \times 1$ input information involving 20 kernels of size $k_1 \times 1$. Subsequently, the layer Cl that has $20 \times n_2 \times 1$ nodes, whereas $n_2 = n_1 - k_1 + 1$. It consists of $20 \times (k_1 + 1)$ trained variables among Cl and input layer. The M2 is considered as the kernel size is $(k_2, 1)$ and 2nd hidden layer.

The M2 layer contains $20 \times n_3 \times 1$ node, and $n_3 = \frac{n_2}{k_2}$. The whole connected layer F3 comprises a set of n_4 nodes while $(20 \times n_3 + 1) \times n_4$ displays readable variables from present to M2 layer. Later, the output layer comprises n_5 node, and $(n_4 + 1) \times n_5$ trained variables from the existing layer F3. Therefore, the structure of proposed CNN classifier comprises group of $20 \times (k_1 + 1) + (20 \times n_3 + 1) \times n_4 + (n_4 + 1) \times n_5$ training variables. Figure 3 demonstrates the structure of Inception v4 model (Szegedy et al., 2017).

The classifier of a specific CRC pixel requires a related CNN involving an abovementioned variable, in which n_5 and n_1 indicate spectral channel size and total extremal class of datasets, respectively. Nearly, k_1 may be $[n_1/9]$, in which $n_2 = n_1 - k_1 + 1$. Here, n_3 indicates values between 30 and 40 and $k_2 = [n_2/n_3]$.n_4 stay constant to hundred. Henceforth, the choices are optimum and effective for general CRC data. In this structure, it is obviously noted that the layer Cl and M2 is noted as trained feature extraction for input CRC information, whereas layer F3 is considered as a trained classifier for feature extraction. Here, the CNN based Inception v4 methods are employed as feature extraction (Shankar et al., 2021). The concluding result of sub sample is the actual feature of original data.

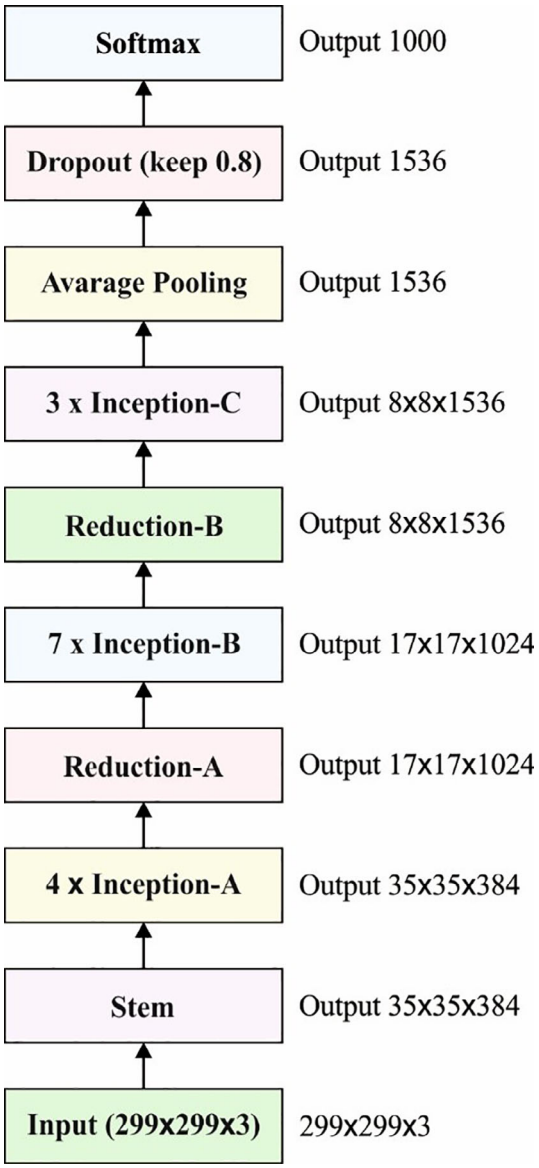

FIGURE 3 Structure of inception v4

3.4 | Image classification

When the fusion methods have extracted a useful set of feature vectors, the DSVM model is applied to carry out the classification process (Le et al., 2021). This technique categorizes a binary problem using linear hyperplane via assuming trained set by *n*-training instances, such as $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$, where $x_i \in \mathfrak{R}^N$ denotes *N* dimensional vector belonging to each class of $y_i \in \{-1,+1\}$. The binary classifier problem gets separated by linear decision functions which are given by,

$$f(x) = w \cdot x + b \tag{9}$$

where $w \in \mathfrak{R}^N$ indicates vector, determining the orientation of the chosen hyperplane have to separate, and $b\mathfrak{R}$ called as "bias." An optimum hyperplane is vital for partitioning two objects which are represented by,

$$y_i(w \cdot x + b) \geq 1 \tag{10}$$

The resolution to this problem is via resolving the limited optimize problem (either primal problem) as defined by,

$$minimise\frac{1}{2}w\cdot w+C\sum_{i=1}^{n}\xi_i \tag{11}$$

subjected to: $y_i(w\cdot x+b)\geq 1-\xi_i,\xi_i>0$, and $\forall i=1,n$, where $C,0<C<\infty$, called as penalty values/ regulation variable; however,ξ_i denotes slack parameter. In nonlinear state, the optimization problem is defined as,

$$maximize\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i,x_j) \tag{12}$$

subjected to: $\sum_{i=1}^{n}\alpha_i y_i=0$, and, $0\leq\alpha_i\leq C$, for $i=1,\ldots,n$. Since output decision function is given by,

$$f(x)=sign\left[\sum_{i=1}^{n}y_i\alpha_i^0K(x_i,x)+b^0\right] \tag{13}$$

α_i^0 denotes support vector as $K(x_i,x)$ implies kernel function/kernel trick. The DSVM is expressed using multiple-layer structure which has many hidden layers.

$X_1,X_2,\ldots,X_n$ indicates input layer data point. The multi hidden layer contains $SVM_{11},SVM_{12},\ldots,SVM_{1k},SVM_{21},SVM_{22},\ldots,SVM_{2k}$ and $SVM_{n1},SVM_{n2},\ldots,SVM_{nk}$ but $F_1(X),F_2(X),\ldots,F_n(X)$ denotes resulting layer point. To X_1, the outcome is trained $SVM_{11},SVM_{12},\ldots,SVM_{1k}$ as $F_1(X)$. To X_2, the result of training $SVM_{21},SVM_{22},\ldots,SVM_{2k}$ is $F_2(X)$. To X_n, the result of training $SVM_{n1},SVM_{n2},\ldots,SVM_{nk}$ as $F_n(X)$. The network weight is demonstrated as $f(x)$. All $f(x)$ are calculated in the hidden layer by multi-layer connecting all input neurons with final neuron. The net input on hidden layer neurons is expressed as,

$$\begin{aligned}
net_{h1}&=f_{11_1}(x)\cdot X_1+f_{11_2}(x)\cdot X_2+\cdots+f_{11_n}(x)\cdot X_n+b1\\
net_{h2}&=f_{12_1}(x)\cdot X_1+f_{12_2}(x)\cdot X_2+\cdots+f_{11_n}(x)\cdot X_n+\mathrm{b1}\\
&\vdots\\
net_{hn}&=f_{1n_1}(x)\cdot X_1+f_{1n_2}\cdot X_2+\cdots+f_{11_n}(x)\cdot X_n+b1
\end{aligned} \tag{14}$$

The logistic activation function is used to compute the outcome for every input neuron is expressed as,

$$\begin{aligned}
out_{h1}&=\frac{1}{1+e^{-net_{h1}}}\\
out_{h2}&=\frac{1}{1+e^{-net_{h2}}}\\
&\vdots\\
out_{hn}&=\frac{1}{1+e^{-net_{hn}}}
\end{aligned} \tag{15}$$

The outcome of hidden layer neuron is used as input to compute resulting layer neuron $net_{o1_1}\ldots,net_{o1_n},net_{o2_1}\ldots,net_{o2_n}$, and $net_{on_1}\ldots,net_{on_n}$ is expressed as

$$\begin{aligned}
net_{o1_1}&=f_{21_1}(x).out_{h1}+f_{21_2}(x).out_{h2}+\cdots+f_{21_n}(x).out_{hn}+b_2\\
&\vdots\\
net_{o1_n}&=f_{k1_1}(x).out_{h1}+f_{k1_2}(x).out_{h2}+\cdots+f_{k1_n}(x).out_{hn}+b_2
\end{aligned} \tag{16}$$

$$\begin{aligned}
net_{o2_1}&=f_{22_1}(x).out_{h1}+f_{22_2}(x).out_{h2}+\cdots+f_{22_n}(x)\cdot out_{hn}+b_2\\
&\vdots\\
net_{o2_n}&=f_{k2_1}(x).out_{h1}+f_{k2_2}(x).out_{h2}+\cdots+f_{k2_n}(x).out_{hn}+b_2
\end{aligned} \tag{17}$$

$$\begin{aligned}
net_{on_1}&=f_{2n_1}(x).out_{h1}+f_{2n_2}(x).out_{h2}+\cdots+f_{2n_n}(x).out_{hn}+b_k\\
&\vdots\\
net_{on_n}&=f_{kn_1}(x).out_{h1}+f_{kn_2}(x).out_{h2}+\cdots+f_{kn_n}(x)\cdot out_{hn}+b_k
\end{aligned} \tag{18}$$

Consider the instance of $net_{o1_1}\ldots,net_{o1_n}$. The outcome is estimated by logistic activation function is given by,

$$\begin{aligned} out_{o1_1} &= \frac{1}{1+e^{-net_{o1_1}}} \\ &\vdots \\ out_{o1_n} &= \frac{1}{1+e^{-net_{o1_n}}} \end{aligned} \tag{19}$$

The error compute output $output_{o1}$ to X_1, is calculated by subtracting computed output $output_{o1}$ from the identified value for $F_1(X)$ is given by,

$$E_{01} = \sum_{i=1}^{n} \frac{1}{2} (F_1(X) - output_{o1_i}) \tag{20}$$

Similarly, the method is estimated by summing all computed error $E_{o1}, E_{o2}, \ldots, E_{on}$ is represented by,

$$E_{total} = E_{o1} + E_{o2} + \cdots + E_{on} \tag{21}$$

Below the application of BP, every $f(x)$ in network guarantees that original output is maximal to the targeted output $F(X)$, so the error decrease from every resultant neuron and entire network. For instance, $f_{11_1}(x)$ is estimated as gradient to ∂E_{total} is given by,

$$\frac{\partial E_{total}}{f_{11_1}(x)} = \frac{\partial net_{o1_1}}{\partial f_{11_1}(x)} * \frac{\partial out_{o1_1}}{\partial net_{o1_1}} * \frac{\partial E_{total}}{\partial out_{o1_1}} \tag{22}$$

The upgraded function $f_{1l}^{(new)}(x)$ is estimated by

$$\frac{f_{11_1}^{(new)}(x) = f_{11_1}(x) - \lambda * \partial E_{total}}{\partial f_{11_1}(x)} \tag{23}$$

where λ denotes learning rate to adapted the weight of network. In same technique, all weights in the network to be precise $f(x)$ is upgraded, and the method is repetitively iterative from Equation (14) until E_{total} converts 0 / ∞.

3.5 | Hyperparameter optimization of DSVM model

For the optimal tuning of hyperparameters in the DSVM model, the WOA is applied to it. At the first stage, the initialization process takes place. During the encircling prey process, the humpback whales have noticed the position of prey and surround them. For unclear places in the searching area, the present optimal solution is considered as the prey. When the optimal searching agent is defined, the other searching agents refresh the situation in the direction of optimal searching agent.

$$\vec{U} = \left| S.\vec{K}^{*}(t) - \vec{K}(t) \right| \tag{24}$$

$$\vec{K}(t+1) = .\vec{K}^{*}(t_{best}) - Y^{*}\vec{U}.$$

where $S = 2.r$ and $Y = 2.I.r - I$. The newly obtained solution for optimal fitness will consolidate the highlights which comprise fewer parameter dependencies. It is not needed to define the initial set of parameters and step sizes for an ideal solution. Based on the functional fitness value over iterations, the coefficient vector $\}y\}$ can be achieved by receiving the adaptable probability function

$$y \Rightarrow Probability = \begin{cases} C_1(f_{max} - f_x)/(f_{max} - F_{avg}), & f_x \geq F_{avg} \\ C_3, & f_x \leq f_{avg}. \end{cases} \tag{25}$$

where, $f_{min}$and f_{max} are the minimum and maximum values of the fitness function, whereas C_1 and C_3 ranges in the interval of 0 and 1. The location in the direction of ideal solution is dynamically changed by the functional fitness. Next, to define the bubble-net nature of the humpback whales, two improved schemes are utilized. The bubble-net mechanism takes place by the use of exploitation and exploration phases.

In spiral form is utilized among the position of whale as well as prey for mimicking the helix-framed progress of humpback whales that is referred as:

$$\vec{K}(t+1) = e^{bt}.\cos(2\pi.y).\vec{U}^{'} + \vec{K}^{*}(t). \tag{26}$$

It can be noticeable humpback whales swim approximately the prey inside the contracting circle and beside the winding moulded approach. In order to illustrate this synchronous performance, the probability of 50% is predictable to opt among the contracting enclosed and spiral system for refreshing the condition of whales amid optimize. The mathematical process is:

$$\vec{K}(t+1) = \begin{cases} \vec{K}(t) - y..\vec{U} & if p < 0.5 \\ \vec{U}^{'}.e^{bs}.\cos(2\pi s) + \vec{K}(t) & if\ p \geq 0.5, \end{cases} \tag{27}$$

where y →random value among −1 to 1, show there synchronous performance that is allowed which there are probabilities of half to pick among together the contract surrounding method and twisted method for reviving the condition of whales in midst of optimizing.

In comparative method, with respect to assortment of the vector, is employed for searching to prey (exploration). Everything to be considered, the humpback whales appear heedlessly as depicted by position of together. Consecutively, to take entire analyser, the request administrator is invigorating by aimlessly-picked search specialist rather than an optimal search operator

$$\vec{U} = \left|\vec{S}.\vec{K}_{rand} - \vec{K}\right| \tag{28}$$

$$K(t+1) = \vec{K}_{rand} - \vec{Y}.\vec{U}. \tag{29}$$

Therefore, the random values are utilized predominantly than 1 or under −1 to make search agent for moving distant in reference whale.

4 | PERFORMANCE VALIDATION

This section assess the CRC diagnostic performance of the proposed AIFM-CRC model on the applied dataset. The proposed model is simulated using Python 3.6.5 tool. The parameter setting of the AIFM-CRC model is given as follows: batch size: 128, learning rate: 0.001, epoch count: 500, and momentum: 0.2.

Table 1 and Figures 4 and 5 investigates the classification results analysis of the AIFM-CRC model under varying validation processes. The obtained values denoted that the AIFM-CRC model performance is increased with an increase in training dataset size. For instance, on the applied 50:50 validation, the AIFM-CRC model has obtained a sensitivity of 92.90%, specificity of 94.20%, accuracy of 94%, positive predicted value (PPV) of 94.60%, and negative predicted value (NPV) of 95.40%. At the same time, on the applied 60:40 validation, the AIFM-CRC method has attained a sensitivity of 92.20%, specificity of 96.30%, accuracy of 95.10%, PPV of 95.30%, and NPV of 96.30%. In line with, the applied 70:30 validation, the AIFM-CRC approach has achieved a sensitivity of 91.90%, specificity of 96.50%, accuracy of 95.80%, PPV of 95.40%, and NPV of 96.60%. Accordingly, on the applied 80:20 validation, the AIFM-CRC manner has reached a sensitivity of 93.40%, specificity of 97.30%, accuracy of 96.50%, PPV of 96.40%, and NPV of 97.10%. Meanwhile, on the applied 90:10 validation, the AIFM-CRC model has achieved a sensitivity of

TABLE 1 Result analysis of proposed AIFM-CRC method on various validation models

Training/testing	Sensitivity	Specificity	Accuracy	PPV	NPV
50:50	92.90	94.20	94.00	94.60	95.40
60:40	92.20	96.30	95.10	95.30	96.30
70:30	91.90	96.50	95.80	95.40	96.60
80:20	93.40	97.30	96.50	96.40	97.10
90:10	94.70	98.70	97.80	97.40	97.80
10-Folds	96.40	98.90	97.90	97.60	98.30
Average	93.58	96.98	96.18	96.12	96.92

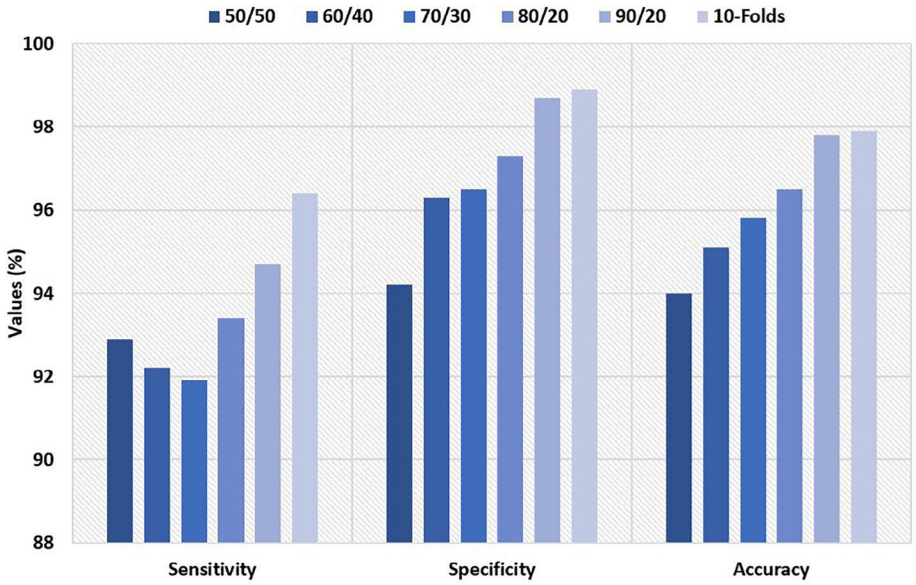

FIGURE 4 Result analysis of AIFM-CRC model with distinct measures

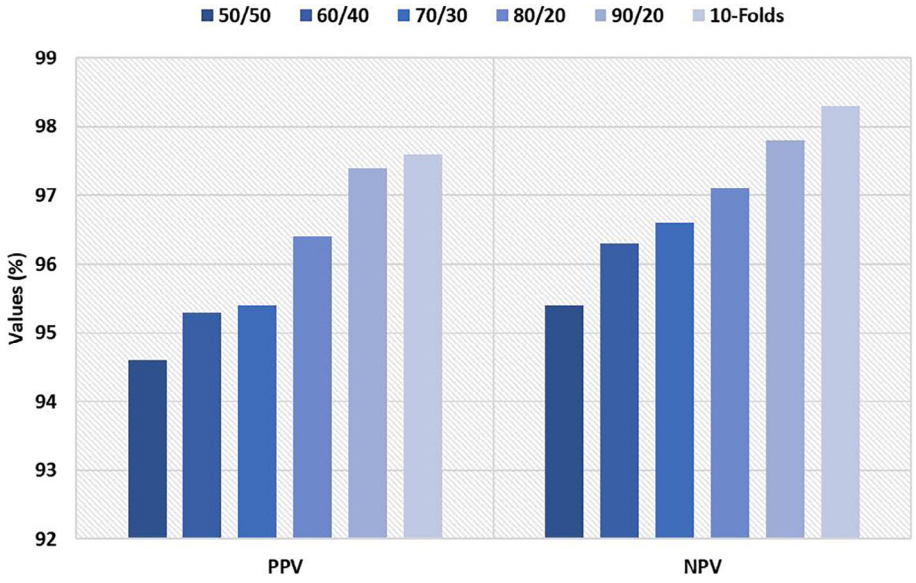

FIGURE 5 PPV and NPV analysis of AIFM-CRC model

94.70%, specificity of 98.70%, accuracy of 97.80%, PPV of 97.40%, and NPV of 97.80%. Eventually, on the applied 10 folds, the AIFM-CRC methodology has obtained a sensitivity of 96.40%, specificity of 98.90%, accuracy of 97.90%, PPV of 97.60%, and NPV of 98.30%. At last, it is noticed that the AIFM-CRC method has attained a maximum average sensitivity of 93.58%, specificity of 96.98%, accuracy of 96.18%, PPV of 96.12%, and NPV of 96.92%.

Table 2 and Figure 6 made a detailed comparative CRC detection results analysis of the AIFM-CRC model (Choi et al., 2020; Sarwinda et al., 2021).

On examining the CRC detection results interms of sensitivity and specificity, the obtained values portrayed that the Endoscopist (novice) model has reached the least sensitivity and specificity of 77.98% and 92.63% respectively. At the same time, the ResNet18 model has achieved slightly increased results over the earlier one with the sensitivity and specificity of 83% and 87% correspondingly. Besides, the ResNet-50 model has showcased closer results to the ResNet18 model with the sensitivity and specificity of 83% and 93% respectively. Followed by, moderate sensitivity and specificity of 85% and 95% have been exhibited by the Endoscopist (expert) whereas a somewhat increased sensitivity and specificity of 86.53% and 96.6% has been accomplished by the Inception-v3 model. Moreover, the ResNet model has depicted manageable outcomes with

TABLE 2 Comparative analysis of proposed AIFM-CRC with existing techniques

Methods	Sensitivity	Specificity	Accuracy	PPV	NPV
Proposed AIFM-CRC	93.58	96.98	96.18	96.12	96.92
Inception-v3	86.53	96.60	89.65	87.39	96.68
ResNet Model	86.64	96.51	89.45	87.10	96.59
DenseNet-161	87.24	96.74	89.95	87.59	96.76
Endoscopist (novice)	77.98	92.63	77.91	78.39	92.76
Endoscopist (expert)	85.00	95.00	85.00	85.74	95.05
ResNet 18	83.00	87.00	85.00	85.42	87.26
ResNet 50	83.00	93.00	88.00	86.79	93.49

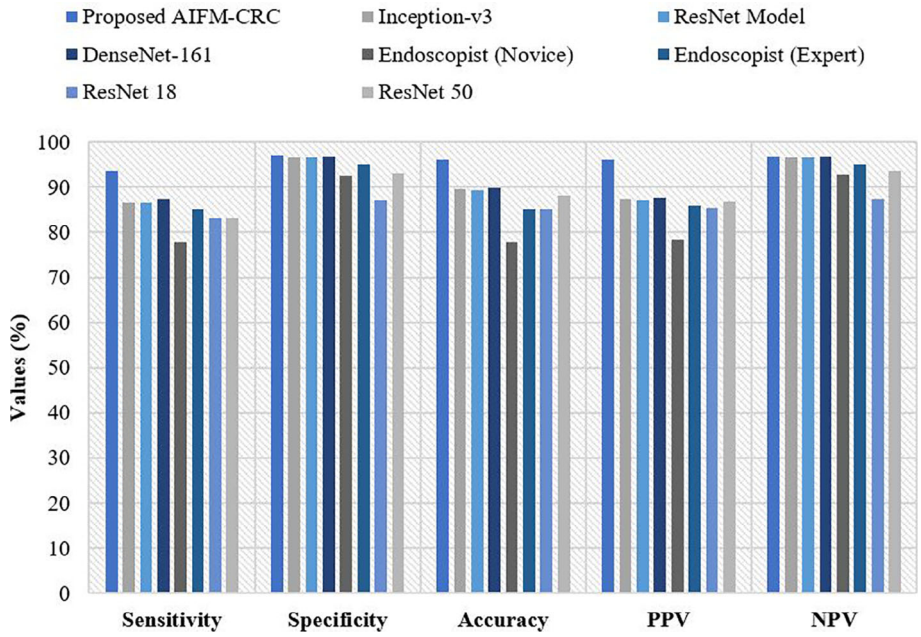

FIGURE 6 Comparative analysis of AIFM-CRC model with existing methods

the sensitivity and specificity of 86.64% and 96.51% correspondingly. Though the DenseNet-161 model has led to a competitive sensitivity and specificity of 87.24% and 96.74%, the AIFM-CRC model has outperformed the other compared methods with the sensitivity and specificity of 93.58% and 96.98% correspondingly.

On investigative the CRC detection outcomes with respect to accuracy, PPV, and NPV, the achieved values exhibited that the Endoscopist (novice) model has achieved to minimum accuracy, PPV, and NPV of 77.91%, 78.39%, and 92.76% correspondingly. Simultaneously, the ResNet18 technique has attained somewhat improved outcomes over the earlier one with the accuracy, PPV, and NPV of 85%, 85.42%, and 87.26% respectively. Besides, the Endoscopist (expert) methodology has illustrated closer outcomes to the ResNet18 model with the accuracy, PPV, and NPV of 85%, 85.74%, and 95.05% correspondingly. On continuing with, a moderate accuracy, PPV, and NPV of 88%, 86.79%, and 93.49% has been depicted by the ResNet50 whereas a somewhat higher accuracy, PPV, and NPV of 89.45%, 87.10%, and 96.59% has been accomplished by the ResNet approach. Also, the Inception-v3 method has showcased manageable outcomes with the accuracy, PPV, and NPV of 89.65%, 87.39%, and 96.68% correspondingly. However, the DenseNet-161 manner has led to a competitive accuracy, PPV, and NPV of 89.95%, 87.59%, and 96.76%, the AIFM-CRC methodology has demonstrated the other compared techniques with the accuracy, PPV, and NPV of 96.18%, 96.12%, and 96.92% correspondingly.

The resulting experimental values pointed out the betterment of the AIFM-CRC model by achieving a maximum accuracy of 96.18%. By looking into the above-mentioned tables and figures, it is evident that the proposed AIFM-CRC model is found to be an effective tool for CRC diagnosis using X-ray images due to the inclusion of GF technique, CLAHE, Inception v4, and optimal parameter tuning using WOA.

5 | CONCLUSION

This paper has presented a novel AIFM-CRC based fusion model for CRC disease diagnosis and classification. At the beginning stage, the proposed AIFM-CRC model performs preprocessing to remove the noise using GF technique and improve the contrast using CLAHE model. The preprocessed image is fed into the feature extractor where the fusion of handcrafted and deep features was taken place. The handcrafted features are derived by the SIFT technique and the deep features are generated by the Inception v4 model. The fusion process takes place in such a way that the classification performance can be increased. When the fusion methods have extracted a useful set of feature vectors, the WOA-DSVM method is applied to carry out the classification process. For the optimal tuning of hyperparameters in the DSVM model, the WOA is applied to it. For assessing the proficient results analysis of the AIFM-CRC model, a complete simulation analysis is performed. The resulting experimental values pointed out the betterment of the AIFM-CRC model by achieving a maximum accuracy of 96.18%. As a part of future extension, the learning rate scheduling technique can be incorporated to the proposed AIFM-CRC model.

ACKNOWLEDGMENT

This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fast-track Research Funding Program.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ETHICS STATEMENT

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

CONSENT TO PARTICIPATE

Not applicable.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Romany F. Mansour https://orcid.org/0000-0001-5857-8495

REFERENCES

Abdulsahib, A. A., Mahmoud, M. A., Mohammed, M. A., Rasheed, H. H., Mostafa, S. A., & Maashi, M. S. (2021). Comprehensive review of retinal blood vessel segmentation and classification techniques: Intelligent solutions for green computing in medical images, current challenges, open issues, and knowledge gaps in fundus medical images. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *10*(1), 1–32.

Abu-Mostafa, Y. S., Magdon-Ismail, M., & Lin, H. T. (2012). *Learning from data* (Vol. 4). AMLBook.

Alagappan, M., Brown, J. R. G., Mori, Y., & Berzin, T. M. (2018). Artificial intelligence in gastrointestinal endoscopy: The future is almost here. *World journal of gastrointestinal endoscopy*, *10*(10), 239.

Al-Waisy, A. S., Mohammed, M. A., Al-Fahdawi, S., Maashi, M. S., Garcia-Zapirain, B., Abdulkareem, K. H., … Le, D. N. (2021). COVID-DeepNet: Hybrid multimodal deep learning system for improving COVID-19 pneumonia detection in chest X-ray images. *Computers, Materials and Continua*, *67*(2), 2409–2429.

Awan, M. J., Rahim, M. S. M., Salim, N., Mohammed, M. A., Garcia-Zapirain, B., & Abdulkareem, K. H. (2021). Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach. *Diagnostics*, *11*(1), 105.

Burt, R. W., Cannon, J. A., David, D. S., Early, D. S., Ford, J. M., Giardiello, F. M., & Freedman-Cass, D. (2013). Colorectal cancer screening. *Journal of the National Comprehensive Cancer Network*, *11*(12), 1538–1575.

Chang, C. C., & Chen, S. H. (2019). Developing a novel machine learning-based classification scheme for predicting SPCs in breast cancer survivors. *Frontiers in Genetics*, *10*, 848.

Chao, W. L., Manickavasagan, H., & Krishna, S. G. (2019). Application of artificial intelligence in the detection and differentiation of colon polyps: A technical review for physicians. *Diagnostics*, *9*(3), 99.

Choi, K., Choi, S. J., & Kim, E. S. (2020). Computer-aided Diagonosis for colorectal cancer using deep learning with visual explanations. In *2020 42nd annual international conference of the IEEE engineering in Medicine & Biology Society (EMBC)* (pp. 1156–1159). IEEE.

Corley, D. A., Jensen, C. D., Marks, A. R., Zhao, W. K., Lee, J. K., Doubeni, C. A., Zauber, A. G., de Boer, J., Fireman, B. H., Schottinger, J. E., Quinn, V. P., Ghai, N. R., Levin, T. R., & Quesenberry, C. P. (2014). Adenoma detection rate and risk of colorectal cancer and death. *New England Journal of Medicine*, *370*(14), 1298–1306.

Division of Cancer Prevention and Control, Colorectal Cancer Statistics. 2019. https://www.cdc.gov/cancer/colorectal/statistics/

Gao, S., Tibiche, C., Zou, J., Zaman, N., Trifiro, M., O'Connor-McCourt, M., & Wang, E. (2016). Identification and construction of combinatory cancer hallmark–based gene signature sets to predict recurrence and chemotherapy benefit in stage II colorectal cancer. *JAMA Oncology*, *2*(1), 37–45.

Jasim Hussein, I., Burhanuddin, M. A., Abed Mohammed, M., Elhoseny, M., Garcia-Zapirain, B., Suliman Maashi, M., & Maashi, S. (2021). Fully automatic segmentation of gynaecological abnormality using a new viola-jones model. *Computers, Materials & Continua*, *66*(3), 3161–3182.

Kopetz, S., Tabernero, J., Rosenberg, R., Jiang, Z. Q., Moreno, V., Bachleitner-Hofmann, T., Lanza, G., Stork-Sloots, L., Maru, D., Simon, I., Capellà, G., & Salazar, R. (2015). Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors. *The Oncologist*, *20*(2), 127.

Kudo, S. E., Mori, Y., Misawa, M., Takeda, K., Kudo, T., Itoh, H., Oda, M., & Mori, K. (2019). Artificial intelligence and colonoscopy: Current status and future perspectives. *Digestive Endoscopy*, *31*(4), 363–371.

Le, D. N., Parvathy, V. S., Gupta, D., Khanna, A., Rodrigues, J. J., & Shankar, K. (2021). IoT enabled depthwise separable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification. *International Journal of Machine Learning and Cybernetics*, 1–14. https://doi.org/10.1007/s13042-020-01248-7

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, *3361*(10), 1995.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, *60*(2), 91–110.

Mohammed, M. A., Elhoseny, M., Abdulkareem, K. H., Mostafa, S. A., & Maashi, M. S. (2021). A multi-agent feature selection and hybrid classification model for Parkinson's disease diagnosis. *ACM Transactions on Multimidia Computing Communications and Applications*, *17*(2s), 1–22.

Mori, Y., Kudo, S. E., Berzin, T. M., Misawa, M., & Takeda, K. (2017). Computer-aided diagnosis for colonoscopy. *Endoscopy*, *49*(8), 813.

Mulenga, M., Kareem, S. A., Sabri, A. Q. M., Seera, M., Govind, S., Samudi, C., & Mohamad, S. B. (2021). Feature extension of gut microbiome data for deep neural network-based colorectal cancer classification. *IEEE Access*, *9*, 23565–23578.

Nandan, D., Kanungo, J., & Mahajan, A. (2018). An error-efficient Gaussian filter for image processing by using the expanded operand decomposition logarithm multiplication. *Journal of ambient intelligence and humanized computing*, 1–8. https://doi.org/10.1007/s12652-018-0933-x

Sarwinda, D., Paradisa, R. H., Bustamam, A., & Anggia, P. (2021). Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Computer Science*, *179*, 423–431.

Shankar, K., Perumal, E., Elhoseny, M., & Nguyen, P. T. (2021). An IoT-cloud based intelligent computer-aided diagnosis of diabetic retinopathy stage classification using deep learning approach. *CMC-Computers Materials & Continua*, *66*(2), 1665–1680.

Simidjievski, N., Bodnar, C., Tariq, I., Scherer, P., Andres Terre, H., Shams, Z., Jamnik, M., & Liò, P. (2019). Variational autoencoders for cancer data integration: Design principles and computational practice. *Frontiers in Genetics*, *10*, 1205.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, no. 1).

Takamatsu, M., Yamamoto, N., Kawachi, H., Chino, A., Saito, S., Ueno, M., … Takeuchi, K. (2019). Prediction of early colorectal cancer metastasis by machine learning using digital slide images. *Computer Methods and Programs in Biomedicine*, *178*, 155–161.

Ting, W. C., Chang, H. R., Chang, C. C., & Lu, C. J. (2020). Developing a novel machine learning-based classification scheme for predicting SPCs in colorectal cancer survivors. *Applied Sciences*, *10*(4), 1355.

Ting, W. C., Lu, Y. C. A., Lu, C. J., Cheewakriangkrai, C., & Chang, C. C. (2018). Recurrence impact of primary site and pathologic stage in patients diagnosed with colorectal cancer. *Journal of Quality*, *25*(3), 166–184.

Topçuoğlu, B. D., Lesniak, N. A., Ruffin, M. T., IV, Wiens, J., & Schloss, P. D. (2020). A framework for effective application of machine learning to microbiome-based classification problems. *MBio*, *11*(3), pp.e00434–20.

Tseng, C. J., Lu, C. J., Chang, C. C., & Chen, G. D. (2014). Application of machine learning to predict the recurrence-proneness for cervical cancer. *Neural Computing and Applications*, *24*(6), 1311–1316.

Tseng, C. J., Lu, C. J., Chang, C. C., Chen, G. D., & Cheewakriangkrai, C. (2017). Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence. *Artificial Intelligence in Medicine*, *78*, 47–54.

Winawer, S. J., Zauber, J., Ho, A. G., O'brien, M. J., Gottlieb, L. S., Sternberg, S. S., & National Polyp Study Workgroup. (1993). Prevention of colorectal cancer by colonoscopic polypectomy. *New England Journal of Medicine*, *329*(27), 1977–1981.

Wirbel, J., Pyl, P. T., Kartal, E., Zych, K., Kashani, A., Milanese, A., Fleck, J. S., Voigt, A. Y., Palleja, A., Ponnudurai, R., Sunagawa, S., Coelho, L. P., Schrotz-King, P., Vogtmann, E., Habermann, N., Niméus, E., Thomas, A. M., Manghi, P., Gandini, S., et al. (2019). Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nature Medicine*, *25*(4), 679–689.

AUTHOR BIOGRAPHIES

Romany F. Mansour, Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt.

Nada M. Alfar, Department of Mathematics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia.

Sayed Abdel-Khalek, Department of Mathematics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia.

Maha Abdelhaq, Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, 84428 Riyadh, Saudi Arabia.

Rashid A. Saeed, Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, 93499 Riyadh, Saudi Arabia.

Raed Alsaqour, Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, 93499 Riyadh, Saudi Arabia.

How to cite this article: Mansour, R. F., Alfar, N. M., Abdel-Khalek, S., Abdelhaq, M., Saeed, R. A., & Alsaqour, R. (2021). Optimal deep learning based fusion model for biomedical image classification. *Expert Systems*, e12764. https://doi.org/10.1111/exsy.12764