Original Article

Increment Learning for Acute Lymphoblastic Leukemia Classification

G. Mercy Bai[1], P. Venkadesh[2]

[1,2]*Department of Computer Science and Engineering, Noorul Islam Centre for Higher Education, Tamil Nadu, India*

[1]*Corresponding Author : mercybaicse@gmail.com*

Abstract - *Acute Leukemia is a dangerous blood disease originating from bone marrow, that is commonly seen both in adults and children. Acute Lymphoblastic Leukemia (ALL) is a prominent one in Leukemia disease as it resists the body to fight pathogens and spread rapidly throughout the body. ALL is identified manually from bone marrow examination and blood smear but, this manual diagnosis of ALL is a slow process and is inaccurate which tends to immediate death. Moreover, various other techniques are also carried out for diagnosis that is costlier and suffer from catastrophic forgetting when learning new classes incrementally. To rectify these issues, an effectual strategy is developed for incremental classification of ALL. Here, the classification is done by a Deep Convolutional Neural Network (Deep CNN) trained by the Fractional Horse Whale Optimization Algorithm (Fractional-HWOA), which is the integration of the Fractional concept into the Horse Herd Optimization Algorithm (HOA) and Whale Optimization Algorithm (WOA) respectively. Various stages included in this paper are pre-processing, segmenting, extracting features, and classifying the image. Here, the Gaussian filter is used for pre-processing and Generative Adversarial Network (GAN) perform the process of segmentation. Alternatively, incremental classification is accomplished using Deep CNN where the network classifier is trained using the proposed Fractional-HWOA. Finally, weights are bounded by the entropy function based on an error condition. This Proposed method is evaluated using various parameters such as accuracy, sensitivity, and specificity and the values attained are 0.961, 0.957, and 0.958, accordingly.*

Keywords - *Increment Learning, Acute Lymphoblastic Leukemia, Whale Optimization Algorithm, Deep Convolutional Neural Network, Horse Herd Optimization.*

1. Introduction

Acute Lymphocytic Leukemia (ALL) or acute lymphoblastic leukemia, is a disease related to blood or hemic. It is the type of cancer in the bone marrow that vigorously spreads to the blood. ALL, thus rapidly spread into the stream of blood and affect vital organs of the human system. It is necessary to treat ALL earlier and if it is not treated properly can cause death. It is most importantly seen in adults aged above fifty and in young children. For children, death is avoided by diagnosing it in earlier stages before spreading it to vital organs. But diagnosis is not much easier as ALLs symptoms like weight loss, fever etc. are common in other diseases that have similar symptoms [9]. The cause for ALL is not known yet, but there is a possibility of disease generated by direct contact with heavy chemicals like benzene and radiation which boosts the disease formation in humans [27]. ALL is classified into many categories. Based on French-American-British (FAB) classification, ALL is categorized into L1, L2, and L3 subtypes [6]. The First L1 type is smaller in size than the other two types. The nuclear shape of L1 is regular and they have homogenous chromatin with tiny nucleoli. The basophilic cytoplasm of L1 is scanty and small. The next type L2 have the irregular shape of nucleoli and are larger in size. They have clefting with variable chromatin and large nucleoli. The third L3 type is medium in size with two or more nucleoli's having prominent cytoplasmic vacuoles. The shape of the nucleus in L3 is round or oval [11].

Diagnosis of ALL is an inspection of peripheral blood flow through the microscope. If cancer is present in the blood, there remain malformations in white cells and inspection of these white cells helps in finding the confirmation of the disease. The identification task of cancer is done by experienced operators, by performing blood analysis based on the classification of cells and counting them. Now, this process is done by cytometers, which identify characteristics and the number of cells. Moreover, other than the physical analysis of the blood smear, morphological analysis of the blood image is also carried out. For this morphological analysis, there is no need for blood samples and hence carried out nowadays by a single image. Analysis based on a single image is very important because the cost of identification is low and is highly accurate [9]. Expression profiling based on image identification of blood smears is not the only robust approach. In modern cancer therapy, Pediatric acute lymphoblastic leukemia is another diagnostic method. The conventional laboratory method for ALL identification is replaced by profiling a single platform of gene expression

in leukemic blasts. 360 pediatric ALL patients of over 12,600 genes are diagnosed with this leukemic blast using oligonucleotide microarrays. These analyses of ALL identify the molecular subtypes very accurately [10]. ALL diagnosis is also done by utilizing chromosome-based tests, like fluorescent in situ hybridization, polymerase chain reaction, and cytogenetics where chromosomes are detected to know unusual blood cells. Moreover, an automated Computer aided Prognosis (CAP) tool which is an image-based method also diagnoses ALL in the early stages as this increase the survival rate to 90% [27].

Neural network-based methods have various applications, for instance in medical prognosis as they have great learning and simplification skills [15]. Machine learning analytic techniques is used especially in cancer research that identifies various risk factors of multiple clinical and linked parameters with the intention to decrease multidimensional facts retrieved from various sources and permit proper prediction [16]. Increment Learning is also one of the Machine learning techniques that provide auspicious attainments on lesion detection and segmentation [3]. Generally, the identification of disease involves dual stages of quantification which evaluate texture, histogram, and so on, along machine learning mechanism for classification [14]. Increment learning comprises incremental classification and incremental segmentation, where both are different from each other. In an image, old classes along with new classes are present with a high range of probability. But these newly added classes of the image in this incremental segmentation may lead to impractical annotations in which only old classes are clearly identified [3]. Catastrophic forgetting is always a big problem in Increment learning systems. Catastrophic forgetting is an inability to define the model for classification before proper knowledge of new task learning. To eradicate these issues, contrastive learning [33] and knowledge distillation [13] related strategies are followed by many researchers. This can boost performance knowledge in classification and resist catastrophic forgetting as they eradicate interdependencies and similarities in structures between various representations of knowledge [1]. Deep learning train machines are also another model for the classification of images, which has a vast number of learn features and label data that are processed for feature extraction. This extraction excludes the manual method of extraction as it is widely combined with CNN for the processing of biological image classification [11].

This paper is focused to develop ALL classification methods using Fractional-HWOA-based Deep CNN connected with Increment learning. The proposed model has various stages like pre-processing, segmentation, extraction of features and ALL classification. The input blood image is fed to the pre-processing stage, where the image is effectually pre-processed by Gaussian Filter for denoising the image. Pre-processed output is further allowed for the segmentation stage, where the image is segmented using GAN. After the segmentation of the image, the features are extracted that include texture,

statistical, and grid features. After that, the extracted features are fed to the classification stage, where Deep CNN trained by Proposed Fractional-HWOA is used to classify into normal or abnormal that developed an error. Conversely, incremental data blood image is employed in preprocessing stage, followed by segmentation and classification stages that generated another error. Both errors are related and if incremental learning's error is more than the proposed method's error, then weight bounding based Renyi entropy is carried out, else no retraining is required.

The major contribution of this paper is enlightened below:

1.1. Proposed Fractional-HWOA-based Deep CNN for ALL Classification

An effective ALL classification method is developed using the proposed Fractional-HWOA-based Deep CNN. Also, GAN is utilized for image segmentation and Deep CNN is used to perform ALL classification by considering extracted features obtained from the segmented image. Moreover, the proposed Fractional-HWOA is used for training Deep CNN to classify ALL.

The remaining paper is followed as below: section 2 enumerates a literature review of multiple existing ALL classification methods and section 3 enhances the proposed Fractional-HWOA based Deep CNN. Section 4 elaborates results and discussions of the proposed method and section 5 includes the conclusion of the paper.

2. Motivation

Classification of ALL diseases is essential for recovery from an early death. Different methods are available for this identification and classification however they are highly expensive and consume more time for a proper prediction. To eradicate these issues of cost, time and prediction, there is a need for a better methodology for ALL classifications. This section presents literature reviews of existing ALL classification methods along with their merits, demerits, and challenges to develop a vigorous ALL classification method that inspires the researchers.

2.1. Literature Review

The literature review of various leukemia classifications is addressed as followed in this section. Sirshar, M., *et al.* [1] presented Incremental Learning Loss Function for screening various pulmonary diseases from chest X-ray (CXR) scans without regarding scanner specifications. This method utilized only a few trainings for the identification of chest pathologies without catastrophically forgetting its earlier attained knowledge. However, this method was not adapted for other disease identification and failed to show its grading severity. To compensate severity in grading, Al-Qudah, R. and Suen, C.Y., [2] proposed Enhanced Incremental Training in which blood smear dataset-based deep neural networks were utilized for detecting morphological abnormalities, classification of fifteen various white blood cells and subtypes of platelets with cross-validation was performed.

This method improved macro-average precision and accuracy of classification but failed to check with other validation approaches namely the K-fold cross-validation system. The drawback in [2] was overcome by He, W., *et al.* [3], who proposed an incremental learning method that focussed on the segmentation of Diabetic Retinopathy (DR) lesions of exudate and hemorrhage on fundus images. This method distilled the previous model's knowledge that improvised the current model thus increasing its efficiency but the drawback remained in choosing the dataset of the previous model. In order to rectify the drawback of [3], Zhao, J., *et al.* [4] proposed a detection method trained by CNN for automatic identification and classification of White Blood Cells (WBC). This method, therefore, helped to find the abnormality based on WBC classification, however, some WBCs were not detected properly from peripheral images. Das, P.K. and Meher, S [5] proposed hybrid deep CNN for efficient ALL detection and classification. Here, a new weight factor was introduced that helped for easy identification of CLL. This method is computationally efficient and performed well in small datasets. However, this method showed poor performance when the dataset is split into testing and training data. Choosing a dataset was overcome in [6] as Rawat, J., *et al.* [6] proposed an automated hybrid hierarchical classification method for 3-class and 2-class. This method helped to discriminate with a high accuracy rate as the classification depend on the cytoplasm and nucleus of the cells. Further, this method was not extended to analyse other disorders like Acute myeloid leukemia (AML) and was not applied to a larger database. To eradicate these issues Srisukkham, W., *et al.* [7] proposed two modified Bare-bones Particle Swarm Optimization (BBPSO) algorithms that characterized healthy and blast cells for significant discrimination of ALL. Also. A logistic chaotic map and cross-domain sonar data set were used for further enhancement of ALL classification, which helped for quick identification. But this method was not tested with many other medical image data sets to further diagnose the problems. These issues were overcome by Li, Y., *et al.* [8] who proposed a dual-threshold method which supported classifying the ALL images with WBC. This method proved to be better than a single-threshold method in the accuracy range and they performed well in Hue, saturation, and value (HSV) and Red, Green and Blue (RGB) images. However, in this method, expert oncologies were utilized and hence this method remains a tedious process.

2.2. Challenges

The prevailing methodologies for the classification of leukemia undergo the following challenges,

➢ The method in [1] screened and recognized multiple lung diseases from images extracted from CXR scans that resisted catastrophic forgetting phenomena. However, this technique was not used for the identification of Corona Virus Disease (COVID)-19 and its severity.

➢ Abnormalities in blood cells, fifteen WBC types and subtypes of platelets are easily and automatically identified as in [4] from synthetic blood smears dataset as in [2] which helped for recognition and prediction of disease type. But the challenge remained in the proper identification of disease for non-medical datasets and large-scale medical datasets.

➢ The method in [3] improvised the present SSmodel by training knowledge of the previous model for DR lesions segmentation very efficiently by conducting experiments. However, middle feature maps were not utilized for the transfer of knowledge from the previous model to the current model.

➢ The drawback in [2] was overcome by an automatic recognition model for WBC as in [4] that detected WBCs directly and classified them without manual operation. But the limitation persisted in the wrong identification of non-WBCs as WBCs that tend to improper recognition.

➢ Acute Lymphoblastic Leukemia is one of the most life-threatening and fatal cancer diseases which is affecting people around the globe. Despite the recent development in the field of cancer detection and prevention, it is still a challenging task to automatically detect and segment Leukemia in the earlier stages.

3. Introduced Fractional-HWOA based Deep CNN for Incremental ALL Classification

ALL is cancer of the body's blood-forming tissues that start first in an immature bone marrow which is a soft tissue. It then spreads throughout the body affecting blood and finally leading to death if untreated. This section elucidates the process of incremental classification of ALL by Deep CNN trained by Fractional-HWOA.

Initially, an input image is taken from the Acute Lymphoblastic Leukemia Image database (ALL-IDB) [32] that is allowed for pre-processing, which is performed using a Gaussian filter [25] for denoising the image. Next to pre-processing, the segmentation is carried out using GAN [17] which divides the image into various segments. Once the segmentation is carried out, the features namely texture, statistical and grid features are extracted. Here, texture features include a Significant Local Binary Pattern (SLBP) [18], Local Optimal Oriented Pattern (LOOP) [24], Local Gabor Binary Pattern (LGBP) [19], and Local directional Ternary Pattern (LDTP) [20]; statistical features includes mean, standard deviation, varience, kurtosis, energy, entropy, and skewness; and grid features includes shape, size, and area. After the extraction of features, the classification process is carried out by the proposed Fractional-HWOA-based Deep CNN [21], which generates an error. Simultaneously, incremental data images are allowed for classification using the proposed model-based Deep CNN that further generates another error. However, the proposed Fractional-HWOA is derived by the integration of the Fractional concept into HOA [22] and WOA [23]. Then, both the errors are

compared and if the error generated from the input image is lesser than the error generated from the incremental data image, then the weights are bound using Renyi-entropy [29], else no retraining is required. Figure 1 shows the schematic view of the introduced HWOA-enabled Deep CNN technique for incremental ALL classification.
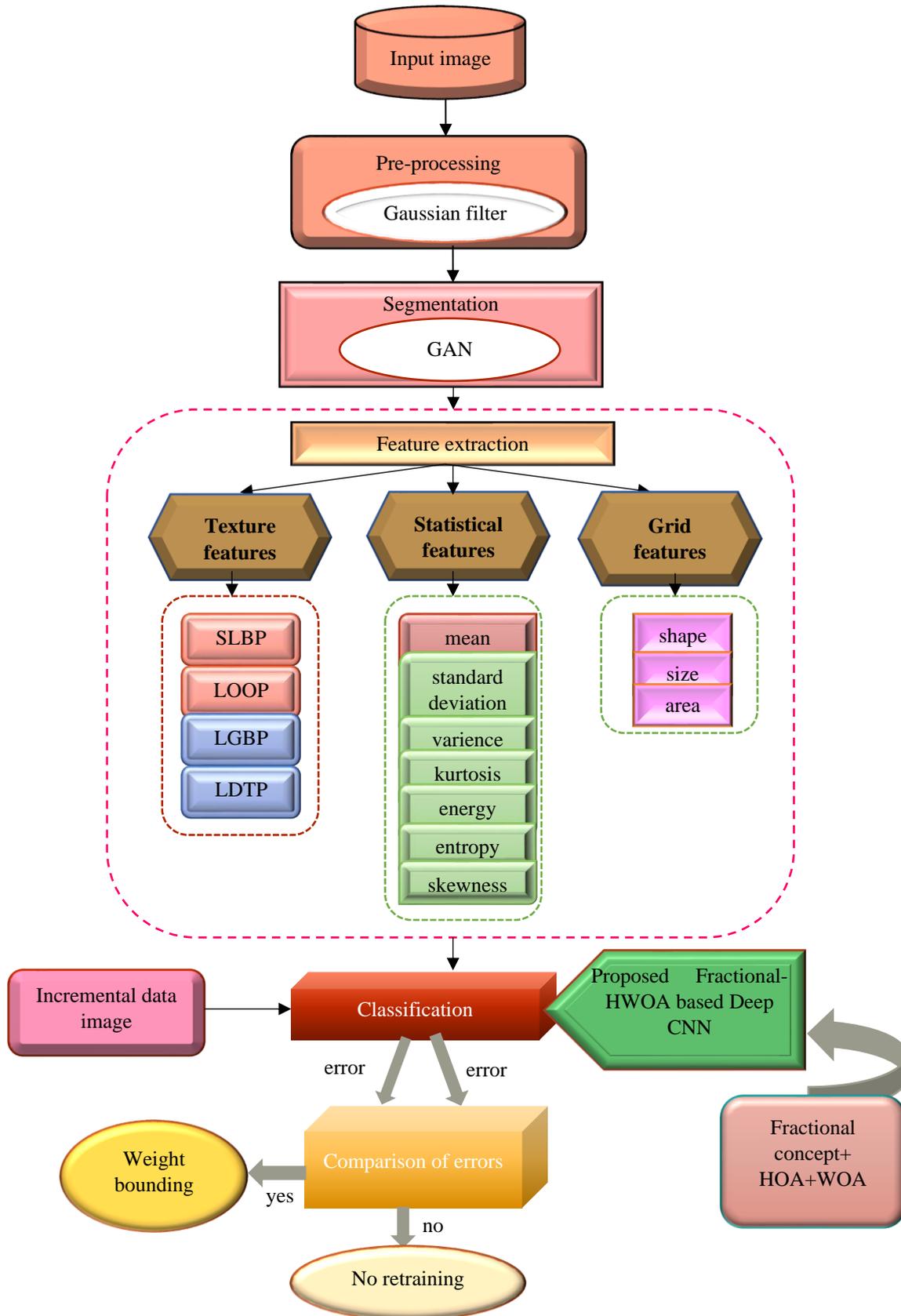

Fig. 1 Schematic view of introduced HWOA-enabled Deep CNN technique for incremental ALL classification
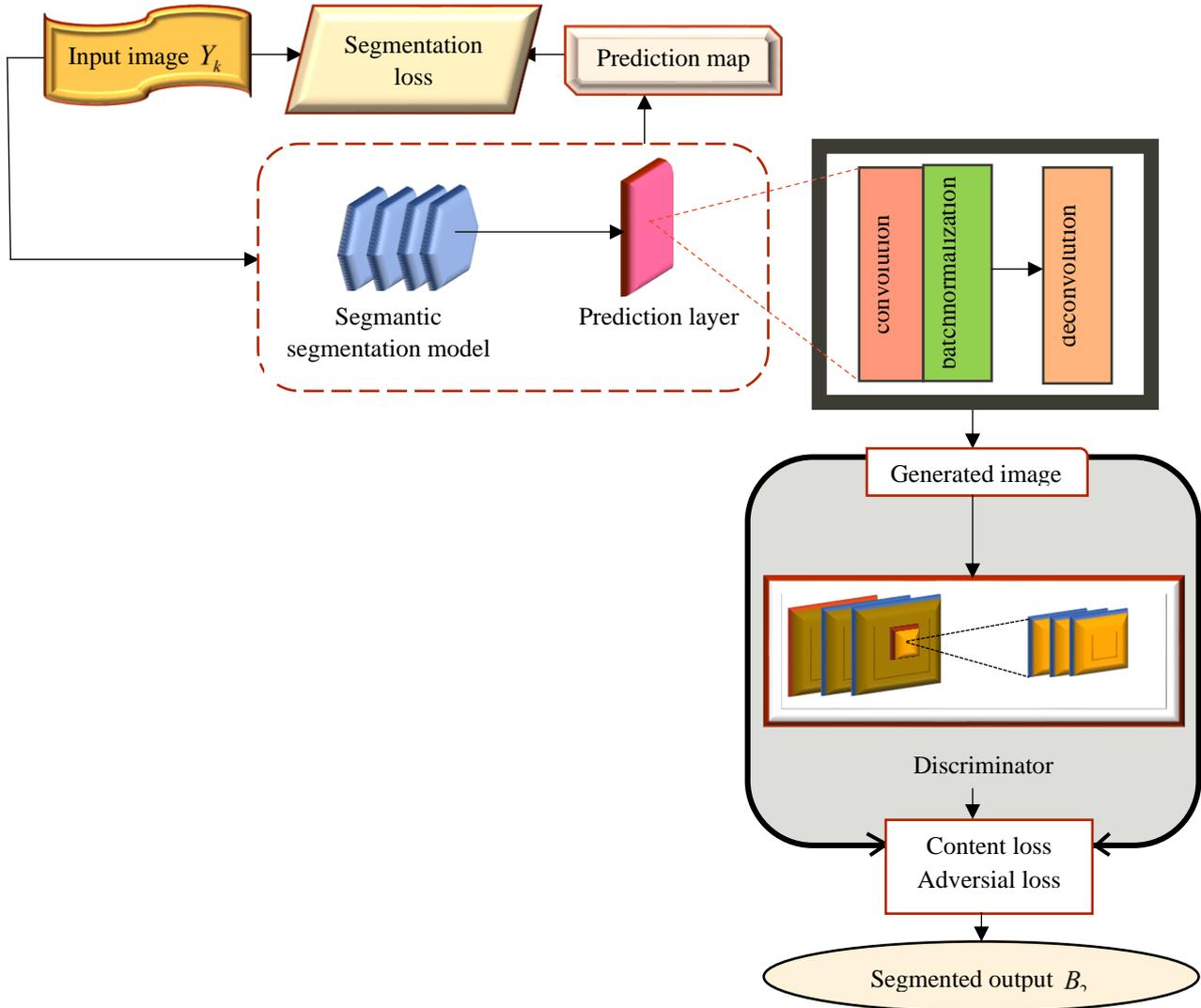
Fig. 2 Schematic architecture of GAN

3.1. Pre-Processing

The input image acquired from the database has some noises present in it and the removal of these noises along with artifacts is known as pre-processing. Before pre-processing, the image is first selected from the database. First let us consider dataset A with z number of input images, which is expressed as follows,

$$A = \{Y_1, Y_2, \dots Y_k, \dots Y_z\} \tag{1}$$

where, A is image database Y_k is the image represented on the kth index, and Y_z is the total images in the database. Here, Y_k an image is selected for pre-processing. For better classification of ALL, it is necessary to filter noises present in the image. Removing noises from the image is the first step after data acquisition. The pre-processing is carried out using a gaussian filter which denoises the image and enhances its quality of the image. The output generated from pre-processing is denoted as B_1 that is sent for the segmentation process.

3.2. Segmentation using GAN

Segmentation of the image is carried out as they easily verify ALL affected parts from the original image. This segmentation process is done after the above pre-processing step. The pre-processed image Y_k is segmented by GAN for converting the pre-processed image into various images. The GAN model is incorporated with a network of segmentation that formed Segmented GAN (SegGAN) [17], which is a differentiable method. This is composed of a semantic segmentation model, discriminator, and generator. The generator network G_e aims at generating an image based on the prediction layer of the segmentation model. Here, the discriminator network D_e performs as a supervisor for a generator G_e in the min-max optimization method. The synthetic image S^i is generated by the original image by the generator, and this reconstructed image is compared with the original image by the discriminator. Thus, the loss function based on the discriminator and synthetic image of GAN is as follows,

$$\min_{\theta G_e} \max_{\theta D_e} log(D_e(S)) + log(1 - D_e(S^i)) \tag{2}$$

where S is the original image. The generator conversely generates the same image from the prediction layer l_{seg} and the processing is termed as $G_e(l_{seg})$. The fake and original images are allowed to discriminator so that selecting images randomly is avoided. Based on simplicity and computation performance, DeepLab is chosen as the model for segmentation. The aim of this model is the generation of confidence maps $l_{seg} \in i^{d*w*h}$, where the dataset class number is indicated as d, the width of the prediction map is w, and h is the height of the prediction map. The segmentation process thus gives an output B_2. The resultant output is segmented into multiple image segments and is depicted as,

$$B_2 = \{B_2^1, B_2^2, \dots, B_2^n\} \qquad (3)$$

where, B_2^n signifies n number of image segments of B_2. Figure 2 depicts the architectural process of GAN.

3.3. Feature Extraction

The extraction of features in the processing of the image is a technique for reducing large dimensional redundant data to a reduced set of feature vectors [28]. This transformation of input data to a set of features is termed feature extraction. After the process of segmentation, the features are extracted. The features extracted in this paper are of three categories namely texture, statistical, and grid features. This extraction reduces the overfitting issues and hence leads to improve accuracy. The segmented output B_2 is fed to the feature extraction process, where the below-given features are extracted.

3.3.1. Texture Features

Nucleus texture measurements are performed on images of the nucleus on the version of the grayscale. The texture features comprise SLBP, LOOP, LGBP, and LDTP and are explained as follows,

SLBP

This feature describes the significant factor according to Weber's Law [18] of the object. This law demonstrates that the threshold's absolute value is not compared with local significance. The significant factor of a middle pixel is calculated as,

$$C(b_c) = arctan \sum_{g=0}^{\alpha} \frac{f_g - f_t}{f_t} + \frac{\pi}{2} \qquad (4)$$

where, the value of the middle pixel is denoted as b_c, the gray value of gth pixel is represented as f_g, and the gray value of tth pixel is represented as f_t. The LBP feature represents the local texture of the image and it is utilized everywhere as it is simple in the calculation, and severe for changes in the illumination. The LBP feature is represented by the below formula,

$$LBP(b_c) = \sum_{g=0}^{7} a(f_g, f_t).2^g \qquad (5)$$

where, $a(f_g, f_t)$ is the function of comparison which is represented as,

$$a(m,n) = \begin{cases} 1, if m - n > H \\ 0, otherwise \end{cases} \qquad (6)$$

where, the threshold is denoted as H. Thus, the SLBP feature is formed by the fusion of the LBP feature and the significant factor, thus the output of SLBP is represented by $T_1^1(1 \times 24)$

LOOP

This pattern is formed by the integration of the Local Directional pattern (LDP) and Local Binary pattern (LBP) and hence this is formed to reduce the drawbacks of both LDP and LBP [24]. The incorporation is done based on the Kirsch mask and the strength of this Kirsch output help to assign binarization weights to nearby pixels in its direction. The value of LOOP for a pixel is formulated as,

$$LOOP(r_p, s_p) = \sum j(q_o - q_p).2^{u}{}_o \qquad (7)$$

$$j(x) = \begin{cases} 1, if x \geq 0 \\ 0, otherwise \end{cases} \qquad (8)$$

where, q_o and q_p are the central pixels' grey level value. The LOOP output is signified by the term T_1^2.

LGBP

Displacement in the Gabor transform does not affect the magnitude; hence they are again encoded. To highlight Gabor Magnitude Picture (GMP), the LBP operator is encoded with the magnitude [19]. The binary numbers outcome is represented by a$F(y_a - y_b) = \begin{cases} 1, y_a \geq y_b \\ 0, y_a < y_b \end{cases}$ (9)

where $y_a (a = 0,1,\dots,7)$ is the thresholding neighborhood of each pixel and y_b is the centre value. The LBP pattern is assigned by applying 2^a to each $F(y_a - y_b)$ and this is formulated as,

$$LBP = \sum_{a=0}^{7} F(y_a - y_b)2^a \qquad (10)$$

Hence GMP is operated by LBP, which is denoted as an LGBP operator. LGBP term's output is signified as $T_1^3(1 \times 20)$.

LDTP

This operator comprises powerful textural information extracted from the principal axis [20]. Kirsch masks help to obtain LDTP code by including eight values of edge response, which is represented as,

$$K_r = |M_i * L_r| \qquad (11)$$

where, image is denoted as M_i, r^{th} Kirsch mask is noted as L_r, and operation of convolution is represented as $*$. Thus, LDTP is represented as $T_1^4(1 \times 20)$.

Hence, texture features are concluded as $TF_1 = \{T_1^1, T_1^2, T_1^3, T_1^4\}$.

where, image is denoted as M_i, r^{th} Kirsch mask is noted as L_r, and operation of convolution is represented as $*$. Thus, LDTP is represented as $T_1^4(1 \times 20)$.

Hence, texture features are concluded as $TF_1 = \{T_1^1, T_1^2, T_1^3, T_1^4\}$.

3.3.2. Statistical Features
The statistical features included are mean, standard deviation, variance, kurtosis, energy, entropy, and skewness. Each statistical feature is described below,

Mean

Mean feature has pixels related to brightness [26]. All the pixels of an image are counted and these pixels are also added together for the calculation of the mean as the sum of pixels by the count of pixels is the mean.

$$M_1 = \left(\frac{1}{o \times p}\right) \sum_{x=0}^{o-1} \sum_{y=0}^{p-1} f(x,y) \qquad (12)$$

where (o, p) is a pixel value in the maximum range, $(o - 1) and (p - 1)$ indicates decrement in one point, and $(x = 0) and (y = 0)$ indicates that pixels start from zero points. Thus, the output means is represented as $M_1(1 \times 1)$.

Standard Deviation

It is defined as variance square or mean values difference [26].

$$sd_1 = \sqrt{\left\{\left(\frac{1}{o \times p}\right) \sum_{x=0}^{o-1} \sum_{y=0}^{p-1} (f(x,y) - M_1)^2\right\}}$$
$$(13)$$

The output value of the standard deviation is represented by sd_1.

Variance

The value of gray level allocation is defined as the variance. It is the average square value in accordance with mean and single-pixel [26].

$$V_1 = \left(\frac{1}{o \times p}\right) \sum_{x=0}^{o-1} \sum_{y=0}^{p-1} (f(x,y) - M_1)^2 \qquad (14)$$

where the variance value is indicated as $V_1(1 \times 1)$.

Kurtosis

This represents peak values which are represented as the ratio of moments [30]. This value is represented as $K_1(1 \times 1)$.

Energy

The similarity of every pixel enhances energy in an image. The gray level is significant to find energy [26].

$$E_1 = \sum_{x=0}^{o-1} \sum_{y=0}^{p-1} f^2(x,y) \qquad (15)$$

Thus E_1 is the energy representation in an image in the order (1×1).

Entropy

It helps to identify the heterogeneity in an image [26]. The entropy is represented Ey_1 in the order of (1×1) and formulated as follows,

$$Ey_1 = -\sum_{x=0}^{o-1} \sum_{y=0}^{p-1} f(x,y) \log_2 f(x,y) \qquad (16)$$

Skewness

This represents the surface of the image by its brightness and dullness and depends on the variance and mean [30]. The skewness is denoted as $sk_1(1 \times 1)$.

The statistical features are expressed as $SF_1 = \{M_1, sd_1, V_1, K_1, E_1, Ey_1, sk_1\}$.

3.3.3. Grid features
The grid-enabled features, namely shape, area, and size are elaborated below as follows,

Shape

The blasts are discriminated by shape. It exhibits the region of the nucleus with zero-pixel value taken from a segmented image. This feature is exhibited as $sh_1(1 \times 64)$.

Size

The segmented image is clearly identified from its size as this denotes the dimensions. The size feature is denoted as $si_1(1 \times 2)$.

Area

The region in the image is checked for the identification of non-zero pixels and their total value is noted as the area of an image. The area is represented as $Ar_1(1 \times 1)$.

The grid features are finally denoted as the sum of the above three features and are represented as $GF_1 = \{sh_1, si_1, Ar_1\}$.

The overall extracted features are denoted $B_3 = \{TF_1, SF_1, GF_1\}$ in the order of (260×140).

3.4. Classification using Proposed Fractional-HWOA-based Deep CNN
After the extraction of features, the process of classification is performed by the proposed Fractional-HWOA-enabled Deep CNN. Using Deep CNN is very efficient as this classifies the disease of patients into two categories namely normal or abnormal. The efficiency and performance of the overall model are increased as the classification of the system is done after the feature extraction. The structure of Deep CNN along with its schematic diagram and process of training are given below,

3.4.1. Deep CNN Structure
Deep CNN [21] is adapted to learn discriminative denoisers by considering performance, discriminative color image, and speed before modelling. Deep CNNs architecture comprises three layers like convolutional (conv) layers, pooling (POOL) layers and a fully connected (FC) layer. The accuracy of the detection is increased by maximizing the convolution layer. As the convolution layers allow the features to the fully connected layers, then the prediction is made possible in a very accurate manner. Figure 3 shows the schematic architecture of Deep CNN.
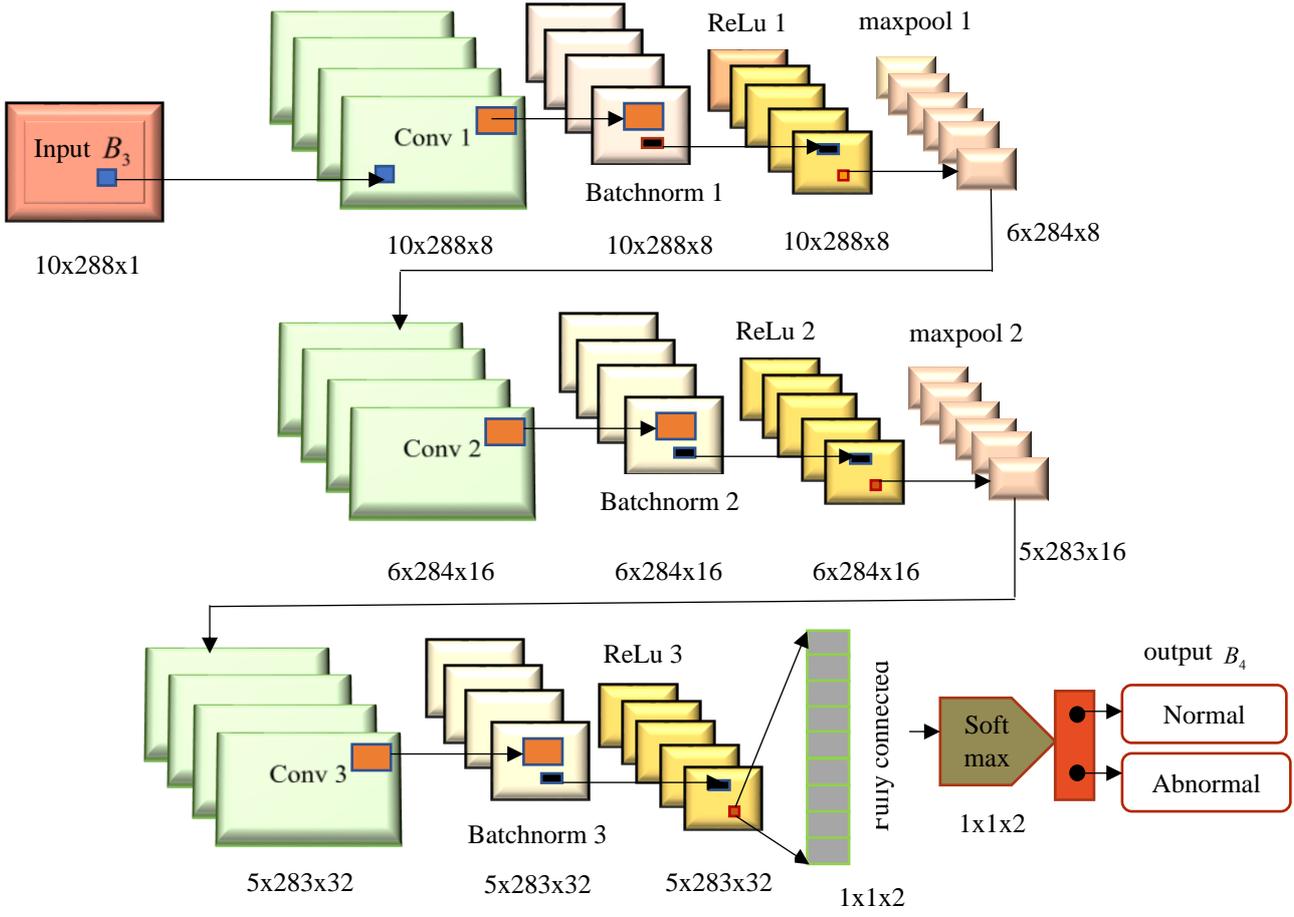
Fig. 3 Schematic architecture of Deep CNN

Convolutional Layers

They act as feature extractors and help to identify the feature representations of the input images. These layers have multiple neurons organized into feature maps. Every neuron in the map consists of trainable weights with a receptive field, which constantly forms a filter bank. The weights in the neurons are arranged in an equal manner. Also, if the neurons in the same convolution layer have various weights, then the features extracted are of multiple categories [31]. Many neurons are further arranged to extract variable features located in various locations. Normally, the feature map remains the same with equally weighing neurons [21]. The output B_3 from the feature fusion is taken as input for this convolutional layer and the representation of the convolutional layer is formulated as,

$$B_3 = \{I_1, I_2, \ldots, I_y, \ldots, I_l\}; 1 < y \leq l \quad (17)$$

where B_3 represents the convolution set with total convolutional layers represented as l. This set of convolutional layers acts as an input. The output is formulated below with the unit centered at (u,v),

$$(N_y^{\Im})_{u,v} = (O_y^{\Im})_{u,v} + \sum_{s=1}^{t_1^{y-1}} \sum_{w=-\omega_1^y}^{\omega_1^y} \sum_{x=-\omega_2^y}^{\omega_2^y} (k_{\Im,z}^y)_{w,x} * (N_y^{\Im-1})_{u+w,v+x} \quad (18)$$

where, $(N_y^{\Im})_{u,v}$ is the fixed feature map from y the convolutional layer, $(N_y^{\Im-1})_{u+w,v+x}$ is the feature map corresponding previous layer, $k_{\Im,z}^y$ is the kernel function in y^{th} a convolutional layer, $O_y^{\Im}$ is the bias of yth the convolutional layer, and * denotes convolutional operation.

The Rectified Linear Unit (ReLU) computation is a compulsory and easy step as there may be some nonlinearity in each layer. The yth nonlinear with feature map and result from the ReLU layer is formulated as,

$$N_y^{\Im} = fun(N_z^{\Im-1}) \quad (19)$$

where the convolution layer's activation function is $fun()$.

POOL Layers

To simplify computation, global and local POOL are included in convolutional layers [21]. POOL aims to reduce the special resolution of maps to get special invariance to distortions of input. The pooling aggregation is done to all small neighborhood features in an image. They propagate the highest or maximum values within the same receptive field [31].

$$\Gamma_{\Phi m_1 n_2} = \max_{(j1,j2) \in \aleph_{m_1 n_2}} \chi_{\Phi r_1 s_2} \qquad (20)$$

where output is $\Gamma_{\Phi m_1 n_2}$ with Φ^{th} feature and element in location $(j1, j2)$ is $\chi_{\Phi r_1 s_2}$ with pooling region of $\aleph_{m_1 n_2}$.

Fully Connected Layers

The features are detected from the fully connected layer and the prediction module as the features from the convolutional layers are fed into it. The resultant output from these fully connected layers is formulated below,

$$P_y^{\Im} = \gamma(Q_y^{\Im}) with Q_y^{\Im} = \sum_{s=1}^{t_1^{y-1}} \sum_{w=-\omega_1^y}^{\omega_1^y} \sum_{x=-\omega_2^y}^{\omega_2^y} (k_{\Im,z}^y)_{w,x} * (N_y^{\Im-1})_{u+w,v+x} \qquad (21)$$

where the unit linked by weight is $k_{\Im,z_{w,x}}^y$ at (u, v) of layer $(y-1)$ in z^{th} the map feature.

The stacking of the POOL and convolutional layer extracts more features while moving via the network. The softmax operator is used to improve the accuracy of classification. Thus, the output achieved by Deep CNN is represented as B_4 that represents normal or abnormal cells.

3.4.2. Training using proposed Fractional-HWOA

Proposed Fractional-HWOA is used for the training process of Deep CNN. This proposed Fractional-HWOA is derived from the integration of the Fractional concept into HOA and WOA. HOA imitates the horse performance at various stages forming a meta-heuristic algorithm. This horse herd algorithm solves problems in high dimensions having complex characters producing good accuracy and efficiency with less computational cost. WOA is a nature-inspired algorithm. This algorithm imitates the character of humpback whales forming a meta-heuristic algorithm. The integration of both characters is seen in this proposed method. The algorithmic stages of developed HWOA are explained as given below,

Initialization

Initialization is the first step of the training process in which the dataset is denoted R with ∂ the number of solutions. The solution set is given as,

$$R = \{R_1, R_2, \ldots, R_e, \ldots, R_\partial\}, (1 < e \le \partial)$$

where R_e is the element represented in e^{th} place and ∂ signifies the total elements in the set.

Computation of Fitness

The fitness computation is significant in finding the best value of fitness based on the formula and the fitness function is formulated as,

$$Fitness = \frac{1}{\chi} \sum_{h=1}^{\chi} [L_{a_\alpha} - H_{a_\alpha}] \qquad (23)$$

where L_{a_α} is the output targeted, H_{a_α} is the result obtained from Deep CNN, and h depicts the number of training samples used for processing.

Horse Movement Analysis

The horse movement is represented in the below equation,

$$J_\forall^{age}(\Re + 1) = \Delta_\forall^{age}(\Re + 1) + J_\forall^{age}(\Re) \qquad (24)$$

At a minimal time, the results are obtained in an optimized way by integrating the fractional concept of WOA and HOA. According to WOA, the standard equation is given as,

$$J_\forall^{age}(\Re + 1) = J_\forall^*(\Re) - \hbar.\psi \qquad (25)$$

$$\psi = |W.J_\forall^*(\Re) - J_\forall^{age}(\Re)| \qquad (26)$$

$$J_\forall^{age}(\Re + 1) = J_\forall^*(\Re) - \hbar.|W.J_\forall^*(\Re) - J_\forall^{age}(\Re)| \qquad (27)$$

Considering $J_\forall^*(\Re) > J_\forall^{age}(\Re)$,

$$J_\forall^{age}(\Re + 1) = J_\forall^*(\Re) - B[W.J_\forall^*(\Re) - J_\forall^{age}(\Re)] \qquad (28)$$

$$J_\forall^{age}(\Re + 1) = J_\forall^*(\Re) - \hbar W.J_\forall^*(\Re) + \hbar J_\forall^{age}(\Re) \qquad (29)$$

$$J_\forall^{age}(\Re + 1) = J_\forall^*(\Re)[1 - \hbar W] + \hbar J_\forall^{age}(\Re) \qquad (30)$$

$$\hbar J_\forall^{age}(\Re) = J_\forall^{age}(\Re + 1) - J_\forall^*(\Re)[1 - \hbar W] \qquad (31)$$

$$J_\forall^{age}(\Re) = \frac{J_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{\hbar} \qquad (32)$$

Substituting equation (32) in (24), the equation becomes,

$$J_\forall^{age}(\Re + 1) = \Delta_\forall^{age}(\Re + 1) + \frac{J_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{\hbar} \qquad (33)$$

$$J_\forall^{age}(\Re + 1) = \frac{\hbar \Delta_\forall^{age}(\Re+1) + J_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{\hbar} \qquad (34)$$

$$J_\forall^{age}(\Re + 1) - \frac{J_\forall^{age}(\Re+1)}{\hbar} = \frac{\hbar \Delta_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{\hbar} \qquad (35)$$

$$J_\forall^{age}(\Re + 1) - \frac{J_\forall^{age}(\Re+1)}{\hbar} = \frac{\hbar \Delta_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{\hbar} \qquad (36)$$

$$\frac{\hbar J_\forall^{age}(\Re+1) - J_\forall^{age}(\Re+1)}{\hbar} = \frac{\hbar \Delta_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{\hbar} \qquad (37)$$

$$J_\forall^{age}(\Re + 1)[\hbar - 1] = \hbar \Delta_\forall^{age}(\Re + 1) - J_\forall^*(\Re)[1 - \hbar W] \qquad (38)$$

$$J_\forall^{age}(\Re + 1) = \frac{\hbar \Delta_\forall^{age}(\Re+1) - J_\forall^*(\Re)[1-\hbar W]}{[\hbar - 1]} \qquad (39)$$

Subtracting $J_\forall^{age}(\Re)$ on both sides,

$$J_\forall^{age}(\Re + 1) - J_\forall^{age}(\Re) = \frac{\hbar\Delta_\forall^{age}(\Re+1)-J_\forall^*(\Re)[1-\hbar W]}{[\hbar-1]} - $$
$$J_\forall^{age}(\Re) \qquad (40)$$

From the Fractional concept,

$$A^\Omega\left(J_\forall^{age}(\Re + 1)\right) = \frac{\hbar\Delta_\forall^{age}(\Re+1)-J_\forall^*(\Re)[1-\hbar W]}{[\hbar-1]} - J_\forall^{age}(\Re)$$
$$(41)$$

$$J_\forall^{age}(\Re + 1) - \Omega J_\forall^{age}(\Re) - \frac{1}{2}\Omega J_\forall^{age}(\Re - 1) - \frac{1}{6}(1$$
$$- \Omega)J_\forall^{age}(\Re - 2) - \frac{1}{24}\Omega(1 - \Omega)(2$$
$$- \Omega)J_\forall^{age}(\Re - 3)$$
$$= \frac{\hbar\Delta_\forall^{age}(\Re+1)-J_\forall^*(\Re)[1-\hbar W]}{[\hbar-1]} - J_\forall^{age}(\Re) \qquad (42)$$

$$J_\forall^{age}(\Re + 1) = \Omega J_\forall^{age}(\Re) + \frac{1}{2}\Omega J_\forall^{age}(\Re - 1) + \frac{1}{6}(1$$
$$- \Omega)J_\forall^{age}(\Re - 2) + \frac{1}{24}\Omega(1 - \Omega)(2$$
$$- \Omega)J_\forall^{age}(\Re - 3)$$
$$+ \frac{\hbar\Delta_\forall^{age}(\Re + 1) - J_\forall^*(\Re)[1 - \hbar W]}{[\hbar - 1]} - J_\forall^{age}(\Re)$$
$$(43)$$

$$J_\forall^{age}(\Re + 1) = (\Omega - 1)J_\forall^{age}(\Re) + \frac{1}{2}\Omega J_\forall^{age}(\Re - 1) + \frac{1}{6}(1$$
$$- \Omega)J_\forall^{age}(\Re - 2) + \frac{1}{24}\Omega(1 - \Omega)(2$$
$$- \Omega)J_\forall^{age}(\Re - 3)$$
$$+ \frac{\hbar\Delta_\forall^{age}(\Re+1)-J_\forall^*(\Re)[1-\hbar W]}{[\hbar-1]} \qquad (44)$$

where, ages of horses are $\ell, \theta, \Omega, \varepsilon$; coefficient vectors are $\hbar, W$; Ω is in range of $[0,1]$; $J_\forall^*(\Re)$ is the best solution's position; $J_\forall^{age}(\Re)$ is solution's position at iteration$\Re$, $J_\forall^{age}(\Re - 1)$ is solution's position at iteration $(\Re - 1)$;$J_\forall^{age}(\Re - 2)$ is solution's position at iteration$(\Re - 2)$;$J_\forall^{age}(\Re - 3)$ is solution's position at iteration $(\Re - 3)$;$\Delta_\forall^{age}(\Re + 1)$ its velocity vector.

Grazing Analysis

To find the accurate area of grazing, the region of grazing is accurately modelled by applying a coefficient to each horse. Hence, the equation is given below,

$$B_\forall^{age}(\Re + 1) = \kappa_{iter}(\partial + \varpi Y)[J_\forall(\Re)] \qquad (45)$$

$$\kappa_\forall^{age}(\Re + 1) = \kappa_\forall^{age}(\Re) \times \lambda_\kappa \qquad (46)$$

where, the parameter for movement is indicated as $B_\forall^{age}(\Re + 1)$, linearity is denoted as λ_K, the upper bound is indicated as ∂, random number is indicated as ϖ, and the lower bound is defined as Y.

Hierarchy Analysis

The horses follow always a leader and this representation is given as,

$$E_\forall^{age}(\Re + 1) = \infty_\forall^{age}(\Re + 1)[J_*(\Re) - J_\forall(\Re)] \qquad (47)$$

$$\infty_\forall^{age}(\Re + 1) = \infty_\forall^{age}(\Re) \times \lambda_K \qquad (48)$$

where, the effects of the position of the finest horse is $\infty_\forall^{age}(\Re + 1)$, and $J_*(\Re)$ symbolizes the best horse's position.

Sociability Analysis

This stage is maintaining a social relationship or associating with the average position of other horses. This expression is formulated as follows,

$$\zeta_\forall^{age}(\Re + 1) = \phi_\forall^{age}(\Re + 1)\left[\left(\frac{1}{M}\sum_{v=1}^{M}J_v(\Re)\right) - J_\forall(\Re)\right]$$
$$(49)$$

$$\phi_\forall^{age}(\Re + 1) = \phi_\forall^{age}(\Re) \times \sigma_\phi \qquad (50)$$

where, the associated horse movement vector is noted as$\zeta_\forall^{age}(\Re + 1)$, the orientation of a particular horse to herd is formulated as$\phi_\forall^{age}(\Re + 1)$, and the overall horse weight factor is symbolized as σ_ϕ.

Imitation Analysis

Imitation of grazing is seen in horses after noticing the proper location for grazing. This identification is expressed as follows,

$$A_\forall^{age}(\wp + 1) = \ell^{\circ age}_\forall(\Re + 1)\left[\left(\frac{1}{XM}\sum_{v=1}^{XM}\hat{J}_v(\Re)\right) - J(\Re)\right]$$
$$(51)$$

$$\ell^{\circ age}_\forall(\Re + 1) = \ell^{\circ age}_\forall(\Re) \times \sigma_n \qquad (52)$$

where, the horse movement vector in the direction of the average of the best horse with the finest position is $A_\forall^{age}(\wp + 1)$, an overall horse with the finest position is denoted asXM, and each iteration cycle's reduction factor is σ_n.

Defence Strategy Analysis

The disturbances and unwanted responses are kept away from the horse herd and it is formulated as,

$$H_\forall^{age}(\Re + 1) = -\overrightarrow{\leftrightarrows} \$_\forall^{age}(\Re + 1)\left[\left(\frac{1}{p^\circ M}\sum_{v=1}^{p^\circ V}\hat{J}_v(\Re)\right) - J(\Re)\right] \qquad (53)$$

$$\$_\forall^{age}(\Re + 1) = \$_\forall^{age}(\Re) \times \sigma_\$ \qquad (54)$$

where the vector of escapism from horses average with poor positions is given as $H_\forall^{age}(\Re + 1)$ worst position of horse $p^\circ M$, and the parameter for reduction per cycle is $\sigma_\$$.

Roaming Analysis

At younger stages, the horses roam and after attaining maturity, this roaming character is disappeared this analysis is formulated as,

$$I_\forall^{age}(\Re + 1) = \xi_\forall^{age}(\Re + 1)\omega J(\Re) \qquad (55)$$

$$\xi_\forall^{age}(\Re + 1) = \xi_\forall^{age}(\Re) \times \sigma_\xi \qquad (56)$$

where horse velocity is indicated randomly $I_\forall^{age}(\Re + 1)$ and the limit of reduction $\xi_\forall^{age}(\Re + 1)$ per cycle is indicated as σ_ξ.

Updated Solutions Feasibility Assessment

Feasibility assessment is carried out by the consideration of fitness value to find the best solution as soon as the updated solution is derived. When superiority exists in a new solution while compared with the old solution, then the old solution is replaced with a newly updated solution.

End

Iterations are done for the above steps till the superior maximum solution is attained. Pseudo-code of the proposed Fractional-HWOA is depicted in table 1.

Table 1. Pseudocode of developed Fractional-HWOA

Input: Solution J
Output: Maximum solution
Start
Calculate the equation of fitness (23)
If the circumstances of stopping are not satisfied **then**
Calculate horse ages $\ell, \theta, \Omega, \varepsilon$
Calculate each horse's velocity by equation (24)
Calculate the position of the horse including the fractional concept by equation (44)
Update new horse position with grazing by equation (45)
Update new horse position with Hierarchy by equation (47)
Update new horse position with sociability by equation (49)
Update new horse position with imitation by equation (51)
Update new horse position with defence mechanism by equation (53)
Update new horse position with roaming by equation (55)
End if;
Recalculate error by equation (23)
$\Re = \Re + 1$;
End

Hence, the developed fractional-HWOA-based Deep CNN is efficient for ALL classification to identify as normal or abnormal. Also, the error by this proposed model from the input image is obtained as $\bar{\tau}$ and then this method is again followed for the incremental image.

3.5. Incremental Learning using Proposed Fractional-HWOA based Deep CNN for ALL Classification

The incremental learning method is very helpful for the classification of disease while simultaneously retaining its prior obtained knowledge. They exploit structure-based and semantic-based similarities which help for better identification of diseases [1]. The input image Y_k is acquired from the dataset and the classification is carried out for this input image Y_k by proposed fractional-HWOA based Deep CNN and the error attained from this process is denoted as $\bar{\tau}$. Similarly, the same process is carried out for the incremental image $(Y_k + 1)$ as followed in above section 3 by proposed fractional-HWOA enabled Deep CNN where the error obtained is denoted as $\bar{\tau}_e$.

If $\bar{\tau}_e > \bar{\tau}$, then weight bounding based on Renyi entropy is done.

If $\bar{\tau}_e < \bar{\tau}$, then no retraining is required.

where, $\bar{\tau}$ is an error obtained from the input image Y_k, and $\bar{\tau}_e$ is an error obtained from the incremental input image $(Y_k + 1)$.

3.5.1. Weight Bounding

Renyi entropy is used to attain an optimal solution with a low computational burden via the sequential process. In this process first a new factor for weight is developed that assembles other factors to balance local exploitation and global exploration by quantifying the variation degree of obtained value. This helps to assess the relative discrepancy between uncertainty and predicted value [29]. Weight update is based on the bounded weights as,

$$\cup^{\tau_e} = \cup^\tau + \cup^{RE} \qquad (57)$$

where $\cup^{\tau_e}$ is the weight obtained from the incremental image, $\cup^\tau$ is the weight obtained from the input image, and $\cup^{RE}$ is the Renyi entropy-based weight coefficient which is formulated as,

$$\cup^{RE} = \frac{1 - N(O_j)}{\sum_{j=1}^{n}(1 - N(O_j))} \qquad (58)$$

$$N(O_j) = \frac{1}{1-\eta} log\left(\sum_{i=1}^{m} O_{ij}{}^\eta\right); \eta \geq 0 \, and \, \eta \neq 1 \qquad (59)$$

where, the order of renyi entropy is represented by η, and the entropy value of each index is represented as $N(O_j)$.

4. Results and Discussion

This section suggests the results based on the metrics used for evaluation and the discussions are done based on the values obtained from the performance metrics.

4.1. Experimental Setup

The implementation of the proposed method is done in MATLAB tool having a PC with an Intel I3 core processor, 2 GB RAM, and windows 10 OS.

4.2. Description of Dataset

The dataset used in this method is ALL-IDB, which is an image-based dataset. This dataset consists of images of microscopic samples of blood with 108 images of 39000 blood samples. These samples are important for the identification and classification of ALL. Also, the images of ALL-IDB are obtained using an optical laboratory microscope with a Canon Powershot G5 camera in Joint Photographic Experts Group (JPG) format with the resolution $2592*1944$ having the intensity of 24-bit color.

4.3. Evaluation Metrics

The metrics assessed for this method are accuracy, sensitivity, and specificity. The details of metrics performed in this proposed method are given below,

4.3.1. Testing Accuracy

The accurate measurement of classification of ALL from the healthy person is identified based on this metric and is formulated as,

$$A_{cc} = \frac{u_{cc}+v_{cc}}{u_{cc}+a_{cc}+v_{cc}+b_{cc}} \tag{60}$$

where true positives are noted as v_{cc} true negatives are u_{cc}, and false positives and negatives are depicted as b_{cc} and a_{cc} respectively.

4.3.2. Sensitivity

This metric suggests the positive measurement it means the abnormal disease classification in patients is easily identified by this formula,

$$Sen = \frac{v_{cc}}{a_{cc}+v_{cc}} \tag{61}$$

4.3.3. Specificity

This metric suggests the negative measurement it means the normal disease classification in patients is easily identified by the below formula,

$$Spe = \frac{u_{cc}}{u_{cc}+b_{cc}} \tag{62}$$

4.4. Experimental Outcomes

Figure 4 portrays the experimental results of the proposed fractional-HWOA-based Deep CNN method in performing ALL classification evaluation. Figure 4a) depicts the input image, figure 4b) presents pre-processed image, figure 4c) shows the segmented image, figure 4d) depicts the SLBP feature extracted image, figure 4e) represents the LOOP feature extracted image, and 4f) represents the LDTP feature extracted image.

4.5. Performance Analysis

The various techniques taken for the performance assessment comprise, Fractional HWOA-based Deep CNN with epoch 20, Fractional HWOA-based Deep CNN with epoch 40, Fractional HWOA-based Deep CNN with epoch 60, Fractional HWOA-based

Deep CNN with epoch 80, and Fractional HWOA-based Deep CNN with epoch 100.

4.5.1. Performance Analysis with Respect to Chunk Size

Figure 5 depicts the performance analysis of fractional HWOA-based Deep CNN with chunk size. Figure 5 a) represents the performance analysis of fractional HWOA-based Deep CNN in terms of accuracy. When the chunk size is 3, the accuracy for Fractional

HWOA-based Deep CNN with epoch 100 is 0.92, whereas the accuracy for epoch 80, epoch 60, epoch 40, and epoch 20 are 0.9, 0.88, 0.86, and 0.85. Figure 5 b) illustrates the analysis of the performance of the proposed method in terms of sensitivity. When chunk size is 4, the proposed method's sensitivity for epoch 100 is 0.98 and the sensitivity values are 0.94, 0.92, 0.9, and 0.86 for epoch 80, epoch 60, epoch 40, and epoch 20. The performance analysis of specificity for the proposed fractional HWOA-based Deep CNN is given in figure 5 c), and the specificity in terms of epochs 100, 80, 60, 40, and 20 are 0.9, 0.88, 0.87, 0.85, and 0.84.

4.6. Algorithmic Analysis

The various techniques taken for the algorithmic analysis comprise Deep CNN + Particle Swarm Optimization (PSO), Deep CNN + Grey Wolf Optimization (GWO), Deep CNN + Whale Optimization Algorithm (WOA), Deep CNN + Harris Hawks Optimization (HHO), and Deep CNN + Horse Whale Optimization Algorithm (HWOA).

4.6.1. Algorithmic Analysis with Varying Iterations

Figure 6 illustrates the Algorithmic analysis of Deep CNN with varying iterations. Figure 6 a) enlightens the accuracy-based algorithmic analysis of Deep CNN with altering iterations. For iteration 50, the accuracy value for Deep CNN + HWOA is 0.93, whereas the accuracy values are 0.78, 0.8, 0.82, 0.85, and 0.9 for Deep CNN + PSO, Deep CNN + GWO, Deep CNN + WOA, and Deep CNN + HHO. The algorithmic analysis of Deep CNN with altering iterations in terms of sensitivity is given in figure 6 b). For iteration 250, the sensitivity value for Deep CNN + HWOA is 0.98. The values of sensitivity are 0.83, 0.85, 0.87, 0.9, and 0.92 for Deep CNN + PSO, Deep CNN + GWO, Deep CNN + WOA, and Deep CNN + HHO. Figure 6 c) illustrates the algorithmic analysis of Deep CNN with altering iterations in terms of specificity. The specificity value is 0.95 for Deep CNN + HWOA, whereas the specificity values are 0.8, 0.82, 0.83, 0.88, and 0.91 for Deep CNN + PSO, Deep CNN + GWO, Deep CNN + WOA, and Deep CNN + HHO.

4.7. Comparative Techniques

The various comparative methods taken for the assessment include, Incremental scheme, Incremental learning loss function, Hybrid deep learning, and proposed Fractional HWOA-based Deep CNN.

(a)

(b)

(c)
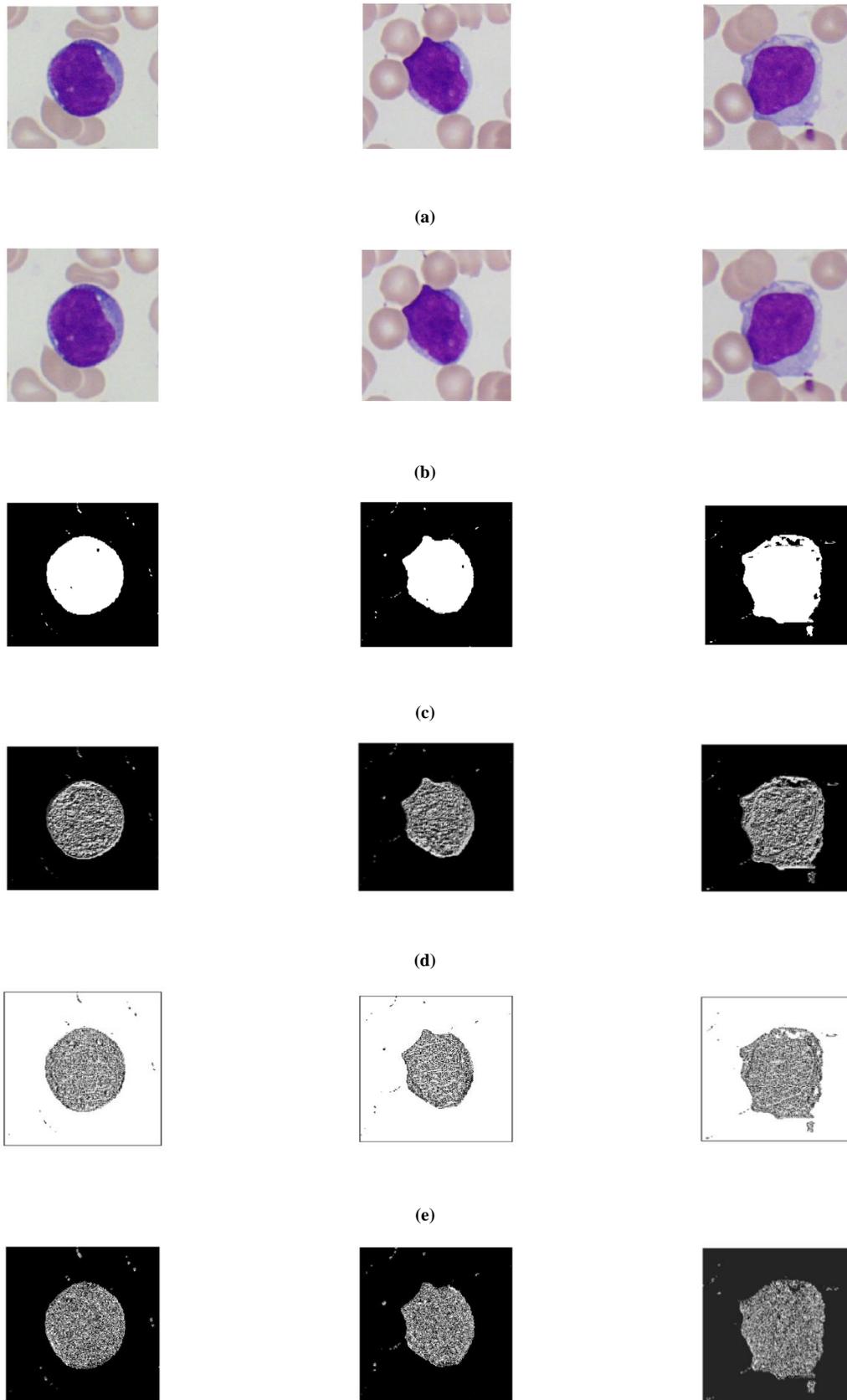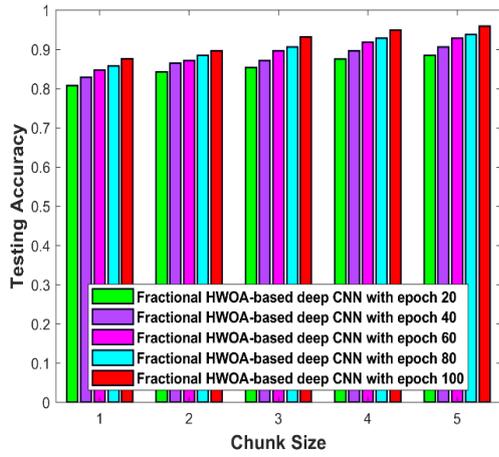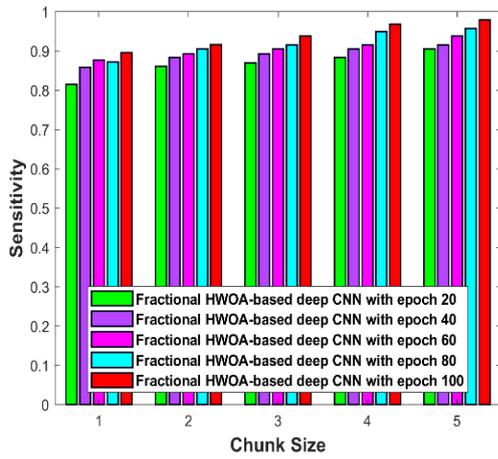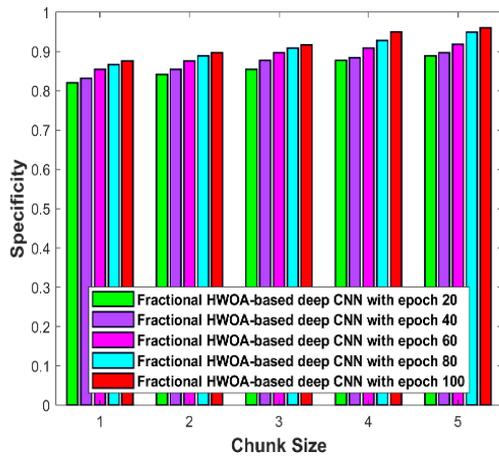

(d)

(e)

(f)

Fig. 4 Experimental results of, a) input image, b) pre-processed image, c) segmented image, d) SLBP image, e) LOOP image, f) LDTP image
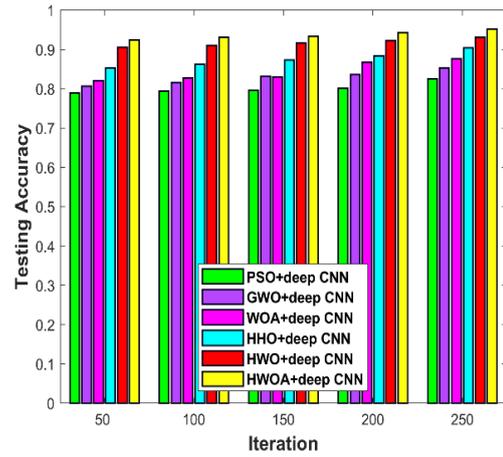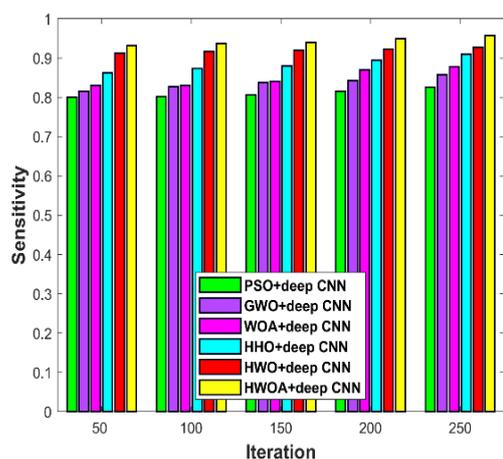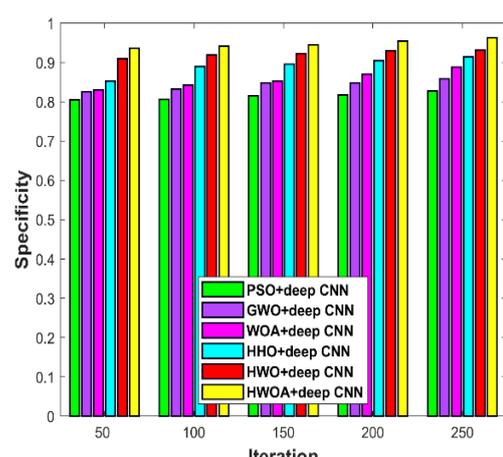
(a)

(b)

(c)

Fig. 5 Performance analysis of fractional HWOA-based Deep CNN with chunk size, a) accuracy, b) sensitivity, c) specificity

(a)

(b)

(c)

Fig. 6 Algorithmic analysis of Deep CNN with varying iterations, a) accuracy, b) sensitivity, c) specificity
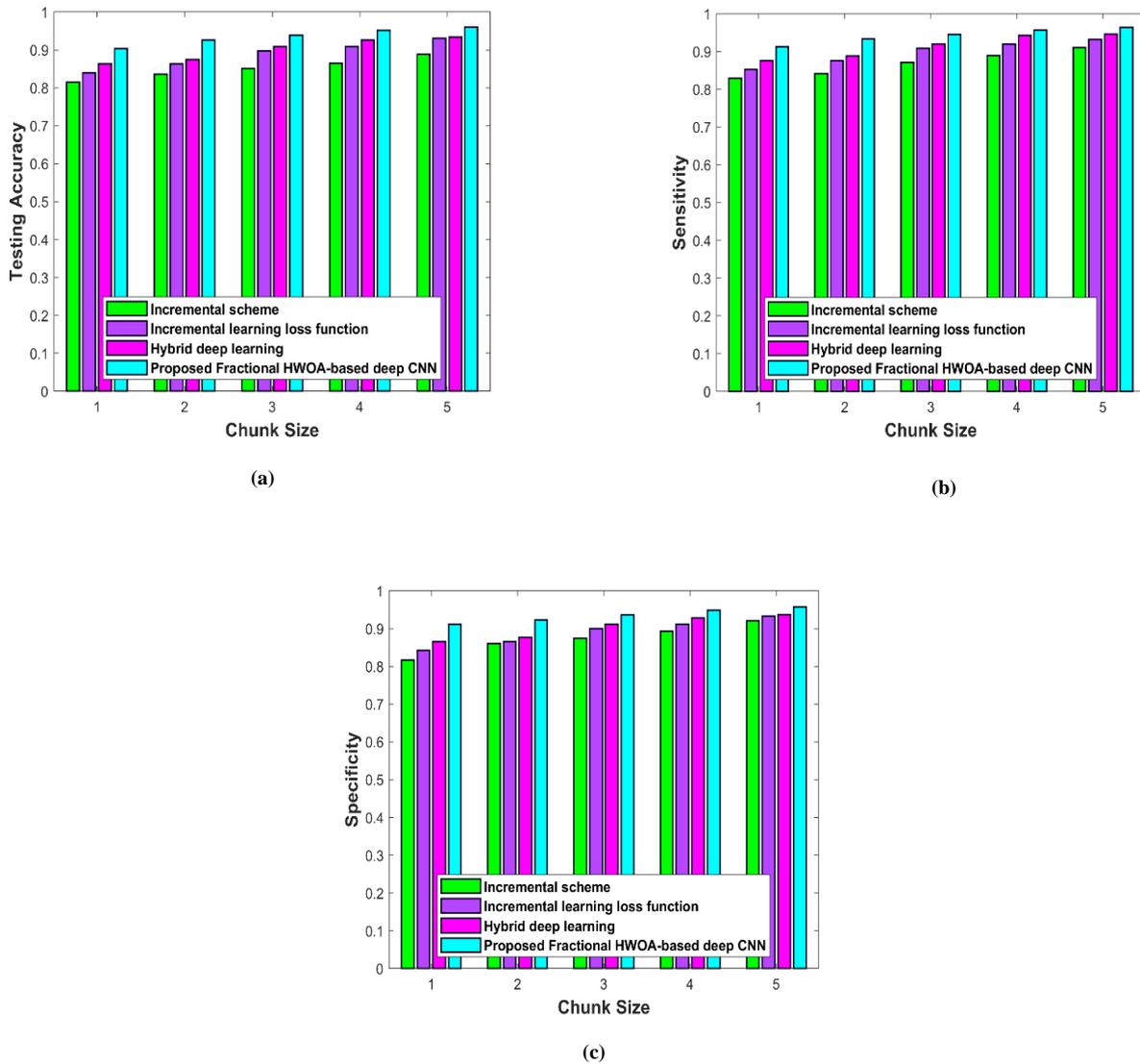
(a)

(b)

(c)

Fig. 7 Comparative analysis of fractional HWOA-based Deep CNN a) accuracy, b) sensitivity, c) specificity

4.9. Comparative Discussion

Table 2 enumerates the comparative discussion of the proposed Fractional-HWOA-based Deep CNN with existing methods. For the comparative discussion, the accuracy for introduced Fractional HWOA-enabled Deep CNN is 0.98, whereas the accuracy values of Incremental scheme, Incremental learning loss function, and Hybrid deep learning are 0.85, 0.9, and 0.9. Also, maximal sensitivity of 0.98, and maximal specificity of 0.98, respectively are attained for chunk size 5. The incorporation of algorithms and feature extraction tends to maximise the values of three evaluation metrics like accuracy, sensitivity, and specificity. This signifies the effectiveness of the developed Fractional HWOA-based Deep CNN method in ALL classification.

5. Conclusion

ALL classification in the earlier stage is not much easier to predict. Hence, this research work represents an example of effective incremental classification of ALL, in accordance with the Proposed Fractional HWOA-based Deep CNN. The pre-processing carried out by the gaussian filter followed by GANs segmentation helped to extract possible features. The classification is done by Deep CNN trained by Fractional HWOA for extracted features. The designed fractional HWOA is the hybridization of the fractional concept of HOA and WOA. Also, Incremental image classification is carried out based on our proposed model, where the errors of the input image and incremental image are compared to apply weight bounding. The experiments demonstrated that the devised fractional HWOA-based Deep CNN model can precisely classify ALL cells with maximal accuracy of 0.961, maximum sensitivity of 0.957, and maximum specificity of 0.958. However, future work will consider a more efficient model that enables professionals of medical related and patients to give extra confidence that exactly predict ALL type. Moreover, the training model can be changed to get more accurate prediction results when compared with the present model.

Table 2. Comparative discussion of the proposed method with existing techniques

Classification level	Metrics / Methods	Incremental scheme	Incremental learning loss function	Hybrid deep learning	Proposed Fractional HWOA-based Deep CNN
	Accuracy	0.889	0.930	0.934	**0.961**
Chunk size 5	**Sensitivity**	0.911	0.932	0.946	**0.957**
	Specificity	0.921	0.933	0.937	**0.958**

References

[1] Sirshar M, Hassan T, Akram M.U, and Khan S.A, "An Incremental Learning Approach to Automatically Recognize Pulmonary Diseases from the Multi-Vendor Chest Radiographs," *Computers in Biology and Medicine,* vol. 134, 2021. Crossref, https://doi.org/10.1016/j.compbiomed.2021.104435

[2] Al-Qudah R, and Suen C.Y, "Improving Blood Cells Classification in Peripheral Blood Smears using Enhanced Incremental Training," *Computers in Biology and Medicine,* vol. 131, 2021. Crossref, https://doi.org/10.1016/j.compbiomed.2021.104265

[3] He W, Wang X, Wang L, Huang Y, Yang Z, Yao X, Zhao X, Ju L, Wu L, Wu L, and Lu H, "Incremental Learning for Exudate and Hemorrhage Segmentation on Fundus Images," *Information Fusion*, vol. 73, pp. 157-164, 2021. Crossref, https://doi.org/10.1016/j.inffus.2021.02.017

[4] Zhao J, Zhang M, Zhou Z, Chu J, and Cao F, "Automatic Detection and Classification of Leukocytes Using Convolutional Neural Networks," *Medical & Biological Engineering & Computing,* vol. 55, no. 8, pp. 1287-1301, 2017. Crossref, https://doi.org/10.1007/s11517-016-1590-x

[5] Das P.K, and Meher S, "An Efficient Deep Convolutional Neural Network Based Detection and Classification of Acute Lymphoblastic Leukemia," *Expert Systems with Applications*, vol. 183, 2021. Crossref, https://doi.org/10.1016/j.eswa.2021.115311

[6] Rawat J, Singh A, Bhadauria H.S, Virmani J, and Devgun J.S, "Classification of Acute Lymphoblastic Leukaemia Using Hybrid Hierarchical Classifiers," *Multimedia Tools and Applications,* vol. 76, no. 18, pp. 19057-19085, 2017. Crossref, https://doi.org/10.1007/s11042-017-4478-3

[7] Srisukkham W, Zhang L, Neoh S.C, Todryk S, and Lim C.P, "Intelligent Leukaemia Diagnosis with Bare-Bones PSO Based Feature Optimization," *Applied Soft Computing*, vol. 56, pp. 405-419, 2017. Crossref, https://doi.org/10.1016/j.asoc.2017.03.024

[8] Li Y, Zhu R, Mi L, Cao Y, and Yao D, "Segmentation of White Blood Cell from Acute Lymphoblastic Leukemia Images using Dual-Threshold Method," *Computational and Mathematical Methods in Medicine*, 2016. Crossref, https://doi.org/10.1155/2016/9514707

[9] Labati R.D, Piuri V, and Scotti F, "All-IDB: The Acute Lymphoblastic Leukemia Image Database for Image Processing," *In Proceedings of 18th IEEE International Conference on Image Processing*, IEEE, pp. 2045-2048, 2011. Crossref, https://doi.org/10.1109/ICIP.2011.6115881

[10] Eng-Juh Yeoh, Mary E Ross, Sheila A Shurtleff, W Kent Williams, Divyen Patel, Rami Mahfouz, Fred G Behm, Susana C Raimondi, Mary V Relling, Anami Patel, Cheng Cheng, Dario Campana, Dawn Wilkins, Xiaodong Zhou, Jinyan Li, Huiqing Liu, Ching-Hon Pui, William E Evans, Clayton Naeve, Limsoon Wong, and James R Downing, "Classification, Subtype Discovery, and Prediction of Outcome in Pediatric Acute Lymphoblastic Leukemia by Gene Expression Profiling," *Cancer Cell*, vol. 1, no. 2, pp. 133-143, 2002. Crossref, https://doi.org/10.1016/s1535-6108(02)00032-6

[11] Rehman A, Abbas N, Saba T, Rahman S.I.U, Mehmood Z, and Kolivand H, "Classification of Acute Lymphoblastic Leukemia using Deep Learning", *Microscopy Research and Technique*, vol. 81, no. 11, pp. 1310-1317, 2018. Crossref, https://doi.org/10.1002/jemt.23139

[12] V. Shalini, K. S. Angel Viji, "Integration of Convolutional Features and Residual Neural Network for the Detection and Classification of Leukemia from Blood Smear Images," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 176-184, 2022. Crossref, https://doi.org/10.14445/22315381/IJETT-V70I9P218

[13] Rebuffi S.A, Kolesnikov A, Sperl G, and Lampert C.H, "iCaRl: Incremental Classifier and Representation Learning", *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5533-5542, 2017. Crossref, https://doi.org/10.1109/CVPR.2017.587

[14] Kim G.B, Jung K.H, Lee Y, Kim H.J, Kim N, Jun S, Seo J.B, and Lynch D.A, "Comparison of Shallow and Deep Learning Methods on Classifying the Regional Pattern of Diffuse Lung Disease," *Journal of Digital Imaging,* vol. 31, no. 4, pp. 415-424, 2018. Crossref, https://doi.org/10.1007/s10278-017-0028-9

[15] Ma X, He X, and Tu Z.C, "Prediction of Fatigue–Crack Growth with Neural Network-Based Increment Learning Scheme," *Engineering Fracture Mechanics,* vol. 241, 2021. Crossref, https://doi.org/10.1016/j.engfracmech.2020.107402

[16] Mahmood N, Shahid S, Bakhshi T, Riaz S, Ghufran H, and Yaqoob M, "Identification of Significant Risks InPediatric Acute Lymphoblastic Leukemia (ALL) Through Machine Learning (ML) Approach," *Medical & Biological Engineering & Computing,* vol. 58, no. 11, pp. 2631-2640, 2020. Crossref, https://doi.org/10.1007/s11517-020-02245-2

[17] Zhang X, Zhu X, Zhang N, Li P, and Wang L, "Seggan: Semantic Segmentation with Generative Adversarial Network," *In 2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM),* pp. 1-5, 2018. Crossref, https://doi.org/10.1109/BigMM.2018.8499105

[18] Tan Feigang, "SLBP: An Improved Texture Feature for Pedestrian Detection," *In Proceedings of International Conference on Smart City and Systems Engineering (ICSCSE), IEEE,* pp. 202-204, 2017. Crossref, https://doi.org/10.1109/ICSCSE.2017.57

[19] Zhang W, Shan S, Gao W, Chen X, and Zhang H, "Local Gabor Binary Pattern Histogram Sequence (Lgbphs): A Novel Non-Statistical Model for Face Representation and Recognition," *In Tenth IEEE International Conference on Computer Vision (ICCV'05),* vol. 1, pp. 786-791, 2005. Crossref, https://doi.org/10.1109/ICCV.2005.147

[20] Zeng H, Zhang R, Huang M, and Wang X, "Compact Local Directional Texture Pattern for Local Image Description," *Advances in Multimedia*, 2015. Crossref, https://doi.org/10.1155/2015/360186

[21] Sugave S, and Jagdale B, "Monarch-EWA: Monarch-Earthworm-Based Secure Routing Protocol in IoT," *The Computer Journal,* vol. 63, no. 6, pp. 817-831, 2020. Crossref, https://doi.org/10.1093/comjnl/bxz135

[22] Farid Miar Naeimi, Gholamreza Azizyan, and Mohsen Rashki, "Horse Herd Optimization Algorithm: A Nature-Inspired Algorithm for High-Dimensional Optimization Problems," *Knowledge-Based Systems*, vol. 213, 2021. Crossref, https://doi.org/10.1016/j.knosys.2020.106711

[23] Seyedali Mirjalili, and Andrew Lewis, "The Whale Optimization Algorithm," *Advances in Engineering Software*, vol. 95, pp. 51-67, 2016. Crossref, https://doi.org/10.1016/j.advengsoft.2016.01.008

[24] Narain Ponraj D, Christy E, Aneesha G, Susmitha G, and Sharu M, "Analysis of LBP and LOOP based Textural Feature Extraction for the Classification of CT Lung Images," *In Proceedings of 2018 4th International Conference on Devices, Circuits and Systems (ICDCS),IEEE,* pp. 309-312, 2018. Crossref, https://doi.org/10.1109/ICDCSyst.2018.8605138

[25] Kumar A, and Sodhi S.S, "Comparative Analysis of Gaussian Filter, Median Filter and Denoise Autoenocoder," *In Proceedings of 2020 7th International Conference on Computing for Sustainable Global Development (INDIACom),* pp. 45-51, 2020. Crossref, https://doi.org/10.23919/INDIACom49435.2020.9083712

[26] Khairandish M.O, Sharma M, Jain V, Chatterjee J.M, and Jhanjhi N.Z, "A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images*," Intermediate-Range Ballistic Missile,* vol. 43, no. 4, pp. 290-299, 2021. Crossref, https://doi.org/10.1016/j.irbm.2021.06.003

[27] Mondal C, Hasan M, Jawad M, Dutta A, Islam M, Awal M, and Ahmad M, "Acute Lymphoblastic Leukemia Detection from Microscopic Images using Weighted Ensemble of Convolutional Neural Networks," Arxiv Preprint arXiv:2105.03995, 2021. Crossref, https://doi.org/10.48550/arXiv.2105.03995

[28] Mohapatra S, Patra D, and Satpathi S, "Image Analysis of Blood Microscopic Images for Acute Leukemia Detection," *In IEEE International Conference on Industrial Electronics, Control and Robotics*, pp. 215-219, 2010. Crossref, https://doi.org/10.1109/IECR.2010.5720171

[29] Qian J, Yi J, Zhang J, Cheng Y, and Liu J, "An Entropy Weight-Based Lower Confidence Bounding Optimization Approach for Engineering Product Design," *Applied Sciences,* vol. 10, no. 10, pp. 3554, 2020. Crossref, https://doi.org/10.3390/app10103554

[30] Hemanth D.J, "EEG Signal Based Modified Kohonen Neural Networks for Classification of Human Mental Emotions," *Journal of Artificial Intelligence and Systems,* vol. 2, pp. 1-13, 2020. Crossref, https://doi.org/10.33969/AIS.2020.21001.

[31] Rawat W, and Wang Z, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," *Neural Computation*, vol. 29, no. 9, pp. 2352-2449, 2017. Crossref, https://doi.org/10.1162/NECO_a_00990

[32] ALL-IDB Acute Lymphoblastic Leukemia Image Database for Image Processing, 2022. [Online]. Available: https://homes.di.unimi.it/scotti/all/

[33] Tian Y, Krishnan D, and Isola P, "Contrastive Representation Distillation," arXiv Preprint arXiv:1910.10699, 2019. Crossref, https://doi.org/10.48550/arXiv.1910.10699