International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Intellectual Acute Lymphoblastic Leukemia (ALL) Detection Model for Diagnosis of Blood Cancer from microscopic images using Hybrid Convolutional Neural Network

Anuj Sharma, Bala Buksh

Abstract: ***Blood cell malignantly growth has been accounted for to be one of the most transcendent types of disease maladies. ALL (Acute Lymphoblastic Leukaemias) is the malignant types of blood cancer and their detection and classification in earlier stage is biggest issue. Automatic detection and classification of ALL from microscopic images is a challenging and intellectual assignment in medical science. Existing techniques for ALL detection and classification are an understandable alternative for real-time dermoscopic data analysis. Existing microscopic image processing approaches are unable to analyze the ALL data with non-stationary nature. In this perspective, the focus of this research is to design hybrid Convolutional Neural Network (CNN) architecture by utilizing Firefly Optimization Algorithm (FOA/FFA) to detect the ALL from microscopic images of human blood cell into malignant or normal blood cell. Methods: For training and testing of proposed ALL Detection and Classification (ALL-DC) Model, Standard ALL-IDB (Acute-Lymphoblastic-Leukaemias Image Database for Image Processing) is used with hybrid CNN architecture based on the FOA. Here, Histogram of Oriented Gradients (HOG) descriptor with FOA is used as feature extraction and selection mechanism from the Region of Blood Cell (ROBC).Feature extraction approach plays an important responsibility to classify lots of blood diseases. On the way to achieve this goal, we proposed ALL-DC model that combines recent developments in deep learning with fuzzy based CNN structure and for ROBC segmentation, hybridization of K-means segmentation algorithm with FOA that are capable to segment the accurate blood cell region from microscopic images. Using k-means segmentation technique, the foreground and background component is separated into two regions and after that to improve the segmentation results; FOA is used with the novel concept of image enhancement approach.***

Results: The proposed ALL-DC system is evaluated using the largest publicly accessible standard ALL-IDB dataset, containing 600 training and 400 testing microscopic images. When the evaluation parameters of proposed work is compared with a number of other state-of-art schemes, the proposed scheme achieves the most excellent performance of 98.5% in terms of accuracy which also known as area under the curve (AUC) in differentiating ALL from benign cell using only the extracted and optimized HOG feature. Conclusion: When the proposed model is tested on different microscopic images, evaluation parameters is calculated and compared with a few other state-of-art methods and we obtained the proposed method achieves the best performance in terms of classification accuracy. ALL-DC model is implemented and constructed using the concept of Image Processing and Neural Network Toolbox within MATLAB Software.

Keywords: ALL Detection and Classification Model-means, Histogram of Oriented Gradients Descriptor, Firefly Optimization Algorithm, Fuzzy Logic, Convolutional Neural Network

I. INTRODUCTION

Human blood samples visual assessment is a most important decisive factor for the detection and classification of leukemia from microscopic images [1]. There are two different types of Leukemia first in known as Acute Lymphoblastic Leukemia (ALL) and second is Acute Myeloid Leukemia (AML) which can go ahead to human death if not treated at the accurate time [2].Myeloid organs of human are affected by AML, whereas bone marrow is affected by the ALL types of leukemia. The collection of abnormal white blood cells (WBCs) is the main reason of ALL and it is a most hazard hematopoietic cancer disease. The skirmishing potential of the human body is reduced with the increase in the number of malignant WBCs and the foreign material gets diminished. Detection and classification of ALL in early stage can considerably get better chance of recovery, particularly in the case of 5 to 10 years old children [3]. The detection and categorization of blast blood cell (also known as unhealthy WBCs) in the human bone marrow is also an imperative footstep for the prevention from ALL hematopoietic disease. For detection and classification of the proper stage of the ALL, blood cell blast percentage is a major concern and it is also cooperative in the appropriate treatment of the cancer patients. Three distinctive types of ALL disease are characterized according to FAB (French-American-British) standard based on their structure and the morphological differences among the lymphoblast [4]. The detection and classification of the ALL disease highly depends upon the rightness of the segmentation of human blood cell region (Region of Interest—ROI) from microscopic images for hematologists and pathologists.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Mr. Anuj Sharma,* Ph. D from Career Point University, Kota (Rajasthan).
Dr. Bala Buksh , Ph. D. (Comp. Sc.) - Birmingham, United Kingdom.

To support the medical science and accurate detection with classification of ALL, CAD (Computer-Aided Diagnosis) is a basic necessity [5]. The most important step in a CAD based ALL Detection and Classification (ALL-DC) model is to generate a set of unique features of ROI of WBCs which will help to classify the cells as healthy or affected. The ROI of WBCs is shown in the figure 1 with the illustration of marginal blood smear having a lymphocyte and a lymphoblast.

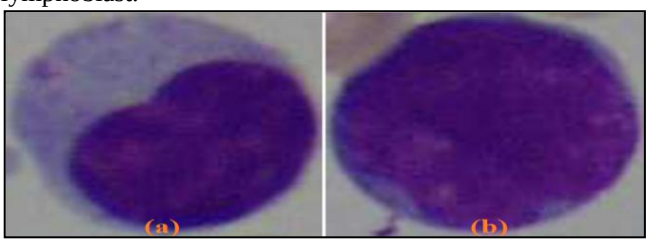


Figure 1: Peripheral blood smears

Above figure 1 demonstrate the blood smear where (a) represents a lymphocyte (healthy) and (b) represents lymphoblast (affected) peripheral blood smears. The morphological, statistical, and textural features are the most specialized feature of a normal blood cells and it is also compulsory to detect and categorize other blast types. In this paper, we have proposed Histogram of Oriented Gradients (HOG) based feature in microscopic images using Firefly Optimization Algorithm (FOA/FFA) and the designed system in known as ALL-DC model. The extracted and optimized HOG feature defines the characteristic of a microscopic ALL image based on the hybrid segmentation approach using K-means with FOA.

Inspiration & Contributions: Human blood cell image is types of microscopic image and it is generally used for detection and classification of the myeloid organs and disease in bone marrow which is caused by the ALL. The most important function of this research is to improve an intellectual ALL-DC model which can help to human utilizing swarm intelligence based optimization approach in the Artificial intelligence/Deep learning mechanism. Because the amount of demise in increases time by year due to ALL problems and the main reason of this death is lack of better detection and classification model availability. From these types of challenging task, research presents an intellectual ALL-DC for diagnosis of blood disease from microscopic images using hybridization of HOG feature with FOA and CNN. In straightforward words, this paper makes the following contributions.

- We used pre-processing of microscopic images using image enhancement approach to improve the quality of images.
- We developed an improved blood cell region segmentation approach (FOA based K-means algorithm) for microscopic images. So the HOG descriptor can extract only relevant feature set and then FAO again used to select best feature set among them.
- To train the ALL-DC model, fuzzy based CNN classifier is used which helps to achieve better detection and classification accuracy.
- At the last of evaluation, a comparative analysis is done to validate the performance of system based on existing work.

Rest of the research paper is organized as follows. The related work and their analysis are described in the Section 2 which consider as background survey. Section 3 summarizes the proposed methodology for the ALL-DC model and Section 4 of article presents the dataset description with a relative comparison of the proposed methodology. The conclusion and future scope of this research is given in Section 5.

II. BACKGROUND SURVEY

In this section, we present the survey of existing work based on the ALL detection and classification using different algorithms and techniques.

Saif S Al-jaboriy et al. [1] had proposed an acute lymphoblastic leukemia segmentation utilizing neighborhood pixel data. Creators have received diverse computational techniques to distinguish the idea of impact cells; be that as it may, these strategies are unequipped for precisely fragmenting leukocyte cells because of some real weaknesses, for example, absence of complexity among articles and foundation, affectability to dim scale, affectability to clamor in pictures, and huge computational size. In this manner, it is imperative to build up a better than ever method for leukocyte cell segmentation. In the present research, a programmed leukocyte cell segmentation procedure was presented that depends on machine learning approach and picture preparing method. They had proposed a programmed leukocyte cell segmentation procedure dependent on a machine learning approach and picture preparing strategy. The segmentation of WBCs was performed dependent on factual highlights chosen by GA, and ANN was embraced to channel all ROI and non-ROI. ***Vasundhara Acharya and Preetham Kumar [2]*** had planned a model for detection of acute lymphoblastic leukemia utilizing picture division and information mining calculations. Creator's intends to review distinctive PC supported framework methods used to portion the blood smear picture. The essential target here is to get learning from the various techniques utilized for separating highlights from white platelets and build up a framework that would precisely section the blood smear picture by conquering the downsides of the past works. The goal referenced above is accomplished in two different ways. The presentation of the model is assessed utilizing the test pictures re-colored with different stains. The proposed calculation accomplished a general exactness of 98.6%. The promising outcomes demonstrate that it very well may be utilized as a symptomatic instrument by the pathologists. An epic methodology for leukemia detection is proposed. In examination with different calculations, the creative methodology separates the cytoplasm from the pictures without including manual editing system. Fringe cells that were situated at extraordinary corners are not precisely managed and need to exact extraction of outskirt cells and furthermore the sub arrangement of acute lymphoblastic leukemia into its particular phenotypes. ***Sonali Mishra et al. [3]*** structured a texture highlight put together arrangement with respect to minuscule blood smear for acute lymphoblastic leukemia detection. They present a powerful plan for grouping of the ordinary white platelets from the influenced cells in a minute picture.

The proposed technique at first pre-forms the info pictures utilizing Y part of the CMYK picture and a triangle strategy for thresholding. They have proposed a DOST+PCA+LDA based CAD framework for the detection of ALL. PCA+LDA are tackled to diminish the element of the separated highlights. The huge highlights chose are the provided to a mix of AdaBoost and irregular woodland classifier for accomplishing better execution.

Sachin Kumar et al. [5] structured a computerized model for detection of acute leukemia utilizing k-mean bunching calculation. Analyst's exhibits a calculation for mechanized picture based acute leukemia detection frameworks. The strategy actualized utilizes essential improvement, morphology, sifting and portioning system to extricate district of enthusiasm utilizing k-implies grouping calculation. This manual examination procedure is tedious and blunder inclined, therefore a PC based framework for computerized detection of ALL may give an assistive demonstrative apparatus to pathologists. The proposed calculation accomplished an exactness of 92.8% and is tried with Nearest Neighbor (k-NN) and Naïve Bayes Classifier on the dataset of 60 tests. ***Hossain Abedy et al. [5]*** proposed a leukemia expectation from minute pictures of human platelet utilizing pig highlight descriptor and calculated relapse. The principle challenge of this exploration is to accomplish a worthy precision with an adaptable technique. Nonetheless, information irregularity, missing qualities and information inadequacy made the scientists' activity considerably more troublesome. In these outcomes, creators proposed a versatile leukemia expectation strategy dependent on an openly accessible ALL-IDB dataset utilizing the HOG highlight descriptor and Logistic Regression. At first, they proposed strategy utilized watchful edge indicator and commotion decrease administrators to distinguish the careful state of Lymphocytes. At that point, Principal Component Analysis (PCA) is connected to the identified picture shapes. The PCA lessens the information measurements without losing any profitable data and in this manner enormously limits the short time later computational expense. At last, a classifier based model is delivered for unexpected occasions and it is tried. The outcomes are approved utilizing n-overlay cross-approval system, where n is a positive whole number more prominent than or equivalent to three. The greatest normal exactness of the proposed model is 96% which is a lot higher than the cutting edge plans.

Based on the survey of related, we conclude some most essential key point which helps to short out existing problem of ALL detection and classification system. Main concern in microscopic ALL image interpretation models is the elimination of unwanted noise to improve the quality and segmentation of blood cell ROI which can helps to achieve better accuracy of ALL-DC system. Therefore, it is concluded that, image enhancement and segmentation are mandatory for the proposed model. The enhancement approaches used in the existing systems are directly used on the entire image but right way to use it with blood cell indentified portion. After that HOG descriptor and FOA is applied and to train the model, CNN is used as a classifier. These combination in the proposed work helps to achieve better detection and classification accuracy of proposed ALL-DC model and their methodology is described on the next section.

III. PROPOSED METHODOLOGY

The proposed optimized artificial intelligence technique based detection of acute lymphoblastic leukaemia (ALL) cells using microscopic image consists of five most important steps. Firstly, microscope based blood cell images will be captured and a database will be prepared for ALL, at image acquisition step. In this step we used the ALL-IDB dataset for training and classification of system. Thereafter various image pre-processing techniques will be applied to improve image quality and suitable segmentation technique will be applied to separate out background from the image. Then features will be extracted from the segmented region and reduced further (if required) to feed as input to the model. Finally, detection and classification will be performed using soft computing models to produce desired output for the ALL-DC system (Figure 2).

Figure 2: Steps of Proposed ALL-DC Model

The problem of this research work is to classify the ALL type of cancer from microscopic images using optimization algorithm i.e. Firefly Optimization Algorithm (FOA) and train the system using Convolutional Neural Network (CNN) on the basis of the features extracted from the ROI of blood cell. The subsequent steps demonstrate the variety of phases that need to be accomplished:

Step 1: Design a simulator to simulate the proposed ALL detection and classification system from the microscopic images. To plan a test system we utilized Graphical User Interface (GUI) which gives the user an intelligent situation. When the GUI is made the user need not know anything about the coding segment.

Step 2: Upload the microscopic image database of different classes to train and test the proposed system. In both section, training and testing microscopic image is uploaded for processing. In training the more number of microscopic images is uploaded but in testing single test fill is uploaded. The data uploading process is known as Microscopic Image Data Acquisition (MIDA) and the algorithm of MIDA is given as:

Algorithm: MIDA Algorithm

Input Attributes: N →Number of training and testing images
Output Attributes: Microscopic Imaged

1) **Begin**
2) Define image browsing option for the image acquisition
3) Pathname = Browse (Image format (jpeg, png, bmp, tif), Title of uploading)

4) **For i→ N**
5) Full-path = string concatenation (Pathname, Filename with sequence)
6) Images (i) = Read (Full-path)
7) **End**
8) **Return:** Images (i) as a hand gesture image
9) **End**

Step 3: After the image acquisition we apply pre-processing on the uploaded microscopic images to segment the Region of Interest (ROI) of cell. In proposed mode, pre-processing is done to remove various type of noise that is inherited in the microscopic images to enhance the quality of uploaded microscopic image. In every ALL detection and classification system, selection of popper cell region is the major factor and need to remove extra part from the uploaded microscopic images. Image enhancement is used to get better pixel points of an uploaded microscopic image. In the proposed work, Lighting based Microscopic Image Enhancement (LMIE) technique is used to improve the quality of uploaded microscopic image which helps to minimize the noisy portion from image so that the extracted attributes of microscopic ALL images should be better and appropriate. The LCIE algorithm is given as:

Algorithm: **LCIE**

Input Attributes: Microscopic-Img→Microscopic Image
Output Attributes: Enh-Microscopic Img→Enhanced Microscopic Image

1) **Begin**
2) Microscopic image in double precision, Microscopic-Img = double (Microscopic-Img)
3) [Height, Width, Plane] = size of Microscopic-Img
4) Patch Size = P // selection of the small regions
5) Pad Size = p // to add extra bit to create a border
6) Pmat = Pad array (I, [Pad Size, Pad Size]) // It is blank mask according to the Pad Size
7) **For imgRow → 1 to Height of Microscopic-Img**
8) **For imgCol→ 1 to Width of Img**
9) $\text{Patch} = \text{Pmat}(\text{imgCol}: (\text{imgCol} + \text{PatchSize} - 1), \text{imgRow} : (\text{imgRow} + \text{patchSize} - 1), \text{All})$
10) $\text{Zmat}(\text{imgCol}, \text{imgRow}) = \min(\text{Patch}(:))$
11) **End**
12) **End**
13) A= intensity (Iimg, Zmat, Plane) // Apply intensity adjustment approach
14) CoeffImg=1-Zmat
15) Enh-Img = zeros (size (Img)) // Blank matrix with size of Img
16) **For ind→ 1to Plane of Img**
17) $\text{Eimg}(\text{All rows}, \text{All Columns}, \text{ind}) = \text{A}(\text{ind}) + (\text{Img}(:,:,\text{ind}) - \text{A}(\text{ind}))./\max(\text{CoeffImg}, 0.1)$
18) Where, 0.1 is the coefficient of lighting
19) **End**
20) Structure Img Elements = strel('disk', 3) // for in shape of disk with 3 element
21) CloseImg = imclose (Enh-Img, Structure Img Elements)
22) Image Error = double (CloseImg) - double (Cloth-Img)
23) Dilated Image = (Image Error >5)
24) Enh-Microscopic Img = Dilated Image
25) **Return:** Enh-Microscopic Img as an Enhanced-Microscopic Image
26) **End**

Microscopic image enhancement techniques have been widely used in many medical science applications using the concept of image processing where the subjective quality of microscopic images is important for interpretation. Contrast is an important factor in any subjective evaluation of microscopic image quality. Contrast is created by the difference in luminance reflected from two adjacent surfaces. In other words, image contrast is the difference in visual properties that makes a call distinguishable from other cells and the background in microscopic images. In visual perception, contrast is determined by the difference in the color and brightness of the object with other objects. Our visual system is more sensitive to contrast than absolute luminance; therefore, we can perceive the world similarly regardless of the considerable changes in illumination conditions. Many algorithms for accomplishing contrast enhancement have been developed and applied to problems in image processing. In every recognition system, selection of popper image region is the major factor and need to remove extra part form the images. In pre-processing after enhancement process, segmentation is used to find out the better cell region of interest (ROI) which helps to achieve better detection accuracy. For the selection of microscopic cell ROI from images, FOA based K-means (Optimized K-means) is used which is based on the morphological operations like binarization, thinning, etc. Morphological operation is a collection of non-linear operations related to the shape or morphology of features in an image. Apply Morphological operations on the binary image to find out the exact region of hand within the image using the some basic operations. In the proposed work, morphological operations help to find out the exact cell region and separate the extra region from microscopic images. The used equations for morphological operations are given described in the below section.

Binarization: The binarization process is carried out using the following equation with threshold value.

$$B - Microscopicimg(i,j) = \begin{cases} 1 & if Enh - MicroscopicImg(i,j) \geq Threshold \\ 0 & else \end{cases}$$

As shown in the above equation, B-Microscopic img is binary image and Enh-Microscopic Img is enhanced image with row (i) and columns (j). The binarization algorithm of proposed work is given as:

Algorithm: Binarization Algorithm

Input Attributes: Enh-Microscopic Img → Enhanced Microscopic Image
Output Attributes: B- Microscopic img→Binary

Microscopic Image

1) **Begin**
2) Define Row (R) and Columns (C) of Enh-Microscopic Img
3) **For i=1→R**
4) **For j=1→C**
5) Using equation (4)
6) **If Enh-Microscopic Img (i, j) > Average (Enh-Microscopic Img)**
7) B-Microscopic img(i,j) = 1
8) **Else if Enh-Microscopic Img (i, j) < Average (Enh-Microscopic Img)**
9) B-Microscopic img (i,j) = 0
10) **End**
11) **End**
12) **End**
13) **Return**: B-Microscopic img as a binary image
14) **End**

Area Opening: The area opening is performed to remove the pixels from the boundary of cell region in image so we can find out the well appropriate region. The opening is performed by using the given equation:

$$OpenImage = \begin{cases} \mathbf{0}, & \boldsymbol{Pixels < Threshold} \\ \mathbf{1}, & \boldsymbol{Pixels \geq Threshold} \end{cases}$$

Step 4: Here we proposed K-mean with FOA as an optimization approach and develop a code with objective function for the FOA for minimization of irrelevant area from the segmented ROI of microscopic image. On the basis for morphological operations, Optimized K-means is designed and the algorithm of Optimized K-means is given as:

Algorithm: Optimized K-means (FOA)

Input Attributes: M-Image → Microscopic Image
Output Attributes: ROI→ Cell ROI Microscopic Image

1) **Begin**
2) [R, C, P]=size (M-Image)
3) M-Image =double (M-Image)
4) Number of Part = 2
5) SimgIndex=kmeans (M-Image, Number of Part)
6) SegLabelImg=reshape (SimgIndex, R, C)
7) DataPos=find (SegColLabelImg>0)
8) Data=SegColLabelImg(DataPos)
9) **Initialize FOA parameter** – Iterations (T), Pop Size (S)Lower Bound (LB), Upper Bound (UB), Fitness function
10) Calculate T = Size (M-Image Image)
11) Fitness function:
12) $f(fit) = \begin{cases} 1 & if pixel is less \\ 0 & otherwise \end{cases}$
13) **For → T**
14) $fs = \sum_{i=1}^{P} Data(i)$
15) $ft = \frac{\sum_{i=1}^{P} Data(i)}{Length of feature}$
16) $f(fit)$ = fitness function which define by above given equation
17) $Threshold_{value} = FOA(P, T, LB, UB, N, f(fit))$
18) **End**
19) **While T ~= Maximum**
20) $\boldsymbol{Threshold} = Threshold_{value}$
21) MaskImg=Morphological (SimgIndex, Threshold)
22) Boundaries = bwboundaries (MaskImg)
23) Segmented Cell Region = Boundaries
24) **For i→1: P**
25) ROI = M-Image X Segmented Cell Region
26) **End**
27) **Return:** ROI of Microscopic Image as Cell ROI Microscopic Image
28) **End**

After the Optimized K-means applied on the enhanced dermoscopy skin lesion images, we obtained below given results which are useful in next process of proposed work. Below figure 3 represent the ROI extraction process with pre-processing steps. In figure 3, (a) is the original microscopic image, (b) is the enhanced microscopic image, (c) is the mask of cell ROI, and (e) is the region of cell in microscopic image using the Optimized K-means algorithm with FOA.

Step 5: Extract feature from the ROI of cell portion of microscopic image based on the HOG feature extraction algorithm. After the feature extraction algorithm, a set of feature is return by the HOG algorithm in terms of feature points. The HOG algorithm is given as:

Algorithm: HOG Algorithm

Input Attributes: ROI → Cell ROI Microscopic Image
Output Attributes: Fpoints → Feature points

1) **Begin**
2) Load ROI data of microscopic images
3) **For I →1 to all sets**
4) Gradient Image = ROI (I)
5) Binning =Oriented Histogram (Gradient Image (I))
6) **If Binning need orientation**
7) Orientated Histogram = Binning (I, Angle)
8) **End**
9) Feature_descriptor=All best Orientated Histogram
10) **End**
11) **Return:** Feature_descriptor as Fpoints
12) **End**

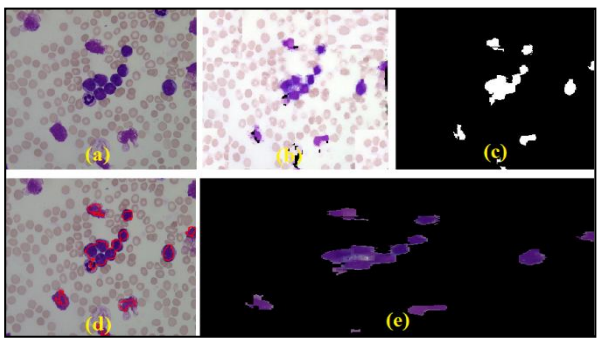


Figure 3: ROI Extraction Process using Optimized K-means (FOA)

After the feature extraction algorithm applied on the cell ROI of microscopic image, we obtained below given results which are useful in training as well as classification process of proposed work.

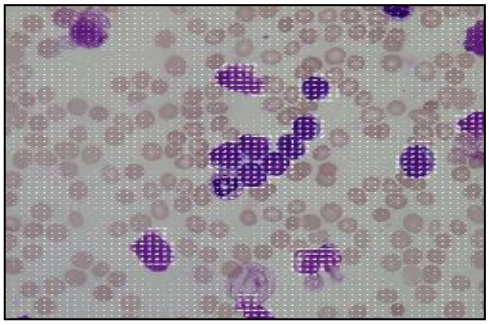


Figure 4: HOG Feature of ROI

Step 6: After the feature extraction we apply FAA as feature optimization technique with fitness function. The algorithm of FFA for feature optimization is written follows:

Algorithm 6: FOA Feature Optimization
Input: Fpoints → Feature points of HOG
Output: OFpoints →Optimized Feature points

1) **Initialize FOA parameter** – Iterations (T), Population Size (S), Lower Bound (LB), Upper Bound (UB) and Fitness function
2) Calculate T = Size (T)
3) Fitness function:
4) $f(fit) = \begin{cases} Selected & if feature is appropriate \\ Rejected & otherwise \end{cases}$
5) **for each data in T from 1 → T length**
6) $fs = \sum_{i=1}^{P} Data(i)$
7) $ft = \frac{\sum_{i=1}^{P} Data(i)}{Length of feature}$
8) $f(fit)$ = fitness function which define by above given equation
9) $Optimized_{Feature}, OFpoints = FFA(P, T, LB, UB, N, f(fit))$
10) **End**
11) **Return:** OFpoints as a set of Optimized Feature points
12) **End**

Step 7: Initialize CNN for classification purpose using two phases, namely, training and testing. After the training of proposed system, we save the trained structure which is use in the classification section to classify the ALL and non ALL data from the macroscopic images. In the testing phase, the test microscopic image is uploaded and repeats the steps from 2 to 5. In the classification section, test microscopic ROI feature is matched with trained CNN structure and return result type and the CNN algorithm is given as:

Algorithm: Fuzzy based CNN Algorithm
Input Attributes: OFpoints → Optimized Feature Key Points as training data (T-Data), Target as categories of ALL type (G) and Number of Neurons (N)
Output Attributes: Classified results with trained CNN structure

1) **Set up the CNN model with basic parameters**: Epochs/Iterations (E), Training performance constraints: Cross Entropy, Gradient, Mutation and Validation, Training methods: Scaled Conjugate Gradient (Trainscg), and Training data division: Randomly
2) **For each set of T**
3) **If T-Data belongs to ALL then**
4) $G(1) = T - data(1)$ based on the fuzzy rule set
5) **Else if Training Data belongs to Non ALL then**
6) $G(2) = T - data(2)$ based on the fuzzy rule set
7) **Else**
8) $G(3) = Extra$
9) **End**
10) Call CNN,
11) Net = patternnet (N)
12) Set the training performance parameters according to the requirements
13) Train the system,
14) Net = Train ($Net, T - Data, G$)
15) Classification Results = simulate (Net, Test cloth key points)
16) **If Classification Results match with trained structure's group**
17) Show classified results in terms of the ALL types
18) Calculate the model performance parameters
19) **End**
20) **Return:** Classification results as a type of ALL with trained CNN structure
21) **End**

Algorithm HOG descriptor is used to extract the important key point of cell ROI which helps to train the proposed model with fuzzy based CNN. The extracted feature is passes to the CNN as an input training data and store which is used in the classification process to classify the type of ALL data with is taken from ALL-IDB Dataset. The structure if used CNN is shown in the figure 5.

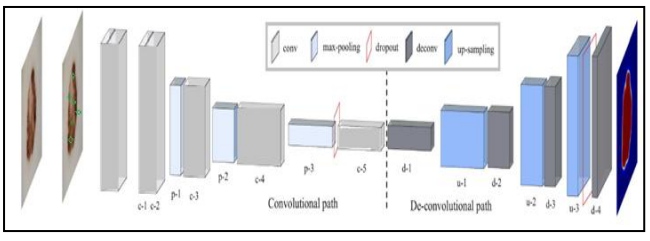


Figure 5: CNN Training Structure

Above figure 5 represents the CPU based CNN model for proposed optimized artificial intelligence technique based detection of acute lymphoblastic leukaemia (ALL) cells using microscopic images with different input, output and hidden layers. The layers of CNN are given as:

Input Layer: It is used to provide input data to CNN for training of proposed model in terms of weight value. Extracted **HOG** key points are consider as input data of CNN and there are **two** types of **HOG** key points are passes to CNN in input layers. For each type of **ALL** attributes, **HOG** key points are extracted.

Grouping of Input Data: Here the concept of fuzzy logic is used to separate the **HOG** key feature points into different target or group. The designed fuzzy interface is described in the below section.

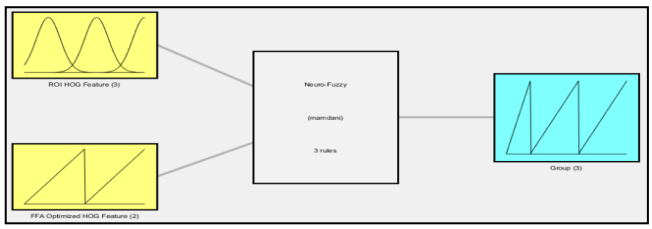


Figure 6: Designed Fuzzy Interface

Above figure 6 represents the designed fuzzy interface of CNN to categorization of data. Here input HOG feature is divided in three different categories like Target 1, Target 2 and Target Extra. The rule set of designed model is given below figure 6.

1. If (ROI HOG Feature is Good) and (FFA Optimized HOG Feature is Selected as Target 1) then (Group is Group 1) (1)
2. If (ROI HOG Feature is Non Good) and (FFA Optimized HOG Feature is Selected as Target 2 either Extra) then (Group is Group 2) (1)
3. If (ROI HOG Feature is Non understandable) and (FFA Optimized HOG Feature is Selected as Target 2 either Extra) then (Group is Extra Group) (1)

Figure 6: Rule sets of proposed Fuzzy interface model

Based on these rule set be designed a rule viewer which is shown in the figure 6 with HOG optimized feature which is used in the CNN training.

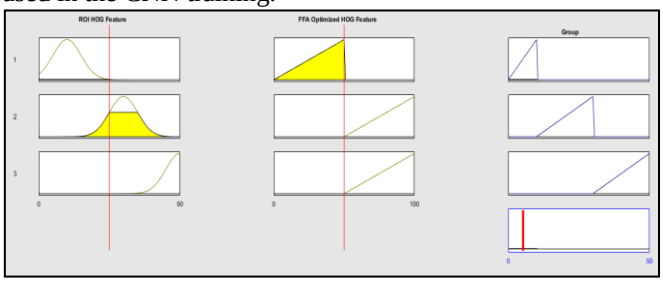


Figure 7: Rule Viewer of Fuzzy Interface

***Convolution Layers*:** In CNN, the convolution layer is a basic element and the objective of convolution is to extract best features from the input **HOG** key points. It consists of a group of learnable square filters which helps to find out the appropriate and unique feature sets. Each filter is applied to the raw values of the **HOG** key points.

Max-pooling Layer: In the CNN architecture of proposed model, convolution layers are followed by sub-sampling layers and act as a unique feature extraction approach. A layer of max-pooling is an alternative of feature selection to increase the chances of extracted feature uniqueness. The output of max-pooling layer is passes to the classification model which is used as a maximum activation value and creates a structure of model and return by output layer.

Output Layer: It is used to return a trained structure by CNN which is used in the classification phase. In the proposed two different categories are retuned by CNN named as 1 for ALL Above figure 9 represents the training performance of fuzzy based CNN and figure (a) represents the plotting of the training, validation, and test based on Cross-Entropy as a and 2 for non ALL type. The CNN structure of proposed model is shown in the figure 8 with their algorithms, progress and their graphical representations.

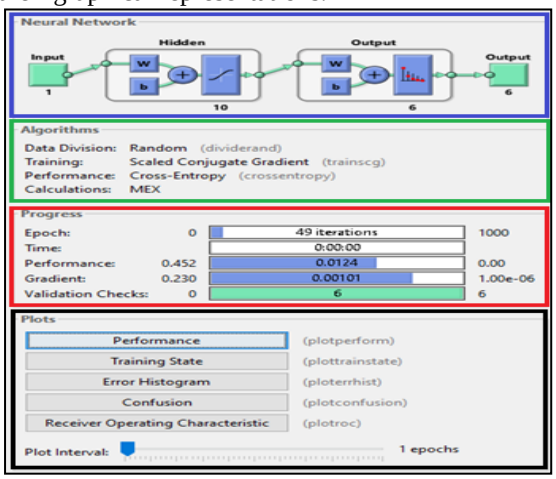


Figure 8: Training using fuzzy based CNN

After uploading all the two types of microscopic images, training is performed using fuzzy based CNN algorithm. For each iteration the dataset folder is uploaded and passed to the intermediate layer like convolutional layer, pooling layer etc and the weight of the HOG feature is adjusted as per the need, and then the image feature is passed to the output layer, where we obtained two number of output each corresponding to the different ALL image type attributes. The performance of fuzzy based CNN is examined on the basis of the following parameters.

- Cross Entropy
- Gradient
- Validation
- Error Histogram
- Confusion Matrix
- ROC of training data

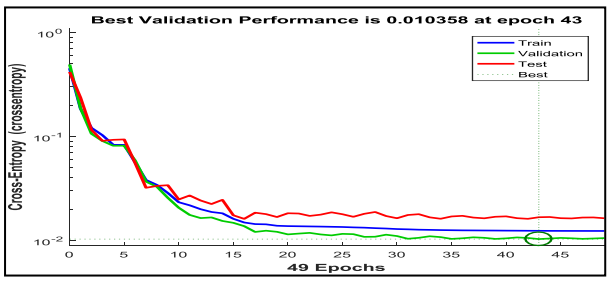
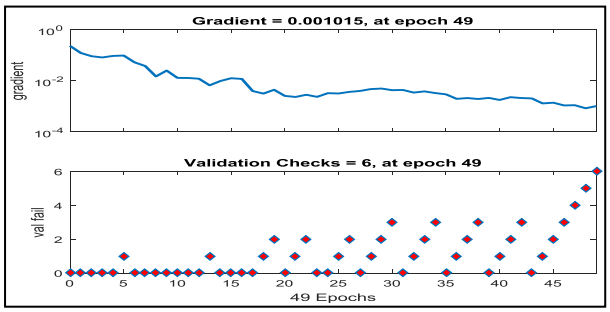


Figure 9: (a) Cross-Entropy and (b) Training State



performance parameters. Basically above figure represents three types of data used by CNN, first is training data which is presented by the blue color line, second is testing data which is represented by the green color line and last is the green color line which denotes the valid data. The best validation of data is represented by circle which is obtained after 43 iterations. Figure (b) represents the validation fails during the training of system in terms of graphically representation. The minimum validation fails is 0 and the maximum validation fails is 6 in CNN and both gradient and validation fails represents the training status with respect to the epochs/iterations.

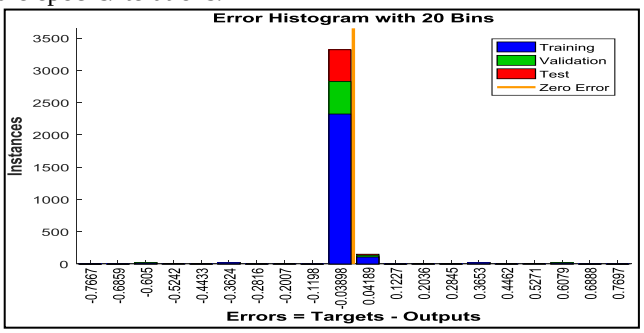


Figure 10: Error Histogram

The bars on the figure 10 represents the bins and the total error that has been measured for the fuzzy based CNN during experiment is lies between -0.7667 (leftmost bin) to 0.7697 (rightmost bin). The entire range is sub-divided into 20 numbers of smaller elements and the width of each bin is measured using the formula written below.

$$(0.7697 - (-0.7667))/20 = 0.0768$$

The training data that comprises of different cloth attributes are represented by vertical bars. For an example for first vertical bin, the height is about 0 with the error value of -0.7667. This indicates that there are two numbers of features and the error of these feature lies in the following range.

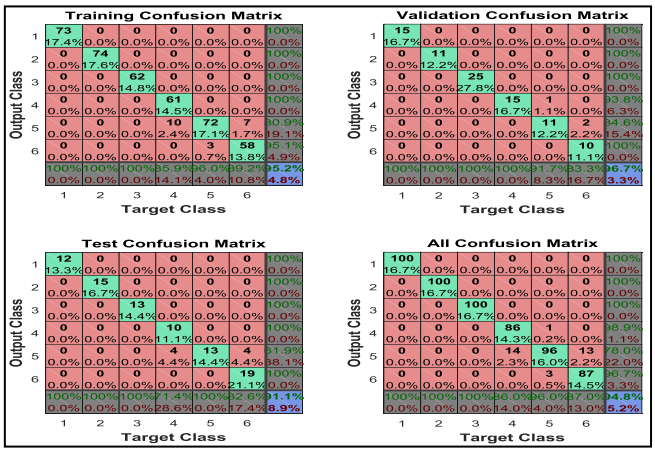


Figure 11: Confusion Matrix

Above figure 11 represents the confusion matrix of fuzzy based CNN training with four different data types such as training, testing, validation and combination of all. From the figure, there are total six types of categories in training data and its present the training of system is good because all confusion matrix show good performances and their ROC is shown in the figure 12.

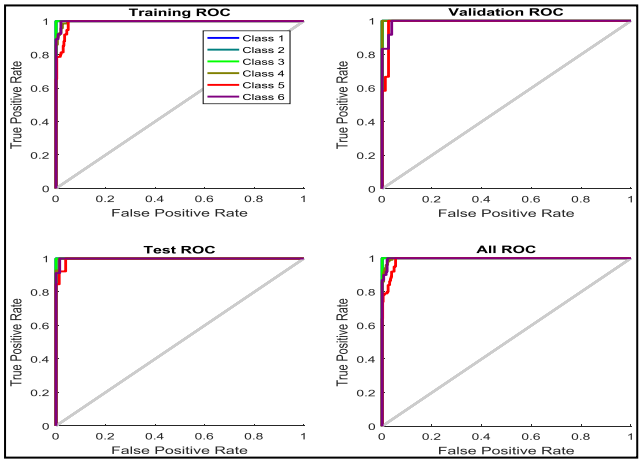


Figure 12: RoC of fuzzy based CNN Training

Above figure 12 represents the ROC curve of fuzzy based CNN training with four different data types such as training, testing, validation and combination of all similar to the confusion matrix.

From the figure, the training of the proposed fuzzy based CNN with optimized HOG key points achieve better performance and training accuracy is also high. Based on the designed methodology, we achieved better efficiency of proposed scenario with use of fuzzy based CNN.

Step 7: At the last, calculate the performance metrics like precision, recall, f-measure, error and accuracy to check proposed system efficiency.

IV. RESULTS AND DISCUSSION

In this section, the simulation results of proposed ALL-DC using a hybrid approach by CNN with FOA are discussed and the efficiency of proposed work is compared with previous ALL detection and classification model.

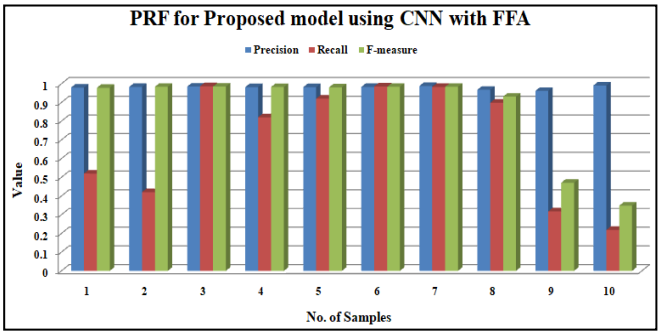


Figure 13: Precision, Recall and F-measure (PRF) for fuzzy based CNN with FOA/FFA

In this section the simulation results for the ALL detection and classification using fuzzy based CNN with FFA is described in the below section with explanation. The above figure 13 represents the comparison graph obtained for (i) Precision (ii) Recall and (iii) F-measure for the classification model. X-axis shows the number of test sample data that are uploaded to detect the type of ALL. Y-axis depicts the value of different computation parameters such as precision, recall and f-measure. From the above graph it is clear that the code is simulated ten times with ten different samples and the average value of precision, recall and F-measure are 0.9802, 0.707 and 0.862 respectively.

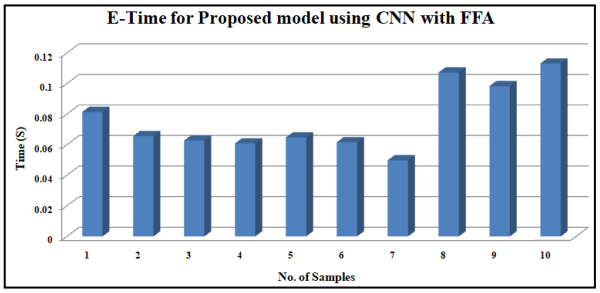


Figure 14: Execution Time for fuzzy based CNN with FFA/FOA

The execution time which is required to detect the ALL types for 10 iterations is shown in figure 14. X-axis and Y-axis signifies the number of test microscopic samples and execution time respectively. From the graph, it is understandable that average execution time required for the proposed ALL detection and classification model using fuzzy based CNN with FFA is 0.113 seconds.

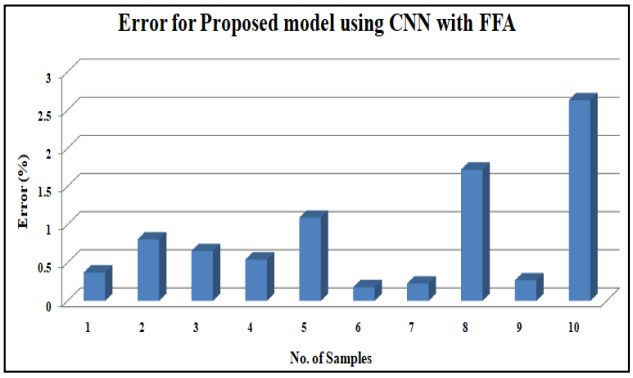


Figure 15: Error Rate for fuzzy based CNN with FFA

The error obtained while detecting the leaf disease is shown in figure 15. The average error measured for 10 tests microscopic ALL data is 0.849 % and it is very low because the use of FOA as feature optimization approach to optimized the HOG feature.

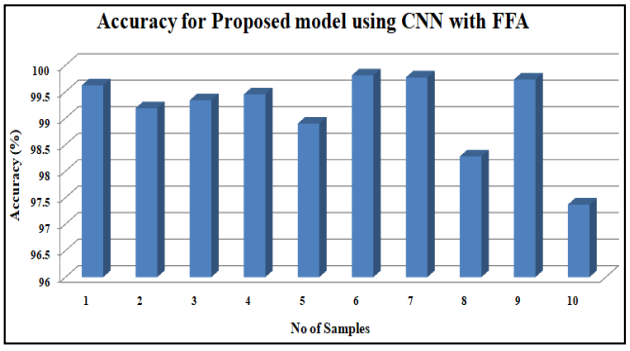


Figure 16: Classification accuracy of proposed ALL-DC using for fuzzy based CNN with FFA

The above figure 16 shows the value of classification accuracy obtained for ten tests microscopic ALL data from different types. The average value of classification accuracy using the concept of CNN with FOA is 99.14 % which represents that the ALL data detected with high accuracy. The main reason to achieve better classification accuracy is the use of feature selection as an optimization technique with CNN. In the proposed model FFA is used as a feat selection algorithm from the HOG key points. The comparison of proposed work with some other existing work, which is considered in scrutiny of proposed work, is described in below table.

Table 2: Comparison of accuracy of proposed work with existing works

Authors	Accuracy (%)
Al-jaboriy [1]	97.01
Vasundhara Acharya [2]	93.15
Sonali Mishra [4]	91.32
Sachin Kumar [9]	95.74
Hossain Abedy [13]	96.35
Proposed work	99.14

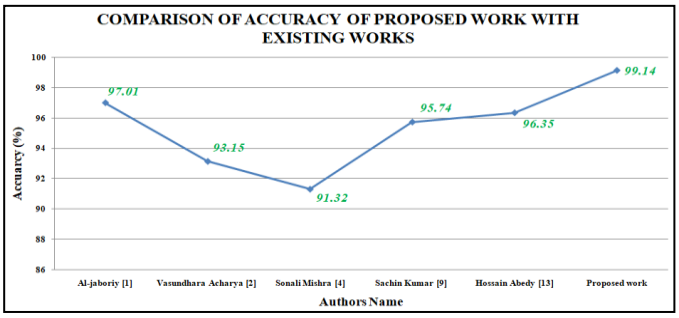


Figure 17Comparison of accuracy of proposed work with existing works

Figure 17 represents the comparative exploration of existing work based on the classification accuracy. From the figure we observe that the accuracy achieve by proposed work is better than other author by using the hybridization of FOA as an optimization technique with fuzzy based CNN classifier.

V. CONCLUSION AND FUTURE WORK

In this research work, an optimized artificial intelligence technique based detection of ALL cells using microscopic image using a convolutional neural network based on the hybrid segmentation technique is proposed which is known as ALL-DC model. It provides a detailed view of the different applications and potential challenges of segmentation and classification of cancer from microscopic images which a difficult task in medical science. For the detection and classification of ALL type of cancer, segmentation of cell region is major task which is known as the cell ROI and it is performed by hybridization of K-means with FOA. After that, in this research, we present a fuzzy based CNN with FFA based HOG descriptor for the segmentation and classification of ALL and Non-ALL data and ALL-IDB-2016 dataset is used for validation of proposed model. Utmost classification accuracy is reported when proposed work is simulated on ALL-IDB dataset using the concept of CNN. With proposed method, the accuracy is 99.14% whereas with the existing work, the accuracy is less. In the future work, hybridization artificial intelligence techniques can be used to enhance the performance of ALL detection and classification from microscopic images. Feature extraction algorithms can also be used to classify the region of cell with more accurately so that the efficiency of ALL detection system can be improved.

REFERENCES

1. Al-jaboriy, Saif S., et al. "Acute lymphoblastic leukemia segmentation using local pixel information." Pattern Recognition Letters 125 (2019): 85-90.
2. Acharya, Vasundhara, and Preetham Kumar. "Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms." Medical & biological engineering & computing (2019): 1-29.
3. Mishra, Sonali, Banshidhar Majhi, and Pankaj Kumar Sa. "Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection." Biomedical Signal Processing and Control 47 (2019): 303-311.
4. Kumar, Sachin, Sumita Mishra, and Pallavi Asthana. "Automated Detection of Acute Leukemia Using K-mean Clustering Algorithm." Advances in Computer and Computational Sciences. Springer, Singapore, 2018. 655-670.
5. Abedy, Hossain, et al. "Leukemia Prediction from Microscopic Images of Human Blood Cell Using HOG Feature Descriptor and Logistic Regression." 2018 16th International Conference on ICT and Knowledge Engineering (ICT&KE). IEEE, 2018.
6. Tuba, Eva, et al. "Acute Lymphoblastic Leukemia Cell Detection in Microscopic Digital Images Based on Shape and Texture Features." International Conference on Swarm Intelligence. Springer, Cham, 2019.
7. Patil, Neha Dattatray, and Soniya Bastawade. "Leukaemia Detection using Image Processing Techniques." International Journal Of Emerging Technology and Computer Science 3.3 (2018): 29-33.
8. Laosai, Jakkrich, and Kosin Chamnongthai. "Classification of acute leukemia using medical-knowledge-based morphology and CD marker." Biomedical Signal Processing and Control 44 (2018): 127-137.
9. Ben-Suliman, Karima, and Adam Krzyżak. "Computerized Counting-Based System for Acute Lymphoblastic Leukemia Detection in Microscopic Blood Images." International Conference on Artificial Neural Networks. Springer, Cham, 2018.
10. Sahlol, Ahmed T., Ahmed M. Abdeldaim, and Aboul Ella Hassanien. "Automatic acute lymphoblastic leukemia classification model using social spider optimization algorithm." Soft Computing 23.15 (2019): 6345-6360.
11. Nikitaev, V. G., et al. "Textural characteristics of bone marrow blast nucleus images with different variants of acute lymphoblastic leukemia." Journal of Physics: Conference Series. Vol. 945. No. 1. IOP Publishing, 2018.
12. Shafique, Sarmad, and Samabia Tehsin. "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks." Technology in cancer research & treatment 17 (2018): 1533033818802789.
13. Moshavash, Zeinab, Habibollah Danyali, and Mohammad Sadegh Helfroush. "An Automatic and robust decision support system for accurate acute leukemia diagnosis from blood microscopic images." Journal of digital imaging 31.5 (2018): 702-717.
14. Mishra, Sonali, Banshidhar Majhi, and Pankaj Kumar Sa. "GLRLM-Based Feature Extraction for Acute Lymphoblastic Leukemia (ALL) Detection." Recent Findings in Intelligent Computing Techniques. Springer, Singapore, 2018. 399-407.
15. Rehman, Amjad, et al. "Classification of acute lymphoblastic leukemia using deep learning." Microscopy Research and Technique 81.11 (2018): 1310-1317.
16. Talukdar, Nur Alom, Daizy Deb, and Sudipta Roy. "Automated blood cancer detection using image processing based on fuzzy system." IJARSC, ISSN-2277-128X 4 (2014).
17. Jothi, G., et al. "Rough set theory with Jaya optimization for acute lymphoblastic leukemia classification." Neural Computing and Applications (2018): 1-20.
18. Matek, Christian, et al. "Human-level recognition of blast cells in acute myeloid leukemia with convolutional neural networks." bioRxiv (2019): 564039.

AUTHORS PROFILE

Mr. Anuj Sharma is pursuing Ph. D from Career Point University, Kota (Rajasthan). He received MCA master degree from AMC Engineering College Bangalore under Visvesvaraya Technological University, Belagavi, Karnataka. He obtained his Bachelor degree in Computer Application from MCM DAV College, Kangra (H.P) under Himachal Pradesh University, Shimla (H.P). He is having life time membership of Indian Science Congress Association (ISCA). His research area of interests is in the field of Image Processing, Computer Programming, Artificial Intelligence and Database.

Dr. Bala Buksh pursued his **Ph. D.** (Comp. Sc.) - **Birmingham, United Kingdom.** He was Director in R.N Engineering College till March, 2017. Superannuated from ONGC as General Manager after rendering 30 years of service. Implemented National level projects in the areas of Information Technology for Oil Industry. This include, Real Time (SCADA) Project for Bombay Offshore Production Platforms, EPINET (Exploration & Production Network) system, ERP system, LIBNET system connecting 35 Libraries of ONGC for sharing resources, Drilling Operation Monitoring Information System and Digitization & soft copying of E&P Physical Intellectual Assets for sharing of information across the Organization. Awarded by CMD. Had been head of various offices including Chief EPINET Coordinator. Life member of CSI and has been and Chairman and Vice Chairman of Computer Society of India at Dehra Dun & Vadodara. Worked as **General Manager (Projects**) at Educational Consultants India Ltd., New Delhi (**under Ministry of HR**) and headed about 40 various projects operating in India and African Countries.