ODRNN: Optimized Deep Recurrent Neural Networks for Automatic Detection of Leukaemia

K. Dhana Shree

dhanu1304@gmail.com

Sri Ramakrishna Engineering College

S. Logeswari

Karpagam College of Engineering

Research Article

Keywords: Leukaemia, deep learning, Convolutional Neural Networks, Optimized Deep Recurrent Neural Network and Red Deer Optimization

Posted Date: January 16th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3857793/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Signal, Image and Video Processing on March 16th, 2024. See the published version at https://doi.org/10.1007/s11760-024-03062-y.

ODRNN: Optimized Deep Recurrent Neural Networks for Automatic Detection of Leukaemia

[1*]K.Dhana Shree, [2] S.Logeswari

[1]Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India

[2]Department of Information Technology, Karpagam College of Engineering, Coimbatore, Tamil Nadu, India

[*]Corresponding author: dhanu1304@gmail.com

Abstract: Leukaemia, a kind of cancer that may occur in individuals of all ages, including kids and adults, is a significant contributor to worldwide death rates. This illness is currently diagnosed by manual evaluation of blood samples obtained using microscopic imaging, which is frequently slower, lengthy, imprecise. Additionally, inspection under a microscope, leukemic cells look and develop similarly to normal cells, making identification more difficult. Convolutional Neural Networks (CNN) for Deep Learning has provided cutting-edge techniques for picture classification challenges throughout the previous several decades, there is still potential for development with regard to performance, effectiveness, and learning technique. As a consequence, the study provided a unique deep learning approach known as Optimized Deep Recurrent Neural Network (ODRNN) for identifying Leukaemia sickness by analysing microscopic images of blood samples. Deep recurrent neural networks (DRNN) are used in the recommended strategy for diagnosing Leukaemia, then the Red Deer Optimization algorithm (RDOA) applies to optimize the weight gained by DRNN. The mass of DRNN from RDOA will be tuned on the deer roaring rate behavior. The model that has been proposed is evaluated on two openly accessible Leukaemia blood sample datasets, AML, ALL_IDB1 and ALL_IDB2. It is possible to create an accurate computer-aided diagnosis for Leukaemia malignancy by using the proposed deep learning model, which shows encouraging results. The research work uses statistical metrics related to disease including specificity, recall, accuracy, precision and F1 score to assess the effectiveness of the proposed model for identification and classification. The proposed method achieves highly impressive results, with scores of 98.96%, 99.85%, 99.98%, 99.23%, and 99.98%, respectively.

Keywords: Leukaemia, deep learning, Convolutional Neural Networks, Optimized Deep Recurrent Neural Network and Red Deer Optimization

1. Introduction

Lympho-haematopoietic cancers make a significant contribution to cancer cases in India. Specifically, men comprise 9.5% of all malignancies, whereas women account for 5.5%, as reported in reference [1]. Population-based surveys in India show 0.8/1,000,000 Leukaemia cases in Barshi, a rural area of Maharashtra, and 5/1,000,000 in Delhi. Beginning in the cancellous (marrow), Leukaemia creates an unusually high number of blood cells. These blood cells, also called as immature blood cells (blast cells) or Leukaemia cells, are yet to finish their maturation. Symptoms include bruising, bleeding, bone stiffness, fatigue, fever, and an increased risk of infection. The deficiency of normal blood cells is responsible for these symptoms. Typically, a cancellous sample or blood tests are used to determine the examination [2]. Scientists are unaware of what causes Leukaemia. The phenomenon seems to result from a confluence of hereditary and environmental factors.

The reliable proof of malignancy WBC's in the early stages of sickness at an affordable price is a serious concern in medical examination. The equipment for flow cytometry are not available in laboratory medical facilities, and the procedures are complex [3]. A very common kind of blood cancer is Leukaemia. That occurs in individuals of all ages, particularly children. This abnormal occurrence in excessive Stimulation of blood cells and premature development, and it can affect the haematids (RBC), the marrow of the bone, and the immune system. Leukaemia accounted for over 3.5 percent of all newly identified malignancies in the United States in 2018, including over 60,000 new instances recognized [4]. Lymphoblast's, or abnormal white blood cells, migrate via the circulation to various organs, such example the pancreas, brain, kidneys, liver, where they metastasis to essential internal tissue [5].

Haematologists at transplanted cells centres can identify and diagnose several types of Leukaemia using microscopic images. The microscope slide is appropriately coloured, certain types of Leukaemia are easy to detect and identify; nevertheless, discovering underlying Leukaemia takes more technology. One of the most prevalent forms of Leukaemia is seen on stained slides in Figure 1. AML, which stands for acute myeloid Leukaemia, is

shown in Figure 1a. which stands for acute lymphoblastic Leukaemia, is shown in Figure 1. The four most frequent types of Leukaemia are 1b, CML (chronic myeloid Leukaemia) shown in Fig. 1c, and CLL (chronic lymphocytic Leukaemia) shown in Fig. 1d [4].

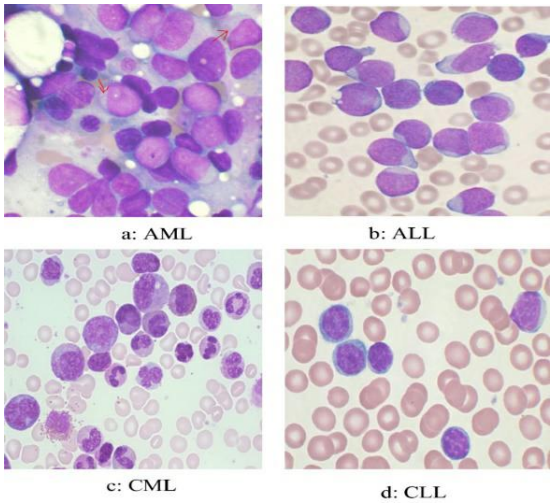

Figure 1. The four major forms of Leukaemia [4]

Haematologists utilize microscopes to examine blood smears to make the basic diagnosis of Leukaemia [6]. The amount of various blood cell types presents, and their morphologic properties are used to classify Leukaemia. A sudden increase in number of underdeveloped white blood cells, Leukaemia is indicated by a decrease in the quantity of normal blood cells. Pathologists might next propose a marrow of bone testing to and diagnose the kind of Leukaemia. Critical cytochemical examination, cytometry flow, chromosomal analysis, and other methods are needed to typical validation and identification Leukaemia's other types. These sophisticated procedures are time-consuming and expensive, and they are impacted by the medical team's expertise and competency in sample collection, preparation, and testing [7].

A conventional examination would also be problematic for the sophisticated laboratory testing. Pathologists may apply image processing techniques to examine microscopic pictures in order to establish an accurate and timely diagnosis of Leukaemia. Previously, many researchers concentrated on automated methods for identifying a wide range of pathological illnesses, including malaria, allergies, lymphoma, and Leukaemia. The initial automated strategies in the literature concentrated on leukocyte identification and categorization [8], with systems for detecting other disorders coming later. When compared to various advanced clinical procedures, image processing-based alternatives are simpler, faster, and less costly. There have been several studies on computer-aided Leukaemia diagnosis, with researchers concentrating on various ways to identify and diagnose Leukaemia [9].

Image filtering, the process of segmentation, extraction of feature, and categorization using Machine Learning (ML) techniques such as Support Vector Machine (SVM) and Artificial Neural Networks (ANN's) are often used in automated detection-based studies for identifying and classifying cases of Leukaemia available in the literature. Pattern recognition algorithms are classified into two types: unsupervised and supervised learning. This area includes supervised learning such as linear regression, SVM, ANN, and others, as well as unsupervised approaches like as grouping by k-means, Fuzzy C-Means (FCM), and more. There has been few research that use Deep Learning (DL) for automated Leukaemia categorization and recognition.

Thresholding, region-based techniques, the case of clustering, watershed segmentation, active contours, ANN-based methods, and more are among the segmentation methods accessible in the scientific literature. The authors' evaluation [10] of the segmentation algorithms utilized to separating microscopic stain pictures in Leukaemia diagnosis and similar studies indicated that a large number of relevant publications (about 41%) employed clustering-based image segmentation. However, all forms of separation techniques are hard to implement since they need considerable calculations. The suggested study attempts to detect leukemic cells in contrast to healthy

cells without the use of such segregation methods. The planned study for automated identification of Leukaemia will make the following significant contributions:

- Automated computerized assistance Deep Learning Networks are being used to identify Leukaemia in general, rather than sophisticated and computationally expensive segmentation and Extracting features methods.
- This study used microscopic images of blood samples to construct a unique deep learning algorithm called Optimized Deep Recurrent Neural Network (ODRNN) for diagnosing Leukaemia. The recommended approach for detecting Leukaemia makes use of a deep recurrent neural network (DRNN), and to optimize the weight from DRNN, the Red Deer Optimization (RDOA) method is used. The weight of the DRNN from RDOA will be modified based on the behavior of the deer roaring rate.

As indicated, this article continues Section 1.2 reviews pertinent material on ML and DL Leukaemia detection. Sect.2 describes methods as well as resources, including an overview of current work and an overview of CNNs and pretrained networks. Sect. 3 explains the outcomes, and sect.4 gives a summary for findings. Sect. 5 presents the findings of the current study.

2. Related work

Das and Meher [11] established a unique probability-based weight parameter that is important for effectively combining mobilenetv2 with resnet18 while keeping the advantages of each method. Its efficiency is proven through the utilization of two standard datasets: ALLIDB1 and ALLIDB2. According to the experimental findings, the recommended method obtains the maximum accuracy in the ALLIDB1 and ALLIDB2 datasets, with corresponding values of 99.39% and 97.18% (training rate of 70% and testing rate of 30%). With 50% training and 50% testing, it also achieves the highest accuracy in ALLIDB1 and ALLIDB2 datasets, with results of 96.00% and 97.92%, respectively. Additionally, it performs better in both datasets than contemporary transfer learning-based algorithms in terms of F1 score, ROC (receiver operating characteristic), specificity, sensitivity, dependability, and accuracy. In any case, massive databases are typically necessary to achieve exceptional results.

Agustin et al. [12] proposed a two-phase Artificial Neural Network technique paired with Particle Swarm Optimization to identify immature WBC in every patient. The suggested technique begins with a binary categorization of lymphoid cells and progresses to a binary categorization of lymphoblast cells. To construct the suggested model, five peripheral blood specimens from the Sarbjito Hospital ALL databases were employed. The proposed technique includes preliminary data processing, feature selection, feature extraction, and two-stage sorting. The implementation of the technique is equated to the performance of multiple-class neural networks with back propagation classifying (multiclass NN-BP) and multiclass neural networks particle swarm optimization (multiclass NN-PSO). The findings reveal that the aimed approach outperforms the other models evaluated in respect to correctness (86.92%) and may be developed and applied in different scenarios. However, based on the label encode method employed, they could not be sufficiently read for data that is categorical.

Acharya et al. [13] described a novel computer-aided diagnostic technique for segmenting blood smear images and assessing AML stage. This paper's technique is broken down into many stages: image acquisition, segmentation of the image, extraction/selection of features, and classifications. The method was evaluated on two sets of images: a total of two hundred microscopic peripheral blood cell images from the collection, together with eight hundred blood smear images from Kasturba Medical College Manipal, were used for the development of automated identification methods. Multiple lobes and nucleated red blood cells (nrbcs) may be seen, together with Auer rods and granules. The info-gain attribute evaluation and ranker search techniques are used to choose features. However, some of these systems' major flaws, like poor efficiency, learning technique, accurate detection rate, and more, need to be solved.

Claro et al. [14] analyse the influence of frequently used strategies in deep learning-based CAD systems, including augmenting data, as well as multilevel and ensemble setups. Three binary classification tasks and two multiclass classification questions were included in the examination. The study used 3,536 photos from 18 datasets to determine if data augmentation strategies increase the efficiency of convolutional neural networks (CNNs). Furthermore, the classification results increase when CNNs are combined. The ensemble design outperformed the multilayer configuration for binary problems. However, in multiclass settings, the outcomes were statistically equivalent. The findings were encouraging, using multilevel and ensemble setups in a four-class setting,

accuracies of 94.73% and 94.59% were achieved. Traditional CNNs, on the other hand, often require massive databases for effective training in order to get outstanding performance.

Jawahar et al. [15] created an ALNett deep neural network-based model which uses convolutions of depth with dissimilar dilation rates to categorize microscopic WBC pictures. In the cluster layers, convolution and max pooling are applied, along with normalization method that gives deeper architectural and contextual information in order to acquire healthy global and local attributes from microscopic pictures for true forecast of acute lymphoblastic Leukaemia. The model was compared to VGG16, resnet-50, google net, and AlexNet using recall, accuracy, precision, F1 score, loss accuracy, and ROC curves. The suggested ALNett model had the lowest processing complexity and maximum classification accuracy, 91.13%, with an F1 score of 0.96. However, massive amounts of computer power and information remain needed to develop and modify neural networks to do the specified purpose.

Abhishek et al. [16] suggested employing a heterogeneous dataset for the automation binary classification task, being part of the work's main obstacles. The proposed study solves binary and three-class classification problems utilizing cutting-edge machine learning and deep learning techniques. After joining all layers and fine-tuning VGG16's final three convolution layers, the suggested method achieves 97% binary classification accuracy and 98% densenet121 and supports vector machine precision. For assignments requiring three-class categorization, resnet50 and support vector machine obtain an accuracy of 95%. The innovative dataset is being created with the help of many specialists to aid scientists in medical research using machine learning algorithms. Advanced studies, on the other hand demand a big and well-annotated dataset.

Das et al. [17] presented an Acute Leukaemia detection methodology based on Orthogonal softmax Layer (OSL) that comprises deep feature extraction utilizing ResNet 18 and then accurate OSL-based classification. OSL and resnet18 are coupled to create weight vectors that are orthogonal to one another, which improves classification execution. Consequently, it combines the benefits of OSL-based classification with those of ResNet. To build a quicker and more effective network, the design also included additional ejection and ReLU layers. But it's crucial to remember that when using this method on small medical imaging datasets, overfitting may result.

Saeed et al. [18] use a CNN technology to offer an automated solution for identifying Leukaemia with acute lymph illness. Simulations used Acute Lymphoblastic Leukaemia-IDB 1 and Leukaemia-lb 2. Evaluate the proposed model's efficiency, a qualitative investigation was conducted by analysing the triggering middle layer, Convnet filters and heat-map layers. Furthermore, the study was compared with existing methodologies for a comprehensive evaluation. The proposed approach, on the other hand, obtained 99.61% accuracy in Leukaemia with acute lymph diagnosis. To solve the model's overfitting problem, data augmentation techniques have been utilized to produce pictures. This method, however, is reliant on manual observation and usually yields incorrect answers.

Batool and Byun [19] using DL methods to develop an accurate computer-aided methodology for diagnosing ALL. Based on efficientnet-B3, this paper aimed a lightweight DL-assisted robust model and uses depth wise separable convolutions to differentiate acute lymphoblastic Leukaemia and normal cells in a data set of WBC images. The aimed lightweight efficientnet-B3 employs fewer trainable attributes to increase the Effectiveness as well as performance of Leukaemia classification. In addition, the efficiency and ability to be generalized of the proposal lightweight efficientnet-B3 are evaluated using two publicly accessible datasets. Furthermore, the proposed and baseline classifiers are evaluated using accuracy, precision, recall, and F1-score. However, other pre-trained DL classifiers are lightweight and advised for a Leukaemia diagnosis, and explainable AI (such as SHAP) should be used to improve the transparency and flexibility of the efficient net-B3.

The previously mentioned pre-trained models, on the other hand, have a number of limitations, including complicated designs with several additional hyper-parameters, ineffectiveness as a function of processing time, and feature discrimination between normal and ALL pictures. To overcome these problems and establish a viable solution for Leukaemia cancer categorization, a novel DL technique as per swarm-based methods is required.

3. Proposed Methodology

This section describes the planned research study's comprehensive approach. Figure 2 shows the suggested method. First acquire the openly available dataset's data set available for free. Pre-process the data after obtaining the dataset. Following pre-processing, the suggested model efficientnet-B3 for Leukaemia classification must be

created. The classification model made use of a pre-trained ODRNN that was trained with DRNN. Finally, these models have been assessed using evaluation metrics.

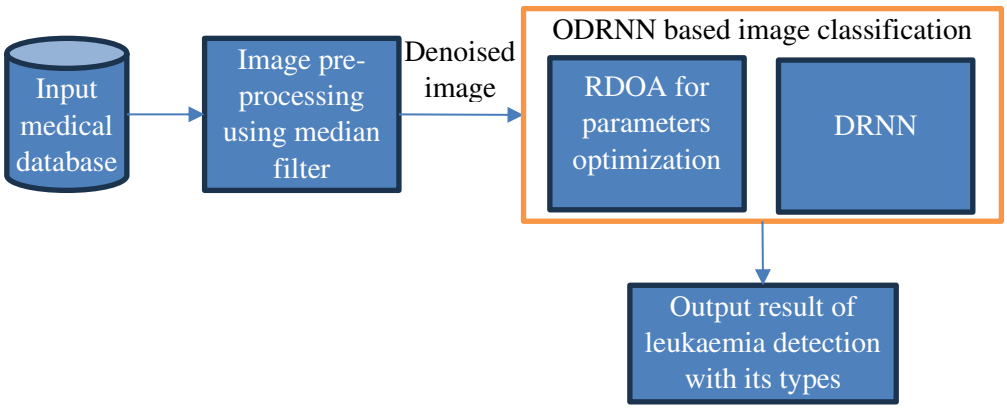

Figure 2. Enlightenment of proposed methodology.

3.1 Input Data Description

This work used publicly accessible datasets in this investigation, including C_NMC_2019 [21], Acute Myeloid Leukaemia (AML) [22], and the acute lymphoblastic Leukaemia – ALL database [23]. The C_NMC_2019 dataset contains 15,114 ALL pictures from 118 patients. Every image has a black backdrop and is 450450 pixels in size. Similarly, the dataset includes all subtypes of Leukaemia, including healthy, early, pre, and pro. All of the pictures in the information have been captured using. A Zeiss camera and classified using a color threshold-based technique. To emphasize it importance of deep learning (DL)-based robust classifiers, multi-class and binary classifications are used to diagnose cancer. The AML dataset comprises 10,000 single-cell pictures (64x64 pixels) acquired from AML (Blood Cancer) patients' peripheral blood smears.

3.2 Image Preprocessing Methods

Preprocessing is required to enhance the quality of source image and make the following technique to be processed easier. The initial processing processes in this investigation are depicted in Figure 3.

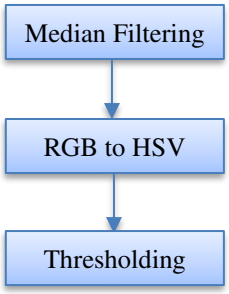

Figure 3. Image Preprocessing stages

To improve the image, the median filtering procedure will be used to rectify the white blood tumor cells. Opencv has a separate module for picture enhancement using a median filtering method. If the "cvmedianblur" technique and the value of the matrix input via an image pixel are used, the system will automatically fix the image. The center value that is necessary for the pixel matrix was represented by the pixel matrix used in this investigation, which had a value of 7. The final image's average filtration pixel value is calculated using equation (1).

$$F(x,y) = \frac{1}{m \times n} \sum_{x-1}^{m} \quad \sum_{y-1}^{n} \quad G(x,y) \qquad (1)$$

$F\ (x,y)$ = pixel filter value result

$G\ (x,y)$ = input pixel in the image

m,n = average size of the matrix

Median filter outcome is shown in Figure 4. After the median filtering procedure, images of blood cells become smoother.

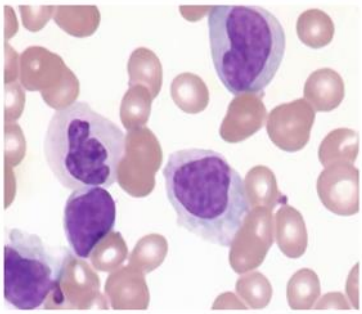

Figure 4. Median filter-based noise filtering results

3.3 Process of RGB to HSV

After initial processing, the process results will be converted from RGB to HSV. The shades are defined by HSV in terms of color, saturation, hue, and values. HSV has the advantage of including shades that are comparable to those perceived by the human eye. Colors are made by combining primary colors from different categories, such as RGB. Use the following equation to convert RGB pictures [25] to HSV images:

$$r = \frac{R}{(R+G+B)}, g = \frac{G}{(R+G+B)}, b = \frac{B}{(R+G+B)} \quad (2)$$

$$V = (R, G, B) \quad (3)$$

$$s = 0 \text{ if } v = 0 \; s = 1 - \frac{(r,g,b)}{v} \text{ If } v > 0 \quad (4)$$

$$H = H + 360; H < 0 \quad (5)$$

Where:

R, G, B = red, green, blue

H, S, V = Hue, Saturation, Value

When the process is completed, the HSV image will be generated from the modified RGB image. This happens because the colors in the HSV spectrum are identical to those experienced by the senses of humans. The color conversion outcomes RGB to HSV conversion are shown in Figure 5. During color conversion from RGB to HSV, the image of blood cells changes considerably [25].

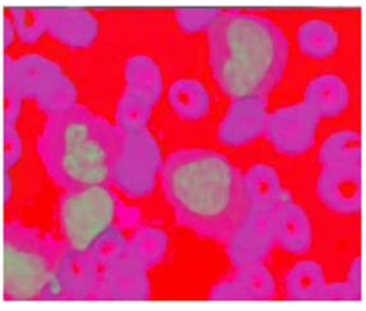

Figure 5. The result of HSV

3.4 Thresholding

For objects and background pixels, there are two main alternatives. To remove things from the background, utilize the (Th) threshold value that divides the two types. Any point in the coordinates (x, y) that fulfills f (x, y)> Th an object's point or it is the background point. This quality is the histogram partitioning technique determines the success of this strategy. When a pixel value in an image exceeds the threshold value of 0, it is swapped with maxval value of 255 (white). If the shade of gray pixel count is greater than in the picture is less than the threshold,

the pixel value replaces it with 0 (black). As a result, the blood smear image area is known to be employed as a segmented object. The threshold's value was calculated via a slider in this study, with the lowest and highest values determined using opencv's "in range" function. After the pattern has been obtained, the original data will be masked in accordance with the pattern obtained during the holding operation [26]. Figure 6 depicts the thresholding findings. Following the thresholding operation, the formerly red blood cell-only image now includes white illness objects.

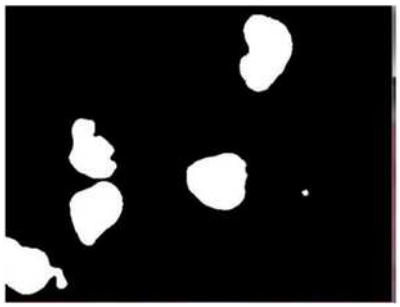

Figure 6. Result of Thresholding process

3.5 Leukaemia Classification Using ODRNN

3.5.1 Deep Recurrent Neural Network (DRNN) Model: DRNN which makes use of a previous phase's input as the current phase's input. Neural networks model the action of neurons used in pattern recognition. The feed forward network is known as DRNN if it is a direct cycle. The DRNN employs three different levels: input, output, and hidden. The primary part of the DRNN that retains information about the received signal is the buried layer. RNN incorporates a number of layers that are hidden based on the issue specification; RNN is sometimes called as DRNN owing to the vast number of hidden layers. F hidden layers are used in this study to increase performance. A directed cycle may be seen in DRNN networks with nodes in the buried layer [24]. The signal parameters that have been analysed are inputted into the DRNN, which then analyses them using a certain number of hidden layers. The input signals are compared to the previous outputs before being transferred to the hidden layers. The DRNN incorporates several hidden layers that are determined by the issue specification, resulting in increased complexity of the network. The frequency and amplitude of the buried layer in the DRNN make it an S-matrix value. The input matrix becomes hidden layers in DRNN. To get to the optimal layer, the RDOA technique is employed to lower the DRNN's network complexity [27]. In figure 7, the DRNN is illustrated using a comparable unfold network.

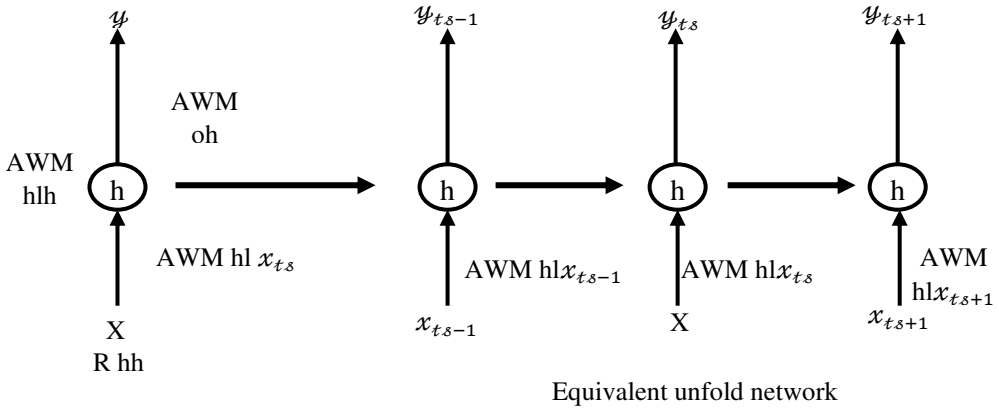

Figure 7. DRNN Model using a comparable unfold network

Figure 8 depicts the network input with the series of times ts as $(\ldots x_{ts-1}, x_{ts}, x_{ts+1} \ldots.)$ The connection is assigned using the weight matrix AWM_{1h}. If the layers that are hidden with n units are identified as hl_{ts} $(hl_1, hl_2, \ldots, hl_n)$, then a network with recurrent connections connects these hidden layers. To investigate the present situation, apply the following equation [27]:

$$cs_{ts} = f_{tf}(ps_{ts-1}, x_{ts}), \quad (6)$$

Where, cs_{ts} Is the current state, ps_{ts1}Is the prior state, x_{ts} Is the input state, and f_{tf} Is the state transition function [26]. The equation for the state at ts time is,

$$cs_{ts}ps(AWM_{hh} \cdot ps_{ts1} \cdot x_{ts}) \quad (7)$$

Where WM_{hh} Is the recurrent neuron weight, AWM_{xh} Is the input neuron weight, and tan p is the activating parameter. In the preceding equation, the recurrent neuron took into consideration the prior state. Only when the ultimate state was computed was the output obtained [27]. The final step in the outcome of the computation was accomplished by,

$$y_{ts} = f_0(cs_{ts}) \quad (8)$$

$$y_{ts} = A_{hy}cs_{ts} \quad (9)$$

Where y_{ts} Is the output, AWM_{hy} Is the mass of the output layer, and f_0 - function of output. A collection of N" training sequences is provided by,

$$S = \{((x_1^n, y_1^n) \dots (x_T^n, y_T^n))\}_{n-1}^N \quad (10)$$

The cost function is given by

$$ct(\theta) = \frac{1}{N}\sum_{n=1}^{N} \quad \sum_{ts=1}^{T_n} \quad d\left(y_{ts}^n, f_0(cs_{ts}^n)\right) \quad (11)$$

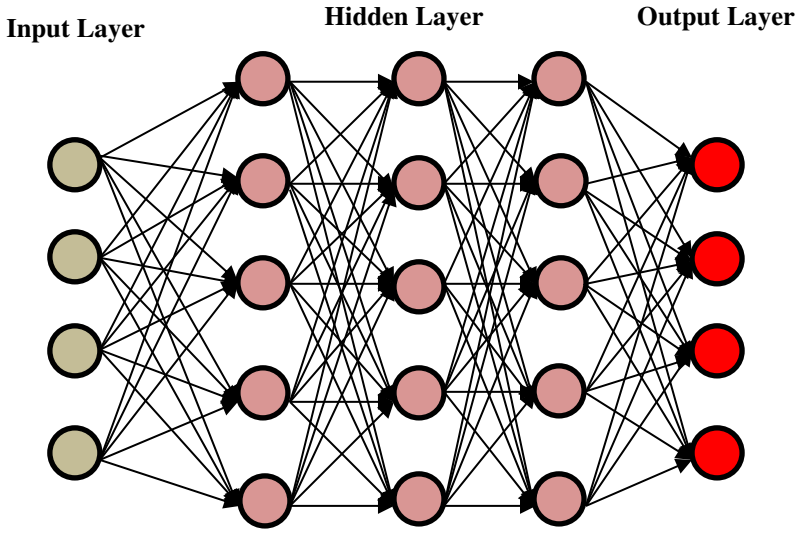

Figure 8. Deep Neural Network (DRNN) Structure

In this study, the RDOA is used to optimize the three Layers that are hidden. This type of network is referred to as a deep network because to the number of hidden levels. Fig. 8 displays the deep neural network (DRNN) construction [28] – [34].

Model Training and Evaluation: The RNN model is trained on labeled datasets where each data point belongs to a specific Leukaemia subtype. During training, the model learns to associate specific sequence patterns with corresponding subtypes. Once trained, the model's performance is evaluated on a separate test dataset to assess its generalizability and accuracy.

Classification and Interpretation: The trained RNN model can classify new data points (gene expression profiles or time-series measurements) into different Leukaemia subtypes. Additionally, the model can identify crucial features within the sequences that contribute to its classification, providing insights into the underlying biological mechanisms of each subtype. Advantages of using RNNs for Leukaemia classification are given as follows:

- High accuracy: RNNs can achieve high classification accuracy, often surpassing traditional machine learning methods.
- Long-term dependencies: RNNs can capture long-range relationships within sequences, providing a more comprehensive understanding of the data.
- Feature extraction: RNNs can identify key features within the data that contribute to classification, offering valuable insights into disease biology.

3.5.2 Red Deer Optimization Algorithm (RDOA): A RDOA method is used to discover the approximation layer that is closest to the ideal solution, which is determined by layer weighting. In DRNN, RDOA looks for a no. Of hidden layers. Based on layer weights, the RDOA algorithm improves the overall number of layers in DRNN. The layer weights correspond to the pitch of the deer's roar, and choosing deer for breeding corresponds to the layers segment in DRNN. The RDOA algorithm [25] was inspired by variations in mating behaviour of Scottish Red Deer (RD) throughout the breeding session. There are two types of deer in the RDOA: hinds and men RD. Assume that weight represents male RD and biases indicate hinds. Figure 9 depicts the RDOA algorithm procedure.

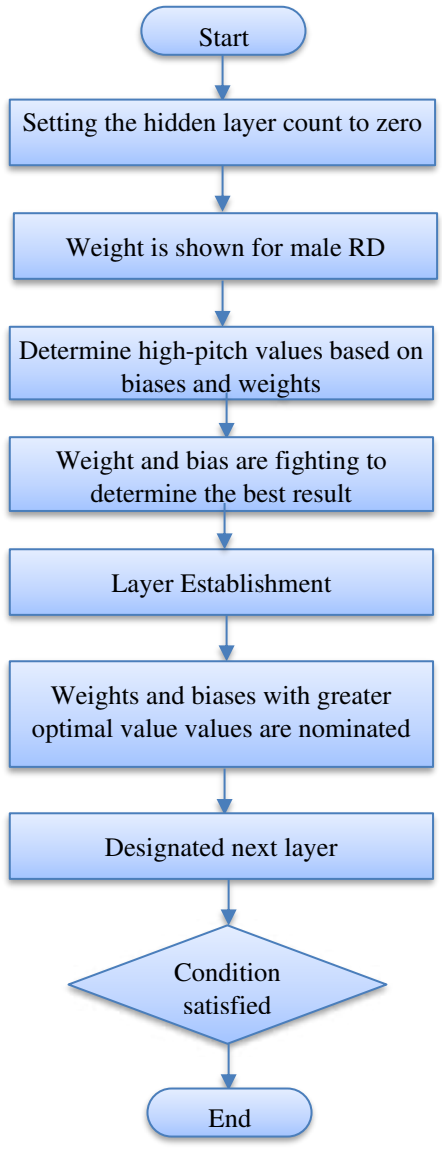

Figure 9. Flowchart of RDOA

Starting of RD: The RDOA begins by building the starting number of RD (i.e., layer), followed by the formation of an ordered set of values for optimization efforts. Each Layer is supplied as an optimization solution. The array is defined in the Q" dimension problem [25].

$$RD\{D_1, D_2, D_3, \ldots \ldots D_{Q-1}\} \quad (12)$$

The function VF's value is tested using,

$$VF\ f(RD) \quad (13)$$

'$Qpopulation$' denotes the initial population's matrix size.

Position Shifting: At each level, the male RD (i.e., layer) shifts positions and improves their objectives function over the previous one. The biases gravitate toward the highest booming rate (i.e., value).

Male Commanders (MC) Assortment: High-pitch roaring weight commanders are distinguished from low-pitch biases, which produce the most effective weights, by their roaring rate [25].

$$Q.MC\ round\{.Q_M\} \quad (14)$$

The stag population SP is computed using the formula,

$$Q.SP\ Q_M\ Q.MC \quad (15)$$

Where, $Q.MC$ stands for weight commander, Q_M Stands for weight, and $Q.SP$ stands for biases.

Hostile Between Commanders and Stags: Weight commanders, or male commanders, now operate with randomized biases, meaning they engage in combat only when an objective function surpasses its previous value.

$$U1 = \frac{CS}{2} + f_1 \times \big((up - lr) \times f_2\big) + lr \quad (16)$$

$$U2 = \frac{CS}{2} - f_2 \times \big((up - lr) \times f_2\big) + lr$$

Where CS represents the commander of the stage and f_1, f_2 Are uniformly distributed ranging from 0 to 1. The new value generation is represented by $U1$ and $U2$. In the equation, up and lr represent the upper and lower bounds of area for investigation.

Harem Establishment: As stated before, a harem is a group of hinds who combat and cry under the rule of the male lead (i.e., the power of weight commander). The male commander's normalized value is computed as follows:

$$Norm_n\ b_n\ max\ (b_i) \quad (17)$$

b_n Is the number of n^{th} Male commanders in the next equation, and $Norm_n$ Is the normalization value. The commander pc is normalized power is given by,

$$pc_n = \left| \frac{Norm_n}{\sum_{i=1}^{MC} V_{il}} \right| \quad (18)$$

The total number of hinds within a harem is determined by

$$Q.harem_n\ round(\ pc_n \cdot Q_{hind})\ . \quad (19)$$

Wherein Q_{hind} is the total number of hinds. Male commander and hinds are parents in the cross-over process of mating, and their progeny supply the new solution of

$$Q.haerm_n^M\ round(.Q.haerm_n) \quad (20)$$

Where is mate the number of hinds in the n^{th} Harem who're ready to mate with the male rds is denoted by $Q.haerm_n^{mate}$. The commander mates with the $\lfloor$ proportion of hinds in the harem, which is picked at random. That is,

$$Q.harem_n^M\ round(\ .Q.harem_n) \quad (21)$$

Where, mate $Q.harem_n$ is the total number of hinds mated to a male deer leader in the harem. The most recent solution, LS is employed to build the mating formulation.

$$LS = \frac{(C+hind)}{2} + (up - lr) \times f \quad (22)$$

Stag Mating: The male RD's distance from all hinds is calculated as

$$D_i = \left(\sum_{k \epsilon K} \left(S_k - hindk_k^i\right)^2\right)^{\frac{1}{2}} \quad (23)$$

The hind for the following generation is chosen by regular selection from the subsequent generation of male RD. The convergence might be due to several iterations or the excellent quality of the solution produced. Figure 10 shown DRNN's workflow while optimizing the layers with RDOA.

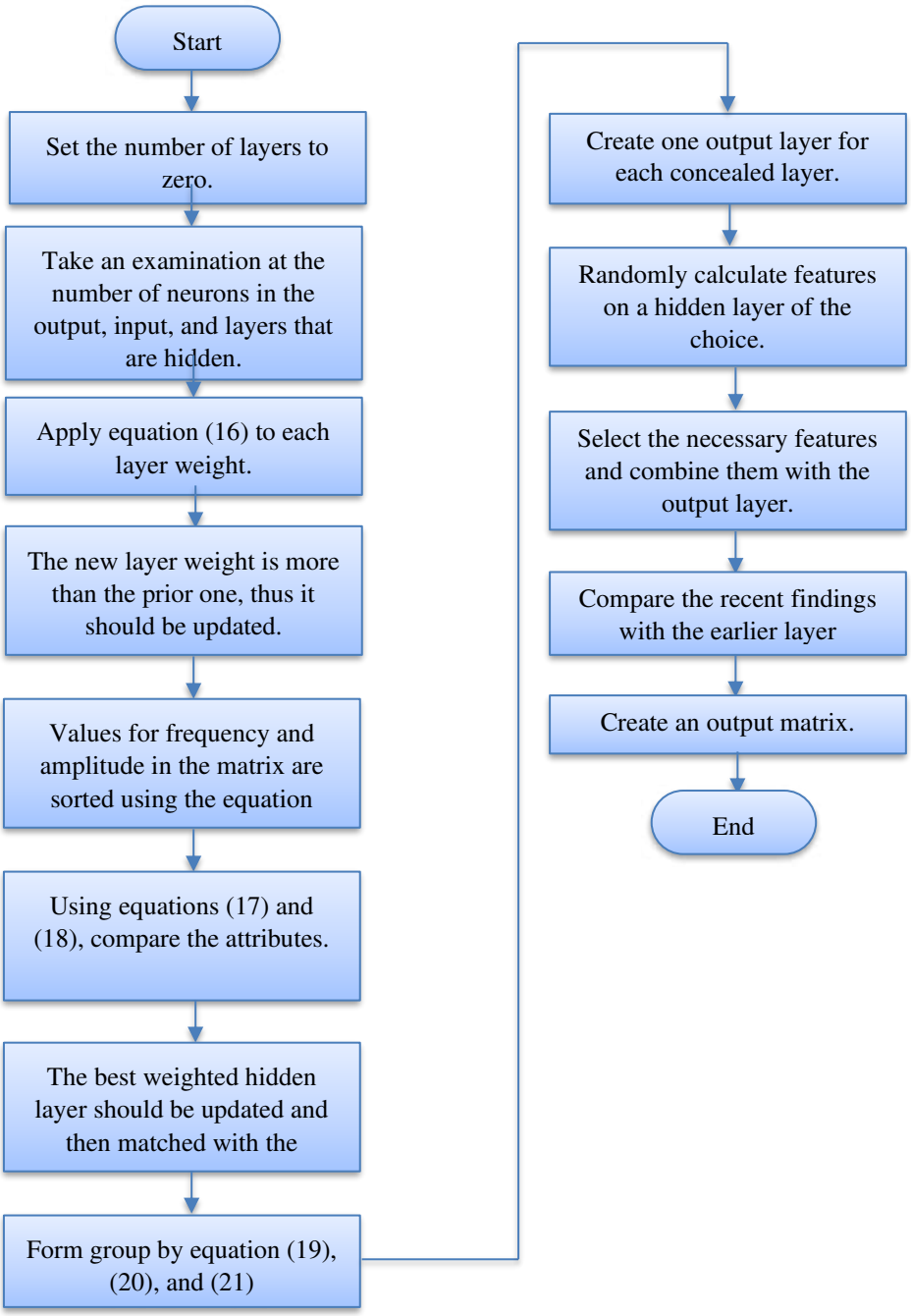

Figure 10. Workflow of proposed method

4. Experimental Results and Discussion

The researchers assess the suggested identification and classification model using statistical measures linked to diseases, including specificity, recall, accuracy, precision, and F1 score. The TP (true positive), FP (false positive),

TN (true negative), and FN (false negative) rates are used to represent positive classifications. The recall of a certain category of Leukaemia subtype, conversely, reflects its level of predictability. There is agreement between the predicted and actual subtypes of Leukaemia. While precision is the percentage of correctly classified positive leukaemia subtype predictions, the recall and accuracy measurements translate to the F1 score. Additional performance measures must be evaluated in order to put the model to the test. These acts are defined shortly below.

Accuracy: This statistic indicates how well a strategy works by computing the overall count of successfully predicted cases. It may be stated numerically as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (24)$$

Precision: The ratio of all properly predicted occurrences to all accurately expected events that happened is how it is defined. High accuracy and low false positives are linked. As stated below:

$$Precision = \frac{TP}{TP+FP} \quad (25)$$

Specificity: The result depends on the ratio of absolutely negative observations to negative conditions. This measure shows the frequency with which different disease kinds are accurately identified.

$$Sepecificity = \frac{TP}{TP+FP} \quad (26)$$

Recall: The proportion of all class observations with positive explanations anticipated correctly is shown by this ratio.

$$\text{(Recall) } sensitivity = \frac{TP}{TP+FN} \quad (27)$$

F1 score: The F1 score is used to measure classifier performance. Although the F1 score is susceptible to data circulation, it is an excellent metric for classification problems on unstable data sets.

The suggested ODRNN model identifies the categorization of four leukaemia subtypes from blood smear pictures, including ALL, CLL, AML, and CML. Performance criteria like as precision, recall, F1 score, and accuracy are used to assess the provided models. The F1 score, accuracy, precision, recall, and other metrics for all kind of Leukaemia were determined using the standard dataset mentioned in the dataset description section image bank. According to Table 1, the ODRNN is 100% accurate for ALL, CLL, CML, AML, and healthy individuals. ODRNN is 99.91% accurate in predicting CLL. However, the accuracy, recall, and F1 score are all 0.99%. AML accuracy is 98.99% with a precision of 0.99%, recall of 1.0%, and F1 score of 1.0%, compared to ODRNN's CML accuracy of 98.96%. ODRNN forecasts the F1 score, recall, accuracy, and precision for healthy and ALL with 100%.

Table 1: Performance of the ODRNN model for classifying Leukaemia subtypes

Leukaemia type	Precision	Recall	Specificity	F1 score	Accuracy
ALL	100	100	100	100	100
AML	98	99	98	99	98.99
CLL	99	99	99	100	99.91
CML	99	100	100	100	98.96
Healthy	100	100	100	100	100

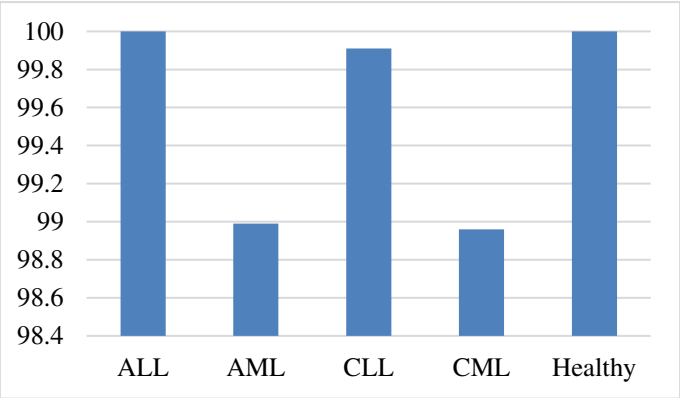

Figure 11: Comparison results of the ODRNN model for classifying Leukaemia subtypes

To demonstrate their effectiveness, the proposed models are compared to earlier approaches such as Apriori [20], ALNett [13], and CNN [35-37]. AML, ALL, and healthy samples are identified using GA with SVM, and the kind of Leukaemia present in each sample is identified using CNN. ODRNN is utilized to detect ALL and healthy samples. However, Figure 11 reveals that the employed models, namely ODRNN, outperform the present models, including Apriori, ALNett, CNN, and the suggested ODRNN. Figure 12 and table 2 shows that ODRNN achieves 99.98% accuracy, while current techniques such as Apriori, ALNett, and CNN achieve 92.41%, 93.50%, and 94.61%, respectively.

Table 2: Performance metrics comparison

Leukaemia classification	Accuracy	Precision	Specificity	Recall	F1 score
Apriori [20]	92.41	92.56	91.41	92.21	93.32
ALNett [13]	93.50	93.65	92.69	93.45	94.65
CNN [16]	94.61	94.95	93.74	95.62	95.45
Proposed ODRNN	**99.98**	**99.23**	**98.96**	**99.85**	**99.98**

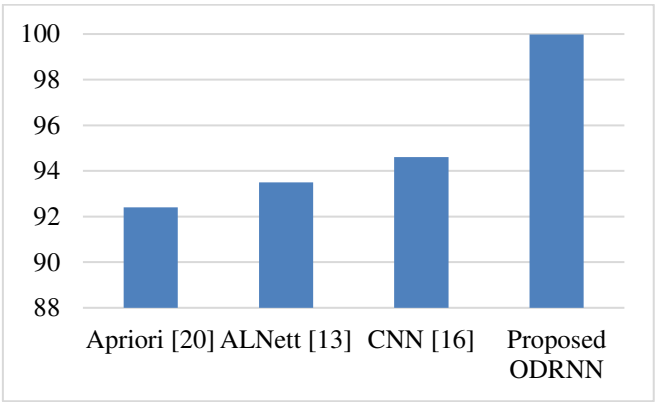

Figure 12. Accuracy Comparison of Leukaemia classifier

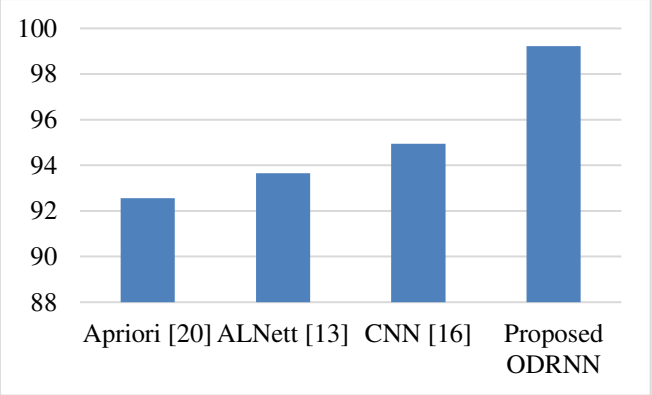

Figure 13. Precision Comparison of Leukaemia classifier

According on the comparative study, the suggested classifier has a 99.98% efficiency and a precision of 99.23%, much outperforming cutting-edge classifiers as Apriori [20], ALNett [13], and CNN [16], shown in Figure 13. According to the comparative research, the proposed classifier beats the present classifier in terms of detection rate. The analysis also reveals that the suggested model has a slightly increased rate of detection for multi-class Leukaemia classification than for binary Leukaemia classification. Furthermore, detection rate of all produced categories for the multi-class classification of Leukaemia is enhanced to the comparison to the previous categorization.

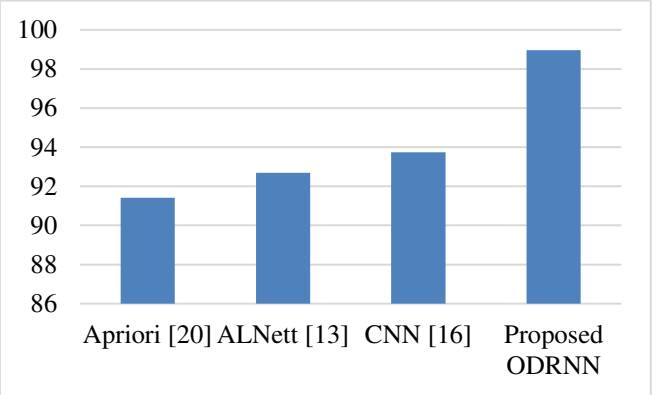

Figure 14. Specificity Comparison of Leukaemia classifier

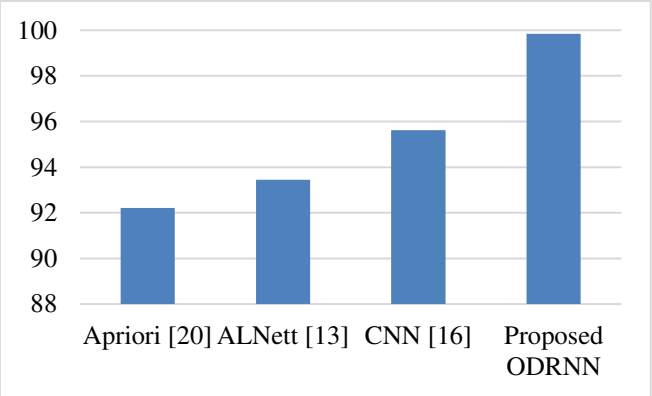

Figure 15. Recall Comparison of Leukaemia classifier

According to the comparison analysis, the proposed classifier achieved 99.23% specificity and 98.96% recall that as shown in figure 14 and figure 15, much outperforming state-of-the-art classifiers as Apriori [20], ALNett [13], and CNN [16]. Weights and biases are two trainable characteristics that are used to assess the effectiveness of recommended and baseline classifier. According on the findings, the recommended ODRNN classifier required less parameters for training than prior developed classifiers. As a result, the suggested classifier is shown to be

efficient since it considerably decreases the cost of computation in the form of parameters that can be trained to boost its efficiency for the proposed research study for Leukaemia diagnosis.

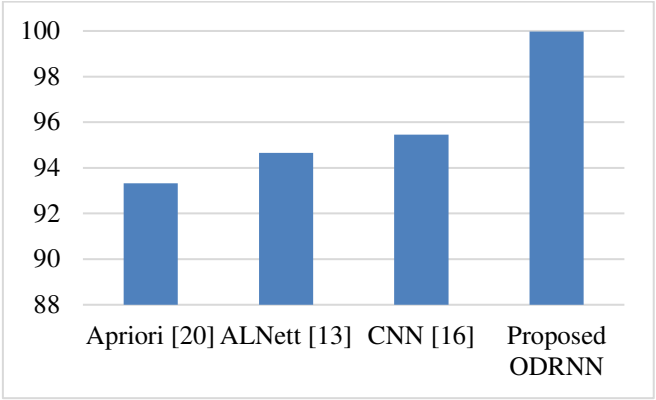

Figure 16. F1 score Comparison of Leukaemia classifier

The suggested approach has the highest classifying f1 score of 99.98%, according to the experiment results shown in figure 16. According to the findings of the efficacy research, the proposed classifier performs more effectively for binary classification data than for classification of multiple classes datasets. However, the proposed classifier beat the present classifiers in the multi-class categorization of Leukaemia. Furthermore, ODRNN beat other existing pre-trained DL classifiers not only in terms of efficiency, but also in terms of efficiency, by utilizing a smaller number of parameters for training to reduce processing complexity.

To demonstrate their effectiveness on blurred images, the proposed models are compared to earlier approaches such as Apriori [20], ALNett [13], and CNN [35-37]. Figure 17 and Table 3 shows that ODRNN achieves 91.23% accuracy, while current techniques such as Apriori, ALNett, and CNN achieve 83.66%, 84.75%, and 85.86%, respectively.

Table 3: Performance metrics comparison based on blurred leukaemia images

Leukaemia classification	Accuracy	Precision	Specificity	Recall	F1 score
Apriori [20]	83.66	83.81	82.66	83.46	84.57
ALNett [13]	84.75	84.9	83.94	84.7	85.9
CNN [16]	85.86	86.2	84.99	86.87	86.7
Proposed ODRNN	**91.23**	**90.48**	**90.21**	**91.1**	**91.23**

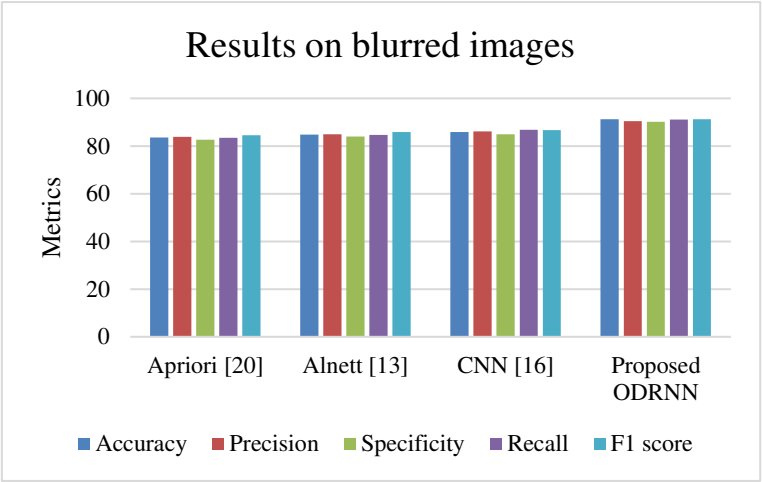

Figure 17. Performance Comparison of Leukaemia classifier for blurred image

To showcase their efficacy on blurry images, the suggested models are contrasted with previous methodologies, including Apriori [20], ALNett [13], and CNN [35-37]. Table 4 reveals that the employed models, namely ODRNN, outperform the present models, including Apriori, ALNett, CNN, and the suggested ODRNN. Figure 18 shows that ODRNN achieves 90.13% accuracy, while current techniques such as Apriori, ALNett, and CNN achieve 82.56%, 83.65%, and 84.76%, respectively.

Table 4: Performance metrics comparison based on fuzzy image

Leukaemia classification	**Accuracy**	**Precision**	**Specificity**	**Recall**	**F1 score**
Apriori [20]	82.56	82.71	81.56	82.36	83.47
ALNett [13]	83.65	83.8	82.84	83.6	84.8
CNN [16]	84.76	85.1	83.89	85.77	85.6
Proposed ODRNN	**90.13**	**89.38**	**89.11**	**90**	**90.13**

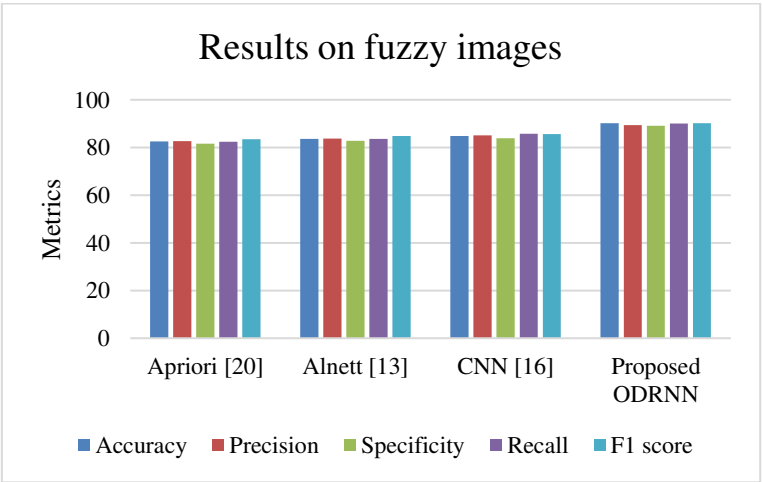

Figure 18. Performance Comparison of Leukaemia classifier for fuzzy image

To highlight their effectiveness on partly covered images, the proposed models are compared to earlier methodologies, such as Apriori [20], ALNett [13], and CNN [35-37]. According to Table 5, the employed models,

specifically ODRNN, demonstrate superior performance compared to current models, including Apriori, ALNett, CNN, and the suggested ODRNN. Figure 19 illustrates that ODRNN attains an 89.48% accuracy, surpassing the accuracy of existing techniques such as Apriori (81.91%), ALNett (83%), and CNN (84.11%).

Table 5. Performance metrics comparison based on partly covered images

Leukaemia classification	Accuracy	Precision	Specificity	Recall	F1 score
Apriori [20]	81.91	82.06	80.91	81.71	82.82
ALNett [13]	83	83.15	82.19	82.95	84.15
CNN [16]	84.11	84.45	83.24	85.12	84.95
Proposed ODRNN	**89.48**	**88.73**	**88.46**	**89.35**	**89.48**

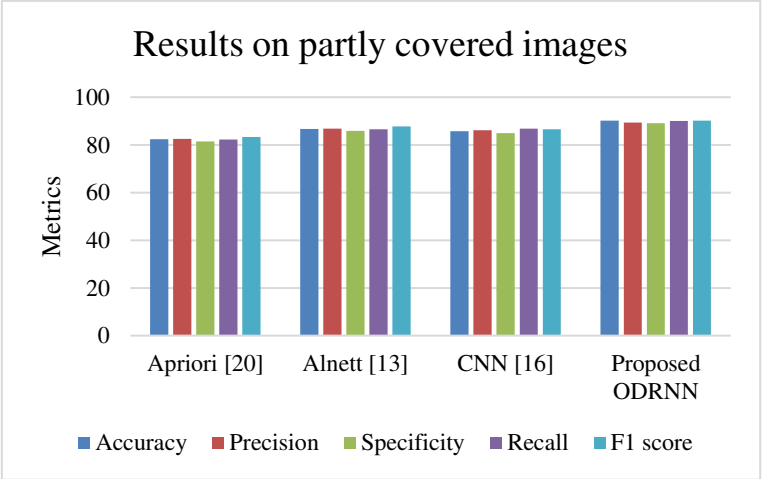

Figure 19. Performance Comparison of Leukaemia classifier for partly covered image

5. Conclusion and future work

The proposed ODRNN model in this study detects the categorisation of four leukaemia types from blood smear images, includes AML, ALL, CLL, and CML. The ODRNN model are utilized to make a Leukaemia diagnosis. The research incorporates disease-related statistical metrics, such as specificity, recall, accuracy, precision, and F1 score, and ODRNN method attains remarkable results with scores of 98.96%, 99.85%, 99.98%, 99.23%, and 99.98%, respectively. The effectiveness of the categorization is compared to that of other systems such as Apriori, ALNett, and CNN. Experiment findings show that the ODRNN methodology beats existing approaches in both terms of efficiency and other performance metrics. Results were obtained including the subtyping of leukaemia to increase performance. Furthermore, the efficiency of the deep learning classifier improved via choosing the optimal combination of hyper parameters. For instance, the number of concealed layers, the size of the batch, function activations, and so on. A realistic technique to improve the image classification performance metrics in blur images, fuzzy images and partly covered images is still required. The future work includes:

- Integrate various data modalities such as combining imaging data with clinical data, genomic data, or other relevant information. This can provide a more holistic view for accurate diagnosis.
- Explore the effectiveness of ensemble models, combining multiple deep learning architectures to enhance overall performance and robustness.
- Explore advanced data augmentation techniques tailored to medical imaging, considering the sensitivity of the data and the potential impact on diagnosis accuracy.

COMPETING INTERESTS

The authors state that none of their known financial interests or relationships might have had an impact on the work presented in this publication.

FUNDING

Not Applicable

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are available in the UCI repository.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

K.Dhana Shree: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. S.Logeswari: Investigation, Formal analysis, Supervision

References

1. Ahirwar, D. R., Nigam, R. K., & Parmar, D. (2018). A study of Leukaemias profile in central India. Tropical Journal of Pathology & Microbiology, 4(2), 2456-1487.
2. Van Zwieten, R., Verhoeven, A. J., & Roos, D. (2014). Inborn defects in the antioxidant systems of human red blood cells. Free Radical Biology and Medicine, 67, 377-386.
3. Nolan, J. P., & Jones, J. C. (2017). Detection of platelet vesicles by flow cytometry. Platelets, 28(3), 256-262.
4. Talaat, F. M., & Gamel, S. A. (2023). Machine learning in detection and classification of Leukaemia using C-NMC_Leukaemia. Multimedia Tools and Applications, 1-14.
5. Hegde RB, Prasad K, Hebbar H, Singh BMK, Sandhya I (2019) Automated decision support system for detection of Leukaemia from peripheral blood smear images. J Digit Imaging 33:361–374.
6. Das NN et al (2022) Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays.". Irbm 43(2):114–119
7. Ehrenstein V, Nielsen H, Pedersen AB, Johnsen SP, Pedersen L (2017) Clinical epidemiology in the era of big data: new opportunities, familiar challenges. Clin Epidemiol 9:245–250
8. Anilkumar, K. K., Manoj, V. J., & Sagi, T. M. (2020). A survey on image segmentation of blood and bone marrow smear images with emphasis to automated detection of Leukaemia. Biocybernetics and Biomedical Engineering, 40(4), 1406-1420.
9. Ratley, A., Minj, J., & Patre, P. (2020, January). Leukaemia disease detection and classification using machine learning approaches: a review. In 2020 First International Conference on Power, Control and Computing Technologies (ICPC2T) (pp. 161-165). IEEE.
10. Kumar N et al. (2021) Efficient automated disease diagnosis using machine learning models. J Healthcare Eng 2021.
11. Das, P. K., & Meher, S. (2021). An efficient deep convolutional neural network based detection and classification of acute lymphoblastic Leukaemia. Expert Systems with Applications, 183, 115311.
12. Agustin, R. I., Arif, A., & Sukorini, U. (2021). Classification of immature white blood cells in acute lymphoblastic Leukaemia L1 using neural networks particle swarm optimization. Neural Computing and Applications, 33(17), 10869-10880.
13. Acharya, V., Ravi, V., Pham, T. D., & Chakraborty, C. (2021). Peripheral blood smear analysis using automated computer-aided diagnosis system to identify acute myeloid Leukaemia. IEEE Transactions on Engineering Management.
14. Claro, M. L., de MS Veras, R., Santana, A. M., Vogado, L. H. S., Junior, G. B., de Medeiros, F. N., & Tavares, J. M. R. (2022). Assessing the impact of data augmentation and a combination of cnns on Leukaemia classification. Information sciences, 609, 1010-1029.

15. Jawahar, M., Sharen, H., & Gandomi, A. H. (2022). ALNett: A cluster layer deep convolutional neural network for acute lymphoblastic Leukaemia classification. Computers in Biology and Medicine, 148, 105894.
16. Abhishek, A., Jha, R. K., Sinha, R., & Jha, K. (2022). Automated classification of acute Leukaemia on a heterogeneous dataset using machine learning and deep learning techniques. Biomedical Signal Processing and Control, 72, 103341.
17. Das, P. K., Sahoo, B., & Meher, S. (2022). An efficient detection and classification of acute Leukaemia using transfer learning and orthogonal softmax layer-based model. IEEE/ACM Transactions on Computational Biology and Bioinformatics.
18. Saeed, U., Kumar, K., Khuhro, M. A., Laghari, A. A., Shaikh, A. A., & Rai, A. (2023). Deepleuknet—A CNN based microscopy adaptation model for acute lymphoblastic Leukaemia classification. Multimedia Tools and Applications, 1-25.
19. Batool, A., & Byun, Y. C. (2023). Lightweight efficientnetb3 Model based on Depthwise Separable Convolutions for Enhancing Classification of Leukaemia White Blood Cell Images. IEEE Access.
20. Hossain, M. A., Islam, A. M., Islam, S., Shatabda, S., & Ahmed, A. (2022). Symptom based explainable artificial intelligence model for Leukaemia detection. IEEE Access, 10, 57283-57298.
21. Https://www.kaggle.com/datasets/avk256/cnmc-Leukaemia
22. Https://www.kaggle.com/datasets/akhiljethwa/blood-cancer-image-dataset
23. Https://www.kaggle.com/datasets/andrewmvd/Leukaemia-classification
24. Hermans, M., & Schrauwen, B. (2013). Training and analysing deep recurrent neural networks. Advances in neural information processing systems, 26.
25. Fard, A. F., & Hajiaghaei-Keshteli, M. (2016). Red Deer Algorithm (RDA); a new optimization algorithm inspired by Red Deers' mating. In International Conference on Industrial Engineering, IEEE (Vol. 12, pp. 331-342).
26. Shree, K. D., & Janani, B. (2019). Classification of Leucocytes for Leukaemia Detection. Research Journal of Engineering and Technology, 10(2), 59-66.
27. Kumar, A., Priyanka, S., Dhanashree, K., Praveen, V., Rekha, R. (2022). Efficient binary grasshopper optimization based neural network algorithm for bitcoin value prediction. International Journal of Nonlinear Analysis and Applications, 13(Special Issue for selected papers of ICDACT-2021), 53-60. Doi: 10.22075/ijnaa.2022.6330.
28. Arunachalam, S. K., & Rekha, R. (2022). A novel approach for cardiovascular disease prediction using machine learning algorithms. Concurrency and Computation: Practice and Experience, 34(19), e7027.
29. Dhanashree, K., Jayabal, P., Kumar, A., Logeswari, S., Priya, K. (2022). Fingernail analysis for early detection and diagnosis of diseases using machine learning techniques. International Journal of Nonlinear Analysis and Applications, 13(Special Issue for selected papers of ICDACT-2021), 61-69. Doi: 10.22075/ijnaa.2022.6331.
30. Liu, J., Hua, J., Chellappa, V., Petrick, N., Sahiner, B., Farooqui, M., ... & Summers, R. M. (2012, February). Automatic detection of axillary lymphadenopathy on CT scans of untreated chronic lymphocytic Leukaemia patients. In Medical Imaging 2012: Computer-Aided Diagnosis (Vol. 8315, pp. 107-113). SPIE.
31. Tharsanee, R. M., Soundariya, R. S., Kumar, A. S., Karthiga, M., & Sountharrajan, S. (2021). Deep convolutional neural network–based image classification for COVID-19 diagnosis. In Data Science for COVID-19 (pp. 117-145). Academic Press.
32. Aghamaleki, F. S., Mollashahi, B., Nosrati, M., Moradi, A., Sheikhpour, M., & Movafagh, A. (2019). Application of an artificial neural network in the diagnosis of chronic lymphocytic Leukaemia. Cureus, 11(2).
33. Kumar, A. S., & Rekha, R. (2023). An improved hawks optimizer-based learning algorithms for cardiovascular disease prediction. Biomedical Signal Processing and Control, 81, 104442.

34. Priyanka, S., Praveen, V., & Sivapriya, G. (2022). Hindrance Detection and Avoidance in Driverless Cars Through Deep Learning Techniques. In Advances in Deep Learning Applications for Smart Cities (pp. 69-100). IGI Global.
35. Rao, G. E., Rajitha, B., Srinivasu, P. N., Ijaz, M. F., & Woźniak, M. (2024). Hybrid framework for respiratory lung diseases detection based on classical CNN and quantum classifiers from chest X-rays. *Biomedical Signal Processing and Control*, *88*, 105567.
36. Chaki, J., & Woźniak, M. (2023). A deep learning based four-fold approach to classify brain MRI: BTSCNet. *Biomedical Signal Processing and Control*, *85*, 104902.
37. Chaki, J., & Woźniak, M. (2023). Deep learning for neurodegenerative disorder (2016 to 2022): A systematic review. *Biomedical Signal Processing and Control*, *80*, 104223.