



## An Optimal Detection for Leukaemia Cancer Based On RNS-Metaheuristic Technique in Micro Array Dataset

<sup>1</sup>Bamidele, S. A., <sup>2</sup>Asaju-Gbolagade A.W. and <sup>3</sup>Gbolagade, K. A.

<sup>1,3</sup>Kwara State University, Malete, Ilorin, Nigeria

<sup>2</sup>University of Ilorin, Nigeria.

<sup>1\*</sup>[simirado@gmail.com](mailto:simirado@gmail.com)

<sup>2</sup>[aishatwurala@gmail.com](mailto:aishatwurala@gmail.com)

<sup>3</sup>[kazeemgbolagade@kwasu.edu.ng](mailto:kazeemgbolagade@kwasu.edu.ng)

### Abstract

This paper addresses the critical challenge of leukaemia cancer detection through the integration of Residue Number System (RNS) and Convolutional Neural Network (CNN) Deep Learning Framework using a Microarray dataset. Leveraging a dataset obtained from the Kaggle machine learning repository, the study employs a comprehensive image processing pipeline, encompassing grayscale conversion, data augmentation, contrast enhancement, geometry normalization, and OTSU segmentation. The subsequent stages involve feature extraction using Histogram of Gradient (HOG) and comparative feature selection through Ant Colony Optimization (ACO) and an optimized ACO+RNS approaches. Results indicates that incorporating ACO+RNS outperforms the ACO-only in terms of classification accuracy, sensitivity, specificity, precision, and F1-score. Notably, the ACO+RNS model achieves a lower error rate and reduced training time, emphasizing the efficiency of incorporating Residue Number System encoding in feature selection.

**Keywords:** *Ant Colony Optimization, Convolutional Neural Network, Histogram of Gradient, Leukaemia Cancer, Residue Number System*

### 1. Introduction

Leukaemia is the most prevalent childhood and adult blood cancer. It starts in bone marrow and creates abnormal blood cells. Blood blasts or leukaemia cells are immature. Blood, bruising, bone discomfort, weariness, fever, and infection risk are symptoms. Lack of regular blood cells causes these symptoms [1, 2].

A bone marrow biopsy or blood test is often done to diagnose. Exactly what causes leukaemia is unknown to scientists. Probably genetic and environmental malignant leukocytes are difficult to differentiate at cheap cost in early illness diagnosis. Laboratory diagnostic centers lack flow

cytometry technology and lengthy procedures [3]. Hematologists in cell transplant centers use microscopic pictures to diagnose leukaemia .

Microscopy of PBS is used to diagnose leukaemia ; however, bone marrow samples are best [4]. Several research efforts have used machine learning (ML) and computer-aided diagnostic techniques for laboratory image analysis to overcome late leukaemia detection restrictions and discover subgroups in the recent two decades. The key to detecting acute lymphoblastic leukaemia is separating cancer cells from B-lymphoid progenitors. Cancer cells look like normal cells in microscopic pictures, making differentiation difficult. Quantitative blood sample analysis by machine learning or deep learning-based CAD systems addresses such challenges. Several studies recommend segmenting lymphocyte images to get accurate characteristics from the region of interest. K-means, watershed, and HSV color-based segmentation are examples [5].

---

Bamidele, S. A., Asaju-Gbolagade A.W. and Gbolagade, K. A. (2024). An Optimal Detection for Leukaemia Cancer Based On RNS-Metaheuristic Technique in Micro Array Dataset, *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)*, Vol. 11 No. 1, pp. 1 – 13

These segmentations remove unnecessary blood constituents, leaving lymphocytes and lymphoblasts in WBC [6]. Pixel, area, and shape-based segmentations are employed for leukaemia. K-means and edge-based segmentation are used to separate blast cells from blood smears [7, 8].

Ant Colony Optimization (ACO) handles discrete optimization problems using ant colonies as inspiration. Dorigo introduced ACO as the ant system. AS was created for computational intelligence combinatorial optimization [9]. It was first applied to the traveling salesman problem, then to other tough problems. Pheromone on the ground by ants marks a suitable colony path, which inspired AS. A colony finds the fastest route between food and nest. A self-organized ant colony is governed by positive feedback (pheromone deposit) and negative feedback (evaporation). If all edges have constant pheromone levels, no path is determined. Feedback allows minimal edge selection. The approach moves from being unstable with no stronger edges to stable with the strongest.

Numerous scholars have put forth methodologies grounded in machine learning and deep learning algorithms in order to facilitate the timely identification of acute lymphoblastic leukaemia. However, previous studies have shown that these methods have limitations in terms of achieving optimal performance. Additionally, these studies did not take into account the use of feature extraction and selection strategies, as highlighted by [10]. Given the aforementioned constraints, this study presents a methodology that utilizes various image processing techniques, application of Histogram of Gradient for feature extraction, fusion of Ant Colony Optimization and Residual Number System for feature selection and the implementation of a deep learning model for detection and classification of Leukaemia Cancer.

## 2. Literature Review

Leukocytes are classified into lymphoid and myeloid stem cells using a powerful classifier system to separate and classifies blood [11]. According to the investigation, 100 sample photos may be lymphoid or myeloid stem cells

based on leukaemia detection. The approach was evaluated using KMeans clustering. The partitioned cytoplasm and nucleus provide aesthetic form and structure indicators.

A non-parametric method was used with KNN algorithm for pattern recognition and statistical estimation in the early 1970s [12]. Due to the fact that it classifies new states based on similarity (e.g., distance functions) and saves all accessible state. Out of 72 samples, 66 were identified. System accuracy is 91.66%.

An acute leukaemia detection algorithm was developed [13]. Simple improvements, anatomy, filters, and segmentation are used with a K-mean clustering approach. A Naïve Bayes Classifier and Nearest Neighbor (k-NN) algorithm with 92.8% accuracy is evaluated over a range of 60.

A neural network (NN)-based method to differentiate normal and abnormal blood cell pictures was presented [14]. The proposed system has 96.6 percent accuracy. on test the NN architecture's precision and reliability, the authors used it on a much-increased dataset. Only acute lymphoblastic leukaemia (ALL) was included in the dataset.

Random forest classifier was employed to classify lymphocytes as blastocysts or normal cells [15]. An enthusiastic learner, the Random Forest Classifier generates decision trees and class output, or class mode, during training. It works well with large, stable databases in high-dimensional areas. Subsets were randomly selected using the random forest algorithm. This arrangement obtained 95% accuracy; outstanding.

A decision support system and clustering algorithm were used to simulate differential measurements (SDMs) in lymphocyte nucleus and cytoplasm segmentation. Researchers used a fresh data set of four leukaemia kinds to evaluate the architecture. Further research is needed to optimize weight initialization and activation algorithms to improve neural network design [16].

Researchers proposed two goals, first by removing function-containing cells and second is to distinguish ordinary and used shape and histogram features to analyze using k-nearest

neighbor categorization [17]. Sibling values 1, 3, 5, 7, 9, 11, 13, and 15 had k variance. Precision, sensitivity, and specificity were best with  $k=7$  utilizing area-perimeter-mean-standard deviation characteristics of 90%, 90%, and 90%.

Naïve Bayes method was employed to detect blast cells using extracted attributes from fragments. The suggested classifier analyzes blast cell data to detect their presence. Merging [18]. the Naïve Bayes classifier and gene index is recommended. The classifier uses Naïve Bayes indicator results to classify. It is clear and effective, due of its fast convergence and inequality tolerance.

### 3. Research Methodology

Implementing an effective RNS-Based Deep Learning Framework for Leukaemia Cancer Detection in Microarray dataset involves two primary phases for the experimental set of the analytical framework: Training (Enrolment) and Testing (Detection and Recognition). The Kaggle dataset repository was used for training (Enrolment). Image pre-processing includes geometric and photometric normalization after acquisition. Image pre-processing improves clarity, controls lighting, resizes, and crops images.

The output of pre-processed images is data augmented through random sampling, then the augmented leukaemia images are segmented to extract the region of interest, which is then passed to the feature extraction phase to extract and store data as a template using the Histogram of Gradient. After the extracted template is passed into the feature engineering phase, ACO was used to reduce high dimensionality, but RNS was added to improve subset feature selection.

The ACO calculations like pheromone update and solution fitness evaluation which can be computationally demanding.

- i. ACO employs pheromone evaporation and deposition rates. These settings are sensitive to problem magnitude, making setting challenging. RNS' modular arithmetic may reduce scaling factor sensitivity, improving resilience.
- ii. RNS can conserve memory space and improve pheromone storage and retrieval in ACO, where huge amounts of data must be saved.

The selected features are introduced to the Convolutional Neural Network (CNN), which adapted, trained, and created experimental knowledge from the reduced features. The model is saved at the CNN model accuracy peak. The 25%-fold hold out data triggers unfamiliar leukaemia images during classification. Testing results is used to evaluate system performance and validate its results. This adapted approach is described in Figure 1.

#### 3.1 Image processing and Photometric Normalization

The leukaemia -coloured images was first converted to gray scale, followed by the Contrast limited adaptive histogram equalization which was used for the photometric illumination control on the image while the geometric resized the image to an appropriate smaller and undistorted dimension window size of 350 X 350.

##### i. RGB to GrayScale

Grayscale photographs are more straightforward to manipulate and can aid in reducing the complexity of the data while preserving significant visual details.



Figure 1: Sample Dataset of the Acquired Image

The system flow chart of the proposed model is given in the figure.2.

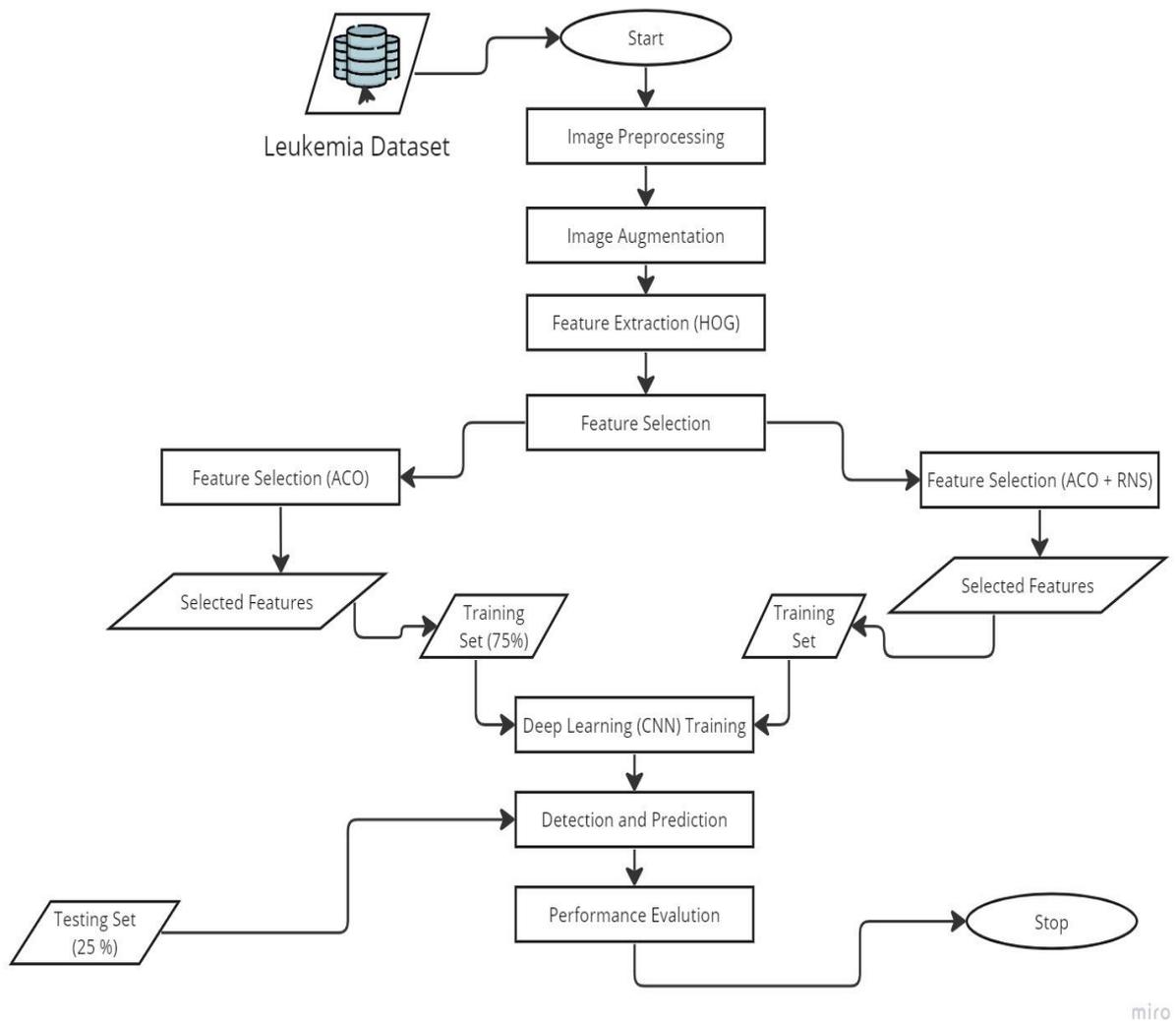


Figure 2: System Flowchart

## ii. Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE, which stands for Contrast Limited Adaptive Histogram Equalization, is a method employed in image processing to improve the contrast of images while also restricting the amplification of noise. CLAHE utilizes a technique where an image is partitioned into small, overlapping sections and then performs histogram equalization to each section, so enhancing the contrast in a localized manner. This technique aids in preserving fine features of an image while preventing excessive amplification of noise. Here is a description of the functioning of CLAHE and its corresponding mathematical formula:

$$I'(x, y) = C(I(x, y)) * L_{max} \quad (1)$$

Where,  $I'(x, y)$  is the transformed pixel value,  $C(I(x, y))$  is the CDF value corresponding to  $I(x, y)$ , and  $L_{max}$  is the maximum pixel value.

## 3.2 Data Augmentation (Random Sampling)

Data augmentation is a common machine learning and computer vision approach that artificially increases a training dataset by making various changes to it. These modifications help machine learning models anticipate accurately and endure data changes. The Random Sampling Method adds data points by randomly selecting samples from the original dataset. The dataset

was randomized within the training data folds and testing folder sets.

### 3.3 Image Segmentation (OTSU Method)

The third stage of the leukaemia detection is to separate the suspicious regions that may contain masses from the background parenchyma. The image will be segmented using Otsu thresholding. Otsu's method, named after Nobuyuki Otsu is used to automatically perform clustering-based image thresholding, or, the reduction of a gray level image to a binary image. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal.

### 3.4 Feature Extraction using Histogram of Oriented Gradients (HOG) Descriptor

The Histogram of Oriented Gradient feature (HOG) was used to extract the feature of the segmented region of the leukaemia images. The HOG descriptor approach quantifies the frequency of gradient orientations in certain areas of an image, known as the detection window or region of interest (ROI).

Recommended values for the HOG parameters are [19]:

- i. 1D centered derivative mask [-1, 0, +1]
- ii. Detection window size is 64x128
- iii. Cell size is 8x8
- iv. Block size is 16x16 (2x2 cells)

### 3.5 Feature Selection using Ant Colony Optimization

The feature engineering phase comprises of a two comparative phase, the Ant Colony Optimization (ACO) was used to select optimal subset only while the Residual Number System was used alongside with the Residual Number System to optimize the performance of ACO. Ant colony optimization is a branch of swarm intelligence inspired by ant foraging [20].

Ants can discover the shortest way between food and colony without direct contact. Ants deposit pheromone between the food source and nest. This pheromone underpins ant communication. The shortest path with the most pheromone is chosen by other ants. Pheromone concentration and heuristic understanding help ants navigate. For shortest path problems like the traveling salesman problem, the ACO algorithm uses the goodness measure to choose the path. A graph shows the issue. Each TSP node represents a city. An agent must complete its tour using the fastest path. After the source, an agent (ant) must choose the next node at each level using a goodness metric. Calculate the selection probability from  $i$ th to  $j$ th node using Equation (2).

$$P_{i,j} = \frac{(\tau_{i,j}^\alpha)(\eta_{i,j}^\beta)}{\sum_{k \in S} (\tau_{i,j}^\alpha)(\eta_{i,j}^\beta)}, \quad (2)$$

The variable  $\tau_{i,j}$  denotes pheromone concentration along the journey from  $i$  to  $j$ , as per equation (2) Value  $\eta_{i,j}$  represents the heuristic function determining the value of choosing  $j$ . The parameters  $\alpha$  and  $\beta$  determine the relative relevance of the pheromone and heuristic function. The equation is divided by the sum of the pheromone and heuristic values of all connected nodes to the  $i$ th node. The next node is dependent on its  $P_{i,j}$  value. After an ant's tour, the pheromone is updated along the path.

$$\tau'_{i,j} = \tau_{i,j} + \tau_{i,j} \cdot \left[ \left( 1 - \frac{1}{1 + P_{i,j}} \right) \right], \quad (3)$$

The symbol  $\Delta_{i,j}$  stands for ant fitness on a certain path, while  $\tau_{i,j}$  represents the pheromone's prior value. Pheromone concentration rises as ants follow a track. To prevent local convergence, the ant algorithm alters pheromone concentration and evaporates some on each

$$\tau_{i,j} = (1 - \rho) \cdot \tau'_{i,j} + \Delta_{i,j}, \quad (4)$$

Here,  $\rho$  is the proportion of the pheromone that will undergo evaporation.

### 3.6 Improved Ant Colony System by Residual Number System (Hybridized System)

Residue number systems (RNS) represent numbers by their remainders when split by numerous pairwise coprime integers, called moduli. The Chinese remainder theorem states that if  $N$  is the moduli multiplication, there is exactly one integer within an interval of length  $N$  with any required modular values.

#### *Forward Conversion*

RNS represents numbers by their residues or remainders after dividing by pairwise coprime moduli. Forward conversion in Residue Number System (RNS) converts integers to residues. In mathematics, the forward conversion of a number  $X$  into its Residues  $R_1, R_2, R_3, \dots, R_n$  with respect to the moduli  $M_1, M_2, M_3, \dots, M_n$  is:

For each modulus  $M_i$ :

$$R_i = X \bmod M_i \quad (5)$$

$X$  is the starting integer to be represented by the Residue Number System. The modulus " $M_i$ " represents the  $i^{\text{th}}$  coprime modulus. The remainder of  $X$  divided by  $M_i$  is  $R_i$ . This produces a set of residues  $R_1, R_2, R_3, \dots, R_n$  that reflect the beginning number  $X$  in the Residue Number System (RNS). Residue Number System (RNS) may perform arithmetic operations autonomously on each residue, speeding up and improving computations, especially for large number.

The Hold-Out approach is a straightforward and widely used method for dividing a dataset into partitions. The process of the data splitting for the reduced leukaemia dataset adopted 75% for training and 25% for testing.

The mathematical construct represents and updates pheromone levels using RNS. RNS-encoded values show pheromone quantities for each decision

variable combination. RNS arithmetic is used to efficiently compute and update RNS-encoded pheromone levels during the update phase. RNS advantages like effective modular arithmetic can accelerate the convergence of the ACO algorithm by allowing ants to compute more efficiently during solution space exploration and pheromone information update. RNS can be used to update pheromones, a crucial part of ACO.

#### *Pseudo-Code of the Hybridized RNS+ACO*

##### **1. Define the problem space:**

- Define the decision variables and their respective domains.

##### **2. Initialize ACO parameters:**

- Initialize the number of ants, maximum iterations, pheromone evaporation rate, and other ACO-specific parameters.

##### **3. Initialize RNS parameters:**

- Choose suitable moduli ( $m_1, m_2, \dots, m_k$ ) for the RNS representation. These moduli should be coprime to each other and typically span a wide range of values.
- Define the bases ( $b_1, b_2, \dots, b_k$ ) for each modulus.

##### **4. Initialize pheromone matrix P:**

- Initialize the pheromone levels for each combination of decision variables using the RNS representation. For each combination ( $x_1, x_2, \dots, x_n$ ), initialize  $P(x_1, x_2, \dots, x_n)$  using RNS-encoded values.

## Main Loop: 5. For each iteration:

### a. Ant Construction Phase:

- For each ant, construct a solution using RNS arithmetic:
- Initialize an empty solution  $S$ .
- For each decision variable  $x_i$ :
- Compute RNS-encoded values  $r_i$  based on the available pheromone information.
- Decode RNS-encoded values to obtain  $x_i$  in the original domain.
- Add  $x_i$  to the solution  $S$ .

### b. Evaluate solutions:

- Evaluate the fitness of each ant's solution  $S$  based on the problem's objective function.

### c. Pheromone Update Phase:

- Evaporate pheromone levels: Decrease all pheromone values by a factor  $(1 - \text{evaporation rate})$ .

### d. Select the best solution found so far.

## 6. Termination Criteria:

- Continue iterating until a termination criterion is met (e.g., a maximum number of iterations or convergence criteria).

## 7. Result:

- Return the best solution found by the ACO algorithm.

### 3.7 CNN Architecture

Convolutional and subsampling levels may be followed by fully connected layers in a CNN. Convolutional layers receive images having dimensions of  $m \times m \times r$ , where  $m$  is the image's height and width and  $r$  is its channels. RGB images have  $r = 3$ . The convolutional layer will include  $k$  filters (or kernels) with dimensions  $n \times n \times q$ , where  $n$  is smaller than the image's dimension and  $q$  can be equal to or smaller than  $r$ , depending on the kernel. Locally connected arrangement is determined by filter dimensions, where each filter

convolves with the picture to yield  $k$  feature maps of size  $m-n+1$ . Further, each map is downsampled using mean or max pooling over continuous patches of size  $p \times p$  smaller images (e.g., MNIST) have a  $p$  value of 2 while larger inputs rarely surpass 5. Each feature map receives a bias term and sigmoidal nonlinearity function before or after the subsampling layer. Convolutional and subsampling sublayers make up a CNN layer, as shown in Figure 3. Color-matched units weigh the same.

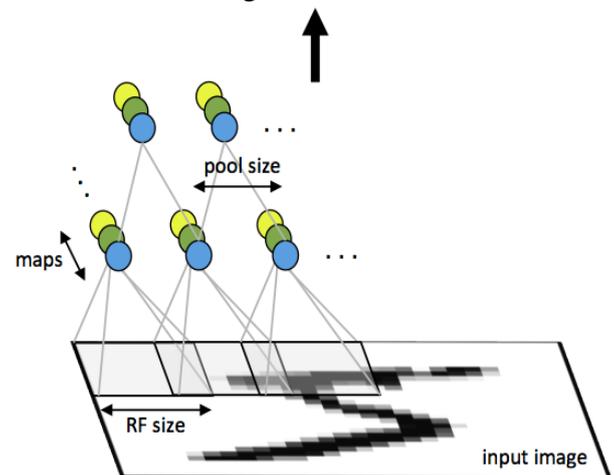


Figure 3: First layer of a convolutional neural network with pooling. Units of the same color have tied weights and units of different color represent different filter maps. Source: [7].

After convolutional layers, any number of fully linked layers can follow. The densely connected layers match a multilayer neural network. CNNs have input, convolutional, pooling, and fully connected layers.

## 4. Results and Discussion

The grayscale image's normalized histogram and the CLAHE-enhanced displayed a well-balanced since the value for the entire bin, which displays pixel frequency, is equally distributed. The "C-NMC\_Leukaemia" biomedical dataset was sampled using the Random Sampling Method. 2600 malignant ALL-Leukaemia cases and 1400 non-cancerous cases were randomly selected. Otsu's thresholding approach minimizes intraclass variation between black and white pixels in the threshold image. This

approach projects a two-dimensional histogram onto the diagonal and uses a 2D Otsu algorithm to find the best threshold value. The investigation showed that it directly influences the gray point histogram, increasing threshold speed and noise resistance. Figure.4 shows the masked region where the tumour is located which is achieved by the morphological operation that took place while Figure. 5 shows the sample region of abnormality of the leukaemia image.

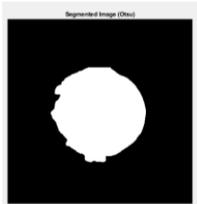


Figure 4: Otsu Threshold Image

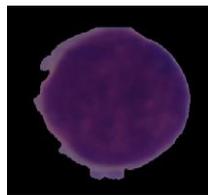


Figure 5: Region of Abnormality

HOG counts gradient orientation in localized brain tumor images to find significant things in the segmented brain picture using its feature descriptor. The lower-dimensional images that produced the leukaemia image feature description are shown in Figure 4. The features extracted 108900 descriptors for each leukaemia segmented image, leading in a 1X108900 dimension size, which will impair system performance. Ant Colony Optimization and RNS were added for optimal feature subset selection.

Feature Selection process with ACO reduced the histogram of gradient features to 12526 descriptors, a very significant dimensionality reduction from the original dimensional space, allowing the features to extract 12526 descriptors for each leukaemia segmented image, resulting in a 1X12526 dimension size.

Furthermore, RNS math was utilized to efficiently compute and update RNS-encoded pheromone levels. The RNS used modular arithmetic to speed up ACO algorithm convergence by allowing ants to compute faster during solution space search and pheromone information update. RNS moduli set:  $S_2 = \{2n-1, 2n, 2n+1\}$ , yielding  $m_1=9$ ,  $m_2=10$ , and  $m_3=11$ . The ACO algorithm performed better after the RNs provided features with fewer descriptors of 3752, resulting in 1X 3752-dimension size for each Leukaemia image in the dataset. The Convolution Neural Network received these

last characteristics as indicated in Figures 6 and 7 respectively.

Figure 6: ACO Selected Features

Figure 7: ACO+RNS Selected Features

The CNN was used to build the classifier mode: the model assigns class labels to problem occurrences, which are represented as vectors of feature values, with the class labels selected from a finite set, an instance of a correctly classified Leukaemia

The optimal features as reduced by the RNS+ACO algorithm was passed into the CNN at testing percentage of 25%, the CNN is accessed through the Deep Network Designer toolbox in MATLAB and was used to test the selected features.

#### 4.1 Performance Evaluation

Performance evaluation metrics must accurately represent the task's objectives and account for false positives and negatives. True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), Error rate, Classification Accuracy, Sensitivity, Specificity, and Error Rate evaluated performance for leukaemia detection or any medical diagnostic assignment.

This section provides the performance evaluation having subject the testing data to a partition value of 25%. The Evaluation was carried out for two major instances highlighted below:

**Instance 1:** Image Pre-processing + HOG + ACO + CNN

**Instance 2:** Image Pre-processing + HOG + ACO+RNS + CNN

It is observed that the **Instance 2:** Image Pre-processing + HOG + ACO + RNS + CNN outperformed the **Instance 1:** Image

Table 1: Comparative Evaluation of Instance 1 & 2.

| Techniques                                  | F-score | Precision | Specificity | Sensitivity | Accuracy (%) | Error Rate (%) | Training Time (secs) |
|---|---------|-----------|-------------|-------------|--------------|----------------|----------------------|
| Image Pre-processing + HOG + ACO + CNN      | 0.9374  | 0.9629    | 0.9874      | 0.9133      | 96.79        | 0.0321         | 154.672              |
| Image Preprocessing + HOG + ACO + RNS + CNN | 0.9672  | 0.9686    | 0.9895      | 0.9658      | 98.36        | 0.0164         | 83.124               |

Pre-processing + HOG + ACO + CNN both in terms of positive rate and negative rate detection.

The training time shows the time taken by the model to create knowledge retention of the data supplied to the CNN classifier is shown in fig.8. The total training time taken for the CNN model to memorize the optimal dataset is seen to be more efficient with the Instance 2 model: Image Pre-processing + HOG + ACO + RNS + CNN having a better optimal training time of 83.124 secs as compared to the instance 1 which had a longer training time of 154.672 secs.

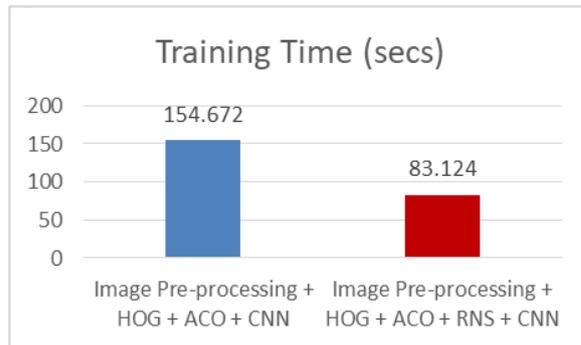


Figure 8: Training time for both instances

The classification accuracy shows the correct classification rate attained by the CNN deep learning Classifier. The classification accuracy in percentage shows the percentage of instances that were classified correctly. The classification accuracy shows the percentage of instances that was classified correctly, furthermore, the higher

classification accuracy rate is with Instance 2 model **Image Pre-processing + HOG + ACO + RNS + CNN** as revealed in figure 9. Therefore, validating the importance of RNS to enhance the Ant Colony Optimization Algorithm

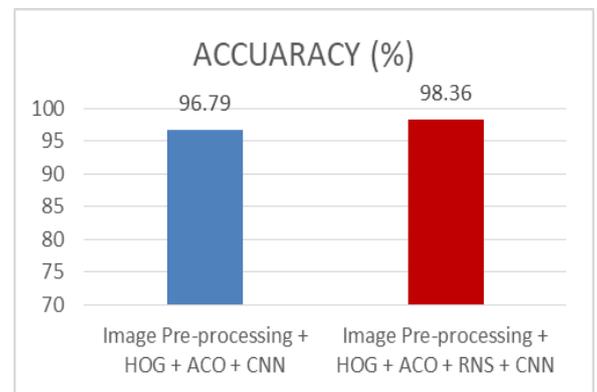


Figure 9: Classification Accuracy

The Sensitivity (SN) is calculated as the number of correct positive predictions divided by the total number of positives, Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. The best sensitivity and specificity fall at 1. From the obtained results shows the sensitivity and the specificity rate has value close 1 for Instance 2 model: Image Pre-processing + HOG + ACO + RNS + CNN as compared to the instance 1 model: Image Pre-processing + HOG + ACO + CNN which is indicative of better positive and negative predictive rate strength as indicated in figure 10.

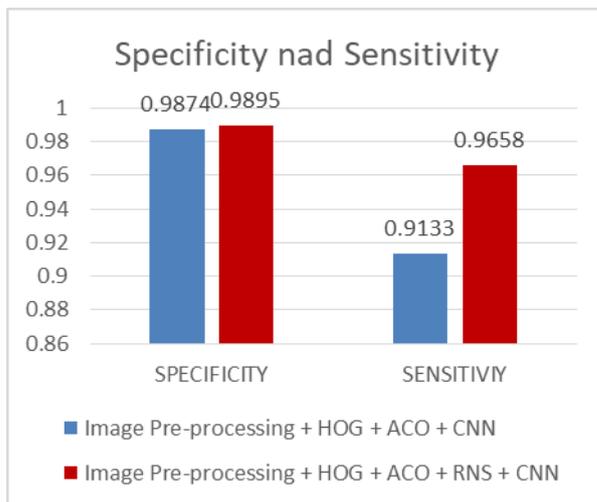


Figure 10: Specificity and Sensitivity

The error rate shows the lowest possible error rate for any classifier in a random outcome during the classification. The **Instance 2 model:** Image Pre-processing + HOG + ACO + RNS + CNN shows a very low error rate of 0.0164 which tends towards zero more than the instance 1 model.

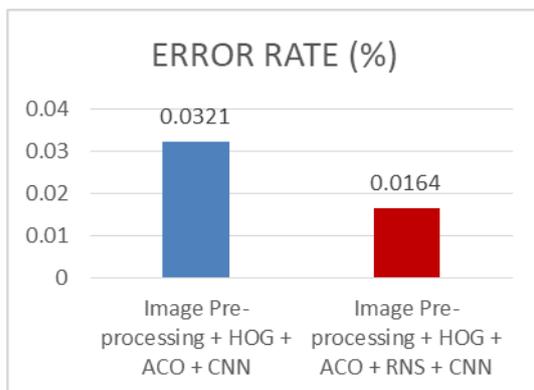


Figure 11: Error Rate

The F1 Score is calculated as the harmonic mean of precision and recall, serving as a balanced measure of both measurements. It is particularly advantageous when seeking to achieve an equilibrium between incorrect positive results and incorrect negative results. Precision is a metric that evaluates the accuracy of a model in correctly identifying positive instances while minimizing the occurrence of incorrectly categorizing negative examples as positive. It is seen that the **Instance 2 model:** Image Pre-processing + HOG + ACO + RNS + CNN also outperformed

the **Instance 1 model:** Image Pre-processing + HOG + ACO + CNN.

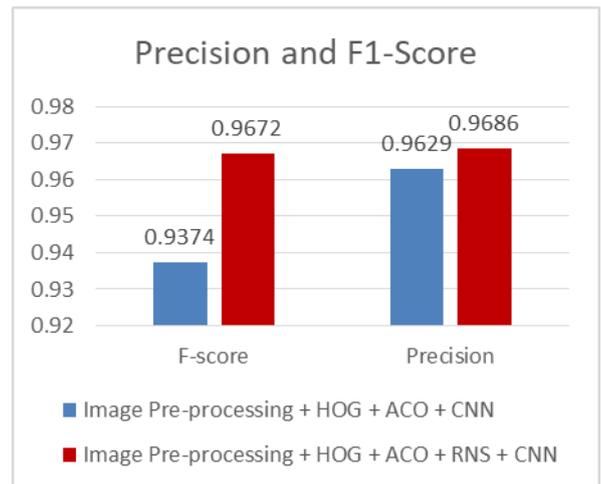


Figure 12: F1-Score and Precision Value

## 4.2 Discussion

It is well known that RNS cannot stand on its own rather, it is good in enhancing the process of some existing metaheuristic algorithm in simulating the behaviour of some existing features selection. Ant colony optimization (ACO) in terms of accuracy results in 96.79%, error rate of 0.0321% and training time of 154.672sec., with splitting rate of 75% and 25% for the training phase and the testing phase respectively.

When subjected to this process (Image Pre-processing + HOG + ACO + CNN) in predicting the detection of leukaemia. ACO only help in calculating the distance, time and reducing a robust system that can solve issues with large search space and high quality solutions, i.e. it reduces the dimensionality of extract descriptions (images or dataset). Image Pre-processing + HOG + ACO + RNS + CNN, enhance the process in terms of accuracy results of 98.36%, error rate 0.0164% and training time of 83.124sec.

## 5. Conclusion

Ultimately, the thorough experimentation and analysis conducted in this study emphasize the efficacy of the suggested methodology for detecting leukaemia malignancy. The application of a

combined method incorporating Residual Number System (RNS) and Deep Learning Framework, particularly Convolutional Neural Network (CNN), demonstrated encouraging outcomes. The systematic image processing steps, starting from converting the image to grayscale and enhancing the contrast, followed by extracting features using Histogram of Oriented Gradients (HOG), formed the basis for reliable analysis of leukaemia images. The incorporation of Ant Colony Optimization (ACO) and RNS during the feature selection stage was crucial in reducing the number of dimensions, hence enhancing the efficiency of the dataset for training Convolutional Neural Networks (CNN).

The comparative evaluations yielded valuable insights, repeatedly showing that Instance 2 outperformed Instance 1. The graphical analyses provided additional evidence of Instance 2's superior performance in terms of classification accuracy, sensitivity, specificity, and computational training time. The adapted methodology, incorporating unique feature selection techniques and a deep learning architecture, shows significant potential for precise and quick diagnosis of leukaemia malignancy in reported speech. The study findings reported here make a substantial contribution to the developing field of medical image analysis and computational techniques in healthcare.

## References

- [1.] Gehlot, S., Gupta, A., & Gupta, R. (2020). SDCT-auxNet $\theta$ : DCT Augmented Stain Deconvolutional CNN with Auxiliary Classifier for Cancer Diagnosis. *Med. Image Anal.* 61, 101661.
- [2.] Mondal, C., Hasan, K., Ahmad, M., Awal, A., Jawad, T., Dutta, A., Islam, R., & Moni, M.A. (2021). Ensemble of Convolutional Neural Networks to Diagnose Acute Lymphoblastic Leukaemia from Microscopic Images. *Inform. Med. Unlocked* 27, 100794
- [3.] Hegde, R. B., Prasad, K., Hebbar, H., Singh, B. M. K., & Sandhya, I. (2019) Automated decision support system for detection of leukaemia from peripheral blood smear images. *J Digit Imaging* 33:361–374
- [4.] Namayandeh, S. M., Khazaei, Z., LariNajafi, M., Goodarzi, E., & Moslem, A. (2020) GLOBAL Leukaemia in children 0–14 statistics 2018, incidence and mortality and human development index (HDI): GLOBOCAN sources and methods. *Asian Pac J Cancer Prevent* 21(5):1487–1494
- [5.] Jagadev, P., & Virani, H. (2017). “Detection of leukaemia and its types using image processing and machine learning,” in *Proceedings of the 2017 International Conference on Trends in Electronics and Informatics (ICEI)*, pp. 522–526, IEEE, Tirunelveli, India.
- [6.] Ghaderzadeh, M., Asadi, F., Hosseini A., Bashash, D., Abolghasemi, H., & Roshanpour, A., (2021) “Machine learning in detection and classification of leukaemia using smear blood images: a systematic review,” *Scientific Programming*.
- [7.] Alamgir Sarder, M., Maniruzzaman, M., & Ahammed, B. (2020) “Feature selection and classification of leukaemia cancer using machine learning techniques,” *Machine Learning Research*, vol. 5, no. 2, p. 18
- [8.] Al-jaboriy, S. S., Sjarif, N. N. A., Chuprat, S., & Abdulllah, W. M. (2019). “Acute lymphoblastic leukaemia segmentation using local pixel information,” *Pattern Recognition Letters*, vol. 125, pp. 85–90
- [9.] Dorigo, M. (1992). *Optimization, learning and natural algorithms* [Ph.D. thesis], Politecnico di Milano, Milan, Italy.
- [10.] Hossain, M. A., Sabik, M. I., Rahman, M., Sakiba, S. N., Muzahidul Islam, A., Shatabda, S., et al. (2021). “An effective leukaemia prediction technique using supervised machine learning classification algorithm,” in *Proceedings of International Conference on Trends in Computational and Cognitive Engineering* (Springer), 219–229. doi: 10.1007/978-981-33-4673-4\_19
- [11.] Engineering, T. (2014). “Acute Leukaemia Classification by Using SVM and K-Means Clustering,” pp. 1–4.
- [12.] Gumble, P. M. (2017). “Analysis & Classification of Acute Lymphoblastic Leukaemia using KNN Algorithm,” pp. 94–98.
- [13.] Kumar, S., Mishra, S., & Asthana, P. (2018). Automated Detection of Acute Leukaemia Using K-mean Clustering Algorithm.
- [14.] Thanh, T. T. P., Vununu, C., Atoev, S., Lee, S., & Kwon, K. (2018). “Leukaemia Blood Cell Image Classification Using Convolutional Neural Network,” vol. 10, no. 2.

- [15.] Souza, R. D., & Fernandes, R. (2018). "LEUKAEMIA PREDICTION USING RANDOM FOREST ALGORITHM," vol. 8, no. 3, pp. 1–8.
- [16.] Umamaheswari, D. & Geetha, S. (2018) "Segmentation and classification of acute lymphoblastic leukaemia cells tooled with digital image processing and ML techniques," in Proc. 2nd Int. Conf. Intell. Comput. Control Syst. (ICICCS), Jun. 2018, pp. 1336–1341.
- [17.] Wisesty, U. N., Warastri, R. S., & Puspitasari, S. Y. (2020) "Detection of acute lymphocyte leukaemia using knearest neighbor algorithm based on shape and histogram features Detection of acute lymphocyte leukaemia using k-nearest neighbor algorithm based on shape and histogram features."
- [18.] Das, B. K. & Dutta, H. S. (2020). GFNB: Gini index – based Fuzzy Naive Bayes and blast cell segmentation for leukaemia detection using multi-cell blood smear images,".
- [19.] Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. In Proc. of INRIA Rhone-Alps, 655 avenue de l'Europe, Montbonnot 38334, France.
- [20.] Khan, M. A., Khan, T. M., Bailey, D. G., & Kittaneh, O. (2018) "Deriving scale normalisation factors for a glog detector," IET Image Processing, vol. 12, no. 9, pp. 1673–1682