Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Classification of Acute Leukemia Based on Multilayer Perceptron

To cite this article: Nurul Hazwani Abd Halim *et al* 2019 *J. Phys.: Conf. Ser.* **1372** 012044

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

This content was downloaded from IP address 103.86.131.118 on 16/12/2020 at 06:19

Classification of Acute Leukemia Based on Multilayer Perceptron

Nurul Hazwani Abd Halim[1*] Mohd Yusoff Mashor[1] Rosline Hassan[2]

[1]Electronic & Biomedical Intelligent Systems (EBItS) Research Group,
School of Mechatronics Engineering, University Malaysia Perlis,
02600 Kampus Pauh Putra, Perlis, Malaysia.
[2]Department of Hematology, School of Medical Science, University Science Malaysia,
Kubang Kerian, Kelantan, Malaysia.

[*]nurul.hazwani43@yahoo.com.

Abstract. In general, various artificial neural network have been applied in many areas such as modelling, pattern recognition, signal processing, diagnostic and prognostic. In this paper, artificial neural network are used to detect and classify the white blood cell (WBC) inside the acute leukemia blood samples. There are 25 features have been extracted from segmented WBC, which consist of shape, color and texture based features. Then, it have been fed up as the neural network inputs for the classification process in order to classify the segmented regions into two classes either B or T. The training algorithm for MLP network is Levenberg-Marquardt (LM). The MLP network achieves the highest testing accuracy of 96.99% for 4 hidden nodes at state of 5 by using the overall 25 input features. Thus, MLP network trained by using LM algorithm is suitable for acute leukemia cells detection in blood sample.

1. Introduction

Leukemia is a blood cancer, which can affect both children and adults. Acute leukemia can be divided into two types: Acute Lymphoblastic Leukemia (ALL) and Acute Myelogenous Leukemia (AML) [1]. It is wise to identify the types of leukemia early and rapidly assist the patients by giving appropriate treatments. In addition, the process to classify the type of leukemia is depending on the skills and experience of the hematologist. However, the manual microscopic examination has several drawbacks. Even for an expert, this traditional method can be very tedious and furthermore time-consuming [2].

Several development of semi and automated detection system related to leukemia have been proposed. Shafique *et*.al [3] proposed blood detection that utilizes thresholding based on Zack algorithm and shape and color based features were extracted from the segmented white blood cell (WBC).It able to achieve accuracy of 93.70% by using support vector machine (SVM) in order to classify either normal or blast cells. Meanwhile, Asadi *et.*al used Backpropagation Neural Network to classify ALL and AML, which produced 86.66% of accuracy [4]. Rajpurohit *et*.al proposed various classifiers such as convolutional neural network (CNN), feed-forward neural network (FNN), support vector machine (SVM) and K-Nearest Neighbor (KNN) with their following accuracies respectively, 98.33%, 95.40%, 91.40% and 93.30%. CNN provided a highest accuracy compare to other classifiers [5]. In this study, a multilayer perceptron by training using Levenberg-Marquardt (LM) has been proposed in order to classify the acute lymphoblastic leukemia either it is B or T. The motivation to do this classify because the hematologist face some difficulty to distinguish these cells unless do some specific test such as flow cytometry. The following section will be discussed about several image processing and how to implement this classifier effectively.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

2. Methodology

The block diagram below shown the proposed classification method for acute leukemia images which consist of several image processing techniques such as image enhancement, color thresholding, feature extraction and classification.

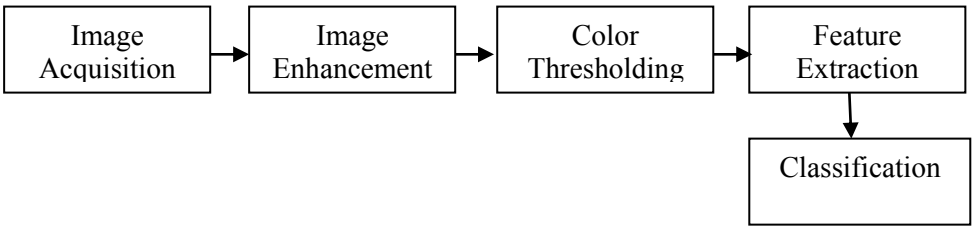
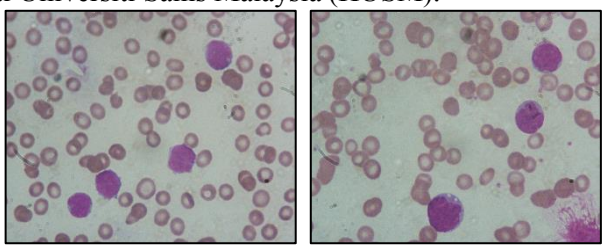


Figure 1: The proposed method for classification of acute lymphoblastic leukemia.

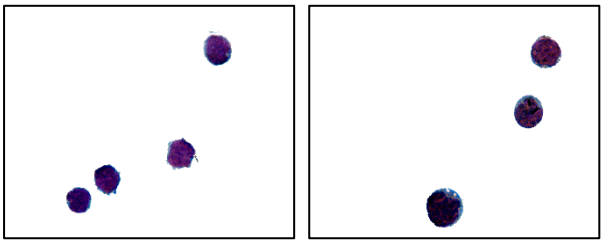
2.1. Image Acquisition

Figure 2 (a) and (b) were some examples of the captured images of acute lymphoblastic leukemia for type B and T, respectively, which, were captured under 100X magnification and at a resolution of 1280×960 pixels in Hospital Universiti Sains Malaysia (HUSM).



(a) Type B (b) Type T

Figure 2. Sample images of acute lymphoblastic leukemia (ALL).



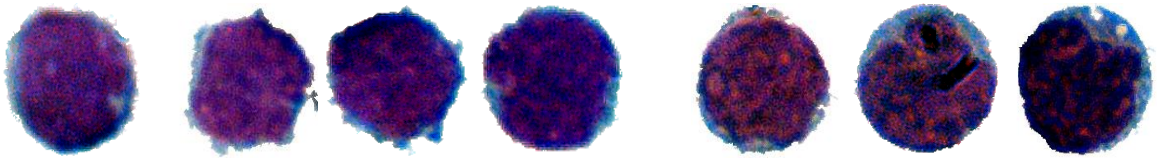
(a) Type B (b) Type T

Figure 3. Segmented images of acute lymphoblastic leukemia (ALL).

2.2. *Image Processing*

The effectiveness of the image enhancement process will make it easier for image segmentation, features extraction and classification of the blood sample slide images to identify leukemia. Thus, local contrast stretching has been implemented on B and T types of ALL images. By implementing this algorithm, each red, green and blue color space will be distributed linearly over the whole histogram so that the dynamic range of the histogram is fulfilled (0 – 255)[6][7]. After that, automatic color thresholding based on HSI (Hue, Saturation, Intensity) color space has been applied. In order to segment white blood cell (WBC) from the background and red blood cell (RBC), Hue component has been extracted from HSI color space. Based on the previous study, Hue component can provided a

fully information about WBC while saturation component contains information about nucleus only [8]. The threshold value is set to 0.5. Any region which has greater than 0.5 will be considered as the WBC and the rest will be eliminated from the images as shown in Figure 3.



(a) Type B (b) Type T

Figure 4. Zooms of Segmented images of acute lymphoblastic leukemia (ALL).

2.3. *Feature Extraction*

Figure 4 shown some individual cells from the segmented images for each category. In order to distinguish WBC either it is type B or type T, it must have fulfilled the characteristic as below [9]:

- Type B has rather small, uniform blast cells with scanty cytoplasm and rounded with usually a single nucleolus and have a smudgy homogenous chromatin without prominent nucleoli.
- Type T has varied lymphoblasts, including numerous larger blasts with more open chromatin, prominent nucleoli and abundant cytoplasm.

After that, feature extraction is used to measure the properties of WBC. A number of approaches have been developed for feature extraction in acute leukemia identification system, such as the geometrical features [10][11], texture features [12][13], and the combination of geometrical ,texture and color features[14][15][16]. Before that, all of the WBC should be crop manually in order to extract features efficiently. There are some shape and geometrical based features that have been extracted such as area, perimeter, convex area, eccentricity, solidity, circularity and Affine Moment Invariant. For texture-based features such as contrast, correlation, energy and entropy have been extracted from segmented WBC images. Finally, color-based features are also extracted such as standard deviation and mean of RGB (Red, Green,Blue) color space. In total, there are 25 features have been extracted from the segmented WBC of both types of ALL which were then fed up as the neural network inputs for the classification. Beforehand, the features must be normalized between 0 to 1 in order to achieve high performance of classification. Noted that these features have been implemented on whole images in this study and the choice of the features has been driven by suggestions of the experts in HUSM and validate by them.

2.4. Classification using Multilayer Perceptron

Multilayer Perceptrons are gaining popularity in classification task due to its flexibility, robustness and high computational rates [17]. The classification performance of the MLP network depends on the structure of the network and training algorithm. The MLP network consists of an input layer that accepts the input data used in the classification, hidden layers and an output layer as shown in Figure 5. In this study, the Levenberg-Marquardt (LM) was used as a training algorithm due to it has a much better learning rate and can keep the relative stability.Detailed about LM algorithm can be found here [18].

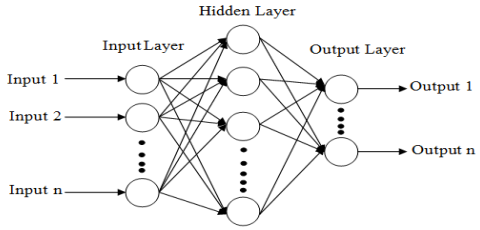


Figure 5. Schematic diagram of MLP model with 1 hidden layer

For the classification of overall WBC inside acute leukemia blood samples, there are a total of 1062 WBCs have been segmented from B-type and T-type images. The data division is set into 70% for training phase while 30% for the testing phase as shown in Table 1 below. In this study, the comparison has been done in five states in order to obtain the optimum number of hidden nodes for training and testing data. For the hidden node, the numbers used during the training phase are varied from 1 to 30 nodes, with the interval of 1 until the best classification result is achieved.The states are referring to the initial value assigned to the function when a random number generator (RNG) is called. The structure of the MLP network is set to 25:10:1 (input node: hidden node: output node). The other parameters are set such as training algorithm = LM, goal = 0.001, the number of epochs is set to 1000 and activation function is tan-sigmoid. Noted that,the evaluation is based on accuracy only due to the analysis is based on positive samples,which are B and T, hence, sensitivity and specificity are neglected in this context.

Table 1. Input data division for classification method

Overall Dataset	Training Dataset (70%)	Testing Dataset (30%)
B = 601 T = 461	B = 421 T = 323	B = 180 T = 138
Total = 1062	Total = 744	Total = 318

3. Results&Discussions

In this section, the classification performance based on overall features will be elaborated here. For this purpose, a total of 25 input features have been fed into the MLP network. There are two different analyses that have been conducted which are analysis of finding the best number of states and the best number of hidden nodes. All of this can be obtained when the MLP network achieved the highest testing result. The states refer to the initial values assigned to the function when a random number of generator is called. Different initial random values will produce different results. Hence, the testing has been done in five states to choose the best structure for MLP network.

. In order to avoid the problem of over-fitting, it is necessary to determine the best number of hidden nodes. Table 2 and Figure 6 show the analysis of number of hidden nodes for classification between B and T using MLP(LM) network.Based on Table 2, the best classification performance is obtained at state 5 and number of hidden nodes of 4 with testing accuracy of 96.99%.The results also show that the testing accuracy better than 90% has been archieved for other state.

Table 2. Analysis of number of hidden node for classification between T-type and B-type WBC using MLP_LM network

State	Number of Hidden Node	Training Accuracy (%)	Testing Accuracy (%)
1	29	99.61	95.32
2	29	99.48	95.65
3	30	99.48	95.97
4	21	99.74	94.31
5	**4**	**97.51**	**96.99**

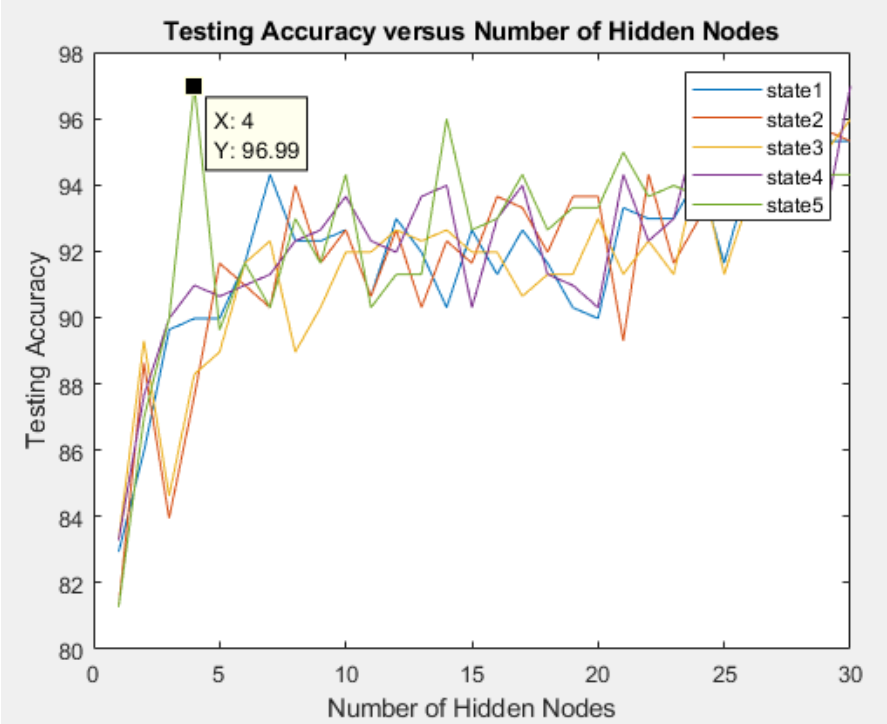


Figure 6. Analysis of Number of Hidden Nodes for classification between B and T using MLP(LM) network

4. Conclusion

It is shown that MLP_LM network produces the best classification performance of 96.99% at state of 5 using 4 hidden nodes. Hence, MLP network is suitable for detection of acute leukemia cells in order to classify either it is B or T type of leukemia.

Acknowledgments
The authors gratefully acknowledge to the National Science Fund (NSF) team members and University Science Malaysia (USM).

References

[1] Rajendran S 2007 Image Acquisition and Retrieval Systems for Leukaemia Cells *Master's thesis, Department of Computer Science, UM*.

[2] Thung F and Suwardi I S 2011 Blood parasite identitication using feature based recognition *International Conference on Electrical Engineering and Informatics (ICEEJ)* 1-4.

[3] Shafique S, Samabia T, Anas T 2019 Computer-assisted Acute Lymphoblastic Leukemia detection and diagnosis *2nd International Conference on Communication, Computing and Digital Systems (C-CODE)*, 184–189.

[4] Asadi F, Putra F M, Sakinatunnisa M I and Marzuki I 2017 Implementation of Backpropagation Neural Network and Blood Cells Imagery Extraction for Acute Leukemia Classification. *5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)* 106–110.

[5] S.Rajpurohit, S.Patil, & N.Choudhary(2018). Microscopic Blood Image Using Image Processing, (Cll), 2359–2363.

[6] Aimi Salihah A N, Mashor M Y, Harun N H, Abdullah A A and Rosline H 2010 Improving colour image segmentation on acute myelogenous leukaemia images using contrast enhancement techniques *Proceedings of 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES* 246–251.

[7] Mat Isa N A, Mashor M Y and Othman N H 2003 Contrast Enhancement Image Processing on Segmented Pap Smear Cytology Images *International Conferences on Robotic, Vision, Information and Signal Processing* 118 – 125.

[8] Huang C Y and Wu M J 2006 Image Segmentation *Final Report, University of Winconsin-Madison*.

[9] Miwa S 1998 Atlas of Blood Cells. Tokyo, Japan: *Bunkodo Co., Ltd.*.

[10] Chatap N and Shibu S 2014 Analysis of blood samples for counting leukemia cells using Support vector machine and nearest neighbour *IOSR Journal of Computer Engineering (IOSR-JCE)* **16** 79-87.

[11] Faivdullah L, Azahar F, Htike Z Z and Naing W N 2015 Leukemia Detection from Blood Smears *Journal of Medical and Bioengineering* **4** 488-491.

[12] Jyoti Rawat, Annapurna Singh, Bhadauria HS, Jitendra Virmani and Jagtar Singh Devgun 2017 Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia *Biocybernetics and Biomedical Engineering* **37** 637–654.

[13] Morteza Moradi Amin, Saeed Kermani, Ardeshir Talebi and Mostafa Ghelich Oghli 2015 Automatic recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and multiclass support vector machine classifier *Journal of Medical Signals & Sensors* **5**.

[14] Srisukkham W, Zhang L, Neoh S C, Todryk S and Lim C P 2017 Intelligent leukaemia diagnosis with bare-bones PSO based feature optimization *Applied Soft Computing* **56** 405–419.

[15] Lim H N, Francis E U, Mashor M Y and Hassan R 2017 Classification of bone marrow acute leukemia cells using multilayer perceptron network *3rd International Conference on Electronic Design, ICED 2016* 486–490.

[16] Kumar P and Udwadia S M 2017 Automatic detection of Acute Myeloid Leukemia from microscopic blood smear image *2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017* 1803–1808.

[17] Lim T Y, Ratnam M M and Khalid M A 2007 Automatic Classification of Weld Defects using Simulated Data and An MLP Neural Network *Journal INSIGHT, Learned & Professional Society Publisher* **49**(4) 154-159.

[18] Nasir A S A, Mashor M Y and Rosline H 2011 Detection of Acute Leukaemia Cells Using Variety of Features and Neural Networks *5th Kuala Lumpur International Conference on Biomedical Engineering 2011. IFMBE Proceedings* **35** 40–46.