Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Leukemia Blood Cells Detection using Neural Network Classifier

Muhammad Naufal Mansor[1,*], Mohd Zamri Hasan[1,3], Wan Azani Mustafa[1], Farah Hanan Mohd Faudzi[1], Syahrul Affandi Saidi[1], Mohd Aminudin Jamlos[2], Noor Anida Abu Talib[1], Ahmad Kadri Junoh[4]

[1] Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus, Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia
[2] Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, Malaysia
[3] Centre of Excellence for Intelligent Robotics & Autonomous Systems (CIRAS), Universiti Malaysia Perlis, Arau, 02600, Perlis, Malaysia
[4] Institute of Engineering Mathematics, Universiti Malaysia Perlis, Arau, 02600, Perlis, Malaysia

ARTICLE INFO

Article history:
Received 25 May 2023
Received in revised form 20 September 2023
Accepted 29 September 2023
Available online 18 October 2023

Keywords:
Leukemia; blood cells; neural network

ABSTRACT

Image segmentation is an image processing operation performed on the image in order to partition the image into some images based on the information contained in the original image. Image segmentation plays an important role in many medical imaging applications, image segmentation facilitates the anatomy process in a particular body of human body. Classification and clustering are the methods used un data mining for analyzing the data sets and divide them on the basis of some particular classification rules. There are many image segmentation tools that used for medical purpose, so it is necessary to define and/or to improve the image segmentation methods in order to get the best method. In this study, the image of leukemia and red blood cells will be used as samples to determine the best algorithm in image segmentation. The procedure for doing segmentation itself is clustering image, edge detection on image, and image classification. The clustering is to extract important information from an image. The edge detection is to determine the existence of edges of lines in image in order to investigate and localize the desired edge features. Moreover, the classification analyzes the properties of some images and organizes the information into certain categories. In this study, the Neural Network and K-Nearest Neighbor are used for image classification by paired with Local Binary Pattern and Principal Component Analysis. The results revealed that the best method of proven in classifying images is from Local Binary Pattern feature extraction with the average accuracy of 94%.

1. Introduction

Image segmentation has been used in a variety of methods in the scientific field. Medical image segmentation plays an essential function in many clinical diagnoses and pathological researches. In the past decades, many promising strategies have been proposed to handle the segmentation hassle [1-8]. With the fast increase in the population there is a call for an improved prognosis in the field of

* *Corresponding author.*
E-mail address: naufal@unimap.edu.my

https://doi.org/10.37934/araset.33.1.152162

medicine. Thus, an environmentally friendly and extra fantastic approach is required for higher analysis of clinical image information. Segmentation is the primary section in the automated analysis of a disorder where a gadget is given an input such that it segments the enter image and further diagnoses based totally on the brain [9-12].

Image data is one piece of substantially sensible significance in scientific informatics. Medical images such as the Computed Axial Tomography (CAT), Magnetic Resonance Imaging (MRI), Ultrasound, and X-Ray, in general DICOM format are often stored in Picture Archiving and Communication Systems (PACS) and associated with different clinical information in the medical EHR administration system. Research efforts have been dedicated to processing and analyzing the clinical images to extract important data such as the volume, shape, organ movements, to detect abnormalities, and to measure coordination in follow-up research. In the visualization and analysis of medical images, segmentation is a necessary step in the image processing. Due to the imaging modality characteristics for use in the clinic, the resulting image shows different tissue properties for diagnosis. In addition, the anatomical structure is difficult and diverse.

Therefore, segmentation will be a difficult but important issue in biomedical applications. It is a technique for dividing the image into sub-regions or important objects with the same attribution, and generally images and applications depend on some algorithm-based approaches such as histogram analysis, terrain growth, edge detection, and pixel welding which was proposed in the past [13-15]. In general, this method uses local information (i.e., the gray level of the pixels of the neighboring pixels) and/or global information (i.e., distribution of the whole gray image) for the image segment. Some algorithms using neural network approaches have also been investigated in image segmentation problems [16].

Red blood cells, or erythrocytes, are biconcave disks that are important for gas exchange. Their essential product is hemoglobin, carrier of Oxygen and Carbon dioxide in body travel. Red blood cells are produced from pluripotent haemopoietic cells. The stem cells are in the bone marrow and when there is an appropriate stimulation of various hormones or cytokines, it affects the sequence of maturation activities which then produces a fully aimed erythrocyte. Furthermore, there are a large number of cells that allow many erythrocytes to be produced by the individual stem cell activation [17].

Leukemia period refers to a group of cancerous blood cells. In leukemia, white blood cells grow abnormally, dividing and developing in an uncontrolled manner. There are two main types of leukemia: acute leukemia and chronic leukemia. Acute words mean the progress of disease and rapid progress. The acute image of acute leukemia cells based on color images is one of the most difficult tasks in image processing [18].

Clustering is an unsupervised learning method used to group similar cases on the foundation of features.

Clustering is a beneficial approach for the discovery of some expertise from a dataset. It is an exploratory technique for supporting to resolve classification problems. Its use is suitable when little or nothing is recognized about the category shape in a body of facts [19].

K-means is a center-based clustering algorithm which is effectively employed for clustering giant databases and excessive dimensional databases. The goal of a middle based total algorithm is to decrease its objective feature and is nicely applicable for convex form clusters and fails extensively for clusters of arbitrary shapes.

The most straight-forward method to identify clusters used to be the k-means clustering algorithm. The basic thought of k-means clustering is to group statistics factors into k separate clusters (where k is an input parameter), such that we minimize the sum of the distances between each factor and the recommended respective cluster [20]. To do this, the algorithm iterates between

2 main steps: calculate the mean for each cluster and then reassign the data points to clusters using the newly calculated means. This iteration is continued until no new assignments of data points occur.

Clustering is a simple but essential problem in scientific image processing, which is extensively used in clinical applications such as surgical planning, abnormality detection and therapy growth monitoring. One of the benefits of this strategy is that it takes less time. As a weakness, we cannot refer to spatial information. Like the classification approach, these algorithms now do not take into account space information; Therefore, they can be sensitive to sound and intensity in homogeneity [21-27].

Classification is the undertaking of assigning a label from one of the recognized lessons to the unknown check vector. The problem of object classification can be targeted as a hassle to identify the class or category that the new observations belong to the base of a training dataset containing observations which category or class is known. Generally, classification works by way of first plotting the training data into multidimensional space. Then every classifier plot checks out the statistics into the same multidimensional house as the training information and compares the data factors to decide on the correct class for every individual question point.

2. System Overview

The software design and development of a system for red blood monitoring presented in this chapter (Figure 1). After inputting the red blood image, the images are going through some preprocessing methods such as noise removal images, clustering image and edge detection method from the original image. Then the images are extracted with the Principal Component Analysis and Local Binary Pattern. Finally, the Neural Network was employed, to determine the image conditions. MATLAB is used to run the program that build in this research. This can be considered as a pilot preparation to derive the required parameters. In subsequent subsections, these classifier parameters are analysed one by one.

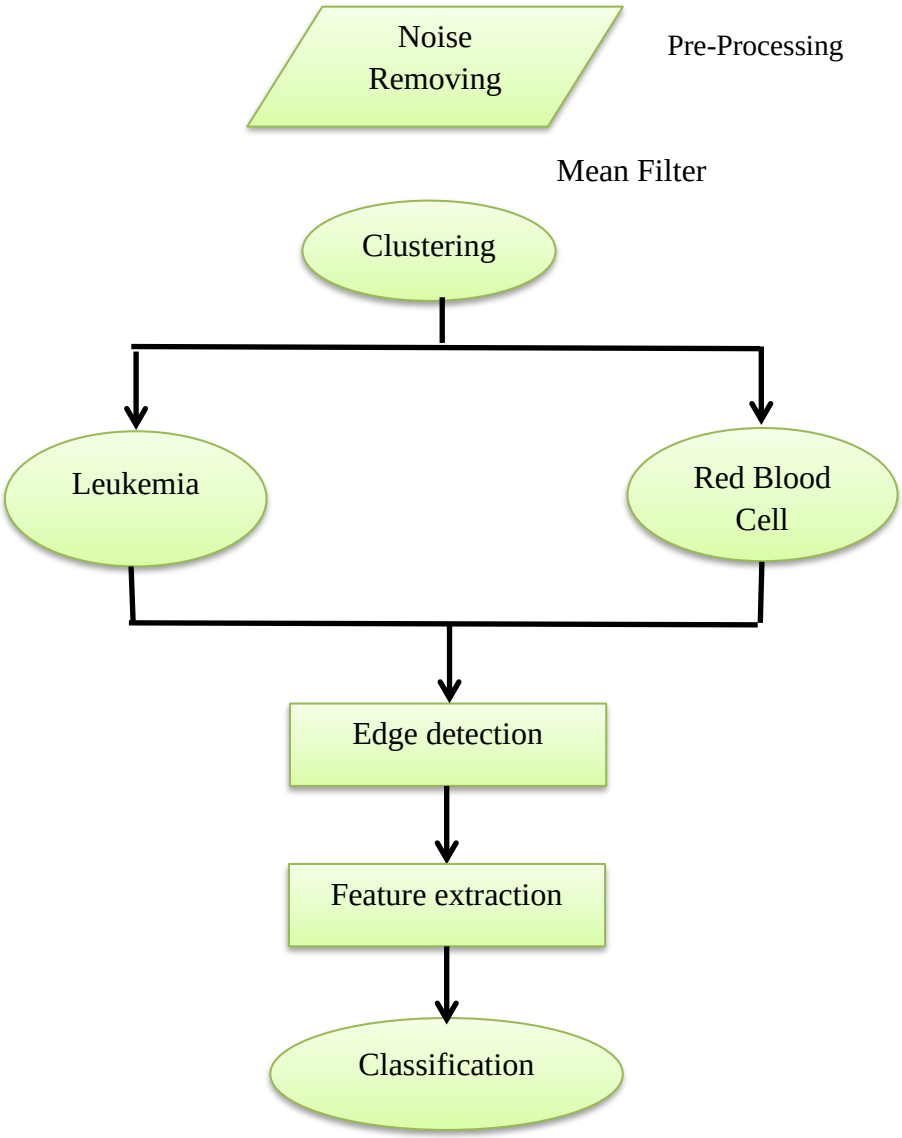


Fig. 1. Flowchart

2.1 Original Image

This dataset in Figure 2 contains images of blood cells images extracted using the negative stain transmission electron microscopy. The data was conducted by Hospital University Sains Malaysia, Kelantan.

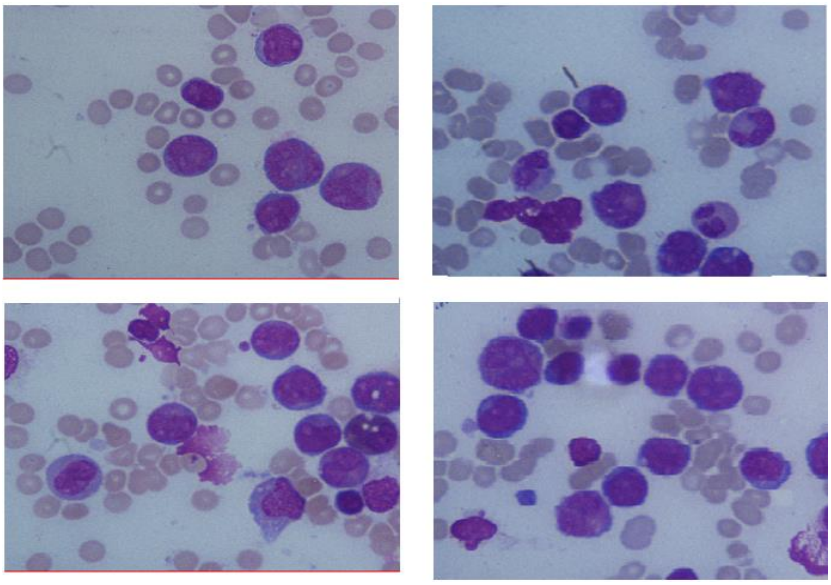


Fig. 2. Show the original image of a blood cell which from Hospital University Sains Malaysia, Kelantan database

2.2 Pre-processing

Pre-processing is the first stage of analyzing the image in the image processing field where it involves various algorithms, techniques and operators which serve to perform an initial processing such as filtering and normalization. More specifically, pre-processing aims to convert the raw data and signals into obvious useful data representation to reduce noises. Pre-processing is a stage where the requirements are normally very straightforward, such as removing the image information regarded as unnecessary for the application as shown in Figure 3.

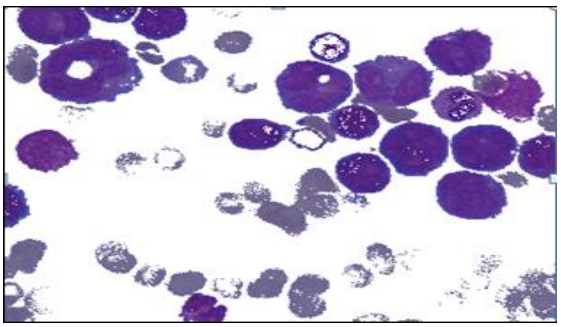


Fig. 3. The Result Images after pre-processing

2.3 Noise Removing

One goal in image restoration is to remove the noise from the image. Nonetheless, one approach is to determine that features that exist on a very small scale in the image are noise and that casting these off whilst preserving large elements may help ease things up [21]. Mean filter; basically, filters come from two types of linear filtering and nonlinear filtering. The linear filter comprises of the mean filters and also known as the averaging filters. In signal processing, the mean filter applies a mask above each pixel in the signal. The average pixel factor that comes below the mask is calculated collectively to form a single pixel that explains why the filter is distinguished in any other case as a filter average. The image is shown in Figure 4.

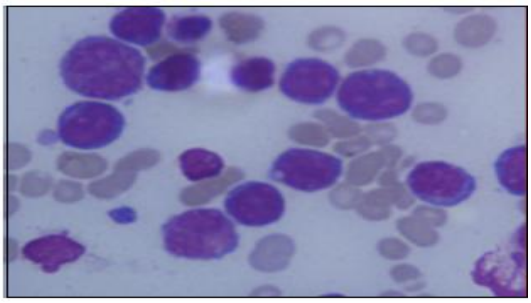
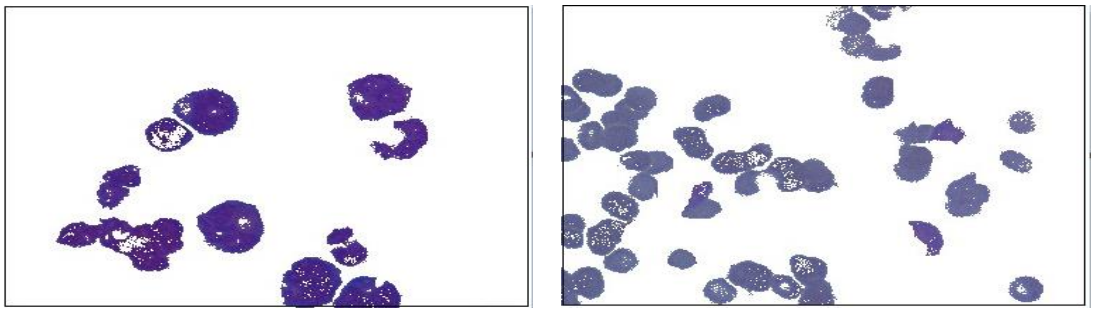


Fig. 4. Mean Filter Result Image

2.4 Clustering

K-means is a center-based clustering algorithm which is effectively used to compile large databases and remote-dimension databases. The purpose of the algorithm based on all is to reduce its objective function and can be applied well to clusters formed in convex form, especially failure in random form clusters. The traditional K-means algorithm was first proposed with the help of Hazlyna *et al.,* [17]. This technique clusters the facts in a fixed range of clusters, and the average of one group is positioned as far as possible from another group. Each information point is connected to the nearest means and belongs to one of the groups. Many variations correspond to subjects that are comparable to literature that generally depends on change or concentration [20]. Clustering result image is shown in Figure 5.

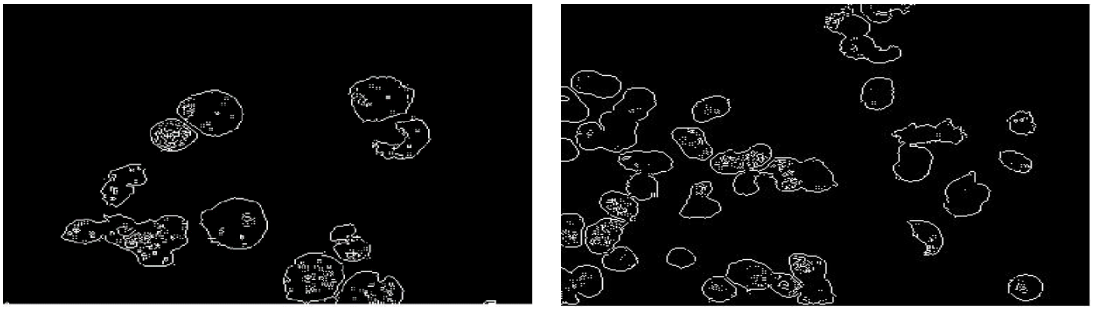


(a) (b)

Fig. 5. Clustering Result Image; (a) Leukemia image, (b) Red blood cells image

2.5 Edge Detection

Edge detection plays a significant role in image processing and image segmentations. Extracting edge features is considered the basic process in edge detection. Commonly, edge features of an image are contained by information and patterns for obtaining specific characteristics or behavior of an image. The term edge detection is defined as a process to identify and locate the sharp discontinuities in digital images such as the photometrical images, physical and geometrical characteristics segments or regions. Figure 6 shows the image of edge result.



(a) (b)

Fig. 6. Edge Result Image; (a) Leukemia image edge, (b) Red blood cells image edge

2.6 Feature Extraction

2.6.1 Principal component analysis (PCA)

PCA is a multivariate technique which extracts important information from data and represents it as a new set of variables called principle components. PCA is a type of factor analysis that is often employed for dimension reduction in a dataset. PCA is often found in the research regarding "data mining, pattern recognition and information retrieval for unsupervised dimensionality reduction". Additionally, in the processing of patient's blood PCA uses smear images to identify Plasmodium parasites for malaria. The results were successful and provide a foundation for further exploratory work [23].

2.6.2 Local binary pattern (LBP)

The LBP properties can be extracted faster in one scan over the raw image and reside in a low-resolution area, even as nonetheless protecting facial information efficiently. The player records the pixels of the image by selecting the three areas for every pixel with the average graphic and consider that binary number as the final result [3].

2.7 Neural Network

Neural network back proliferation is a multilayered, neural network feeding forward. This grid has three layers, such as input layer, output layer, and hidden layer. Hidden neurons and outputs have a bias of 1. Bias is given to perform better [27]. Bias will not change throughout the training. Two hidden layers were found sufficient for this task after a set of experiments. The neural network model is proposed here to classify the features obtained by the optimum value of the IKEA, LBC and DCT. Two conditions are derived from the neural network model as shown in Figure 7. The input to the neural network is the optimal values of the three inputs in the first layer. Two other hidden layers consist of (16 neurons in the second layer, 16 neurons in the third layer) and finally, two outputs are considered in the production layer as shown in Figure 7.

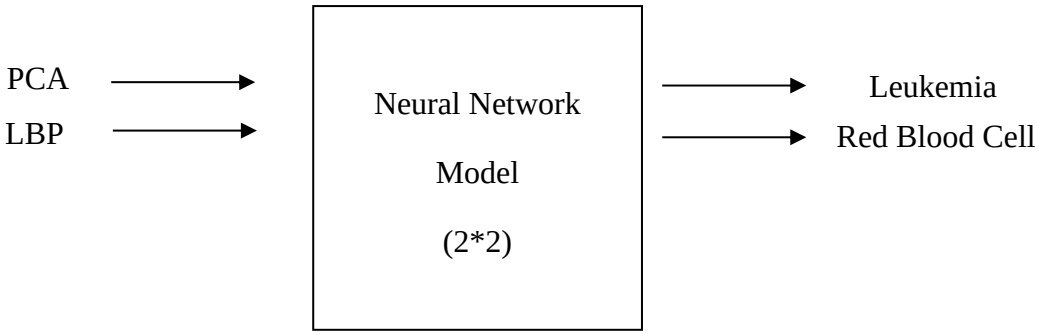


Fig. 7. Neural network model

3. Results and Discussion

The goal of the current study is to partition an image into meaningful areas with an appreciation for a unique application. Medical image segmentation plays an essential function in many medical diagnoses and pathological research. The modern study also targets to improve the pictorial facts for human interpretation and to partition an image into significant areas- thanks to a unique utility which can be beneficial for the cell counting and limit the time of cell counting and disease prognosis effectively.

Therefore, this dissertation investigates the use of various features from the image processing approach with the aim of identifying reliable classifiers. The feasible combos of the features are employed in Figure 8 to Figure 9. This method permits the evaluation of the nature of the features clearly, and it helps to discover the most suitable feature and classifiers.

Figure 8 is the result obtained when combining the extraction method of the Principal Component Analysis with the Neural Network classification method. The highest accuracy is obtained for this 93% accuracy for the first trial. This is despite the fact that the lowest accuracy is 90%. The average accuracy was 91.4%.

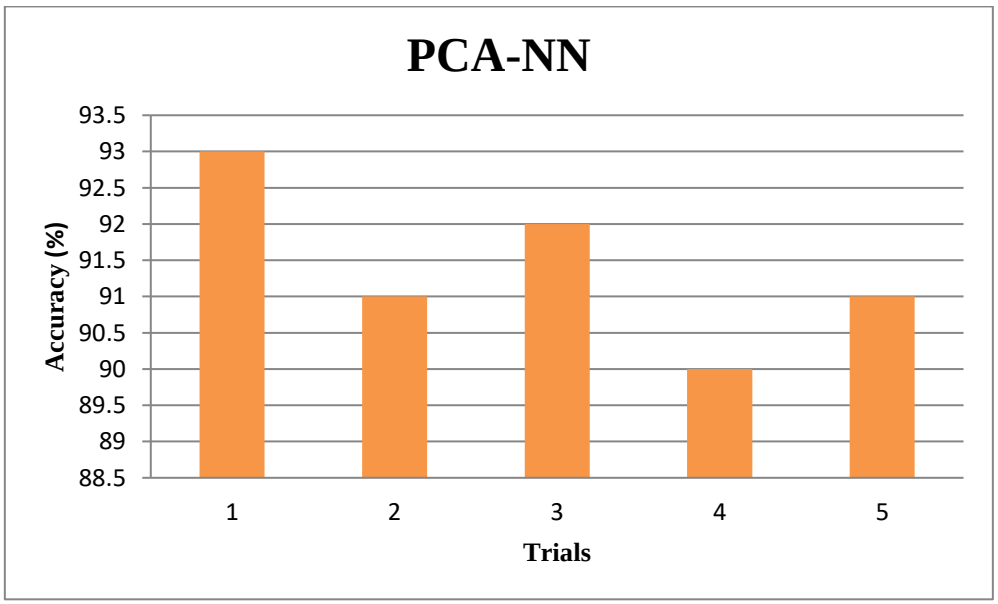


Fig. 8. The accuracy of the results in the PCA with NN was tested at different Trials

Figure 9 is the result obtained when combining the extraction method of the Local Binary Pattern Analysis with the Neural Network classification method. The highest accuracy is obtained for this 94% accuracy for the first trial. This is despite the fact that the lowest accuracy is 90%. The average accuracy was 91.8%.

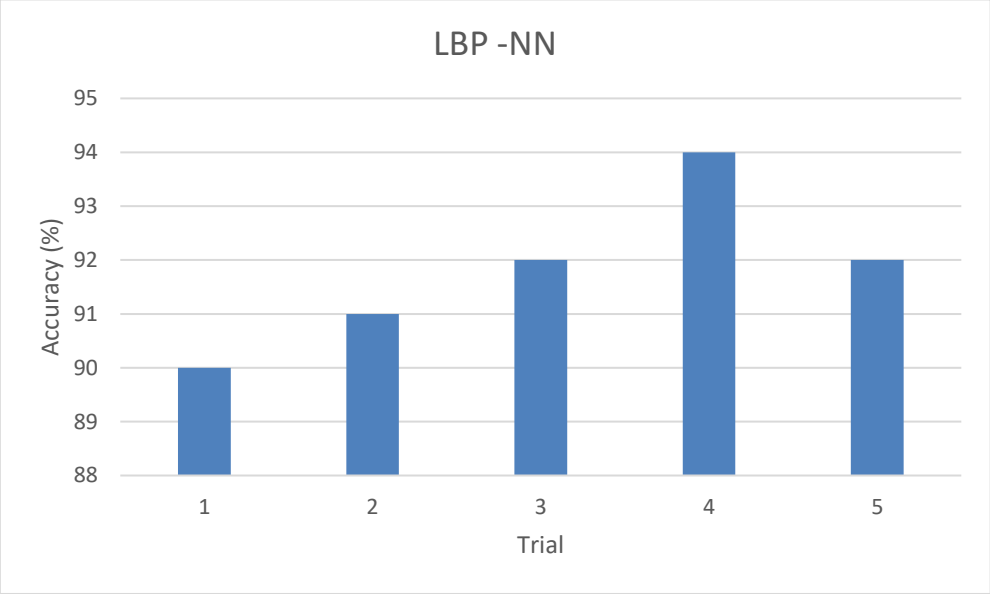


Fig. 9. The accuracy of the results in the LBP with NN was tested at different Trials

4. Conclusions

In conclusion, the aims of Image segmentation are in many medical diagnoses and pathological researches. Image segmentation performs a critical position in many scientific imaging applications, via automating or facilitating the delineation of anatomical constructions and other areas of interest. One of the vital aspects of picture segmentation is to partition an image into significant areas with an appreciation over its unique application. Image processing methods are beneficial for cell counting as they can minimize the time of cell counting and ailment diagnosis effectively. Proper recognition of cell is essential for cell counting. This detects the disorder and diagnosis. Using this method helps diagnose the disorder in less time, no use of chemicals, and no human error and the accuracy of this method gives more than 90% accuracy.

Acknowledgement

The authors would like to acknowledge the financial support in the form of publication incentive grant from Universiti Malaysia Perlis (UNIMAP).

References

[1] Alizadeh, Hosein, Behrouz Minaei-Bidgoli, and Saeed K. Amirgholipour. "A New Method for Improving the Performance of K Nearest Neighbor using Clustering Technique." *Journal of Convergence Information Technology* 4, no. 2 (2009): 84-92. https://doi.org/10.4156/jcit.vol4.issue2.alizadeh

[2] Amelio, Alessia, and Clara Pizzuti. "Skin lesion image segmentation using a color genetic algorithm." In *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation*, pp. 1471-1478. 2013. https://doi.org/10.1145/2464576.2466810

[3] Anusha, Bamini A. M., and T. Kavitha. "Dominant local binary pattern based face feature selection and detection." *National Journal on Advances in Computing and Management* 1, no. 2 (2010).

[4] Bandyopadhyay, Sanghamitra, and Ujjwal Maulik. "An evolutionary technique based on K-means algorithm for optimal clustering in RN." *Information Sciences* 146, no. 1-4 (2002): 221-237. https://doi.org/10.1016/S0020-0255(02)00208-6
[5] Bremner, David, Erik Demaine, Jeff Erickson, John Iacono, Stefan Langerman, Pat Morin, and Godfried Toussaint. "Output-sensitive algorithms for computing nearest-neighbour decision boundaries." *Discrete & Computational Geometry* 33 (2005): 593-604. https://doi.org/10.1007/s00454-004-1152-0
[6] Cabrera, Reymond Joseph A., Criselle Amor P. Legaspi, Erika Jasmine G. Papa, Reden D. Samonte, and Donata D. Acula. "HeMatic: An automated leukemia detector with separation of overlapping blood cells through Image Processing and Genetic Algorithm." In *2017 International Conference on Applied System Innovation (ICASI)*, pp. 985-987. IEEE, 2017. https://doi.org/10.1109/ICASI.2017.7988618
[7] Cheng, Kuo-Sheng, Jzau-Sheng Lin, and Chi-Wu Mao. "The application of competitive Hopfield neural network to medical image segmentation." *IEEE Transactions on Medical Imaging* 15, no. 4 (1996): 560-567. https://doi.org/10.1109/42.511759
[8] Cover, Thomas, and Peter Hart. "Nearest neighbor pattern classification." *IEEE Transactions on Information Theory* 13, no. 1 (1967): 21-27. https://doi.org/10.1109/TIT.1967.1053964
[9] Cui, Wenchao, Jian He, Guoqiang Gong, Ke Lu, and Shuifa Sun. "Local kernel mapping based piecewise constant model for medical image segmentation." In *Eighth International Conference on Digital Image Processing (ICDIP 2016)*, vol. 10033, pp. 320-325. SPIE, 2016. https://doi.org/10.1117/12.2244918
[10] Fausett, Laurene V. *Fundamentals of neural networks: architectures, algorithms and applications*. Pearson Education, 1994.
[11] Luger, George F., and William A. Stubblefield. *Artificial intelligence: structures and strategies for complex problem solving*. Addison-Wesley, 1998.
[12] Gharehchopogh, Farhad Soleimanian, Seyyed Reza Khaze, and Isa Maleki. "A new approach in bloggers classification with hybrid of k-nearest neighbor and artificial neural network algorithms." *Indian Journal of Science and Technology* 8, no. 3 (2015): 237-246. https://doi.org/10.17485/ijst/2015/v8i3/59570
[13] Grau, Vicente, A. U. J. Mewes, M. Alcaniz, Ron Kikinis, and Simon K. Warfield. "Improved watershed transform for medical image segmentation using prior information." *IEEE Transactions on Medical Imaging* 23, no. 4 (2004): 447-458. https://doi.org/10.1109/TMI.2004.824224
[14] Harisinghani, Mukesh G., Jelle Barentsz, Peter F. Hahn, Willem M. Deserno, Shahin Tabatabaei, Christine Hulsbergen van de Kaa, Jean de la Rosette, and Ralph Weissleder. "Noninvasive detection of clinically occult lymph-node metastases in prostate cancer." *New England Journal of Medicine* 348, no. 25 (2003): 2491-2499. https://doi.org/10.1056/NEJMoa022749
[15] Hartigan, John A., and Manchek A. Wong. "Algorithm AS 136: A k-means clustering algorithm." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, no. 1 (1979): 100-108. https://doi.org/10.2307/2346830
[16] Hall, Peter, Byeong U. Park, and Richard J. Samworth. "Choice of neighbor order in nearest-neighbor classification." *The Annals of Statistics* 36, no. 5 (2008): 2135-2152. https://doi.org/10.1214/07-AOS537
[17] Hazlyna, Harun Nor, Mohd Yusoff Mashor, Naematurroziah R. Mokhtar, AN Aimi Salihah, Rosline Hassan, Rafikha Aliana A. Raof, and Muhammad Khusairi Osman. "Comparison of acute leukemia Image segmentation using HSI and RGB color space." In *10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010)*, pp. 749-752. IEEE, 2010.
[18] Jung, Ki-Hyun. "A Data Hiding Method of Binary Images Using Pixel-Value Weighting." *Journal of the Korea Institute of Military Science and Technology* 11, no. 4 (2008): 68-75. https://doi.org/10.1109/ICHIT.2008.293
[19] Chiu, Stephen L. "Fuzzy model identification based on cluster estimation." *Journal of Intelligent & Fuzzy Systems* 2, no. 3 (1994): 267-278. https://doi.org/10.3233/IFS-1994-2306
[20] Karthikeyan, T., and Narayana Poornima. "Microscopic image segmentation using fuzzy C means for leukemia diagnosis." *Leukemia* 4, no. 1 (2017): 3136-3142.
[21] Khashman, Adnan, and Esam Al-Zgoul. "Image segmentation of blood cells in leukemia patients." *Recent Advances in Computer Engineering and Applications* 2, no. 1 (2010): 104-109.
[22] Kim, Jinho, Byung-Soo Kim, and Silvio Savarese. "Comparing image classification methods: K-nearest-neighbor and support-vector-machines." In *Proceedings of the 6th WSEAS international conference on Computer Engineering and Applications, and Proceedings of the 2012 American conference on Applied Mathematics*, pp. 133-138. 2012.
[23] Klinken, S. Peter. "Red blood cells." *The International Journal of Biochemistry & Cell Biology* 34, no. 12 (2002): 1513-1518. https://doi.org/10.1016/S1357-2725(02)00087-0
[24] Liao, Qingmin, and Yingying Deng. "An accurate segmentation method for white blood cell images." In *Proceedings IEEE International Symposium on Biomedical Imaging*, pp. 245-248. IEEE, 2002.
[25] McQueen, R. G., S. P. Marsh, and J. N. Fritz. "Hugoniot equation of state of twelve rocks." *Journal of Geophysical Research* 72, no. 20 (1967): 4999-5036. https://doi.org/10.1029/JZ072i020p04999

[26] Malik, Hina, Roopali Randiwe, Jyotsna Patankar, and Priya Bhure. "Disease Diagnosis Using RBCs & WBCs Cell Structure by Image Processing." *International Journal of Scientific Research in Science and Technology* 3, no. 2 (2017): 120-123.
[27] Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." In *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565-571. IEEE, 2016. https://doi.org/10.1109/3DV.2016.79