IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Review on Implementation of ACO technique for leukaemia detection

Jeenu Garg

Master of Technology, Department of IT, CGC, Landran, India

Abstract: Automatic advance of leukaemia detection is planned. A physical technique of LEUKAEMIA DETECTION, specialist checks microscopic images. Leukaemia detection is generates in the bone marrow. Lengthy and time taking process which depends on human's ability and not having accuracy. Each bone contains a thin material inside is known bone marrow. The components of erythrocytes and leucocytes and platelets. Basically Leukaemia is detected only by analysing the white blood cells. Focused only on WBC, Leukaemia Detection system analyses the microscopic image and conquer these Drawbacks. It extracts the necessary parts of images and direct applies some techniques. K-mean clustering is used only WBC (WHITE BLOOD CELL) detection. In this paper we describe a system for medical data processing that mainly uses ACO (Ant Colony Optimization) technique that provides the consequences for leukaemia detection and classification through multi-layer neural network (BPNN).

Keywords: Data Mining, ACO (Ant Colony Optimization, BPNN (Back Propagation Neural Network), K-mean clustering

1. INTRODUCTION

A. Data Mining

Data mining also known as knowledge-discovery in databases (KDD) is process of extracting potentially useful information from raw data. A software engine can scan large amounts of data and automatically report interesting patterns without requiring human intervention. Other knowledge discovery technologies are Statistical Analysis, OLAP, Data Visualization, and Ad hoc queries. Unlike these technologies, data mining does not require a human to ask specific questions. Medical imaging has become one of the most significant conception and explanation methods in ecology and medicine over the previous decade. This time has perceived incredible expansion of new, prevailing apparatuses for detecting, packing, conducting, analysing, and exhibiting medical images. This has led to enormous growth in the application of digital image dispensation techniques for cracking medical difficulties.

B. LEUKAEMIA

Leukaemia is the cancer of the blood. It starts in the bone marrow, it is the area where blood cells are made. When you have leukaemia, the bone marrow starts to make a lot of abnormal white blood cubicles, called leukaemia cells. They don't do the exertion of normal white blood cells. They grow faster than normal cells, and they don't break increasing when they should. Over time, leukaemia cells can crowd out the normal blood cells. This cans chief to serious difficulties such as anaemia, bleeding, and infections. Leukaemia cells can also spread to the lymph nodes or other organs and origin bulge or pain.

There are numerous different types of leukaemia. In general, leukaemia is collected by how fast it gets poorer and what kind of white blood cell it affects.

The microscopic images of the blood cells are experiential to find out numerous diseases. Variations in the blood condition show the development of diseases in an individual. Leukaemia can central to demise if it is left unprocessed. Based on some statistics it is found that the leukaemia is the fifth cause of death in men and sixth cause of death in women. Leukaemia originates in the bone marrow. The cells in the bone marrow start changing and they get infected and become leukaemia or infected cells. These leukaemia cells are having strange properties than the normal cells.

Their development is irregular and existence time is more than the normal cells. They interrupt normal cells to carry out their work. After a certain amount of time normal cells die while leukaemia cells don't. The old leukaemia cells last for a long time and new leukaemia cells produce in an abnormal way. The rate at which the leukaemia cells progress is different according to the type of leukaemia.

C. WHITE BLOOD CELL

White blood cells are bigger in size than the red blood cells. The concentration and composition of the white blood cells provide some important information which helps us to find out many diseases .

D. FEATURE EXTRACTION USING PRINCIPLE COMPONENT ANALYSIS

It is a way of classifying patterns in data, and expressing the data in such a way as to highlight their resemblances and changes. First of all, we had to create the data set. The aim is to choose a good number of pictures and a good purpose of these in order to have the best recognition with the smallest database. Then, the next step is to deduct the

mean from each of the data dimensions. The mean subtracted is simply the average across each dimension. The step three is to compute the covariance matrix of the database. We could not calculate the covariance matrix of the first medium, since it was too huge. So we had to find a way to find out the principal eigenvectors without calculating the big covariance matrix. The technique consists in choosing a new covariance matrix. Our covariance medium for A was called D and D is defined by:

$$C = D^* D'$$

The Eigenvectors and the Eigenvalues of C are the principal mechanisms of our data set.

E. Ant Colony Optimization Algorithm for Optimization Technique

- Swarm intelligence studies the cooperative performance of unsophisticated agents that interact locally through their situation.
- It is motivated by social insects, such as ants and termites, or other animal societies, such as fish schools and bird flocks.
- Although each separate has only limited capabilities, the complete swarm exhibitions complex overall behaviour. Therefore, the intelligent behaviour can be seen as an emergent distinguishing of the swarm.
- When focusing on ant colonies, it can be perceived that ants communicate only in an indirect manner through their environment by dropping a substance called pheromone.
- Paths with higher pheromone levels will more likely be preferred and thus reinforced, while the pheromone intensity of pathways that are not chosen is decreased by evaporation.
- This form of indirect statement is known as stigmergy, and offers the ant colony shortest-path finding capabilities.
- ACO employs reproduction ants that collaborate to find good solutions for discrete optimization difficulties. These software agents mimic the foraging behaviour of their biological complements in finding the shortest-path to the food source.

F. Back Propagation Neural Network

The Back Propagation neural network is artificial neural network based on error back propagation algorithm. The Back Propagation Neural Network model consists of an input layer, some hidden layers and an output layer. Each connection connecting neurons has a distinctive weighting value. In training the network, the nodes in the BP neural network obtain input information from exterior sources, and then go by to hidden layer which is an interior information processing layer and is answerable [6] for the information conversion, and then the nodes in the output layer supply the required output material. After that, the anti-propagation of error is transported by distinct the actual output with wanted output. Each weight is reviewed and back propagated layer by layer from output layer to hidden layer and input layer. This process will be sustained until the output error of network is reduced to an acceptable level or the predetermined time of learning is realized. The processing consequences of information are exported by output layers to the outside.

BP neural network consists of many neurons that are arranged in a form of three layers: input, hidden and output. The neurons are linked by weights W y In training the network with a given architecture, the back propagation approach, finds a single best set weight values by minimization of suitable error function. In a multi-layer feed forward neural network, the processing elements are arranged in layers and only the rudiments in adjacent layers are connected. It has a minimum of three layers of elements (i.e., input layer, the central or hidden layer, and the output layer). The name “back propagation” (BP) derives from the fact that computations are passed feed forward from the input layer to the output layer, following which calculated errors are propagated back in other direction to change the weights to obtain a better performance.

BP algorithm is an extension of the smallest mean square algorithm that can be used to train multi-layer networks. The three-layered free forward neural network is displayed in Fig. 3 which is comprised by input layer, hidden layer and output layer.

2. LITERATURE SURVEY

[1] Madhloom performed some image arithmetic operations and some threshold operations to find out the white blood cell nuclei. The challenging task in developing this system was the selection of an appropriate threshold for the segmentation. The threshold used in the system developed is not providing efficient result for the segmentation of white blood cells nuclei.

[2] Kovalev developed a system to classify five types of leucocytes from the blood image. In this process, he detected nuclei first and then applied a region growing techniques to find out the entire membrane. The results achieved were quite good.

[3] Scotti used some threshold operations, low-pass filter for the removal of background and clustering for white blood cells segmentation. Scotti tried to achieve a good segmentation results for the images which are taken under different lightening conditions.

[4] Piuri performed white blood cell segmentation. He used edge detection technique for the each leucocyte. He used a neural network for the classification. He trained neural network by morphological features to recognize lymphoblast.

[5] Halim proposed an automated system which counts number of blasts in the microscopic blood image. He applied some threshold operation on S component of HSV colour space to detect white blood cells from the image. The results achieved were quite amazing but the problem

in the system is that no method is mentioned for selecting the optimum threshold for the better segmentation. There is no features extracted and no classifier has been used.

[6] Mohapatra applied clustering for white blood cells segmentation and extracted the features like shape, colour, texture, fractal, Fourier descriptors and contour. The system 15 was trained to recognize Leukaemia. The results achieved were quite good but the proprietary data set was used to achieve it therefore, it cannot be compared with other methods.

[7] Donida Labati proposed the data set including the blood samples of the normal patients and leukaemia patients which found very helpful for our proposed system testing. This data set is publicly available for the research purpose.

[8] Liao and Deng developed a system to segment white blood cells from the image. They applied some threshold techniques and then after they applied contour identification. They have assumed that all cells are circular to make the algorithm works efficiently. Due to the assumption of the circular shaped cells, this system is not suitable for the lymphoblast cells which are irregular in shape.

[9] Angulo et al. proposed a system in which he proposed "two-stage blood image segmentation algorithm". They are using dualistic sifting and some automatic threshold techniques. This system performs well for extracting the centre, cytoplasm and nucleolus from the lymphocyte descriptions. The two stage segmentation process has been applied here and because of this the computation time is higher. The images are taken under different lightening condition which makes difficult to choose the optimum threshold for segmentation.

[10] Sinha et al. proposed a scheme which segments the leucocytes automatically. He used EM algorithm and Gaussian mixture modelling. In this method, parameter tuning is not required. This is unsupervised approach. This scheme is not work for all stains.

[11] Umpon invented a technique for the white blood cell nucleus segmentation. He applied fuzzy clustering. This technique works well for the nucleus segmentation but the 16 cytoplasm extraction is also as important as the nucleus segmentation which is not taken care in this technique.

[12] Dorini et al. proposed a scheme for the nucleus extraction. The watershed transform has been used in this scheme which is based on the image forest transform. He has extracted cytoplasm by using the size distribution information. This scheme is not working well if the cytoplasm isn't round.

[13] Wenlong Tang et.al, in this paper described as, a novel compressive sensing based method for the subtyping of leukaemia. The CS method is an emerging method in statistics and mathematical signal analysis, which enables the reconstruction of signals from a small set of incoherent projections. They developed a CS based detector to classify ALL and AML, based on ours selected genes out of7129 samples. This work demonstrates that the CS method can be effectively used to detect subtypes of leukaemia subjects, implying improved accuracy of diagnosing leukaemia patients.

[14] Subrajeet Mohapatra et.al, in this Paper planned as the general nature of the signs and symptoms of ALL often leads to wrong diagnosis. Diagnostic confusion is also postured due to simulated of similar signs by other disorders. Careful microscopic examination of stained blood mark or bone marrow enunciates is the only way to effective diagnosis of leukaemia. Techniques such as fluorescence in situ hybridization, immune phenol-typing, cytogenetic analysis and cytochemistry are also employed for specific leukaemia finding. The need for mechanization of leukaemia detection arises since the above specific tests are time consuming and costly.

[15] Yong Jiang et.al, in this purposed as learning was to study the extract of Rumex root which had an inhibitory action on the cell propagation of human leukaemia cell line THP-1. The combination of percolation and abstraction was used to separate and quotation the main chemical compositions of Rumex; MIT was used to assay the curve of inhibition ratio.

3. FUTURE SCOPE

There are so many ways to make this system better in future. We can improve the segmentation scheme which can segment the overlapped cells also. There were found the use of optimization techniques in some systems. We can also use Back Propagation Neural network classifiers to improve the accuracy of the classification. Doing so will increase the cost but accuracy will also be improved. We can use parallel algorithm for the execution so that the execution time can be decreased

4. CONCLUSION

From the study of leukocyte segmentation, it is noticed that large number of methods are only working on the extraction of nucleus but there are very few methods available which are extracting the cytoplasm and even with less accurately. The main reason behind the less accuracy in the cytoplasm extraction is that most of the researchers are using the grey level colour for the extraction of cytoplasm which is not easily separable from the other colours. It is noticed in the literatures that different approaches are used for the white blood cells detection. Some have used KNN approach, threshold techniques, EM algorithm, Fuzzy rules, watershed transform, GVF model, trained neural network, Fuzzy c-mean clustering, computer morph metric system and many more. From the literature studied, it has been observed that there are many ways we can make a better system for the

identification of leukaemia from the microscopic blood image. None of the researchers has used the K-mean clustering for the segmentation of the white blood cells from the microscopic blood image.

In this paper, K-mean clustering approach has been used on the clean microscopic blood image followed by image cleaning and the extraction of the nucleus and cytoplasm with a good accuracy. In our work survey helps to classify the data using BPNN (Back Propagation Neural Network) to find the accuracy.

ACKNOWLEDGEMENT

Firstly thank you to my strength of my life, my parents for their encouragement and support. I would like to place on record my deep sense of gratitude to my supporters my teachers for his generous guidance, help and useful suggestions.

REFERENCES

[1] Tang, Wenlong, Hongbao Cao, and Yu-Ping Wang. "Subtyping of Leukemia with Gene Expression Analysis Using Compressive Sensing Method."Healthcare Informatics, Imaging and Systems Biology (HISB), 2011 First IEEE International Conference on. IEEE, 2011.

[2] Mohapatra, Subrajeet, et al. "Fuzzy based blood image segmentation for automated leukemia detection." Devices and Communications (ICDeCom), 2011 International Conference on. IEEE, 2011.

[3] Jiang, Yong, et al. "Extractives of Rumex restrain the proliferation of leukemia cell THP-1." IT in Medicine and Education (ITME), 2011 International Symposium on. Vol. 2. IEEE, 2011.

[4] Halim, NH Abd, et al. "Nucleus segmentation technique for acute leukaemia. "Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on. IEEE, 2011.

[5] Madhloom, Hayan T., Sameem Abdul Kareem, and Hany Ariffin. "A Robust Feature Extraction and Selection Method for the Recognition of Lymphocytes versus Acute Lymphoblastic Leukemia." Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference on. IEEE, 2012.

[6] Ning, Yuan, et al. "Comparison of different BP neural network models for short-term load forecasting." Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on. Vol. 3. IEEE, 2010.

[7] Madhloom HT, Kareem SA, Ariffin H, Zaidan AA, Alanazi HO, Zaidan BB. "Automated white blood cell nucleus localization and segmentation using image arithmetic and automated threshold." J Apple Sci 2010;10(11):959–66.

[8] Kovalev VA, Gregory AY, Ahn "H. Robust recognition of white blood cell images".In: Kavanaugh ME, Werner B, editors. Proceedings of the 13th international conference on pattern recognition, August 25–29. Vienna, Austria: IEEE Publisher;1996. p. 371–5.

[9] Scotti F. Robust "segmentation and measurements techniques of white cells in blood microscope images. "In: Daponte P, Linnenbrink T, editors. Proceedings of the IEEE instrumentation and measurement technology conference, 24–27April. Sorrento, Italy: IEEE Publisher; 2006. p. 43–8.

[10] Piuri V, Scotti F. "Morphological classification of blood leucocytes by microscope images". In: Proceedings of the IEEE international conference on computational intelligence for measurement systems and applications, 14–16 July. Boston, MA, USA: IEEE Publisher; 2004. p. 103–8.

[11] Halim NHA, Mashor MY, Hassan R. "Automatic blasts counting for acute leukemia based on blood samples". Int J Res Rev Comput Sci 2011;2(August(4)). 35

[12] Mohapatra S, Patra D, Satpathy S." An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images." J Neural Comput Appl 2014; 24:1887–904.

[13] Donida Labati R, Piuri V, Scotti F." ALL-IDB: the acute lymphoblastic leukemia image Database for image processing". In: Macq Benot, Schelkens Peter, editors. Proceedings of the 18th IEEE ICIP international conference on image processing, September 11–14. Brussels, Belgium: IEEE Publisher; 2011. p. 2045–8.

[14] Q. Liao and Y. Deng. "An accurate segmentation method for white blood cell images". In Proceedings of the IEEE International Symposium on Biomedical Imaging, pages 245 – 248, 2002.

[15] J. Angulo and G. Flandrin. "Microscopic image analysis using mathematical morphology: Application to haematological cytology". In A. Mendez-Vilas, editor, Science, Technology and Education of Microscopy: An overview, pages 304–312. FORMATEX, Badajoz, Spain, 2003.

[16] N. Sinha and A.G. Ramakrishna. "Blood cell segmentation using algorithm." In Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, 2002.

[17] N. T. Umpon. "Patch based white blood cell nucleus segmentation using fuzzy clustering". ECTI Transaction Electrical Electronics Communications, 3(1):5–10, 2005.

[18] L.B. Dorini, R. Minette, and N.J. Leite. "White blood cell segmentation using morphological operators and scale-space analysis." In Proceedings of the Brazilian Symposium on Computer Graphics and Image Processing, pages 294–304, October 2007.

[19] Huang, De-Shuang, and Chun-Hou Zheng. "Independent component analysis-based penalized discriminant method for tumour classification using gene expression data." Bioinformatics 22.15 (2006): 1855-1862.

[20] Liu, Zhenqiu, Dechang Chen, and Halima Bensmail. "Gene expression data classification with kernel principal component analysis." BioMed Research International 2005.2 (2005): 155-159.

[21] Martens, David, et al. "Classification with ant colony optimization."Evolutionary Computation, IEEE Transactions on 11.5 (2007): 651-665.

[22] Supardi, N. Z., et al. "Classification of blasts in acute leukemia blood samples using k-nearest neighbour." Signal Processing and its Applications (CSPA), 2015 IEEE 8th International Colloquium on. IEEE, 2015.

[23]Badea, Liviu. "Unsupervised analysis of leukemia and normal haematopoiesis by joint clustering of gene expression data." Bioinformatics & Bioengineering (BIBE), 2012 IEEE 12th International Conference on. IEEE, 2012.

[24] Tong, Dong Ling, and Graham R. Ball. "Exploration of leukemia gene regulatory networks using a systems biology approach." Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on. IEEE, 2014.

[25] Raje, Chaitali, and Jyoti Rangole. "Detection of Leukemia in microscopic images using image processing." Communications and Signal Processing (ICCSP), 2014 International Conference on. IEEE, 2014.

BIOGRAPHY

Jeenu Garg ,having B.Tech Degree in Computer Science Engineering from Punjab Technical University, Jalandhar, India, in 2014, Persuing M.tech (Final Year) in Infirmation Technology from Chandigarh Engineering College, Landran, India. Her research area include Data Mining in medical field finding implementation of ACO technique in leukaemia cancer detection.

DOI 10.17148/IJARCCE.2016.54211