International Journal of Engineering Sciences & Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

RESEARCHERID
THOMSON REUTERS

[Jain* *et al.*, 8(1): January, 2019]
IC™ Value: 3.00

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

CLASSIFICATION OF BLOOD LEUKEMIA USING NEURAL NETWORKS AND IMAGE PRE-PROCESSING

Saloni Jain*[1] & Divya Gupta [2]

* Department of Computer Science and Engineering, VITM, Indore, Indore

DOI: 10.5281/zenodo.2549887

ABSTRACT

In this paper, we have presented an efficient mechanism for the classification of blood leukemia using microscopic images employing a Probabilistic Neural Network (PNN). Probabilistic Neural Network is based on Bayes's theorem of Conditional Probability and is a famed paradigm for data classification for systems employing artificial intelligence. Pre-processing has been achieved using gray scaling and thresholding. Discrete Wavelet Transform (DWT) has been used as a tool to remove the abrupt variations in the calculated feature values. Principal component analysis (PCA) has been used to find particular trends in the computed feature data and minimize the redundancy. We have shown that the proposed technique achieves 98.33% percent classification accuracy which can be attributed to the highly rigorous pre-processing and feature extraction mechanisms which culminates to training a Probabilistic Neural Network which is used for the final classification of the data.

KEYWORDS: Artificial Intelligence (AI), Probabilistic Neural Network (PNN), Discrete Wavelet Transform (DWT), Principal Component Analysis (PCA), Accuracy, Leukemia

1. INTRODUCTION

The classification of blood cancer is a challenging task since the cancer positive and cancer negative cases show a high level of similarity. The need for accurate classification of heart diseases is necessary for physicians to classify blood cases with accuracy. A new paradigm of classification of blood cancer classification using artificial intelligence based approaches has been under an active area of research.

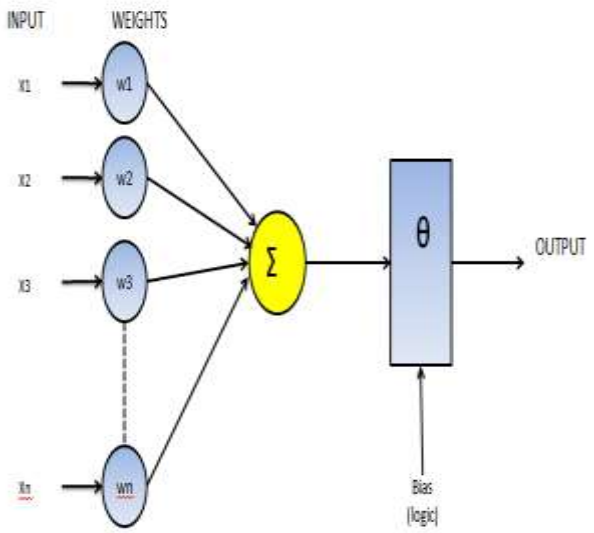


Fig.1 Mathematical Formulation of Neural Network

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

To implement the artificial intelligence based approaches, there is a need to design artificial neural networks. The figure above shows the mathematical model for artificial neural networks. [1] The figure assumes that the neural network behaves similar to the human brain with the following two attributes:
a) Parallel data processing structure
b) Learning based on weights and continuous adapting capability
The above two attributes result in the design of the mathematical equation of the output of the neural network which is given by:

$$\mathbf{Y = \sum_{i=1}^{n} XiWi + \Theta} \qquad \textbf{(i)}$$

Here X represents the signal
W represents the weight
Θ represents the bias.

The internal diagram for the perception of the neural network is given below which basically comprises of the following three layers viz.
1) Input layer
2) Hidden layer
3) Output layer

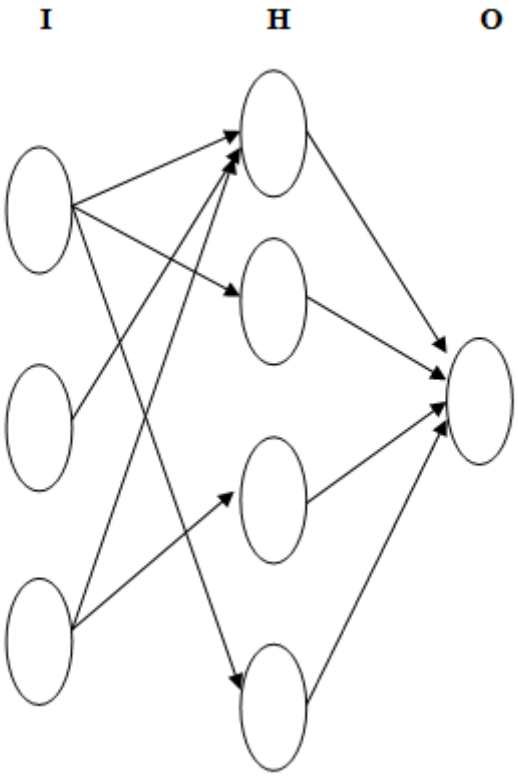


Fig.2 Model of PNN with Input Feature Values

The figure above describes the internal structure of the neural architecture. It is necessary to note that the number of neurons in the hidden layer is generally based on a trial and error basis and there is no fixed rule to ascertain the number of hidden layer neurons. The subsequent section deals with the methodology of classification of actual microscopic images using the neural network architecture.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

2. PREVIOUS WORK

There are different approaches which have proposed differently and stark techniques in terms of classification of blood cancer images. The summary of such techniques have been presented below.

Topic: Classification of Leukemia Blood Samples Using Neural Networks, Springer 2010[1]
Authors: Malek Adjouadi et.al.
The authors proposed a technique based on the classification of blood leukemia samples that often tend to have overlapping image attributes. The authors resorted to the use of artificial neural networks that could classify irregular or even overlapping patterns for the blood leukemia samples. The algorithm produced very high sensitivity results that improved up to 96.67%.

Topic: Classification of blasts in acute leukemia blood samples using K-Nearest Neighbor, IEEE 2012[2]
Authors: N.Z Supardi et.al.
In this paper, the authors have presented an approach for acute myelogenous leukemia (AML) and acute lymphocytic leukemia (ALL) by the use of the *k*-NN based approach. The major features extracted were 12 fundamental features which are colour and texture based features. The approach used could attain a classication accuracy which could reach up to 80%.

Topic: Detection of Leukemia in Microscopic Images Using Image Processing, IEEE 2014[3]
Authors: Chaitali Raje et.al.
In this paper, the authors used the tolls of MATLAB and Lab View for the classification of blood cancer images. The approach used was the segmentation and further classification of images using the pre-defined tools. The classifier used was the Euclidean distance classifier. The classifier was based on the distance comparison of classification vectors form a standard data set.

Topic: Computer Aided Diagnostic System for Detection of Leukemia using Microscopic Images, Elsevier 2015[4]
Authors: Jyoti Rawat et.al
This paper proposes a novel technique to differentiate ALL (acute lymphoblastic leukemia) lymphoblast cells from healthy lymphocytes using the Bray Scale Covariance Matrix and the Support Vector Machine (SVM). Te method used the determination of the appropriate hyper-plane that could classify the cases with maximum accuracy. The classification accuracy of combined texture-shape feature is 89.8%.

Topic: Image segmentation and classification of white blood cells with the extreme learning machine and the fast relevance vector machine, Elsevier 2015[5]
Author: S. Ravikumar et.al
In this paper, it was discussed that white blood cells (WBCs) or leukocytes are an important part of the body and that classification of blood cancer based on machine learning techniques was a critical factor. The approach used was extreme machine learning. The focus was on deep learning techniques which are nothing but the use of multiple hidden layers for the design of the neural network. It can be shown using the proposed technique that artificial neural network owing to its fitting capability of complex data sets can prove to be a potent tool for the classification of blood sample. There are various stage of image pre-processing needed for the classification of the blood cancer cases.

Topic: Large Scale Image Feature Extraction from Medical Image Analysis, IJAERS 2016[6]
Authors: Desai Devanshi Manojbhai et.al.
In this paper, research in big data is focused on deriving knowledge from seemingly complex data sets. The approach was the use of a probabilistic approach that is used to compute the probability of the data set to belong to the positive category or the negative category. The probabilistic approach is often needed since the overlapping attributes of the images make it difficult to classify with high accuracy.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Topic: Efficacy of Azacitidine in autoimmune and inflammatory disorders associated with myelodysplastic syndromes and chronic myelomonocytic leukemia, Elsevier 2016[7]
Authors: Jean-Baptiste Fraisona et. al
This study describes efficacy of the discrete wavelet transform for the classification of image based blood cancer classification. Since Fourier based techniques were competent only for the classification and analysis of slowly changing signals, hence they are often not appropriate for the classification of suddenly changing or abrupt signals such as images. The approach here was to use the discrete wavelet transform to remove noise and disturbance effects so as to extract features with high accuracy.

Topic: Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review, IEEE 2016[8]
Authors: Fuyong Xing and Lin Yang
This paper explains the various techniques for the classification of images based on image processing and neural networks. The approach also discusses dimensional reduction techniques. This is used to remove the redundancy in the training data samples to be fed to the neural network. The approach is to minimize redundancy and make the training of the neural based system as effective as possible. The performance metrics are the accuracy and the sensitivity of the system.

3. MATERIALS AND METHODS

The proposed system can be understood using the following approach that uses the probabilistic neural network along with different image pre-processing techniques. The images are to be extracted and processed based on the following stages:
a) RGB to Gray Scale conversion: This process is necessary for the ease of the computation of the features and further training of the neural network. It maps a color image to a grayscale image. Mathematically:

$$\boldsymbol{I(R,G,B) \rightarrow I'} \qquad \textbf{(ii)}$$

Here,
I(R,G,B) is the color image
I' is the gray scale image
A transformation maps one image to the other.
b) The Discrete Wavelet Transform: In this process, the noise effects can be removed from the image. This results in better training of the network subsequently. . Since Fourier based techniques were competent only for the classification and analysis of slowly changing signals, hence they are often not appropriate for the classification of suddenly changing or abrupt signals such as images. The approach here was to use the discrete wavelet transform to remove noise and disturbance effects so as to extract features with high accuracy.
c) Use of Dimensional Reduction using Principal Component Analysis (PCA): The principal component analysis is a technique that is used for the dimensional reduction of redundant data. The approach is to minimize redundancy and make the training of the neural based system as effective as possible. The performance metrics are the accuracy and the sensitivity of the system. It is useful for making overlapping data into non-overlapping data.
d) Use of Probabilistic Neural Network: The probabilistic neural network is by far the most effective technique where there is no clearly visible boundary to discriminate the data set. The technique proposed was that of a probabilistic approach that is used to compute the probability of the data set to belong to the positive category or the negative category. The probabilistic approach is often needed since the overlapping attributes of the images make it difficult to classify with high accuracy. Mathematically,

$$\boldsymbol{P\left(\frac{A}{B}\right) = \frac{P\left(B/A\right)P(A)}{P(B)}} \qquad \textbf{(iii)}$$

Here,
P(A/B) is the probability of A given B is true
P(B/A) is the probability of B given A is true
P(A) is the probability of A
P(B) is the probability of B

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

We assume here that the data sample used for training purposes is B and hence are considered true. The probability of a test samples needs to be checked for probability of a particular case. The category with the maximum probability is adjudged as the category of the data sample. The fundamental attribute to be computed once the classification is over is the accuracy of the proposed system. The classification accuracy can be crudely understood as:

$$Accuracy = \frac{No.of\ Correctly\ classified\ cased}{Total\ number\ of\ cases\ tested} \quad \textbf{(iv)}$$

The accuracy is however computed based on the definitions of true positive, true negative, false positive and false negative. The concept has been discussed in detail in a subsequent section. The overall flow of the proposed approach is given in the flowchart below.

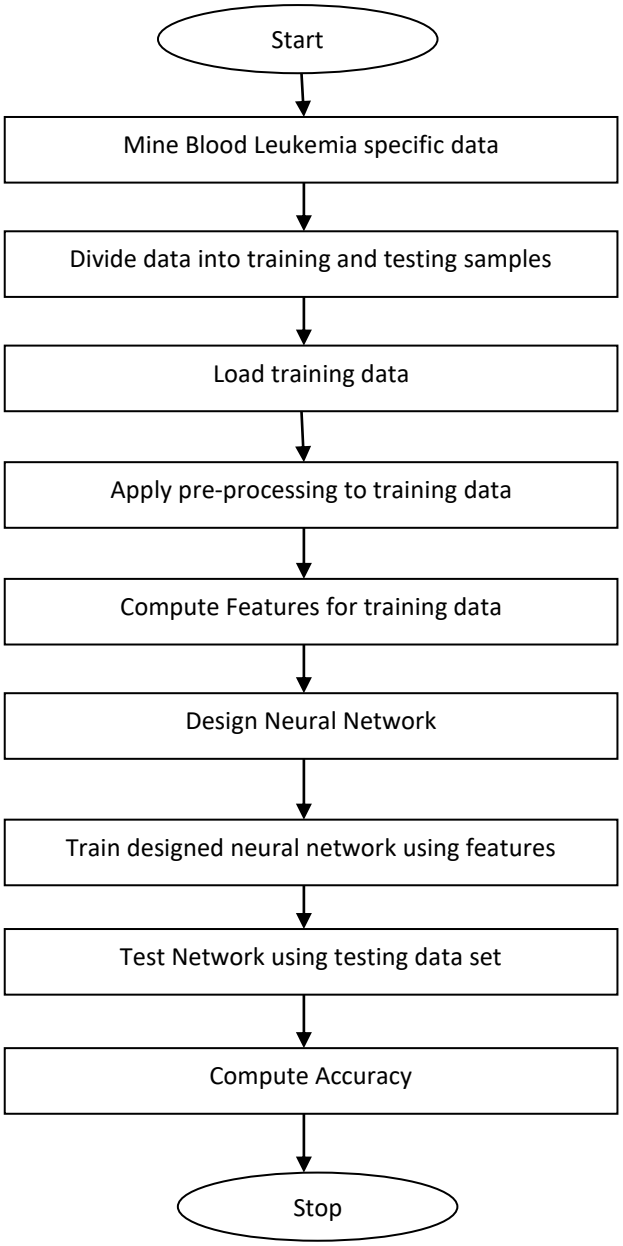


Fig.3 Proposed Ssytem's Flow Chart

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

4. RESULTS AND DISCUSSION

Evaluation Parameters:

The various parameters for the classification are:

1. **True Positive (TP):** It is the case when a sample belongs to category and the test also predicts its belongingness.
2. **True Negative (TN):** It is the case when a sample does not belong to category and the test also predicts its non-belongingness.
3. **False Positive (FP):** It is the case when a sample does not belong to category and the test predicts its belongingness.
4. **False Negative (FN):** It is the case when a sample belongs to category and the test predicts its non-belongingness.

Accuracy (Ac): It is mathematically defined as:

$$\frac{TP+TN}{TP+TN+FP+FN} \qquad \text{(v)}$$

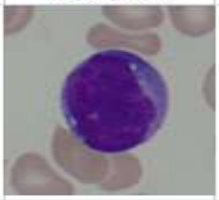
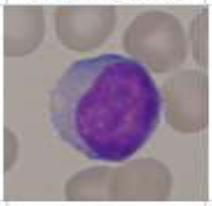
HEALTHY	LEUKEMIA
	
0.566477273	0.403409091
0.134990779	0.079914238
0.805264075	0.689903151
0.938386769	0.910672348
0.004660435	0.009083011
0.106527261	0.106241438
2.208658726	3.764806736
0.106600358	0.106600358
0.011328815	0.011261729
0.895968612	0.943774119
20.05759459	7.15389322
2.159108057	0.788548077

Fig.4: Tabulation of feature values for Normal and Leukemia Blood samples

A total of 200 images have been used for training and 60 images have been used for testing. It has been found that the accuracy is 98.33%.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

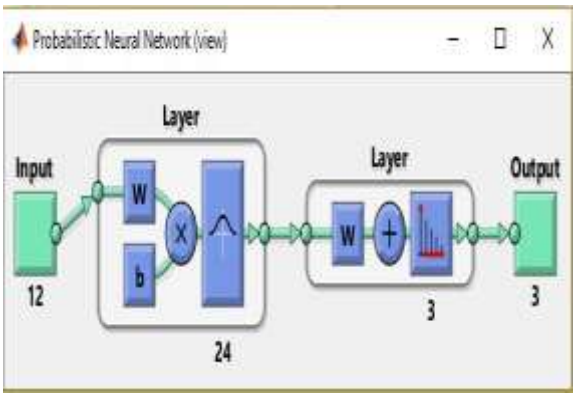


Fig.5: Designed Neural Network

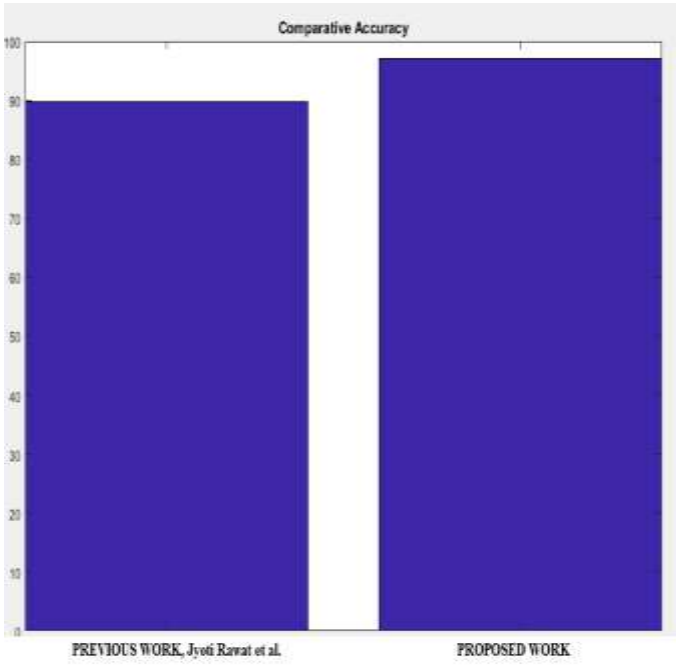


Fig.6: Comparative accuracy with previous work [4]

5. CONCLUSION

In this paper, a holistic approach for blood leukemia classification based of image processing and neural networks has been presented. It has been shown that the proposed approach used here can classify with relatively high accuracy owing to the effectiveness of the individual techniques. It has also been shown that the proposed approach can classify a large data set of images which again turns out as an indicative point towards the versatility of classification of the proposed method

REFERENCES

[1] Classification of Leukemia Blood Samples Using Neural Networks, by Malek Adjouadi, Melvin Ayala, Mercedes Cabrerizo, Nuannuan Zong, Gabriel Lizarraga, Mark Rossman, Springer 2010
[2] Classification of blasts in acute leukemia blood samples using K-Nearest Neighbor, by N.Z Supardi, M.Y Mashor, N.H Harun, F.A. Bakri, R.Hassan, IEEE 2011
[3] Detection of Leukemia in Microscopic Images Using Image Processing, by Chaitali Raje, Jyoti Rangole, IEEE 2014
[4] Computer Aided Diagnostic System for Detection of Leukemia using Microscopic Images, by Jyoti Rawat, Annapurna Singh, H S Bhadauria, Jitendra Virmani, Elsevier 2015

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[5] Image segmentation and classification of white blood cells with the extreme learning machine and the fast relevance vector machine by S. Ravikumar, Elsevier 2015
[6] Large Scale Image Feature Extraction from Medical Image Analysis, by Desai Devanshi Manojbhai and R.Rajamenakshi IJAERS 2016
[7] Efficacy of Azacitidine in autoimmune and inflammatory disorders associated with myelodysplastic syndromes and chronic myelomonocytic leukemia by Jean Baptiste et.al, Elsevier 2016
[8] Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review, Fuyong Xing and Lin Yang, IEEE 2016
[9] Detection of Blood Cancer in Microscopic Images of Human Blood Samples: A Review,: M.Saritha,Prakash.B.B,Sukesh.K and Shrinivas. B, IEEE 2016
[10] Design of New Architecture to Detect Leukemia Cancer from Medical Images,by Gurpreet Singh,Gaurav Bathla and SharanPreet Kaur IJAER 2016 using Microscopic Images

CITE AN ARTICLE

Jain, S., & Gupta, D. (2019). CLASSIFICATION OF BLOOD LEUKEMIA USING NEURAL NETWORKS AND IMAGE PRE-PROCESSING. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 8*(1), 223-230.