JCST Journal of Data Acquisition and Processing

ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7765811

PROBABILISTIC NEURAL NETWORK WITH LASSO REGRESSION BASED PREVENTIVE ANALYSIS USING DETECTION OF LUKEMIA FROM SMEAR BLOOD IMAGES

G. Jeyakumar[1], Dr.T. Kamalakannan[2]
[1]Research Scholar, Vel's University VISTAS
[2]Associate Professor, Department of Information Technology (BCA&IT)
School of Computing Sciences, VISTAS, Chennai,
jeyakumarmsc@gmail.com, kkannan.scs@velsuniv.ac.in

ABSTRACT:

The exact distinction of malignant leukocytes at minimal expenses during earlier phase of illness, which is a significant difficulty in disease diagnosis, is a real concern in the field of disease detection. Flow cytometry equipment is few, and the procedures available at laboratory diagnosis institutes are time- consuming, despite the high frequency of leukaemia. The current systematic review was undertaken to examine the works aimed at discovering and classifying leukaemia using machine learning, which was motivated by the possibilities of machine learning (ML) in disease detection. This research proposenovel technique in detection of Leukemia from smear blood images based on deep learning techniques. Here the input image has been processed for noise removal and image resize with smoothening the image. Then this processed image has been extracted using probabilistic neural network with lasso regression (PNN_LR) from which the images with leukemia has been extracted. The extracted image classification has been utilizing Support vector machine (SVM) in which the accuracy has been Enhanced from the classification output. For the test datasets, the method was capable of classifying leukaemia type containing accuracy, specificity, and sensitivity of 97.69 percent, 97.86 percent, and 100 percent, respectively, and for the validation datasets, 97.5 percent, 98.55 percent, and 100 percent, respectively. In addition, the system had a 94.75 percent accuracy rate for WBC counts, which included both lymphocytes and monocytes. In comparison to unsupported manual procedures, the computer-assisted diagnostic machine takes below 1 min to process and assign the leukaemia kinds.

Keywords: leukemia, **smear blood images, deep learning**, classification, PNN_LR, SVM

1. Introduction:

A doctor uses finding to determine the existence or absence of a given illness in a patient based on a particular database that may include symptoms, side effects, clinical photos, and tests. An incorrect diagnosis can have negative results. With increasing the cost of therapy, erroneous analyses might muddle treatment systems [1]. To assist doctors with accomplishing high indicative exactness, numerous associate frameworks were proposed. Numerous infections, including glaucoma, skin disease, bosom malignant growth, and leukemia, are as of now tended to by such frameworks. Early and precise conclusions could really decrease Treatments expenses rise, the possibility of abatement rises, and patients' lives are extended [2]. Leukemia is a common fatal infection that threatens the lives of many teenagers and children. Children under the age of five who are born are at a higher risk. According to a recent research, roughly 352k adults [3]. There are several reasons for this illness, including exposure to radiation and some synthetic substances, as well as familial heritage. Analyses can be carried out using a variety of techniques, such as a physical examination, a blood test, a blood count, and a bone marrow biopsy. Infinitesimal research is regarded as the most practical approach for generating preliminary results, but it is typically carried out physically by a weak administrator to the point of tiredness, which might occur as a result of doing several tests in a single day. Furthermore, such hand judgements are inherently inconsistent, as they are dull, laborious, and likely to differ amongst observer types. As a result, there is a need to develop robotized, low-cost frameworks that can distinguish between sound and unpleasant blood smear images with high precision yet without the need for personal intervention. The primary approach for determining the presence of leukaemia is microscopic blood testing [4]. The most well-known method of detecting leukaemia is the examination of blood spreads, although it is by no means the only 1.Interventional radiology is elective procedure for analysing leukemia. Be that as it may, radiological strategies, like percutaneous desire, biopsy, and catheter waste, experience the ill effects of acquiring constraints of imaging methodology responsiveness and goal of the radio pictures. Besides, different procedures, To differentiate leukaemia kinds, methods [5]. Minute well-known tools in recognizing leukaemia subtypes because to the time and cost constraints of these procedures. When an enormous preparation set is free, an AI(ML)computation will aid in distinguishing the platelets with leukaemia from the HEALTHY cells. The ALL-IDB leukaemia image warehouse is 1 of the datasets used as a standard byclinical analyzers. Another leukaemia dataset is available online first from American Society of Hematology(ASH)[6].

1.1. Related works:

This part covers a portion of the exploration led in the field of sicknesses, particularly blood illness identification and finding. A portion of these investigations relate to conventional strategies [7], which comprise of a few stages, for example, pre-handling, division, include extraction, and order, though different

techniques relate to profound learning-based strategies which utilize profound brain networks for start to finish learning assignments. A few ML calculations help to arrange and perceive leukemia illness from minute pictures. For instance, work [8] utilized SVM with K-Nearest Neighbor (k-NN) were used to describe the AML leukaemia subtype and achieved an accuracy of 83 percent. [9] used SVM to rank ALL leukaemia

subtypes with a 93 percent accuracy rate. Work [10] also used SVM and the c-mean grouping approach to extract WBC from the foundation, and they were able to do so with 90% accuracy. Although using a Deep

Learning (DL) technique looks to be more enticing, its effectiveness is highly dependent on the size and type

of the dataset used. Convolutional Neural Network (CNN) is a type of neural network that is commonly used

to deal with image organisation and enrolment challenges. CNN was used to recognise ALL leukaemia

subtypes in a study [11]. Their results were nearly 100 percent for paired order, comparing ALL and HEALTHY cases, and 96 percent for further classification of ALL subtypes. Creator [12] used a CNN model
using 5 convolutional to create a two-dimensional arrangement of ALL leukaemia subtypes with 96.6 percent accuracy. Unfortunately, with this type of brain network, presentation characterisation is required, an
enormous preparation information to figure out how to distinguish significant articles from the entire picture.

Notwithstanding, fostering an enormous preparation dataset is extremely tedious and is an exceptionally work escalated task. To stay away from this issue, we recommend extending the predetermined number of tests by picture increase. Utilizing an insufficient number of picture tests in the preparation dataset may prompt an overfitting issue [13]. Henceforth, the greater part of the scientists in the writing rely upon applying some picture change procedures to build the quantity of preparing set examples artificially to keep away from an overfitting issue. Work [14] applied middle and Wiener channels to eliminate commotion and fogginess. In the writing, a considerable lot of the picture change procedures have been utilized like picture turn and reflecting, histogram leveling, picture interpretation, dim scale change, obscuring pictures, and picture shearing [15,16]. The use of image expansion allows a DL strategy to be used, which requires a large number in the preparation dataset. In their research, Creator [17] used the ASH data collection to detect AML leukaemia. Google is another another source of unexplained leukaemia images, with images culled carelessly from many websites. In their study for identifying leukaemia, Work [18] used small photographs acquired from Google, where producers remarked on pictures without anyone else&#39;s knowledge. The successful implementation of Machine Learning-based leukaemia detection can be predicated on the use of a well-defined image dataset.

System Model

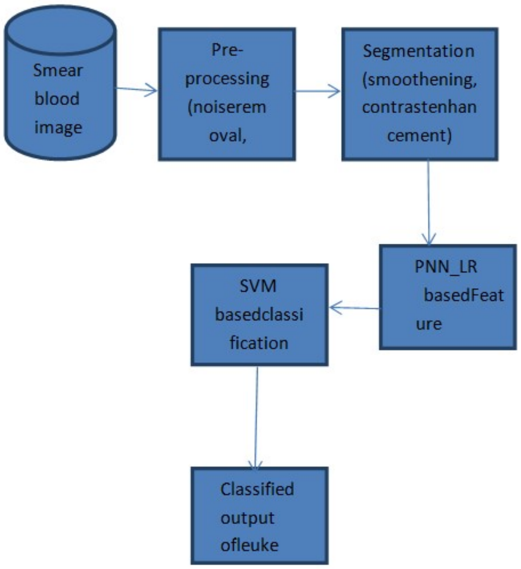

Smear blood image
Pre-processing (noiserem oval,
Segmentation (smoothening, contrastenhan cement)
PNN_LR basedFeat ure
SVM basedclassi fication
Classified output ofleuke

Figure-1 Over all Proposed Architecture

This section discusses the novel technique in detection of Leukemia from smear blood images based on deep learning techniques. Here the aim is collect the smear blood image dataset and pre-process, segment the image for noise removal, image resize and smoothening the image. Then the feature extraction is carried out using probabilistic neural network with lasso regression (PNN_LR) from which the images with leukemia has been extracted. The extracted deep features classified using Support vector machine based classification. The overall proposed architecture is shown in figure-1.

2.1. Dataset Description

The dataset for the proposed study was obtained from 2 distinct subgroups of a database collection. The first portion of the collection contains photographs of patients with B-ALL, or B-Lineage Acute Lymphoblastic Leukaemia, with a total of ninety photographs. The base veil and inner covering of the corresponding ALL image are shown in Figure 2. The 2nd part of the dataset, on the other hand, is made up of images of individuals who have been diagnosed with Multiple Myeloma, with a total of hundred images. For the related MM image, images with the base cover, core veil, and veil again for cytoplasmic of the cells. Pre- handling incorporates dealing with invalid qualities, one-hot encoding, standardization, multi-collinearity, scaling the information, rearranging and dividing information, and so forth In the proposed review, the changed information acquired after include choice is first standardized and afterward isolated into preparing and testing sets in the wake of rearranging.

2.2. Probabilistic neural network with lasso regression (pnn_lr) based feature extraction

A Bayes - Parzen classification is the PNN. The methodology's foundation dates back many years; but, due to a lack of suitable computing power not long ago, the technique was limited in its use. The PNN was the first to be shown, and it countless simple cycles completed in a multi-facet NN, Bayes – Parzen characterisation, it&#39;s worth taking a look at both the Bayes hypothesis for limiting likelihood and Parzen's approach for determining the probability thickening capability of arbitrary components.

Consider the following example x=[x1, x2,..., xp] derived from a collection of tests performed on multiple distinct populations (1,2,..., k,..., K). The Bayes hypothesis characterises an obscure example into the ith populace if the (earlier) percentage chance that an example has a place with the kth populace (class) is hk, the cost of misclassifying that example is ck, and the genuine likelihood thickness capacity of all populaces f1(x), f2(x),..., fk(x),..., fK(x) is known.

$$h_i c_i f_i(\boldsymbol{x}) > h_j c_j f_j(\boldsymbol{x})$$

The density function fk represents the concentration of class k examples around the unknown example(x). As indicated in Eq. 1, Bayes' theorem chooses a class with a high density in the specimen region, or if the cost of classification or prior probability is high. The multivariate PDF estimation, g(x), given by:

$$g(x_1, x_2, \ldots x_p) = \frac{1}{n\sigma_1\sigma_2 \ldots \sigma_p}\sum_{i=1}^{n} W \left(\frac{x_1 - x_{1,i}}{\sigma_1}, \frac{x_2 - x_{2,i}}{\sigma_2}, \ldots, \frac{x_p - x_{p,i}}{\sigma_p}\right)$$

r1, r2,..., and rp are smoothing parameters that reflect the confidence interval (often called as windows or kernels size) all around the mean of p uncertain variables. x1, x2,..., xp, xp, xp, xp, xp, xp W stands for a weighting factor that may be selected based on certain parameters, and n is the total no., of training samples. If all smoothing factors are taken to be same (i.e., r1 = r2 = r3 =:: = rp = r), and W would be a bell shaped Gaussian function, a reduced form of Eq.(2)is obtained.

$$g(\boldsymbol{x}) = \frac{1}{(2\pi)^{p/2} n\sigma^{p}} \sum_{i=1}^{n} \exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|^2}{2\sigma^2}\right)$$

Where in x is the ith preparation vector and xi would be the vectors of irregularity features(informative factors). The standard of the multidimensional colloidal is addressed by Eq.(3), in which each dissemination is centred on a single preparation model. It's worth assuming Decent weighing ability would be Gaussian(typical); alternative weighting capacities, such as the identical capability [W(d) = 1/1 + d2], might be used instead. The Parzen's PDF assessors asymptotically approaches toward fundamental thickness work as the example size, n, grows.

Consider the basic organization engineering shown in Figure 2, which includes four information hubs(p=4)in the information layer, 2 population classes(1 as well as 2), 5 preparation models(n1=5), and 3 models in class 2 (n2 = 3). This sample layering (in Figure. 1) is designed to include 1 neuron (hub) to every available preparation scenario, with the neurons split into 2 classes. Each class has one neuron performs an insignificant limit segregation; it merely holds the limit of the 2 summation neurons.

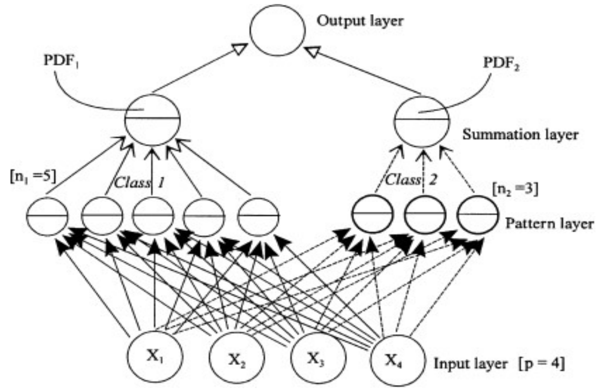


Figure 2. PNN having 4 inputs, 2 classes, with 8 trained example.

A preparation case is executed by PNN by first presenting it to other example layer neurons. In the example layer, each neuron calculates the distance here between supplied input vector and the prepared models specified by such a design neuron. PNN then applies Parzen windows (weighting capacity, W) to thisdistance metric, yielding effective beginning of any neuron inside the sample layer. As a result, each class'sinitiation is handled by comparison summary layer neuron, which sums all of the results in a single classtogether. The evaluated likelihood thickness work worth of a population of a certain class is acquired by using there sidual component of Parzen's assessor conditions(e.g.,constant multiplication in Eq.(3))to initiate each summation neuron. The commencement of the summing up neurons would be equitable to the back probability of each class if misclassified expense and previous largest is passed on to the result neuron, resulting in processing category and the probability that this models will fit into that class. The smoothing borders(r1,r2,...,and rp)of an irregularity factors are the major boundary that has not been completely resolved in order to obtain an ideal PNN. Selecting an unpredictable upsides of r's, creating the organisation, then evaluating this on a test(approved)set of models are all part of a direct system. This technique is repeated for different r's, and the combination of r's with the lowest misclassification rate (number of models misclassified) is chosen.

Previously, it was considered that the probabilities of development(p1)and no development(p0)were comparable to the independent breadth of growth and no growth model in the entire data set; specifically, p1=0.55 and p0=1-p1=0.45. The back probability of development (P1) for a given data vector x is calculated using Bayes' hypothesis in selected and comprehensive classes as follows:

$$P_1(\boldsymbol{x}) = \frac{p_1 g_1(\boldsymbol{x})}{p_1 g_1(\boldsymbol{x}) + p_0 g_0(\boldsymbol{x})}$$

g1(x) and g0(x) are growth and no-growth classes(Eq.(3)).(4).The confusion matrix was utilized to develop a method for calculating the accuracy of the produced PNN, as well as a method for comparing the different PNNs created using the four distinct training procedures. Because the PNN output(probability of occurrence)is basically continuous and the constructed PNN comprises classification, the probabilities were converted tosharp value one for development and Zero for no growth, with .5 as threshold possibility between the 2 states. You may then calculate the confusion matrix(C-matrix), that is expressed as:

$$\mathrm{C-matrix} = \begin{pmatrix} \mathrm{a} & \mathrm{b} \\ \mathrm{c} & \mathrm{d} \end{pmatrix} = \begin{pmatrix} 0 \rightarrow 0 & 0 \rightarrow 1 \\ 1 \rightarrow 0 & 1 \rightarrow 1 \end{pmatrix}$$

In which a, b, c, and d are network components that handle the number of instances as determined by the PNN. an is the number of non-event (i.e., no-development) instances set up by the PNN as nonissue andalluded to in Eq. (5) as 0 - 0, b is the number of non-event cases arranged by the PNN as circumstance (i.e., development) instances symbolised by 1 and made reference to in Eq. (5) as 0 - 1 (i.e., number of phoney problem cases), c is the number of occasion cases char arranged. The sum of a and d is the total number ofinstances that were correctly categorised, whereas (b + c) is the total no., of cases that were misclassified. The quality of the group's prediction was checked using 2 execution measurements obtained from the C-framework. FC=(a + d)/N as Well as FAR = b/(a + b), in which FC is part correct and FAR seems to be the misrepresentative problem rate(bogus positive) and N is the total number of instances(N=a+b+c+d), respectively. For each PNN generated using the four plans, these boundaries were obtained both for preparations and test collections.

With b = c = 0, the optimal classification system is unified, implying FC is hundred percent and FAR is zero percent. Hundred and seventy nine(T, aw)advancement/no advancement models were arbitrarily divided into hundred and forty three (seventy three developed and seventy not developed) categories used for PNN preparation, and thirty six (twenty six developed and ten not developed) models never used in PNN development but used to assess the validity of created PNN. Approximately twenty percent of the total dataset was made up of testing information. Figure 2 shows the average PNN for this model, which has 2 neurons in the info layer, hundred and forty three neurons in the illustration layer divided into seventy three neuron for class 1(developed) and seventy neuron for class 2(not developed), 2 summation neurons(1 per each class), and 1 result neuron that returns the possibility of development given by P1.

The LassoNet objective function is defined as

$$\underset{\theta,W}{\text{minimize}}\ L(\theta,W)+\lambda\,\|\theta\|_1$$

$$\text{subject to}\ \ \|W_j^{(1)}\|_\infty \le M|\theta_j|,\ j=1,\ldots,d.$$

Where L(,W) is the loss function, and W(1)j is the weights for feature j in the first hidden layer, as specified in Section 3.1. We stress that our objective is to sparsify the network in an organised manner that picks the important input characteristics for the whole network. We don't need to punish the weights in the remaining hidden levels since the network is feed-forward. The limitation is the essential concept in our method.

$$|W_{jk}^{(1)}| \le M\cdot|\theta_j|,\ k=1,\ldots,K$$

There are two steps to learning LassoNet. To begin, all model parameters are subjected to a standard gradient descent step. The input layer pair (W(1)) is then subjected to a hierarchical proximal operator. This sequential action simplifies the procedure’s implementation in common machine learning frameworks, requiring only a few lines of code modification from a conventional residual network. The method’s computational efficiency is an extra bonus. The cost of training the LassoNet regularisation approach is similar to that of training a single model.

SVM based classification:

Structural Risk Minimization (SVM) is a strategy for reducing structural risk (SRM). The input vector is mapped to a greater environment where a maximum separation hyperplane is constructed using SVM. 2 parallel hyperplanes are formed along either side of the plane to separate the data. The separation of the 2 parallel hyperplanes is maximised by the separating hyperplane. The concept is that the greater the separation between such parallel hyperplanes, the smaller the generalisation error of the classifier. The form’s data points are taken into account.

$$\{(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)\ldots\ldots\ldots,(x_n,y_n)\}.$$

Here yn=1 / -1 is just a variable that denotes the class to which xn belongs. A no., of sample taken is n. The p-dimensional actual vector is represented by each x n. To prevent factors (characteristics) with such a lot of fluctuation, scaling is required. The splitting (or separating) hyperplane, which separates the

data into two
half, may be used to observe this Training data..

Where w is then subjected to a hierarchical proximal operator. This sequential action simplifies the procedure's $w.x+b=o$ implementation in common machine learning frameworks, requiring only a few lines of code modification from a conventional residual network. The method’s computational efficiency is an extra bonus. The cost of training the LassoNet regularisation approach is similar to that of training a single model.

$$\begin{aligned} w.x + b &= 1 \\ w.x + b &= -1 \end{aligned}$$

The commencement of the summing up neurons would be equitable to the back probability of each class if misclassified expense and previous largest is passed on to the result neuron, resulting in processing category and the probability that this models will fit into that class.

$$w.\, x_i - b \geq 1 \quad \text{or} \quad w.\, x_i - b \leq -1$$

That is given as

Yi (w. xi–b) ≥ 1, 1≤i ≤n
Support Vectors are samples that run along the hyperplanes (SVs). The training data points closest to a hyper
plane with the biggest margin specified by M = 2 / w that defines support vectors.

$$y_j\,[w^T .\, x_j + b] = 1 \quad , i=1$$

A canonical Hyperplane with a maximum margin is called an Optimal Canonical Hyperplane (OCH). The following restrictions should be met by OCH for all data.

$$y_i[w^T .\, x_i + b] \geq 1 \quad ; \ i=1,2...l$$

Where n is the no., of training points, and l is the no., of training data points. A learning machine should minimise w2 subject to inequality constraints in order to identify the optimal separation hyperplane with a maximal margin.

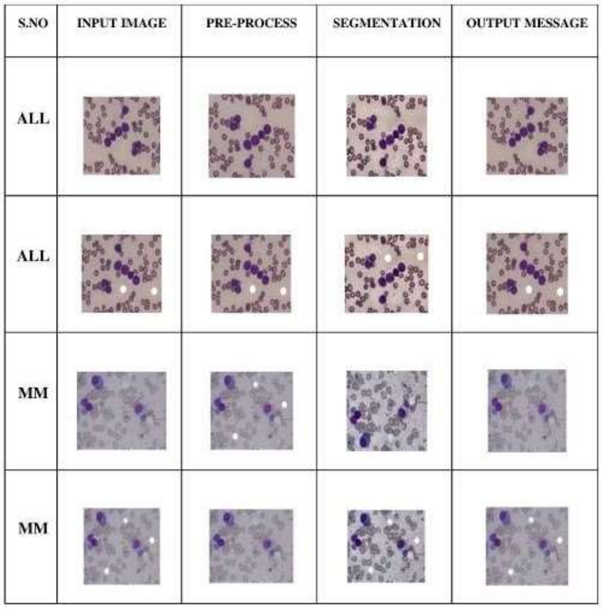
$$y_i\,[w^T .\, x_i + b] \geq 1 \quad ; \ i=1,2.......l$$

The capacity plans the preparing vectors xi into a higher (perhaps infinite) layered space. Then, in the this higher aspect space, SVM selects a straighter isolated hyperplane with the maximum edge. The error term’s reformative nature threshold is C > 0. In addition, the component. Because there are various part capabilities in SVM, deciding on a good bit task is also a test concern. However, there are a few well-known bit capacities for widespread use. In SVM, model selection is also a key challenge. SVM has recently demonstrated excellent performance in terms of information order. Its success is dependent on the fine adjustment of a few parameters that determine the speculative mistake. This boundary tuning approach is commonly referred to as model determination. If you’re using the direct SVM, you’ll just need to adjust the expenditure boundary C.

2. Performance Analysis

Tensor Flow, an end-to-end open-source stage, is used to build the arrangement model. On four hundred and thirty two images with 1k emphases, a paired characterisation model was created. Every emphasis improves the misfortune work with Adam Optimizer, resulting in the least amount of misfortune at the end of the cycle. The developed model was then used to predict the type of disease shown in the images. The model is prepared using GPU. The suggested model outputs are initially depicted in the next section. The following boundaries were used to make the correlation: Reliability, Precise, Memory, Selectivity, and F1 Measure. The positive class do not, while FN stands for the number of tests where the fully realise the negative class incorrectly. The table-1 depicts the suggested procedure's processing of information blood images.

Table-1 Processing of Input blood image based on various classes of cancer.



S.NO	INPUT IMAGE	PRE-PROCESS	SEGMENTATION	OUTPUT MESSAGE
ALL				
ALL				
MM				
MM				

2.1. Performance Metrics

The evaluation of the confusion matrices is dependent on the examination of other characteristics including reliability, sharpness, memory, and F1 measure. The estimated True Positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP) values are used to assess the provided parameters.

Accuracy: Number of accurately predicted values based on no., of forecasts is known as accuracy. It's written as an equation(14)

$$Accuracy\square\square\square \quad \frac{TP\square TN}{} \qquad (12)$$

TP□*TN*□*FP*□*FN*

Recallor Sensitivity: It's calculated as the proportion of accurately predicted values to total projected values. It's written as an equation(15)

TP
TP□
Re *call* □ ———
□□□□□□□(13)
FN

Specificity: The ratio of genuine positive values to total anticipated values is calculated. It's written as an equation(16)

$$\text{Precision}= \frac{TPT}{P+FP}$$

(14)

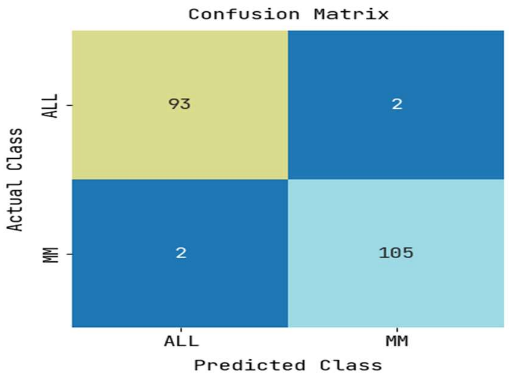
Confusion Matrix: It evaluates the proposed model's performance using a comparison of actual and anticipated values. The study is based on TP, TN, FP, and FN estimations. It's written as an equation(18)

$$\text{Confusion Matrix}=\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$$

(16)

True Positive (TP) is defined as a prediction number that is expected to be positive in an AI model.
False Positive (FP) is a prediction value that is first estimated as negative but is later predicted as positive in
an AI model.
True Negative (TN) showed that the AI model's predicted value was negative and that it was expected to be
unfavourable.
False Negative (FN) is a prediction value that is originally estimated as positive but is later predicted as negative in an AI model.

2.2. Confusion matrix for cancer detection using proposed method



The above figure gives confusion matrix for proposed leukaemia detection based on deep learning. Here the

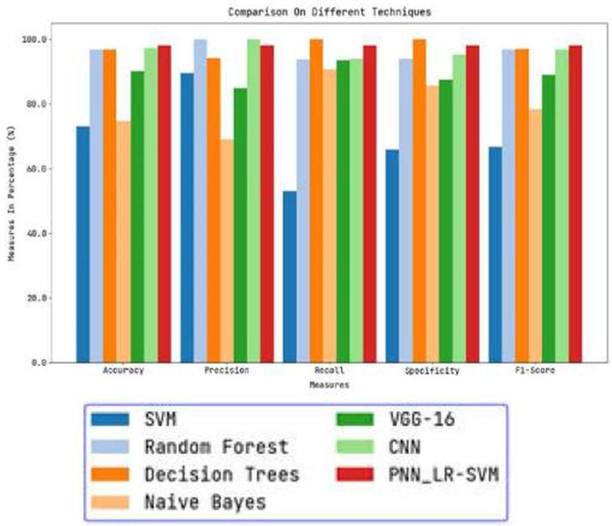
confusion matrix has been taken for predicted class and actual class of blood cancer detection. The table-1

shows overall parametric comparison based on tumor classes.

Table-2 Comparative Analysis Based On Classifiers

Parameters	SVM	Random Forest	Decision tree	Naïve Bayes	VGG	CNN	PNN_LR-SVM
Accuracy	73.02	96.83	96.77	74.6	90.1	97.25	98.02
Precision	89.47	100	94.11	69.05	84.88	100	98.02
Recall	53.12	93.75	100	90.65	93.58	93.97	98.01
Specificity	65.9	93.93	100	85.71	95.19	95.19	98.01
F-1 score	66.66	96.77	96.96	89.01	89.01	96.89	98.01

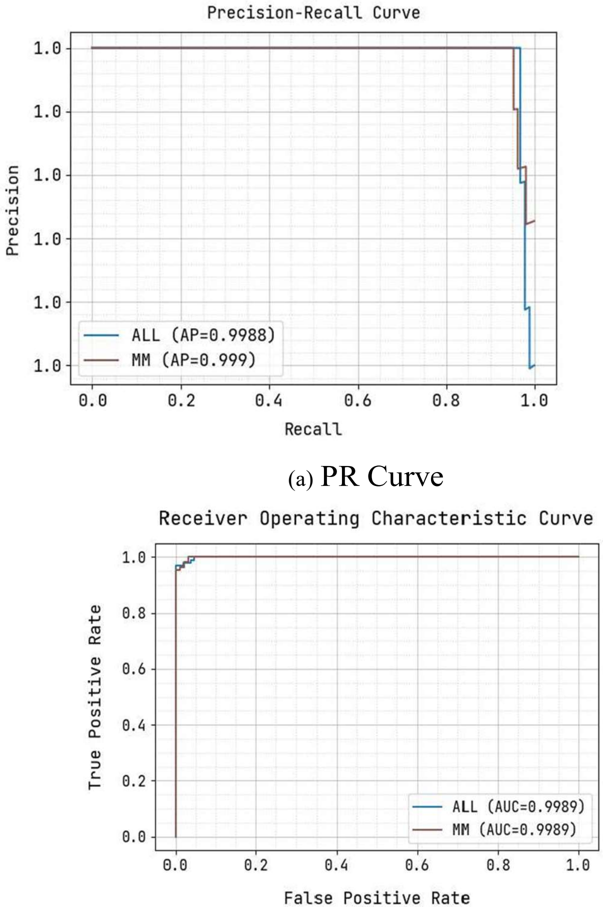
The below Figure-3 shows the comparative analysis of proposed PNN_LR-SVM with SVM, RF, DT, NB, VGG and CNN. Here the proposed technique



.

Figure-3 Comparative Analysis of Proposed Technique with Existing Technique

From the above graphical analysis the proposed technique obtained accuracy of 98.02%, Precision of 98.2 %, Recall of 98.01%, specificity of 98.01% and F-1 score of 98.01%. The enhanced results has been shown by below ROC curve and PR curve in figure-4 (a) and (b).



(a) PR Curve

(b) ROC curve

3. Discussion and Conclusion:

Leukaemia can be treated more successfully if it is detected early. Two classification methods were suggested in this work to discriminate among leukemia-free and leukemia-affected blood Smear pictures. These researches propose novel technique in detection of Leukemia from smear blood images based on deep learning techniques. Here the input image has been processed for noise removal and image resize with smoothening the image. Then this processed image has been extracted using probabilistic neural network with lasso regression (PNN_LR) from which the images with leukemia has been extracted. The retrieved picture was categorized using a Support vector machine (SVM), which improved the accuracy of the classification results. For the training and validation datasets, the system was able to classify leukemia types with an accuracy of 98.02 percent, precision of 98.02 percent, recall of 98.01 percent, specificity of 98.01 percent, and F-1 score of 98.01 percent. In addition, the system had a 98.02

percent accuracy rate for WBC counts, which included both lymphocytes and monocytes.

References

1. R. B. Hegde, K. Prasad, H. Hebbar, B. M. K. Singh, and I. Sandhya(2019), “Automated decision

support system for detection of leukemia from peripheral blood smear images,” Journal of Digital Imaging, vol. 33, pp. 361–374.

2. S.M.Namayandeh, Z.Khazaei, M.Lari Najafi, E.Goodarzi, and A.Moslem(2020). “GLOBAL Leukemia in children 0–14 statistics 2018, incidence and mortality and human development index (HDI):GLOBOCAN sources and methods,” Asian Pacific Journal of Cancer Prevention, vol. 21, no. 5, pp. 1487–1494.

3. QiaoN(2019).“ “A systematic review on machine learning in sellar region diseases: quality and

reporting items,” Endocrine Connections,,vol.8, no.7, pp.952–960.

4. E. Fathi, M.J. Rezaee,R. Tavakkoli-Moghaddam, A. Alizadeh,and A.Montazer(2020), “Design ofan integrated model for diagnosis and classification of pediatric acute leukemia using machine learning,”Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, vol. 234, no.10, pp.1051–1069.

5. Boldú, L., Merino, A., Acevedo, A., Molina, A., & Rodellar, J. (2021). A deep learning model (ALNet)for the diagnosis of acute leukaemia lineage using peripheral blood cell images. *Computer Methods and Programs in Biomedicine*, *202*, 105999.

6. Rastogi, P., Khanna, K., & Singh, V. (2022). LeuFeatx: Deep learning–based feature extractor for thediagnosis of acute leukemia from microscopic images of peripheral blood smear. *Computers in Biology and Medicine*, 105236.

7. Anilkumar, K. K., Manoj, V. J., &Sagi, T. M. (2021). Automated detection of b cell and t cell acute lymphoblastic leukaemia using deep learning. *IRBM*.

8. Ghaderzadeh, M., Asadi, F., Hosseini, A.,Bashash, D., Abolghasemi, H.,& Roshanpour, A.(2021).

Machine learning in detection and classification of leukemia using smear blood images: a systematic review. *Scientific Programming*, *2021*.

9. Dhalla, S., Mittal, A.,Gupta, S.,& Singh, H.(2021, January). Multi-model Ensemble to Classify Acute Lymphoblastic Leukemia in Blood Smear Images. In *International Conference on Pattern Recognition* (pp.243-253).Springer, Cham.

10. Patil, A. M., Patil, M. D., &Birajdar, G. K. (2021). White blood cells image classification using deep learning with canonical correlation analysis. *IRBM*, *42*(5), 378-389.

11. Alagu, S., & Bagan, K. B. (2021). Computer assisted classification framework for detection of acute myeloid leukemia in peripheral blood smear images. In *Innovations in Computational Intelligence and Computer Vision* (pp.403-410). Springer, Singapore.

12. Eckardt, J.N., Middeke, J.M., Riechert, S., Schmittmann, T., Sulaiman, A.S., Kramer, M.,

...& Bornhäuser,M. (2022).Deep learning detects acute myeloid leukemia and predicts NPM1 mutation status from bone marrows mears. *Leukemia*, *36*(1), 111-118.

13. Khandekar, R., Shastry, P., Jaishankar, S., Faust, O., & Sampathila, N. (2021). Automated blast cell detection for Acute Lymphoblastic Leukemia diagnosis. *Biomedical Signal Processing and Control*, *68*, 102690.

14. Sidhom, J. W., Siddarthan, I. J., Lai, B. S., Luo, A., Hambley, B. C., Bynum, J., ... & Shenderov, E.(2021). Deep learning for diagnosis of acute promyelocyticleukemia via recognition of genomically imprinted morphologic features. *NPJ precision oncology*, *5*(1), 1-8.

15. Patil, A.P.(2022). A Concise Review of Acute Myeloid Leukemia Recognition Using Machine Learning Techniques. *Cyber Intelligence and Information Retrieval*, 417-424.

16. Vogado, L., Veras, R., Aires, K., Araújo, F.,Silva, R., Ponti, M.,& Tavares, J.M.R.(2021).

Diagnosis of leukaemia in blood slides based on a fine-tuned and highly generalisable deep learning model. *Sensors*, *21*(9), 2989.

17. Anilkumar, K. K., Manoj, V. J., & Sagi, T. M. (2021). Automated detection of leukemia by pretrained

deep neural networks and transfer learning: A comparison. Medical Engineering & Physics, 98, 8-19.

18. Das, P.K., Pradhan, A., & Meher, S.(2021). Detection of acute lymphoblastic leukemia using machine learning techniques. In *Machine learning, deep learning and computational intelligence for wireless communication* (pp.425-437). Springer, Singapore.