International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 8 Issue 08 Series. II || Aug 2019 || PP 73-78

Detection and Classification of Blood samples using Soft Computing Techniques

RUCHI JOGI[1]
M.Tech. Scholar, Department of CSE, RITEE Raipur, (C.G). India
ruchi64jogi64@gmail.com

AVINASH DHOLE[2]
Asst.Prof., Department of CSE, RITEE Raipur (C.G.) India
avi_dhole33@rediffmail.com

Abstract- *Blood group of a person is always be a very important part of medical science. It has been detected by diluting the blood sample with some specific mixture then taking it on a glass plate and observing the pattern on microscope. But at the same time skilled person needed to detect it correctly. So, computer can play a very important role in detection of blood group if we can able to train the system with the images of all types of blood groups. In this paper, we are proposing a machine learning based approach to detect the blood group, where have created a dataset of all types of blood group images, then feature extraction have been done by using GLCM method, finally feature matrix is send to the SVM, Decision Tree and LDA classifiers for analysis of performance of machine learning algorithms and we found that LDA algorithm performs very well with more than 99% accuracy in classification.*

Keywords: *Blood Group, Feature Extraction, Machine Learning, Support Vector Machine.*

Date of Submission: 08-08-2019 Date of acceptance: 25-08-2019

I. Introduction

Blood is one of the significant fluid in human body, which transport oxygen and sustenance to body. In any case, next to this blood performs pH guidelines, diverse immunological capacity. Any type of blood is made out of three noteworthy sorts, red platelets (RBC) oxygen transporter and (WBS) which helps battle disease and help in the invulnerable procedure. Third significant segment is Platelets which is one more significant substance in blood piece. It serves to cluster the draining if any damage happens. On account of crisis if patient experiences basic damage and largepart of blood is misfortune, a blood transfusion is required[1,2].

The field of Image processing has experienced colossal valuable changes in the ongoing couple of years. Therapeutic imaging plays an indispensable part nowadays in different application utilized in the field of medicinal and is giving an extraordinary help to the individuals related to this field regarding time and the endeavors utilized. Different sickness have been recognized utilizing picture preparing strategies [1-3] and along these lines have given early stage identification and helped the specialists to fix the malady. In [1] Alireza, et al have dealt with Automated distinguishing proof of diabetic retinal exudates in computerized shading pictures, in [2] Veropoulos, K., et al have chipped away at The mechanized recognizable proof of tubercle bacill utilizing picture handling and neural figuring systems

II. Literature Survey

S.M. NaziaFathima[4], Classification of blood classification by tiny shading pictures. In this semi-mechanized procedure, the blood gathering is dictated by drops of shading pictures. At first, picture preprocessing is finished by histogram evening out and shading rectification. Next shading space interpretation is accomplished for changing over the RBC to HIS. Next, itextracts the shading and surface element of the picture utilizing the total histogram and hard lick strategy individually. Mismatch of the blood arrangement can incite the agglutination, and the reaction of the blood can cause sudden going of the patient. Regardless of this risk can be verified by transfusing 2 units of comprehensive supplier's 0 adverse blood just in emergencies to any remarkable blood pack individuals. Since minimal human misstep can be deadly if there ought to be an event of blood transfusion. So it is indispensable to get automate these blood bundle ID strategies and get definite results if there ought to emerge an event of emergency. [5].A couple of structures have been made

Autoamted approach for ID of blood gathering utilizing different picture handling steps and delicate figuring procedures [6].

Picture division is one of the most basic strategies of picture preparing. In division, a greater picture is isolated into various sub pictures. While the calculations run separately on the sub-isolated pictures, the computations happen all the more explicitly and the outcome turns out to be increasingly exact. There are a few different ways of picture division. Otsu strategy is one of them. Otsu is a programmed edge choice locale based division technique [7].

Another Significant and significant picture preparing procedure is thresholding. Thresholding does binarization on any picture. Some unique thresholding systems likewise does denoising. Sometimes, some fragmented picture ends up shady and the significant data which is should have been removed turned out to be muddled to recover. In such circumstances thresholding is exceptionally useful [8]. In this way, essentially, thresholding methods makes a picture in highly contrasting and it makes the picture much more clear. One mechanized structure was raised where the scientist proposed [9] the entire test was done dependent on slide test for deciding blood classifications and a product created utilizing picture preparing systems. The picture was prepared by picture handling strategies created with the IMAQ Vision programming from National Instruments [10].

III. Methodology

Methodology used in this paper can be comprise in following steps(shown in figure 1) :

A. Dataset Preparation:

Here, we have created our local dataset from various sources like some images are taken form research papers, some images are taken from local pathology lab and some are collected from internet. Then we have categories the images into 9 categories i.e. Invalid image, A positive, A negative, B positive, B negative, AB positive, AB negative, O positive and O negative.

B. Segmentation

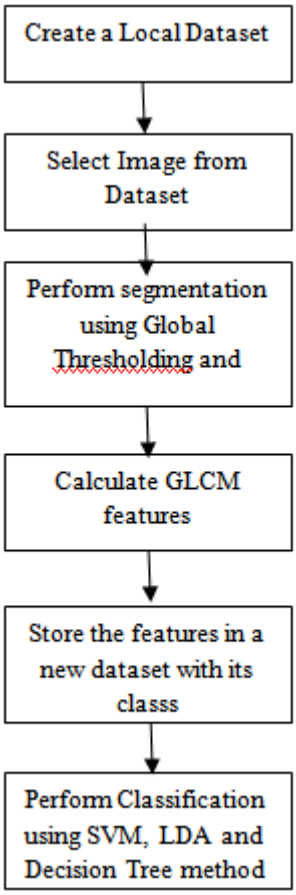
For the purpose of segmentation, we have applied the global thresholding method, which follows following steps:

Step 1: Divide the image in n×n parts.

Step2: Calculate the average intensity values of each block

Step3: now select this average value as threshold value T.

Step4: All the intensities above T will set as a fore ground object and all intensities less than T will treat as back ground object



E

Fi**gure** 1: Flowchart of methodology

C. Calculate GLCM Features:

GLCM is a gray level co relation matrix features, which extracts the texture feature of input image. Here, we have applied this feature extraction technique because every blood group image contains different texture information. So it can be easily classify on using these features. Here, we have calculated the homogeneity, contrast, entropy and correlation features of image in 4 different directions. So, for each direction we got 4 features and total 16 features of each image.

$$Correlation = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2} \quad (1)$$

$$Contrast = \sum_{i,j=0}^{N-1} P_{ij} (i-j)^2 \quad (2)$$

$$Entropy = \sum_{i,j=0}^{N-1} -\ln(P_{ij}) P_{ij} \quad (3)$$

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} \quad (4)$$

Where,

P_{ij} = Element i,j of the normalized symmetrical GLCM

N = Number of gray levels in the image as specified by Number of levels in under Quantization on the GLCM texture page of the Variable Properties dialog box.

μ =the GLCM mean (being an estimate of the intensity of all pixels in the relationships that contributed to the GLCM)

D. Create a feature dataset:

Then features dataset id created by adding all these features to a single dataset, and at last column class is added to make it suitable for training of any neural network or machine learning model.

E. Classification using Soft Computing Technique:

For classification purpose , we have used many classifiers but performance of decision Tree, support vector machine and Linear discriminate analyzer(LDA) was good so they are considered as a selected classifiers for our work[11].

Decision tree: A decision tree may be a tree within which every branch node represents an alternative between variety of alternatives and every leaf node represents a choice.It is a sort of supervised learning classifier that's largely utilized in classification issues and works for each categorical and continuous input and output variables. It's one in all the foremost wide used and sensible ways for inductive illation[12].

Step1: Place the best attribute of the dataset at the root of the tree
Step2: Split the training set into subsets.
Step3: Repeat step1 and step2 on each subset until you find the leaf nodes in all the branches of the tree

Algorithm of Support Vector Machine:
Input: initialize subset S = { 1,2,3.....}
Output: Rank list according to smallest weight R
Step1: Initially defined R = { }.
Step2: Repeat step 3 to 8 until G is not empty.
Step3: Train support vector machine model using G.
Step4:Compute weight W vector for SVM
Step5: Compute Rank R= W*W
Step6: Rank features and sort accordingly
Rank $_{new}$ = Sort (Rank);
Step7: Update feature rank list
Update R = R + G (Rank $_{new)}$
Step8: Eliminate feature with smallest rank
Update G = G- G(Rank $_{new}$)
Step9: End

Linear discriminate Analyzer (LDA):It consists of statistical properties of your data, calculated for each class. For a single input variable (x) this is the mean and the variance of the variable for each class. For multiple variables, this is the same properties calculated over the multivariate Gaussian, namely the means and the covariance matrix. These statistical properties are estimated from your data and plug into the LDA equation to make predictions. These are the model values that you would save to file for your model [13].The LDA model estimates the mean and variance from your data for each class. It is easy to think about this in the univariate (single input variable) case with two classes.

The mean (mu) value of each input (x) for each class (k) can be estimated in the normal way by dividing the sum of values by the total number of values.

muk = 1/nk * sum(x) (5)

Where muk is the mean value of x for the class k, nk is the number of instances with class k. The variance is calculated across all classes as the average squared difference of each value from the mean.

sigma^2 = 1 / (n-K) * sum((x – mu)^2) (6)

Where sigma^2 is the variance across all inputs (x), n is the number of instances, K is the number of classes and mu is the mean for input x.

IV. Result Analysis

On applying combination of both image processing and soft computing techniques, we could get the highest accuracy more than 99 percent. First, we are showing the extracted features in figure 2 and description of features in figure3. Here, it shows the features distribution of data of 9 classes.

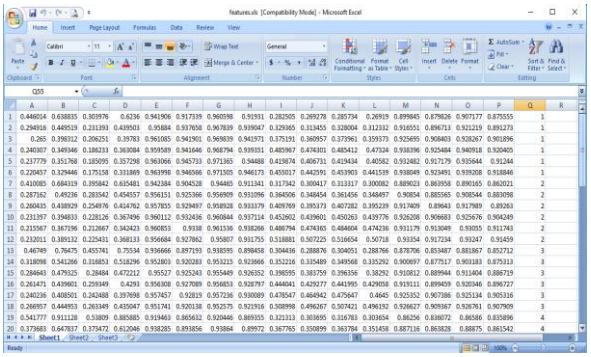


Figure 2: 16 features extracted from each image

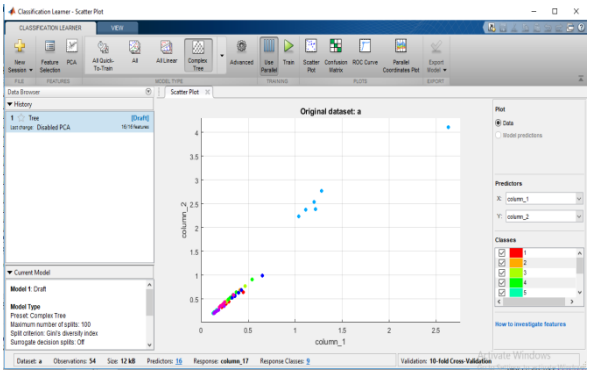


Figure 3: dataset description

After applying classification techniques on 54 images randomaly selected from our dataset the resultant confusion matrix is shown in figure 4

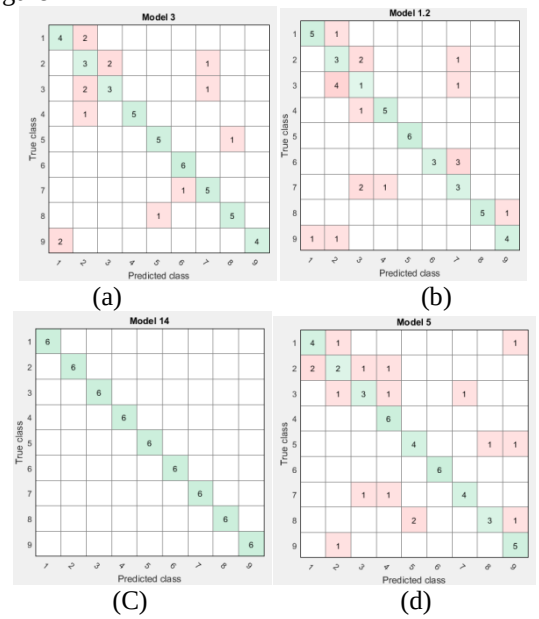


Figure 4: Confusion matrix of classification using Quadratic SVM in figure (a), Liner SVM in (b) LDA in figure (c) and Decision Tree in figure (d)

Accuracy of linear SVM (in percentage)= (sum of diagonal /total)*100 = (35/54)*100 = 64.8%
Accuracy of Quaderitic SVM (in percentage) = (sum of diagonal /total)*100 =(40/54)*100 = 74.07%
Accuracy of LDA (in percentage) = (sum of diagonal /total)*100 =(53/54)*100 = 99.12%
Accuracy of Decision Tree (in percentage) = (sum of diagonal /total)*100 =(37/54)*100 = 68.5%
Above result is shown in table 1 and graphcal representation of results is shown in figure 5 as :

Table 1: Accuracy of different classifiers

S.No.	Classifier	Accuracy(in percent)
1	Linear SVM	64.8
2	Quaderitic SVM	74.07
3	Decision Tree	68.50
4	LDA	99.12

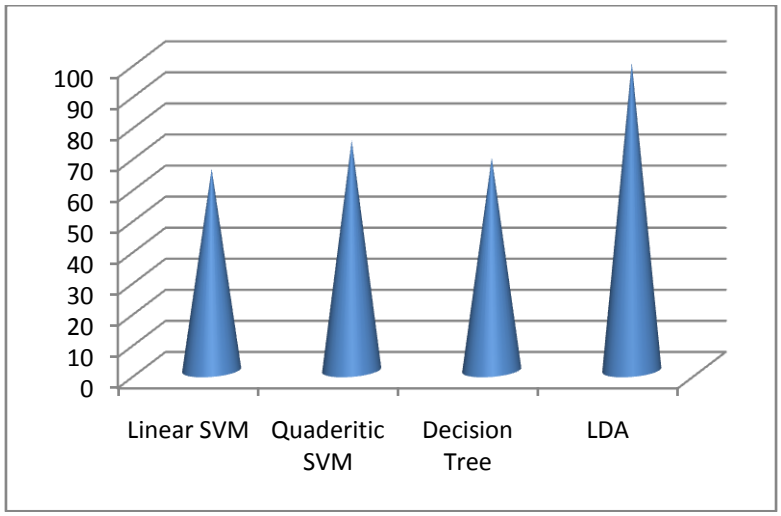


Figure 5: accuracy comparisons

V. Conclusion

Nowadays, we have seen the importance of right prediction of blood group for various purposes. So, to increase the accuracy computer can play an important role in prediction of blood group. This motivated us to design a machine learning based model for accurately classify the different blood groups. here, we have created the dataset of 9 classes and features are extracted to create a feature matrix, then on applying various classification method we come to know that, linear discriminate classifier performs well with more than 99% of accuracy of classification in all classes. We recommend the use of this classifier and for future work some features may be reduce to reduce the time needed.

References

[1]. Osareh, Alireza. "Automated identification of diabetic retinal exudates in digital color images." British Journal of Ophthalmology 87.10 (2003): 1220-1223.
[2]. Veropoulos, K and Campbell, C and Learmonth, G and Knight, B and Simpson, J. "The automated identification of tubercle bacilli using image processing and neural computing techniques." ICANN 98. Springer London, 1998. 797-802.
[3]. Yun, Wong Li and Acharya, U Rajendra and Venkatesh, YV and Chee, Caroline and Min, Lim Choo and Ng, EYK. "Identification of different stages of diabetic retinopathy using retinal optical images." Information Sciences 178.1 (2008): 106-121.
[4]. Naziya Fatima S.M. "Classification of blood types by microscopic color images machine learning and computing.vol 3, no. 4, pp.376-379, 2013.
[5]. Ferraz, Ana. "Automatic system for determination of blood types using image processing techniques." Bioengineering (ENBENG), 2013 IEEE 3rd Portuguese Meeting in. IEEE, 2013. 8
[6]. IP. Sturgeon, "Automation: its introduction to the field ofblood group serology," Immunohematology Journal ofBlood Group Serology and Education, vol. 17, no. 4,2001.
[7]. M. H. J. Vala and P. A. Baxi, "A Review on Otsu Image Segmentation Algorithm," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 2, no. 2, pp. 387–389, Feb. 2013.
[8]. D. T. R. Singh, S. Roy, and O. I. Singh, "A New Local Adaptive Thresholding Technique in Binarization," IJCSI International Journal of Computer Science Issues, vol. 8, no. 6, no.2, Nov. 2011.
[9]. A. Ferraz, "Automatic system for determination of blood types using image processing techniques," 2013 IEEE 3rd Portuguese Meeting in Bioengineering (ENBENG), 2013.
[10]. IMAQ, "IMAQ Vision Concepts Manual", National Instruments, Austin, 2004.
[11]. SumanPandey, Deepak Kumar Gour, Vivek Sharma" Comparative Study on Classification of Thyroid Diseases" International Journal of Engineering Trends and Technology (IJETT) – Volume 28 Number 9 - October 2015.
[12]. G. RasithaBanu**"** A Role of decision Tree classification data Mining Technique in Diagnosing Thyroid disease" International Journal of Computer Sciences and EngineeringVolume-4, Issue-11 2016.
[13]. M. Nazari, F.Seiti, and M. Teshnehlab. "An Experimental Comparative Study on Thyroid Disease Diagnosis Based on Feature Subset Selection and classification" International Journal of Electrical & Computer Sciences IJECS-IJENS ,(2012) Vol: 12 No: 01 pp.13-19.