Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi
Dokuz Eylul University Faculty of Engineering Journal of Science and Engineering

Basılı/Printed ISSN: 1302-9304. Elektronik/Online ISSN: 2547-958X

A Comparative Study of Automatic Detection of Acute Lymphocytic Leukemia with Machine Learning Methods

Akut Lenfositik Löseminin Makine Öğrenimi Yöntemleriyle Otomatik Tespitine İlişkin Karşılaştırmalı Bir Çalışma

Canan Kocatürk [1], Cemre Candemir [1*], İlker Kocabaş [1]

[1] Ege Üniversitesi Uluslararası Bilgisayar Enstitüsü, İzmir, TÜRKİYE
*Sorumlu Yazar / Corresponding Author**: cemre.candemir@ege.edu.tr

Geliş Tarihi /Received: 17.02.2022 *Araştırma Makalesi/Research Article*
Kabul Tarihi / Accepted: 31.05.2022 *DOI:10.21205/deufmd.2022247229*
Atıf şekli/ How to cite: KOCATÜRK, C., CANDEMİR, C., KOCABAŞ, İ.(2022). A Comparative Study of Automatic Detection of Acute Lymphocytic Leukemia with Machine Learning Methods. DEÜ FMD 24(72), 1021-1032.

Abstract

Acute Lymphocytic Leukemia (ALL) is one of the most prevalent types of leukemia which has the risk of death of children is relatively higher than adults. The early diagnosis of this disease is crucial and it can be detected by examining the morphological changes of the blood cells. In this study, we exhibit a comparative study on the automatic classification and identification of the ALL with machine learning methodologies. Acute Lymphoblastic Challange Database (ALL-CDB) served by the Cancer Imaging Archive, which consists of 6500 digital microscopic pathology images from 118 subjects, is used. As the first step, the geometric features are extracted and after, the feature selection was performed with Principal Component Analysis (PCA). Finally, the classification process on the selected features was carried out by using Naive Bayes, k-Nearest Neighbor (k-NN), Linear Discriminant Analysis (LDA), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) neural network methods. The results between the methodologies have been analyzed in terms of accuracy, precision, recall, and F1-score metrics. According to the results, MLP gives the both highest accuracy and F1-score with 97% to classify the ALL cells for leukemia.

Keywords: *Acute Lymphocytic Leukemia, machine learning, classification, multilayer perceptron, support vector machine, decision tree, linear discriminant analysis*

Öz

Akut Lenfositik Lösemi (ALL) en sık görülen lösemi tiplerinden biridir ve çocukların ölüm riski yetişkinlere göre nispeten daha yüksektir. Bu hastalığın erken teşhisi çok kritik olup, kan hücrelerinin morfolojik değişiklikleri incelenerek tespit edilebilir. Bu çalışmada, ALL'nin makine öğrenmesi metodolojileri ile otomatik olarak sınıflandırılması ve tanımlanması üzerine karşılaştırmalı bir çalışma sunuyoruz. Çalışmada, 118 deneğe ait 6500 dijital mikroskobik patoloji görüntüsünden oluşan Kanser Görüntüleme Arşivi tarafından sunulan Akut Lenfoblastik Görüntü Veritabanı (ALL-CDB) kullanılmaktadır. İlk adım olarak geometrik öznitelikler çıkarılmıştır ve ardından Temel Bileşen

Analizi (PCA) ile öznitelik seçimi yapılmıştır. Son olarak Naive Bayes, k-En Yakın Komşu (k-NN), Lineer Diskriminant Analizi (LDA), Karar Ağacı (DT), Rastgele Orman (RF), Destek Vektör Makinesi (SVM) ve Çok Katmanlı Algılayıcı (MLP) yöntemleri kullanılarak seçilen öznitelikler üzerinde sınıflandırma işlemi gerçekleştirilmiştir. Metodolojiler arasındaki sonuçlar, doğruluk, kesinlik, hatırlama ve F1-skor metrikleri açısından analiz edilmiştir. Sonuçlara göre MLP, ALL hücrelerini sınıflandırmak için %97 ile hem en yüksek doğruluk hem de F1-skorunu vermektedir.

Anahtar Kelimeler : *Akut Lenfositik Lösemi, makine öğrenmesi, sınıflandırma, karar ağacı, çok katmanlı algılayıcı, destek vektör makinesi, lineer diskriminant analizi*

1. Introduction

Leukemia is a common type of cancer, especially seen in childhood. Acute Lymphocytic Leukemia (ALL), which is a sub-type of leukemia, accounts for the majority of leukemia in children. ALL is a type of blood cancer caused by the uncontrolled and rapid increase of lymphoblast (immature lymph) as a result of the inability of normal blood cells to fulfill their duties. Although most cases of ALL are seen in children, the risk of death is relatively higher in adults [1]. According to the 2021 report of Leukemia and Lymphoma Society, approximately 158 people each day or more than six people every hour, someone in the US dies from blood cancer[2]. Moreover, the expected total number of living with leukemia is 397,501 people, and on the other hand, the report asserts that 23,660 people are expected to die from leukemia in 2021, only in the US. Globally, it is reported that ALL diseases increased from 49.1 thousand in 1990 to 64.3 thousand in 2017, during the 27 years. In 2018, it is also reported that there were 437,000 new leukemia cases and 309,000 deaths worldwide in total [3].

As is known, early diagnosis is the most important factor in ALL, as in many types of cancer. However, it is a challenging goal for several aspects. First, it should be found out the cancer markers to be able to seperate the cancer and healthy cells. However, different types of cancer differ in markers and this variation is caused by pathogenesis, origin, prognosis, age, sex, and races [4]. On the other hand, they may be both be affected by internal factors such as genetic and external factors such as radiation, as well.

Another challenging point is classifying the cancer type according to the cell form and the appearance. Depending on the type, cancer cells may split at different speed, exhibit different behaviour in spreading and require different medical treatments. In order to the treatment method to be chosen correctly by the physicians, the analysis of the cells should be done correctly. The difficulty in distinguishing the cancerous cell from other cells by morphological examination manually may lead to the progression of the disease and delay in its treatment. And in such cases, the delay has vital importance on the patient. In addition to the abovementioned challenges, the reason why the detection of cancerous cells is difficult is due to the similarity of these cells to normal cells and their slight differences.

At this point, machine learning methodologies present an effective solution in detecting cancer cells with many advantages. With the rapid developments in machine learning studies, it is possible to get more accurate results from imaging techniques every day [5]. By computer-based approach, it is aimed to improve the conditions, to reduce the rate of misdiagnosis in test results by modeling the treatment method, and to increase the quality of life of both physicians and patients by putting the results to a certain standard. At the same time, both information that may be overlooked and the time spent by the expert are reduced. For these reasons, the importance of image processing and machine learning infrastructure, and diagnostic applications has increased [5-8].

The main aim of this study is to find an efficient and robust method for detecting ALL cancer cells from healthy ones. There are various types of classifier methods and so, matching the efficient and suitable methodology and the problem leads to another question. To answer this, we followed a comparative point of view by implementing six well-known methodologies on an ALL dataset. After that, we propose a multilayer neural network model for classifying cell images as cancerous and non-cancerous. Finally, we implemented and discussed each method on a collection of 6500 images from Cancer Image

Archive [5] for observing the advantage and disadvantage points of each method.

The rest of the manuscript is organized as follows: In Section 2, the related studies in the literature are summarized. In Section 3, the data set and preprocessing steps are mentioned in materials and the classification methods are presented in methods. The comparative results and discussions about findings are given in Section 4 and conclusions and feature works are presented in Section 5.

2. Related Work

Since it is a crucial and challenging problem, there are various studies investigating the diagnosis of leukemia with different methodologies. The Acute Lymphoblastic Challange Database (ALL-CDB) is commonly used in most studies [5]- [6] for this problem.

In [5], the images were first preprocessed and then features such as color-based, geometric amd statistical were extracted.Then, the image dataset, whose features were extracted, was classified by Gradient Boosting Decision Tree (GBDT) and SVM algorithms. When the algorithms are compared in terms of performance values, GBDT performed better than the SVM algorithm with 85.2% accuracy.

In addition to the dataset used in our approach, in [6], a new approach was developed for the detection of leukemia subtypes from blood cell images using the American Society of Hematology (ASH) Image Bank dataset. Convolutional Neural Networks (CNN) method was used for this. According to the results obtained, this method has 85.25% and 81.74% accuracy in classifying all subtypes as healthy and multiclass leukemia subtypes, respectively.

There are various studies using different datasets apart from ALL-CDB. Microscopic images in [7], [8] and [9] were preprocessed to prepare them for classification. The k-means clustering algorithm was used to extract the cell images at [10] and [11]. In [12], Otsu's Thresholding method was used as the segmentation technique. In [13], Watershed used the segmentation technique.

In [7], the Histogram of Oriented Gradients (HOG) feature descriptor was used to extract the features. Since the size of the feature vector obtained by extracting the features is quite large, the size of the data set has been reduced by PCA method. In [14], openCV and skimage were used to extract the relevant features from the blood image.

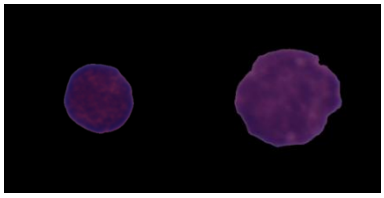
In [13] and [15], texture and color features were used. In [15], shape attributes are also used.

The classification process was carried out using different algorithms. Successful results were obtained with the k-NN algorithm in [10], [12] and [15]. In [10], 92.8% accuracy was obtained with the kNN algorithm, and in [12], 96.25% accuracy was obtained. Apart from the k-NN algorithm in [10], Naive Bayes was used and in [15], SVM, Naive Bayes and DT were used. In [7] and [13], they proposed a high-fidelity model with the SVM algorithm.The SVM method was the most successful algorithm with 99.05% accuracy in [7] and 96.93% in [13]. In addition to the SVM method, RF, Logistic Regression (LR) and DT were used in [7], and Artificial Neural Networks were used in [13]. In [11], 98.88% accuracy was obtained by using the MLP method. In [8], decisions regarding ALL subtypes were made based on the Fuzzy System. With this method, an accuracy of 93.75% was obtained. In [14], Convolutional Neural Networks (CNN), Feed Forward Neural Network (FNN), SVM and k-NN methods were used. Among these methods, CNN was the most successful algorithm with an accuracy rate of 98.33%. In [9] and [16], leukemia subtypes were classified using the Convolutional Neural Network of Deep Learning.With this method, an accuracy of 97.78% in [9] and 99.50% in [16] was obtained. Apart from this method in [9], SVM, k-NN, Artificial Neural Networks and Naive Bayes methods were used.

3. Materials and Methods

3.1. Data Set

In this study, the Acute Lymphoblastic Challenge Database (ALL-CDB) database provided by Cancer Image Archive was used [17]. The ALL-CDB dataset was presented in ISBI 2019 challenge for the detection of ALL cells [13], and it is a commonly used database in classification problems. The whole dataset contains 6500 real-world microscopicimages from 118 patients (69 diseased / 49 healthy) in total, where the number of cancerous and healthy cell images is equal. Each of these cells is at 450x450 pixel image scale and labeled as *'all'* for cancerous and



(a) (b)

Figure 1. Image examples included in the ALL dataset: healthy cell (a) from patients without ALL, probable lymphoblast from ALL patients (b).

'hem' for healthy. In Figure 1, a *'hem'* (Fig.1a) and *'all'* (Fig.1.b) cell image examples are presented. As shown in Figure 1, healthy and cancerous cell images differ both in size and shape. By considering these changes, some geometric features can be determined for defining healthy and cancerous cells (see Sec.3.3 for further detail).

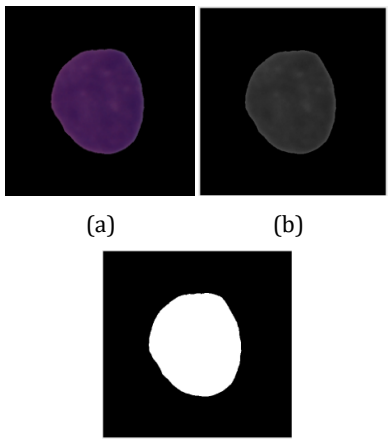
3.2. Data Preprocessing

Once the images are acquired, they should be preprocessed before fed into the classification algorithms. For all models applied on this study, following preprocessing steps were followed:

The first step is converting all images from RGB format to gray-scale images for further processing. Then, segmentation process was applied. Segmentation is a partitioning process that converts digital images into multiple images. The main purpose is to obtain more representative images for the analysis. The segmentation process is typically used to determine the location and boundaries of objects in the image[18]. Here, the global thresholding-based segmentation technique, one of the most commonly used techniques, is followed which divides an image according to the intensity of the pixel value[19]. Global thresholding can be defined as shown in 1,

$$\left.\begin{aligned} g(x,y) = 1 \; if \quad f(x,y) > T \\ g(x,y) = 0 \; if \quad f(x,y) \le T \end{aligned}\right\} \qquad (1)$$

where $f(x,y)$ is an input image and $g(x,y)$is a binary image, generated depending on the threshold value $T(x,y)$.



(a) (b)

(c)

Figure 2. (a) Digital representation of microscopic blood cells (b) The grayed out image (c) The image formed after thresholding.

This method is more suitable for images with a dual-mode histogram. With the thresholding process, a binary image is obtained from the gray image. The advantage of this process is that it simplifies classification and recognition processes by reducing the complex information in the image. The threshold value can be set manually or automatically using information from the image's attributes [19].

Figure 2 illustrates the abovementioned preprocessing steps. Microscopic blood cells are presented in RGB image format in Fig. 2(a), the grey-scale state of the histopathological image is given in Fig.2(b) and in Fig.2(c), the preprocessed image form can be seen after segmented by global thresholding.

In our approach, microscopic images (*T=50*) were segmented into multiple regions of interests (ROIs) and background based on the threshold value. After this process was completed, morphological operations were applied to the images obtained as the third step.

Morphological opening is a technique for removing defects from photographs that impact their texture and shape. As a result, morphological processes play a crucial role in image processing, particularly in picture segmentation. Erosion and dilation are combined in the opening operator. It erodes the

image first using the structuring element, then dilates it using the same structural element. Opening smoothes an object's boundaries and removes little undesirable objects from within the image [19]. This process is defined with 2.

$$A \circ B = (A - B) + B \qquad (2)$$

Here, A is an image and B is a structural element. In this study, morphological opening process was applied to the binary images obtained after thresholding.

3.3. Feature Extraction and Selection

Hematologists believe that the geometry of the nucleus is one of the most important factors that may be used to characterize cells. The size and shape of a nucleus can be determined using geometric features. These features are calculated from the nucleus binary image [20]. A total of 10 geometric features were obtained for the nucleus and cytoplasm. These features are listed below.

- f1: the total number of pixels in the nucleus.
- f2: the length, in pixels, of the major axis of the ellipse surrounding the nucleus.
- f3:the length, in pixels, of the minor axis of the ellipse surrounding the nucleus.
- f4: a measure of how far the object has ceased to be circular. Healthy lymphocytes are more circular than other diseased cells, so this feature is very important.
- f5: the angle between the ellipse's major axis length and the x-axis.
- f6: area of the smallest convex polygon enclosing the nucleus.
- f7: the diameter of the circle whose area is the same as the nucleus.
- f8: the ratio of the number of pixels in the nucleus to the area of the convex polygon containing the nucleus.
- f9: the ratio of the pixels in the nucleus region to the total pixels in the limiting frame and is obtained by dividing the nucleus area by the area of the limiting frame.
- f10: the distance between each adjacent pixel on the border of the nucleus.

After determining the geometric features, the principal component analysis (PCA) method has been applied on the feature set. PCA is a multivariate statistical method used in the domains of recognition, classification, and image compression that uses linear combinations of variables to describe the variance-covariance structure of a set of variables, providing dimension reduction and interpretation [21]. PCA, which is the most widely used algorithm for feature extraction, finds a new set of dimensions where all dimensions are orthogonal and ordered according to the variance data between them [22]. Number of p variables with the number of n measurements show interdependence structure with PCA; linear, orthogonal and independent k variables called principal components are transformed into new variables [21]. Thus, the size of the data set has been reduced and features that do not contribute to the performance of classification algorithms are removed.

3.4.Classification Methods

The classification process is the correct distribution of the extracted data over various classes. In order to the data distribution to be successful, the classification algorithms are trained by the given training set and thus they learn the distribution of the data. The trained classifier models determine which class it belongs to by looking at the characteristics of the incoming data and perform the assignment [23]. According to this process, digital pathology data is passed through various classification algorithms and classified as cancerous and healthy. Seven different machine learning algorithms were used in the study.

3.4.1. Naïve Bayes

Naive Bayes (NB) is a probabilistic classifier that builds a probability set by measuring the frequency and combinations of values in a dataset [24]. The approach is based on Bayes' theorem, and all variables in the dataset are assumed to be independent. This assumption is rarely encountered in real-world problems; hence it is called Naive Bayes, but the algorithm learns quickly in controlled classification problems.

3.4.2. K-Nearest Neighbors

The k-Nearest Neighbors (*k-NN*) method is a supervised machine learning model that can be used for various classification and regression problems [25]. Here k is an integer value that represents the number of classes. The *k-NN*

method is used to predict which class a dataset belongs to by looking at other datasets around it [26].The *k-NN* algorithm performs the training process by calculating the similarities of the closest k-data within the framework of a certain distance criterion. One of the Minkowski, Euclidean, Chebyshev and Cosine equations can be used for the distance measures[27]. Among them, the Euclidean distance is the most frequently used metric in the literature, and it is also preferred for this study.

3.4.3. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) algorithm is a linear classifier used to separate samples belonging to two or more classes [28]. This method tries to find linear combination of features. In other words, LDA tries to find the vectors belonging to the space that can best distinguish the classes from each other. While it distances data points from different classes from each other, it brings data points from the same class closer together, thus producing a new variable that is the result of existing data. It aims to maximize the differences between the first defined classes according to the new variable.

3.4.4. Decision Tree

The decision tree (DT) method is a very popular and practical approach that is used for pattern classification. A decision tree is a flowchart that looks like a tree, with each internal node representing an attribute test, each branch representing a test result, and each leaf node carrying a class label [29]. To find feature threshold pairs that maximize the purity of the resulting two or more classes of data samples, different decision tree models use different approaches [29]. Some of them are ID3, C4.5, C5.0, CART, CHAID and QUEST algorithms.ID3 decision tree algorithm ischosen in this study for its common usage. In this algorithm, entropy and information gain calculations are used when deciding which feature the root will be in the tree structure and how the division will take place after the root. Entropy is the measurement of the uncertainty of a system that is calculated as in 3. Another value obtained based on the entropy value is the information gain value and is calculated as in 4[30].

$$H(S) = \sum_{i=1}^{n} -p(c)\log_2 p(c) \qquad (3)$$

$$IG(A,S) = H(S) - \sum_{t \in T} p(t)H(T) \quad (4)$$

3.4.5. Random Forest

Random Forest (RF) is a popular and efficient pattern recognition and machine learning algorithm that has shown to be one of the most successful ensemble learning techniques [31]. Random forest is a technique that combines many classification trees and gives high accuracy results. When a new sample is to be classified, the input vector of that sample is classified individually by each tree in the forest. This is called tree voting [32].The random forest algorithm provides good results in large data sets where the input variable is very large, as well as making good estimations in missing data.

3.4.6. Support Vector Machine

SVM is a supervised learning model used in machine learning. When doing classification, SVM creates one or more hyperplanes in the feature space. This hyperplane is used to classify data. If the hyperplane has the longest distance to the nearest data point of any class, then a good split is obtained. A larger margin means less generalization error. Therefore, an appropriate hyperplane should be chosen to reduce data point errors along the class border line. These data points are called support points or support vectors [33].

3.4.7. Multilayer Perceptron Neural Network

The most frequently used artificial neural networks are Multilayer Perceptron Neural Networks (MLP). There are three layers in MLP: input, hidden, and output. Data is sent from the input layer to the hidden layer. Information is received, processed, and then conveyed to the output layer in the hidden layer. The MLP neural network's output information is sent to the output layer. Poor generalization and overfitting problems are caused by an MLP neural network with numerous nodes in the hidden layer. As a result, determining the number of hidden nodes is usually performed through trial and error [34]. Figure 3 shows an MLP neural network structure.

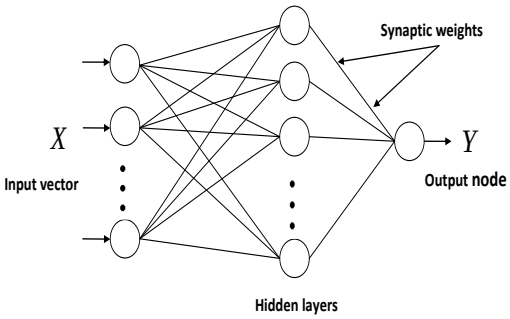


Figure 3.MLP neural network structure

MLP is a feed forward neural network model that converts data into a set of outputs. With a common method known as the error back propagation algorithm, MLP has efficiently handled certain challenging and diverse issues. There are two steps to the error back propagation technique. An input vector is applied to the network's sensory nodes during the first step, the forward pass, and its influence is propagated throughout the entire layer. As a result, the network's real answer is a series of multiples. At this point, all synaptic weights are constant. All synaptic weights are updated according to an error correction rule during the second stage, the backward transition. By subtracting the network's actual response from the target response, an error signal is generated. "Error back propagation" refers to the fact that the error signal propagates backwards through the network in the opposite direction of the synaptic weights. Synaptic weights are modified so that the network's actual response approaches the desired response [35].

4. Results and Discussion

In this study, the image data were preprocessed using the MATLAB R2017b platform before being given to the classification algorithms (see sections 3.2 and 3.3). The preprocessed images were later fed into each classification algorithm that was implemented in Python on the Jupyter Notebook platform, which supports many libraries such as Sklearn, Numpy, and Matplotlib.

In this manuscript, a comparative study is presented using Naive Bayes, k-NN, LDA, Decision Tree, Random Forest, SVM and MLP algorithms, performed on a personal computer with Intel Core i5 configuration with 8 GB RAM. A set of different parameters were used in feature selection and classification stages. Cross-Validation technique is used to estimate how each model performs out of sample to a new dataset, which is also defined as test data. For all classification algorithms, 10-cross validation technique was used and 80% of total 6500 images are allocated to the training set and 20% to the test dataset to train the model. As given in Table 1, in the PCA method, the optimal number of principal components was determined as 3. GaussianNB model is used in the Naive Bayes algorithm. In the k-NN algorithm, the number of neighbors is determined as k=2. In the decision tree algorithm, the maximum tree depth is taken as 6. In the random forest algorithm, the number of trees was determined as 100. The kernel feature used in SVM is linear, the C value is determined as 1, and gamma value is determined by its default value which uses $1/n\_features$ as the value of gamma. The success rates of the algorithms vary according to the selected parameter values. Therefore, different parameters were tested and it was determined which parameter provided the maximum success.

Table 1. Optimal parameters of the methods.

Methods	***Optimal Parameters***
NB	GaussianNB, Alpha=0.3
KNN	Number of neighbors=2, Metric= ”euclidean”, Alpha=0.3
LDA	Alpha=0.3
RF	Number of estimators=100, Alpha=0.3
DT	Criterion=”entropy”, Maximum_depth= 6, Alpha=0.3
SVM	C=1, Gamma=”auto”, Kernel=”linear”, Alpha=0.3
MLP	Hidden_layer_sizes=50, Maximum_iteration=2000, Activation=”relu”, Solver=”adam”, Random_state=0, Alpha=0.3

Figure 4. Confusion matrix values of MLP neural network classifier for the node numbers of hidden layers 10, 20, 30, 40 and 50, respectively.

In the MLP neural network method, to be able to reveal the effect of the number of hidden nodes in the hidden layer, the algorithm was run with 10, 20, 30, 40, and 50 nodes. The results of the network complexity were shown in complexity matrix in Figure 4. Since there is a slight difference between the sizes of the nodes in the hidden layer, a simple network structure is sufficient which reduces the network complexity for the classification problem. Otherwise, increasing the node size in the hidden layer results in increasing the cost and complexity at the same time.

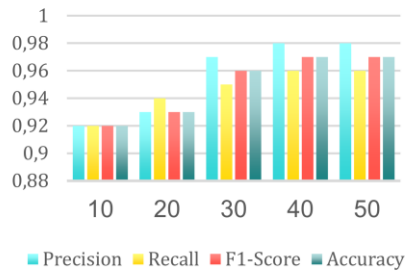


Figure 5. Performance analysis of MLP neural network classifier on different size of nodes in hidden layer.

Figure 5 also shows the comparative accuracy metrics of the number of hidden nodes in terms of precision, recall, accuracy, and F1-score. These metric values were calculated from the actual and predicted results, which are given in confusion matrix in Figure 4.

Similar to the given results in Figure 5, to determine the performance levels of all the above-mentioned algorithms, similar variables were calculated for each of them. The algorithms were run with the optimal parameters that given in Table 1.

Finally, in Figure 6, the performance metrics of the algorithms in terms of precision, recall, F1-score, and accuracy metrics were compared.

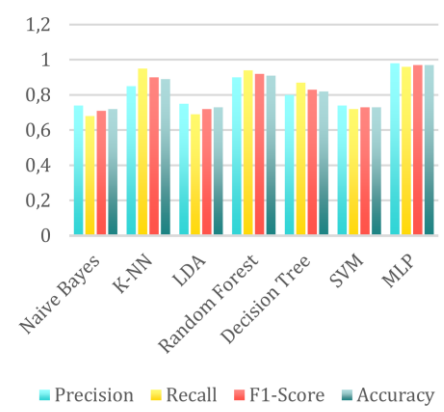


Figure 6. Performance analysis of classifiers.

Table 2. Performance metrics for all methods.

Methods	**Precision**	**Recall**	**F1**	**Accuracy**
NB	*0.74*	*0.68*	*0.71*	*0.72*
KNN	*0.85*	*0.95*	*0.90*	*0.89*
LDA	*0.75*	*0.69*	*0.72*	*0.73*
RF	*0.90*	*0.94*	*0.92*	*0.91*
DT	*0.80*	*0.87*	*0.83*	*0.82*
SVM	*0.74*	*0.72*	*0.73*	*0.73*
MLP	*0.98*	*0.96*	*0.97*	*0.97*

The comparative results are also given in Table 2. When the algorithms were evaluated according to the accuracy, the most effective algorithm is seen as MLP with an accuracy rate of 97% and 97% F1-score. Random Forest also has a higher classification rate with 91% accuracy and 92% F1-score among the other methods. k-NN is another method which can be preferred for classification with a high success rate.

In terms of precision values, which measure how accurate predictions were obtained from the existing classes, Random forest achieved the best success rate of 90%. The other metric, recall, expresses how accurately it was predicted from all positive classes and according to this metric, Random Forest shows the most success rate with %94.

For the classification of ALL problem, a broad perspective comparison, including the results of other applied methods in the literature and the results of studies on different data sets[1], is presented in Table 3. In all cases, it has been observed that the selected classifiers can classify with over 80% accuracy. The overall results indicate that independent of the dataset, MLP is observed as a method that results with high accuracy.

5. Conclusion

ALL is a very common type of leukemia, especially in children. As in other types of cancer, early diagnosis is very crucial in the treatment of the disease. ALL can be distinguished from healthy ones by morphological changes in blood cells, but the manual-visual search may cause

Table 3. Comparative results with different classifiers on different datasets.

Studies	**Classifiers**	**Datasets**	**Accuracy (%)**
Mandal et al. [5]	DT	ALL-CDB	85.20
Ahmed et. al. [6]	CNN	ALL-IDB, ASH Image Bank	85.25, 81.74
Bhuiyan et. al. [7]	PCA-SVM	ALL-IDB	99.05
Khosrosereshki & Menhaj [8]	Fuzzy Clustering	Set of 32 blood smears	93.75
Rehman et. al. [9]	CNN	ALL-IDB	97.78
Kumar et. al. [10]	KNN	60 pretested samples	92.80
Parvaresh et. al. [11]	Chain Tabu-MLP	ALL-IDB2	98.88
Umammaheswari& Geetha[12]	KNN	ALL-IDB2	95.96
Wahhab [13]	SVM	UMMC	96.93
Rahpurohit et al. [14]	CNN	ALL-IDB	99.50
Implemented Methods	NB, KNN, LDA,RF, DT, SVM, MLP	ALL-CDB	72.00, 89.00 73.00, 91.00 82.00, 73.00 97.00

[1] ALL-CDB : Acute Lymphoblastic Leukemia Challenge Database (the newest one)
ALL-IDB : Acute Lymphoblastic Leukemia Image Database (with versions IDB and IDB2)
ASH : American Society of Hemotology Image Bank
UMMC : University of Malaya Medical Center Dataset

various human-induced errors and delays in its treatment. On the other hand, this challenging point can be exceeded by the machine learning algorithms with high success. However, there are various types of machine learning methodologies and it is also another question that which algorithm is more suitable for the problem.

In this study, we addressed the challenging points of distinguishing the classification problem of cancerous cells from the healthy ones on ALL diseases. After that, to answer the abovementioned question, we proposed to implement the 7 well-known machine learning methodologies by using a commonly used dataset. The ALL dataset consists of 6500 digital microscopic pathology images from 118 subjects and each image is labeled as healthy or cancerous. Before running the algorithms, a feature extraction step has been performed by applying image processing algorithms to the data set. The PCA method was applied for feature selection on the data set whose features were extracted. Thus, the size of the data set has been reduced and features that do not contribute to the performance of classification algorithms are removed. Finally, the acquired features of the dataset have been fed into the 7-different classification algorithms and classified as cancerous and non-cancerous. When the performance values of the classification algorithms were compared, the most successful was MLP neural network with 97% accuracy. The results show us that computer-aided systems can be used in the field of pathology.
Here, it is also worth mentioning the limitations of machine learning studies. As we know, image processing and machine learning-based systems have begun to play a key role in medicine in recent years. However, it should keep in mind that, these kinds of methods may not provide a certain diagnosis but can assist medical specialists in making appropriate decisions.

References

[1] https://www.cancer.org/acutelymphocytic-lukemia/about/key-statistics.html (access:10.12.2021).

[2] "PS80 FactsBook_2020_2021_FINAL.pdf". https://www.lls.org/sites/default/files/2021-08/PS80%20FactsBook_2020_2021_FINAL.pdf (access:10.12.2021).

[3] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, ve A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries", *CA: a cancer journal for clinicians*, c. 68, v 6, ss. 394-424, 2018.

[4] Y. Dong *vd.*, "Leukemia incidence trends at the global, regional, and national level between 1990 and 2017", *Exp Hematol Oncol*, c. 9, sy 1, s. 14, Ara. 2020, doi: 10.1186/s40164-020-00170-6.

[5] S. Mandal, V. Daivajna, ve R. V., "Machine Learning based System for Automatic Detection of Leukemia Cancer Cell", içinde *2019 IEEE 16th India Council International Conference (INDICON)*, Rajkot, India, Ara. 2019, ss. 1-4. doi: 10.1109/INDICON47234.2019.9029034.

[6] Ahmed, Yigit, Isik, ve Alpkocak, "Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network", *Diagnostics*, c. 9, sy 3, s. 104, Ağu. 2019, doi: 10.3390/diagnostics9030104.

[7] Md. N. Q. Bhuiyan, S. K. Rahut, R. A. Tanvir, ve S. Ripon, "Automatic Acute Lymphoblastic Leukemia Detection and Comparative Analysis from Images", içinde *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, Paris, France, Nis. 2019, ss. 1144-1149. doi: 10.1109/CoDIT.2019.8820299.

[8] M. A. Khosrosereshki ve M. B. Menhaj, "A fuzzy based classifier for diagnosis of acute lymphoblastic leukemia using blood smear image processing", içinde *2017 5th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, Qazvin, Iran, Mar. 2017, ss. 13-18. doi: 10.1109/CFIS.2017.8003589.

[9] Maria, Italia Joseph; DEVI, T.; RAVI, D. Machine learning algorithms for diagnosis of leukemia. Int J Sci Technol Res, 2020, 9.1.

[10] S. Kumar, et al. Automated detection of acute leukemia using k-mean clustering algorithm. In: Advances in computer and computational sciences. Springer, Singapore, 2018. p. 655-670.

[11] H. Parvaresh, H. Sajedi, ve S. A. Rahimi, "Leukemia Diagnosis Using Image Processing and Computational Intelligence", içinde *2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES)*, Las Palmas de Gran Canaria, Haz. 2018, ss. 000305-000310. doi: 10.1109/INES.2018.8523900.

[12] D. Umamaheswari ve S. Geetha, "Segmentation and Classification of Acute Lymphoblastic Leukemia Cells Tooled with Digital Image Processing and ML Techniques", içinde *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, Haz. 2018, ss. 1336-1341. doi: 10.1109/ICCONS.2018.8662950.

[13] Wahhab, Hayan Tareq Abdul. Classification of acute leukemia using image processing and

machine learning techniques. 2015. PhD Thesis. University of Malaya.

[14] S. Rajpurohit, S. Patil, N. Choudhary, S. Gavasane, ve P. Kosamkar, "Identification of Acute Lymphoblastic Leukemia in Microscopic Blood Image Using Image Processing and Machine Learning Algorithms", içinde *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, Eyl. 2018, ss. 2359-2363. doi: 10.1109/ICACCI.2018.8554576.

[15] A. M. Abdeldaim, A. T. Sahlol, M. Elhoseny, ve A. E. Hassanien, "Computer-Aided Acute Lymphoblastic Leukemia Diagnosis System Based on Image Analysis", içinde *Advances in Soft Computing and Machine Learning in Image Processing*, c. 730, A. E. Hassanien ve D. A. Oliva, Ed. Cham: Springer International Publishing, 2018, ss. 131-147. doi: 10.1007/978-3-319-63754-9_7.

[16] S. Shafique ve S. Tehsin, "Acute Lymphoblastic Leukemia Detection and Classification of Its Subtypes Using Pretrained Deep Convolutional Neural Networks", *Technol Cancer Res Treat*, c. 17, s. 153303381880278, Oca. 2018, doi: 10.1177/1533033818802789.

[17] S. Mourya, S. Kant, P. Kumar, A. Gupta, ve R. Gupta, "ALL Challenge dataset of ISBI 2019". The Cancer Imaging Archive, 2019. doi: 10.7937/TCIA.2019.DC64I46R.

[18] F.Çam,A.Güven, "Methods Used In The Classification of White Blood Cells from Blood Cell Images Taken under a Digital Microscope", s.21, 2019.

[19] Z. F. Mohammed ve A. A. Abdulla, "Thresholding-based White Blood Cells Segmentation from Microscopic Blood Images", *UHD J SCI TECH*, c. 4, pp1,s.9,Şub.2020.doi: 10.21928/uhdjst.v4n1y2020.pp9-17.

[20] M. MoradiAmin, A. Memari, N. Samadzadehaghdam, S. Kermani, ve A. Talebi, "Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis", *Microscopy Research and Technique*, c. 79, sy 10, ss. 908-916, 2016, doi: https://doi.org/10.1002/jemt.22718.

[21] K. Yildiz, A. Çamurcu, ve B. Dogan, *A Comperative Analize of Principal Component Analysis and Non-Negative Matrix Factorization Techniques in Data Mining*. 2010.

[22] A. Jamal, A. Handayani, A. Septiandri, E. Ripmiatin, ve Y. Effendi, "Dimensionality Reduction using PCA and K-Means Clustering for Breast Cancer Prediction", *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, s. 192, Ara. 2018, doi: 10.24843/LKJITI.2018.v09.i03.p08.

[23] A. B. Varol ve İ. İşeri, "Lenf Kanserine İlişkin Patoloji Görüntülerinin Makine Öğrenimi Yöntemleri ile Sınıflandırılması", *European Journal of Science and Technology*, ss. 404-410, Eki. 2019, doi: 10.31590/ejosat.638372.

[24] M. M. Saritas, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification", *ijisae*, c. 7, sy 2, ss. 88-91, Oca. 2019, doi: 10.18201/ijisae.2019252786.

[25] J. Gupta, "The Accuracy of Supervised Machine Learning Algorithms in Predicting Cardiovascular Disease", içinde *2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST)*, Yogyakarta, Indonesia, Haz. 2021, ss. 234-239. doi: 10.1109/ICAICST53116.2021.9497837.

[26] S. Sharma, A. Aggarwal, ve T. Choudhury, "Breast Cancer Detection Using Machine Learning Algorithms", içinde *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, Ara. 2018, ss. 114-118. doi: 10.1109/CTEMS.2018.8769187.

[27] M. A. Pala, M. E. Çimen, Ö. F. Boyraz, M. Z. Yildiz, ve A. F. Boz, "Meme Kanserinin Teşhis Edilmesinde Karar Ağacı Ve KNN Algoritmalarının Karşılaştırmalı Başarım Analizi", *acperpro*, c. 2, sy 3, ss. 544-552, Kas. 2019.

[28] C. Oral, A. Aydın Yurdusev, ve E. Bergil, "Mamogramların Sınıflandırılmasında Dokusal Özelliklerin Etkileri", *DÜMF Mühendislik Dergisi*, c. 10, sy 1, ss. 23-33, Mar. 2019, doi: 10.24012/dumf.403657.

[29] A. M. Elsayad, H.A. Elsalamony, Diagnosis of breast cancer using decision tree models and SVM. International Journal of Computer Applications, 2013, 83.5.

[30] H. Güldal, Karar ağacı algoritmalarının eğitsel veriler üzerindeki performanslarının incelenmesi The analysing of desicion alghoritms' performance on educational data. In: 13th International Balkan Education and Science Congress. 2018. p. 6.

[31] J. Thongkam, G. Xu, ve Y. Zhang, "AdaBoost algorithm with random forests for predicting breast cancer survivability", içinde *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Hong Kong, China, Haz. 2008, ss. 3062-3069. doi: 10.1109/IJCNN.2008.4634231.

[32] B. Özlüer Başer, M. Yangin, ve E. S. Saridaş, "Makine Öğrenmesi Teknikleriyle Diyabet Hastalığının Sınıflandırılması", *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, Şub. 2021, doi: 10.19113/sdufenbed.842460.

[33] S. Ghosh, S. Mondal, ve B. Ghosh, "A comparative study of breast cancer detection based on SVM and MLP BPN classifier", içinde *2014 First International Conference on Automation, Control, Energy and Systems (ACES)*, India, Şub. 2014, ss. 1-4. doi: 10.1109/ACES.2014.6808002.

[34] A. O. Ibrahim, S. M. Shamsuddin, A. yahya Saleh, A. Abdelmaboud, ve A. Ali, "Intelligent multi-

objective classifier for breast cancer diagnosis based on multilayer perceptron neural network and Differential Evolution", içinde *2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering (ICCNEEE)*, Khartoum, Sudan, Eyl. 2015, ss. 422-427. doi: 10.1109/ICCNEEE.2015.7381405.

[35] D. Soria, J. M. Garibaldi, E. Biganzoli, ve I. O. Ellis, "A Comparison of Three Different Methods for Classification of Breast Cancer Data", içinde *2008 Seventh International Conference on Machine Learning and Applications*, San Diego, CA, USA, 2008, ss. 619-624. doi: 10.1109/ICMLA.2008.97.