Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230479

An Artificial Intelligence-Based Diagnostic System for Acute Lymphoblastic Leukemia Detection

Yousra EL ALAOUI[a,1], Regina PADMANABHAN[a], Adel ELOMRI[a], Marwa K. QARAQE[a], Halima EL OMRI[b], Ruba YASIN TAHA[b]
[a] *College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*
[b] *Medical Oncology-Hematology Department, National Centre for Cancer Care and Research (NCCCR), Hamad Medical Corporation (HMC), Doha, Qatar*

Abstract This study suggests a novel Acute Lymphoblastic Leukemia (ALL) diagnostic model, built solely on complete blood count (CBC) records. Using a dataset comprised of CBC records of 86 ALL and 86 control patients respectively, we identified the most ALL-specific parameters using a feature selection approach. Next, Grid Search-based hyperparameter tuning with a five-fold cross-validation scheme was adopted to build classifiers using Random Forest, XGBoost, and Decision Tree algorithms. A comparison between the performances of the three models demonstrates that Decision Tree classifier outperformed XGBoost and Random Forest algorithms in ALL detection using CBC-based records.

Keywords. ALL, Early detection, Machine learning, CBC

1. Introduction

ALL is among the deadliest and most prevalent types of acute leukemia. It starts in the bone marrow, particularly affecting children and elderly. If not diagnosed and treated early, ALL disease progression can lead to death in both adults and children [1]. In fact, it was proven that 90% of the cases are treatable if detected early [2]. Therefore, a need for an accurate diagnostic system to early spot the signs of ALL is crucial.

Currently, most ALL detection procedures are manual microscopic examination techniques that mainly rely on a physician's medical expertise to identify any abnormalities within the blood cells [1]. Although medical imaging methods have been successful in extracting prevalent ALL features from peripheral blood smear images in the initial screening phase, the overall diagnostic process is still considered time-consuming, tedious for the physician, and highly prone to human errors [2]. With the aim to timely diagnose ALL and minimize the practitioner's burden, intelligent diagnosis systems have been introduced to replace manual procedures. For instance, Escalante et al. [3] have developed an ensemble model based on particle swarm optimization for binary and multi-categorical classifications of acute leukemia cells including ALL cells

[1] Corresponding Author: Yousra El Alaoui, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar; Email: yoel32553@hbku.edu.qa.

in digitized bone marrow images. Similarly, Mohapatra et al. [4] suggested a leukemia diagnostic model using a fuzzy-segmentation technique and a Support Vector Machine (SVM) classifier resulting in an accuracy of 93%. Moreover, Israet et al. [5] employed Convolutional Neural Networks (CNNs) for effective blood image classification and leukocyte segmentation, which Roy et al. [6] also achieved by alternatively developing a system based on DeepLabv3 and ResNet-50 to extract deep feature maps from the rest of the blood images. While the power of automation and Artificial Intelligence (AI) was greatly exploited through different applications of Machine Learning (ML) and Deep Learning (DL) on blood and bone marrow images for abnormality identification, microscopic examination was still considered challenging and computationally expensive [7]. In majority of cases, the input blood cell raw images require intensive enhancement and preprocessing to achieve higher quality images for future classifications. Hence, our study aims to expedite and facilitate the ALL diagnosis process by making use of patients' basic electronic health information for an accurate and timely detection.

2. Methods

2.1. Study Design

A dataset including CBCs of 86 confirmed ALL and 86 control cases was collected from the National Center for Cancer Care and Research (NCCCR) at Hamad Medical Corporation (HMC) in Doha, Qatar. The data was retrieved as per the NCCCR-HMC regulations over the period extending from 2016 to 2021, and included records of ALL and healthy patients corresponding to values of 20 CBC parameters as follows: WBC, RBC, Hgb, Hct, MCV, MCH, MCHC, RDW-CV, Platelets, MPV, ANC, Lymphocyte count, Monocyte count, Eosinophil count, Basophil count, Lymphocytes %, Monocytes %, Eosinophils %, Basophils % and Neutrophils %.

Feature Selection: To select the most ALL-discriminatory features, we performed a Forward Feature Selection (FFS) approach on the 20 CBC parameters [8]. An iterative evaluation of all the possible CBC feature combinations used in predicting the target class, where 0 and 1 correspond to benign and ALL patients respectively, was performed using FFS. The final set of features included: ANC, Hct, RBCs, MCV, MCH, Neutrophils %, Basophils %, Lymphocyte count, MPV and Platelets.

2.2. ALL Classification Models

Using the selected features, we trained three ML models on our dataset to compare the different model performances in ALL classification (see Figure 1).

Data Training: After cleaning our data and filling out the few missing values using K-Nearest Neighbor imputation technique [9], we proceeded to perform an 80%-20% data split, resulting in a training set containing 137 patients, and a testing set consisting of 35 samples. Our dataset exhibited a balanced number of samples for each class and set (see Table 1).

Model Design: In this study, we trained and tested three ML models. (1) Random Forest Algorithm (RF): To find the best hyperparameters for our best RF estimator, we performed a Grid Search hyperparameter tuning technique [10]. The latter accounted for the following ranges: [10, 15] and [0, 14], for the maximum depth and the maximum

number of features, respectively. (2) XGBoost Algorithm: Similarly, a Grid Search hyperparameter tuning approach was applied on XGBoost, that searched for the best combination of hyperparameters in the following ranges: [1, 10], [40, 220] and [0.01, 0.1], corresponding to the maximum depth, number of estimators and the learning rate of the model, respectively. (3) Decision Tree (DT) Algorithm: A Grid Search was identically performed on the DT model to optimize the minimum number of samples required to be at a leaf node and the maximum depth of a tree, given a list of different combinations in the range corresponding to [1, 3] for both parameters.

Figure 1. The proposed study pipeline.

Table 1. Dataset statistics.

Dataset		172
Training Set	137	68 ALL 69 control
Testing Set	35	18 ALL 17 control

3. Results and Discussion

A 5-fold cross validation was performed on the suggested models to calculate the average accuracy, recall, precision and F1 score for each model. The three algorithms performed well on our dataset composed of patient electronic health records only, with DT achieving the highest accuracy corresponding to 91.4% (see Table 2).

Table 2. Performance evaluation of the three ML models for ALL classification.

Metric	Random Forest	XGBoost	Decision Tree
Accuracy	88.6%	88.6%	91.4%
Recall	82.3%	76.5%	88.2.%
Precision	93.3%	100.0%	93.7%
F1 Score	87.5%	86.7%	90.9%

Our results showcase the power of a simple CBC test in detecting the warning signs of ALL, using a selected combination of features. This finding constitutes a very big leap in the field of ALL classification, as it indicates the importance of correct reading and efficient utilization of basic patient health tests in optimizing diagnosis test outcomes. Similarly, several AI techniques were applied on clinical data and have yielded excellent results, for instance Chen et al. [11] predicted cancer survival and Kim et al. [12] developed a prediction model of breast cancer recurrence. Although the need for human expertise is indispensable in the field of hematology detection, our suggested system can be used as a support for physicians, namely in the triage process, to reduce the number of unnecessary referrals to specialized hospitals, leaving the room for actual ALL suspected cases in need of further investigations. Thus, the impact of this study lies in providing an additional confirmation layer for the physician that can contribute to early ALL detection and timely treatment.

4. Conclusion

Our study consisted of applying three ML models on 172 patients for ALL classification. The suggested models such as RF, DT and XGBoost, were trained using the scikit-learn library and evaluated in a five-fold cross-validation scheme in terms of accuracy, recall, precision and F1 score. Our final results demonstrated the strength of a basic test like CBC in accurately classifying ALL patients, where DT outperformed the rest of the ML methods. Although these findings are considered novel and contribute to optimizing the overall ALL diagnosis process, an improved study version is still needed to incorporate a larger dataset and the diversity in patient gender, nationality and patient health status at the time of diagnosis, that comes with it.

Acknowledgments

This article was made possible by the National Priorities Research Program-Standard (NPRP-S) Twelfth (12th) Cycle grant (NPRP12S-0219-190108) from the Qatar National Research Fund (a member of Qatar Foundation). The findings herein reflect the work, and are solely the responsibility, of the author[s].

References

[1] Abunadi I, Senan EM. Multi-method diagnosis of blood microscopic sample for early detection of acute lymphoblastic leukemia based on Deep Learning and hybrid techniques. Sensors. 2022;22(4):1629.
[2] Rezayi S, Mohammadzadeh N, Bouraghi H, Saeedi S, Mohammadpour A. Timely diagnosis of acute lymphoblastic leukemia using artificial intelligence-oriented deep learning methods. Computational Intelligence and Neuroscience. 2021;2021:1–12.
[3] Escalante HJ, Montes-y-Gómez M, González JA, Gómez-Gil P, Altamirano L, Reyes CA, et al. Acute leukemia classification by Ensemble Particle Swarm Model Selection. Artificial Intelligence in Medicine. 2012;55(3):163–75.
[4] Mohapatra S, Samanta SS, Patra D, Satpathi S. Fuzzy based blood image segmentation for automated leukemia detection. 2011 International Conference on Devices and Communications (ICDeCom). 2011
[5] Isra N, Nazeer M, Mussarat Y, Muhammad S, Jamalhussain S, Stevenlawrence F. Robust discrimination of leukocytes protuberant types for early diagnosis of leukemia. Journal of Mechanics in Medicine and Biology. 2019;19(06):1950055.
[6] Roy RM, Ameer PM. Segmentation of leukocyte by Semantic Segmentation Model: A deep learning approach. Biomedical Signal Processing and Control. 2021;65:102385.
[7] El Alaoui Y, Elomri A, Qaraqe M, Padmanabhan R, Yasin Taha R, El Omri H, et al. A review of Artificial Intelligence Applications in hematology management: Current practices and future prospects. Journal of Medical Internet Research. 2022;24(7).
[8] Kumar V. Feature selection: A literature review. The Smart Computing Review. 2014;4(3).
[9] Huang SF, Cheng CH. (2020) "A safe-region imputation method for handling medical data with missing values," Symmetry. 2020;12(11):1792. Available at: https://doi.org/10.3390/sym12111792.
[10] Shekar BH, Dagnew G. Grid search-based hyperparameter tuning and classification of Microarray Cancer Data. 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP). 2019;
[11] Chen Y-C, Ke W-C, Chiu H-W. Risk classification of cancer survival using ANN with gene expression data from multiple laboratories. Computers in Biology and Medicine. 2014;48:1–7.
[12] Kim W, Kim KS, Lee JE, Noh D-Y, Kim S-W, Jung YS, et al. Development of novel breast cancer recurrence prediction model using support vector machine. Journal of Breast Cancer. 2012;15(2):230.