Al-Rafidain Engineering Journal (AREJ) Vol. 30, No.1, March 2025, pp. 175-185

Comparison of Deep Learning and Machine Learning Techniques for Automated Diagnosis of Acute Lymphoblastic Leukemia

Zeina Mahmoud
zeina.mahmoud2022@stu.uoninevah.edu.iq

Mohammed A. M. Abdullah
mohammmed.abdulmuttaleb@uoninevah.edu.iq

Computer and Information Engineering Department, College of Electronics Engineering, Nineveh University, Mosul, Iraq

Received: August 2nd, 2024 Received in revised form: October 27th, 2024 Accepted: November 28th, 2024

ABSTRACT

Acute lymphoblastic leukemia (ALL) presents major diagnostic and categorization challenges due to its wide range of clinical manifestations and the invasive nature of existing diagnostic procedures. In this work, we examine how Deep Learning (DL) and Machine Learning (ML) approaches can be used to enhance ALL diagnosis and classification using bone marrow images. We do an extensive investigation of the performance of various DL and ML models, such as Convolutional Neural Network (CNN), Feed-forward neural networks (FNN), Naive Bayes (NB), and Decision Trees (DT), using the ALL-IDB1 dataset. A developed feature extraction method based on VGG16 is proposed with feature selection based on recursive feature elimination. Our study includes fine-tuning pre-trained models, feature extraction with VGG16, and model optimization. F1 score, accuracy, and recall measures are used to assess the performance of the model. The investigation produced encouraging results, with both DL and ML models recording 100\% accuracy and excellent classification scores. Additionally, it is confirmed that automated systems based on DL and ML models have the potential to improve patient outcomes by speeding up and improving diagnosis accuracy.

Keywords:

Acute Lymphoblastic Leukemia; Deep Learning; Machine Learning; Feature Extraction; Convolutional Neural Network.

https://rengj.mosuljournals.com
Email: alrafidain_engjournal3@uomosul.edu.iq
==========

1. INTRODUCTION

Acute lymphoblastic leukemia (ALL) is a broad class of lymphoid tumors originating from B and T cell progenitors. These neoplasms can be presented as a widespread leukemic process, including the peripheral blood and bone marrow, or as localized tissue invasion known as lymphoblastic lymphoma (LBL) [1]. Even though ALL and LBLs have a variety of clinical presentations, they are thought to form a continuous biological spectrum, and they are classified as B- or T-LBL by the World Health Organization [2]. Even though the majority of instances of ALL are in children, adult cases might develop differently and have different prognoses because of advancements in treatment, including new medications [2]. ALL is frequently identified through the analysis of bone marrow samples, as it comes from acute myeloid leukemia, which is crucial for efficient treatment planning. Additional diagnostic procedures are added to bone marrow testing procedures such as peripheral blood smear analysis and flow cytometric immunophenotyping [3]. However, bone marrow biopsy and aspiration continue to be the benchmark for determining ALL diagnoses because they enable a full evaluation of cellular shape and structure, which can assist in evaluating prognosis and disease progression [4].

Generally, hospitals use manual methods to detect ALL, and they use immunohistochemistry labeling to count the leukocyte cells on peripheral blood smears [5]. Because staining is done manually by lab personnel, this procedure is labor-intensive and prone to mistakes [6]. Therefore, using computer-aided detection (CAD) systems can effectively diagnose ALL and minimize the toxicity levels of cancer patients and can assist the doctor in proper disease diagnosis and treatment, aiming to save valuable lives [7].

Recently, breakthroughs in artificial intelligence, ML, and CNNs, have revolutionized medical image processing, including the interpretation of hematological data. While DL techniques have shown potential in analyzing peripheral blood smears, their applicability in interpreting bone marrow aspirates and biopsies, the gold standard for ALL diagnosis, is still restricted [8].

This work investigates employing DL and ML approaches to improve the clinical identification and classification of ALL using bone marrow pictures. In addition, a developed feature extraction method based on VGG16 is proposed along with feature selection utilizing the recursive feature elimination method. Given the comprehensive nature of the ALL-IDB1 dataset, our investigation covers the breadth of ALL diagnostic subjects. Multiple DL and ML models are thoroughly analyzed, including CNN, FNN, NB, and DT. The paper is structured as follows: The literature review is presented in Section 2. Details of the suggested methodology, including picture pre-processing, feature extraction, and classification, are covered in Section 3. The results and discussion of this article are presented in Section 4. Section 5 contains the paper's conclusion and future scope.

2. RELATED WORK

Several works are available in the litereture for ALL classification. As a case in point, Researchers in [9] used CNN and FNN with the ALL-IDB2 dataset, which consists of 260 images, divided into 156 images as training data and 104 images as test data. The achieved accuracy rates from 98.33% to 95.40% for CNN and FNN, respectively. CNN outperformed FNN and some ML algorithms. However, the research was not able to apply CNN to ALL-IDB1 data, while the rest of the algorithms were applied to ALL-IDB1 data.

The authors in [10] used traditional image processing techniques to separate lymphocytes from white blood cells and then extracted the shape and color features and fed them into a Support Vector Machine (SVM) classifier to achieve 93.70% accuracy in detecting leukemia on the ALL-IDB1 dataset. Forty images were used as training data and the remaining 68 were used as testing data. However, the paper relied on a single algorithm.

Researchers in [11] used the Otsu threshold method to identify white blood cells and then used geometric feature extraction for classification using SVM on the ALL-IDB1 dataset consisting of 108 images, achieving an accuracy of 98.40%. The method was based on a single algorithm and relied on image processing techniques.

The work in [12] proposed an automated method for detecting ALL, where they used feature detection by smart edge detection and directed gradient mapping, then used feature reduction by principal component analysis, and ML algorithms on the ALL-IDB1 dataset, which was divided into 5 folds, where 4 folds were used as training data and 1 fold was used as test data. Various ML algorithms achieved high accuracy, which contributed to the progress of medical technology. However, the study focused on specific algorithms and the lack of discussion on preprocessing may be limitations.

In [13], the researchers used ML and CNN algorithms on two datasets namely: ALL-IDB and ASH, each dataset consisting of 354 and 549 images. After applying the data augmentation technique, the number of images in each dataset increases to 2478 and 3843, respectively, using 5-fold cross-validation (CV). The highest accuracy of 88.25% was achieved using the CNN model. Their method provided a comprehensive diagnosis. However, its reliance on a large training dataset and focus on well-known ML algorithms may limit its wider applicability.

In [14], the researchers highlighted the classification of leukemia using the ALL-IDB1dataset. After extracting geometric features and feeding them into ELM and SVM classifiers using 5-fold CV, they obtained an accuracy of 92.24% and 86.36%, respectively. The method's reliance on specific algorithms and the absence of a more comprehensive description of ML techniques may be drawbacks, although it provided insight into early-stage diagnosis and the use of public datasets.

Researchers in [15] presented a CNN model for detecting ALL. Using the ALL-IDB1 and ALL-IDB2 datasets, data augmentation technology was used to increase the amount of training data, which effectively alleviated the overtraining problem. The number of data became 736. After the data augmentation technology, the model was trained on 515 images, achieving an accuracy of 95.54%. Although high accuracy was achieved, two data sets were used and the dimensions of the ALL-IDB1 data set were changed, which resulted in the loss of image information and may suffer from the problem of overfitting.

Also, in [16], the researchers used data processing and feature extraction of shape and color and then fed them into an SVM classifier to classify all types of blood cancer. A dataset of

520 images was used and divided into 390 images as training data and 130 images as test data, where they obtained an accuracy rate of 97.69%. The approach enabled the identification of blood disorders more accurately. However, the paper relied on a single algorithm.

The work in [6] proposed a DL approach for real-time detection of acute ALL cells, by combining both traditional CNN models and the You Only Look Once (YOLO) model using the ALL-IDB1 dataset where 15% of which were selected as test data. Both CNN and YOLOv5s achieved an accuracy of 99.22% and 97.2%, respectively. Their work highlights the need for AI techniques to accurately and efficiently identify leukemia cells from microscopic images. However, insufficient data caused overfitting problems.

In reference [17], researchers presented DeepLeukNet, a CNN-based model specifically designed for ALL classification. They used the ALL-IDB1 and ALL-IDB2 datasets and applied the data augmentation technique. The employed images from the ALL-IDB1 dataset were divided into 741 images as training data and 259 images as test data. Their method was remarkably accurate in identifying ALL, with a diagnosis rate of 99.61%. Compared to previous methods, their study offered advantages such as automated diagnosis, less reliance on manual observation, and increased efficiency. Despite achieving high accuracy in classification, it suffers from the problem of overfitting.

The research work in [18] aims to develop an approach for diagnosing ALL using deep learning and machine learning. They used feature extraction using DenseNet201 and feature selection using Random Forest genetic algorithm with binary ant colony optimization, and a set of classifiers. The proposed approach was evaluated on the C-NMC 2019 dataset consisting of 15,114 images and yielded an accuracy of 90.55% and sensitivity of 95.94%. In this study, data augmentation techniques that could have improved accuracy were not used.

In [19], an end-to-end system for automated diagnosis of ALL and Acute Myeloid Leukemia (AML) classification was proposed using a deep learning model based on graph theory and CNN. The ALL and AML dataset consisting of 670 images was used and divided into 70% training data, 20% validation data, and 10% test data. It was able to classify the samples with an accuracy of 99.85%. The results indicate that the model can be a reliable tool for clinicians in diagnosing leukemia. However, the system depends on high-resolution images which may consume long time in real-time processing due to complexity.

From the above-mentioned works, key limitations can be summarized as:

- Single Algorithm Dependency: Several studies relied on one algorithm (e.g., SVM), limiting method generalization.
- Real-Time Constraints: High-accuracy models may face difficulties in real-time processing due to complexity.
- Limited Applicability: Focus on specific algorithms and datasets reduced flexibility in broader medical applications.
- Overfitting Risks: training was performed on a limited number of images highlighting potential issues.

3. METHODOLOGY

3.1. Dataset

All experiments are performed on the ALL-IDB dataset, which has been obtained through permission from the Department of Information Technology - Universita degli Studi di Milano, Italy (https:/scotti.di.unimi.it/all/). This dataset contains two parts, ALL-IDB1 and ALL-IDB2. In this work, we focus on ALL-IDB1, which is considered more complex than ALL-IDB2 because it includes all blood components while ALL-IDB2 includes only lymphocytes. ALL-IDB1 contains 108 images in total. 49 out of 108 images are of all patients (Blast cell), and the remaining 59 are of healthy individuals (Normal cell) as shown in Fig. 1. It comprises approximately 39,000 blood components, and lymphocytes are identified by skilled oncologists. In order to increase the number of images, data augmentation is used. Table 1 gives the comparison between the two datasets ALL-IDB1 and ALL-IDB2 [11].

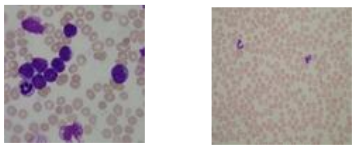

Fig. 1 Blast cell (left) and Normal cell (right)

Table 1: Comparison between ALL-IDB1 and ALL-IDB2.

Comparison in terms of	ALL-IDB1	ALL-IDB2
Dataset components	not segmented	segmented
Images	108	260
Resolution	2592×1944	257×257
Elements	39000	260
Lymphoblasts	510 cells	130 cells

3.2. Overall model block diagram

The methodological strategy described in this article attempts to improve the diagnosis and classification of ALL by combining DL and ML techniques. The proposed methodology employed CNN models for leukemia classification using innovative data reduction techniques. The proposed design is shown in Fig. 2 which consists of several basic elements, including data entry, feature extraction, feature selection, and algorithm blocks.

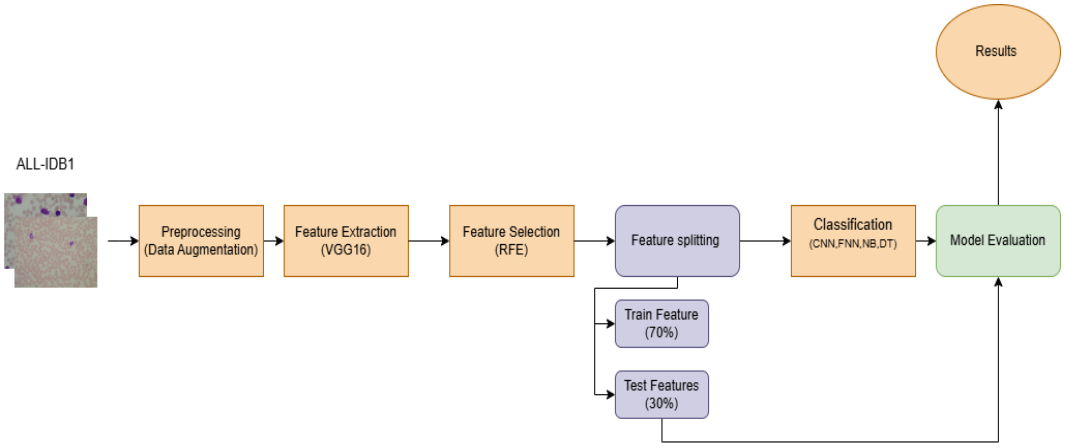

Fig. 2 Block diagram of the proposed scheme of ML with feature reduction based on deep learning.

3.2.1. Data processing

In this study, utilizing a data augmentation method, one can raise the amount of input data. The importance of the data augmentation technique focuses on minimizing overfitting and improving the generalization capabilities of deep learning networks, in addition to improving the data classification. Data augmentation is performed by applying different transformations to the original images. By producing several augmented samples for each original image. In the current study, several operations are used to increase the number of images such as rotation, shift, zoom, shear, and flip. Details of augmentation operations are given in Table 2 below. The augmented dataset is formed by combining the original and transformed images. The number of images after the data augmentation technique became 540 images.

Table 2: Details of augmentation operations.

Operation	Details
rotation	10o
width shift, height shift	10% of width of height
zoom, and shear	10% of image size
horizontal and vertical flip	true

3.2.2. Feature extraction

Feature extraction refers to the process of extracting useful and instructive features from raw data. Feature extraction is used in DL and computer vision in order to reduce complexity. After features are extracted, they can be fed into another machine learning model, like a regression or classifier model, to carry out certain tasks like object detection, image production, or image classification. In this paper, six CNN models were used to extract features, namely: ResNet152, VGG16, DenseNet121, MobileNetV2, InceptionV3, and EfficientNetB0. All of these models work well with ALL-IDB1, but the VGG16 model is chosen for feature extraction among the aforementioned models because it achieves high accuracy and less time on the ALL-IDB1 datasets. Using feature extraction, the number of features was reduced to 512.

3.2.3. Feature selection

Feature selection, as a dimensionality reduction method, seeks to remove unnecessary, redundant, or noisy features from the original features in order to choose a small subset of the relevant features. Generally speaking, feature selection can improve the interpretability of the model, reduce computing costs, increase learning accuracy, and improve learning performance.

In this work, we employed Recursive Feature Elimination (RFE) for feature selection. RFE is one of the ML feature selection methods that helps in finding the most important features within a dataset. The main concept of RFE is to iteratively remove the least important features to identify the most important features. This improves the model's performance by reducing the dimensions of the input data. After using RFE, 256 features were selected from 512 features.

3.2.4. Classification

After data processing, extraction, and selection, data is fed into algorithm blocks such as CNN, FNN, DT, and NB. These models are briefly described in the following subsection, focusing on their distinctive design and benefits in image classification tasks.

3.3. Employed models

Gaussian Naive Bayes (GNB): An NB classifier relies on the application of the Bayes theorem and is a straightforward probabilistic classifier. Every attribute variable is seen by NB as an independent variable. This classifier may be used in challenging real-world scenarios and trained quite well in supervised learning environments [20].

Decision Trees (DTs): DTs are well-known ML models. DTs are a basic yet effective tree-based model for classification and regression applications. They recursively divide the feature space into zones based on feature thresholds, To minimize impurity (such as Gini impurity or entropy) [21].

CNN: falls within a category of deep neural networks created especially for processing pictures and structured grid data. It is essential to gain knowledge about the CNN architecture and its different components and how they are being utilized to know about the advancements in the CNN structure. It consists of convolutional, Pooling, and Fully Connected layers. The convolutional layer is a fundamental module that contributes to the architecture of CNNs. This is comprised of many filters, which can otherwise be referred to as kernels that are applied to the input data. Each kernel is used to extract features from the input data and has a unique width, height, and weight, and CNNs employ the pooling layer, to minimize the dimensionality of feature maps while maintaining crucial information while the fully connected layer, groups of neurons are arranged to match those in traditional neural networks. Every node in a completely linked layer is correspondingly directly connected to every other node in the layer above and below [22].

FNN: An FNN is a type of artificial neural network in which information only moves via any hidden layers from the input layer to the output layer and back again. It is the most popular and straightforward kind of neural network architecture [23].

3.4. Proposed method

Our proposed method includes the DL part and ML part and is explained below.

3.4.1. Deep learning

The DL component of our methodology uses CNN and FNN models. The proposed CNN model, which is shown in Fig. 3, consists of two max-pooling layers, two convolutional layers, and flattened, dense, and dropout layers. Where the number of filters for convolution layers is (64,128) in sequence. The first Dense layer has 256 neurons where the dropout rate is set to 0.5. The output layer is the second Dense layer, which has 1 neuron and represents binary classification. The number of the parameters of the proposed model is (58,049) with a model size of (226.75 KB).

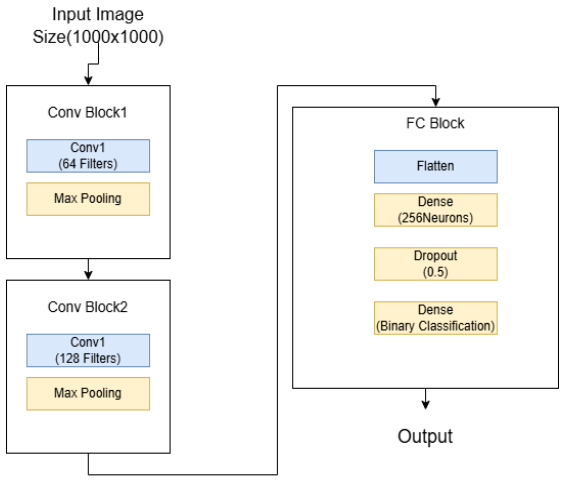

Fig. 3 The architecture of the proposed CNN.

The proposed FNN model, which is shown in Fig. 4, consists of three dense layers where the number of neurons for these layers is (64,128,256) in sequence, and consists of dropout where the dropout rate is set to 0.5. The output layer is the final Dense layer, which has 1 neuron that represents binary classification. The model that has been suggested has 58,049 parameters

and a 226.75 KB model size. When choosing the best method for a particular application, it is essential to take the dataset's features and unique requirements into account. To optimize the model for the particular difficulties presented by the available data, hyperparameter adjustments are required. Table 3 lists the best hyperparameters for the methods used.

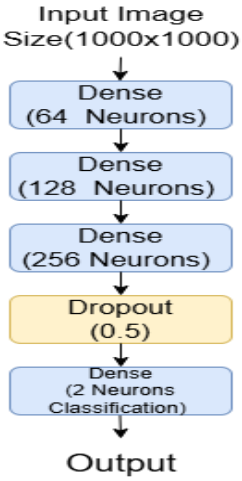

Fig. 4 The architecture of the proposed FNN scheme.

Table 3: Hyperparameters of different algorithms.

Algorithm	Hyperparameters
CNN and FNN	Batch size =64, Epoch=150 CNN consists 4 convolutional and max pooling layers and 2 dense and dropout layers.
Naive Bayes	Was selected NB model is Gaussian NB. Through which the best accuracy was obtained.
DT	The best accuracy was obtained by setting the max depth to 5 and min sample split to 2

3.4.2. Machine learning

The ML component of our methodology uses NB and DT models. These models are fine-tuned using the training subset, which includes tuning their parameters to match the specific properties of ALL-IDB1. The grid search technique was used for the ML algorithms to choose the best parameters that achieve the best accuracy. As for the NB algorithm, it showed improved accuracy using the GNB model. DT also resulted in the highest level of accuracy when setting the maximum depth to 5 and the minimum number of samples (minimum samples split) to 2.

4. RESULTS AND DISCUSSION

4.1. Experimental results

In order to evaluate the proposed models, different performance measures are used: precision, recall, F1 score, and accuracy as shown in Eqs. (1), (2), (3), and (4), respectively. The variables used in the equations are TP, FN, FP, and TN, Where TP stands for True Positive. TP is the number of ALL cells correctly classified as All cells. FN stands for false negative (FN) and is the number of ALL cells classified as normal cells. FP, which means false positive (FP), is the number of normal cells classified as ALL cells. TN stands for True Negative (TN), which is the number of normal cells classified as normal cells [24].

$$Accuracy = \frac{Tp+Tn}{TP+TN+FP+FN} \quad \text{.....(1)}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{.....(2)}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{.....(3)}$$

$$F1 - Score = 2 \times \frac{percision \times recall}{Percison+recall} \quad \text{.....(4)}$$

In this work, the All-IDB1 dataset was divided into 70% as training data, and the remaining 30% was used for testing.

4.2. Performance without data augmentation

Table 4 shows the performance of different algorithms when data augmentation is not used, which include Naive Bayes (NB), Decision Tree (DT), Convolutional Neural Network (CNN), and Feed-forward Neural Network (FNN). The results of CNN and FNN reveal lower accuracy and suffer from overfitting problem when data augmentation is not used. Also, for NB and DT, the accuracy is lower when data augmentation is not used.

Table 4: Performance without data augmentation

	Without data augmentation		With data augmentation	
Algorithms	Accuracy	F1-Score	Accuracy	F1-Score
NB	93.2	92.3	98.7	98.5
DT	93.9	90.9	95.6	94.8
CNN	96.9	96	100	100
FNN	96.9	96	100	100

4.3. Performance using different models

Six CNN models were used for feature extraction, namely ResNet152, VGG16, DenseNet121, MobileNetV2, InceptionV3, and EfficientNetB0. All of these models perform well on ALL-IDB1, but VGG16 achieved high accuracy and less time on ALL-IDB1 datasets as shown in Table 5.

Table 5: Performance using different models.

Classification algorithm	Feature extraction Algorithms	Accuracy	F1-Score
Naive Bayes	VGG16	98.7	98.5
	Inceptionv3	90.1	89.0
	Mobilenetv2	97.5	97.1
	Densenet121	96.2	95.7
	Resnet152	85.8	82.9
	Efficienetb0	98.4	98.4

4.4. Performance with augmentation, feature extraction using VGG16, and features selection

In this work data augmentation and feature extraction techniques were used using the VGG16 model and feature selection which achieved good results.

Firstly, the proposed CNN model achieved an accuracy of 100%, precision of 100%, recall of 100%, and an F1-score of 100%. Fig. 5 shows the accuracy curve while Fig. 6 shows the loss curve of the CNN model over 150 epochs. Similarly, the FNN model also achieved an accuracy of 100%. Other ML algorithms have a lower accuracy compared to CNN. The DT and NB algorithms achieved an accuracy of 95.6% and 98.7%, respectively, as shown in Table 6. On the other hand, when using CV at a rate of 5-fold, CNN, and FNN also achieved an accuracy of 100%. However, the DT and NB algorithms had an accuracy of 99.3%, and 98.5% respectively as shown in Table 7.

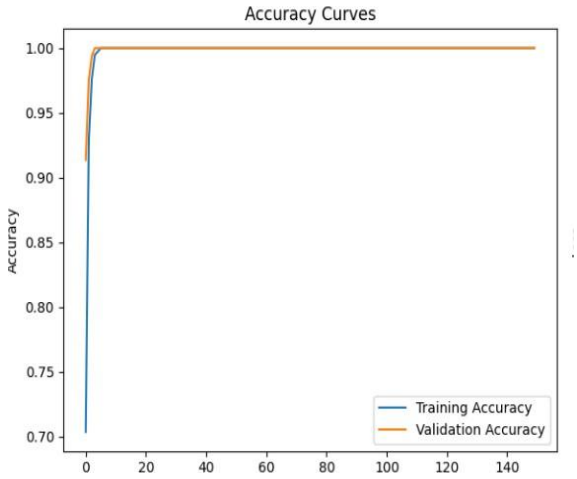

Fig. 5 Accuracy Curve

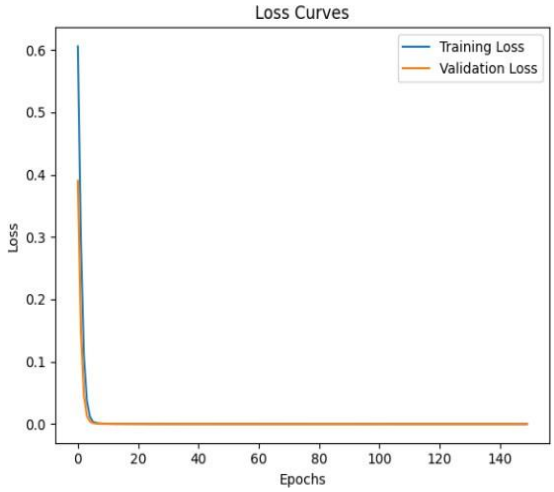

Fig. 6 Loss Curve

Table 6: Results of algorithms without CV.

Algorithm	Acc.	Prec.	Recall	F1-score
CNN	100	100	100	100
FNN	100	100	100	100
Naive Bayes	98.7	98.5	98.5	98.5
DT	95.6	92.8	97	94.8

Table 7: Results of algorithms using CV.

Algorithm	Acc.	Prec.	Recall	F1-score
CNN	100	100	100	100
FNN	100	100	100	100
Naive Bayes	98.5	97.7	99.2	98.4
DT	99.3	98.9	99.6	99.2

Our work was implemented in Google Colab and Kaggle environments, leveraging the powerful Kaggle P100 GPU for execution. The training times for CNN and FNN are 16.461, and 14.905 seconds, respectively without using CV, and the training times for NB and DT are 0.534, and 1.628 seconds, respectively, without using CV. A higher computational load is recorded using CV. The varying computation times of different algorithms demonstrate how important it is to consider predictive performance and computational economy when selecting a model, particularly when resource optimization or real-time decision-making are required. Table 8 presents training time without CV while Table 9 shows training time using CV.

Table 8: Time Performance for proposed DL and ML without CV.

Algorithm	Time without CV (s)	Accuracy without CV (%)
CNN	16.461	100
FNN	14.905	100
Naive Bayes	0.534	98.7
DT	1.628	95.6

Table 9: Time Performance for the proposed DL and ML using CV.

Algorithm	Time with CV (s)	Accuracy with CV (%)
CNN	78.061	100
FNN	75.798	100
Naive Bayes	0.572	98.5
DT	1.543	99.3

4.5. Discussion

Table 10 shows the general summary of the results we obtained from each algorithm compared with different research works. Due to the use of the ALL-IDB1 dataset, one of the problems with this dataset is that it is made up of a few pictures. To boost the amount of photos in this dataset, a data augmentation technique was used. As it is known, CNN requires large datasets because a small training data set may cause the problem of overfitting. To overcome this problem, we used data augmentation. The second is large because it includes all components of the blood without removing any part of these components. For this reason, we used feature extraction using the VGG16 method, after which the features were selected, and then the ML and DL algorithms were applied to obtain good results. The problem is that the image size of this data set is compared to that of previous studies. We also evaluated our approach using a 5-fold CV.

Table 10 presents that the results obtained were more accurate and less complex than previous efforts.

5. CONCLUSION

Our findings demonstrated the efficacy of DL and ML approaches in automating the detection and categorization of ALL using bone marrow images. Our proposed method proved its efficiency in dealing with the limited number of

sampled images by applying data augmentation techniques that solved the overfitting problem.

We used VGG16 to extract features and features selected using the RFE method. As a result, 100% accuracy was achieved using the ML and DL algorithms. The tested models showed excellent accuracy and robust classification metrics, indicating their benefit in clinical practice

Table 10: Models Implementation results.

Researcher	Algorithm	datasets	Use Data Augmentation	Training and Testing data.	Use CV or not	Accuracy (%)
[6]	CNN	ALL_IDB1	YES(The data was augmented by randomly flipping, rotating and zooming the images.)	85% for training data 15% for test data.	NO	99.22
[9]	FNN CNN	FNN applied to ALL_IDB1 CNN applied to ALL_IDB2	NO	FNN(81 images for training 27 images for testing.) CNN(156 images for training 104 images for testing.)	NO NO	95.40 98.33
[12]	DT	ALL-IDB1	NO	4 folds are used to train and 1 fold used to test	YES	98
[15]	CNN	ALL_IDB1 and ALL_IDB2 (368 images, out of them 108 and 260 from ALL-IDB1 and ALL-IDB2, respectively).	YES(736 images have been generated)	515 images for training 221 images for testing.	YES	99.5
[17]	CNN	ALL_IDB1 and ALL_IDB2 (268 images, out of them 108 and 160 from ALL-IDB1 and ALL-IDB2, respectively).	YES(1000 images have been generated)	741 images for training 259 images for testing.	YES	99.61
Present Study	CNN FNN NB DT	ALL_IDB1	YES(540 images)	378 images for training 162 images for testing.	NO NO NO NO	100 100 98.7 95.6
Present Study	CNN FNN NB DT	ALL_IDB1	YES(540 images)	378 images for training 162 images for testing.	YES YES YES YES	100 100 98.5 99.3

REFERENCES

[1] A. Pastorczak, K. Domka, K. Fidyt, M. Poprzeczko, and M. Firczuk, "Mechanisms of immune evasion in acute lymphoblastic leukemia," Cancers, vol. 13, no. 07, p. 1536, 2021, doi:10.3390/cancers13071536.

[2] R. Kansal, "Diagnosis and Molecular Pathology of Lymphoblastic Leukemias and Lymphomas in the Era of Genomics and Precision Medicine: Historical Evolution and Current Concepts—Part 1: Lymphoid Neoplasms," Lymphatics, vol. 1, no. 2, pp. 55-76, 2023 ,doi:10.3390/lymphatics1020007

[3] M.-E. Percival, C. Lai, E. Estey, and C. S. Hourigan, "Bone marrow evaluation for diagnosis and monitoring of acute myeloid leukemia," Blood reviews, vol. 31, no. 4, pp. 185-192, 2017. doi:10.1016/j.blre.2017.01.003

[4] L. Malcovati et al., "Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet," Blood, The Journal of the American Society of Hematology, vol. 122, no. 17, pp. 2943-2964, 2013, doi.:10.1182/blood-2013-03-492884

[5] D. Umamaheswari and S. Geetha, "Segmentation and classification of acute lymphoblastic leukemia cells tooled with digital image processing and ML techniques," in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018: IEEE, pp. 1336-1341.

[6] E. Chen, R. Liao, M. Y. Shalaginov, and T. H. Zeng, "Real-time detection of acute lymphoblastic leukemia cells using deep learning," in 2022 IEEE International Conference on Bioinformatics and

Biomedicine (BIBM), 2022: IEEE, pp. 3788-3790. doi:10.1101/2022.10.22.513362

[7] S. S. Al-jaboriy, N. N. A. Sjarif, S. Chuprat, and W. M. Abduallah, "Acute lymphoblastic leukemia segmentation using local pixel information," Pattern Recognition Letters, vol. 125, pp. 85-90, 2019, doi:10.1016/j.patrec.2019.03.024

[8] P. K. Das, V. Diya, S. Meher, R. Panda, and A. Abraham, "A systematic review on recent advancements in deep and machine learning based detection and classification of acute lymphoblastic leukemia," IEEE access, vol. 10, pp. 81741-81763, 2022.

[9] S. Rajpurohit, S. Patil, N. Choudhary, S. Gavasane, and P. Kosamkar, "Identification of acute lymphoblastic leukemia in microscopic blood image using image processing and machine learning algorithms," in 2018 International conference on advances in computing, communications and informatics (ICACCI), 2018: IEEE, pp. 2359-2363.

[10] S. Shafique, S. Tehsin, S. Anas, and F. Masud, "Computer-assisted acute lymphoblastic leukemia detection and diagnosis," in 2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), 2019: IEEE, pp. 184-189.

[11] R. R. Putri, E. P. Mandyartha, and A. S. Indrawanti, "Automatic Identification of Acute Lymphoblastic Leukemia on Blood Cell An image Using Geometric Features," in IOP Conference Series: Materials Science and Engineering, 2019, vol. 462, no. 1: IOP Publishing, p. 012018, doi:10.1088/1757-899X/462/1/012018

[12] M. N. Q. Bhuiyan, S. K. Rahut, R. A. Tanvir, and S. Ripon, "Automatic acute lymphoblastic leukemia detection and comparative analysis from images," in 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT), 2019: IEEE, pp. 1144-1149.

[13] N. Ahmed, A. Yigit, Z. Isik, and A. Alpkocak, "Identification of leukemia subtypes from microscopic images using convolutional neural network," Diagnostics, vol. 9, no. 3, p. 104, 2019, doi:10.3390/diagnostics9030104

[14] S. Chand and V. Vishwakarma, "Leukemia diagnosis using computational intelligence," in 2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), 2019, vol. 1: IEEE, pp. 1-7.

[15] S. Anwar and A. Alam, "A convolutional neural network–based learning approach to acute lymphoblastic leukaemia detection with automated feature extraction," Medical & biological engineering & computing, vol. 58, no. 12, pp. 3113-3121,2020,doi:10.1007/s11517-020-02282-x

[16] K. Dese et al., "Accurate machine-learning-based classification of leukemia from blood smear images," Clinical Lymphoma Myeloma and Leukemia, vol. 21, no. 11, pp. e903-e914, 2021, doi: 10.1016/j.clml.2021.06.025

[17] U. Saeed, K. Kumar, M. A. Khuhro, A. A. Laghari, A. A. Shaikh, and A. Rai, "DeepLeukNet—A CNN based microscopy adaptation model for acute lymphoblastic leukemia classification," Multimedia Tools and Applications, vol. 83, no. 7, pp. 21019-21043, 2024,doi:10.1007/s11042-023-16191-2

[18] M. Hosseinzadeh et al., "A Diagnostic model for acute lymphoblastic leukemia using metaheuristics and deep learning methods," arXiv preprint arXiv:2406.18568, 2024.

[19] L. Zare, M. Rahmani, N. Khaleghi, S. Sheykhivand, and S. Danishvar, "Automatic Detection of Acute Leukemia (ALL and AML) Utilizing Customized Deep Graph Convolutional Neural Networks," Bioengineering, vol. 11, no. 7, p. 644, 2024. doi:10.3390/bioengineering11070644

[20] H. Kamel, D. Abdulah, and J. M. Al-Tuwaijari, "Cancer classification using gaussian naive bayes algorithm," in 2019 international engineering conference (IEC), 2019: IEEE, pp. 165-170.

[21] Y. Izza, A. Ignatiev, and J. Marques-Silva, "On explaining decision trees," arXiv preprint arXiv:2010.11034, 2020.

[22] M. M. Taye, "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions," Computation, vol. 11, no. 3, p. 52, 2023. doi:10.3390/computation11030052

[23] J. Yang and J. Ma, "Feed-forward neural network training using sparse representation," Expert Systems with Applications, vol. 116, pp. 255-264, 2019,doi:10.1016/j.eswa.2018.08.038

[24] N. W. S. Wardhani, M. Y. Rochayani, A. Iriany, A. D. Sulistyono, and P. Lestantyo, "Cross-validation metrics for evaluating classification performance on imbalanced data," in 2019 international conference on computer, control, informatics and its applications (IC3INA), 2019: IEEE, pp. 14-18.

مقارنة بين تقنيات التعلم العميق والتعلم الآلي للتشخيص الآلي لسرطان الدم الليمفاوي الحاد

زينة محمود zeina.mahmoud2022@stu.uoninevah.edu.iq

محمد عبد الله محمد عبد المطلب mohammmed.abdulmuttaleb@uoninevah.edu.iq

قسم هندسة الحاسوب والمعلومات، كلية هندسة الالكترونيات، جامعة نينوى، الموصل، العراق

تاريخ الاستلام: أغسطس 2024 استلم بصيغته المنقحة: 27 أكتوبر 2024 تاريخ القبول: 28 نوفمبر 2024

الملخص

يمثل سرطان الدم الليمفاوي الحاد (ALL) تحديات تشخيصية وتصنيفية كبيرة بسبب نطاقه الواسع من المظاهر السريرية والطبيعة الغازية لإجراءات التشخيص الحالية. في هذا العمل، ندرس كيف يمكن استخدام نهجي التعلم العميق (DL) والتعلم الآلي (ML) لتحسين تشخيص وتصنيف سرطان الدم الليمفاوي الحاد باستخدام صور نخاع العظم. نقوم بإجراء تحقيق مكثف لأداء نماذج التعلم العميق والتعلم الآلي المختلفة، مثل الشبكة العصبية التلافيفية (CNN)، والشبكات العصبية ذات التغذية الأمامية (FNN)، وNB، وDT، باستخدام مجموعة بيانات ALL-IDB1. تم اقتراح طريقة استخراج ميزات متطورة تعتمد على VGG16 مع اختيار الميزات بناءً على إزالة الميزات المتكررة. تتضمن دراستنا ضبط النماذج المدربة مسبقًا، واستخراج الميزات باستخدام VGG16، وتحسين النموذج. يتم استخدام درجة F1 والدقة ومقاييس التذكر لتقييم أداء النموذج. أنتج التحقيق نتائج مشجعة، حيث سجل كل من نموذجي التعلم العميق والتعلم الآلي دقة 100٪ ودرجات تصنيف ممتازة. بالإضافة إلى ذلك، تم التأكيد على أن الأنظمة الآلية القائمة على نماذج التعلم العميق والتعلم الآلي لديها القدرة على تحسين نتائج المرضى من خلال تسريع وتحسين دقة التشخيص.

الكلمات الدالة:

سرطان الدم الليمفاوي الحاد، التعلم العميق، التعلم الالي، استخراج الميرات، الشبكة العصبية التلافيفية.