Analysis of LEUKEMIA classification using Microarray Data with Deep Learning models

Nithin Kandi[a], Dorai Sai Charan[a], Vijay Gosu[a], Vasavi C.S[b,*], Karthikeyan B[c,*] and Hemprasad Y. Patil[c]

[a]Department of Computer Science and Engineering, Amrita School of Artificial Intelligence, Bengaluru, Amrita Vishwa Vidyapeetham, India
[b]Department of Artificial Intelligence, Amrita School of Artificial Intelligence, Bengaluru, Amrita Vishwa Vidyapeetham, India
[c]Department of Embedded Technology School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

ARTICLE INFO

Keywords:
Microarray data
Neural Networks
Hybrid model
Leukemia
Deep Learning
Classification

ABSTRACT

Leukemia is a severe type of cancer. Basic classification methods have constraints in diagnosis and treatment. Leukemia classification using Microarray data with deep learning models, including Convolutional Neural Networks (CNNs), Multi-Layer Perceptrons (MLPs), and Recurrent Neural Networks (RNNs), improves classification accuracy. A hybrid deep learning model combining CNN+LSTM and MLP enhances classification further. The study aims to accurately classify leukemia subtypes and normal samples through gene expression profiling. Extensive experiments prove that deep learning models, particularly CNNs, outperform traditional classification methods. Deep-learning-based techniques improve leukemia diagnosis accuracy and make treatment more effective and personalized. The study highlights a paradigm shift toward leukemia diagnosis using deep learning in modern healthcare.The Hybrid model (CNN+LSTM and MLP) achieves 96.6% accuracy, showing a remarkable improvement in leukemia classification. This precision is key to developing tailored treatment plans. The results validate deep learning's role in medical diagnostics and open the way to future innovations.The adoption of deep-learning-based models in clinical practice increases diagnostic precision and improves patient outcomes with personalized therapeutic interventions. This work confirms the significance of deep learning in medical diagnostics, supporting AI-driven leukemia detection. The findings reinforce deep learning's potential to revolutionize leukemia diagnosis and treatment. Implementing AI in clinical workflows enhances early detection, improving survival rates and patient care through precise, individualized therapies.

1. Introduction

Deep learning has transformed leukemia diagnosis, with models such as MobileNetV2 performing outstandingly on datasets such as ALL and ASH, far surpassing traditional methods [45]. CNNs have automated white blood cell classification, thus providing faster and more accurate leukemia detection [46]. Over the last decade, deep learning, especially CNNs, has been successful in achieving high accuracy while overcoming challenges such as data imbalance and computational complexity ([47],[53]). Machine learning and deep learning have also been improved in the early detection of cancers such as leukemia. This has helped in increased accuracy and decreased errors during diagnosis [48]. Frameworks like ResNet-34 improved the accuracy of leukemia classification, making them applicable to clinical use [49]. Techniques like CNNs, SVM, and K-NN have been optimized; hence, the accuracy in these techniques has improved a lot [50]. Further improvements in classification accuracy were made by combining nature-inspired optimization methods such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Grey Wolf Optimization (GWO) with stacked ensemble learning [51]. Ensemble approaches using Deep Neural Networks (DNNs) and gradient-boosting meta-classifiers have achieved high accuracy rates, up to 98% [52].

ORCID(s):

Innovative solutions such as GAN-based data augmentation have enhanced the classification efficiency of Acute Leukemia (AL) based on subtypes like normals, AML, or ALL with accuracy ranging between 73% and 86% [54]. Technologies like DDRNet have begun to bridge the gaps encountered in the detection of ALL [55]. The ability to integrate GAN, hybrid models, and optimal algorithms has improved accuracy coupled with addressing challenges like sparsity and complexity of data under consideration [56]. Bone marrow sample studies have explored the application of deep learning in the diagnosis of acute lymphoblastic leukemia by overcoming validation and computational challenges ([57],[58]). Gene selection techniques; hybrid methods and genetic algorithms have improved classification accuracy, with reduced computational cost [59]. Microarray technology is applied to profiling gene expression in different subtypes of leukemia, identification of biomarkers, and even molecular diagnosis ([60],[62]). Distributed computing frameworks, for example, MapReduce-based K-NN classifiers have been proposed to deal with high-dimensional datasets derived from Microarray analyses [61]. Data mining methods like Decision Trees, SVM and K-NN have been promising with Decision Trees producing the highest accuracy [63]. Efficient gene selection with PCA has identified significant genetic markers that overlap with some markers identified in previous studies [64].

Leukemia is a cancerous disease of the blood and bone marrow with the rapid proliferation of irregular white blood

cells in the bloodstream [22]. Precise classification into the sub-types of leukemia is very important for determining the best treatment strategy and achieving the best possible outcomes for patients. Traditionally, methods to classify leukemia have depended on morphological examination and immunophenotyping [21]. However, these methods often lack sensitivity and specificity. Microarray technology helps study gene expression profiles in connection with various leukemias from an entirely new, highly detailed point of view. Now, machine-learning and deep-network approaches offer the promise of advances in medical diagnostic areas and are probably only termed as nominal because they describe the data-intensive characterization of systems that can capture complex patterns. As CNN, MLP, and RNN are some of the famous models in the deep learning world used for both image and sequence data analysis, the models are experienced for the concerned field in Microarray data classification ([23], [25]). This paper depicts the advanced techniques adopted to reach for improved classification accuracy while classifying leukemia concerning the Microarray data. Leukemia is a type of malignancy where both blood and bone marrow are implicated directly. The uncontrolled growth culminates in the overproduction of white blood cells, thereby interrupting the average blood cell production [26]. Accurate classification is associated with better prognoses, and effective treatment programs and subgroups form a requirement for precise classification. Though traditional classification is mainly based on morphological examination and immunophenotyping, these methods are, however, not sensitive and specific. Tremendous insights into the molecular differences between various sub-types of leukemias can be overwhelmingly done using Microarray technology. In a different context, medical diagnostics have also had breakthroughs with the discovery of machine-learning and deep-learning technologies. These models extract complex patterns from massive numbers of data units, which usually escape classical methods. These CNNs, MLPs, and RNNs are strong models for the analysis of image and sequence data [24] and, hence, are good contenders for Microarray data in the research of leukemia. This paper also uses an advanced hybrid model with the combination of CNN+LSTM and MLP deep learning models which gives more accurate results and is easy to understand compared to the normal deep learning models. Whether it is a classifier designed to subgroup the leukemic diseases according to some characteristic or a predictor of patient mortality, the aim of this paper is to use such state-of-the-art classifiers in order to better the current scenario of leukemia classification based on Microarray data.

[1] analyzed Microarray data across various cancers, identifying genes critical for accurate classification through classifiers like SVM and Naïve Bayes. Genetic and Multi-Objective Evolutionary Algorithms have been employed to select minimal but informative gene sets that enhance classifier performance. [2] designed a model for leukemia classification using Principal Component Analysis (PCA) for dimensionality reduction alongside a stacked ensemble learning algorithm. ([3],[6]) emphasized the effectiveness of data mining and artificial intelligence in gene selection, enhancing leukemia diagnosis accuracy. K-Means clustering has been applied in [9] to identify significant genes for leukemia, reducing the number of required genes for classification. [7] introduced ML-Pred-CLL, a method capable of predicting chronic lymphocytic leukemia with up to 90% accuracy by analyzing frequently mutated genes. [8] developed a deep neural network (DNN)-based convergent learning model for classifying acute lymphocytic leukemia (ALL) and acute myeloid leukemia (AML) using gene expression data, achieving high accuracy, sensitivity, and specificity. Ensemble machine learning methodologies, as described in [14], have shown improved diagnostic potential by outperforming simpler classifiers in Microarray datasets. In [13] the authors combined deep learning with refined feature selection, achieving high accuracy in classifying ALL and AML using autoencoders and deep neural networks. [5] underscored the significance of methods such as Mutual Information and Pearson Correlation Coefficient for selecting relevant genes from thousands.[4] demonstrated the Kruskal-Wallis test with Bonferroni correction as an effective feature selection method, leading to improved cancer classification accuracy. Advanced classifiers, including Random Forest, Decision Trees, Logistic Regression, and hybrid models, have been benchmarked across databases like CuMiDa, which consists of thousands of genes, to advance leukemia research ([6],[43]). Deep learning approaches have achieved notable success in leukemia diagnosis. Optimized Convolutional Neural Networks (CNNs), as seen in studies utilizing PCSO, BiCNN-CML, and ResNet101, have demonstrated high diagnostic accuracy, overcoming challenges associated with manual diagnosis ([26],[27],[30]). MobileNetV2 and ANCOM CNN models have excelled in classifying blood cancer, with the former achieving up to 96% validation accuracy on blood smear images ([32],[33]). Hybrid Ensemble Deep Learning models have automated the analysis of blood cell abnormalities, consistently achieving over 95% accuracy ([31]).

PCA has played a crucial role in dimensionality reduction, as observed in applications for leukemia and ovarian cancer detection ([34]). Affinity propagation clustering has also been instrumental in selecting feature genes in Microarray data, improving classification performance in leukemia cases ([36]). Studies have compared neural network algorithms, finding cascade correlation with resilient backpropagation to be particularly effective ([37]). Random Subspace ensemble methods have addressed high-dimensional Microarray data challenges, yielding strong performance in cancer detection across multiple datasets ([35]). Advances in leukemia detection have also incorporated traditional and hybrid ML techniques, including YOLO models and feature selection frameworks such as FCM-SVM-RFE and Lasso regression, improving subtype prediction and classification ([38],[42],[43]). Gene expression data remains central to these advancements, as researchers continue to refine and

Table 1
Comparative study analysis with existing methods

Author	No of classes in dataset	Method	Classification Accuracy (%)
Shermin Begum, et.al. [5]	2	KNN, DT, RF	97%
Quang Huy, et.al. [2]	4	RF, LR, KNN, DT; K-means	95.5%
Sara Haddou, et.al. [3]	2	SVM, KNN, LDA, Fisher, SNR	97%
Tasnim, et.al. [15]	2	LR, SVM, KNN, RF; Naive Bayes, ANN	95%
Rita Francese, et.al. [13]	2	SVN, DNN, LIMMA; Genetic Algorithm	92%
Ilyas Mahwish et.al.[18]	5	FLD classifier	96%
Ilyas Mahwish et.al.[19]	5	Linear Programming	98%
Rupara Vaibhav et.al.[20]	7	ETC, LVTrees, KNN,ADA, SVC, DT, RF	97%
Sunil Sharath er.al.[10]	3	KNN, SVM, NB	60.89%
Prashanth Palaniswamy et.al.[9]	5	Kmeans clustering	94%
Shadnam Sakib et.al.[12]	2	SVM, LR, RF, KNN, NB, DT	98%
Alrefai Nashat [14]	2	DT, NN, NB, KNN, SVM, Rule Learner	98%

benchmark algorithms for robust prediction accuracy, sensitivity, and specificity ([12],[15]). Challenges such as small datasets and generalizability persist, but recent proposals, including hybrid hierarchical classifiers and AI-based detection frameworks, offer promising directions ([18],[40]).

1.1. Comparative study analysis

Table 1 presents a comparative study of classification models used in various research papers. The highest accuracy of 98% was achieved on 2-class datasets using methods such as K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Rule Learner. Another study achieved 95.5% accuracy on a 4-class dataset using Random Forest (RF), Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), and K-means Clustering. The models showcase the efficacy of various machine-learning techniques across different classification tasks. In [19],[12] and [14] they achieved an accuracy of 98% which is the best among the above-mentioned works because they only classified cancer into 5,2 and 2 classes respectively and also they used the basic machine learning models like SVM, LR, RF, KNN, NB, DT etc. But in this work, the dataset is divided into 13 classes and also the models that are used in this work are advanced deep-learning models including CNN, RNN, MLP and a Hybrid model which is a combination of all three deep-learning models

This is of great importance in the classification of leukemia cases with the use of Microarray data using a hybrid deep learning model; it can be applied for improved early and accurate recognition of the leukemia sub-type. It leverages the best from CNN and LSTM in combining spatial pattern exploitation capabilities of CNNs and sequential data handling capabilities from LSTMs, ensuring thorough analysis of high-dimensional Microarray data. It enhances feature extraction and learning, which makes the model robust and generalizable. The project uses various model performance metrics for its effectiveness and reliability: accuracy, precision, recall, F1 score, and confusion matrices. This project, therefore, has the potential to impact clinical practices providing a reliable tool for the classification of leukemia that health professionals can use in making important decisions concerning the enhancement of patient outcomes. In addition, it adds value to the domain of AI in medicine by demonstrating the potential of advanced deep-learning techniques in solving complex medical challenges.

2. Methodology

2.1. Data Description

The dataset consists of Microarray gene expression profiles of patients with the diagnosis of 13 subtypes of leukemia as mentioned in Table 3, including B-cell Acute Lymphoblastic Leukemia(B cell- ALL), Juvenile Myelomonocytic Leukemia(JMML), Acute Myeloid Leukemia(AML), Adult T-cell Leukemia/Lymphoma(ATL), and Chronic Lymphocytic Leukemia(CLL). In order to compare, normal, non-leukemic samples were used as controls. The dataset contains 16382 features representing gene expression levels of samples and class labels 635 as mentioned in Table 3 specifying a sample's subclass affiliation or the group but the dataset contains imbalance, so to balance the dataset SMOTE technique is used. After SMOTE is used the new dataset contains 1027 class labels as mentioned in Table 3. It has been preprocessed to reject missing values and standardized to make samples more comparable.

Gene expression levels indicate the amount of a particular gene product, RNA, or protein produced in any given condition. The gene expression level gives a quantitative level of activity of the genes and it is very diverse among different genes, cell types, as well as environmental conditions. Quantification of the gene expression level has been very important to most areas of biological as well as medical

research, as it helps in the determination of gene functions, gene regulation, and molecular mechanisms of diseases.

Microarray is a very popular technique used for the measurement of gene expression levels. The process includes RNA isolation from the cell samples and its further conversion to complementary DNA (cDNA) with the help of reverse transcription. For example, the cDNA tagged with fluorescent dyes acts on a Microarray chip; usually, each probe corresponds to just one specific gene. Hybridization is followed by the washing-off of any cDNA that is non-specifically bounded in order to enable the scanner to detect the fluorescence intensity on each spot of that Microarray. The intensity is related to the amount of the mRNA in the sample, reflecting the gene expression level. After normalization, such data are collected and the level of different genes' relative expression is plotted and analyzed. It is more indirect and complex to extract gene expression data directly from single-cell images under a microscope. This is usually done by immunofluorescence or in situ hybridization, where the antibodies or probes, labelled with fluorescent markers, bind to those molecules, proteins, or nucleic acids (mRNA) that are targeted within the cells. High-resolution images are then captured of labelled cells with the camera-mounted microscope. Then using image analysis software, the fluorescence intensity in various areas of the images is measured. This comprises the steps of segmentation (where the cells or some structures within the cells are delineated) and quantification (which is the measurement of the fluorescence intensity). The fluorescence intensity is then converted into numbers that represent the relative expression levels of target genes and further normalized to the background fluorescence and sample-to-sample variability. Quite a few of such processes regularly rely on several tools and software packages. For instance, in the case of Microarray data analysis, normal applications are R/Bioconductor, GEO2R (an online tool by NCBI), and Agilent GeneSpring including other commercial ones. Image analysis could be done by the use of tools such as the ImageJ/ Fiji, the CellProfiler, the MetaMorph, or the Imaris for commercial analyses. With such methods and tools, the way has opened in that it is possible to measure gene expression levels, either directly or indirectly through fluorescence-based imaging techniques.

Table 2
Description of selected datasets from CuMiDa

Dataset	Classes	Samples	Repository
GSE14317	2	25	CuMiDa
GSE71449	4	45	CuMiDa
GSE33615	2	71	CuMiDa
GSE63270	2	101	CuMiDa
GSE28497	7	281	CuMiDa
GSE22529_U133A	2	52	CuMiDa
GSE22529_U133B	2	52	CuMiDa

Table 2 describes the datasets that are used in this work. The data used for this work is obtained by combining all the datasets in the Table 2. All these datasets contain the same number of features which is 16384. Many of the classes are repeated in these datasets. The combined dataset obtained contains 13 classes which in detail are discussed further.

The information on 13 classes of leukemia used in the data is

B-Cell Acute Lymphoblastic Leukemia (B-Cell ALL) B-CELL_ALL_TCF3-PBX1- A distinctive type of B-cell ALL that contains the chimaera gene TCF3-PBX1.

B-CELL_ALL_HYPERDIP- B-cell ALL 46 chromosomes. Hyperdiploidy ALL has a relatively good prognosis among the ALL subtypes.

B-CELL_ALL_HYPO- Hypodiploid expressed B-cell ALL is indicative that the patient has less than 46 chromosomes.

B-CELL_ALL_MLL- This is the subtype involving the MLL gene rearrangements (Mixed Lineage Leukemia), most commonly caused by a translocation.

B-CELL_ALL_T-ALL- This reflects the presence of T-cell markers in what is otherwise a dominant B-lineage ALL.

B-CELL_ALL_ETV6-RUNX1- One such gene is the fusion gene ETV6-RUNX1 due to a translocation between chromosomes 12 and 21.

Juvenile Myelomonocytic Leukemia (JMML) JMML- Juvenile Myelomonocytic Leukemia: Rare, aggressive leukemia in young children due to the overproduction of myelomonocytic cells.

JMML_LIN28_low- This JMML is characterized by low levels of an important gene LIN28.

JMML_LIN28_high- This sub-type of JMML is characterized by high expression levels of the LIN28 gene.

Acute Myeloid Leukemia (AML) AML- Acute Myeloid Leukemia is an aggressive form of leukemia that affects myeloid cells.

Adult T-cell Leukemia/Lymphoma (ATL) ATL is a rare subtype of T-cell cancer affects human T-cell infection by the human T-cell lymphotropic virus type 1 (HTLV-1).

Chronic Lymphocytic Leukemia (CLL) CLL is a kind of slow-process leukemia where dysfunctional B lymphocytes tend to accumulate.

Normal Normal- Represents normal, healthy cells without any leukaemia-associated abnormalities.

Table 3 shows the no of samples of each class before and after applying SMOTE. The sample count of each class remained the same after using SMOTE making the dataset a balanced one.

SMOTE

The class imbalance is dealt with by using SMOTE (Synthetic Minority Over-sampling Technique) in generating synthetic samples for the minority class; new data points are created through interpolation between a minority class sample xi and one of its k-nearest neighbours xj. The detailed formula is mentioned in Table 4. The synthetic sample x_new is generated as:

x_new = xi + lambda.(xj - xi)

Table 3
Number of samples for each class before and applying applying SMOTE technique

Target	count(Before SMOTE)	count(After SMOTE)
B-CELL_ALL_T-ALL	46	79
JMML_LIN28_high	21	79
B-CELL_ALL_HYPERDIP	50	79
B-CELL_ALL_TCF3-PBX1	22	79
B-CELL_ALL_ETV6-RUNX1	53	79
AML	58	79
ATL	65	79
B-CELL_ALL	73	79
CLL	81	79
normal	103	79
B-CELL_ALL_HYPO	18	79
B-CELL_ALL_MLL	17	79
JMML_LIN28_low	16	79
TOTAL	623	1027

where lambda is a random value between 0 and 1. This helps to find the balance of the dataset in a way and, as a result, enhances the performance of the model, especially for imbalanced data.

2.2. Auto-Encoder training

Autoencoding is an unsupervised learning technique; it is widely used for dimensionality reduction, anomaly detection, and noise reduction. In the present work, an autoencoder architecture has been implemented. There are three main components to the architecture: the input layer, the encoder, the latent space, the decoder, and then there is the output layer. The primary goal of an autoencoder is to learn a representation in the latent space that can best reconstruct the input and output features with minimal reconstruction error. In this work for the autoencoding technique, the input to the autoencoder is a dataset consisting of numerical features, represented as ‘x1, x2,., xn‘. These features are passed through the input layer of the autoencoder model. The goal is to compress these features into a lower-dimensional representation while minimizing reconstruction error. The encoder consists of a single dense layer with 256 neurons. Activation Function This layer utilizes the activation function ReLU. An encoder takes high dimensional input data to project a compact representation that is referred to as the "latent space". Latent space captures what is important for the given data and ignores noise as well as redundancy. Neurons in latent space include 256. This is a bottleneck layer where the dimension is reduced to force the autoencoder to learn a meaningful representation of the data. The decoder is symmetric to the encoder and consists of a dense layer with 256 neurons using the "Sigmoid activation function". The decoder has the role of reconstructing the original input data from the latent space representation. Ideally, this reconstruction should be as close to the input data as possible. The decoder’s output is the reconstructed data, ‘x’1, x’2,., x’256‘, and it is compared to the original input data in order to calculate the reconstruction error. This error is minimized in training to make the autoencoder better learn the underlying structure of the data. The autoencoder is trained using the Mean Squared Error (MSE) loss function, which measures the squared differences between the original and reconstructed data. An optimizer, such as Adam, is used to update the weights of the encoder and decoder during training to minimize the reconstruction error. The model is trained for a fixed number of epochs, and both training and validation losses are monitored. The training process updates the weights of the model to decrease the reconstruction error. Plotting the losses over epochs assesses the model’s performance and potential overfitting. Once trained, the autoencoder’s performance is measured by comparing original data and reconstructed data. Samples are plotted to look at how well the autoencoder reconstructs the input features. All the samples’ reconstruction errors are calculated and are then presented in a histogram for their distribution. From the above distribution, one can detect anomalies or outliers in data points due to high reconstruction errors showing data points not according to the patterns learned. The Fig 1 illustrates an autoencoder showing the structure of the layers: an input layer followed by encoder 256 neurons with the ReLU activation, followed by a latent space where the decoder starts with a 256 neurons Sigmoid activation, finally the reconstruction output and then End.

The graphs in Fig 2 are training and validation losses, showing how well the model is performing during the training phase. The training loss was decreased from an initial value of around 0.521 down to 0.516 over 50 epochs, indicating how well the model learned reconstruction tasks and optimized its capabilities for reconstruction. A steadily dropping training loss suggests that the autoencoder was doing an effective job of minimizing reconstruction errors for the training data. On the other hand, validation loss fluctuates between 0.738 and does not grow much in the process of training, which indicates the model did not overfit to the training data set but its generalization ability to the unseen data is not improved during the process. The small

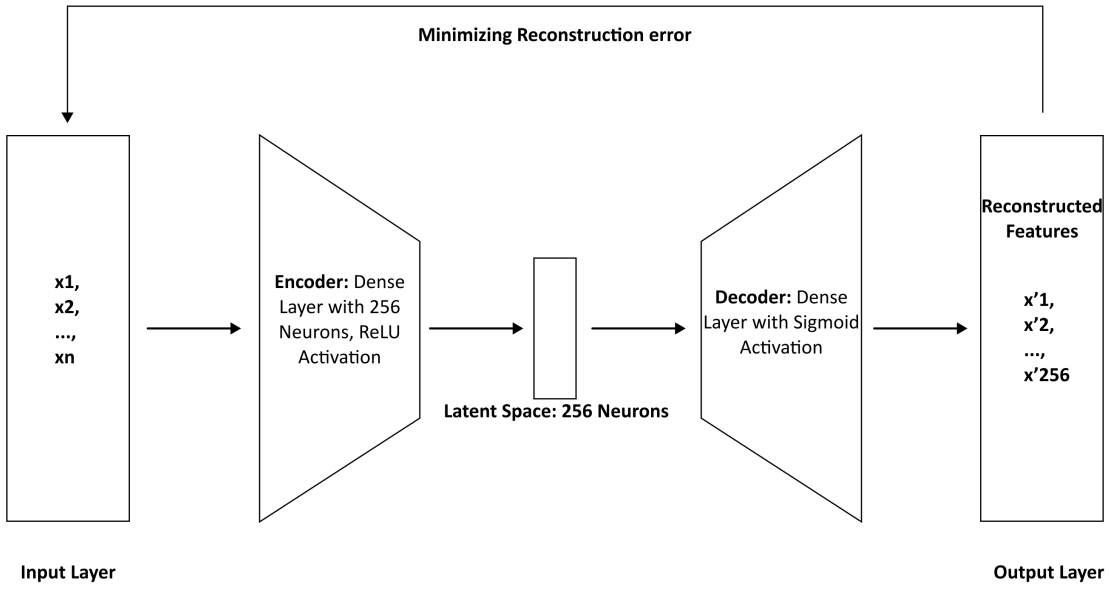

Figure 1: The architecture of Autoencoder used in this work

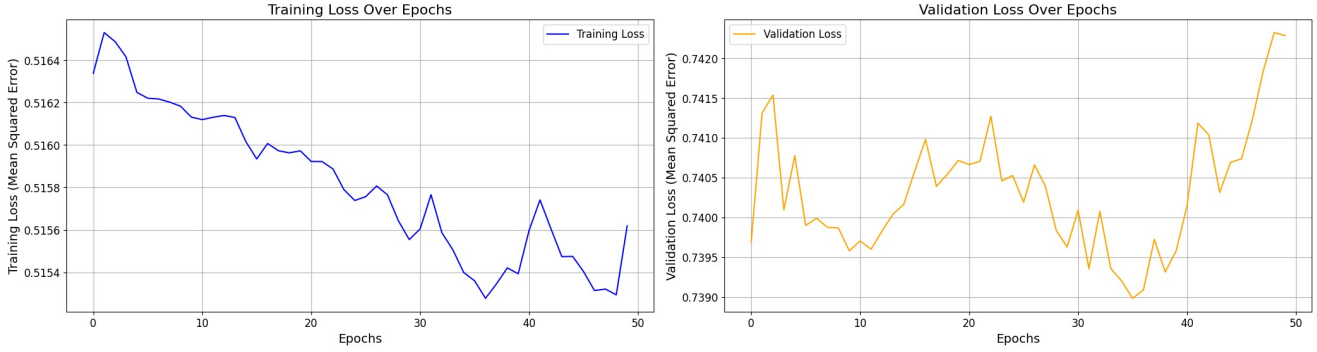

Figure 2: Training and validation loss of autoencoder data after each epoch

gap between training and validation loss suggests that the model is robust and stable but also indicates that it may need further fine-tuning, such as tweaking its architecture or hyperparameters, to improve its reconstruction of data outside the training set.

Fig 3 represents the graph of original and reconstructed values clearly demonstrating that the autoencoder successfully learned the structure of relationships within the dataset. Reconstructed values closely correspond to the original input patterns and validate the model's capacity to preserve and replicate essential features. However, in the decoding layer, it employs a sigmoid activation function that will result in output bounded between [0, 1]. This compresses the extreme values existing in the original input data. Thus, it causes some discrepancies in the reconstructed values mainly for the inputs with enough magnitudes. Despite these, alignment between the original and reconstructed data is very good for most samples, indicating a proper performance of the model on dimensionality reduction and reconstruction. The performance further suggests that this autoencoder well retains significant information but perhaps does not adequately pick extreme values or even more intricate patterns.

The Fig 4 represents a graph that could be reconstructed from the distribution of error and offers insights on how an autoencoder could perform based on the samples drawn. In this case, it spans a minimum value of 0.139 to a maximum of 1.421 with a mean reconstruction error of 0.560.

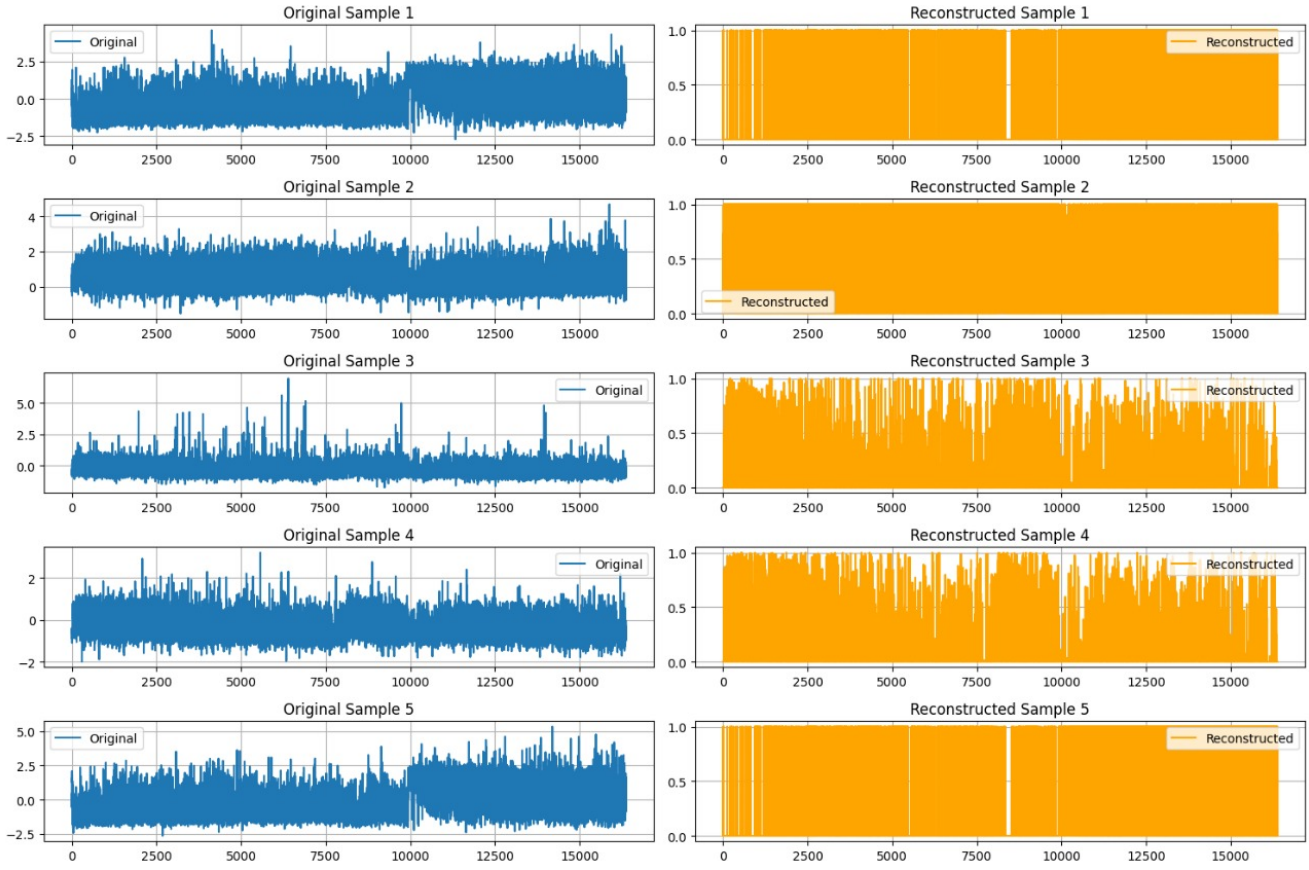

Figure 3: Original and reconstructed values of data

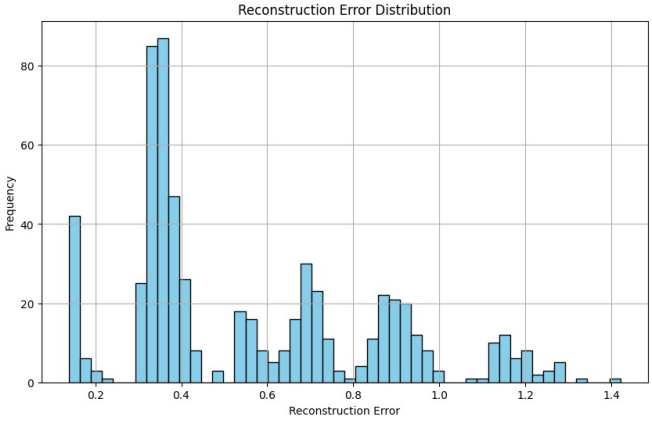

Figure 4: Reconstruction error distribution for autoencoder

This indicates that the majority of samples are placed in the low error range between 0.3 and 0.8. However, the distribution also shows a long tail, and a few samples have significantly higher reconstruction errors. These most likely correspond to data points that have complex or unusual patterns, which the autoencoder finds hard to reproduce. Overall low error values for most of the samples prove the reliability and consistency of the model, whereas high-error instances reveal scopes of further improvement in terms of including extra layers or the size of latent space. This graph illustrates the fact that autoencoders possess excellent real-world capability since they generalize well to varied patterns while pointing out outliers.

2.3. Data Collection and Pre-processing

The process of classifying leukemia using CNN, RNN, MLP and CNN+LSTM hybrid model starts from

Fig 5 describes the flow of work for data pre-processing. The dataset is a Microarray dataset of leukemia patients and is loaded into a pandas Data Frame, in which the pre-processing consists of the extraction of features and labels. The features present in the dataset represented the values of Microarray expression, while the labels represented the types of leukemia. To begin with, features and labels were separated; then, basic data analysis was carried out to understand the structure, shape, and class distribution of the dataset. Then conversion to the categorical values of the classes into numerical values and changed them back to categorical values using Label Encoder, for them to be readable by deep learning models. Standard Scaler is the main feature data to force it to have zero mean and one as standard deviation. This is an important transformation

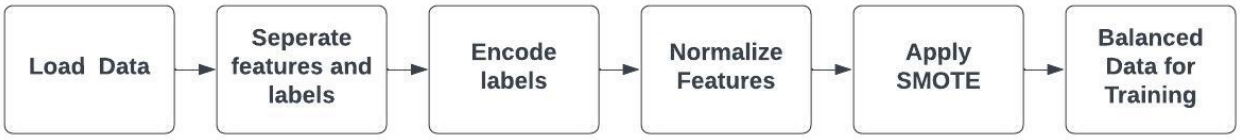

Figure 5: Flowchart of data processing

for effective training of neural networks. Then SMOTE technique is used to balance the dataset which has a class imbalance. The SMOTE technique is used to avoid class imbalance by generating new samples for minority classes.

2.4. Model Architecture and Training

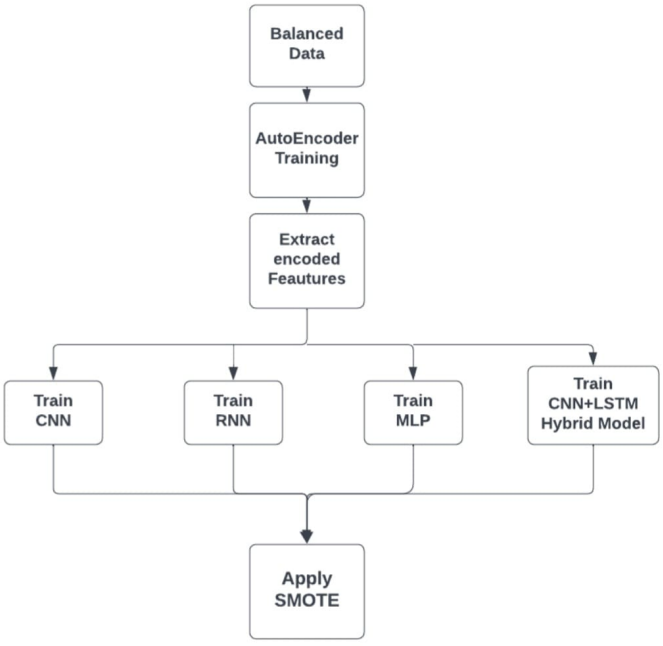

Figure 6: Flowchart of Model Training

The Fig 6 describes the flow of training the model. The work describes a three-architecture deep learning model, i.e., CNN, MLP, and RNN. Utilize an auto-encoder for feature extraction which improves the performance mainly for classification models. It was followed by an encoding layer with 256 neurons and a decoding layer to reconstruct the original input. The auto-encoder was trained to reduce the mean squared error for the output. Once trained, the auto-encoder was put to reducing dimensions in a lower-dimensional encoded representation of the input features.

CNN: The features are reshaped. The reshaping is done to correspond to the class of CNN input shape. The CNN architecture comprises the convolutional layer, activation function relu, max-pooling, flattening, and fully connected dense layers. Inclusion of dropout for regularization. The CNN model is trained with the help of loss function, categorical cross entropy and Adam optimizer. MLP: The model had multiple dense and dropout layers to avoid the problem of overfitting with an MLP. The model was trained using the same loss function and optimizer as the CNN model. RNN: The model was implemented by embedding SimpleRjson layers to help capture the temporal dependencies within the sequential data. This model was trained similarly to all other models. Model evaluation Classifiers were evaluated using various evaluation metrics, such as accuracy, confusion matrices, and classification reports, with a separate test set. The evaluation metrics are very informative on how every model performed with respect to the classification of leukemia types and point out strengths and weaknesses due to the criteria used for classification. The Fig 7 represents the architecture of the CNN model developed for classification starting with an Input Layer that accepts data in shape (16, 16, 1). The first layer is the Conv Layer 1, which has 32 filters of size 3x3 using the ReLU activation function followed by the Max Pooling Layer 1, having a 2x2 pool size for the reduction of spatial dimensions, and then the Dropout Layer 1 with a rate of 0.25. After that, Conv Layer 2 applied 64 filters of size 3x3 with a ReLU activation. Then, Max Pooling Layer 2 used the pool size 2x2. After this, Dropout Layer 2 applied a rate of 25%. After these layers, the flattened layer flattens out the data. This vector is passed to Dense Layer 1, which has 128 neurons and uses ReLU as its activation function. Here, it applies Dropout Layer 3 with a rate of 0.5. The Output Layer has 13 neurons with the Softmax activation function; categorical cross-entropy is used as its loss function, and the optimizer is Adam.

The RNN model developed for classification starts with an Input Layer that accepts data in shape (16, 16,). The first layer is the simpleRNN Layer, which has 128 neurons and uses the ReLU activation function. After this, the Dropout Layer was applied at a rate of 5%. The Output Layer has 13 neurons with the Softmax activation function; categorical cross-entropy is used as its loss function, and the optimizer is Adam.

The MLP model developed for classification starts with an Input Layer that accepts data in shape (256,). The first layer is the dense layer 1, which has 512 neurons and uses the ReLU activation function. After this, dense layer 2 has 256 neurons and uses the ReLU activation function. Then Dropout Layer 1 applied a rate of 5%. After that Dropout Layer 2 applied

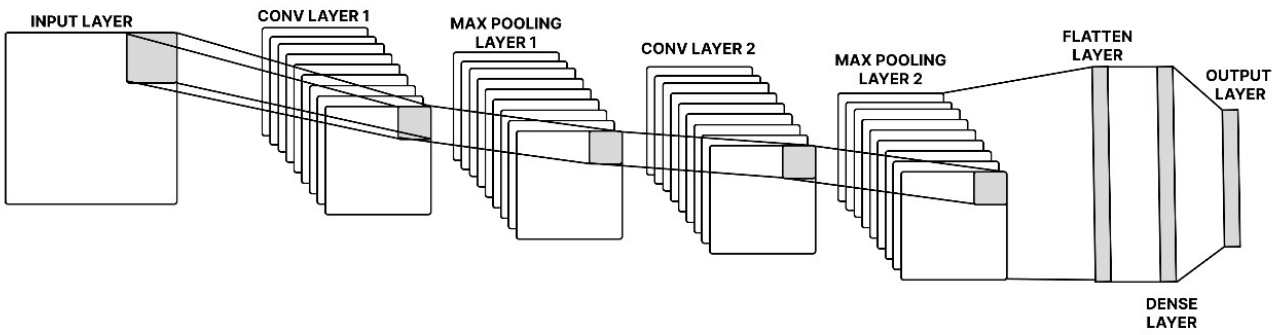

Figure 7: CNN model architecture

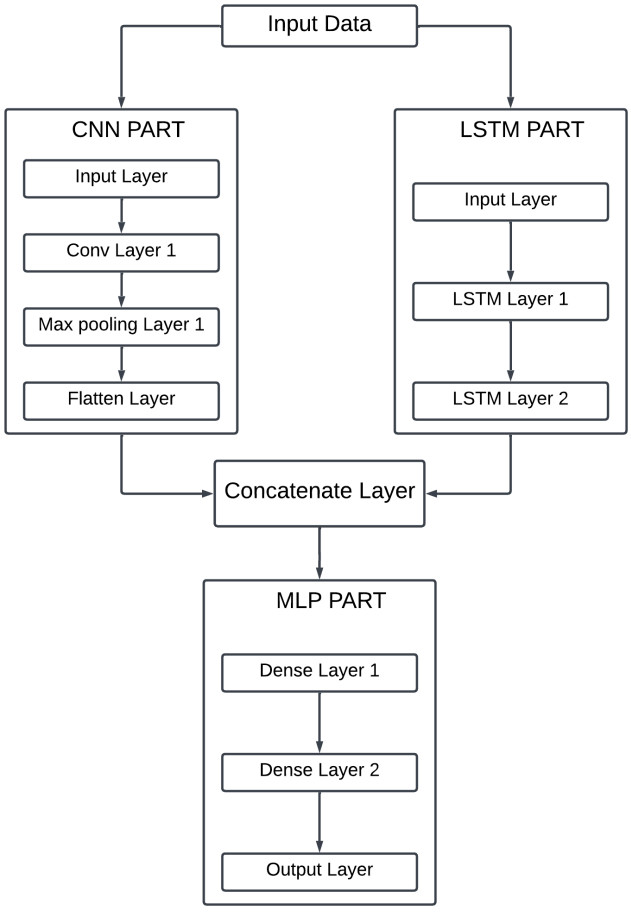

Figure 8: CNN+LSTM MLP Hybrid model architectural details

a rate of 5%. The Output Layer has 13 neurons with the Softmax activation function; categorical cross-entropy is used as its loss function, and the optimizer is Adam.

Fig 8 represents the architecture of the hybrid model developed for classification starting by inputting data. The hybrid model consists of three parts which are CNN part, LSTM part, and MLP part. CNN part starts with an input layer that accepts data in shape (16, 16, 1). The first layer is the Conv Layer 1, which has 32 filters of size 3x3 using the ReLU activation function after that Max Pooling Layer 1, having a 2x2 pool size for the reduction of spatial dimensions, and then the Flatten Layer flattens out the data. LSTM part starts with an input layer that accepts data in shape (16, 16). The first layer is LSTM layer 1, which has 32 units. After that LSTM layer 2, which has 32 units. Table 4 shows all the activation functions used here. Now outputs of the CNN part and LSTM are combined in concatenate Layers. After this MLP part, the first layer is the dense layer 1, which has 64 neurons and uses the ReLU activation function. After this, dense layer 2 has 32 neurons and uses the ReLU activation function. The Output Layer has 13 neurons with the Softmax activation function; categorical cross-entropy is used as its loss function, and the optimizer is Adam.

Function developed for the prediction of types of Leukemia for new samples. The developed function uses the trained auto-encoder for coding new data, and reshaping back to go through the respective deep learning models: CNN, MLP, RNN. The output arrived at all these models is then decoded to drive the final class labels. Fig.8 describes the structure of the CNN+LSTM hybrid deep learning model to classify data. The target variable is encoded using LabelEncoder, and features are normalized using StandardScaler. After normalization, the features are reshaped to suit the inputs of the CNN and RNN models. The model's architecture mainly consists of two parts: the CNN and the LSTM which is a technique of RNN. Spatial features are extracted using convolutional and max-pooling layers through the CNN component, followed by the flattening of the feature maps. The RNN component of the architecture captures temporal dependencies through the LSTM layers. These components are concatenated, and the merged features are passed through fully connected layers for classification. It is trained using Adam optimizer along with sparse categorical cross-entropy loss.

Table 4
Formulae used in this work

Z-score Normalization	$X_{\text{norm}} = \frac{X-\mu}{\sigma}$
SMOTE	$X_{\text{syn}} = X_i + \lambda \cdot (X_j - X_i)$
Mean Squared Error (MSE) for Autoencoder	$L = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$
Categorical Cross-Entropy Loss for Classification	$L = -\sum_{i=1}^{n} y_i \log(\hat{y}_i)$
CNN: Convolution Operation	$\text{Output}_{i,j} = \sum_{m=1}^{M}\sum_{n=1}^{N} \text{Input}_{i+m-1,j+n-1} \cdot \text{Kernel}_{m,n} + b$
CNN: ReLU Activation	$\text{ReLU}(x) = \max(0, x)$
CNN: Max Pooling	$\text{Output}_{i,j} = \max_{(m,n)\in\text{Window}} \text{Input}_{i+m,j+n}$
MLP: Dense Layer	$\text{Output} = \text{ReLU}(W \cdot \text{Input} + b)$
MLP: Dropout	$\text{Dropout}(x) = \begin{cases} 0 & \text{with probability } p \\ \frac{x}{1-p} & \text{with probability } 1-p \end{cases}$
RNN: RNN Cell	$h_t = \text{ReLU}(W_{hx}x_t + W_{hh}h_{t-1} + b_h)$

3. Results and discussion

The results, this study obtained are the evaluation metrics i.e. how well the model is performing under different conditions. ROC curve is a graph showing the performance of the model at all classification thresholds i.e. it plots all the classes in the curve. As shown below Fig 9 plot False Positive Rate as a horizontal label and True Positive Rate as a vertical Label.

3.1. ROC curve of all models

The curves in Fig 9 illustrate the ROC curves of CNN, RNN, MLP, CNN+LSTM MLP Hybrid model. ROC curve is a graphical representation of the diagnostic ability of a binary classifier system as the discrimination threshold is varied. In this multi-class classification scenario, the line in the graph represents ROC for a particular class of the feature set. The X-axis of the graph is False Positive Rate (FPR) which states the total number of negatives incorrectly classified as positive. On the Y-axis there is an illustrated True Positive Rate (TPR) or sensitivity, which outlines the percentage of authentic positive instances that get correctly classified. The range of the two axes is on the scale of zero to one and the best classifier exits in the top-left corner of the graph. The diagonal dashed line on the graph is placed as the reference line which displays the performance of the classifier that has no discrimination ability. The diagonal line will cross the plane at the point (12, 4), and the area under this diagonal line is calculated as 0. 5, and be used as the ground truth of the classifier's performance.

Every thin coloured line in the graph is a plot of the ROC for each of the classes ranging from Class 0 to Class 12. The legend offers the Area Under the Curve (AUC) of each class's Receiver Operating Characteristic (ROC) curve, a single real number quantifying the classifier. The AUC value varies from 0 to 1, typically, the higher AUC implies a higher model's performance. An AUC of 1. The higher index indicates a more suitable classification as 0 perfectly classifies all the inputs. 5 means at par with the rating and does not show any better performance than the random guess. The higher the ROC point of a class closer to (0, 100), the better the classifier performance of the class.

Thus, according to the data on the graph, Class 0, Class 2, and Class 12 have slightly lower AUC values of 0. 98 and 0. 94 for the CNN model which demonstrates that it has more difficulty in distinguishing the classes. However, many of the other classes like Class 1, Class 3, Class 4, and so on, have AUC values of 1, which is a perfect score. In this context, it can also be noted that by analyzing the values obtained in this work, it can be seen that the CNN model has indeed high accuracy in differentiating these classes from the other classes. This is an implication that a given CNN model works well to classify instances correctly in the said classes without many false positives or negatives.

In general, the tested CNN model shows high accuracy in all the classes and relatively high values of AUC that are equal to or slightly lower than 1. This gives a hint that this particular model is reasonable and efficient for this classifying problem.

In the RNN ROC curve, to compare the performance of classes, most classes give a perfect AUC capacity of 1. 00, indicating excellent performance. Nonetheless, there is some degree of classification problem as represented by; Class 0 with AUC of 0. 99, Class 2 with an AUC of 0. 95, and Class 12 with AUC of 0. 97.

In the MLP ROC curve, Almost all the experiments get an AUC score of 1 which perfectly represents the value of class A. 00, indicating excellent performance. But Class 0, Class 2: AUC = 0.97; Class 12: AUC = 0.96 repesent slightly less AUC values.

As it is observed for classes of this hybrid model, all classes reach a mark of 1 in terms of AUC. which means that the proposed model had a very high accuracy in separating the classes. Comparing to all the other model the hybrid model has high performance accuracy.

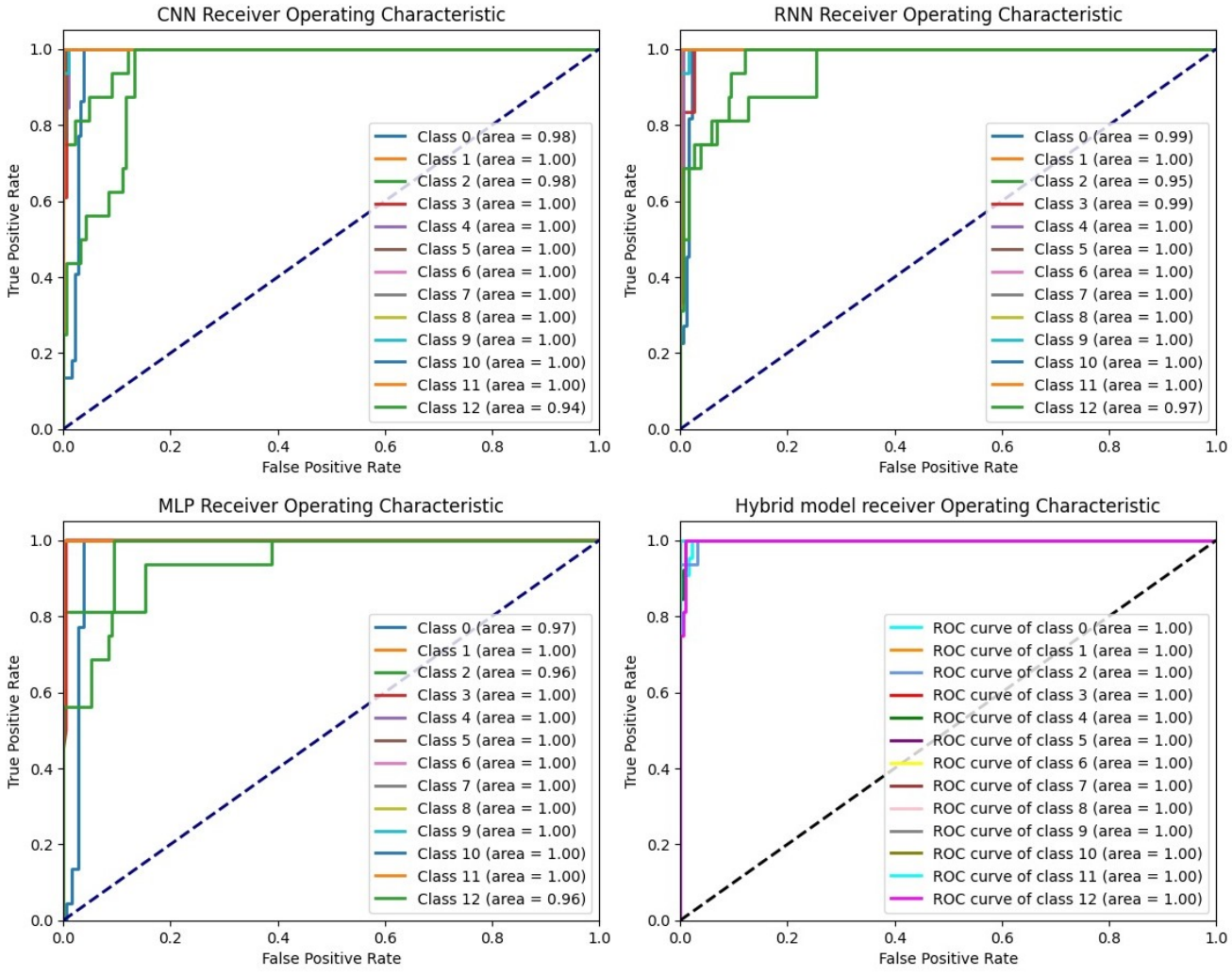

Figure 9: ROC curve of CNN, RNN, MLP, Hybrid models

3.2. Confusion matrix of Hybrid model

The confusion matrix for the Hybrid model shown in Fig 10 represents good results as most of the predicted values are under the diagonal. This means that each model operates and categorizes the disease classes in most of the cases with a certain level of credibility. Nevertheless, certain pitfalls are typical of all the models, namely, the misclassification of instances belonging to the B-CELL_ALL class. This indicates that the models CNN, RNN, and MLP are somewhat challenged in differentiating this category from other similar classes making it quite probable that the features used have some significant overlap to make the differentiation impossible with the available data. It also describes that each model has varying difficulty determining the whereabouts of the normal class. CNN and RNN's classification results are somewhat ambiguous; normal samples are recognized as belonging to AML or other diseases. On the contrary, there is a slightly better classification of healthier samples mostly because the MLP model can recognize multidimensional patterns. The Hybrid model also exhibits some gains over the CNN in this respect owing to the combination of several models to increase the chances of classification success.

However, these difficulties explain why the Hybrid model performs slightly better comparatively overall and particularly in the normal class, as seen and evidenced in the importance of harnessing the qualities of more than one model to overcome the tendencies of classifying samples into erroneous categories. This approach reduces the weaknesses that are found in individual models and gives the following stronger solution to classify disease samples accurately. Table 5 describes the evaluation metrics Accuracy, Precision, Recall, F1score of four Deep learning models CNN, RNN, MLP and CNN+LNN Hybrid model before using the class balancing technique SMOTE.

Table 6 describes the evaluation metrics Accuracy, Precision, Recall, F1score of four Deep learning models CNN, RNN, MLP and CNN+LNN Hybrid model after using the class balancing technique SMOTE. In which the Evaluation metric values are high around 95%'s. The usage of SMOTE has increased the performance of these models.

Table 5
The Evaluation metrics of models before applying SMOTE

Model	Accuracy	Precision	Recall	F1 score
CNN	76%	75%	76%	74%
RNN	78%	76%	78%	76%
MLP	80%	79%	80%	78%
CNN+LSTM Hybrid model	86%	84%	86%	85%

Table 6
The Evaluation metrics of models after applying SMOTE

Model	Accuracy	Precision	Recall	F1 score
CNN	93%	93%	93%	92%
RNN	90%	90%	90%	90%
MLP	94%	96%	95%	94%
CNN+LSTM Hybrid model	96.6%	97%	97%	96.5%

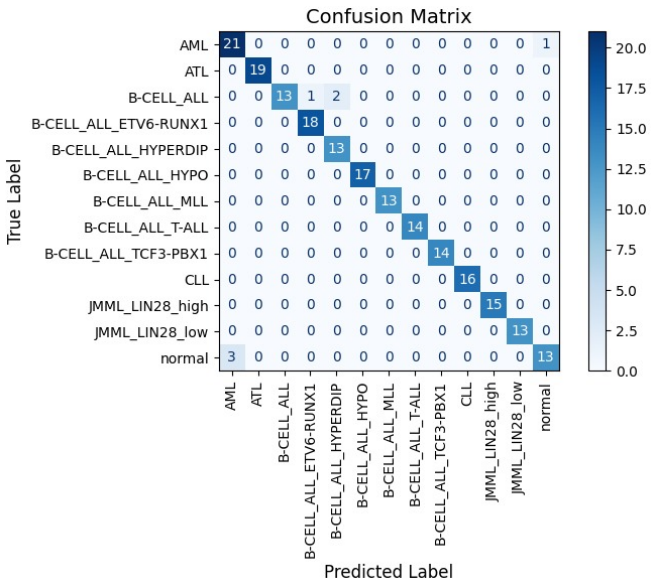

Figure 10: Confusion matrix of CNN+LSTM hybrid model

4. Conclusion

Accurate classification of leukemia subtypes using gene expression data is important to ensure precise diagnosis, effective prognosis, and tailored treatment protocols. Using advanced machine learning and deep learning models, the study has shown substantial improvement in diagnostic accuracy. With a high classification performance of 96.6% with the hybrid CNN+LSTM+MLP model, these improvements reduce the number of misdiagnoses, optimize the best-personalized treatment strategies, and ultimately improve patient outcomes by reducing adverse effects and improving quality of life.

Integration of feature selection techniques, such as RFE and LASSO, and deep learning architectures, underlines the potential of these tools to identify complex gene expression patterns. This approach not only enhances the accuracy of classification but also supports further biomedical research into leukemia pathogenesis and the development of innovative therapies. While challenges such as dataset limitation and subtype differentiation remain, this study's findings pave the way for future exploration of integrating multi-omics data, developing interpretable models, and addressing real-world clinical constraints. In a nutshell, this work underscores the transformative potential of deep learning models in medical diagnostics with a reliable framework for classifying leukemia and reinforcing the role of personalized medicine in patient care.

Code availabity

The code is available from the first author upon reasonable request.

References

[1] Mohiuddin, Tasniah, Shamima Naznin, and Paramita Basak Upama. "Classification and Performance Analysis of Cancer Microarrays Using Relevant Genes." 2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT). IEEE, 2021.

[2] Hoangp, Quang Huy, et al. "Leukemia Classification using Principal Component Analysis and Ensemble Learning on Gene Expression Data." 2023 International Conference on Advanced Technologies for Communications (ATC). IEEE, 2023.

[3] Bouazza, Sara Haddou, Layla Wakrim, and Lwaa Gafer Benedress. "Classifying Leukemia through DNA Expression Data Mining Techniques." 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA). IEEE, 2023.

[4] Das, Utsha, Md Al Mehedi Hasan, and Julia Rahman. "Influential gene identification for cancer classification." 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, 2019.

[5] Begum, Shemim, et al. "Microarray Data Analysis for Diagnosis of Cancer Diseases by Machine Learning algorithm." 2023 5th International Conference on Recent Advances in Information Technology (RAIT). IEEE, 2023.

[6] Patel, Sachin, et al. "Multi-Classifier Analysis of Leukemia Gene Expression From Curated Microarray Database (CuMiDa)." 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC). IEEE, 2021.

[7] Ahmad, Waqar, et al. "ML-Pred-CLL: Machine Learning based prediction of Chronic Lymphocytic Leukemia using protein sequential data." 2022 International Conference on Recent Advances in Electrical Engineering & Computer Sciences (RAEE & CS). IEEE, 2022.

[8] Mallick, Pradeep Kumar, et al. "Convergent learning–based model for leukemia classification from gene expression." Personal and Ubiquitous Computing 27.3 (2023): 1103-1110.

[9] Palanisamy, Prasath, K. Thangavel, and R. Manavalan. "A novel approach to select significant genes of leukemia cancer data using K-Means clustering." 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering. IEEE, 2013.

[10] Sunil, Sharath, and P. Sonu. "An effective approach for detecting acute lymphoblastic leukemia using deep convolutional neural networks." International conference on computational techniques and applications. Singapore: Springer Nature Singapore, 2021.

[11] Cho, Sung-Bae, and Hong-Hee Won. "Machine learning in DNA Microarray analysis for cancer classification." Proceedings of the First Asia-Pacific Bioinformatics Conference on Bioinformatics 2003-Volume 19. 2003.

[12] Sakib, Shadman, et al. "Blood cancer recognition based on discriminant gene expressions: A comparative analysis of optimized machine learning algorithms." 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2021.

[13] Francese, Rita, et al. "A deep learning and genetic algorithm based feature selection processes on Leukemia Data." 2022 26th International Conference Information Visualisation (IV). IEEE, 2022.

[14] Alrefai, Nashat. "Ensemble machine learning for leukemia cancer diagnosis based on Microarray datasets." International Journal of Applied Engineering Research 14.21 (2019): 4077-4084.

[15] Tasnim, Anika, and Bayezid Islam. "Analysis of Different Feature Selection Techniques on Different Machine Learning Approaches to Classify Leukemia Subclasses." 2022 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). IEEE, 2022.

[16] Kalmady, Kaushik S., et al. ”Improved transfer learning through shallow network embedding for classification of leukemia cells.” 2017 ninth international conference on advances in pattern recognition (ICAPR). IEEE, 2017.

[17] Thanmayi A, V. Lakshmi, Sunku Dharahas Reddy, and Sreeja Kochuvila. "Detection of leukemia using K-means clustering and machine learning." Ubiquitous Communications and Network Computing: 4th EAI International Conference, UBICNET 2021, Virtual Event, March 2021, Proceedings. Springer International Publishing, 2021.

[18] Sunil, Sharath, et al. ”An effective approach for classifying acute lymphoblastic leukemia using hybrid hierarchical classifiers.” Advances in Computing and Data Sciences: 5th International Conference, ICACDS 2021, Nashik, India, April 23–24, 2021, Revised Selected Papers, Part I 5. Springer International Publishing, 2021.

[19] Kavitha, K. R., Ajith Gopinath, and Midhun Gopi. ”Applying improved svm classifier for leukemia cancer classification using FCBF.” 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2017.

[20] Ananthu, K. S., et al. ”Acute lymphoblastic leukemia detection using transfer learning techniques.” Intelligent Sustainable Systems: Proceedings of ICISS 2021. Springer Singapore, 2022

[21] G. Sharma, P. S. Rana and S. Bawa, "Hybrid Machine Learning Models for Predicting Types of Human T-cell Lymphotropic Virus," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 18, no. 4, pp. 1524-1534, 1 July-Aug. 2021, doi: 10.1109/TCBB.2019.2944610.

[22] S. Rajeswari, C. S. Vasanth, C. Bhavana and K. S. S. Chowdary, "Detection and Classification of Various Types of Leukemia Using Image Processing, Transfer Learning and Ensemble Averaging Techniques," 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, 2022, pp. 1-6, doi: 10.1109/ASIANCON55314.2022.9909377.

[23] P. Vinayagam, N. V. V. Reddy, P. T. K. Kishore and A. Royappa, "Systematic Development of Modified Hybrid Learning based Detection Scheme for White Blood Cancer Cells using Medical Image Processing Methodology with IoT Alert Mechanism," 2023 Second International Conference on Advances in Computational Intelligence and Communication (ICACIC), Puducherry, India, 2023, pp. 1-6, doi: 10.1109/ICACIC59454.2023.10435291.

[24] A. Dibouliya and V. Jotwani, "Detection and Estimation of Leukemia Using Artificial Intelligence and Fuzzy Logic," 2023 2nd International Conference on Ambient Intelligence in Health Care (ICAIHC), Bhubaneswar, India, 2023, pp. 1-6, doi: 10.1109/ICAIHC59020.2023.10431474..

[25] N. H. Harun, J. A. Bakar, Z. A. Wahab, M. K. Osman and H. Harun, "Color Image Enhancement of Acute Leukemia Cells in Blood Microscopic Image for Leukemia Detection Sample," 2020 IEEE 10th Symposium on Computer Applications Industrial Electronics (ISCAIE), Malaysia, 2020, pp. 24-29, doi: 10.1109/ISCAIE47305.2020.9108810.

[26] G. Himaja et al., "Leukemia Cancer Detection in Microscopic Blood Samples using Optimized Convolutional Neural Network," 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Krishnankoil, Virudhunagar district, Tamil Nadu, India, 2024, pp. 1-5, doi: 10.1109/INCOS59338.2024.10527583.

[27] V. Malik, R. Mittal and A. Rana, "BiCNN-CML: Hybrid Deep Learning Approach for Chronic Myeloid Leukemia," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1771-1777, doi: 10.1109/IC3I56241.2022.10073460.

[28] A. Shivathaya, A. M. S, K. V. Bhat, P. N and Roopashree, "Investigative Study for Identification Categorization of Leukaemia Cell," 2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Erode, India, 2023, pp. 1-7, doi: 10.1109/ICECCT56650.2023.10179792.

[29] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider and R. Deters, "A Hybrid Deep Learning Architecture for Leukemic B-lymphoblast Classification," 2019 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2019, pp. 271-276, doi: 10.1109/ICTC46691.2019.8939959.

[30] B. Saju, V. Asha, N. Tressa, R. A. Suhas, S. Y. Jagesh and V. Sandhya, "OHDLL:Optimized Hybrid Deep Learning Model for Classification of Leukaemia images," 2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM), Bangalore, India, 2023, pp. 1-7, doi: 10.1109/IC-RVITM60032.2023.10435325.

[31] J. Jayachitra and N. Umarkathaf, "Blood Cancer Identification using Hybrid Ensemble Deep Learning Technique," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1194-1198, doi: 10.1109/ICEARS56392.2023.10084996.

[32] A. S. Khanam et al., "Classification of acute myeloid leukemia using convolutional neural network," International Conference on Green Energy, Computing and Intelligent Technology (GEn-CITy 2023), Hybrid Conference, Iskandar Puteri, Malaysia, 2023, pp. 253-258, doi: 10.1049/icp.2023.1788.

[33] S. Selvakumaran, C. Poonguzhali, E. Baskaran, A. Arya, D. K. Nayak and S. Gupta, "Integrated Hybrid Deep Learning Approach for Precise Blood Cancer Detection," 2024 10th International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 2024, pp. 496-501, doi: 10.1109/ICCSP60870.2024.10544027.

[34] A. Khoirunnisa, Adiwijaya and A. A. Rohmawati, "Implementing Principal Component Analysis and Multinomial Logit for Cancer Detection based on Microarray Data Classification," 2019 7th International Conference on Information and Communication Technology (ICoICT), Kuala Lumpur, Malaysia, 2019, pp. 1-6, doi: 10.1109/ICoICT.2019.8835320.

[35] R. Sumara, "Random Subspace Ensemble Learning for Cancer Detection Based on Microarray Data," 2021 3rd International Conference on Electronics Representation and Algorithm (ICERA), Yogyakarta, Indonesia, 2021, pp. 45-50, doi: 10.1109/ICERA53111.2021.9538683.

[36] C. -C. Chuang, Y. -C. Li, J. -T. Jeng, C. -K. Chang and Z. -Q. Wang, "Feature genes selection of adult ALL microarray data with affinity propagation clustering," 2015 IEEE International Conference on Consumer Electronics - Taiwan, Taipei, Taiwan, 2015, pp. 230-231, doi: 10.1109/ICCE-TW.2015.7216871.

[37] S. Prabu, B. R, D. K, E. G and R. K. Karthick, "Leukemia Detection using Computation Intelligence Techniques," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-6, doi: 10.1109/MysuruCon55714.2022.9972436.

[38] E. Chen, R. Liao, M. Y. Shalaginov and T. H. Zeng, "Real-time Detection of Acute Lymphoblastic Leukemia Cells Using Deep Learning," 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 2022, pp. 3788-3790, doi: 10.1109/BIBM55620.2022.9995131.

[39] S. Khobragade, D. D. Mor and C. Y. Patil, "Detection of leukemia in microscopic white blood cell images," 2015 International Conference on Information Processing (ICIP), Pune, India, 2015, pp. 435-440, doi: 10.1109/INFOP.2015.7489422.

[40] P. K. Das, D. V. A, 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," in IEEE Access, vol. 10, pp. 81741-81763, 2022, doi: 10.1109/ACCESS.2022.3196037.

[41] T. Mustaqim, C. Fatichah and N. Suciati, "Deep Learning for the Detection of Acute Lymphoblastic Leukemia Subtypes on Microscopic Images: A Systematic Literature Review," in IEEE Access, vol. 11, pp. 16108-16127, 2023, doi: 10.1109/ACCESS.2023.3245128.

[42] Ilyas, Mahwish, et al. "A Novel Leukemia Gene Features Extraction and Selection Technique for Robust Type Prediction Using Machine Learning." Arabian Journal for Science and Engineering (2024): 1-19.

[43] Ilyas, Mahwish, et al. "Linear programming based computational technique for leukemia classification using gene expression profile." Plos one 18.10 (2023): e0292172.

[44] Rupapara, Vaibhav, et al. "Blood cancer prediction using leukemia microarray gene data and hybrid logistic vector trees model." Scientific reports 12.1 (2022): 1000.

[45] S. A. Syed, H. Nisar, L. Y. Jen, S. M. Qaisar and S. Talpur, "Comparative Analysis of Deep Learning Architectures for Blood Cancer Classification," 2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 2024, pp. 1-6, doi: 10.1109/ICSIPA62061.2024.10686828.

[46] Sudheer, Rhea Vivek, Rapaka Ramesh, Richa Sarvade, Apoorva Kakarla, Sada Jain, Preethi. (2024). "Leukaemia: A Comparative Analysis of Deep Learning Models Using ALL Dataset." 1-7. 10.1109/RAICS61201.2024.10689761.

[47] Rubinos Rodriguez J, Fernandez S, Swartz N, Alonge A, Bhullar F, Betros T, Girdler M, Patel N, Adas S, Cervone A, Jacobs RJ. "A Chronological Overview of Using Deep Learning for Leukemia Detection: A Scoping Review." Cureus. 2024 May 30;16(5):e61379. doi: 10.7759/cureus.61379. PMID: 38947677; PMCID: PMC11214573.

[48] Hari Mohan Rai, Joon Yoo, Abdul Razaque. (2024) . Comparative analysis of machine learning and deep learning models for improved cancer detection: A comprehensive review of recent advancements in diagnostic techniques, Expert Systems with Applications. ISSN 0957-4174, doi:10.1016/j.eswa.2024.124838.

[49] Q. Ul Ain, S. Akbar, S. A. Hassan and Z. Naaqvi, "Diagnosis of Leukemia Disease through Deep Learning using Microscopic Images," 2022 2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2), Rawalpindi, Pakistan, 2022, pp. 1-6, doi: 10.1109/ICoDT255437.2022.9787449.

[50] Talaat, F.M., Gamel, S.A. Machine learning in detection and classification of leukemia using C-NMC_Leukemia. Multimed Tools Appl 83, 8063–8076 (2024). https://doi.org/10.1007/s11042-023-15923-8

[51] S. K. Mohapatra, S. K. P and M. N. Mohanty, "Optimizing Leukemia Classification Through Nature-Inspired Feature Reduction and Stacked Ensemble Learning Algorithms," 2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC), Bengaluru, India, 2024, pp. 1-6, doi: 10.1109/ICECCC61767.2024.10593969.

[52] Mohanty, Mihir Mohapatra, Saumendra Das, Abhishek. (2023). "A Deep Stacked Ensemble Model for Microarray Data Classification with Boosted Meta Classifier." International Journal on Artificial Intelligence Tools. 32. 10.1142/S0218213023500744.

[53] Sampathila N, Chadaga K, Goswami N, Chadaga RP, Pandya M, Prabhu S, Bairy MG, Katta SS, Bhat D, Upadya SP."Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images. Healthcare (Basel)." 2022 Sep 20;10(10):1812. doi: 10.3390/healthcare10101812. PMID: 36292259; PMCID: PMC9601337

[54] Jiunn-Woei Lian, Chi-Hung Wei, Mu-Yen Chen, Ching-Chan Lin. (2024) . "Acute leukemia prediction and classification using convolutional neural network and generative adversarial network" Applied Soft Computing, ISSN 1568-4946, doi:10.1016/j.asoc.2024.111819.

[55] Jawahar, M., Anbarasi, L.J., Narayanan, S. et al."An attention-based deep learning for acute lymphoblastic leukemia classification." Sci Rep 14, 17447 (2024). https://doi.org/10.1038/s41598-024-67826-9

[56] Asar, T.O., Ragab, M. "Leukemia detection and classification using computer-aided diagnosis system with falcon optimization algorithm and deep learning." Sci Rep 14, 21755 (2024).https://doi.org/10.1038/s41598-024-72900-3

[57] Elsayed B, Elhadary M, Elshoeibi RM, Elshoeibi AM, Badr A, Metwally O, ElSherif RA, Salem ME, Khadadah F, Alshurafa A, Mudawi D, Yassin M. "Deep learning enhances acute lymphoblastic leukemia diagnosis and classification using bone marrow images." Front Oncol. 2023 Dec 6;13:1330977. doi: 10.3389/fonc.2023.1330977. PMID: 38125946; PMCID: PMC10731043.

[58] Zare, Lida, Mahsan Rahmani, Nastaran Khaleghi, Sobhan Sheykhivand, and Sebelan Danishvar. 2024. "Automatic Detection of Acute Leukemia (ALL and AML) Utilizing Customized Deep Graph Convolutional Neural Networks" Bioengineering 11, no. 7: 644. https://doi.org/10.3390/bioengineering11070644

[59] Bilen, M., Işik, A.H. Yiğit, T. "A New Hybrid and Ensemble Gene Selection Approach with an Enhanced Genetic Algorithm for Classification of Microarray Gene Expression Values on Leukemia Cancer." Int J Comput Intell Syst 13, 1554–1566 (2020). https://doi.org/10.2991/ijcis.d.200928.001

[60] Modak, Harendra Sujayendra, Kulkarni Kulkarni, Suyamindra Pramod, B Umesh, Hallikeri Gai, Prabhanjan. (2014). "Microarray based comparative genome-wide expression profiling of major subtypes of leukemia." African Journal of Biotechnology. 13. 1174-1181. 10.5897/AJB2013.12856.

[61] Mukesh Kumar, Nitish Kumar Rath, Santanu Kumar Rath. (2016) ."Analysis of microarray leukemia data using an efficient MapReduce-based K-nearest-neighbor classifier", Journal of Biomedical Informatics, ISSN 1532-0464, https://doi.org/10.1016/j.jbi.2016.03.002.

[62] Goswami RS, Sukhai MA, Thomas M, Reis PP, Kamel-Reid S. "Applications of microarray technology to Acute Myelogenous Leukemia." Cancer Inform. 2009;7:13-28. doi: 10.4137/cin.s1015. Epub 2008 Dec 22. PMID: 19352456; PMCID: PMC2664704.

[63] Fahd, Sabry, Esmail., M., Badr, Senousy., Mohamed, Ragaie. (2016). "Predication Model for Leukemia Diseases Based on Data Mining Classification Algorithms with Best Accuracy." World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 10(5):842-851. doi: 10.5281/ZENODO.1124005

[64] Garatti, Simone Bittanti, Sergio Liberati, Diego Maffezzoli, Andrea. (2007). "An unsupervised clustering approach for leukaemia classification based on DNA micro-arrays data." Intell. Data Anal.. 11. 175-188. 10.3233/IDA-2007-11205.

Ethics declarations

Conflict of interest

The authors declare that they have no conflict of interest.