









Efficient Transfer Learning Approach for Acute Lymphoblastic Leukemia Diagnosis: Classification of Lymphocytes and Lymphoblastic Cells

Sanjay Kumar Singh¹, Mamoon Rashid^{2*}, Sultan S. Alshamrani³, Mrim M. Alnfai³, Pranshu Saxena⁴,
Aditya Khamparia⁵

¹ University School of Automation and Robotics, Guru Gobind Singh Indraprastha University, East Delhi Campus, Surajmal Vihar, Delhi 110092, India

² School of Information Communication and Technology, Bahrain Polytechnic, Isa Town 33349, Bahrain

³ Department of Information Technology, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

⁴ School of Computer Science Engineering & Technology, Bennett University, Greater Noida 201310, India

⁵ Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Amethi 227405, India

Corresponding Author Email: mamoon.rashid@polytechnic.bh

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410409>

ABSTRACT

Received: 23 June 2023

Revised: 29 December 2023

Accepted: 20 June 2024

Available online: 31 August 2024

Keywords:

transfer learning, acute lymphoblastic leukemia, lymphoblast, principal component analysis, deep learning

Introduction: Acute lymphoblastic leukemia (ALL) is a severe illness that affects children and adults, and it can be fatal when left untreated. This leukemia strikes children and adolescents suddenly, often claiming their lives within just a few weeks after diagnosis. To diagnose ALL, hematologists investigate blood slides and bone marrow samples. Manual blood testing methods, which have been around for a long time, are typically laborious and may result in lower-quality diagnoses. ALL is essentially the unchecked growth of immature cells found in the bone marrow, often referred to as lymphoblasts. **Methods:** This research focuses on the classification of lymphoblast and lymphocyte cells using a computer-assisted method that employs deep learning and image processing techniques. This classification involves several steps. Prior to feature extraction, preprocessing and data augmentation are performed on the ALL-IBD dataset. Features are extracted from this augmented database using transfer learning with pre-trained networks (DenseNet121, ResNet50, InceptionV3, Xception). The selected and transformed features, obtained through principal component analysis (PCA), are then subjected to 5-fold cross-validation for hyper-tuning and training of individual machine learning models (LR, SVM, DT, RF). Finally, a soft voting classification model is proposed to predict lymphocytes and lymphoblasts. **Results:** The suggested ensemble method achieved 98.23% accuracy. SVM and the ensemble model with DenseNet121 and all feature sets reached an AUC of 1.00. LR achieved an AUC of 1.0 with all features and 0.99 with DenseNet121 features. The minimum AUC for DT was 0.64 and for RF was 0.86. AUC with all features was 0.80 for DT and 0.91 for RF. **Conclusion:** The suggested method uses image processing and deep learning to analyze blood cells automatically, avoiding the many limitations of manual analysis. The acquired results demonstrate that the presented approach may be employed as a diagnostic tool for ALL, which is undoubtedly helpful to pathologists. **Observation:** This procedure can also be employed for enumeration, as it offers exceptional efficiency and enables prompt suspicion of a diagnosis, which can subsequently be validated by a hematologist using specialized techniques.

1. INTRODUCTION

Acute lymphoblastic leukemia is a malignancy of white blood cells (WBCs). It is characterized by an excessive generation and ongoing growth of immature and cancerous white blood cells in the bone marrow, often known as lymphoblasts or blasts. Leukemic cells proliferate rapidly in the blood and spread to various organs including the spleen, nervous system, brain, lymph nodes, and liver [1, 2]. ALL primarily affects the bone marrow and blood. Due to an abnormally high number of cancerous and immature white

blood cells, ALL leads to a deficiency of healthy blood cells. The projected incidence of ALL in the US for 2022 (covering both adults and children) is approximately 6,660 new ALL cases (2,920 in females and 3,740 in males) and approximately 1,560 ALL fatalities (680 in females and 880 in males). The likelihood of an individual being diagnosed with ALL is highest in children under the age of five. The risk remains relatively stable until the mid-20s, after which it begins to progressively rise again around age 50. In most cases, adults account for four out of every 10 cases of all forms of the disease. ALL accounts for fewer than 0.5 percent of all cancers

in the United States, making it a relatively uncommon malignancy. The lifetime chance of developing ALL is around 1 in 1,000. Males are at a slightly higher risk than females, while White individuals are at a higher risk than African Americans [3]. Most cases of ALL occur in youngsters, yet the majority of fatalities (approximately 4 out of 5) are in adults. Children may fare significantly better than adults due to differences in treatment (children's bodies often respond better to intense therapy than adults'), differences in the biology of childhood and adult ALL, or a combination of these factors [3]. The presence of more than twenty percent lymphoblasts in the bone marrow is one of the diagnostic criteria for acute leukemia. Conversely, if left undetected and untreated, it can progress rapidly and lead to death in only a few months. Fortunately, early detection of the illness aids in patient recovery, especially in cases involving children [3].

Distinguishing between lymphocytes and lymphoblastic cells is a crucial factor in the diagnosis, treatment, and monitoring of various blood-related conditions, especially leukemia and lymphoma. It enables personalized healthcare and supports ongoing scientific efforts to enhance our understanding and management of these diseases.

Following a confirmed diagnosis, the staging and categorization of Acute Lymphoblastic Leukemia (ALL) play a crucial role in determining the most suitable treatment strategy. This typically involves a combination of chemotherapy, radiation therapy, and, in some cases, stem cell transplantation. Research and treatment regimens have made significant advancements, leading to improved outcomes for individuals with ALL, particularly in infants and children.

1.1 Morphological characteristics of stained WBC

Most of the peripheral blood cells (leukocytes) are comprised of red and white blood cells. Granulocytes are a type of white blood cell that contain granules, including neutrophils, basophils, and eosinophils. Agranulocytes, on the other hand, are white blood cells that lack granules, such as lymphocytes and monocytes. Human blood typically contains the following percentage of leukocytes: 50-70% neutrophils, 2-10% monocytes, 20-45% lymphocytes, 1-5% eosinophils, and 0-1% basophils. Figure 1 illustrates the morphological structure [1, 2, 4-6].

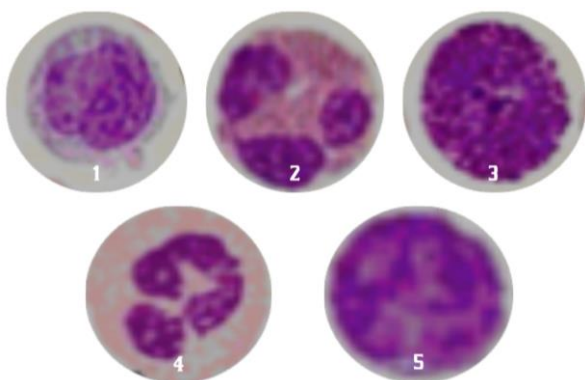


Figure 1. Morphological structure of stained WBC: 1) monocyte 2) neutrophil 3) basophil 4) eosinophil and 5) lymphocyte

There is a connection between the ALL illness and the lymphocytes found in the bone marrow and peripheral blood. Only white blood cells, particularly those with eccentrically

positioned nuclei, tend to take up the stain used in blood preparation. In most cases, white cells are significantly larger than red cells. The FAB (French-American-British) approach, which uses morphological analysis, is the most popular method for classifying leukemia (lymphoblasts). In a collection of lymphocytes, potential lymphoblasts can be identified by examining the cell's morphological distortions [5, 6]. Figure 2 explains the differences between lymphocytes and all classes of lymphoblastic cells.

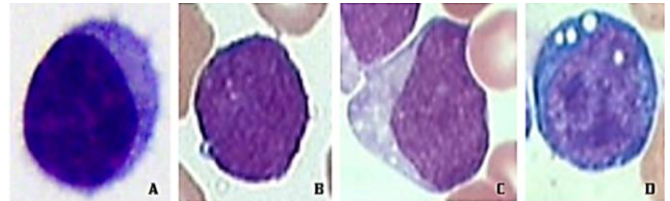


Figure 2. Morphological difference between (a) lymphocyte and (b-d) lymphoblast

For instance, lymphocytes typically have a consistent nucleus with smooth and continuous borders, and a regular shape. In contrast, lymphoblasts exhibit amorphous shapes. When it comes to ALL, potential lymphoblasts are analyzed using the FAB classification, known as L1-small, homogeneous blasts. These blasts have rounded and regular nuclei, with minimal nucleoli and cleft. Vacuoles are typically absent, and the cytoplasm is sparse. In contrast, L2-blasts are larger and more diverse. Their nuclei are disorganized and often clefted, with the possibility of more than one nucleolus. The cytoplasm is often abundant and may contain vacuoles. Finally, L3 describes blasts that are uniform in appearance and range in size from medium to large. Their nuclei are regularly shaped, round-oval, with one or more nucleoli present. Prominent vacuoles and a reasonable amount of cytoplasm are observed overall.

1.2 Limitation and prerequisite of computer assisted diagnosis system

While manual diagnosis of ALL has been crucial in detecting and managing the disease, it comes with certain constraints and difficulties. Several significant limitations include subjective variability among readers, complex cytogenic behavior, limitations in sensitive testing procedures, and reliance on clinical symptoms. Moreover, accurate classification between lymphoblasts and lymphocytes is essential for customizing treatment strategies, predicting outcomes, identifying specific therapies, minimizing adverse effects, enabling participation in clinical trials, optimizing resource distribution, and ensuring efficient long-term monitoring and follow-up care for patients. The proposed work classifies normal (lymphocyte) and lymphoblastic cells through a series of procedures. Firstly, a preprocessing approach is used to enhance image quality and reduce noise caused by various staining procedures. Subsequently, data is augmented to train models with different aspects and angles of the input image. Secondly, transfer learning methods (with pre-trained networks) are employed to extract features. These features are then input to PCA in the third step. PCA reduces the dimensionality of the extracted features, enhancing interpretability while minimizing information loss. Half of the features are selected after PCA. In the fourth step, selected features are properly tuned using cross-validation techniques.

Finally, various classification techniques based on machine learning approaches are deployed to assess the classification accuracy between lymphocytes and lymphoblasts.

The remainder of the paper is outlined as follows: Section 2 addresses the literature review. Section 3 summarizes the proposed methodology, experimental setup, and data collection procedure. Section 4 presents the results of the experiment. Section 5 discusses these results. The conclusion and future work are outlined in Section 6.

2. RELATED WORK

In 2005, Scotti introduced a system for recognizing blast or normal lymphocytes. This paper categorizes blasts and regular cells by a K-nearest neighbor (KNN) classifier based on geometric features. Later, accuracy is tested on Linear and feed-forward neural networks (NN) [6]. A method to categorize WBCs based on nucleus information was proposed by Theera-Umpon and Dhompongsa [1], in which the pattern spectrum of each nucleus is determined. Initially, the area and high location of the spectrum are taken as the two aspects of two selected granulometric measurements. Classifiers based on neural networks and Bayes theorem are employed. In 2008, Adollah et al. [2] examined the procedure of blood cell segmentation to diagnose various diseases and treat pathological conditions. Another way to segment white blood cells is a circular histogram-based Otsu algorithm. A higher degree of entropy is a feature for threshold over a two-dimensional histogram, and this feature is present on two color models: RGB and HSI. When it comes to segmentation, several distinct approaches, including gray-level thresholding, morphological processes, a variety of filtering strategies, color matching, and color thresholding, are evaluated and ranked based on their effectiveness. Differently, the shape information is first found by binarization after the maximum intensity is made, and then it is utilized. Finally, seeded watershed and gradient vector flow methods are employed for cell segmentation and identification.

Sadeghian et al. [3] suggested a segmentation approach to isolate leukocytes from other blood components. The images were first converted to grayscale, and then the WBCs were imaged using a sub-imaging approach. To locate the nuclei, a gradient vector flow model is utilized. The nucleus is created using hole filling after the cytoplasm has been segmented using Zack thresholding, and the nucleus has been removed from the grayscale image. The suggested method is 92% accurate at segmenting the nucleus and 78% accurate at segmenting the cytoplasm. To classify white blood cells into distinct kinds in 2010, Tabrizi et al. [4] collected morphological, textural, and color data from segmented nuclei and cytoplasm. They used them to train learning vector quantization neural networks. The classifications are accurate in 96% of cases. Later, in 2011, Rezatofighi et al. [7] developed a technique that recognizes five distinct blood cell types with three stages each. To differentiate basophils, the nucleus is initially segmented using the Gram-Schmidt approach, which is based on the co-occurrence matrix of characteristics and the local binary pattern. The snake approach is used to segment the cytoplasm in the second stage of image processing after the images have been preprocessed in grayscale and the S component of the HSI color model. In this stage, the images are also preprocessed using the HSI color model. In conclusion, the remaining four subtypes of

WBC are categorized according to their morphological characteristics.

Abd Halim et al. [8] provided a novel approach to the nucleus segmentation problem to distinguish between acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) for acute leukemia. All the images are projected onto the HSI color space to enhance the global contrast and make the categorization zone more visible. The S component from the HSI color space is used to segment the nucleus, and a defined threshold is then applied for more precise segmentation. Region-growing techniques are employed to identify blasts and nuclei. Nee et al. [9] utilized the S component of the HSV color model. This method, which is used for AML and its subtypes, achieves a 94.5 percent accuracy rate and relies on edge segmentation for the watershed transformation and edge identification using erosion, dilation, and magnitude gradient. Pan et al. [10] segment leukocytes using the extreme learning machine (ELM). The peak gradient pixel is selected for sampling using a gradient threshold. The multi-colored object is then examined for segmentation using maximal entropy. Leukocyte images are converted to HSI before the cytoplasm is extracted using the Otsu process. The edge pixels on ELM are considered as categories.

Madhloom et al. [11] extracted the H and S bands from RGB images and converted them to binary. Fifteen disk-shaped structuring elements were utilized to open the H band and erode the S band. Subsequently, the images were reconstructed using a morphological operator to enable the classification of blast cells based on their shape's centroid and axis length. This approach effectively segmented the lymphoblasts, localized them, and achieved a 100% accuracy rate. Jagdeesh et al. [12] proposed a method in 2013 for identifying cancer cells in blood samples. The image is initially converted into binary and then into grayscale form. Subsequently, morphological closure, erosion, and a map depicting the difference between black and white pixels are employed to remove smoothness and distortion from the image. Data segmentation is facilitated by the watershed transformation. When classifying data with SVM, geometric, statistical, and textural considerations are all considered. In comparison, Joshi et al. [13] utilized the Otsu threshold technique for segmentation and the KNN algorithm for classification based on a combination of morphological and textural characteristics of the images.

Blood smear images can be segmented using k-means clustering, as proposed by Mahopatara et al. [14] in 2014. The method extracts RGB color information from the full image before segmenting it using the leukocytes-shadowed C-means algorithm. The nucleus, cytoplasm, and backdrop are separated from the sub-image using color space SCM clustering. Features based on morphology, texture, and color are extracted, normalized, and selected. The classification method employs various classifiers, including radial basis function, naive Bayesian, multilayer perceptron, support vector machines, neural networks, and KNN.

Kulkarni-Joshi and Bhosale [15] proposed a thresholding-based technique for detecting ALL blasts and segmenting nuclei. After removing the background, Otsu thresholding is applied to extract form features for blast identification. They introduced a method for segmenting lymphocyte nuclei to identify leukemia in 2014. To achieve accurate segmentation of nuclei and diagnosis of leukemia with a success rate of 96.5%, the Otsu method is used to obtain nuclei, followed by

the removal of small segments and dilation of the nuclei. This process is repeated iteratively until the desired results are achieved. Additionally, to reduce the high RGB values in the image, the first step involves convolving it with a $2 \times 2/6$ mask.

Vidhya et al. [16] segmented the data using local directional features (LDP) and k-means clustering in 2015. They employed k-means clustering to separate lymphoblasts into various groups in the same year. Subsequently, they gathered statistical and geometrical information to identify ALL and its subtypes, and then they utilized multiclass support vector machines (SVM) to analyze the data. Conversely, the author bases the categorization solely on the characteristics of the nucleus, claiming this method achieves 97% accuracy. Goutam et al. [17] proposed a framework for detecting whether cells are affected by acute myeloid leukemia or are normal by segmenting cell nuclei using grayscale images and k-means clustering. This method was designed to determine whether the cells were impacted by AML or were normal. When employing SVM for classification, LDP with textural features achieves an accuracy of 98%. Bhattacharjee and Saini [18] presented a method for diagnosing acute lymphoblastic leukemia in 2015. To segment the data, watershed transformations were used, followed by morphological procedures. When collecting morphological characteristics for use in a Gaussian mixture model and a binary search tree for classification, the suggested strategy achieves an accuracy of 95.56%.

Using a deep belief network (DBN), Bibin et al. [19] proposed a methodology in 2017 for identifying malaria parasites. The primary objective is to distinguish between parasites and non-parasites. They developed an HSV color space conversion model, partitioned the cells into regions, and then employed DBN for classification using color and texture attributes. Rawat et al. [20] also in 2017, utilized histogram equalization, global thresholding, and morphological patterns to segment lymphoblastic cells. After subtracting this from the preprocessed image, the cytoplasm is obtained. Geometrical, chromatic, statistical, and other characteristics are extracted for classification and fed into various classifiers with PCA.

Although hierarchically applying PCA-ANFIS achieved a maximum accuracy of 97.6%, employing too many classifiers in this combination leads to lengthy processing times.

Mughal et al. [21] proposed an innovative method to remove the pectoral muscle from mammography images. Amjad et al. utilized images of stained bone marrow in 2018 to classify ALL into several subtypes. Using a convolutional neural network and deep learning techniques, the model was trained on bone marrow images to achieve accurate classification results. This allowed for a comparison of the results with those obtained from naive Bayesian, KNN, and SVM classifiers. The experiments demonstrated that the proposed approach resulted in an accuracy of 97.78% [22]. Hegde et al. [23] utilized an SVM classifier in 2020 to classify WBCs as normal or abnormal. NN classifiers were employed to subdivide typical white blood cells into five distinct types. The overall classification accuracy was increased to 98.8% by using both NN and SVM. In 2022, Chand and Vishwakarma [24] introduced a unique deep learning framework for diagnosing ALL based on a CNN. The group developed this framework. When compared to the 41,626 free-tuning parameters in the suggested architecture, the trainable parameters of pre-trained complex networks such as AlexNet, VGG-Net, and ResNet-50 are significantly lower in number.

Table 1 describes various public and private data sets used in the previous study. The entire literature is divided based on region of interest. Various segmentation techniques and classification techniques are used to isolate the region of interest. Literature suggests three types of classification standards. In the First category, authors [1, 8, 9] presented segmentation procedures to isolate nuclei and cytoplasm elements. This category shows accuracy variation ranging from 77% to 95%. In the second category, the authors [4, 5, 7, 24] are classifying five types of white blood cells. The accuracy of this category classification is restricted to 98.62%, and finally, in the third category, researchers [6, 10-12, 20, 23, 25, 26] presented various techniques to classify between centroblast and centrocytes. In this category, accuracy ranges from 89.8 to 98%.

Table 1. Summary of the related literature

Ref.	Dataset	Segmentation & Feature Extraction	Region of Interest	Classification Approaches	Performance Evaluation
[1]	Gray Scale bone marrow Images collected from the University of Missouri Ellis Fischel Cancer Center	Nuclei based features	Extraction of Nuclei	Bayes-based classification followed by ANN	77% of classification is achieved while testing
[4]	251 blood slide containing 302 WBC with magnification 100X and resolution 720X576 pixel acquired form Camera-Sony-Model No. SSC-DC50AP.	Gram-Schmidt and active contour algorithms are used to segment nuclei and cytoplasm followed by PCA for feature selection	Five types of white blood cells in peripheral blood	LVQ neural network is used to classify peripheral blood	Overall, 96% classification accuracy is achieved.
[5]	300 microscopic images of bone marrow with 100X resolution at Amreek Clinical Laboratory Swat KP Pakistan.	A method based on thresholding after changing the color space from RGB to HSV CNN used to extract features	ALL into its sub-type and reactive bone marrow	Convolutional Neural Network is used for classification	97.78% accuracy is achieved
[6]	113 images contain about 8400 blood cells, collected from M. Tettamanti Research Center for Childhood Leukemias and Hematological Diseases, Monza, Italy.	Adaptive prefiltering and segmentation approaches	Centrocytes vs centroblast	linear Bayes Normal classifier, KNN, and Feedforward neural network is used	The mean square error rate are 0.040, 0.0267, and 0.0133 respectively

[7]	251 blood slide containing 302 WBC with magnification 100X and resolution 720X576 pixel acquired from Camera-Sony-Model No. SSC-DC50AP	The textural parameter is extracted by LBP and GLCM	Five types of white blood cells in peripheral blood	SVM classifier	Overall, 93.09% accuracy is achieved
[8]	ALL and AML images are used	A global contrast stretching technique followed by segmentation in HSV color space Extraction of the nucleus and cytoplasm is achieved using features of the neighboring pixels as contextual information to generate homogeneous regions	Identify nucleus region in WBC	No classification techniques are implemented	No evaluation criterion is evaluated
[9]	663 bone marrow leukemia images from the Mexican Social Security Institute at 100X resolution	Extraction of the nucleus and cytoplasm is achieved using features of the neighboring pixels as contextual information to generate homogeneous regions	Isolation of nucleus and cytoplasm	Markov Random Field	Cell separation segmentation accuracy is 95%. While classifying 95% in the diagnosis of leukemia families and 90% in the diagnosis of leukemia
[10]	Leukocyte image	ELM based segmentation	Leukocyte image segmentation	No classification approaches are deployed	Overall, Error using watershed algorithm 0.2797, SVM based ELM based 0.1567
[11]	180 microscopic blood sample	Color feature with morphological reconstruction	Isolation of lymphoblast from microscopic images	No classification criterion is implemented	Isolation of lymphoblast from microscopic image achieved 100% accuracy
[12]	54 samples of size 1024X1024 are collected Department of Hematology, Ispat General Hospital, Rourkela, India	K-means clustering in RGB color space and extraction of cytological components	Lymphocyte vs lymphoblast classification	The ensemble model is used in conjunction with RBF, MLP, KNN, NB	Sensitivity for the ensemble model is 94.93% and specificity is 95%
[20]	ALL IDB Dataset of 260 images	Feature-based algorithm to extract nuclei and cytoplasm followed by PCA for dimensional reduction	Lymphocyte vs lymphoblast classification	Smooth support vector machine, KNN, probabilistic neural network, neuro-fuzzy inference system	Maximum accuracy achieved is 94.6% by PCA-SVM
[23]	1159 images taken from Leishman stained peripheral blood smears	Brightness and color-based features are extracted	Normality between normal and abnormal cell	A combination of NN and SVM classifiers is used	98.8% of accuracy is achieved
[24]	ALL IDB	Deep Learning (CNN) for feature extraction	Diagnosing ALL and their subtypes	CNN-based deep network with a smaller (41626) number of parameters	Experiment A: 98.62% accuracy achieved Experiment B: 97.73% accuracy
[25]	ALL IDB dataset	Histogram equalization to obtain morphological and textural features	ALL vs Healthy	SVM along with the Gaussian radial basis kernel function	The accuracy achieved by the algorithm is 93% while sensitivity is 98%

3. PROPOSED METHODS

The proposed method is divided into seven steps covering data collection, preprocessing, and data augmentation, transfer learning approaches for feature extraction, principal component analysis for feature selection/transformation, grid-search cross-validation method for hyperparameter tuning, training of individual best models, and soft voting classification for prediction. The proposed method of this research paper is shown in Figure 3.

The Proposed soft voting ensemble model is based on four classifiers: random forest, logistic regression (LR), decision tree (DT), and support vector machine (SVM). The hyperparameters of LR, SVM, DT, and RF were properly tuned using 5-fold cross-validation (CV).

3.1 Data collection

ALL-IDB data set consists of normal and lymphoblast cells clipped with their area of interest. There are 260 images containing 130 images of lymphoblasts and 130 images of lymphocytes [6]. All images are true to colour, with a size of $257 \times 257 \times 3$.

3.2 Preprocessing and data augmentation

Preprocessing is the technique to enhance image quality and noise reduction. All images pass by a low pass average filter for noise reduction. Data augmentation helps to train a model with different aspects/angles of the image—parameters used in the data augmentation layer, as shown in Table 2. A total of 14677 images were generated using this technique, which were further used for feature extraction [25].

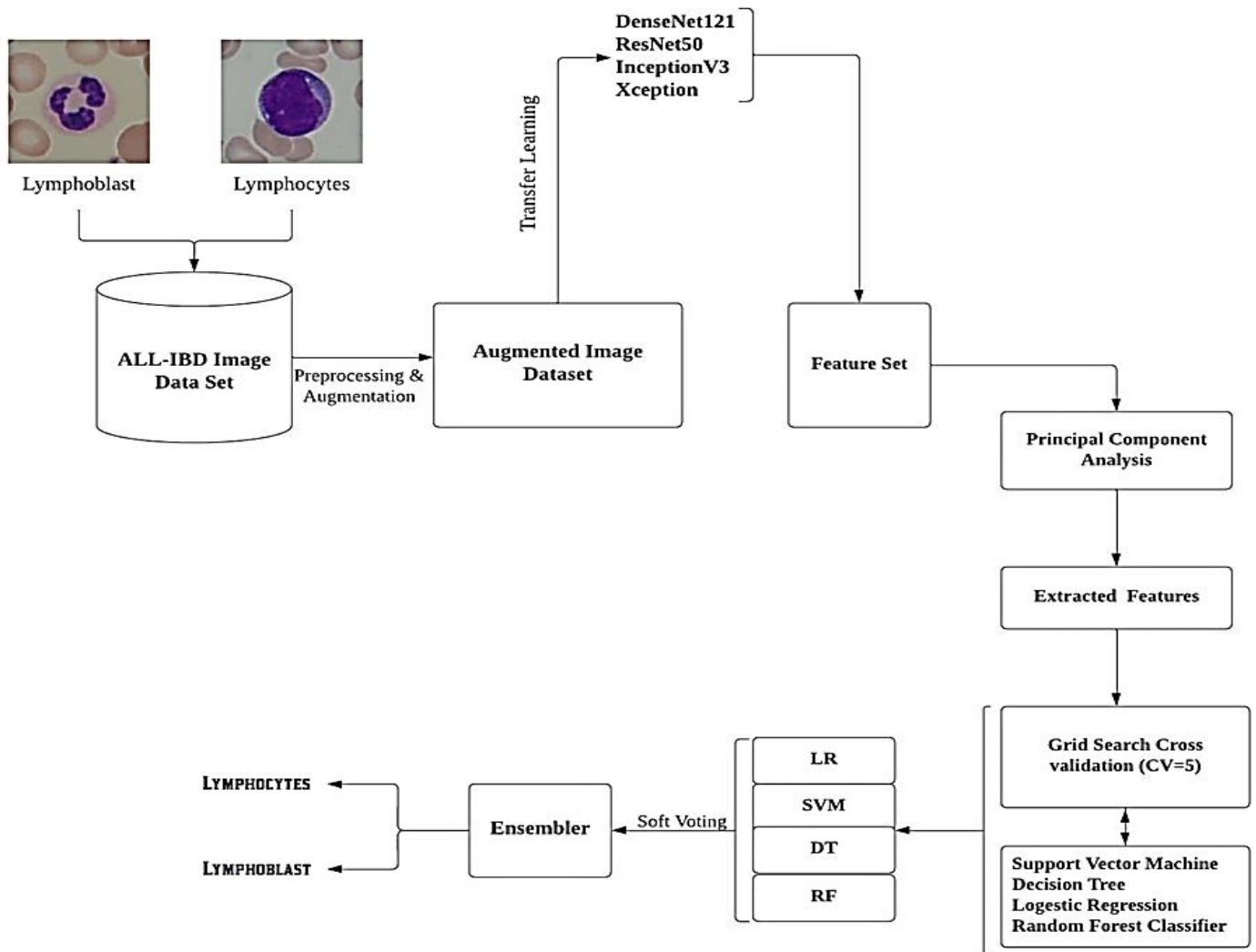


Figure 3. Proposed method for lymphocytes and lymphoblast classification

Table 2. Data augmentation layer parameters

Parameter Name	Values
Range of Rotation	45
The range for the shift in width	0.2
The range for the shift in height	0.2
Range for Zoom	0.1
Shear range	0.1
Image with horizontal flip	True
Image with a vertical flip	True
Fill model in image transformation.	Nearest
Feature Wise center	True
Feature Wise Standard Normalisation	True

3.3 Transfer learning

Transfer learning is the method that uses pre-trained convolutional neural networks directly [27, 28]. Four convolutional neural networks were used to extract the features in this step, namely DenseNet121, ResNet50, InceptionV3, and Xception. These features are saved in .csv file and used in further processes. Feature extraction using transfer learning methods is shown in Figure 4.

Each layer of DenseNet121 is connected to the others by a feed-forward method. When using dense networks, the vanishing-gradient problem in DenseNet121 is resolved,

feature propagation is improved, feature reuse is promoted, and there are fewer parameters. A total of 1024 features are recovered by using the DenseNet121 algorithm. The ResNet50 refers to a convolutional neural network with 50 layers. This network's first training was conducted using more than a million photos from the ImageNet database, each of which represented one of a thousand distinct object categories. As a direct consequence of this, the network now possesses feature representations that are exhaustive for various image types. When tested on the ImageNet database, the image recognition model InceptionV3 has shown that it is capable of achieving an accuracy of more than 78.1%. The model's structural components, including convolutions, average pooling, max pooling, concatenations, dropouts, and link layers, are constructed in a manner that is both symmetric and asymmetric. Batch normalization is often applied to the activation inputs by the model, and softmax is the method that is used to calculate the loss. Approaches like regularisation, factorized convolutions, and parallel processing with dimension reduction are employed in InceptionV3. *Finally*, an Xception neural network with 71 layers categorized images into 1000 classes. The extracted feature size using these pre-trained neural networks is shown in Table 3.

By combining all features, a new feature set of size 14677×7168 has been generated.

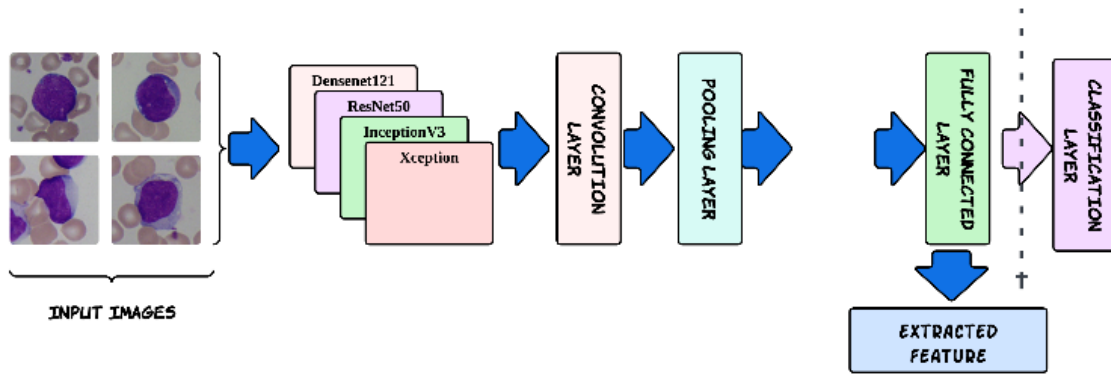


Figure 4. Feature extraction using transfer learning techniques

Table 3. Extracted feature size using a pre-trained neural network

Pre-trained Neural Networks	Number of Layers	Layering Weight	Output Feature Size
DenseNet121	427	604	14677×1024
ResNet50	175	318	14677×2048
InceptionV3	311	376	14677×2048
Xception	132	234	14677×2048

3.4 Principal component analysis

The dimensions of the feature vector significantly impact the effectiveness of machine learning algorithms. While large datasets are more common than ever, they can often be challenging to interpret. Principal Component Analysis (PCA) is a method that can reduce the number of dimensions in such datasets, improve interpretability, and mitigate information loss. This is achieved by generating progressively more uncorrelated variables to maximize the variance [29]. In this study, PCA is utilized to select fifty percent of the characteristics.

The choice of PCA over other techniques is motivated by the high dimensionality of features from deep networks. PCA provides better visualization than other neural-based architectures, making it highly desirable to apply PCA before feeding the data into machine learning architectures.

3.5 Parameter tuning

A classification model's performance mainly depends on the number of features used to build the model. Each model may have a different number of features, which need to be tuned properly for better performance, as seen in Figure 5. This paper uses four classification algorithms, namely LR, SVM, DT, and RF classifiers.

Logistic regression(LR) classification uses the sigmoid function, as shown in Eq. (1), as a logistic function, sometimes known as the sigmoid function.

$$f(n) = \frac{1}{(1+exp^{-n})} \quad (1)$$

where, exponentiation, or exp, serves as the foundation for natural logarithms, and n represents the true numerical value. According to the second line of Eq. (2), the value of the output can be predicted by the linear combination of the values of the inputs using weights. In contrast to linear regression, binary regression models the output value as a binary digit, which can only take on one of two possible values: 0 or 1.

SVM classifier is a popular machine learning supervised learning model [30]. It attempts to separate all data points of various classes. The samples fit within a segment that is divided by the hyperplane. As a result, it is possible to identify the category to which the sample belongs.

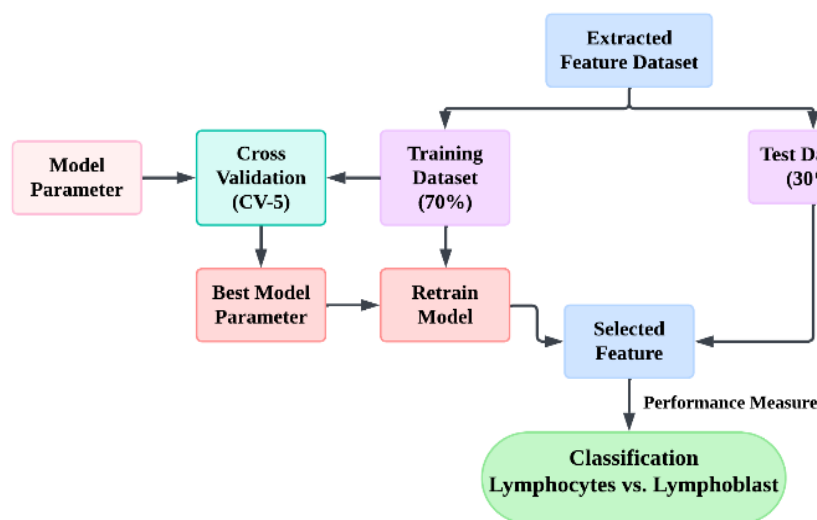


Figure 5. Hyperparameter tuning for classification

Table 4. Parameters and their values for 5-fold cross-validation

Classification Model	Parameter Variables (Variable)	Parameter Range Values
Logistic Regression (LR.)	Regularisation parameter strength (C)	[0.01,0.1, 1, 10, 100, 1000]
	norm of the penalty(penalty)	['None', 'l1', 'l2']
	Algorithm to use in the optimisation problem (solver)	['lbfgs', 'liblinear', 'sag', 'saga']
Support Vector Machine (SVM)	Regularisation parameter strength (C)	[0.01,0.1, 1, 10, 100, 1000]
	Kernel Coefficient(gamma)	[1000,100,10,1, 0.1, 0.01, 0.001, 0.0001]
	Types of Kernel (kernel)	['rbf','linear']
Decision Tree (DT.)	Quality of split measure function(criterion)	['gini','entropy']
	Maximum depth of tree(max_depth)	[4,5,6,7,8,9,10,11,12,15,20,30,40,50,70,90,120,150]
	Number of tree(n_estimators)	[25,50,100,150,200,500]
Random Forest(RF.)	Best Split Maximum Features (max_features)	['auto', 'sqrt', 'log2']
	Maximum depth of tree(max_depth)	[4,5,6,7,8,10,15,20]
	Quality of split measure function(criterion)	['gini', 'entropy']

A decision tree (DT) is a predictor, $h: X \rightarrow Y$, that works by moving from a tree's root node to a leaf to forecast the label associated with an instance x . For the sake of simplicity, we concentrate on the binary classification scenario, where $Y=0$ or 1 ; however, decision trees can also be used for other types of prediction issues [31]. The successor child is determined at each node along the root-to-leaf path by splitting the input space. The splitting is typically based on one of x 's characteristics or a predetermined set of splitting rules. A unique label can be found on a leaf.

RF is a widely used supervised machine learning classification technique. It substitutes a new training subset for the sample training data, and the outcome is determined by majority voting and averaging [31]. The steps involved in the RF algorithm are as follows:

Step 1: In RF, n records are chosen randomly from a set of k records in the data.

Step 2: The decision tree is built separately for each sample.

Step 3: Each DT is analyzed, and the output of each decision tree is generated.

Step 4: A voting or averaging procedure is applied for classifying the data, and the final classification result is recorded.

Parameters of these models are tuned using grid-search 5-fold cross-validation. All the necessary parameters and their values are used to find the best estimator for a given dataset, as shown in Table 4.

3.6 Voting classifier

The majority voting classifier has been used to combine the final predictions of the above classifiers. A Voting classifier uses a majority to favour a decision, as shown in Figure 6. Here final predictions of classifiers LR, SVM, DT, and RF will be combined using soft voting.

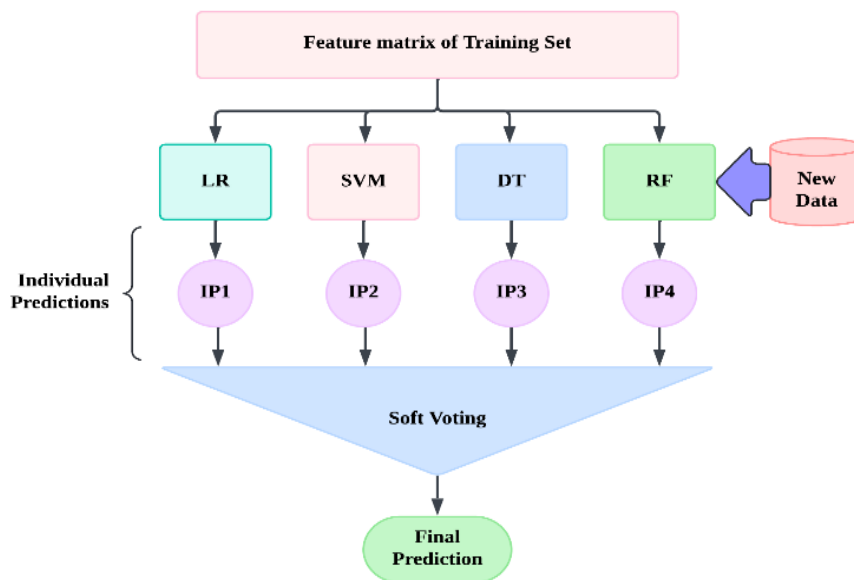


Figure 6. Ensemble method for final prediction

4. EXPERIMENTED RESULTS

All tests were conducted on a computer equipped with a 12th Generation Intel Core i7-12700H processor, with a maximum turbo speed of up to 4.7 GHz, 14 cores, 16 GB of RAM, a dedicated NVIDIA GeForce RTX 3050 Ti graphics card with 4 GB of GDDR6, and 64-bit Windows 10 Home. To

train the pre-trained models, the data was randomly divided into two sets: a 70% training set and a 30% testing set. As shown in Table 5, calculations for Accuracy, Precision, Recall, F1-Score, and ROC-AUC were performed on each feature dataset extracted through the process of transfer learning and modified using principal component analysis.

Let TP, TN, FP, and FN represent the number of samples

classified as true positives, true negatives, false positives, and false negatives, respectively. Calculating and formulating Accuracy, Precision, Recall, and F1-Score are all feasible.

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (6)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

The observation indicates that the combined features perform better than individual isolated features, as demonstrated in Table 5.

Table 5. Detailed performance metric comparison

Pre-Trained Neural Networks (1 for Feature Set Extraction)	Number of Features Extracted/Used	Classifier LR KNN SVM DT RF	Best Parameter (After Cross Validation)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DenseNet121	1024	LR	'C': 0.1, 'penalty': 'l2', 'solver': 'lbfgs'	96.46	96.78	96.17	96.48
		SVM	'C': 100, 'gamma': 0.0001, 'kernel': 'rbf'	98.89	98.96	98.82	98.89
		DT	criterion='entropy', max_depth=5	79.47	85.73	76.27	80.72
		RF	'criterion': 'entropy', 'max_depth': 8, 'max_features': 'auto', 'n_estimators': 50	86.72	89.72	84.69	87.13
		ENSEMBLE	'voting': 'Soft'	98.23	98.46	98.02	98.24
ResNet50	2048	LR	'C': 1, 'penalty': 'l2', 'solver': 'lbfgs'	89.49	90.58	88.69	89.63
		SVM	'C': 100, 'gamma': 0.001, 'kernel': 'rbf'	91.71	92.48	91.12	91.80
		DT	criterion='entropy', max_depth=5	74.89	87.23	70.04	77.69
		RF	'criterion': 'gini', 'max_depth': 8, 'max_features': 'sqrt', 'n_estimators': 200	73.59	82.02	70.28	75.69
		ENSEMBLE	'voting': 'soft'	90.49	92.35	89.08	90.68
InceptionV3	2048	LR	'C': 1000, 'penalty': 'l2', 'solver': 'liblinear'	84.33	85.33	83.73	84.52
		SVM	'C': 100, 'gamma': 1000, 'kernel': 'linear'	83.81	85.24	82.94	84.07
		DT	criterion='gini', max_depth=11	64.08	63.99	64.23	64.11
		RF	'criterion': 'gini', 'max_depth': 8, 'max_features': 'auto', 'n_estimators': 100	78.52	81.34	77.08	79.15
		ENSEMBLE	'voting'='soft'	83.42	84.65	82.70	83.66
Xception	2048	LR	'C': 1000, 'penalty': 'l1', 'solver': 'liblinear'	80.2	82.7	78.84	80.73
		SVM	'C': 1000, 'gamma': 1000, 'kernel': 'rbf'	82.65	84.96	81.28	83.08
		DT	criterion='entropy', max_depth=7	64.35	60.6	65.65	63.02
		RF	'criterion': 'entropy', 'max_depth': 8, 'max_features': 'sqrt', 'n_estimators': 200	77.61	77.99	77.5	77.74
		ENSEMBLE	'voting'='soft'	80.79	82.11	80.08	81.08
Combined All Features	7168	LR	'C': 0.1, 'penalty': 'l2', 'solver': 'liblinear'	97.68	97.87	97.52	97.69
		SVM	'C': 10, 'gamma': 0.0001, 'kernel': 'rbf'	98.98	99.32	98.65	98.98
		DT	criterion='entropy', max_depth=12	79.29	81.3	78.25	79.74
		RF	'criterion': 'entropy', 'max_depth': 8, 'max_features': 'sqrt', 'n_estimators': 200	82.7	78.53	85.76	81.99
		ENSEMBLE	'voting'='soft'	98.23	98.82	97.67	98.24

The proposed ensemble learning techniques outperform the decision tree and random forest classifiers. Figure 7 illustrates the comparison of the area under the receiver operating characteristics curve (AUC) for features extracted using transfer learning methods.

An AUC value of 1.00 was achieved with SVM and the ensemble model using DenseNet121, as well as with the

combination of all feature sets. Logistic regression (LR) also achieved an AUC of 1.0 when combining all features, while with DenseNet121 features, the AUC was 0.99. The minimum AUC achieved by the decision tree (DT) is 0.64, and with random forest (RF), it is 0.86. The AUC with the combined feature set is 0.80 with DT and 0.91 with RF.

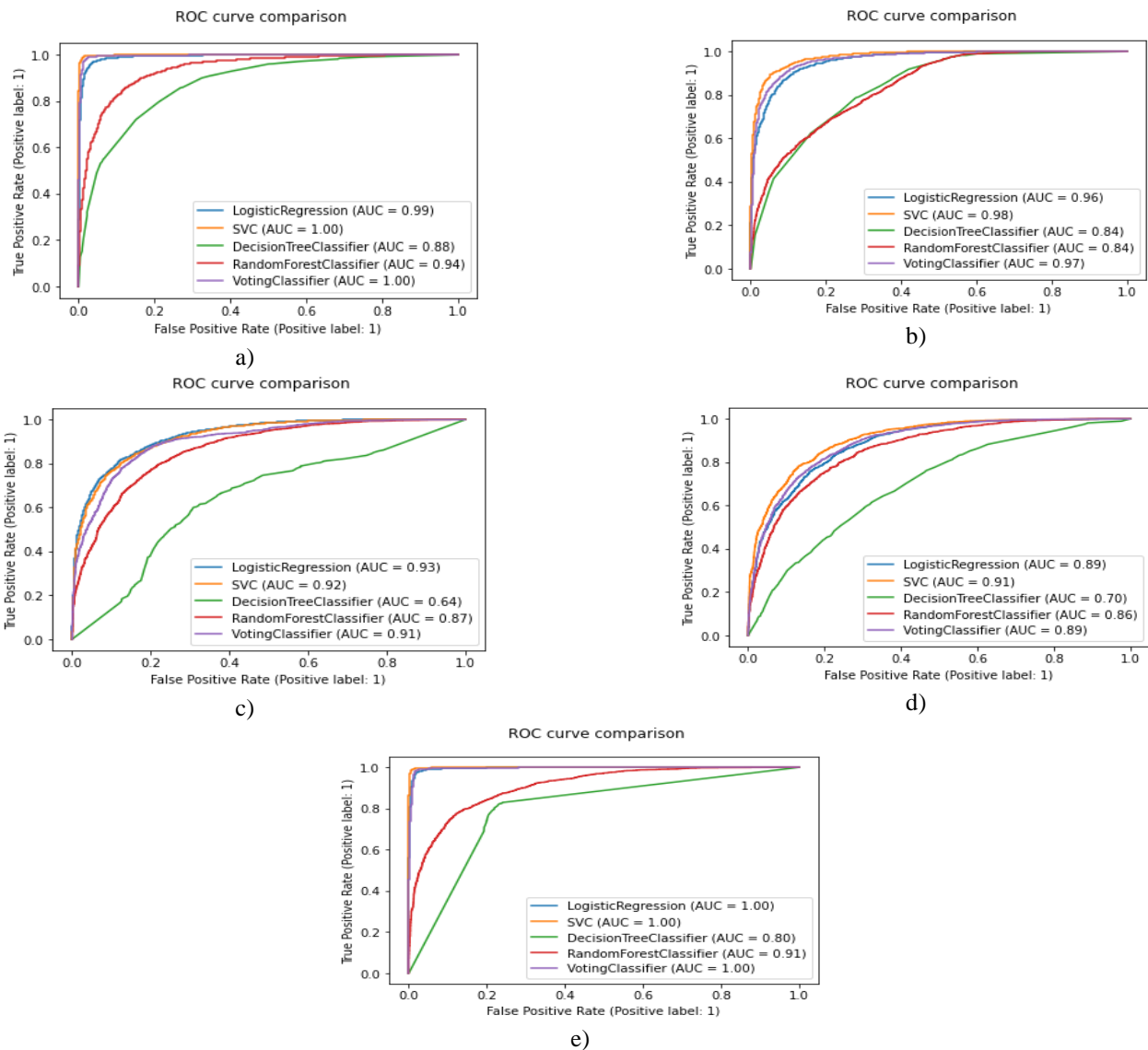


Figure 7. ROC-AUC comparison for a) DenseNet121, b) Resnet50, c) InceptionV3, d) Xception, and e) all features together

5. DISCUSSION

Our implemented work extracted the features from four different deep convolutional neural networks (DenseNet121, ResNet50, InceptionV3, and Xception) and the combination of these features form a new feature set. The best performance is achieved by SVM with DenseNet121 features, reaching 98.89%, outperforming all classifiers on the other feature sets. LR achieves 96.46% accuracy with DenseNet121 features, slightly lower than the performance on the combined feature set. The performance of LR and the ensemble model improves when combining all features compared to individual feature sets. DT and RF exhibit lower performance than LR, SVM, and ensemble models but are still satisfactory, with a minimum accuracy of 64.08% and 73.59%, respectively, and

a maximum of 79.47% and 86.72%, respectively. These classifiers perform particularly well on DenseNet121 features compared to other feature sets. Additionally, SVM achieves the best precision score of 99.32% with all features combined, the highest recall score of 98.82% with DenseNet121 features, and an F1 score of 98.98% with all features combined. Figure 8 provides a visual comparison of the accuracy of various classification methods using different feature extraction algorithms based on transfer learning methods.

Table 6 compares the proposed method for lymphocytes and lymphoblast classification with the previous state-of-the-art. The result shown in Table 5, Figure 7, and Figure 8 highlight that the proposed method can be effective compared to the previous state-of-the-art.

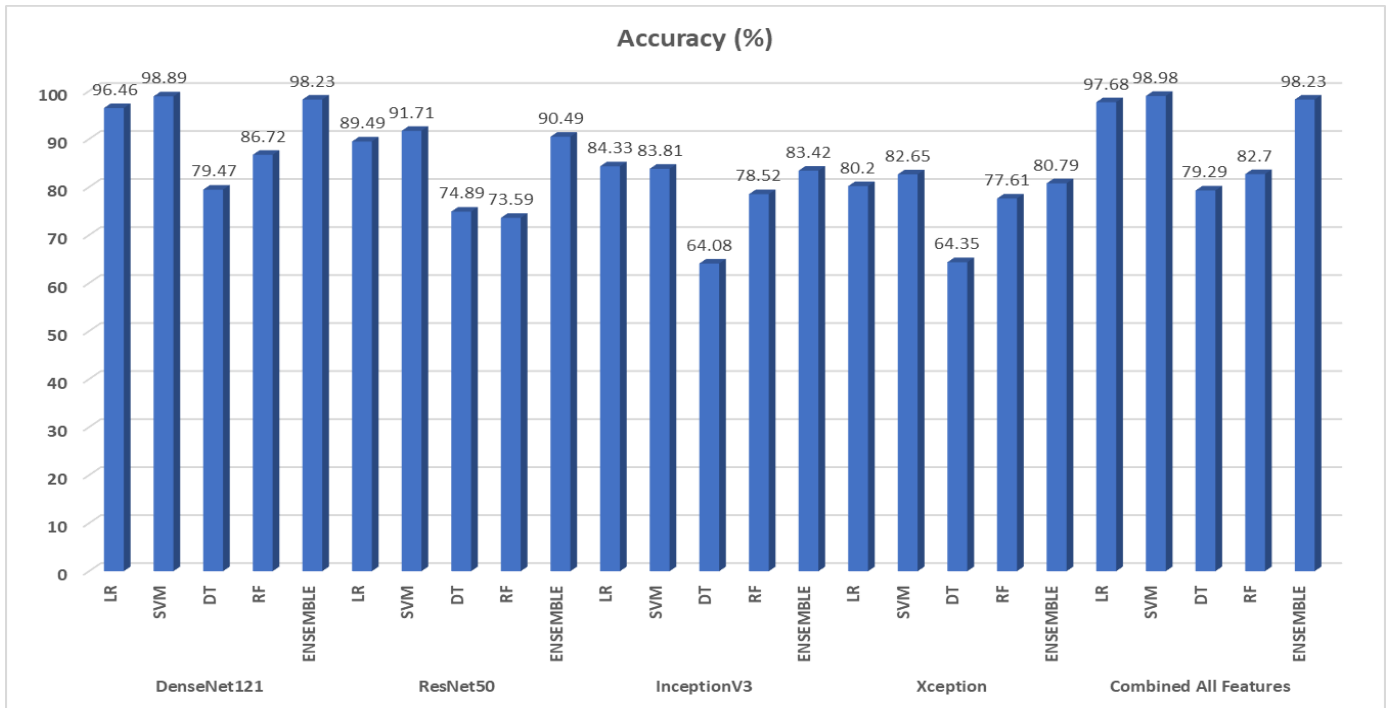


Figure 8. Accuracy graph based on different transfer learning approaches deployed for feature extraction vs machine learning methods used for ALL-IBD data set classification

Table 6. Performance analysis from the previous state-of-the-art

References	Proposed Features	Classification Techniques	Test Data Set	Accuracy
[20]	Morphological feature	PCA with kNN, PNN, SVM, SSVM, and ANFIS	260	97.6
[22]	CNN features	Deep Neural Networks	330	97.78
[25]	Morphological and Textural features	SVM	245	92%
[32]	Texture, as well as shape and colour	SVM	267	92%
[33]	Morphological and Textural features	SVM	196	89.8
[34]	Morphological and Textural features	SVM	260	92.3
[35]	shape feature	ANN	120	95.2
[36]	colour, shape, and texture	Dempster-Shafer	180	96.7
[37]	Shape and texture feature	SVM	21	97
[38]	color, shape, and texture	Fuzzy System	108	98
[39]	Texture and shape based	Neural Network	108	97.2
Proposed Model	Features from transfer learning and selection with PCA	Ensembled Network	260	98.23

6. CONCLUSIONS

The research work has presented a comprehensive architecture based on transfer learning approaches for classifying Lymphocytes and Lymphoblastic cells with an ensemble model. Based on the results obtained, it is evident that the proposed method with an ensemble learning model, combining all features, can effectively be used for ALL-IBD classification. LR, SVM, and the ensemble model with DenseNet121 and combined features outperform previous state-of-the-art methods. Additionally, the performance of DT and RF classifiers is satisfactory. Combining features from DenseNet121, ResNet50, InceptionV3, and Xception networks provides better and more robust features. PCA plays a crucial role in selecting only half of the features, reducing complexity while maintaining high performance. The proposed architecture holds significant clinical implications for leukemia patients, aiding pathologists, and laboratory professionals in accurate diagnosis.

Furthermore, exploring other feature extraction and selection methods such as nature-inspired techniques like genetic algorithms and PSO may further enhance the model's performance by selecting the most suitable feature set.

Additionally, incorporating morphological and textural features alongside the proposed model could potentially improve system performance. While automated analysis offers significant advantages, it is essential to recognize the continued importance of human expertise, especially in complex decision-making scenarios. Integrating computerized analysis with human insight can enhance the effectiveness and thoroughness of diagnostic processes across various fields.

FUNDING

This research was funded by Taif University, Saudi Arabia, Project No. (TU-DSPP-2024-52).

ACKNOWLEDGMENT

The authors extend their appreciation to Taif University, Saudi Arabia, for supporting this work through project number (TU-DSPP-2024-52).

REFERENCES

- [1] Theera-Umpon, N., Dhompongsa, S. (2007). Morphological granulo metric features of nucleus in automatic bone marrow white blood cell classification. *IEEE Transaction on Information Techonology in Biomedicine*, 11(3): 353-359. <https://doi.org/10.1109/TITB.2007.892694>
- [2] Adollah, R., Mashor, M.Y., Mohd Nasir, N.F., Rosline, H., Mahsin, H., Adilah, H. (2008). Blood cell image segmentation: A review. In 4th Kuala Lumpur International Conference on Biomedical Engineering 2008: BIOMED 2008 25–28 June 2008 Kuala Lumpur, Malaysia, pp. 141-144. https://doi.org/10.1007/978-3-540-69139-6_39
- [3] Sadeghian, M.H., Keramati, M.R., Badiei, Z., Ravarian, M., Ayatollahi, H., Rafatpanah, H., Daluei, M.K. (2009). Alloimmunization among transfusion-dependent thalassemia patients. *Asian Journal of Transfusion Science*, 3(2): 95-98. <https://doi.org/10.4103/0973-6247.53884>
- [4] Tabrizi, P.R., Rezatofighi, S.H., Yazdanpanah, M.J. (2010). Using PCA and LVQ neural network for automatic recognition of five types of white blood cells. In 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, pp. 5593-5596. <https://doi.org/10.1109/IEMBS.2010.5626788>
- [5] Rehman, A., Abbas, N., Saba, T., Rahman, S.I.U., Mehmood, Z., Kolivand, H. (2018). Classification of acute lymphoblastic leukemia using deep learning. *Microscopy Research and Technique*, 81(11): 1310-1317. <https://doi.org/10.1002/jemt.23139>
- [6] Scotti, F. (2005). Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images. In CIMS.A. 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, Messian, Italy, pp. 96-101. <https://doi.org/10.1109/CIMS.A.2005.1522835>
- [7] Rezatofighi, S.H., Khaksari, K., Soltanian-Zadeh, H. (2011). Automatic recognition of five types of white blood cells in peripheral blood. *Computerized Medical Imaging and Graphics*, 35(4): 333-343. <https://doi.org/10.1016/j.compmedimag.2011.01.003>
- [8] Abd Halim, N.H., Mashor, M.Y., Nasir, A.A., Mokhtar, N.R., Rosline, H. (2011). Nucleus segmentation technique for acute leukemia. In 2011 IEEE 7th International Colloquium on Signal Processing and Its Applications, pp. 192-197. <http://doi.org/10.1109/CSPA.2011.5759871>
- [9] Nee, L.H., Mashor, M.Y., Hassan, R. (2012). White blood cell segmentation for acute leukaemia bone marrow images. *Journal of Medical Imaging and Health Informatics*, 2: 278-284. <https://doi.org/10.1166/jmihi.2012.1099>
- [10] Pan, C., Park, D.S., Yang, Y., Yoo, H.M. (2012). Leukocyte image segmentation by visual attention and extreme learning machine. *Neural Computing and Applications*, 21: 1217-1227. <https://doi.org/10.1007/s00521-011-0522-9>
- [11] Madhloom, H.T., Kareem, S.A., Ariffin, H. (2012). An image processing application for the localization and segmentation of lymphoblast cell using peripheral blood images. *Journal of Medical Systems*, 36: 2149-2158. <https://doi.org/10.1007/s10916-011-9679-0>
- [12] Jagadeesh, S., Nagabhooshanam, E., Venkatachalam, S. (2013). Image processing based approach to cancer cell prediction in blood samples. *International Journal of Technology and Engineering Sciences*, 1(1): 1-10.
- [13] Joshi, M.M.D., Karode, A.H., Suralkar, S.R. (2013). White blood cells segmentation and classification to detect acute leukaemia. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 2(3): 147-151.
- [14] Mahopatara, S., Patra, D., Satpathy, S. (2014). An ensemble classifier system for early diagnosis of acute lymphoblastic leukaemia in blood smear images. *Neural Computing & Applications*, 24: 1887-1904. <https://doi.org/10.1007/s00521-013-1438-3>
- [15] Kulkarni-Joshi, T.A., Bhosale, D.S. (2014). A fast segmentation scheme for acute lymphoblastic leukaemia detection. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 4(1): 2278-8875.
- [16] Vidhya, C., Kumar, P.S., Keerthika, K., Nagalakshmi, C., Devi, B.M. (2015). Classification of acute lymphoblastic leukaemia in blood microscopic images using SVM. In: *International Conference on Engineering Trends and Science & Humanities (ICETSH)*, pp. 185-189.
- [17] Goutam, D., Sailaja, S. (2015). Classification of acute myelogenous leukaemia in blood microscopic images using supervised classifier. *International Journal of Engineering Research & Technology (IJERT)*, 4(1): 569-574. <https://doi.org/10.1109/ICETECH.2015.7275021>
- [18] Bhattacharjee, R., Saini, L.M. (2015). Detection of Acute Lymphoblastic leukemia using watershed transformation technique. In 2015 International Conference on Signal Processing, Computing and Control (ISPCC), Wagnaghat, India, pp. 383-386. <https://doi.org/10.1109/ISPCC.2015.7375060>
- [19] Bibin, D., Nair, M.S., Punitha, P. (2017). Malaria parasite detection from peripheral blood smear images using deep belief networks. *IEEE Access*, 5: 9099-9108. <https://doi.org/10.1109/ACCESS.2017.2705642>
- [20] Rawat, J., Singh, A., Bhadauria, H.S., Virmani, J., Devgun, J.S. (2017). Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers. *Multimedia Tools and Applications*, 76(18): 19057-19085. <https://doi.org/10.1007/s11042-017-4478-3>
- [21] Mughal, B., Muhammad, N., Sharif, M., Rehman, A., Saba, T. (2018). Removal of pectoral muscle based on topographic map and shape-shifting silhouette. *BMC Cancer*, 18: 778. <https://doi.org/10.1186/s12885-018-4638-5>
- [22] Saxena, P., Goyal, A. (2022). Computer-assisted grading of follicular lymphoma: A classification based on SVM, machine learning, and transfer learning approaches. *The*

- Imaging Science Journal, 70(1): 30-45. <https://doi.org/10.1080/13682199.2022.2162663>
- [23] Hegde, R.B., Prasad, K., Hebbar, H., Singh, B.M.K., Sandhya, I. (2020). Automated decision support system for detection of leukemia from peripheral blood smear images. *Journal of Digital Imaging*, 33: 361-374. <https://doi.org/10.1007/s10278-019-00288-y>
- [24] Chand, S., Vishwakarma, V.P. (2022). A novel deep learning framework (DLF) for classification of acute lymphoblastic leukemia. *Multimedia Tools and Applications*, 81(26): 37243-37262.
- [25] Putzu, L., Caocci, G., Di Ruberto, C. (2014). Leucocyte classification for leukaemia detection using image processing techniques. *Artificial Intelligence in Medicine*, 62(3): 179-191. <https://doi.org/10.1016/j.artmed.2014.09.002>
- [26] Claro, M.L., de MS Veras, R., Santana, A.M., Vogado, L.H.S., Junior, G.B., de Medeiros, F.N., Tavares, J.M.R. (2022). Assessing the impact of data augmentation and a combination of CNNs on leukaemia classification. *Information Sciences*, 609: 1010-1029. <https://doi.org/10.1016/j.ins.2022.07.059>
- [27] Sethuraman, S.K., Malaiyappan, N., Ramalingam, R., Basheer, S., Rashid, M., Ahmad, N. (2023). Predicting Alzheimer's disease using deep neuro-functional networks with resting-state fMRI. *Electronics*, 12(4): 1031. <https://doi.org/10.3390/electronics12041031>
- [28] Saxena, P., Goyal, A., Bivi, M.A., Singh, S.K., Rashid, M. (2023). Segmentation of nucleus and cytoplasm from H&E-Stained follicular lymphoma. *Electronics*, 12(3): 651. <https://doi.org/10.3390/electronics12030651>
- [29] Jolliffe, I.T., Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065): 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- [30] Yang, L., Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415: 295-316. <https://doi.org/10.1016/j.neucom.2020.07.061>
- [31] Singh, S.K., Khamparia, A., Sinha, A. (2022). Explainable machine learning model for diagnosis of Parkinson disorder. *Biomedical Data Analysis and Processing Using Explainable (XAI) and Responsive Artificial Intelligence (RAI)*, pp. 33-41. https://doi.org/10.1007/978-981-19-1476-8_3
- [32] Putzu, L., Di Ruberto, C. (2013). White blood cells identification and classification from leukemic blood image. In *Proceedings of the International Conference on Bioinformatics and Biomedical Engineering*, Copicentro Editorial, Granada, Spain, pp. 99-106.
- [33] Rawat, J., Singh, A., Bhadauria, H.S., Virmani, J., Devgun, J.S. (2017). Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers. *Multimedia Tools and Applications*, 76(18): 19057-19085. <https://doi.org/10.1007/s11042-017-4478-3>
- [34] Singhal, V., Singh, P. (2015). Correlation based feature selection for diagnosis of acute lymphoblastic leukemia. In *Proceedings of the Third International Symposium on Women in Computing and Informatics*, pp. 5-9. <https://doi.org/10.1145/2791405.2791423>
- [35] Bhattacharjee, R., Saini, L.M. (2015). Detection of Acute Lymphoblastic leukemia using watershed transformation technique. In *2015 International Conference on Signal Processing, Computing and Control (ISPPCC)*, Wagnaghat, India, pp. 383-386. <https://doi.org/10.1109/ISPPCC.2015.7375060>
- [36] Neoh, S. C., Srisukham, W., Zhang, L., Todryk, S., Greystoke, B., Lim, C.P., Aslam, N. (2015). An intelligent decision support system for leukaemia diagnosis using microscopic blood images. *Scientific Reports*, 5: 14938. <https://doi.org/10.1038/srep14938>
- [37] Amin, M.M., Kermani, S., Talebi, A., Oghli, M.G. (2015). Recognition of acute lymphoblastic leukaemia cells in microscopic images using K-means clustering and support vector machine classifier. *Journal of Medical Signals and Sensors*, 5(1): 49-58.
- [38] Viswanathan, P. (2015). Fuzzy C means detection of leukaemia based on morphological contour segmentation. *Procedia Computer Science*, 58: 84-90. <https://doi.org/10.1016/j.procs.2015.08.017>
- [39] Singhal, V., Singh, P. (2016). Texture features for the detection of acute lymphoblastic leukemia. In *Proceedings of International Conference on ICT for Sustainable Development: ICT4SD 2015*, pp. 535-543. https://doi.org/10.1007/978-981-10-0135-2_52