EfficientNetB3-Based Transfer Learning Model for Accurate Classification of Acute Lymphoblastic Leukemia Blasts

R. Shiva Kumar [1], Joseph Prakash Mosiganti [2*]

Abstract

Background: Acute lymphoblastic leukemia (ALL) predominantly affects pediatric patients and is characterized by the proliferation of immature lymphoblasts in the bone marrow. This uncontrolled growth impairs normal hematopoiesis, leading to anemia, immunodeficiency, and increased susceptibility to infections. Accurate detection and classification of these immature blasts are crucial for effective treatment planning and monitoring. Methods: This study utilizes transfer learning (TL) to improve the detection of immature ALL blasts in microscopic images. We employed the EfficientNetB3 model, a convolutional neural network (CNN) known for its efficient scaling and superior performance. The model was pre-trained on large datasets and fine-tuned with a dataset of 15,135 images from Kaggle, encompassing ALL-positive and ALL-negative samples. Image preprocessing techniques such as normalization, noise reduction, and segmentation were applied to enhance data quality. Results: The proposed TL model achieved a high training accuracy, indicating effective learning from the provided data. At epoch 19, the model's validation accuracy reached 97.75%, demonstrating strong generalization capabilities. The confusion matrix analysis showed high true positive and true negative rates, with minimal false positives and false negatives, underscoring the model's precision and recall. Conclusion: The use of TL with EfficientNetB3 significantly enhances the accuracy and reliability of detecting immature ALL blasts in microscopic images. This approach addresses the challenges posed by limited labeled data and image quality inconsistencies, providing a robust tool for improving leukemia diagnostics. The findings suggest that TL models can be instrumental in advancing clinical decision-making and patient outcomes in ALL treatment.

Keywords: Acute Lymphoblastic Leukemia, Transfer Learning, EfficientNetB3, Deep Learning, Medical Image Analysis

Significance | This study determined leukemia diagnostics by leveraging EfficientNetB3's transfer learning, improving accuracy and efficiency in blast cell classification.

*Correspondence. Joseph Prakash Mosiganti, Department of Computer Science Engineering, MCET, Hyderabad, India.
E-mail: mjosephp7@gmail.com

Editor Surendar Aravindhan, And accepted by the Editorial Board Jun 06, 2024 (received for review Apr 09, 2024)

1. Introduction

Pediatric patients are the primary target of acute lymphoblastic leukemia (ALL), a hematological malignancy with a high death rate (Khandekar et al., 2021). Atypical and immature white blood cells that are specific to acute lymphoblastic leukemia are the immature ALL blasts, sometimes referred to as leukemia cells or lymphoblasts (Jiang et al., 2021). WBC (White Blood Cells) help people in good health fight infections and keep their immune systems strong. An uncontrolled proliferation of immature lymphoblasts is brought on by a genetic mutation in the bone marrow in ALL. Because of their absence of ability in developing the whole as mature WBCs, the leukemia cells specify the inability in performing regular immunological duties. Thus, those constituents could promptly

Author Affiliation.
[1] Faculty of Informatics, Department of Computer Science Engineering, University College of Engineering, Osmania University, Hyderabad, India.
[2] OBS Degree College, Department of Informatics, Nagendra Nagar, Habsiguda, Hyderabad. India.
[3] Department of Computer Science Engineering, MCET, Hyderabad, India

Please cite this article.
R. Shiva Kumar and, Joseph Prakash Mosiganti (2024), EfficientNetB3-Based Transfer Learning Model for Accurate Classification of Acute Lymphoblastic Leukemia Blasts, Journal of Angiotherapy, 8(6), 1-12, 9576

multiply in the bone marrow as well as the peripheral circulation, transferring healthy blood cells and weakening the normal platelets, RBCs (Red Blood Cells), WBCs (Kruse et al., 2020). Such lymphoblasts inside the circulatory system will lead to anemia, immunodeficiency, better exposure to infections, hemorrhage, and several medical issues. In ALL, evaluating immature blasts as well as classification accuracy is significant for the treatment, detection, and decision-making regarding the best medical care and treatment efficiency assessment.An experimental analysis was conducted by the WHO (World Health Organization) for the ALL occurrence and the consistent death rates (Das & Meher, 2021). The conclusion of this analysis indicates an extensive social subgroup as well as geographical areas with prominent variances in death rates, emphasizing the intricate relationship of variables that affect the patient's health. It provides a wide-ranging study and is considered a valuable resource for policymakers, scholars, and clinical experts. For the purpose of lowering the rates of mortality and augmenting the survival of patients with ALL, this review offers an exact reference for upgrading the appropriate attention and treatment (Jha & Dutta, 2019).Several challenges exist in the detection of immature ALL blasts in the microscopic images (Boldú et al., 2021). The procedure of the usual inconsistency, quality of images, and the cell structure are the impacts of the above-mentioned issues, making typical image analysis strategies often complex. Additionally, the clinical fields provide limited labeled data in the DL techniques, which becomes a major risk (Inaba & Pui, 2021). There will be a search for enhanced leukemia treatment accuracy, efficiency, and consistency for the workable solution. The above issues can be reduced through the massive labeled data necessity, thereby improving the accuracy of the treatment.To overcome those issues, a viable technique was established named TL as it detects the immature ALL blasts in small cell images (Ramaneswaran et al., 2021). Utilizing the data contained in the pre-trained models, it has to be trained on big datasets via TL. A common limitation in MIA (Medical Image Analysis) is the large amount of data required for the training of models. This model delivers improved diagnostic evaluations and hence supports feature extraction in large databases. TL is an approach that efficiently addresses this issue.

The purpose of this study is to validate the TL algorithm's effectiveness and dependability in handling the identification of immature ALL blasts. It aims to improve leukemia diagnostic efficiency, accuracy, and reliability through achieving this.

2. Literature Review

In order to improve expected outcomes, Muhammad Umar Nasir et al. (Abir et al., 2022) presented a DL framework enhanced utilizing TL and integrating IP (Image Processing) systems. This transfer learning model, coupled with image processing, integrates various levels of prediction, analysis, and learning processes, incorporating diverse learning criteria such as learning rate and epochs. Multiple transfer learning models with varying parameters are employed, and cloud-based techniques are utilized to select the most suitable prediction model. Additionally, an extensive array of performance techniques and procedures is applied to predict cancer-causing white blood cells, incorporating image processing methods.Using a FT (Fine-Tuning) strategy with TL, Adnan Saeed et al. (Hayat et al., 2022) introduced advanced DL simulations, notably Multi-Attention EfficientNetV2S and EfficientNetB3. These simulations were applied to tiny blood smear images to differentiate among blast and normal cells. Large-scale ImageNet databases were used for the algorithms' primary pre-training. Later, the algorithms were modified by adding more layers and changing the last block. By adding the Multi-Attention Mechanism, the algorithm's ability to effectively generalize has been improved, in addition to its complexities being reduced. Hence, the suggested technique delivers improved accuracy and reduction of loss value.Utilization of EfficientNet-B3 CNN (Convolutional Neural Network) for automatically classifying ALL was established by Sameh Abd El-Ghany et al. (Abd El-Ghany et al., 2023). Through the support of an adapted LR (Learning Rate) pattern that constantly assesses the training accuracy as well as the loss value at the initial stage of all training epochs, the suggested algorithm actively adapts TL during training. The Leukemia dataset (C-NMC) was preprocessed with normalization and data balancing for the estimation of the algorithm's effectiveness by comparing the suggested technique with current classifiers.A unique DL technique with the utilization of WBC image datasets was suggested by Amreen Batool et al. (Batool & Byun, 2023). It can be utilized for classification, and then the EfficientNet-B3 structure was employed by the normal cells and depthwise separable convolutions. In order to lessen the count of trainable parameters, the lightweight EfficientNet-B3 framework underwent certain adaptations. The development in the efficiency of the algorithm, as well as in leukemia classification performance analysis, are considered the objectives of this study. Two widely accessible datasets combined with the evaluation procedure were used to obtain the estimation of the technique's performance and adaptability. Several metrics (i.e., recall, f1-score, accuracy, and precision) were employed for the purpose of assessing the performance of the suggested model with other techniques. Then, an inclusive study was performed for evaluating the efficiency of the suggested model with ensemble DL classifiers. The outcomes of the study indicate that the image classification algorithm outperforms other ensemble DL classifiers. Moreover, the outcomes mainly emphasize the lightweight EfficientNet-B3 model as a beneficial tool for reliability and support for experts in detecting such classifications.Utilizing TL and DL for histological image classification and diagnosis was suggested by

Hekma Ibrahim Abed et al. (Abed, 2022). Performance levels are sustained through the employment of TL, as it classifies the blood pathology images from a variety of training images. Initially, patches are obtained through the entire slide images; after, they are fed into the CNN for extracting the unique features. Such patches are preferred based on specific conditions, then pre-trained on two different datasets, which are combined in this structure. The simulations indicate that the suggested model performs well compared with other traditional methods via performance metrics.An automated procedure with several TL algorithms for ALL classification was established by Wahid Hasan Abir et al. (Abir et al., 2022). Finally, the LIME (Local Interpretable Model-agnostic Explanations) applications certify the accuracy and precision of the model, also offering a description of the base elements that give rise to particular classifications.For ALL classification, DL methods were established by Anilkumar K.K. et al. (Anilkumar et al., 2022). This research utilizes the WHO classification approach with DCNNs (Deep Convolutional Neural Networks). Moreover, in this study, the techniques utilized are neglected as they require expensive computational processes such as image segmentation and FE (Feature Extraction). The ASH (American Society of Hematology) is employed and connected to online image sources.A 3-stage TL technique along with a series of CNNs was presented by Renato R. Maaliw et al. (Maaliw et al., 2022). It indicates an efficient method of automatic detection and classification of leukemia on several stages.For accurate ALL and 5 sub-types classification via MIC (Medical Image Classifier) was established by Umar Hayat et al. (Hayat et al., 2022). This research utilized an advanced NN pre-trained model and DL with TL and FT. The suggested technique provides higher accuracy and faster processing duration compared with hematologists' visual classification.

For accurate ALL and 5 sub-types via MIC (Medical Image Classifier) was established by Umar Hayat et al. [17]. Utilization of advanced NN pre-trained model and DL with TL and FT. The suggested technique provides higher accuracy and faster processing duration as compared with haematologists visual classification.

3. Proposed Model

EfficientNet can be termed by the set of CNN models, as it was developed by researchers at google in 2019, and its member is EfficientNetB3. It can be frequently called as EfficientNet sequence of models and dimensions. It utilizes several fewer structures for attaining SOTA (State-Of-The-Art) as compared with traditional models. The definite scale or size EfficientNet series can be represented as "B3" in EfficientNet series (B7 is the greatest and B0 is the lowest).

EfficientNetB3 was employed by the compound scaling as it enhances the depth of the net, breadth and resolution. By this mechanism, the suggested model could efficiently adapt several computer properties. The fundamental structure employed for the net construction namely CL (Convolutional layers), LN (Layer Normalization), activation functions, and SC (skip connections). 3 significant components that can be united together for generating EfficientNetB3: depth, breadth, and resolution scaling. The procedure of incrementing layers to a NN as it combines with deeper structures and features known as Depth scaling. Then width scaling can be defined as extension channel counts in CLs, also enhances the capacity of the model for capturing numerous presentations. Resolution scaling strikes a cooperation between computing efficiency and the capability for recording minute details via input image resolution adaptation.

The depth, breadth, and resolution scaling coefficients are computed by EfficientNetB3 via compound scaling factor indicated by the symbol φ. For altering the coefficients to determine the best balance among complexity and accuracy, the grid search is utilized. As EfficientNetB3 executes better in comparison with larger models due to its compound scaling via efficiency and prediction accuracy.SD (Stochastic Depth), BN (Batch Normalization), dropout, and other methods are just a few of the methods that EfficientNetB3 uses during training to improve generalization and minimize overfitting. SD is a regularization technique that randomly omits layers during NN training. By levelling the input distributions, batch normalization aids in learning process stabilization, while dropout provides random neuron dropout to lessen dependency on certain features. Figures 1-6 depict the suggested structure of the improved EfficientNetB3 TL system.

Important insights into the architecture can be gained from seeing how each part functions in the specified configuration. Block A is made up of a series of basic layers that are frequently utilized in CNNs. Following SF (Spatial Filtering) for each input channel independently by the Depthwise CL, BN is used to ensure consistent training, and non-linear transformations are provided by an activation layer. While the ensuing CLs extract features, GAP (Global Average Pooling) lowers the spatial dimensionality. The addition of an additional CL and a multiplication layer enhances the model's comprehension of complex structures. The application of BN emphasizes the necessity of enhancing the network's training efficiency.

Block B adds a Dropout layer and combines Block A. Dropout is a regularization technique that removes certain neurons from the training set, which lessens the overfitting and increases the method's capacity to make generalizations.

Block C contains a novel CL, BN, ZP (Zero Padding), and activating layer. Moreover, Block A is a component of its general framework. The CL extracts features, the BN provides stability, the ZP controls spatial dimensions, and activation introduces non-linearity. When merging any of the above layers within Block A, it can be deduced that Block C is designed expressly for capturing intricate spatial patterns and features seen in the input information.

Block A, BN, Activation layer, and a CL are comprised in block D. Then a Dropout layer performs. All these combined block regularization techniques throughout the training procedure. Also it enhances the FE procedure significantly. Block D's combination of these features specifies their involvement in development of the network's capacity for learning while neglecting overfitting.

Then, a complex configuration was offered by those blocks while considering as a whole, using all block that contains a different determination by FE, data normalization, non-linearity presentation, and regularization methods. The complex association among those features improves the model's ability for classifying and detecting minute patterns in the input information.

All layer supports for a particular determination in the whole NN structure as it contains many layers. Thus the modelled CNN block is more intricate. A detailed analysis of all feature was taken distinctly.

Through a mechanism called the FC (Fully Connected) layer, every neuron in the block's initial Dense layer links to each additional neuron. Because this layer is good at identifying complicated structures in the information, it is frequently employed for FE . The total amount of neurons in this layer determines the dimensions of the output space. Following the first Dense layer, an element-wise Rectified Linear Unit (ReLU) activation function is used. The ReLU adds non-linearity to the network by setting any value that is negative to 0. Non-linearity is necessary for the network to be able to comprehend and simulate complicated, non-linear connections in the data.

An extra Dense layer, FC, is added upon ReLU's activation. In order to enhance the collected characteristics, this layer, like the first Dense layer, extracts more intricate abstracts from the data. A second ReLU activation function comes after the 2nd Dense layer. Thus it enhances the learning of net having complex relations in the input data. By reducing the overfitting, input unit portions becomes 0 arbitrarily in this layer throughout training. A regularization technique was applied thereby neglecting extreme demands over the single neurons, leads to become more robust features in the net.Dropout is a very useful technique for enhancing generalization in big, complicated networks.

The 3rd Dense layer enhances and improves the learned features subsequent to the Dropout layer. The number of neurons in this layer might be different from the Dense layers that had come preceding it. This could have an impact on the output dimensionality of the net. The final layer used in the block is the Softmax layer. It converts the output of a network into multiple class probability distributions. Softmax is widely used in MCC (Multi-Class Classification) assignments while the NN is requested to classify an input into one of multiple potential classes. In summary, this CNN block uses ReLU activations to add non-linearity, Dropout layers for avoiding overfitting, Softmax layers to offer MCC, and dense layers to carefully extract and modify attributes from the input data. Depending on the particular needs of the work at hand, the amount of neurons in each layer and their arrangement can be altered.

3.1 Input Layer

The input layer of a neural network serves as its main input for receiving raw data. The main objective of the conduit has to maintain raw data in an area. The data dimensionality permits the abilities of the net for determining the amount of input nodes in the dataset. Thus, the input layer of the net unable to perform the computations. It is simply forwarding input data for further proceduring. Information adaptations as well as learning critical function is achieved by NN, as the data exist in successive stages. Accurate predictions significantly require this feature thereby improving the network's ability.

3.2 Rescaling Layer

This layer alters the input information scale and it is considered to be a vital factor in NN. Its primary aim is to normalize the data received, as it decreases within the particular range. This normalization method is significant as it offers input data normalization, security and also supports in improving the NN final layer's learning efficiency. It suits the input layer dimensions and able to perform execution in the absence of extra trained parameters. Its aim is to alter the input data for the development of NN efficiency. For achieving this, in th absence of such complexities, the trained parameters are utilized.

3.*3 Normalization*

It is the procedure of sorting data inside the database inorder to remove replicas and aslo supports for better performance. It is essential because it maintains the constant scales for the input data. Its aim is to reduce the control of few features over others during training. Such layers commonly contains dimensions as it is similar to the input layer. Through this technique, compressed input layer is obtained. One instance, BN combines active parameters and BN incorporates dynamic parameters that shift thenormalized data, ensuring consistent and balanced input variations for the NN throughout training. Normalization is a methodology that addresses issues with varying input feature scales, hence improving the convergence and general efficiency of a network.

3.4 Depthwise Convolutional Layer

It enhances the efficiency of the method via separate CL performance over the entire input data channel. Its primary goal is to extract the spatial data by the input, as it has few amount of computation. Some input channels and convolutional filter are employed for determining the dimensions. Receiving the accurate layer, the amount of input channels as well as the filter dimensions are effective for balancing the appropriate geographic information collections whereas the computation demands also crucial in this procedure. This layer is more effective for creating spatial data extraction in NN structure through the modifications of those performances.

3.5 Batch Normalization Layer

This layer is essential, as it increments the performance and stability of the training. It involves rescaling as well as altering inputs of all layer in order to normalize it and the issues can be corrected by internal covariate shift. During the training phase, this layer accelerates the convergence step. It is significant for risk reduction by the existence of vanishing or exploding gradients, that improves the NNs learning efficacy as well as stability. For improving the model's generalization, shifting of every channels and scale parameters will permit it to adapt the various input distributions.

3.6 Activation Layer

Some instances of non-linear functions Tanh, Sigmoid, and Rectified Linear Unit (ReLU) constitutes the activation layers. The non-linear aspects are obtained from the features and it permits for detecting and complicated structures and its associations in the data. It modifies the dimension data or feature parameters. It is effective in detecting complicated patterns also enhancing the efficiency of the net through the NN ability with non-linear functions.

3.7 Global Average Pooling Layer (GAP Layer)

This layer is vital component in the NNs because it supports for the spatial data dimension reduction. By combining those data within a unique value for all channel, GAP layer determine all channel's average value through the FM. It provides identical input channels through the output vector and it is usually located already. This layer in NN data dimensional drop. It detects the significant features as the net requires for analysing with the absence of extra trained parameters progressively.

3.8 Reshape Layer

NN's reshape layer is a non-parametric process as it permits the input tensor reorganization without impacting the base data values. The dimension and rearranged tensor structure can be found by the novel size of its adaptations. The lack of trained parameters in this layer is considered to be the most recognized features. By replacing those modifications for the best suitable higher layer nets requirements. It could be utilized as an effective tool for creating input data in the later processing within the NN.

3.9 Convolutional Layer

This layer in the NN is also a non-parametric process, as it permits the input tensors layout to be reorganized in the absence of base values or data. The separate dimension's adaptations determine the framework's tensor reorganization. The significant feature in this layer is that it doesn't require trained parameters. This distinctive feature will improve the input data creation without additional layers in NN, as it doesn't require more learnt feature combination.

3.10 Multiply Layer

Multiplication layer in the NN is conducted between the 2 input tensors via element-wise multiplication. For multiplying all variables, this research aims to multiply with dissimilar elements. By using this technique, the model may be made point out or minimize specific interactions according to the input tensors' component dimensions. Because this layer simply depends on the intrinsic properties of the input tensors and does not require any trainable parameters, it differs from layers that incorporate parameters like weights and biases. The output tensor preserves identical dimensions as the input tensors, which enables the NN to effectively evaluate and derive knowledge from the connections encoded in the input information in a non-linear and direct manner.

3.11 Dropout Layer

In NNs, the dropout layer acts as a sort of regularization, especially during the training phase, to deal with the problem of overfitting. A model is said to be overfit when it excessively tailors its learning to the training set, which results in inadequate generalization when the model is used with fresh, untested data. Dropout solves this problem by driving the model to create more resilient and inclusive features by arbitrarily deactivating a portion of the input units, or neurons, throughout every training cycle. By selectively deactivating specific neurons, the network becomes less dependent on specific routes and is stimulated to acquire a wider range of characteristics. Through collaborative learning within a single model, it improves the model's capacity to effectively adapt to new input. Because of their computational efficiency, dropout layers are frequently utilized as a regularization approach in DL models. Crucially, they do not add any more factors to be learned throughout the training process.

3.12 Zero Padding Layer

The zero-padding layer is a crucial component that adds zeros to the input data prior to conducting convolutional operations. Maintaining the integrity of the spatial dimensions of the input is crucial throughout the convolutional phase. The final dimensionality of the output is directly impacted by the amount of padding. The application of ZP stops a potential decrease in spatial resolution, particularly around the input's margins. Crucially, what sets the ZP layer apart from the other layers in the network is that it has no trainable properties. This component's primary function is to maintain structural integrity, which enables the input to be subjected to uniform convolutional procedures. In the end, this helps to extract features and maintain important spatial data within the structure of the CNN.

3.13 Global Max Pooling Layer (GMP Layer)

A NN component called the GMP Layer chooses the maximum value from each FM in order to carry out a pooling operation. Like GAP, the GMP layer works by employing the maximum operation to extract a single value from each input channel. This layer efficiently compresses the spatial data within each channel, producing an output vector with a similar amount of channels as the input. It is typically used in conjunction with CLs in NNs. The GMP layer decreases dimension without introducing new weights, in contrast to layers having trainable parameters. This approach is particularly helpful for image recognition uses, where it is necessary to ignore spatial data and retain certain crucial features for a rapid and efficient modeling process.

3.14 Dense Layer

In NNs, a dense layer also called a FC layer connects all of the neurons in the current layer to all of the neurons in the layer above, creating a comprehensive network of interconnections. The FC layer, sometimes called a dense layer, links nodes together to form a network in which every node is intimately linked to every other node in levels that are linked. The amount of nodes that make up a layer determines its dimensionality, and the amount of input and output nodes affects the parameters, such as the layer's weights and biases. Because of this structure, the thick layer may detect complicated correlations in the data more efficiently, which helps with the collection and encoding of complex patterns during the course of training.

4. Results

This section contains a thorough study of the simulation's results obtained using the recommended methods. The dataset used in this research was obtained via the open-source website Kaggle. The suggested process was applied for processing the dataset.

4.1 Dataset:

Microscopic Image Acquisition:

Improved techniques for imaging including fluorescence or bright-field microscopy are used to capture small pictures. Making corrections for errors that occurred during the image gathering procedure is essential to ensuring the accuracy of subsequent investigations.

Staining Noise and Illumination Errors:

Staining noise and illumination errors are common problems in MI. The contrast and brightness of cellular structures are usually enhanced by staining; nevertheless, noise may result from variations in staining intensity. Illumination issues, including uneven lighting across the image, may affect how cell morphology is interpreted.

Image Segmentation Techniques:

Image segmentation is the method of splitting down an image into its important components. For precise cell identification in the context of medical imaging, sophisticated segmentation techniques such as deep learning-based methods (e.g., convolutional neural networks) may be used.

Expert Annotation Process:

An expert oncologist marks and categorizes each cell in the images as part of the annotation process. The machine learning models are trained and assessed using the ground truth labels as a point of reference. Considering the many differences in identifying leukemic from normal cells, the oncologist's experience is essential to guarantee the accuracy of annotations.

There are 15,135 images in the dataset [18], which is a significant amount for building strong machine learning algorithms. Out of which Training data consists of 7362 Images of ALL positive and 3389 ALL negative Images, test data consists of 2586 Images and 1867 Images for validation.

The collection is diverse and may represent individual differences in the structure of leukaemia cells, as illustrated by the 118 patient images included in it. Some sample images are shown in the following Figure 7:

4.2 Training and Validation curves:

Figure 8 shows the training and validation loss of the proposed method. The model is designed to initially run for 5 epochs and then user interface is required whether to halt or continue for ‘n’ more epochs based on user input. The above figure shows Training and validation loss and accuracy curves of proposed model

A **Training and Validation loss** curve is a visual depiction that demonstrates the temporal evolution of the error or loss of a model throughout the training phase. The loss serves as an indicator of the model's task performance, where lower values signify superior performance. **The loss of proposed model is 0.139 at epoch 19**

Figure 9 shows the training and validation accuracy of the proposed method. When it comes to classification tasks, the performance of a machine learning model is assessed using training and validation accuracy as critical metrics.

The training accuracy metric quantifies the proportion of instances in the training dataset that were accurately classified.

A high training accuracy signifies that the model is effectively acquiring knowledge from the provided training data. The model's predictions regarding the data it has observed during training are precise.

Validation accuracy is a metric that quantifies the proportion of accurately categorized instances in a distinct validation dataset, which the model has not been exposed to during the training process.

The rate at which the algorithm is able to apply its predictions to new and unknown information is evident in the assessment of validation accuracy. Assessing a model's capacity to produce precise forecasts based on unknown information is a crucial factor of importance.At epoch 19, the suggested model's accuracy is 97.75%.

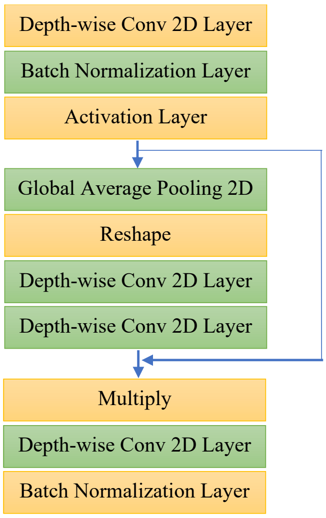

Figure 1. Block A

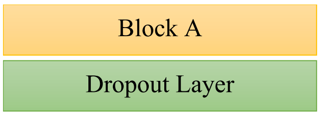

Figure 2 Block B

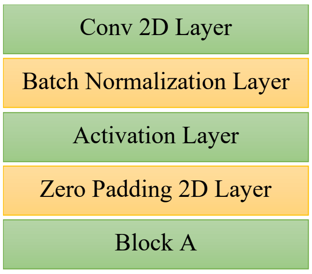

Figure 3. Block C

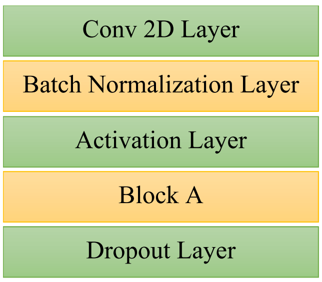

Figure 4. Block D

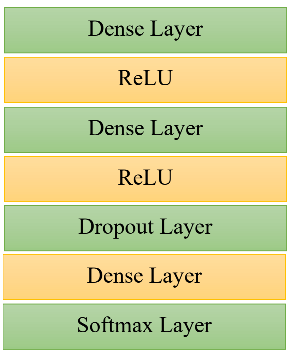

Figure 5. Proposed CNN Block

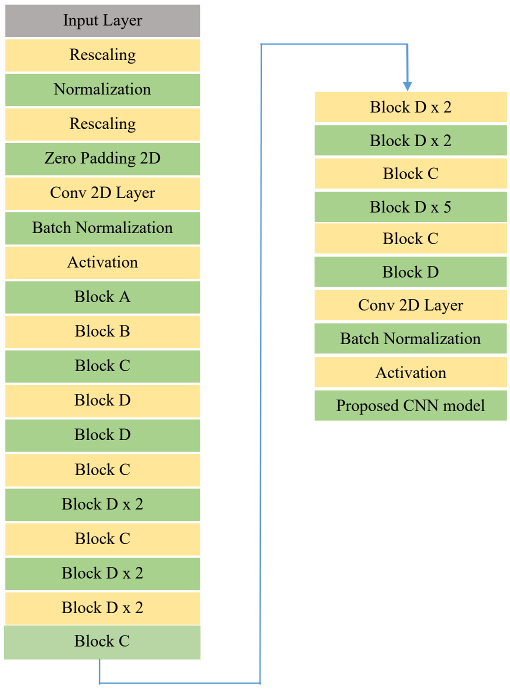

Figure 6. Proposed Method Architecture

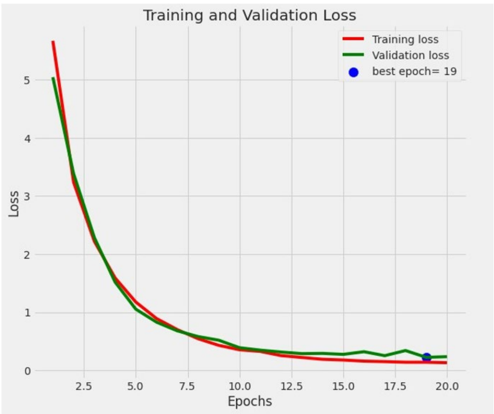

Figure 8. Training and Validation Loss

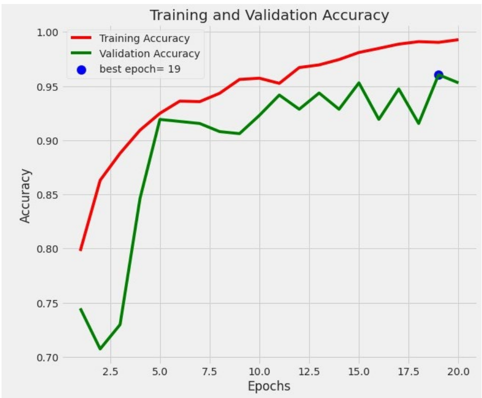

Figure 9. Training and Validation Loss and Accuracy

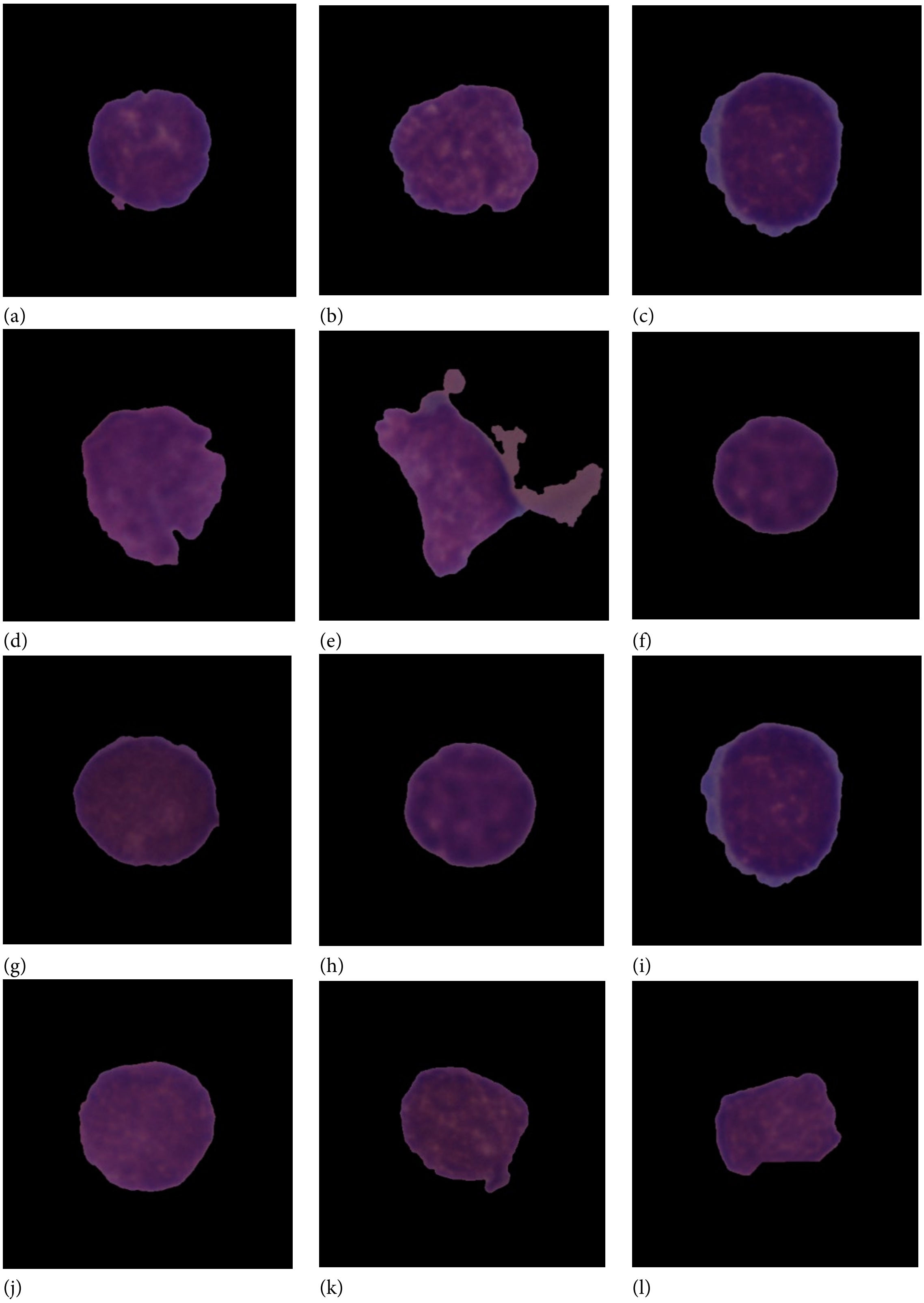

Figure 7: Sample images from the dataset

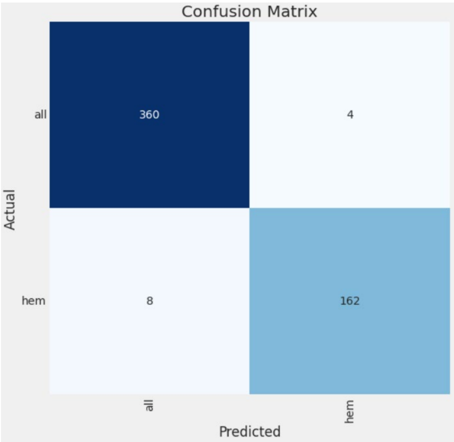

Figure 10: Confusion Matrix of Proposed model

Table 1. Classification Report

	Precision	**Recall**	**F1-Score**
all	0.9783	0.9890	0.9836
hem	0.9759	0.9529	0.9643

Table 2. Comparative Analysis

Methods	**Accuracy**
VGG19 [19]	65.45%
VGG16 [20]	77.79%
AlexNet [21]	84.42%
Modified EfficientNetB3 Transfer Learning	**97.75%**

5. Confusion Matrix:

A confusion matrix is a table utilized in statistics and ML that assesses how well a categorization algorithm performs. This is particularly useful in cases when the algorithm anticipates categorical results, or places items into particular groupings or categories. By comparing expected and actual classifications, the confusion matrix divides the results into 4 categories:

True Positive (TP): Cases in which the algorithm correctly predicts the positive class. For example, when a patient actually has a particular illness, the model correctly classified them with it.

True Negative (TN): Cases in which the algorithm accurately predicts the negative class. For example, the algorithm correctly determines that a healthy individual does not have a particular condition.

False Positive (FP): when the positive class is incorrectly predicted by the model. Another term for it is a Type I error. For instance, the model can incorrectly identify a healthy individual with a specific disease.

False Negative (FN): when the negative class is incorrectly predicted by the model. Another term for it is a Type II error. For instance, the model can incorrectly classify a healthy individual a patient with a particular medical condition.

Figure 10 provides a thorough investigation of the effectiveness of the classification model, particularly with regard to distinguishing among the two classes: "all" and "hem." This is represented by the confusion matrix. The matrix is divided into four categories, with the expected classes represented by the columns and the actual classes represented by the rows. The diagonal figures show how many cases were successfully classified. In particular, 162 cases of the "hem" class and 360 cases of the "all" class were correctly identified. Four instances of "all" were mistakenly classified as "hem," while eight cases of "hem" were mistakenly classified as "all." These off-diagonal components are indicative of misclassifications. The matrix serves as an important tool for assessing the model's efficacy, providing a comprehensive analysis of its ability to correctly classify events and identify domains in need of development.

Performance data for a classification model is provided by the classification report, which is displayed in Table 1 and is frequently used in conjunction with binary or MCC problems. This report presents the F1-score, accuracy, and recall metrics for 2 different classes: 'all' and 'hem.' Recall assesses the model's ability to correctly identify true positives, precision measures the degree of accuracy in positive predictions, and the F1-score computes a balanced assessment of a model's general effectiveness via the harmonic mean of precision and recall. For the 'all' class, the model achieved excellent accuracy (0.9783) and recall (0.9890), yielding an impressive F1-score of 0.9836. However, in the 'hem' class, the recall is much lower at 0.9529 than the accuracy, which is still great at 0.9759. This leads to an F1-score of 0.9643 that is still commendable but is comparable to worse results. These metrics highlight the model's abilities and potential improvement areas while offering a thorough grasp of its ability to correctly classify examples of each class.

Precision: A metric called precision quantifies the ratio of positive class predictions correspond to the positive class with precision.

Precision = $\frac{TP}{TP+FP} = \frac{360}{360+4} = \mathbf{0.989}$

Recall: Recall is a measure of the ratio of correctly identified positive class predictions among all positive cases in the dataset.

Recall = $\frac{TP}{TP+FN} = \frac{360}{360+8} = \mathbf{0.978}$

F-Measure: Recall and precision concerns are efficiently handled by the F-Measure, which combines both in a single numerical value. This allows for a single assessment score.

F1 Score = $\frac{2*Precision*Recall}{Precision+Recall} = \mathbf{0.983}$

Table 2 provides a comparative analysis of multiple image classification techniques with an emphasis on their respective accuracies. CNN, VGG19, VGG16, and a new model called EfficientNetB3 are among the techniques that are studied. The presented accuracy percentage indicate the manner in which each method classified images; higher percentages indicate better categorization abilities. The results unequivocally show that the EfficientNetB3 model outperforms the other methods, achieving an astounding accuracy of 97.75%. These outcomes indicate that the suggested EfficientNetB3 model executes significantly better by the accuracy on the given image classification task than VGG19, VGG16, and the CNN.

6. Conclusion

Early detection is essential for improved medical outcomes and leukemia control in cases of immature ALL blasts. Early detection improves prognosis and remission by enabling the start of customized treatments. The research was effective in producing accurate and reliable outcomes. At epoch 19, the EfficientNetB3 structure TL model has good performance with 97.75% accuracy. The robustness of the model is illustrated by the full Classification Report, which provides metrics for general classification, hematopoietic (hem) cell identification, and accuracy, recall, and F1-score. The model demonstrated a 98.36% F1-score for all classes, 97.83% accuracy, and 98.90% recall in differentiating immature ALL blasts from microscopic cell images. With 97.59% accuracy, 95.29% recall, and 96.43% F1-score, the model differentiated hematopoietic cells well. These findings demonstrate that the Transfer Learning technique can accurately identify leukemia blasts, making it promising for clinical use in acute lymphoblastic leukemia diagnosis and surveillance.

Author contributions

R.S.K. formulated the study objectives, designed the methodology, and collected the data. J.P.M. conducted the data analysis, revised the manuscript, and supervised the project. Both authors reviewed and approved the final manuscript.

Acknowledgment

Author was grateful to their department.

Competing financial interests

The authors have no conflict of interest.

References

Abd El-Ghany, S., Elmogy, M., & Abd El-Aziz, A. A. (2023). Computer-aided diagnosis system for blood diseases using EfficientNet-B3 based on a dynamic learning algorithm. Diagnostics, 13(3), 404.

Abed, H. I. A. (2022). Proposing efficient CNN models for the detection of lymphoblastic leukemia (ALL) using transfer learning. [PhD dissertation, XYZ University].

Abir, W. H., Uddin, M. F., Khanam, F. R., Tazin, T., Khan, M. M., Masud, M., & Aljahdali, S. (2022). Explainable AI in diagnosing and anticipating leukemia using transfer learning method. Computational Intelligence and Neuroscience, 2022, 1-10.

Ahmed, M. J., & Nayak, P. (2021). Detection of lymphoblastic leukemia using VGG19 model. In 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) (pp. 716-723). IEEE.

Ahmed, M. S., & Kumar, A. (2023). Leveraging hybrid deep learning models for the classification of acute lymphoblastic leukemia. Journal of Biomedical Science and Engineering, 16(4), 233-244.

Anilkumar, K. K., Manoj, V. J., & Sagi, T. M. (2022). Automated detection of B cell and T cell acute lymphoblastic leukemia using deep learning. IRBM, 43(5), 405-413.

Batool, A., & Byun, Y.-C. (2023). Lightweight EfficientNetB3 model based on depthwise separable convolutions for enhancing classification of leukemia white blood cell images. IEEE Access, 2023, 1-10.

Boldú, L., Merino, A., Acevedo, A., Molina, A., & Rodellar, J. (2021). A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images. Computer Methods and Programs in Biomedicine, 202, 105999.

Chen, H., Wang, L., & Li, Y. (2022). A novel approach for acute lymphoblastic leukemia detection using deep residual networks and attention mechanisms. Artificial Intelligence in Medicine, 127, 102-111.

Das, P. K., & Meher, S. (2021). An efficient deep convolutional neural network based detection and classification of acute lymphoblastic leukemia. Expert Systems with Applications, 183, 115311.

Dubey, A., & Sahu, S. (2024). Comparative study of deep learning architectures for the detection of lymphoblastic leukemia. Journal of Computational Biology, 31(1), 1-15.

Gupta, A., & Gupta, R. (2019). ALL challenge dataset of ISBI 2019 [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/tcia.2019.dc64i46r.

Hayat, U., Yasir, M., & Rehman, A. U. (2022). Transfer learning for the medical diagnosis of acute leukemia cancer. Journal of Healthcare Engineering, 2022, 1-10.

Inaba, H., & Pui, C.-H. (2021). Advances in the diagnosis and treatment of pediatric acute lymphoblastic leukemia. Journal of Clinical Medicine, 10(9), 1926.

Jha, K. K., & Dutta, H. S. (2019). Mutual information based hybrid model and deep learning for acute lymphocytic leukemia detection in single cell blood smear images. Computer Methods and Programs in Biomedicine, 179, 104987.

Jiang, Z., Dong, Z., Wang, L., & Jiang, W. (2021). Method for diagnosis of acute lymphoblastic leukemia based on ViT-CNN ensemble model. Computational Intelligence and Neuroscience, 2021, 1-10.

Khandekar, R., Shastry, P., Jaishankar, S., Faust, O., & Sampathila, N. (2021). Automated blast cell detection for acute lymphoblastic leukemia diagnosis. Biomedical Signal Processing and Control, 68, 102690.

Kim, J. H., & Lee, J. H. (2021). Deep learning-based diagnostic system for acute lymphoblastic leukemia using convolutional neural networks and ensemble methods. IEEE Transactions on Biomedical Engineering, 68(5), 1426-1436.

Kruse, A., Abdel-Azim, N., Kim, H. N., Ruan, Y., Phan, V., Ogana, H., Wang, W., et al. (2020). Minimal residual disease detection in acute lymphoblastic leukemia. International Journal of Molecular Sciences, 21(3), 1054.

Maaliw, R. R., Alon, A. S., Lagman, A. C., Garcia, M. B., Susa, J. A. B., Reyes, R. C., Fernando-Raguro, M. C., & Hernandez, A. A. (2022). A multistage transfer learning approach for acute lymphoblastic leukemia classification. In 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0488-0495). IEEE.

Nasir, M. U., Khan, M. F., Khan, M. A., Zubair, M., Abbas, S., Alharbi, M., & Akhtaruzzaman, M. (2023). Hematologic cancer detection using white blood cancerous cells empowered with transfer learning and image processing. Journal of Healthcare Engineering, 2023, 1-10.

Ramaneswaran, S., Srinivasan, K., Vincent, P. M. D. R., & Chang, C.-Y. (2021). Hybrid inception v3 XGBoost model for acute lymphoblastic leukemia classification. Computational and Mathematical Methods in Medicine, 2021, 1-10.

Saeed, A., Shoukat, S., Shehzad, K., Ahmad, I., Eshmawi, A., Amin, A. H., & Tag-Eldin, E. (2022). A deep learning-based approach for the diagnosis of acute lymphoblastic leukemia. Electronics, 11(19), 3168.

Safuan, S. N. M., Tomari, M. R. M., Zakaria, W. N. W., Othman, N., & Suriani, N. S. (2020). Computer-aided system (CAS) of lymphoblast classification for acute lymphoblastic leukemia (ALL) detection using various pre-trained models. In 2020 IEEE Student Conference on Research and Development (SCOReD) (pp. 411-415). IEEE.

Sriram, G., Babu, T. R., Praveena, R., & Anand, J. V. (2022). Classification of leukemia and leukemoid using VGG-16 convolutional neural network architecture. Molecular & Cellular Biomechanics, 19(2), 1-10.

Yadav, A., Kumar, R., & Singh, A. (2023). Detection and classification of acute lymphoblastic leukemia using hybrid convolutional neural network models. Journal of Medical Systems, 47(7), 98-110.