Research article

Diagnosis of Blood Microscopic Sample for Early Detection of Acute Lymphoblastic Leukemia Based on Convolutional Neural Network Classification Using GoogleNet , AlexNet and ResNet

By: Maha Emad, Donia Muohamed , Perrehan Essam , Shehab Rashad , Gehad Essam

Department of Medical Informatics, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt

Abstract

Acute lymphoblastic leukemia, a disease that is quite common, necessitates invasive, costly, and time-consuming diagnostic procedures for a final diagnosis. The early classification of cancer cases from non-cancerous samples plays a crucial role in ALL detection utilising peripheral blood smear (PBS) images. Due to the general character of all signs and symptoms, misdiagnosis is a common concern when these PBS images are analyzed by laboratory. Consequently, a lot of researchers used various CNN models to advocate the early diagnosis of leukemia. Moreover, some of the models that many researchers used is InceptionV3, AlexNet, and ResNet. The goal of this work was to enhance the diagnostic tools for the early diagnosis of leukemia using the acute lymphoblastic leukemia image dataset (Acute Lymphoblastic Leukemia (ALL) image dataset). The research shows how the accuracies improved by replacing the fully connected layer with the support vector machine. GoogleNet was 96% and after adding SVM it became 98%, AlexNet accuracy was 95.7% and became 97.8%, and ResNet was 96.3% and became 98.7%.

Keywords: acute lymphoblastic leukemia (ALL); SVM; CNN; GoogleNet; ResNet, AlexNet

1. Introduction

Acute lymphocytic leukemia (ALL) is a malignancy of B or T lymphoblasts characterized by uncontrolled proliferation of abnormal, immature lymphocytes and their progenitors which ultimately leads to the replacement of bone marrow elements and other lymphoid organs resulting in a characteristic disease pattern [1]. To illustrate, Health issues including leukemia, thalassemia, and anemia can result from an increase or reduction in any of the fundamental blood components. A high WBC volume decreases body immunity because it covers both platelets and RBCs. Acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia, acute myeloblastic leukemia, and chronic myeloblastic leukemia are the four categories according to their development, pace, and effects. The most prevalent and lethal of them is ALL, which accounts for 70% of all instances of leukemia. The progression of the illness is also significantly influenced by environmental and genetic variables. ALL is brought on by the bone marrow's excessive and unchecked multiplication of lymphocytes [2]. With about 6500 cases annually in the United States alone, acute lymphoblastic leukemia (ALL) is the second most prevalent acute leukemia in adults [3]. Acute lymphoblastic leukemia (ALL) is becoming more common within the healthcare system every year, with 64.2 thousand new cases of ALL diagnosed globally between 1990 and 2017. The most common form of leukemia in kids is ALL, which also has a high prevalence in adults [4]. The difficulty in making an early diagnosis of lymphocytes stems from the similarities between normal and lymphoid cell types. As a result, lymphocytes were divided into three groups: reactive, atypical, and normal. The characteristics of normal

lymphocytes include homogeneity and round, tiny, and rough nuclei; those of atypical cells include a big size and nucleus as well as the presence of lumpy chromatins; and those of reactive cells include heterogeneity and the presence of red cells around them. The lymphocyte types are identified by microscopic examination, which requires the collection of blood or bone marrow samples for a pathologist to analyse [5]. However, obtaining a sample of bone marrow and doing an analysis on it are necessary for a proper leukemia diagnosis. The analysis is laborious, time-consuming, and sensitive to the various expert viewpoints because it is done manually. Consequently, notwithstanding the possibility of human mistake, an accurate manual diagnosis depends on the pathologist's ability. In order to automatically detect leukemia, some researchers have suggested collecting WBC characteristics from microscopic pictures. Therefore, the quick and accurate diagnosis made possible by the automatic recognition of blood cell pictures will allow for the evaluation of several cells from each individual. Manual diagnostic issues can be resolved by machine and deep learning approaches. It has been demonstrated that the convolutional neural network (CNN), which has a greater capacity to distinguish between normal and blast cells, can assess and address many of the issues associated with manual diagnosis and medical imaging [6].

2. Related work

Numerous researches were conducted for implementing deep learning models specifically CNN techniques to classify various types of cancers. However, one the recent ones is conducted by students from Iran using images from peripheral blood smears to diagnose B-ALL and classify its subtypes [6]. 4 stages of data preprocessing were implemented: segmentation, decoding and resizing, normalization, and augmentation. In segmentation they converted the images to HSV color space then two maximum and minimum thresholds were set for purple then they applied a binary mask [6]. For decoding and resizing, they used DenseNet-201 architecture for feature extraction because of its lower computational costs and its ability for inputting high resolution images. Moreover, pixel intensity normalization was used from range 0 to 1. This step was crucial for enhanced model convergence in the training phase. All of the photos' global means and standard deviations were first determined, and then the values were normalized. Furthermore, six transformations were used to training samples for augmentation. These adjustments to the data's brightness, contrast, JPEG noise, and vertical and horizontal rotation are made at random. Regarding the division of the data, it was divided into 64% training, 20% testing and 16% validation. In addition, parameter tuning was necessary therefore the initial learning rate was 0.0003, batch size was 32, and epochs were 220.

3. Methods

This section displays the materials and methods used to analyze Acute Lymphoblastic Leukemia (ALL) image dataset for early detection of lymphoblastic leukemia. The dataset has samples from people with different stages of lymphoblastic leukemia ALL; Benign, Early, Pre, Pro. All images went through noise reduction, and various types of augmentation techniques then a convolutional neural network model called GooglNet InceptionV3 was used for image classification.

3.1 Dataset description

This study evaluates deep learning, convolutional neural network architecture to be specific on Acute Lymphoblastic Leukemia (ALL) image dataset. The bone marrow laboratory at Taleqani Hospital provided the images for this dataset (Tehran, Iran). This dataset included 3256 PBS pictures from 89 individuals who were thought to be ALL and whose blood samples were properly processed and stained by professional laboratory personnel. This dataset is separated into the benign and malignant classes. Hematogenes make up the first group, whereas the ALL group, which includes the three malignant lymphoblast subtypes Early Pre-B, Pre-B, and Pro-B ALL, makes up the latter. All of the photos were taken with a Zeiss camera at a 100x magnification in a microscope, and they were all saved as JPG files. The specific types and subtypes of these cells were identified by a professional using the flow cytometry instrument. We offer segmented photos

following colour thresholding-based segmentation in the HSV colour space. Figure 1 illustrates some samples from Acute Lymphoblastic Leukemia (ALL) image dataset (accessed 11/3/2023).

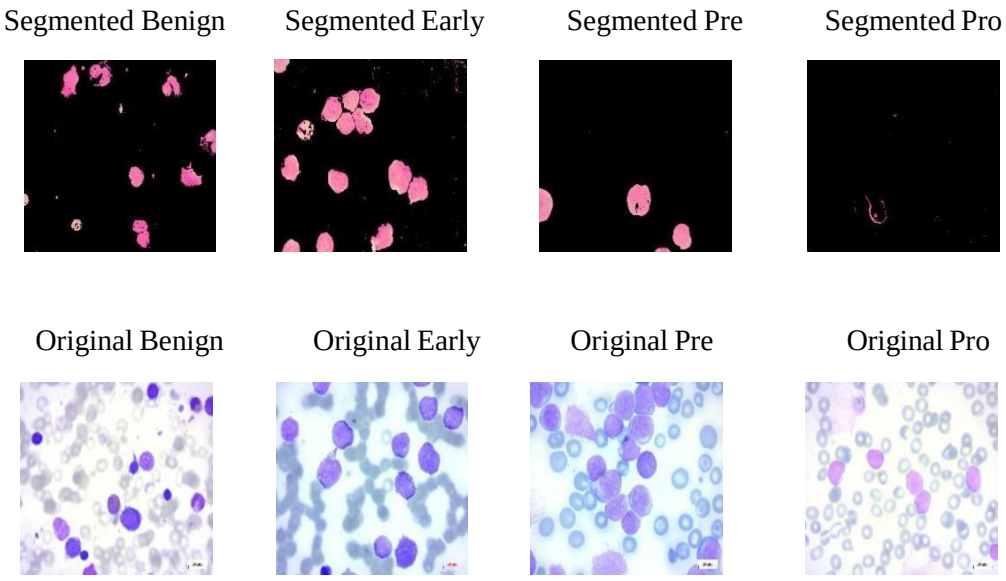

Figure 1. Samples from Acute Lymphoblastic Leukemia (ALL) image dataset

3.2 image enhancement

Image preprocessing and augmentation are two of the most powerful tools that if used correctly it can enhance the image quality hence the results to great extent [7]. Therefore, for original dataset some preprocessing techniques were applied for better accuracy. Zoom 10%, rotation 10, brightness was set randomly from 0.8 to 1.5, and vertical flip and horizontal flip were applied. Moreover, images width and height were set to 224 and rescaled to 1/255.0, Table 1 displays the augmentation and preprocessing used.

Technique used	Range/Type
Zoom range	10%
Rotation range	10
brightness	0.8 – 1.5
rescale	1/255.0
Horizontal flip	True
Vertical flip	True

Table 1. Image preprocessing and augmentation techniques used

3.3 Convolution neural network architectures and support vector machine

Deep learning is a machine learning technique in which many layers of information processing units are used in the unsupervised learning of characteristics and for the analysis or classification of patterns (supervised learning) [20]. The essence of deep learning is to obtain multiple levels of representation, from simple nonlinear modules that transform the representation from one level to a higher and more abstract one. For classification purposes, the higher representation layers amplify the aspects of the input that are most important for discriminating between classes and suppress the variations that are irrelevant. Architecture depth refers to the number of levels of nonlinear operations learned. As the algorithms commonly used in machine learning correspond to superficial architectures, ANN researchers have made efforts to replicate this type of architecture. One of the most critical layers is the convolutional layer, which gives CNNs their

name. This layer performs a linear operation called convolution between filter w(t) and image x(t), and writes (x*w)(t) or s(t), as in Equation (9). There are three parameters that control the convolution layer: filter size, zero padding, and p-step. The larger the filter size, the larger the wrapping around the images. Each filter is designed to detect specific features in the input image. For example, a filter is designed to detect edges, another is designed to detect geometric features, and another is designed to detect textures and colors. Thus, this capability of CNNs is called translation invariance. Zero padding is used to maintain the size of the original input. The size of the zero pad is determined on the basis of the sizes of the convolutional filter and original input. The p-step parameter is used to determine the number of steps taken by the filter on the image at a time. SVM is a powerful computational mathematical model for tasks requiring classifications. SVM is a supervised learning technique applied to classification and regression problems. It has a solid statistical foundation and is quite effective. An SVM performs its classification purpose by building a hyperplane in higher dimensions. A considerable marginal separation between classes is provided by the vector points that make up the decision boundary, which are sought for by the support vector method (SVM). SVM distinguishes classes on the decision plane using the greatest feasible marginal distance.

3.3.1 GoogleNet

Convolutional neural network Inception v3 was developed as a module for GoogLeNet and is used to aid with image analysis and object detection. The third version of Google's Inception Convolutional Neural Network, which was first unveiled for the ImageNet Recognition Competition. Deeper networks were made possible by Inceptionv3's design, which also aimed to limit the number of parameters from exceeding 25 million opposed to 60 million for AlexNet [8].

Similar to how ImageNet may be seen as a database of classified visual objects, Inception aids in the classification of items in the field of computer vision. Inceptionv3 architecture has been applied to a wide range of tasks and is frequently "pre-trained" using ImageNet. It helps in the study of leukemia in the field of life sciences, for example [9]. A high-level diagram of the model is shown in Figure 2 [10].

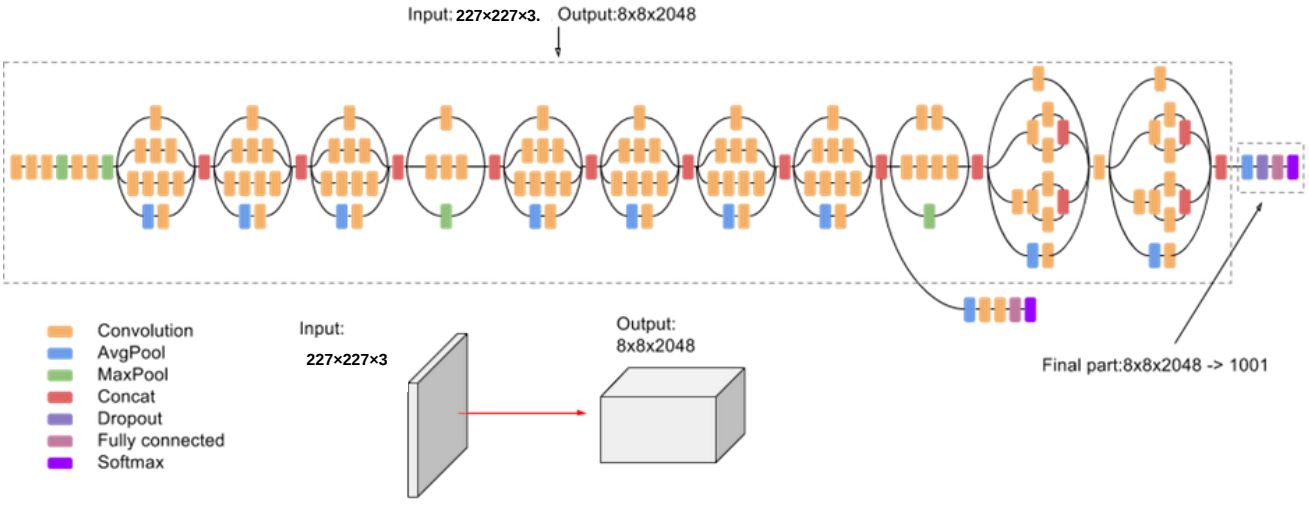

Figure 2. Image illustrates the GoogleNet inception v3 model

3.3.2 GoogleNet with support vector machine algorithm

In this study, a hybrid strategy that combines two algorithms was suggested. The algorithms are SVM and GoogleNet. For hybridization goals, the last layer of GoogleNet Inception V3 was replaced by SVM classifier.

3.3.3 AlexNet

AlexNet is a CNN containing 25 layers divided into many layers from the input of images to the final classification. The most important layers are five convolutional layers,several ReLU layers, three max-pooling layers, two dropout layers, three FCLs, and a softmax layer [23]. AlexNet also contains 650,000 neurons, 62 million parameters, and 630 million connections between neurons. Figure 3 shows the AlexNet architecture and the most critical layers it contains for analyzing the ALL dataset and classifying the data into four classes: Hematogenous, Early Pre-B ALL, Pre-B ALL, and Pro-B ALL.

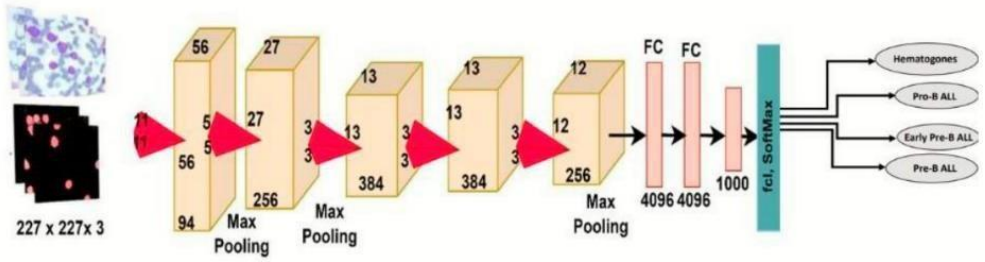

Figure 3. Image illustrates the AlexNet model

3.3.4 AlexNet Model with Support Vector Machine algorithm

Hybrid techniques are proposed because the use of deep learning networks poses some challenges, such as the requirement of high specification computers and the time-consuming process of training the dataset. Thus, the hybrid techniques require

Medium-specification computers, and the process of training the dataset is fast and not time consuming. In this study, hybrid techniques consisting of two blocks each were used. The first block consisted of the CNN models namely AlexNet which extract deep features and feed them into the second block. The second block was the SVM classification algorithm, which classifies the deep feature maps extracted from the CNNs , It can be seen in the figure 4 that the FCLs in the CNNs (the first block) have been replaced by the SVM classifier (the second block)

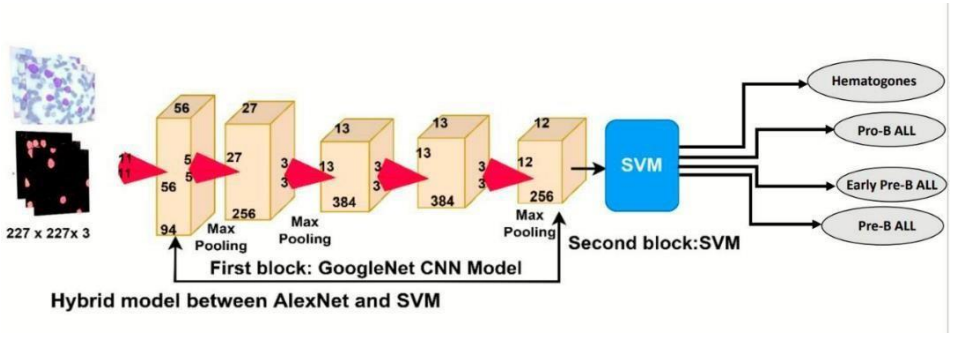

Figure4. Image illustrates the AlexNet model + SVM

3.3.5 ResNet

ResNet-18 is a deep convolutional neural network (CNN) architecture that was introduced in 2015 by Microsoft researchers. It is part of the ResNet (Residual Network) family, which revolutionized image recognition tasks by enabling the training of very deep neural networks. Moreover, resNet-18 has 18 layers, including convolutional layers, pooling layers, and fully connected layers. It is designed to process images of size 224 x 224 pixels and has over 11 million parameters. Alongside this, the network uses skip connections, which allow information to bypass some layers and be directly passed to subsequent layers. This enables the network to learn more complex features and avoid the problem of vanishing gradients that

can occur in deep networks. ResNet-18 is widely used in computer vision tasks such as image classification, object detection, and segmentation. It achieved state-of-the-art performance on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) dataset in 2015, with an error rate of 3.57%. The success of ResNet-18 and its variants has inspired the development of other deep neural network architectures with skip connections, such as DenseNet, Highway Networks, and FractalNet. These architectures have led to significant advances in computer vision and other fields such as natural language processing and speech recognition. In conclusion, ResNet-18 is a powerful and influential deep neural network architecture that has had a significant impact on the field of computer vision. Its innovative use of skip connections has enabled the training of very deep networks and improved performance on image recognition tasks [5].

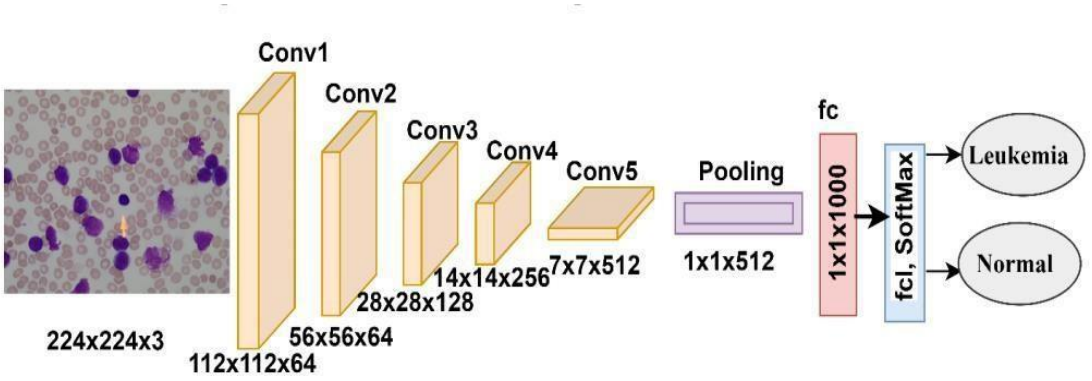

Figure5. ResNet algorithm

3.3.6 ResNet with SVM

ResNet-18 and SVM is to use ResNet-18 as a feature extractor, extracting features from the images and using them as input to SVM. In this approach, we first pretrain the ResNet-18 on a large dataset, such as ImageNet, to learn meaningful features. We then use the trained ResNet-18 as a fixed feature extractor and extract the features from the images in the target dataset. Finally, we use SVM to train a classifier on the extracted features. This approach has been used for various image classification tasks, such as identifying plant species from images and detecting cancerous cells from medical images.

Another approach to combining ResNet-18 and SVM is to use a technique called transfer learning, where we fine-tune the ResNet-18 on the target dataset and then use SVM for classification. In this approach, we start by using the pre-trained ResNet-18 as a feature extractor and extracting features from the target dataset. We then use these features to fine-tune the ResNet-18 on the target dataset. Finally, we use SVM to train a classifier on the fine-tuned features. This approach has been used for various image classification tasks, such as identifying different types of skin lesions from images and detecting different types of crops from aerial images.

In conclusion, combining ResNet-18 and SVM can lead to improved classification accuracy in image recognition tasks. By using ResNet-18 as a feature extractor or fine-tuning it on the target dataset, we can take advantage of its ability to learn meaningful features and then use SVM for classification. The choice of which approach to use depends on the specific task and the characteristics of the dataset [7].

3.4 Training

As mentioned previously, we chose to use segmented cells to train the network to lymphoblast cells, improve the model's accuracy and reliability, and provide a more precise attention map after building and testing various fitting structures and techniques. However, we noticed that much essential information is lost when segmented cells are used instead of the original image. Hence, we considered segmented cells and the original images as two inputs for the network. Each input sample was an image pair that included both the

segmented and PBS versions of the same image. Segmented images and original images were divided into training set, testing set, and validation set as shown in Table 2.

	Train set	Test set	Validation set	Total
Percentage (%)	64	20	16	100
No. of paired samples	2083	652	521	3256

Table 2. How both data from original and segmented datasets were divided between sets [6].

4. Results

The effectiveness of the suggested method is assessed through different techniques and the findings are presented in this section.

4.1 Parameter tuning

All the parameters started operating with random weights for network training. The batch size and number of epochs in the suggested technique were both set to 100 and 32, respectively. The classes were weighted proportionately to the number of samples in order to address the issue of unbalanced data. The use of the Adam optimizer with a 1e 3 starting learning rate helped the gradient descent process perform better. Table 3 displays the techniques used for parameter tuning.

Algorithm / Technique	GoogleNet and GoogleNet+SVM	AlexNet and AlexNet+SVM	ResNet& ResNet+ SVM
Batch size	32	32	16
Epochs	100	220	8
Learning rate	0.001	0.001	0.0001

Table 3. Techniques used for parameter tuning

4.2 Evaluation metrics

In this study, clinically significant statistical parameters like sensitivity, F1 score, accuracy, and precision are used to assess the performance of the proposed detection and classification model. Unfortunately, precision is not always sufficient to assess the model's performance, particularly when dealing with an asymmetrical data set. As a result, different performance measures must be assessed in order to test the model [10]. Table 4 shows the results of each evaluation metric used

Evaluation metric	GoogleNet	GoogleNet + SVM	AlexNet	AlexNet + SVM	ResNet	ResNet + SVM
Accuracy	95.3%	98.4%	95.7%	97.8%	96.3%	98.7%
Sensitivity	95.8%	98.7%	95.3%	97.8%	98.9%	98.6%
F1 score	95.5%	98.5%	95.1%	97.3%	95.9%	98.4%
Precision	95.7%	98.6%	94.9%	96.4%	96.7%	98.7%

Table 4. Comparing results of different algorithms with and without SVM

5. Future directions

There are various things are put into consideration for future enhancement. For instance, it is planned to use 220 epoch and 32 patch size in the future on the ResNet and hybrid ResNet algorithms. Furthermore, due to CPU limitation of the pc regarding GoogleNet and hybrid GoogleNet, no more than 100 epochs could be done. Therefore, it is planned to implement 220 epochs and using remote servers in the future for yielding better results.

6. Conclusion

One of the limitations of this work is that the data set is not big enough to train CNN models, which is necessary to prevent overfitting. The data augmentation technique was used to get around this issue. GoogleNet inception v3, ResNet and AlexNet which are CNN model were used for this study. All three models for identifying ALL and categorizing its subtypes. The current models proposed two image input channels (original images and images segmented using the suggested method), which helped cover more extensive and pertinent characteristics of the feature space. Almost all previous research had employed public data sets with limited data, whose images were based on single cell images which, in terms of hematology, deviated from actual samples at the laboratory (including multicellular slides). As a result, the pattern recognition issue in this instance, which was based on a group of cells, was closer to the hematologist's conclusion. Alongside this, to get better outcome hybrid models was used by adding support vector machine classifier to the applied models. The accuracy after applying GoogleNet using 100 epochs was 95.3%, AlexNet using 220 epochs was 95.7%, and ResNet using 8 epochs was 96.3%, however, when adding SVM the accuracy extremely improved to 98.4% for GoogleNet, 97.8% for AlexNet, and 98.7 for ResNet.

7. References

1. Puckett Y, Chan O. Acute Lymphocytic Leukemia. [Updated 2022 Jun 27]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459149/

2. Terwilliger T, Abdul-Hay M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. Blood Cancer J. 2017 Jun 30;7(6):e577. doi: 10.1038/bcj.2017.53. PMID: 28665419; PMCID: PMC5520400.

3. Sawyers CL, Denny CT, Witte ON. Leukemia and the disruption of normal hematopoiesis. Cell. 1991 Jan 25;64(2):337-50. doi: 10.1016/0092-8674(91)90643-d. PMID: 1988151.

4. Dong Y, Shi O, Zeng Q, et al. Leukemia incidence trends at the global, regional, and national level between 1990 and 2017. Exp Hematol Oncol. 2020;9(1):1-11.

5. Bain BJ. Diagnosis from the blood smear. N Engl J Med. 2005 Aug 4;353(5):498-507. doi: 10.1056/NEJMra043442. PMID: 16079373.

6. Abunadi I, Senan EM. Multi-Method Diagnosis of Blood Microscopic Sample for Early Detection of Acute Lymphoblastic Leukemia Based on Deep Learning and Hybrid Techniques. Sensors (Basel). 2022 Feb 19;22(4):1629. doi: 10.3390/s22041629. PMID: 35214531; PMCID: PMC8876170.

7. Shahriar, Md & Li, Huyue. (2020). A Study of Image Pre-processing for Faster Object Recognition.

8. Szegedy, Christian (2015). "Going deeper with convolutions" (PDF). CVPR2015.

9. Milton-Barker, Adam. "Inception V3 Deep Convolutional Architecture For Classifying Acute Myeloid/Lymphoblastic Leukemia". intel.com. Intel. Retrieved 2 February 2019.

10. Ghaderzadeh M, Asadi F, Jafari R, Bashash D, Abolghasemi H, Aria M. Deep convolutional neural network–based computer-aided detection system for covid-19 using multiple lung scans: design and implementation study. J Med Internet Res. 2021;23(4):e27468. doi:10.2196/27468