Damascus university Journal for engineering sciences
Vol 39 No.2 (2023) :271-279

مجلة جامعة دمشق للعلوم الهندسية
المجلد 39– العدد 2 (2023) 271-279

"Detection of Acute Lymphoblastic Leukemia (ALL)" "Using Deep Learning and Transfer Learning"

Eng. Thouraya G Mouhssiena[1] Eng. Diana T Karbouj[2] Eng. Rasha N Mayhoub[3] PhD. Taghrid Haddad[4]

1: Scientific researcher in an academic master's degree - Biomedical Engineering Department - Faculty of Mechanical and Electrical Engineering - Damascus University.

2,3: Student at Department of Electronics and Communications Engineering - Faculty of Mechanical and Electrical Engineering - Tishreen University.

4: PhD in the Department of Electronics and Communications Engineering - Faculty of Mechanical and Electrical Engineering - Tishreen University.

Abstract

Leukemia results from abnormal and excessive multiplication of white blood cells, and as a result, the immune system in our bodies is destroyed and this leads to death. Annually, an estimated 300,000 new cases of Leukemia are diagnosed, which constitutes 2.8% of the total diagnosed cancers [2]. Of the four different types of Leukemia, Acute Lymphoblastic Leukemia (ALL) is the most serious and deadly type of blood cancer. It spreads very quickly and is fatal within weeks if not treated. Statistically, about 90% of patients can be cured, provided the disease is diagnosed at an early stage [2], so early diagnosis of (ALL) is definitely essential. In manual methods, pathologists diagnose it by a bone marrow biopsy test or an ocular microscopic testing of a blood sample. Although this method is very effective in diagnosis, it takes a long time and requires a lot of experience, so in this case, computer aided diagnosis (CAD) can be considered as a wonderful auxiliary diagnostic tool if computer aided diagnostic systems with high accuracy are developed to diagnose the disease that support Doctor's view. Several supervised and unsupervised machine learning algorithms have been proposed for the detection of (ALL).

In our work, we rely on samples of Acute Lymphocytic Leukemia micrographs from a dataset known as C-NMC being the largest available dataset aiming of building a deep learning model (a convolutional neural network CNN) and using the Transfer Learning for the EfficientNetB3 model to build a system that supports the Diagnostic view of the doctor. At the end of the work, after processing the data and training the model, we obtained a training accuracy of 98.623% and a testing accuracy of 97.75%, which is promising in the development of such systems.

Keywords: Acute Lymphoblastic Leukemia, Leukemia, Deep Learning Dl, Transfer Learning Tl, Convolutional Neural Network Cnn.

Received: 14/4/2023
Accepted: 7/5/2023

" تشخيص ابيضاض الدم اللمفاوي الحاد باستخدام التعلم العميق ونقل التعلم "

م. ثريّا جبران محسنه[1] م. ديانا طارق كربوج[2]

م. رشا نديم ميهوب[3] د. تغريد حداد[4]

1. باحث علمي في ماجستير أكاديمي _ قسم الهندسة الطبية _ كلية الهندسة الميكانيكية والكهربائية _ جامعة دمشق.

2,3. طالبة في قسم هندسة الإلكترونيات والاتصالات _ كلية الهندسة الميكانيكية والكهربائية _ جامعة تشرين.

4. دكتور في قسم هندسة الإلكترونيات والاتصالات _ كلية الهندسة الميكانيكية والكهربائية _ جامعة تشرين.

الملخص

ينتج ابيضاض الدم عن تكاثر شاذ ومفرط لخلايا الدم البيضاء فيتدمر نتيجة لهذا الجهاز المناعي في أجسامنا ويؤدي ذلك للموت. سنويا يشخص ما يقدر بـ 300 ألف حالة جديدة من ابيضاض الدم وهو ما يشكل 2.8% من إجمالي السرطانات المشخصة **[2]**. من بين أربع أنواع مختلفة من ابيضاض الدم يكون ابيضاض الدم اللمفاوي الحاد **acute lymphoblastic leukemia (ALL)** أخطر أنواع سرطانات الدم وأكثرها فتكا فهو سريع جدا بالانتشار وقاتل في غضون أسابيع إذا لم يتم علاجه. إحصائيا يمكن علاج حوالي 90٪ من المرضى بشرط تشخيص المرض في مرحلة مبكرة **[2]** لذا يعد التشخيص المبكر للـ**ALL** ضروري للغاية. في الطرق اليدوية يقوم أخصائيو علم الأمراض بتشخيصه عن طريق اختبار الخذع من نقي العظام أو اختبار مجهري عيني لعينة دموية. وعلى الرغم من أن هذه الطريقة فعالة جدا في التشخيص إلا أنها تستغرق وقتا طويلا وتحتاج لخبرة عالية لذا في هذه الحالة يمكن اعتبار التشخيص بمساعدة الحاسوب **computer aided diagnosis (CAD)** أداة تشخيص مساعدة رائعة إذا ما تم تطوير أنظمة تشخيص حاسوبي ذات دقة عالية لتشخيص المرض تدعم رأي الطبيب وتسانده. تم اقتراح العديد من خوارزميات التعلم الآلي الخاضعة لإشراف وغير الخاضعة للإشراف للكشف عن **ALL**.

في عملنا هذا فإننا نعتمد على عينات من صور مجهرية لابيضاض الدم اللمفاوي الحاد مأخوذة من مجموعة بيانات معروفة باسم **C-NMC** كونها أكبر مجموعة بيانات متاحة بهدف بناء نموذج تعلم عميق (شبكة عصبية التفافية **CNN**) وباستخدام آليات نقل التعلم لنموذج **EfficientNetB3** لبناء نظام يدعم رأي الطبيب في التشخيص في نهاية العمل وبعد معالجة البيانات وتدريب النموذج حصلنا على دقة تدريب بلغت **98.623%** ودقة اختبار **97.75 %**. وهو ما شكل أمر واعد في تطوير هكذا نظم.

الكلمات مفتاحية: ابيضاض الدم، سرطان الدم، اللوكيميا، التعلم العميق، نقل التعلم، الشبكة العصبية الالتفافية.

تاريخ الإيداع:2023/4/14

تاريخ القبول: 2023/5/7

I. Introduction:

Leukemia results from abnormal and excessive multiplication of white blood cells, and as a result, the immune system in our bodies is destroyed, and this leads to death. Leukemia constitutes 2.8% of all diagnosed cancers. According to the statistics of the International Agency for Research on Cancer of the World Health Organization (IARC), the number of leukemia cases in both sexes of all ages in 2018 reached 4,37,033 with a total death toll of 3,03,006. Globally, the infection rate per 100,000 was 5.2, and the death rate per 100,000 was 3.5 [2]. Of the four different types of Leukemia, Acute Lymphoblastic Leukemia (ALL) is the most serious and deadly type of blood cancer. It spreads very quickly and is fatal within weeks if not treated. Statistically, about 90% of patients can be cured, provided the disease is diagnosed at an early stage [2], so early diagnosis of (ALL) is definitely essential. In manual methods, pathologists diagnose it by a bone marrow biopsy test or an ocular microscopic testing of a blood sample. Although this method is very effective in diagnosis, it takes a long time and requires a lot of experience, so in this case, computer aided diagnosis (CAD) can be considered as a wonderful auxiliary diagnostic tool if computer aided diagnostic systems with high accuracy are developed to diagnose the disease that support Doctor's view. This is the aim of this research, as we will use Deep Learning (CNNs) and evaluate the Transfer Learning effectiveness of the EfficientNetB3 model in diagnosing (ALL).

II. Acute lymphoblastic leukemia (ALL):

Blood is a connective vital tissue that consists of red blood cells, white blood cells, platelets, and plasma. All cellular components of blood arise from the differentiation of stem cells, as shown in Figure (1). Blood performs respiratory, nutritional and defensive functions and regulates body temperature, among others.

White blood cells make up 1% of the blood, and they are cells with a very important immune function in the human body, so they protect it from diseases. It consists of three types of granulocytes, lymphocytes, and monocytes, according to the place of stem cell differentiation, which is the:

Either bone marrow, which produces granulocytes that live for several hours only and include:

- Basal: make up 0.5% of white blood cells.
- Acidic: make up 1-4% of white blood cells.
- Neutrophils: make up 60-70% of white blood cells.

Or lymph nodes and tissues, which produce non-granular (not having granules in the cytoplasm) and live for several days or even years. They include:

- Monocytes: make up 3-8% of white blood cells.
- Lymphocytes: make up 20-25% of white blood cells.

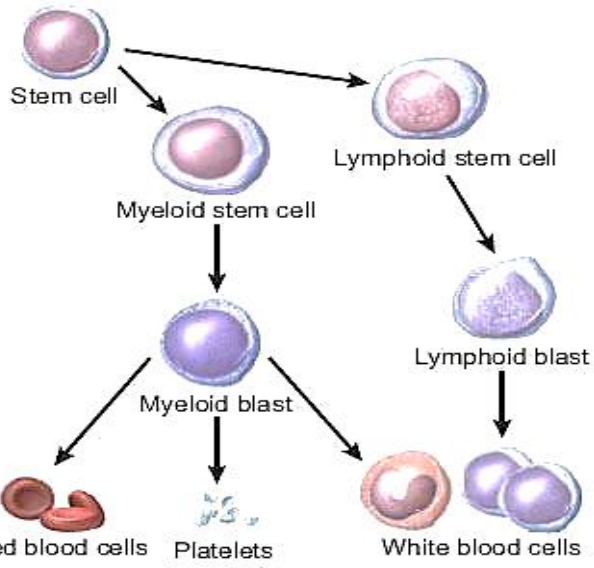


Figure (1) Differentiation of stem cells to give blood components.

Blood diseases and disorders include either an increase in the production of blood components or a decrease in it, or an increase in the degeneration of its components or a deficiency.

In general, the increase in the uncontrolled production of mature or immature white blood cells (they cannot perform their functions as they should) lead to Leukemia. The causes of this type of cancer are unknown to doctors yet, but they use drugs, radiation, and blood transfusions to treat them.

In general, there are two types of white cells, lymphocytes and myeloid cells. Leukemia is also classified in two ways: acute or chronic according to the speed of cell growth. Therefore, according to the French American British (FAB) classification model, the four main forms of Leukemia are:

- Acute Lymphoblastic Leukemia (ALL)
- Acute Myeloid Leukemia (AML)
- Chronic Lymphocytic Leukemia (CLL)
- Chronic Myeloid Leukemia (CML)

(ALL) is defined as a precancerous disease of the lymphocytes, particularly B-cell (B-ALL). The disease is characterized by the abnormal multiplication of lymphocytes and lymph nodes in the bone marrow, then the disease spreads to the rest of the organs. The disease is usually attributed to a molecular molecule that delays or disrupts the process of cell apoptosis.

Microscopic testing of blood samples and bone marrow biopsy is the most common method for diagnosing the disease. In these tests, the specialist checks whether the amount of white blood cells is high enough and other observations are taken into consideration. These manual methods which are with limited experience prone to error because their results depend greatly on the knowledge and skills of the expert conducting the testing. In addition to the fact that the huge number of samples takes a long time, which delays the effectiveness of treatment. Thus, developing accurate and reliable computational methodologies for Leukemia identification is important for early diagnosis and effective treatment.

III. Previous studies:

First study [1]: Tahmina Akter Sumi and his colleagues in 2022 built the CAD-ALL system, and was based on blood micrographs collected from the ALL-IDB database. For this reason, the images were initially processed using an intermediate filter in order to reduce noise and improve image quality. Due to the lack of data and images, dataset was augmented by including slightly modified versions of the images already in the dataset. Then all these images are used to train the CNN model and the following layers were used: Convolution, ReLU, Pooling layer, and a fully connected layer.

The results were: 100% training accuracy and 97.89% testing accuracy.

This study gave good results, but the used database has few data despite the work to increase it.

Second study [2]: Done by Md. Taufiqul Haque Khan Tusar and his colleagues in 2022. In this study, the DNN model was used to train for diagnosing subtypes of (ALL) from microscopic images in a database obtained from the bone marrow laboratory of Taleghani Hospital in Tehran, Iran. It contains images of benign and (ALL) cells from three stages: early, pre, and pro. It consisted of 3256 peripheral blood smear images taken from 89 patients with suspected (ALL), 25 healthy ones, and 64 patients with a definitive diagnosis of all three-stage B-cell (ALL). The studied DNN model was built with four types of CNN model and three models with the aim of evaluating Transfer Learning effectiveness, which is MobileNetV2, ResNet50 and VGG19.

All models' results are shown in Table [1]:

Table [1]: results for second study..

	Accuracy	Loss
MobileNetV2	0.9742	0.2351
VGG19	0.9613	0.099
CNN	0.9128	0.2309
ResNet50	0.8526	0.8412

Analyzing the results in the study, we can conclude that with the exception of ResNet50, other models provide desirable results. The MobileNetV2 model performs higher in terms of

training accuracy and VGG19 provides the lowest loss.

Third study [3]: In 2018 Rehman A. Abbas proposed a method for classifying (ALL) between its subtypes and normal cells from bone marrow stain images. The database was obtained from a laboratory in Pakistan for blood micrographs of normal and (ALL) cases, which are 100 for L1, 100 for L2, 30 for L3 (as L1, L2 and L3 are classifications of (ALL) according to the shape of white cells) and 100 for normal cases. Powerful segmentation and deep learning techniques are used with the Alexnet and CNN models to train the model on images to achieve accurate classification results.

The results were: The experimental results reveal that the proposed method achieved an accuracy of 97.78%.

The results are excellent, but the data were somehow scarce and the images of (ALL) type L3 are less than others, which may cause bias in the model, so the data had to be equalized (Balancing) before training, and this was not mentioned in the study.

IV. Methodology:

To complete the proposed project and build an integrated model for diagnosing (ALL), we worked according to successive stages, as follows:

1) Determine tools needed to do the research:

- ✓ The chosen language is python for its speed, simplicity, and being the world's first language in the field of artificial intelligence and deep learning.
- ✓ The most important libraries are KERAS and TENSORFLOW.
- ✓ For the code editor we used Jupyter Notebook.
- ✓ We identified the necessary dataset which is C-NMC.
- ✓ We will also apply the transfer learning methodology of the EfficientNetsB3 model.

2) Working algorithm:

- ✓ import selected dataset, libraries and Transfer Learning model whose effectiveness we are evaluating in this medical condition.
- ✓ Reviewing, pre-processing and preparing the data to give after the training a model with high reliability.
- ✓ Build, compile, train and test the model.
- ✓ Reviewing and justifying the results.
- ✓ Using the model for diagnosing individual cases.

Figure [2] shows this algorithm in a simplified way.

1st Step (Data Set Selection): In the first step of our work in this research, we identified the dataset that we will deal with. There are many datasets used to develop models for the classification of Leukemia, among which we mention the most famous (ALLIDB1 / ALLIDB2 / Atlas / BCCD / C-NMC), each is described in Table [2].

For this work we chose the C-NMC data set [4] as it is the largest data set ever available. This dataset has been released by the Cancer Image Archive (TCIA) with intent of launching a challenge to develop computer-aided diagnostic (CAD) models to distinguish between normal cells and Acute Lymphoblastic Leukemia (ALL) in processed micrographs of blood cells. Where these cells were segmented from the microscopic images which contain some errors of staining, lighting and noise, although they were processed a lot during the acquisition stage, but these errors were left because they represent images from the reality. This dataset has 15,135 images taken from 118 people classified into two categories: Hematogone and (ALL).

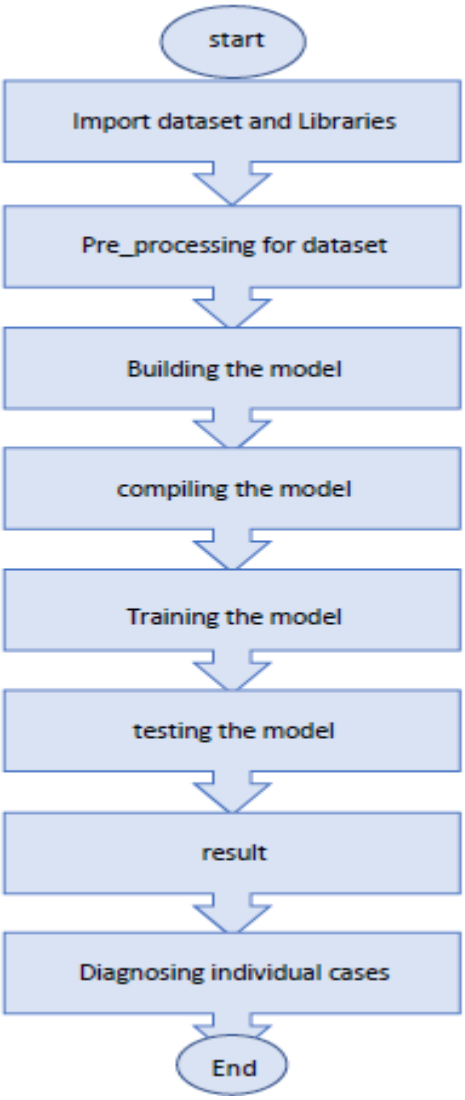


Figure [2]: Search algorithm.

Table [2]: Available databases.

dataset	description
ALLIDB1	It has 108 images that contain multiple leukocytes. Out of these 108 images, 59 are healthy, and 49 are ALL.
ALLIDB2	It has 130 healthy and 130 ALL images.
BCCD	it consists of two sets of data. In first set, it contains 367 microscopic images (without augmentation). In another set, it contains 12444 images (with augmentation).
C-NMC	It contains +15000 cancer cell images of B-ALL along with healthy images.

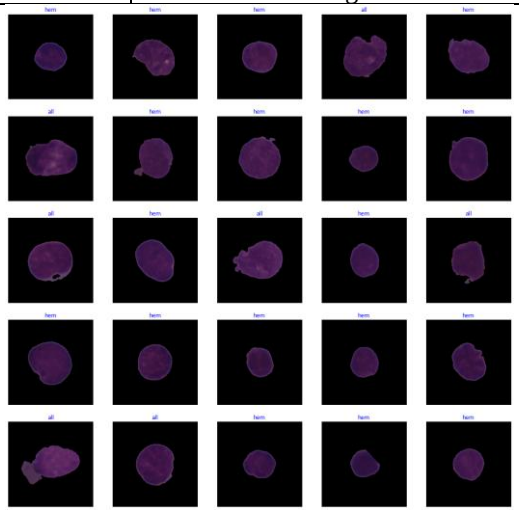


Figure [3]: A sample of this data.

2nd step (browsing, pre-processing and initializing the database): After reviewing the dataset and its description, we found that the images in were pre-processed so that the cells were separated from the microscopic images and had some staining and lighting errors to represent real images. This means that the images do not need any pre-processing to be submitted to the model, but in order to avoid the model to side with a specific category (healthy / ALL), we did the following:

I. Create Data Frames:

Images in the dataset are divided into three training / evaluation / testing sets. In the training group, which has approximately 10,000 images, the images are divided into three volumes. We took the images and their classifications from all these volumes, combined them together in one framework, and then divided them into training, testing, and evaluation images. If we rely on these images only, they are sufficient and there is no need to use all the images in dataset.

At the end of this stage, we got a frame containing about 1000 images divided as follows: 90% for training, 5% for evaluation, and 5% for testing.

- train_df length: 9594
- test_df length: 534
- valid_df length: 533

II. Data Balancing:

By exploring the images in our data frame, we found that they are unbalanced. The healthy images within the training set of the frame are more than the cancerous ones, 6544 normal and 3050 cancerous. Training in this way will make the model biased to diagnose normal cases more accurately.

To prevent this from happening, there are several ways to balance the training data we have, so we reduced the number of healthy cases until it equals the number of cancerous cases, by 3050 images for each training class.

III. Augmentation:

After this, we used automated image generators to improve and increase the education data. Several methods are used to increase the data, including flipping, reversing, rotating, contrasting, and others to obtain different copies of the same image. This gives the algorithm more generalizability later when we build a trained system on the original images and the merged images generated from them. Studies have reported that training with augmented images reduces error rates and provides better generalization.

We used the horizontal-flip operation whose role is to randomly perform horizontal flip operations on the images which means that horizontal flips are not necessarily performed on all images and each epoch is randomly selected to flip.

3rd step (building model): With the aim of transferring learning to the EfficientNetB3 model in order to verify its effectiveness in the case of diagnosing ALL which is the most effective and efficient CNN architecture. Use first as input to our model and then provide output of this model to Dens layer with ReLU activation function. But placing a BatchNormalization layer before it makes the normalization process of these models more generalizable and solves the problems of covariate shift, and also makes neural networks faster and more stable by normalizing or processing the inputs of the next layer by re-centering and re-distributing (standardizing) the output of the previous layer. BatchNormalization affects the output of the previous activation layer by subtracting the Batch average, then dividing by the Batch standard deviation. Then the Droppout layer excludes 45% of the neurons that will enter the next layer, but randomly in order to reduce the possibility of overfitting. Then the last output layer Dens, by two output neurons and a softmax activation function.

For the weight optimization training algorithm we used the Adam optimization algorithm. This algorithm has proven its efficiency for a wide range of neural network architectures. As for the Loss Function, we relied on the categorical-cross entropy function, and the measure adopted during training is the Accuracy. Figure [4] shows the architecture of the proposed model.

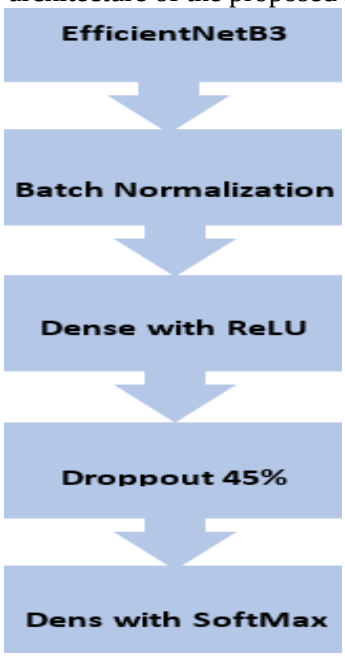


Figure [4]: The architecture of the proposed model.

4th step (start training): The method of training our network is (Supervised) and (Batch training). We asked the model to stop the training if the learning rate did not improve over three consecutive epoch, and the training began. Indeed, the training continued until the 17th epoch, and then stopped because the learning rate did not improve after the 14th epoch. Therefore, the weights of the model 14 were set to be the final weights.

5th step (Reviewing and Interpreting Results): By drawing the curves of the loss function across the different epochs of the training and evaluation data. As well as accuracy, it can simply be noted that the accuracy increases and the loss function decreases, and the best values are at the 14th epoch, as we said. Which gave the data shown in Table [3].

Table [3]: Accuracy training and evaluation results for the final model.

Epoch	Loss	Accuracy	V_loss	V_acc
14	**0.158**	**98.623**	**0.21949**	**95.685**

Figure [5] and [6] show the Loss and Accuracy curves, respectively, during the training process.

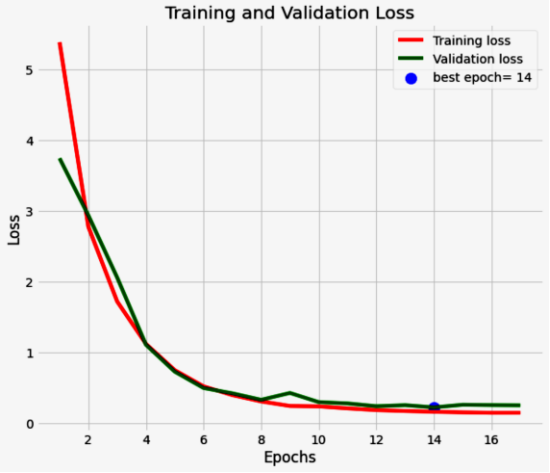


Figure [5]: Loss curves across the 17 epochs.

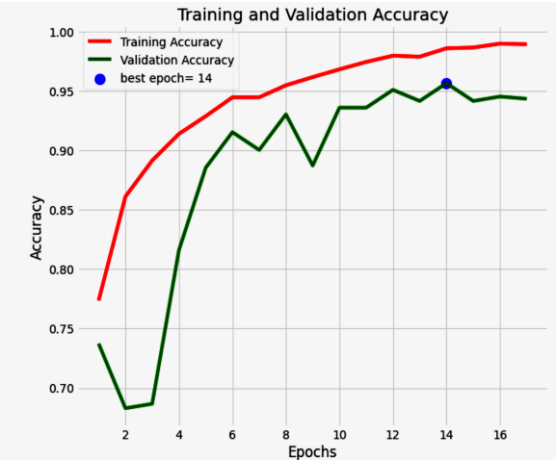
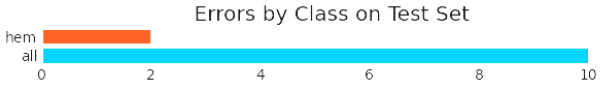


Figure 6: Accuracy function curves across the 17 epochs.

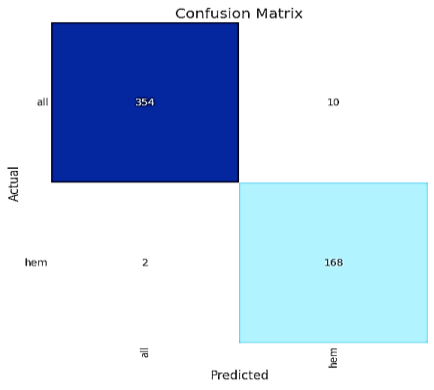
6th step (Testing): After we saved the previous model, we used it with the testing data, which gave the following results for the accuracy of the test:

There were 12 errors in 534 test cases.

Model accuracy = 97.75%.



To find out the accuracy of the model, we used the confusion matrix, which is shown in Figure [7].



	precision	recall	f1-score	support
all	0.9944	0.9725	0.9833	364
hem	0.9438	0.9882	0.9655	170
accuracy			0.9775	534
macro avg	0.9691	0.9804	0.9744	534
weighted avg	0.9783	0.9775	0.9777	534

Figure [7]: Confusion matrix and an approximation of the results of the model testing.

7th step (Building a simple application to use the model in prediction): Finally, a simple application was built to diagnose individual cases by passing an image to it, then it gives the diagnosis. In Figure [8], an illustration of the output of the application and its diagnosis for one of the cases, the diagnosis continued at 280ms/step, and the input image was diagnosed to be (ALL) by 95.74%.

```
input Y if you want to predict or N to stopy
pleas enter the path for your image/content/drive/MyDrive/36.bmp
1/1 [==============================] - 0s 280ms/step
the image is predicted as being all with a probability of  95.74 %
input Y if you want to predict or N to stopn
```

Figure [8]: Using the model in diagnosing cases.

IV. Conclusion:

In this paper, we evaluate the effectiveness of the EfficientNetB3 learning transfer model, which is one of the most important and efficient CNN models. We have achieved our goal of the project by building an acute lymphoblastic leukemia diagnostic system, and we obtained, using CNN and transfer learning models, a model with a training accuracy of 98.623% and a test accuracy of 97.75%. These very promising results lead us to say that the proposed model is generalizable and can now be used to assist the clinician in the diagnostic process. As future results and recommendations, we recommend building a better graphical interface for the application and generalizing its use, for example, by building a suitable web application to help the specialist in diagnosis, and update the model with new data.

V. References:

[1]. Sumi,Tahmina Akter, Automated Acute Lymphocytic Leukemia (ALL) Detection Using Microscopic Images: An Efficient CAD Approach. 2022. Springer. Part of the Lecture Notes in Networks and Systems book series (LNNS,volume 376).

[2]. Taufiqul.T, Anik.R, 'Automated Detection of Acute Lymphoblastic Leukemia Subtypes from Microscopic Blood Smear Images using Deep Neural Networks', Degree of Bachelor Department of CSE City University.2022.

[3]. Rehman, A, Abbas, N, Saba, T, Rahman, SIU, Mehmood, Z and Kolivand, H (2018) Classification of acute lymphoblastic leukemia using deep learning. Microscopy Research and Technique, 81 (11). ISSN 1097-0029

[4]. 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

[5]. Samaan, Suhail, Al-Salem. Sharif, Solomon, Marwan. (2012). Hematology and oncology. The specialized medical encyclopedia. Volume VIII. Arabic Encyclopedia Authority.