Leukemia Cancer Detection Using Various Deep Learning Algorithms

Devanshu Joshi, Rishabh Tater, Priya Yaday, Tripti Jain, Preeti Nagrath

Bharati Vidyapeeth's College of Engineering, New Delhi, India
Emails: joshidev1012@gmail.com ; rishabhtater14@gmail.com ; py9.priya@gmail.com ;
triptijain.tj.jt@gmail.com ; preeti.nagrath@bharatividyapeeth.edu

Abstract

Leukemia is a type of blood cancer. Leukemia is cancer that begins in the blood cells. The lymphocytes and other blood cells are created in the bone marrow. When a person has leukemia the bone marrow does not function properly. Leukemia cells are produced by the bone marrow. Leukemia cells are mainly referred to as "rupture". These naive cancer cells block the cells that create the bone marrow. In this paper, various approaches to the classification & automatic detection of leukemia are described. The experiment was successfully implemented in Kaggle. Deep Learning algorithms were largely used in the treatment of Leukemia for the classification & detection of its presence in a patient. The paper describes Convolutional Neural Networks (CNN) and Visual Geometry Group-16(VGG-16) algorithms that are used to categorize leukemia into its sub-types and presents a comprehensive study of these algorithms.

Keywords: Leukemia; Cancer; WBC; Convolutional Neural Networks; Visual Geometry Group-16; AML; Deep Learning;

1. Introduction

There are three important components of Human Blood, namely Red Blood cells which carry oxygen, White Blood cells which fights infection, and Platelets which help blood clot. Bone marrow produces billions of blood cells every day. When Bone marrow makes more white blood cells rather than Red blood cells, called leukemia. Leukemia is a type of cancer that is in White blood cells due to abnormal blood cells. There are mainly four types of leukemia ALL, AML, CLL, and CML. Bone pain, excessive sweating, and red spot on the skin are the symptoms of Leukemia cancer.

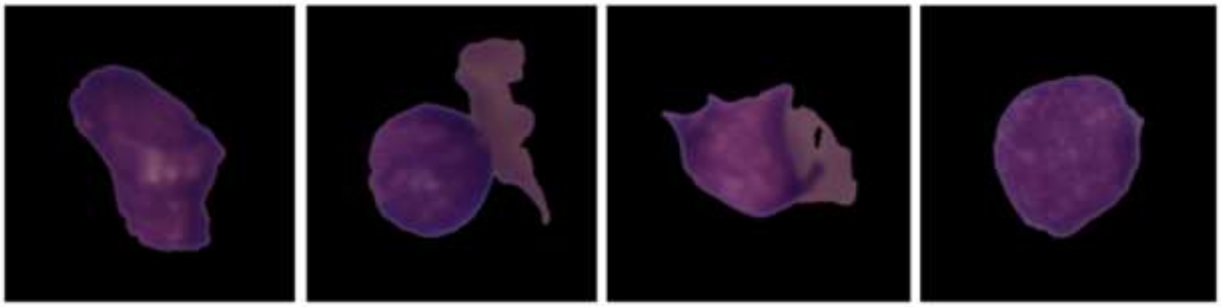

Figure 1: Image of Blood Cells

Doi: https://doi.org/10.54216/FPA.090106
Received: April 10, 2022 Accepted: August 23, 2022

In the Medical field, there is so much research done to solve the problem of leukemia. Treatment of leukemia depends on the stage of cancer and also depends on the type of cancer.

The project is to locate immature cells using special image processing techniques and count the total range of cells. So we want to apply generation that identifies specific forms of blood cells within a quick duration of time in an emergency. Moreover, it's crucial to have a look at how to distinguish the cells and understand them as immature cells and in keeping with it, stumble on leukemia.

A vital and time-ingesting step is the identification of leukemia through an optical blood smears exam monitored by an expert. To clear up such issues, Many CAD (laptop-aided prognosis) methodologies for quantitative evaluation of peripheral blood samples have evolved the use of machine learning and deep learning techniques. but, those tactics have some drawbacks and efficiency and want upgrades in phrases of accuracy, and mastering technique. To summarize and investigate related paintings, in addition to gathering and analyzing patient datasets.

Biomedical picture processing can improve the venture and permit it to be completed in near-real time.

In this paper, Section 2 describes the current leukemia blood cancer analysis. Section 3 will introduce the recommended approach. Section 4 summarizes the experimental outcomes. Section 5 brings the process to a close. The experimental findings demonstrated that the proposed system is both accurate and efficient.

2. Related Work

Blood cancer is a serious disease that has severely impacted people's lives. As a result, the findings of several studies have been established in order to combat this disease and make the world a better place to live. Leukemia cancer detection is accomplished through a series of procedures that begin with obtaining critical information from pre-defined studies and then applying our model to achieve the best results. Using various image processing approaches, Himali P.Vaghela et al.[2] established a model for recognizing immature cells and counting the total number of acute cells. Immature cells are detected using a variety of techniques, including histogram equalization, linear contrast stretching, and morphological techniques such as region opening, region closure, erosion, and expansion. The accuracy of watershed transformation, K-means, histogram equalization, linear contrast stretch, and shape-based features are 72.2%, 73.7%, and 97.8%, respectively. Nimesh Patel and Ashutosh Mishra[12] proposed an automated system that can detect leukemia from a microscopic image, Improves the accuracy required to detect leukemia with a manual approach, and saves time. The main goal of this work is to separate lymphocytes and monocytes from white blood cells in order to detect leukemia. N.Poornima et al.[9] Proposed a method to detect leukemia from blood microscope images using different image processing algorithms. Prasidhi G. Fal Dessai 1 and Geeta Shet2[11] proposed an online database of images of various types of leukemia. All this research shows that image segmentation is the most important step in processing leukemia images. The accuracy of the results of the next step (feature extraction and classification) is based on the segmentation output. Sara Hosseinzadeh Kassani et al.[23] proposed a hybrid deep learning method for distinguishing immature leukemic blasts from normal cells using the ISBI 2019 dataset. Dr. R.JANAKI[3] proposed a method for detecting leukemia and blood cancer. The main goal of this work is core segmentation following feature extraction to detect leukemia. Core shape features such as area and perimeter are taken into account to improve detection accuracy.

Two datasets, ASH image bank, and ALL IDB, Nighat Bibietal.[4] We proposed the discovery of a leukemia subset. The ASH image database can be freely accessed on the Internet. The proposed approach uses an IoT-enabled microscope to send images of a blood smear to the leukemia cloud. The ResNet-34 or DenseNet-121 models are used to diagnose leukemia.

Ahmed T. Sahlol [5] created the model based on data from the Institute of Information Technology at the University of Milan. This paper proposes a hybrid classification strategy for classifying leukemia images. It uses the Deep Convolutional Neural Network (VGGNet) to extract features from WBC images, and the statistically enhanced Salp-Swarm algorithm filters the resulting features to extract only relevant features, and so on. It works by discarding what is not. A proposed review by Siraj Khan et al.[25] The goal is to find the optimal TML and DL approach for MIA, especially for leukocyte classification of blood smear images. Maryam Bukhari et al. [24] This study proposed a new deep-learning alternative for diagnosing leukemia disease by evaluating microscopic images of blood samples. Includes squeeze and excitation learning to recursively recalibrate the functional output per channel by explicitly modeling channel dependencies. At all levels of characterization, the proposed deep learning architecture prioritizes channel allocation. This section contained all previous studies on the detection of leukemia cancer. The next section describes how the model can help detect this deadly disease.

Doi: https://doi.org/10.54216/FPA.090106
Received: April 10, 2022 Accepted: August 23, 2022

3. Leukemia Detection Using Machine Learning Models

A. Methodology

The authors gave an approach that consists of various categories: Firstly, by basic visual techniques, the visual information quality of each inserted image is enhanced. Applied various methods of processing the images and the addition of data are used to increase the visibility of important characteristics of the images. Once the editing process of the images is done and used in feature extraction a section with a proposed hybrid design. We are presently conducting tests using two structures namely, the VGG16 & CNN. The VGG16 hybrid model is one of the simplest but most effective methods.

The VGG16 hybrid model consists of 13 convolutional layers using 3 x 3 convolution filters which are followed by two main layout layers and 4096 full-connected layers, which again act in accordance with a softmax classifier. The architecture of the VGG 16 model is basically designed for image classification. This structure contains deep convolutional layers. The C-NMC-2019 database is being tested with the VGG16 architecture. As each model has its own features, the models were integrated in order to apply the benefits of both structures in sequence to improve the overall accuracy of the prediction. A multi-layer perceptron trains the released features which lead to the detachment of each image in related classroom opportunities. Finally, the performances of the proposed structures are explored in the test images.

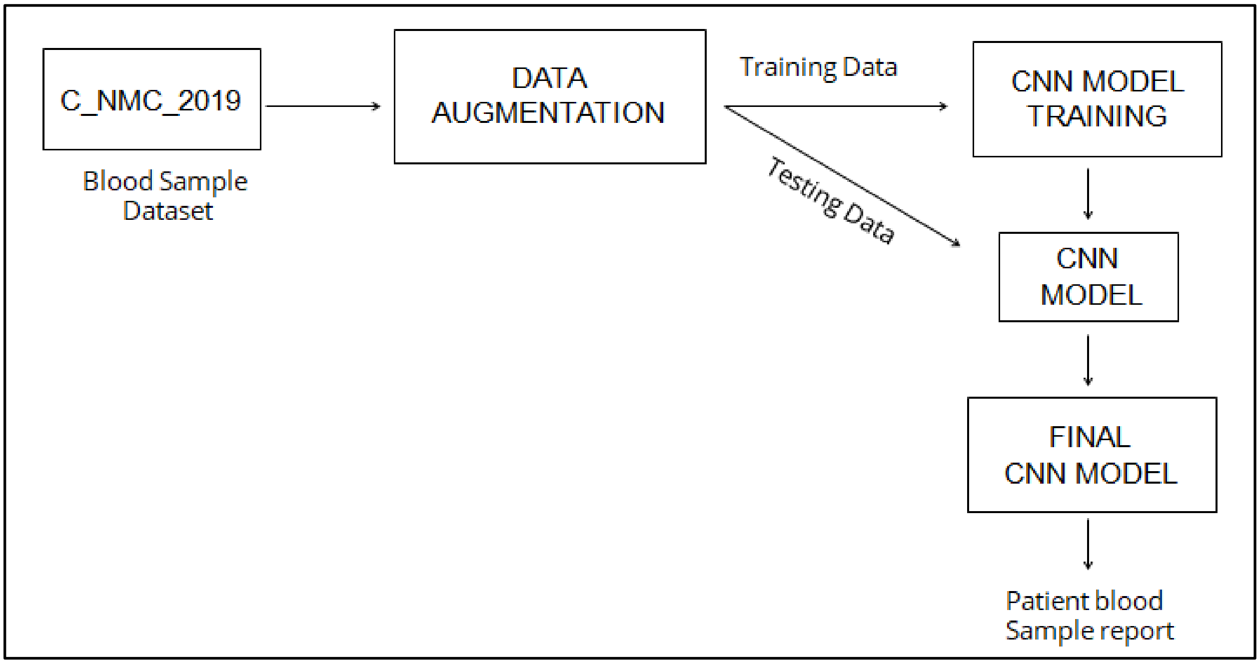

Figure 2: Flow Diagram of the Proposed Methodology

B. Dataset Used

The dataset that was used for the research by the authors was C-NMC-2019, which consists of 118 studies with a number of 49 patients with Hem. Table 1 provides detailed information about the dataset used. In the proposed model, the training database is categorized into 2 subcategories, namely the train and the confirmation sets, the Test set, and the Database image size resolutions of 128 × 128 pixels.

Table 1: C-NMC-2019 database

		Subjects	**Images**

Doi: https://doi.org/10.54216/FPA.090106
Received: April 10, 2022 Accepted: August 23, 2022

States	**Dataset Categories**		**ALL (Cancerous)**	**Hem (Normal)**	**ALL (Cancerous)**	**Hem (Normal)**
1st	**Training Set**	**Training**	32	19	5822	2703
		Validation	15	7	1450	686
2nd	**Test set**	-	13	15	1219	648
3rd	**Final Test set**	-	9	8	1716	825
Total			69	49	10252	4862

Table 1 concludes that the database is inconsistent, and the training images of the cancer cells are approximately 2.15 times more than normal cell images, which makes the separator a little biased in ALL categories. Such bias due to data inequality is processed to reduce in the proposed framework by the application of the next two strategies.

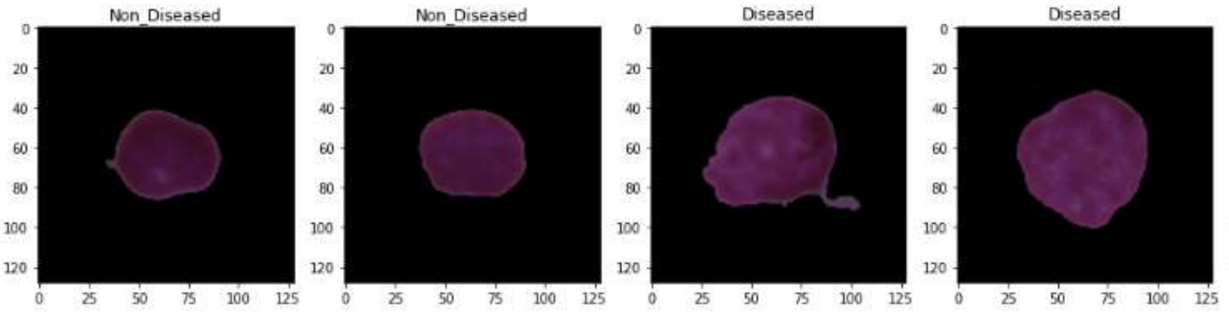

Figure 3: Sample images of the utilized C-NMC-2019 dataset, showing that there are unnecessary black regions around the region of interest.

C. Experimental Setup

The experiment for the proposed work is executed in Python using the Keras package with Tensorflow as the deep learning framework. We ran this code on an online platform named kaggle which offers a no-setup, customizable, Jupyter Notebooks environment. Kaggle provides a GPU of approx 16GB and RAM of around 14GB with disk space of around 80GB.

D. Pre-processing

The key strategies for pre-processing are briefly described, for a better ALL prediction system. Almost all images in the database contain a major interest area, a position with a dark background So, we cut out picture 10 in the middle as a size of 128 × 128 pixels to reduce the size of the input data, which makes learning segments faster by providing an interesting region. Class inequality is as common in the medical imaging field as it is in practice pictures with annotations are very complex and difficult to access. Inequality of such a category can be slightly defeated using different algorithms. This random over-sampling process is used, which includes duplicating samples of the junior class randomly and in addition to training samples measuring unequal data sets. For the proposed model, the Hem category sampled was 5822 images & was trained during training, from a total no. of 11644 images. Various ways are used to add data, such as zoom, rotation, and horizontal and vertical browsing are used for the training process to improve the performance of the model.

Doi: https://doi.org/10.54216/FPA.090106
Received: April 10, 2022 Accepted: August 23, 2022

4. Results and Analysis

In the proposed work, VGG-16 and CNN are used for the detection & classification of leukemia which gives an overall accuracy of 76.8% & 82.3% respectively. Table 2. Describes the comparisons of different algorithms in respect of accuracy and val_ accuracy and Fig 4. Shows the bar chart representation of the results.

Table 2: Comparison of different methods

Method Used	**Accuracy (%)**	**Val_accuracy(%)**
VGG-16	76.8%	72.7%
CNN	82.3%	76.8%

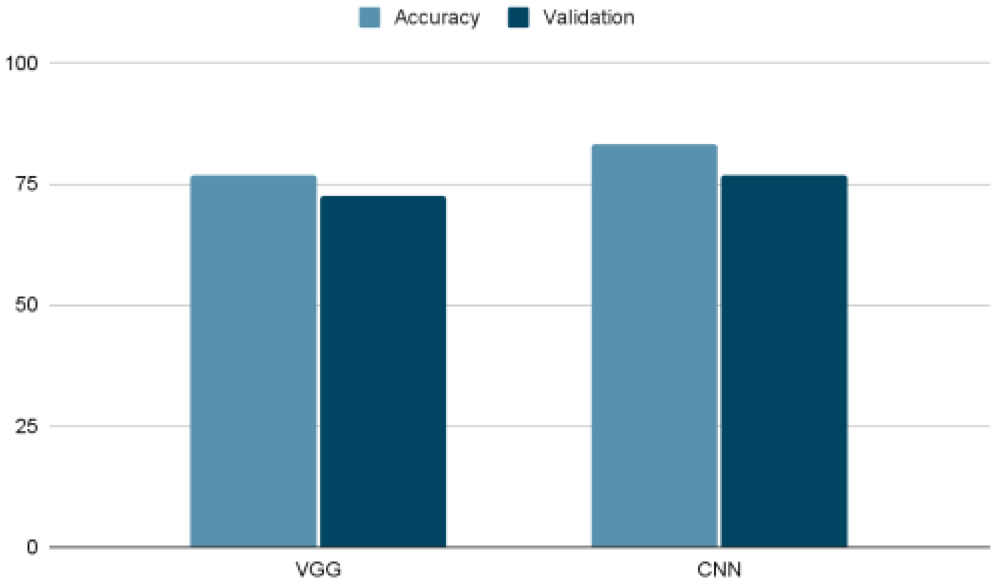

Figure 4: Bar chart displaying accuracy and validation

5. Conclusions

Deep learning algorithms are achieving popularity & this paper is an attempt to analyze the two Deep learning Algorithms used for the classification VGG-16 and CNN. The literature on these leukemia classification and prediction algorithms has been taken for research. Additional work in this study will lead to the identification of cancer patients, through the effective use of appropriate Deep learning algorithms.

References

[1] Maria, I. J., Devi, T., & Ravi, D. (2020). Machine learning algorithms for the diagnosis of leukemia. *Int J Sci Technol Res*, *9*(1).

[2] Vaghela, H., Modi, H., Pandya, M., & Potdar, M. B. (2016). A novel approach to detect chronic leukemia using shape-based feature extraction and identification with digital image processing. *International Journal of Applied Information Systems (IJAIS)*, *11*(5), 9-16.

Doi: https://doi.org/10.54216/FPA.090106
Received: April 10, 2022 Accepted: August 23, 2022

[3] Janaki, R. (2020). Detection of leukemia in microscopic white blood cell images using Gaussian feature convolutional visual recognition algorithm. *J Critic Rev*, *7*(3), 173-180.

[4] Bibi, N., Sikandar, M., Ud Din, I., Almogren, A., & Ali, S. (2020). IoMT-based automated detection and classification of leukemia using deep learning. *Journal of healthcare engineering*, *2020*.

[5] Sahlol, A. T., Kollmannsberger, P., & Ewees, A. A. (2020). Efficient classification of white blood cell leukemia with improved swarm optimization of deep features. *Scientific Reports*, *10*(1), 1-11.

[6] Salah, H. T., Muhsen, I. N., Salama, M. E., Owaidah, T., & Hashmi, S. K. (2019). Machine learning applications in the diagnosis of leukemia: Current trends and future directions. *International journal of laboratory hematology*, *41*(6), 717-725.

[7] Kumar, D., Jain, N., Khurana, A., Mittal, S., Satapathy, S. C., Senkerik, R., & Hemanth, J. D. (2020). Automatic detection of white blood cancer from bone marrow microscopic images using convolutional neural networks. *IEEE Access*, *8*, 142521-142531.

[8] Muthumayil, K., Manikandan, S., Srinivasan, S., Escorcia-Gutierrez, J., Gamarra, M., & Mansour, R. F. (2021). Diagnosis of leukemia disease based on enhanced virtual neural network. *Computers, Materials & Continua*.

[9] Mohammed Shahnawaz, K., & Kumar Saini, M. (2019). Detection Of Leukemia Using Digital Image Processing And Artificial Intelligence. *International Journal of VLSI Design and Technology*, *1*(2), 13-21.

[10] Raje, C., & Rangole, J. (2014, April). Detection of Leukemia in microscopic images using image processing. In *2014 International Conference on Communication and Signal Processing* (pp. 255-259). IEEE.

[11] Ghaderzadeh, M., Asadi, F., Hosseini, A., Bashash, D., Abolghasemi, H., & Roshanpour, A. (2021). Machine learning in detection and classification of leukemia using smear blood images: a systematic review. *Scientific Programming*, *2021*.

[12] Mondal, C., Hasan, M., Jawad, M., Dutta, A., Islam, M., Awal, M., & Ahmad, M. (2021). Acute Lymphoblastic Leukemia detection from microscopic images using weighted ensembles of convolutional neural networks. *arXiv preprint arXiv:2105.03995*.

[13] Ahmed, N., Yigit, A., Isik, Z., & Alpkoçak, A. (2019). Identification of leukemia subtypes from microscopic images using convolutional neural networks. *Diagnostics*, *9*(3), 104.

[14] Shafique, S., & Tehsin, S. (2018). Computer-aided diagnosis of acute lymphoblastic leukemia. *Computational and mathematical methods in medicine*, *2018*.

[15] Paswan, S., & Rathore, Y. K. (2017). Detection and classification of blood cancer from microscopic cell images using SVM KNN and NN classifiers. *Int. J. Adv. Res. Ideas Innov. Technol*, *3*, 315-324.

[16] Billah, M. E., & Javed, F. (2022). Bayesian Convolutional Neural Network-based Models for Diagnosis of Blood Cancer. *Applied Artificial Intelligence*, 1-22.

[17] Shafique, S., & Tehsin, S. (2018). Acute lymphoblastic leukemia detection and classification of its subtypes using pre-trained deep convolutional neural networks. *Technology in cancer research & treatment*, *17*, 1533033818802789.

[18] Madanan, M., Venugopal, A., & Velayudhan, N. C. (2020, July). Designing an artificial intelligence model using machine learning algorithms and applying it to hematology for the detection and classification of various stages of blood cancer. In *Int. Conf. Innovate. Tech. Adv. Dis. Manag*.

[19] Rupapara, V., Rustam, F., Aljedaani, W., Shahzad, H. F., Lee, E., & Ashraf, I. (2022). Blood cancer prediction using leukemia microarray gene data and hybrid logistic vector trees model. *Scientific reports*, *12*(1), 1-15.

[20] Saba, T. (2020). Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *Journal of Infection and Public Health*, *13*(9), 1274-1289.

[21] Kassani, S. H., Kassani, P. H., Wesolowski, M. J., Schneider, K. A., & Deters, R. (2019, October). A hybrid deep learning architecture for leukemic B-lymphoblast classification. In *2019 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 271-276). IEEE.

[22] Bukhari, M., Yasmin, S., Sammad, S., El-Latif, A., & Ahmed, A. (2022). A Deep Learning Framework for Leukemia Cancer Detection in Microscopic Blood Samples Using Squeeze and Excitation Learning. *Mathematical Problems in Engineering*, *2022*.

[23] Khan, S., Sajjad, M., Hussain, T., Ullah, A., & Imran, A. S. (2020). A Review on Traditional Machine Learning and Deep Learning Models for WBCs Classification in Blood Smear Images. *IEEE Access*, *9*, 10657-10673.

Doi: https://doi.org/10.54216/FPA.090106

Received: April 10, 2022 Accepted: August 23, 2022

[24] Gupta, A., & Sharma, P. (2021). A Review of Machine Learning Techniques Being Used For Blood Cancer Detection. *Annals of the Romanian Society for Cell Biology*, 7796-7811

Doi: https://doi.org/10.54216/FPA.090106
Received: April 10, 2022 Accepted: August 23, 2022