Detection of Leukemia Using Convolutional Neural Networks and Support Vector Machine

Amal Bentaher[1], Saleh Albatati[2]

[1]Assistant Professor of Computer Sciences, College of Computers and Information Technology
[2]Assistant Professor of Cell Biology, Histology, and Genetics, College of Medicine
[1,2]Hadhramout University, Mukalla, Yemen

Abstract - *Leukemia is a life-threatening type of cancer, in which early diagnosis is necessary for proper treatment that could help in curing the disease or prolonging patients' survival. Computer Aided Diagnosis (CAD) systems and Machine learning (ML) models are very useful tools for automating the detection of leukemia from blood samples' images, avoiding the limitations of manual screening. In this paper, we propose a method based on pretrained Convolutional Neural Networks (CNN) with Support Vector Machine (SVM) for the detection of Acute Myeloid Leukemia (AML) and classification of blood images into normal and leukemia. In parallel with SVM model, three other models were applied in the classification process for the performance evaluation, K-Nearest Neighbors (KNN), ensemble, and Decision Tree (DT). SVM reached an accuracy of 97.44% which is higher than the other three models (94.87% for ensemble, 87.18% for KNN and 76.92% for DT). These results suggest that the proposed method could be used for the detection of leukemia in blood samples' images.*

Key Words: Leukemia, Diagnosis, Convolutional Neural Networks, Machine Learning, Transfer Learning

1. INTRODUCTION

Leukemia is a life-threatening disease that affects both children and adults. It is a type of cancer that originates in bone marrow and results in the production of immature white blood cells. The rapid growth of this cancer inhibits the production of normal blood cells which is highly required in different body organs for proper immune response. In contrast, the abnormal cells flow to important organs, such as the liver and the brain, causing critical damage. Leukemia is classified, depending on its speed of progression and the phenotype of the cells, into acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML) and chronic myeloid leukemia (CML)[1]–[4]

Artificial intelligence (AI) tools are being deeply integrated in medical practice nowadays, with many AI-based software have been approved for medical use.[5] The digitalization of microscopic images of pathological samples, which paved the way for the use of AI tools, is widely expanded because it eases the process of saving or translocating the sample for further or later observation.[6]

Computer Aided Diagnosis (CAD) systems are developed to help pathologists to overcome the limitations of manual screening of pathological samples, including the relying on the technician's experience and skills, and the long-time function that may affect the sample microscopy.[7]·[8] These system are being widely used in different laboratory devices. Peripheral blood (PB) analyzers and automated cell counters, for example, use the digital image of the PB to recognize and discriminate between blood composing cells[9], count the number of each type of cells, and distinguish between normal and abnormal cells.[10] These system are also used in the classification of chest X-ray images into normal or having a lung disease.[11]

The simplicity of using automated systems in the classification of medical images and its efficacy in the diagnosis process helps in the early detection of many diseases such as glaucoma, skin cancer, and leukemia which is a must for the success of the treatment.[12]

Diagnosis of leukemia, and despite the advancement in the microscopic examination of blood samples and the devices designed for this process, is still suffering delay and should be replaced by a faster and more reliable methods. For that, CAD systems automate the process using image processing machine learning techniques.[1], [7] Many approaches and models were proposed for the using CAD and machine learning (ML) for the detection, classification, and diagnosis of different types of leukemia.[13]–[16]

Convolutional Neural Networks (CNN) are highly accurate models widely used to overcome many problems involving image classification. [7], [17] In this work we approach a method depending on transfer learning CNN with Support Vector Machine (SVM) to improve the accuracy of leukemia image classification.

2. RESEARCH METHODOLOGY

The data set used in this study were taken from the Kaggle Dataset. It consists of 195 images (96 normal and 99 abnormal) and divided into 80% training data and 20% test data.

2.1 Proposed Workflow

The proposed method includes three main steps (Figure 1):

2.1.1 Step 1: Processing

The method proposed in this paper used three operations when training the deep learning model. The following operations are:

A. Convert to RGB: In this operation convert all Leukemia images to RGB color model.
B. Resize all to 227 × 227: All images have different pixel size because they take from different device, so in this operation change the dimension of all images to fixed 227 × 227 pixel.
C. Augmentation

2.1.2 Step 2: Pre-trained deep learning model and direct feature extraction.

Our model used transfer learning Convolutional Neural Network (AlexNet) to be pretrained. The architecture of CNN consists of three types of layers: convolutional, pooling and fully connected layer (Figure 2). The initial (convolutional) layer only learns low-level features. Increasing the layers will increase the features and the network will learn more specific training task patterns. In our work we used the pooling layer to reduce the number of features of the resulting data and to overcome the overfitting problem. Convolutional layers calculate the output by performing the activation function (a rectified linear unit, ReLu).

2.1.3 Step 3: Classification with machine learning models

To classify the features, we chose different machine learning models. First, the method using support vector machine (SVM). This classifier seeks to separate two classes of a Leukemia images using a separation line or hyperplane to classify as normal or abnormal cases. Then data sets are also checked on other classifiers; K-Nearest Neighbors (KNN), ensemble, and Decision Tree (DT) for the performance evaluation.

2.2 Implementation and Experiment

MATLAB 2017b was used to implement our model and several experiments were conducted to evaluate their performance in terms of precision, sensitivity, accuracy, specificity and F1-Score, which are defined below

Precision= TP/(TP+FP), (1)

Sensitivity= TP/(TP+FN), (2)

Specificity= TN/(TN+FP) (3)

Accuracy=(TP+TN)/(TP+TN+FP+FN), (4)

F1-Score = (2× Sensitivity × Precision) / (Sensitivity + Precision) (5)

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

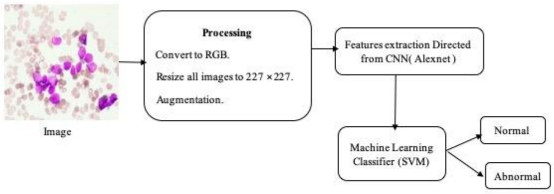

Figure 1: Architecture of Proposed Method.

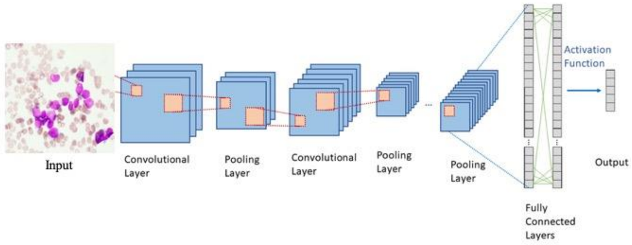

Figure 2: A simplified illustration of the CNN architecture.

3. RESULTS AND DISCUSSION

In this study we proposed a method for leukemia image classification depends on pretrained CNN for feature extraction with four machine learning models (SVM, Ensemble, KNN, DT) for the classification. We used direct feature extraction by CNN (alexNet) to reduce the time of CNN training.

In the four models we used for the classification of images (Table 1 and Figure 4), SVM model showed the highest accuracy (97.44%) followed by Ensemble using model (94.87%). The KNN and DT based model showed much lower accuracy with 87.18% and 76.92%, respectively. These results suggest that SVM is the preferred classification model to be used with direct feature extraction method of CNN.

Table 1: Performance matrices for different ML models

Models	Performance Matrices				
	Sensitivity	Specificity	Precision	F1-Score	Accuracy
SVM	95.00%	100%	100%	97.44%	97.44%
Ensemble	95.00%	94.74%	95.00%	95.00%	94.87%
KNN	75.00%	100%	100%	85.71%	87.18%
DT	75.00%	78.95%	78.95%	76.92%	76.92%

In this work, the use of pretrained CNN and SVM showed the accuracy of 97.44% which is higher than what achieved by

Patel and Mishra[18], Paswan and Rathore[19], and Shafique *et al.*[20], as they all used SVM with different artificial neural networks and reached accuracy of 93.57%, 83.33%, and 93.70%, respectively.

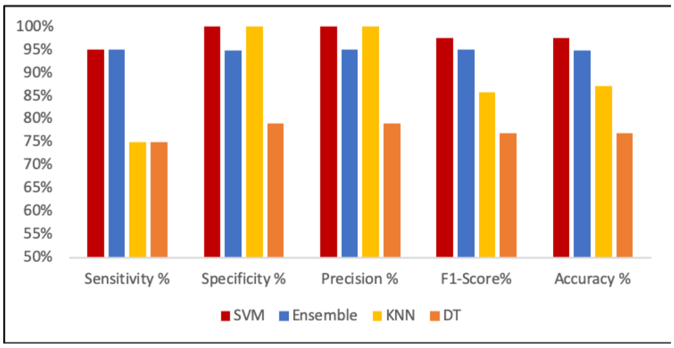

Figure 3: Performance matrices for different ML models

4. CONCLUSION

In this work, we present a simple model for the automation of the classification of blood cell images into normal or leukemia. This model takes the advantage of the high accuracy given by convolutional neural network and in the same time reduces the time of CNN training by direct extraction of features. Our model showed high accuracy compared to other artificial neural network based models. This model could be useful in the detection of leukemia from blood samples' images.

For future work, greater amount of data is needed to be used. And we recommend that sample images should be constantly published.

REFERENCES

[1] M. Madhukar, S. Agaian, and A. T. Chronopoulos, "Deterministic model for Acute Myelogenous Leukemia classification," in 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, Korea (South), Oct. 2012, pp. 433–438. doi: 10.1109/ICSMC.2012.6377762.

[2] M. M. Olsen, L. J. Zitella, and Oncology Nursing Society, Eds., Hematologic malignancies in adults. Pittsburgh, Pa: Oncology Nursing Society, 2013.

[3] S. Shafique and S. Tehsin, "Acute Lymphoblastic Leukemia Detection and Classification of Its Subtypes Using Pretrained Deep Convolutional Neural Networks," Technol. Cancer Res. Treat., vol. 17, p. 153303381880278, Jan. 2018, doi: 10.1177/1533033818802789.

[4] A. A. Ferrando and C. López-Otín, "Clonal evolution in leukemia," Nat. Med., vol. 23, no. 10, pp. 1135–1145, Oct. 2017, doi: 10.1038/nm.4410.

[5] H. T. Salah, I. N. Muhsen, M. E. Salama, T. Owaidah, and S. K. Hashmi, "Machine learning applications in the diagnosis of leukemia: Current trends and future directions," Int. J. Lab. Hematol., vol. 41, no. 6, pp. 717–725, Dec. 2019, doi: 10.1111/ijlh.13089.

[6] M. S. Lasya, A. Mahalakshmi, and A. Suma, "Detection and Classification of Leukemia using Convolutional Neural Network and Updation in IoT," Int. Res. J. Eng. Technol., vol. 07, no. 03, pp. 2395–2399, Mar. 2020.

[7] A. Genovese, M. S. Hosseini, V. Piuri, K. N. Plataniotis, and F. Scotti, "Histopathological Transfer Learning for Acute Lymphoblastic Leukemia Detection," in 2021 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Hong Kong, China, Jun. 2021, pp. 1–6. doi: 10.1109/CIVEMSA52099.2021.9493677.

[8] A. Rehman, N. Abbas, T. Saba, S. I. ur Rahman, Z. Mehmood, and H. Kolivand, "Classification of acute lymphoblastic leukemia using deep learning," Microsc. Res. Tech., vol. 81, no. 11, pp. 1310–1317, Nov. 2018, doi: 10.1002/jemt.23139.

[9] L. Boldú, A. Merino, S. Alférez, A. Molina, A. Acevedo, and J. Rodellar, "Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis," J. Clin. Pathol., vol. 72, no. 11, pp. 755–761, Nov. 2019, doi: 10.1136/jclinpath-2019-205949.

[10] C. Di Ruberto, A. Loddo, and G. Puglisi, "Blob Detection and Deep Learning for Leukemic Blood Image Analysis," Appl. Sci., vol. 10, no. 3, p. 1176, Feb. 2020, doi: 10.3390/app10031176.

[11] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "DeTrac: Transfer Learning of Class Decomposed Medical Images in Convolutional Neural Networks," IEEE Access, vol. 8, pp. 74901–74913, 2020, doi: 10.1109/ACCESS.2020.2989273.

[12] L. H. S. Vogado, R. D. M. S. Veras, A. R. Andrade, F. H. D. de Araujo, R. R. V. Silva, and K. R. T. Aires, "Diagnosing Leukemia in Blood Smear Images Using an Ensemble of Classifiers and Pre-Trained Convolutional Neural Networks," in 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Niteroi, Oct. 2017, pp. 367–373. doi: 10.1109/SIBGRAPI.2017.55.

[13] A. Setiawan, A. Harjoko, T. Ratnaningsih, E. Suryani, Wiharto, and S. Palgunadi, "Classification of cell types in Acute Myeloid Leukemia (AML) of M4, M5 and M7 subtypes with support vector machine classifier," in 2018 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Mar. 2018, pp. 45–49. doi: 10.1109/ICOIACT.2018.8350822.

[14] J. Laosai and K. Chamnongthai, "Acute leukemia classification by using SVM and K-Means clustering," in 2014 International Electrical Engineering Congress (iEECON),

Chonburi, Thailand, Mar. 2014, pp. 1–4. doi: 10.1109/iEECON.2014.6925840.

[15] L. H. S. Vogado, R. M. S. Veras, Flavio. H. D. Araujo, R. R. V. Silva, and K. R. T. Aires, "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification," Eng. Appl. Artif. Intell., vol. 72, pp. 415–422, Jun. 2018, doi: 10.1016/j.engappai.2018.04.024.

[16] D. Goutam and S. Sailaja, "Classification of acute myelogenous leukemia in blood microscopic images using supervised classifier," in 2015 IEEE International Conference on Engineering and Technology (ICETECH), Coimbatore, India, Mar. 2015, pp. 1–5. doi: 10.1109/ICETECH.2015.7275021.

[17] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, "Land Use Classification in Remote Sensing Images by Convolutional Neural Networks," ArXiv150800092 Cs, Aug. 2015, Accessed: Nov. 28, 2021. [Online]. Available: http://arxiv.org/abs/1508.00092

[18] N. Patel and A. Mishra, "Automated Leukaemia Detection Using Microscopic Images," Procedia Comput. Sci., vol. 58, pp. 635–642, 2015, doi: 10.1016/j.procs.2015.08.082.

[19] S. Paswan and Y. Rathore, "Detection and Classification of Blood Cancer from Microscopic Cell Images Using SVM KNN and NN Classifier," Int. J. Adv. Res. Ideas Innov. Technol., vol. 3, no. 6, pp. 315–324, 2017.

[20] S. Shafique, S. Tehsin, S. Anas, and F. Masud, "Computer-assisted Acute Lymphoblastic Leukemia detection and diagnosis," in 2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), Islamabad, Pakistan, Mar. 2019, pp. 184–189. doi: 10.1109/C-CODE.2019.8680972.