Hindawi

Research Article

Automated Detection Model in Classification of B-Lymphoblast Cells from Normal B-Lymphoid Precursors in Blood Smear Microscopic Images Based on the Majority Voting Technique

Mustafa Ghaderzadeh [iD],[1] **Azamossadat Hosseini** [iD],[1] **Farkhondeh Asadi** [iD],[1]
Hassan Abolghasemi [iD],[2] **Davood Bashash** [iD],[3] **and Arash Roshanpoor** [iD][4]

[1]*Department of Health Information Technology and Management, School of Allied Medical Sciences,*
 Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]*Pediatric Congenital Hematologic Disorders Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran*
[3]*Department of Hematology and Blood Banking, School of Allied Medical Sciences,*
 Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4]*Department of Computer Science, Sama Technical and Vocational Training College, Tehran Branch (Tehran),*
 Islamic Azad University (IAU), Tehran, Iran

Correspondence should be addressed to Azamossadat Hosseini; souhosseini@sbmu.ac.ir and Farkhondeh Asadi;
asadifar@sbmu.ac.ir

Introduction. Acute lymphoblastic leukemia (ALL) is the most common type of leukemia, a deadly white blood cell disease that impacts the human bone marrow. ALL detection in its early stages has always been riddled with complexity and difficulty. Peripheral blood smear (PBS) examination, a common method applied at the outset of ALL diagnosis, is a time-consuming and tedious process that largely depends on the specialist's experience. *Materials and Methods.* Herein, a fast, efficient, and comprehensive model based on deep learning (DL) was proposed by implementing eight well-known convolutional neural network (CNN) models for feature extraction on all images and classification of B-ALL lymphoblast and normal cells. After evaluating their performance, four best-performing CNN models were selected to compose an ensemble classifier by combining each classifier's pretrained model capabilities. *Results.* Due to the close similarity of the nuclei of cancerous and normal cells, CNN models alone had low sensitivity and poor performance in diagnosing these two classes. The proposed model based on the majority voting technique was adopted to combine the CNN models. The resulting model achieved a sensitivity of 99.4, specificity of 96.7, AUC of 98.3, and accuracy of 98.5. *Conclusion.* In classifying cancerous blood cells from normal cells, the proposed method can achieve high accuracy without the operator's intervention in cell feature determination. It can thus be recommended as an extraordinary tool for the analysis of blood samples in digital laboratory equipment to assist laboratory specialists.

1. Introduction

Leukemia is one of the most common blood cancers caused by an abnormal increase in the production of immature white blood cells in the bone marrow. In 2018, this disease affected 174,000 people in the United States alone. About 6,000 leukemia cases are diagnosed annually, of which acute lymphoblastic leukemia (ALL) is the second most common type in adults and the most common type of malignancy in children, accounting for about one-third of all pediatric cancers. Based on the most recent World Health Organization (WHO) classification, the purely leukemic presentation, B-lineage ALL (85%), is the most prevalent type of lymphoid cancers. ALL leads to excessive production of cells known as lymphoblasts that are not evolved in mature B and T lymphocytes; gradually displaces normal cells in the bone

marrow; and may spread to vital organs, for example, the liver, lymph nodes, spleen, and central nervous system (CNS). B-acute lymphoblastic leukemia (B-ALL) is a common type of neoplastic disorder with high mortality rates due to immature lymphocyte B-cell proliferation. The identification of the signs and symptoms of childhood cancer is highly challenging because it is not the first diagnosis made for nonspecific complaints, and this causes potential uncertainty in the diagnosis [1–3].

B-ALL can be diagnosed via different techniques, but PBS images play an especially important role in B-ALL detection. Preliminary examination of laboratory blood samples for leukemia and ALL diagnosis is performed manually using a light microscope. On the other hand, primary prevention measures are not effective in averting the development of B-ALL in children, and secondary prevention, i.e., early diagnosis, is essential. In the specific case of ALL, early diagnosis and treatment increase the chances of cure. To diagnose B-ALL, it is usual for the hematologist to examine blood slides; still, due to laboratory specialists' fatigue or other factors that affect them, an exact diagnosis of B-All is prone to errors. The manual examination of blood samples is also susceptible to error due to an unsuitable laboratory environment and contamination of laboratory or ocular microscope slides.

In the last two decades, several studies have adopted machine learning (ML) methods and computer-aided diagnostic methods to analyze laboratory images and overcome the consequences of a late leukemia diagnosis. These studies analyzed leukocyte nuclei in blood smear samples to diagnose and differentiate B-ALL from normal leukocytes. Recently, numerous computer-based methods have been employed to improve the efficiency of medical imaging techniques. One such method is the application of ML algorithms which has achieved remarkable success in medical imaging. Among different types of ML methods, deep learning (DL) has attained a high precision in machine vision tasks in leukemia. Convolutional neural networks (CNNs) as a main DL algorithm have great potentials in feature extraction and image data analysis. In line with the current trend in medical image analysis, these capabilities motivated research into their application and adaptation for blood component classification, especially in B-ALL detection [4–7].

2. Materials and Methods

The researchers reviewed studies that used similar data to provide classification methods. As these methods were limited by the use of a single state-of-the-art model, we decided to adopt a multitrained network-based approach to create an efficient model.

2.1. Dataset. The C-NMC dataset [8–10] comprised 12528 lymphocyte nucleus images, of which 8491 belonged to B-ALL lymphoblast and 4037 to normal B-lymphoid cases. The dataset cell nuclei were segmented from the microscopic images in the real world because they contained some staining noise and illumination errors, although an expert largely corrected these errors via an in-house method of stain color normalization. An expert oncologist marked the ground truth of the dataset images. Figure 1 illustrates samples of B-ALL and healthy cell nuclei. The data are available at https://wiki.cancerimagingarchive.net/display/Public/C_NMC_2019+Dataset%3A+ALL+Challenge+dataset+of+ISBI+2019.

2.2. Data Preparation and Preprocessing. Data standardization and normalization as the first steps of preprocessing for maintaining image integrity play a key role in image analysis and classification [11, 12]. To this end, the pixel-level global mean and standard deviation (STD) were first calculated for all the images. Then, the data were normalized using equation (1) where $\overline{x}$ denotes the global mean X of the image set, σ is the STD, and $\varepsilon = 1e - 10$ indicates the differential value to prevent the denominator from turning zero.

$$X_i = \frac{X_i - \overline{x}}{\sigma + \varepsilon}. \tag{1}$$

The CNN method relies on large volumes of data to improve its efficiency and prevent model overfitting [13, 14]; thus, after normalization, to standardize the image for achieving a uniform ratio for the deep neural network's input, each image's pixel value was mapped to [0, 255] and then converted to the [0, 1] interval. Herein, we were dealing with the nucleus of white blood cells; as such, the hidden features of the WBC core included chromatin density, nucleus open chromatic, etc. After image normalization and standardization, their core was enlarged by cutting their edges so that image processing algorithms could analyze the characteristics of different classes more easily by assessing the lymphocyte nucleus. Figure 2 depicts the two operations of cutting the edge and enlarging the core. Data augmentation was performed for the training dataset by 16 techniques for each image.

All the images in these collections were shuffled so that, during the training process, the network would not see only specific categories of data, and each batch of images would contain different labels belonging to B-ALL and non-B-ALL categories. The input image size was changed to $300 \times 300 \times 3$, but this method can be applied to images of any size.

2.3. Classification Algorithms. In the last two decades, numerous machine learning algorithms have been adopted for classification, each succeeding in specific areas and on specific datasets. For example, in text classification, decision tree algorithms, rule-based methods, and perceptron-based methods have demonstrated extraordinary abilities. However, as feature extraction is highly sensitive in image classification, especially for medical images, there is a need for methods that avoid manual selection based on mathematical methods. Thus, the present study utilizes deep learning algorithms [15].

In medical image analysis and classification, numerous deep CNN-based structures are widely used that
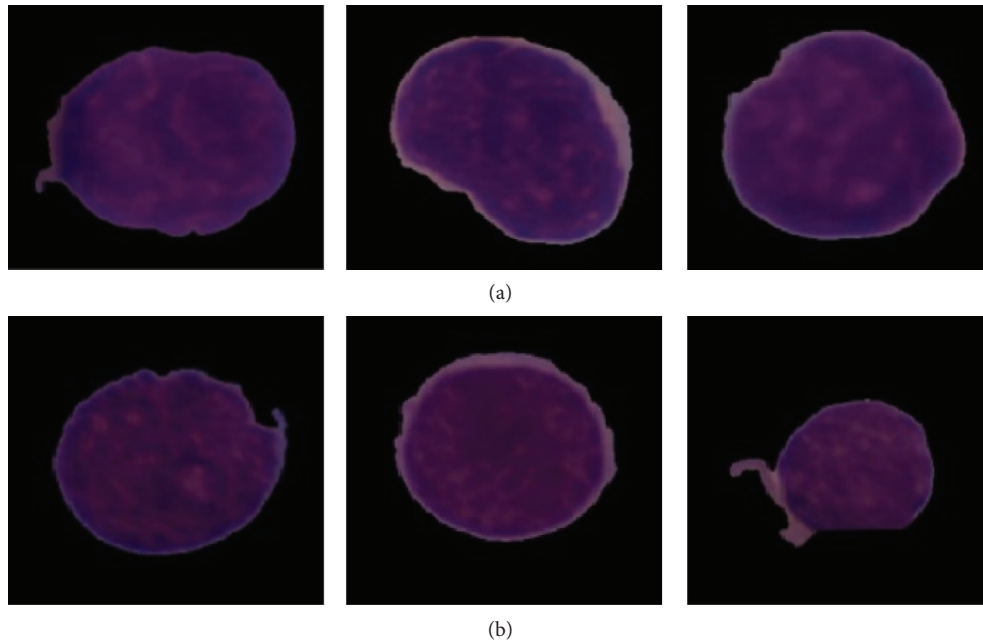
(a)

(b)

FIGURE 1: (a) B-ALL lymphoblast nuclei. (b) Normal cell lymphocyte nuclei.
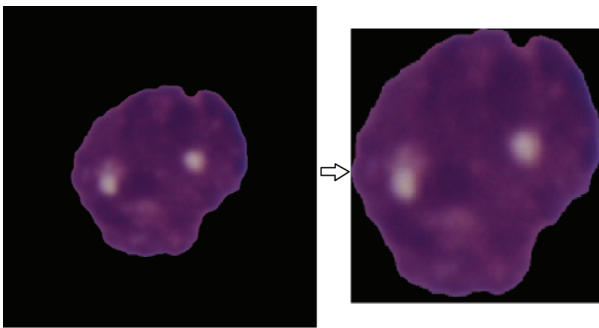

FIGURE 2: Image edge cutting to enlarge the morphological features.

benefit from the powerful feature engineering and representation ability of CNN. Of these methods, the pretrained CNN structure has demonstrated state-of-the-art performance in cells and organ segmentation problems [16–18]. CNN is a multilayer network composed of overlapping convolutional layers (for feature extraction) and downsampling layers (for feature processing). Figure 3 illustrates the structure of a typical CNN. As a perceptron-based mode, a CNN automatically extracts features from images and, thus, has become a hot topic of research [5, 19].

Based on the transfer learning technique, the pretrained models trained with large image collections have yielded extraordinary results in image classification problems. Many studies have utilized these models since they outperform many other models thanks to their image feature extraction. Among the publicly available DL pretraining models, Alexnet, ResNet [20] (ResNet50 and ResNet101), Inception-V3 [21], Inception-ResNet-V2 [22], SqueezeNet [23], and MobileNet-V2 were selected and compared here owing to their higher accuracy than other networks with similar prediction times. These well-known CNN models were pretrained with the ImageNet database; based on their structure, depth, and structural width, each model has unique features in image feature engineering.

2.4. Ensemble Learning Technique. Ensemble methods are algorithms that employ an ensemble of classifiers to predict data labels. The weighted majority voting algorithm, introduced by Littlestone and Warmuth in 1994, relies on the final decision result by taking the weighted majority votes of other algorithms. Littlestone and Warmuth proved that the ensemble methods are robust algorithms with respect to error and can significantly improve the learning system's generalization ability [24–26].

Ensemble algorithms are the most popular research directions in machine learning and computer vision. They aim to combine the predictions of multiple learning algorithms to achieve a classification model with superior predictive performance. Compared to basic learning algorithms, the generalizability of classifier sets has greatly improved. In addition, group learning methods can promote the predictive performance of weak learning algorithms, which is slightly better than random guessing of strong learning algorithms, thereby making very promising predictions [15, 27]. The algorithm forms its prediction by comparing the total weights made for each class and predicts the larger total. In this way, the classification result is voted to obtain the final classification result. The voting method is divided into absolute and relative majority voting methods. In the former, more than half of individual learners output the same classification result which is the final classification of the integrated learning output. Herein, the absolute majority voting was selected as the final method to obtain a comprehensive model and
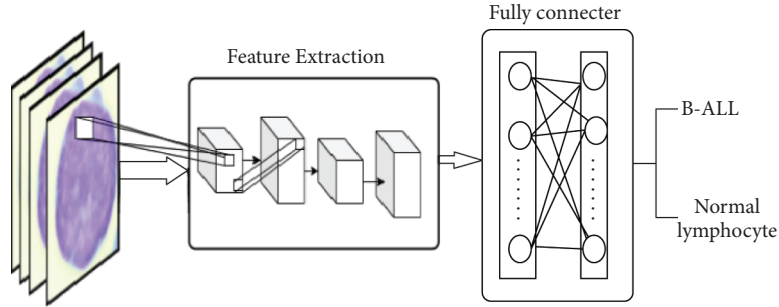
FIGURE 3: An ordinary convolutional neural network structure.

maintain the positive features of the four pretrained networks in dataset feature extraction.

2.5. Performance Evaluation. Four performance metrics were calculated to assess the performance of the CNNs models. We used traditional measures to evaluate the performance of the proposed model based on the confusion matrix The model's performance was determined using a confusion matrix [28, 29]. Generally, sensitivity is the classifier's ability to correctly classify all the cases with the disease (true positive); in other words, sensitivity is defined as the ratio of B-ALL cases accurately detected by the model to all the actual B-ALL lymphoblasts. Specificity is the classifier's ability to correctly identify cases without the disease (true negative rate); in other words, specificity refers to the ratio of normal lymphocytes accurately detected by the model to all the actual non-B-ALL (normal) cases. Moreover, accuracy is defined as the rate of all the B-All and lymphocyte cases correctly classified. The formulas for the metrics are

$$\text{sensitivity} = \text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}, \quad (2)$$

$$\text{specificity} = \frac{\text{true negative}}{\text{true negative} + \text{false positive}}, \quad (3)$$

$$\text{accuracy} = \frac{\text{true negative} + \text{true positive}}{\text{total cases}}. \quad (4)$$

3. Results

After reviewing many studies that used the concept of transitional learning to classify medical images, eight pretrained models were selected. After being customized based on the size of the input images, these models were trained with 80% of the dataset. The selected pretrained models were evaluated by adjusting network parameters with a test dataset. Table 1 lists the results based on the three main evaluation metrics.

After testing these eight models with training datasets and several rounds of parameter tuning, four models achieved the greatest performance in terms of accuracy and calculation time (DenseNet121, Inception V3, Inception-ResNet-v2, and Xception) and were thus selected.

3.1. Ensemble CNN Pretrained Models for Classification of B-ALL from B-Lymphoblast. By examining the performance of the models, we aimed to improve their results. We adopted the majority voting technique as the final decision to improve the model based on four pretrained models. The proposed model scheme first calculated the total number of votes received by each base classifier; then, the majority of votes was calculated by the classification of both output classes. Algorithm 1 presents this model in detail. Let $L = $ {DenseNet121, Inception V3, Inception-ResNet-v2, and Xception} be the set of pretrained models. The four models are fine-tuned with the images from training datasets ($X; Y$), where X is the set of N images, each of size 300×300, and Y contains the image labels. Y is a collection of two classes, including lymphoblast B-ALL and normal B-ALL. A batch size of $n = 256$ was used for implementing the selected models.

In this method, the number of base classifiers should always be odd. In the case of an equal vote, the mode function is applied [30]. The proposed B-ALL disease detection and classification model, which is an ensemble framework of four models, is displayed in Figure 4.

The four selected networks are in parallel and combine the output module with the ensemble technique to improve the classification confidence and accuracy. Figure 5 illustrates the confusion matrix to evaluate the test set for two classes. To quantitatively assess the performance of the proposed method, the sensitivity, specificity, accuracy, and F1 score evaluation criteria were determined based on model performance by using a confusion matrix (Table 2). Evidently, the proposed ensemble model has demonstrated a promising performance that is superior to the previous models. Its success with such a small dataset is attributed to the use of class weights in the training process.

4. Discussion and Conclusion

The analysis of PBS images plays a vital role in the diagnosis of various types of leukemia, anemia, and malaria. Unusual alterations in the color, shape, and size of blood cells indicate an abnormal condition. The results of this PBS assessment, which are often performed manually, depending on the technician's skill and experience. Besides, it is time-consuming and yields poor results [31–33].

With the publication of the ISBI 2019 C-NMC Challenge and its dataset, different methods for classifying images of

TABLE 1: Evaluation metrics of the pretrained models.

Model	Sensitivity (%)	Specificity (%)	Accuracy (%)	GPU time (second)
DenseNet 121	98.5	70.9	85.3	4116
Inception-V3	99	72.4	89.77	8130
AlexNet	88	69	78.6	4966
ResNet-50	79	66	74	5111
ResNet-101	81	68	79	10115
Inception-ResNet-v2	99.92	68.83	89.1	24119
SqueezeNet	89	68.8	81	5211
Xception	99.84	85.65	94.91	9232

Data: lymphoblast images (X, Y), X is the number of dataset images, $Y = \{y \in (normal\ B - ALL, lymphoblast\ B - ALL)\}$
Result: the proposed model based on the majority voting technique classifies the class of lymphocyte images $x \in X$
Applied preprocessing
Step 1: resizing images to dimensions 300×300
Step 2: dataset image normalization, where each image is mapped to [0, 255] and then converted to [0, 1] for $x \in X$.
Step 3: dataset edge detection and tuning brightness and contrast for $x \in X$
Step 4: data augmentation for every $x \in X$ (rotation, flipping, cropping, and color transformation)
Implementing a set of pretrained Model M={DenseNet121, Inception V3, Inception-Reset-v2, ResNet-50, ResNet101-Xception, SqueezeNet, Alexnet}
Models K={DenseNet121, Inception V3, Inception-Reset-v2, ResNet-50, ResNet101-Xception, SqueezeNet, Alexnet} with the problem scale
for each $k \in K$, $\alpha = 0.0017$, epochs = 700
  for $i = 1$ to epochs
    for each batch **do**
      for calculating the output and updating the hyperparameter
      if error is not improving, change hyperparameter
      using dropout, validation, and then feed output of feature extraction block to Softmax
      end
    end
  end
end
for each $x \in xtest$, do
  majority-voting technique's performance for all models
end

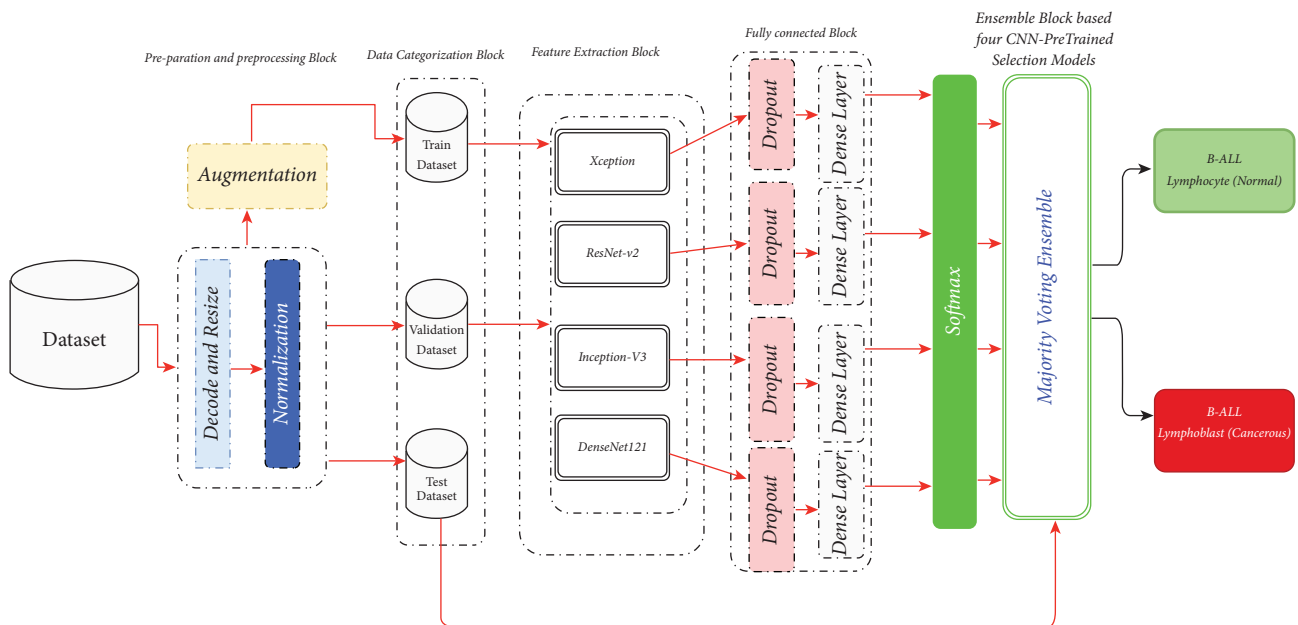ALGORITHM 1: The proposed ensemble algorithm.

FIGURE 4: Proposed model for the classification of B-ALL lymphoblast from B-ALL lymphocytes.
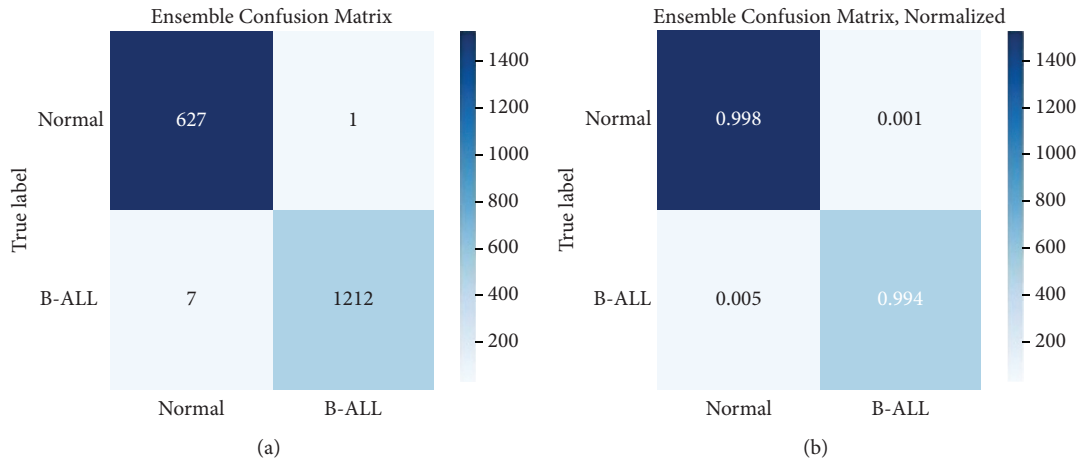
Figure 5: (a). Confusion matrix and (b) normalized confusion matrix of model performance.

Table 2: Evaluation of the proposed model for the test set.

Metrics	Value (%)
Sensitivity = positive predictive value = true positive/positive calls	99.4
Specificity = negative predictive value = true negative/negative calls	96.7
F1 score = (true positive/true positive) + 0.5(false positive + false 99.8% negative)	98.4

Table 3: Top entries of the C-NMC 2019 Challenge [38].

Researcher's Name	Weighted F1-score
Yongsheng Pan et al.	0.91
Ekansh Verma et al.	0.894
Jonas Prellberg et al.	0.889
Fenrui Xiao et al.	0.885
Tian Shi et al.	0.879
Ying Liu et al.	0.876
Salman Shah et al.	0.866
Yifan Ding et al.	0.855
Xinpeng Xie et al.	0.848
Proposed method	**0.984**

this database have been presented by researchers. Almost all of these studies have used deep learning methods and CNN algorithms. In one of these studies, which used VGG-16 with a Tripletloss function concept, they did not get good results [34]. In another study that used ReseNet-18 with an additional regression and advanced data augmentation techniques were applied to solve the large problem of tiny morphological differences between data from two data classes. The result of this study also showed that the F-score was equal to 0.8284 [34]. A custom model based on a fusion of CNN, LSTM was presented which used spectral features of cells by using discrete cosine transform in conjunction with an RNN to extract B-ALL image features. The method they used was an ensemble of convolutional and recurrent neural networks that used the AlexNet and DenseNet pretrained networks [35]. One of the most important techniques that have been highly considered by researchers to solve this challenge is the use of a set of ensemble techniques, given the nature of these techniques that employ common attributes of algorithms. Studies using the ensemble technique to classify B-ALL images from normal precursor blasts have yielded better results. One study assembled SENet and PNASNet-5 including ResNet, VGG, DenseNet, Inception V3, DenseNet, and IncptionReseNetV3 as three pretrained networks employed in an ensemble model [36], and also in another research, ResNeXt50 and ResNeXt101 were assembled to classify images [37]. Investigation on research that deals with C-NMC images can be concluded that ensemble learning can significantly improve the generalization ability of the learning system,

thereby enhancing the performance of the available methods. Hence, the proposed method based on the ensemble majority voting technique presented a framework for the automated classification of leucocyte cell nuclei from microscopic images. This ensemble framework presents a novel combination of imaging methodologies based on the state-of-the-art for B-ALL detection and achieved high superior accuracy in the classification of the two classes. Based on Table 3, the majority voting method is efficient in extracting the characteristics of blood blast images. As such, it can be recommended for medical image feature extraction where the accurate extraction of features is vital.

The method proposed herein is presented for blast cell nuclei. If the research data encounter a colony of blood cell types in the slides, the proposed method will not be effective because, in classifying cancerous lymphoblasts from normal blasts, the blast nucleus must be segmented to extract its features like the denseness rate of nucleolus and chromatin level of the nucleus. Thus, it is strongly suggested that when the data include other blood components such as RBCs, monocytes, and neutrophils, automatic feature extraction methods (e.g., CNN) should not be used because they will extract the features of unrelated components other than blasts. Therefore, to use CNN algorithms in the detection and classification of ALL blasts, segmentation plays a vital role in the performance of the diagnostic method.

In the present method, pretrained network structures as state-of-the-art models without any changes in feature extraction and classification blocks layers and topological entities have been employed. The authors suggest that, in the future, by changing the number of classification block layers including Bach normalization and Dense in the pretrained networks, a higher-accuracy classifier be provided.

Data Availability

Data are available and public in https://wiki.cancerimaging archive.net/pages/viewpage.action?pageId=52758223.

Ethical Approval

This study was approved by the Iran National Committee for Ethics in Biomedical Research (IR.SBMU.RETECH. REC.1399.735).

Disclosure

This study has been published in the MedRxiv database at (https://www.medrxiv.org/content/10.1101/2021.07.10. 21260312v1) [35]. As this research was conducted in line with an international challenge, the authors are presenting it on MedRxiv database to register the idea. This study is part of a Ph.D. project conducted at Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] American Cancer Society, "Childhood leukemia statistics," 2018, http://www.lls.org/http%3A/llsorg.prod.acquia-sit.

[2] S. Chiaretti, G. Zini, and R. Bassan, "Diagnosis and sub-classification of acute lymphoblastic leukemia," *Mediterranean journal of hematology and infectious diseases*, vol. 6, no. 1, Article ID e2014073, 2014.

[3] K. E. Rodrigues and B. d. Camargo, "Diagnóstico precoce do câncer infantil: responsabilidade de todos," *Revista da Associação Médica Brasileira*, vol. 49, no. 1, pp. 29–34, 2003.

[4] M. Ghaderzadeh, M. Aria, and F. Asadi, "X-ray equipped with artificial intelligence: changing the COVID-19 diagnostic paradigm during the pandemic," in *BioMed Research International*, A. Fancellu, Ed., vol. 2021, Article ID 9942873, 16 pages, 2021.

[5] M. Ghaderzadeh, F. Asadi, A. Hosseini, D. Bashash, H. Abolghasemi, and A. Roshanpour, "Machine learning in detection and classification of leukemia using smear blood images: a systematic review," in *Scientific Programming*, P. Wang, Ed., vol. 2021, Article ID 9933481, 14 pages, 2021.

[6] S.1 Shafique, S. Shafique, and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks," *Technology in Cancer Research & Treatment*, vol. 17, pp. 1–7, 2018.

[7] A. Bodzas, P. Kodytek, and J. Zidek, "Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception," *Frontiers in Bioengineering and Biotechnology*, vol. 8, pp. 1–13, 2020.

[8] A. Gupta and R. Gupta, *Isbi 2019 C-Nmc challenge: Classification in Cancer Cell Imaging*, Springer, Berlin, Germany, 2020.

[9] R. Duggal, A. Gupta, R. Gupta, and P. Mallick, "SD-layer: stain deconvolutional layer for CNNs in medical microscopic imaging," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 435–443, Springer, Quebec, Canada, September 2017.

[10] R. Duggal, A. Gupta, R. Gupta, M. Wadhwa, and C. Ahuja, "Overlapping cell nuclei segmentation in microscopic images using deep belief networks," in *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*, pp. 1–8, Guwahati, Assam India, December 2016.

[11] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, Elsevier B.V., vol. 42, pp. 60–88, 2017.

[12] S. Patro and K. K. Sahu, "Normalization: a preprocessing stage," 2015. arXiv Prepr arXiv150306462.

[13] S. Frühwirth-Schnatter, "Data augmentation and dynamic linear models," *Journal of Time Series Analysis*, vol. 15, no. 2, pp. 183–202, 1994.

[14] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, 2019.

[15] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," *Journal of Information Science*, vol. 44, no. 1, pp. 28–47, 2018.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science*, Springer, in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, October 2015.

[17] L. Lu, Y. Zheng, G. Carneiro, and L. Yang, "Deep learning and convolutional neural networks for medical image computing," *Adv Comput Vis Pattern Recognit*, 2017.

[18] M. Ghaderzadeh, M. Aria, A. Hosseini, F. Asadi, D. Bashash, and H. Abolghasemi, "A fast and efficient CNN model for BALL diagnosis and its subtypes classification using peripheral blood smear images," *International Journal of Intelligent Systems*, vol. 2021, 2021.

[19] S. Chen, W. Liu, J. Qin, L. Chen, G. Bin, and Y. Zhou, "Research progress of computer-aided diagnosis in cancer based on deep learning and medical imaging," *Sheng wu yi xue gong cheng xue za zhi= J Biomed Eng Shengwu yixue gongchengxue zazhi.*, vol. 34, no. 2, pp. 314–319, 2017.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[21] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, "Rethinking the inception architecture for computer vision," 2016, https://arxiv.org/abs/1512.00567.

[22] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, February 2017.

[23] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," 2016, https://arxiv.org/abs/1602.07360.

[24] T. Zhou, H. Lu, Z. Yang, S. Qiu, B. Huo, and Y. Dong, "The ensemble deep learning model for novel COVID-19 on CT images," *Applied Soft Computing*, vol. 98, p. 106885, 2021.

[25] J. Z. Kolter and M. A. Maloof, "Dynamic weighted majority: an ensemble method for drifting concepts," *Journal of Machine Learning Research*, vol. 8, pp. 2755–2790, 2007.

[26] J. Z. Kolter and M. A. Maloof, "Using additive expert ensembles to cope with concept drift," in *Proceedings of the 22nd International Conference on Machine Learning*, pp. 449–456, Bonn Germany, August 2005.

[27] A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," *Expert Systems with Applications*, vol. 57, pp. 232–247, 2016.

[28] M. Ghaderzadeh, F. Asadi, R. Jafari, D. Bashash, H. Abolghasemi, and M. Aria, "Deep convolutional neural network-based computer-aided detection system for COVID-19 using multiple lung scans: design and implementation study," *Journal of Medical Internet Research*, vol. 23, no. 4, Article ID e27468, 2021.

[29] M. Ghaderzadeh, "Clinical decision support system for early detection of prostate cancer from benign hyperplasia of prostate," *Studies in Health Technology and Informatics*, vol. 192, p. 928, 2013.

[30] S. P. Kannojia and G. Jaiswal, "Ensemble of hybrid CNN-ELM model for image classification," in *Proceedings of the 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 538–541, IEEE, Noida, India, February 2018.

[31] A. K.-P. in N. Science, *IBCIS: Intelligent Blood Cell Identification System*, 2008.

[32] M. Maity, T. Mungle, D. Dhane, A. K. Maiti, and C. Chakraborty, "An ensemble rule learning approach for automated morphological classification of erythrocytes," *Journal of Medical Systems*, vol. 41, no. 4, p. 56, 2017.

[33] K. Prasad and G. K. Prabhu, "Image analysis tools for evaluation of microscopic views of immunohistochemically stained specimen in medical research-a review," *Journal of Medical Systems*, vol. 36, no. 4, pp. 2621–2631, 2012.

[34] A. Honnalgere and G. Nayak, "Classification of normal versus malignant cells in B-ALL white blood cancer microscopic images," in *ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging*, pp. 1–12, Springer, Berlin, Germany, 2019.

[35] S. Shah, W. Nawaz, B. Jalil, and H. A. Khan, "Classification of normal and leukemic blast cells in B-ALL cancer using a combination of convolutional and recurrent neural networks," in *ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging*, pp. 23–31, Springer, Berlin, Germany, 2019.

[36] T. Shi, L. Wu, C. Zhong, R. Wang, and W. Zheng, "Ensemble convolutional neural networks for cell classification in microscopic images," in *ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging*, pp. 43–51, Springer, Berlin, Germany, 2019.

[37] R. Kulhalli, C. Savadikar, and B. Garware, "Toward automated classification of b-acute lymphoblastic leukemia," in *ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging*, pp. 63–72, Springer, 2019.

[38] "CodaLab-competition," 2021, https://competitions.codalab.org/competitions/20395_learn_the_details-top-entries-of-the-challenge.