International Journal of Computational Intelligence Systems and Applications. Volume 13, Issue 1, (2026) 1-8
http://www.cayley-nielson.com/ijcisa/

Leukemia Medical Image Recognition Based on Deep Learning

Cui Dongyan
School of Artificial Intelligence, North China University of Science and Technology, Tangshan, Hebei, China
E-mail: *cdy_xxz@163.com*

Abstract

Acute Lymphoblastic Leukemia (ALL) is a life-threatening disease that brings immense grief to families and society. The treatment process for leukemia typically consumes substantial medical resources, posing a heavy economic burden on many households. Therefore, timely detection and treatment are crucial to addressing this issue. This study aims to fully explore the application potential of deep learning technology by designing and implementing an efficient and accurate leukemia medical image recognition system. The goal is to address problems such as strong subjectivity, low efficiency, and high misdiagnosis rates associated with manual observation. First, the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique was applied to meticulously optimize the C-NMC dataset, enhancing the visual quality of the images. Next, several mainstream image classification models were compared and discussed, ultimately selecting the ConvNext model for implementation. Then, an end-to-end full network training approach based on transfer learning was used to train the model, achieving a final recognition accuracy of 83.03%.

Keywords: Deep Learning; Acute Lymphoblastic Leukemia (ALL); Transfer Learning; Image Recognition.

1. Introduction

With the rapid advancement of deep learning technology, the application of artificial intelligence in the field of medical diagnosis continues to expand, offering innovative solutions to critical clinical challenges. Particularly in leukemia recognition, deep learning algorithms have demonstrated significant diagnostic efficacy, bringing breakthrough progress to medical image analysis and showing great application prospects [1–3].

In 2021, Chen Lieguang et al. developed a detection model for WT1 gene expression levels based on a deep belief network algorithm and systematically evaluated its application value in the clinical subtyping of acute myeloid leukemia (AML) [4].

In 2022, Raina Rohini et al. conducted a comprehensive survey of deep learning methods for acute leukemia classification and detection. Centered on four key research questions, they systematically summarized cutting-edge research achievements in automated detection and classification of acute leukemia, providing an important reference for subsequent studies in this field [5].

Also in 2022, Zhang et al. successfully constructed an innovative model named CMLcGAN based on cGAN. This model focuses on bone marrow biopsy samples and aims to accurately identify and extract clinically significant megakaryocytes (MK cells) by optimizing the segmentation algorithm, thereby providing efficient technical support for bone marrow cell population analysis [6].

In 2023, Neenavath Veeraiah et al. innovatively integrated the Mayfly optimization strategy into a

Generative Adversarial Network (MayGAN). They adopted the improved MayGAN model for classification and recognition, offering a new technical pathway for the field of leukemia diagnosis [7].

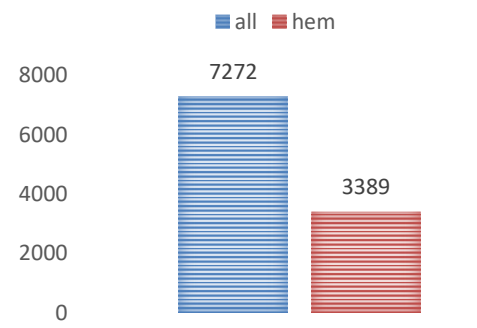
In 2024, Ashish Kumar et al., using an acute lymphoblastic leukemia dataset, performed a refined classification of normal cells and malignant leukemia cells, and systematically analyzed the significant differences in their complex morphological characteristics [8].

Against this backdrop, the core objective of this study is to propose a deep learning-based method for leukemia cell image diagnosis that is highly reliable, efficient, and accurate. This aims to address the long-standing challenge of leukemia image classification, thereby enhancing the model's classification and recognition capabilities and providing a more supportive technical solution for clinical auxiliary diagnosis.

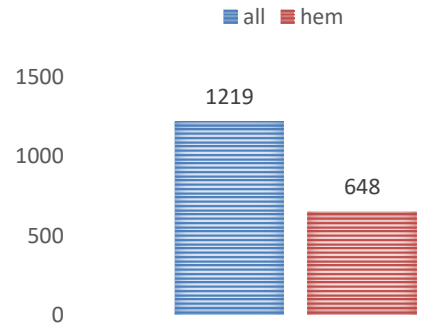
2. Data Preprocessing

2.1. Data Source

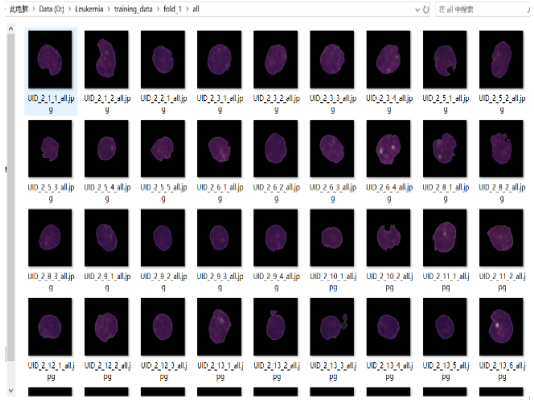
This study utilizes the C-NMC leukemia dataset (Leukemia classification), as shown in Figure 1, which is the most widely used public benchmark in the field of acute lymphoblastic leukemia. The dataset was collected from 118 clinical subjects (including 69 cancer patients and 49 healthy participants). A core characteristic of the dataset is that its training and test sets strictly adhere to the principle of cross-subject separation—meaning the data in these two sets originate from completely distinct individuals.



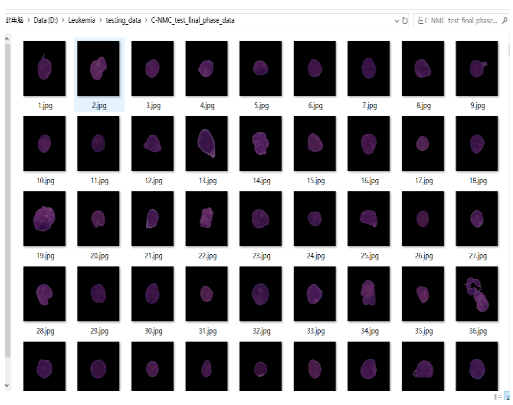
(a) Training Set Distribution



(b) Test Set Distribution



(c) Training Set Folder



(d) Test Set Folder

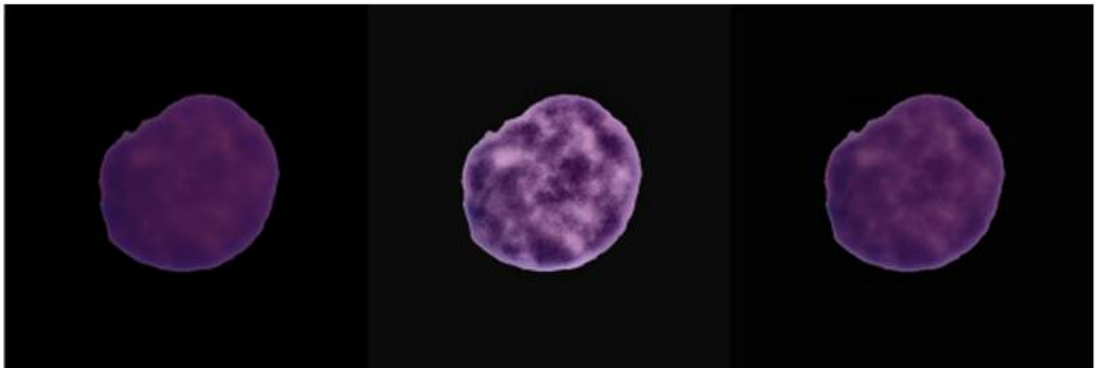
Fig.1 C-NMC Dataset Overview

The dataset comprises two main categories: ALL (cancerous) and HEM (normal). The training set contains a total of 10,661 single-cell microscopic images (7,272 from the ALL class and 3,389 from the

HEM class). The test set consists of 1,867 images (1,219 from the ALL class and 648 from the HEM class). All images are provided in a high resolution of 450×450 pixels.

2.2. Image Histogram Equalization

Given the visual similarity between ALL and HEM samples in the dataset, and considering that clinical practice typically relies on observing key biological indicators—such as cell morphology, chromatin distribution, nuclear count, and volume—to differentiate cell types, this study applies two algorithms for image processing: Histogram Equalization (HE) and Contrast Limited Adaptive Histogram Equalization (CLAHE). These methods enhance and emphasize detailed features such as texture, color, and shape in the images. Sample results of the histogram equalization processing are shown in Figure 2.



Original HE CLAHE

Fig.2 Results of Histogram Equalization Processing

As shown in Figure 2, CLAHE can significantly enhance the local contrast of images, making details more clearly visible—particularly those that might be overlooked under global histogram equalization. By limiting contrast, it effectively avoids excessive noise amplification and over-enhancement of local contrast, thereby improving the quality and stability of image enhancement. It also demonstrates strong adaptability to images with varying lighting conditions and contrast levels, achieving satisfactory enhancement across different scenarios. However, its limitation lies in its relatively high computational complexity, as it requires separate histogram calculation, contrast limiting, and grayscale mapping for each small image region.

2.3. Image Scaling

In research and applications related to image recognition and computer vision, processing large volumes of image data is essential. Cropping images serves to standardize their dimensions and resolution, or to remove irrelevant background information that may not contribute to model training. This reduces data volume and noise interference, thereby improving both training efficiency and model accuracy. Sample results of the image cropping process are shown in Figure 3.

As shown in Figure 3, through the cropping process, the key structures in the cell images become more distinct, while the surrounding uninformative black areas are reduced. Thereby enhancing the model's ability to extract distinctive features beneficial for classification and recognition.

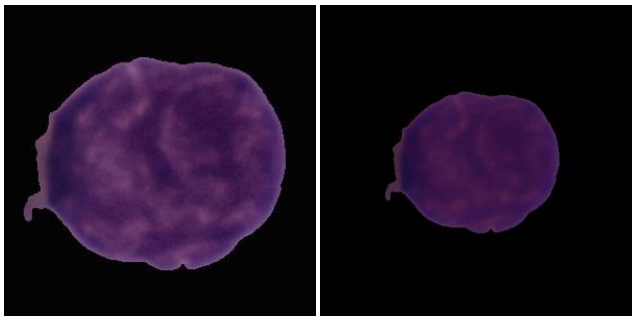


Fig.3 Image Cropping Processing Results

2.4. Image Normalization

The calculation formula is as follows:

$$x = \frac{x - \mu}{\sigma} \tag{1}$$

Where μ is the mean of the dataset and σ is the standard deviation of the dataset.

Furthermore, when performing image normalization, the use of ImageNet's mean and standard deviation for normalization resulted in color shifts. This may cause pixel values to exceed the original [0, 255] range, leading to abnormal color appearance when the images are saved. The initial normalization effect is shown in Figure 4.

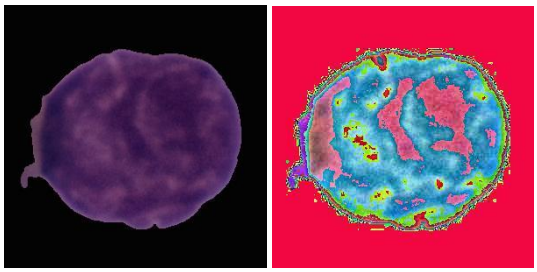


Fig.4. Initial Image Normalization Results

To address this issue, the code was modified to calculate the minimum and maximum pixel values of the normalized images, followed by linearly scaling the pixel values to the range [0, 1] before converting them back to PIL images. The final normalization results are shown in Figure 5.

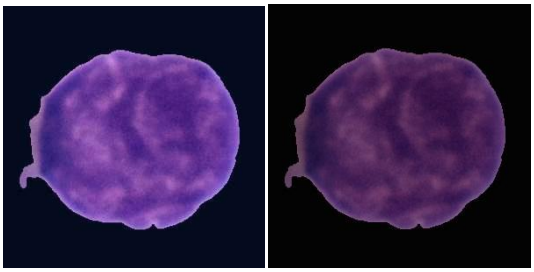


Fig.5 Final Image Normalization Results

3. Experimental Design

3.1. Experimental Environment

The hardware configuration for this study is as follows: an 11th Gen Intel(R) Core(TM) i5-1135G7 processor @ 2.40GHz (2.42 GHz), 16GB RAM, an Intel(R) Iris(R) Xe Graphics card with shared memory, 512GB storage, and a 64-bit Windows 11 operating system. The model was trained for a total of approximately 50 epochs.

3.2. Evaluation Metrics

This study employs four key metrics to evaluate the performance of the leukemia classification network: Accuracy, Precision, Recall, and F1-Score. The formulas for calculating these four metrics are as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+NP} \tag{4}$$

$$F1 = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{5}$$

If the dataset is limited, the trained model is prone to overfitting. Transfer learning was applied to three models: Inception V3, ResNet50, and ConvNext. They were evaluated based on Accuracy, Precision, Recall, and F1-Score, respectively. The specific evaluation results are shown in Table 1.

Tab.1 Evaluation of Each Model Using Transfer Learning

Model	Accuracy（%）	Precision（%）	Recall（%）	F1-Score（%）
Inception V3	75.30	75.52	75.83	75.11
ResNet50	76.65	76.22	75.90	76.02
ConvNext	80.91	80.13	80.57	79.80

As shown in Table 1, after applying transfer learning, the ConvNext model achieved an accuracy of 80.91%, precision of 80.13%, recall of 80.57%, and an F1-score of 79.80%. This performance advantage stems from ConvNext's more suitable architecture for small datasets, whereas the Inception model is deeper with more complex modules, and ResNet relies on traditional convolution operations. Therefore, the ConvNext model demonstrates better suitability for this task.

3.3. Image Scale Experiment

The original image size in the dataset is 450×450 pixels. Sample images are displayed in Figure 6.

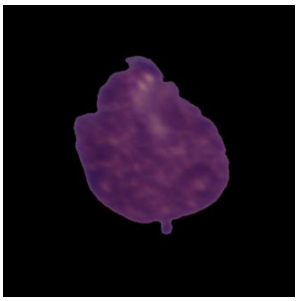
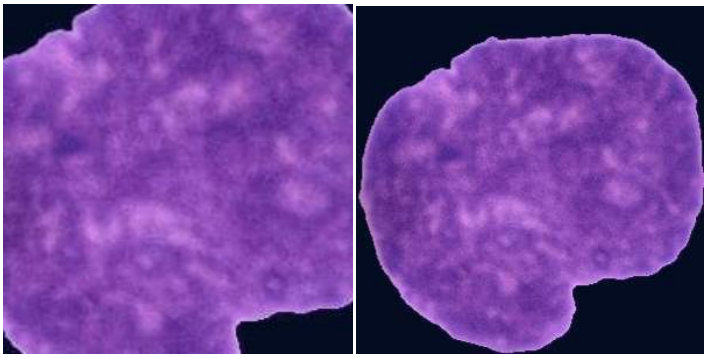


Fig.6 Typical Images from the Leukemia Dataset

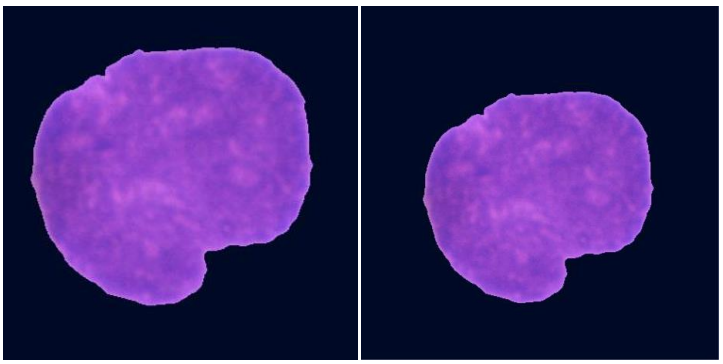
As shown in Figure 6, the original dataset images display cell morphology and characteristics in the central region, surrounded by black padded backgrounds. In the context of this medical image recognition study, the surrounding black backgrounds are non-informative regions. Therefore, image cropping is necessary to reduce data volume and noise interference, thereby improving model training efficiency and accuracy.

Tab.2 Comparative Results of Experiments with Different Image Scales

Size	Accuracy（%）	Precision（%）	Recall（%）	F1-Score（%）
224×224	80.91	80.13	80.57	79.80
299×299	83.03	82.52	82.50	83.01
320×320	81.76	81.22	81.38	81.17
384×384	78.11	78.70	79.15	78.43



(a) 224×224 (b) 299×299



(c) 320×320 (d) 384×384

Fig.7 Leukemia Cell Images at Different Scales

Given the superior performance of the ConvNext model in the previous transfer learning experiments compared to other architectures, and considering that the default input size for ConvNext is 224×224 pixels, initial cropping to this dimension was attempted. However, it was observed that some key cellular features were inadvertently cropped out. To address this, multiple image scales—224×224, 299×299, 320×320, and 384×384—were tested in this section to evaluate model performance across different input dimensions. The experimental results are presented in Table 2.

As shown in Table 2 and Figure 7, when the image size is 224×224, some edge features of the cells are cropped out. Conversely, with image sizes of 320×320 and 384×384, excessive uninformative black regions are retained. Experimental results indicate that the model achieves its highest accuracy (83.03%) when the image size is set to 299×299. Therefore, it can be concluded that appropriately increasing the input image size helps reduce the loss of cellular contour information, thereby enabling the model to extract more distinctive features beneficial for classification. However, since the model's network depth and kernel sizes are fixed, its receptive field remains constant. Excessively large input image sizes may hinder the model's ability to adapt to such dimensional variations.

3.4. Model Training

The study utilized a preprocessed leukemia cell image dataset, with all images resized to 299×299 pixels in RGB three-channel format to meet the input requirements of the ConvNext model. The model architecture is based on ConvNext, which was pre-trained on the large-scale ImageNet dataset, providing robust visual feature extraction capabilities. A transfer learning strategy was adopted: the lower-layer weights of ConvNext were frozen, and only the custom top classifier was trained. This approach leverages the general feature representations of the pre-trained model while reducing computational cost and training time. The specific model structure is as follows: deep features are first extracted by ConvNext, producing a feature map of dimensions 7×7×1024; a global average pooling layer then compresses the spatial dimensions into a 1024-dimensional feature vector; this is followed by a fully connected layer with 256 neurons and ReLU activation for nonlinear feature transformation; finally, a single-neuron output layer with Sigmoid activation produces binary classification probabilities, enabling the discrimination between normal cells and leukemia cells.

4. Conclusion

1) Based on experiments conducted with the Leukemia Classification dataset, this study investigated the performance of diverse data preprocessing methods and comprehensively evaluated the benefits and drawbacks of different image sizes on model training. A rigorous data standardization pipeline was implemented to enhance model convergence efficiency and reduce the time required for stabilization.

2) By evaluating and comparing widely used models such as Inception V3, ResNet50, and ConvNext, the ConvNext model - which demonstrated the best evaluation results—was selected and trained using transfer learning.

3) The model was trained for 50 epochs using the TensorFlow framework for batch data loading and the Adam optimizer. Upon completion, the trained model was exported in PTH format.

4) Experiments demonstrate that the proposed leukemia cell image classification model, leveraging a pre-trained architecture, can automatically classify cell images, providing an intuitive auxiliary tool for preliminary leukemia screening.

References

[1] Zhiwu S,Jie Z,Wanxiang L, et al.A Domain Adversarial Transfer Model with Inception and Attention Network for Rolling Bearing Fault Diagnosis Under Variable Operating Conditions. Journal of Vibration Engineering & Technologies, 2022, 12(1):1-17.

[2] Bose R S,Kumar S V.Efficient inception V2 based deep convolutional neural network for real-time hand action recognition. IET Image Processing, 2020, 14(4):688-696.

[3] Panpan Z,Zhaosheng M,Yingtao Z, et al. Improved Inception V3 method and its effect on radiologists' performance of tumor classification with automated breast ultrasound system. Gland surgery, 2021, 10(7):2232-2245.

[4] Lieguang C, Ying L, Renzhi P, et al. Deep learning in molecular biology marker recognition of patients with acute myeloid leukemia. The Journal of Supercomputing, 2021, 78(9): 11283-11297.

[5] Rohini R, Kumar N G, Chaahat, et al. A Systematic Review on Acute Leukemia Detection Using Deep Learning Techniques. Archives of Computational Methods in Engineering, 2022, 30(1): 251-270.

[6] Zelin Z, Xianqi H, Qi Y, et al. The diagnosis of chronic myeloid leukemia with deep adversarial learning. The American journal of pathology, 2022, 192(7): 1083-1091.

[7] Veeraiah N, Alotaibi Y, Subahi A. MayGAN: Mayfly Optimization with Generative Adversarial Network-Based Deep Learning Method to Classify Leukemia Form Blood Smear Images. Computer Systems Science and Engineering, 2023, 46(2): 2039-2058.

[8] Kumar A, Kumar N, Kuriakose J, et al. A deep transfer learning based approaches for the detection and classification of acute lymphocytic leukemia using microscopic images. Multimedia Tools and Applications, 2024:1-25.