Deep Learning Techniques in Acute Lymphoblastic Leukemia Classification: State-of-the-Art Review

- **sama M. A. Yousif, Department of Computer Science, College of Computer Science and Information Technology, Al Neelain University, Khartoum, Sudan. (Corresponding author).**

  osamamohamed2004@gmail.com
- **Eltyeb Elsamani Abdelgabar Elsamani, Department of Computer Science, College of Computer Science and Information Technology, Al Neelain University, Khartoum, Sudan.**

  tayebsamani@hotmail.com
- **Ibrahim Rizqallah Alzahrani, College of Computer Science and Engineering, University of Hafr Al Batin, Saudi Arabia**

  ialzahrani@uhb.edu.sa
- **Adil Mahmoud Mohamed Mahmoud, Deanship of E-Learning and Digital Transformation, University of Hafr Al Batin, Saudi Arabia.**

  aamohamed@uhb.edu.sa

Abstract

Diagnosing acute lymphoblastic leukemia (ALL) is a major challenge in medical practice, as traditional methods such as microscopy and molecular analysis lack high accuracy and are time-consuming and labor-intensive. In recent years, deep learning techniques, particularly convolutional neural networks (CNNs), have revolutionized blood image analysis and disease diagnosis. This study aims to conduct a systematic review of research published between 2015 and 2023 in specialized databases such as PubMed, IEEE Xplore, Scopus, and Web of Science, focusing on CNN-based models, hybrid models (CNN-LSTM/GRU), and transfer learning and clustering methods. Studies were ranked and compared according to common performance metrics such as precision, recall, F1 coefficient, and ROC-AUC. The review highlights that combining CNNs with GRUs represents a promising improvement over traditional models, combining superior spatial feature extraction capabilities and efficient sequential relationship processing, while reducing computational cost compared to LSTMs. The results demonstrate that this approach provides higher accuracy and reliability, especially when dealing with limited or imbalanced data, making it suitable for clinical applications and implementation in resource-constrained computing environments (Edge AI).

Keywords: Acute Lymphoblastic Leukemia, Convolutional Neural Networks, Transfer Learning, Recurrent Neural Networks, Classification.

المستخلص

يعد تشخيص اللوكيميا اللمفاوية الحادة (ALL) تحدياً كبيراً في الممارسة الطبية، حيث إن الطرق التقليدية مثل الفحص المجهري والتحاليل الجزيئية تفتقر إلى الدقة العالية وتستهلك وقتاً وجهداً كبيرين، في السنوات الأخيرة، أسهمت تقنيات التعلّم العميق، وبخاصة الشبكات العصبية الالتفافية (CNN)، في إحداث نقلة نوعية في تحليل صور الدم وتشخيص المرض. تهدف هذه الدراسة إلى إجراء مراجعة منهجية للأبحاث المنشورة بين عامي 2015 و2023 في قواعد بيانات متخصصة مثل PubMed وIEEE Xplore وScopus وWeb of Science، مع التركيز على النماذج المعتمدة على CNN، النماذج الهجينة (CNN–LSTM/GRU)، وأساليب التعلّم بالنقل والتجميع، تم تصنيف الدراسات ومقارنتها وفق معايير الأداء الشائعة مثل الدقة، الاسترجاع، المعامل التوافقي F1، ومساحة المنحنى ROC-AUC، وتبرز المراجعة أن دمج الـ CNN مع وحدات GRU يشكّل تحسيناً واعداً مقارنة بالنماذج التقليدية، لما يجمعه من قدرة عالية على استخراج السمات المكانية وكفاءة في معالجة العلاقات التسلسلية، مع تقليل الكلفة الحسابية مقارنة بـ LSTM، وتُظهر النتائج أن هذا التوجه يوفر دقة وموثوقية أعلى، خصوصاً عند التعامل مع بيانات محدودة أو غير متوازنة، مما يجعله مناسباً للتطبيقات السريرية وللتنفيذ في بيئات الحوسبة محدودة الموارد (Edge AI).

1. Introduction:

White blood cells in the bone marrow are the primary target of the aggressive blood malignancy known as acute lymphoblastic leukemia (ALL) [1]. It advances swiftly and, if not identified and treated right once, can cause major health issues. Enhancing survival rates and guaranteeing successful treatment depend on early identification. But conventional diagnostic techniques, which mostly depend on manually analyzing blood samples, are frequently laborious and prone to human error [2]. These restrictions make it extremely difficult to provide prompt and precise diagnoses, especially in settings with limited resources [3].

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), has demonstrated significant promise in recent years for improving the precision and effectiveness of medical diagnosis. These tools outperform traditional approaches in identifying irregularities and classifying illnesses by evaluating massive datasets, including medical pictures [7]. Convolutional Neural Networks (CNNs), in particular, have demonstrated the ability to identify intricate patterns in medical imaging, which makes them ideal for the detection and categorization of leukemia [9].

The absence of high-quality medical data and model interpretability are two issues that might hinder the practical use of AI-based diagnostic systems, notwithstanding their potential. Furthermore, the majority of current research focuses on categorizing ALL as "healthy" or "affected," leaving subtype categorization as an area that requires more investigation [5, 6]. With a focus on the application of CNNs, transfer learning, and other complex models like Recurrent Neural Networks (RNNs) and autoencoders, this study examines recent developments in deep learning methods for the diagnosis and classification of ALL. Along with highlighting possible future research avenues to solve these constraints, it also notes important problems, such as interpretability concerns and data shortages.

Section 1 introduces the topic, research problem, research objectives, research methodology, setting the stage for the subsequent discussion.

Section 2 examines the various phases of image classification in deep learning, while Section 3 focuses on datasets integral to this process. Section 4 offers a comprehensive discussion, followed by Section 5, which highlights the challenges encountered in this domain. Sections 6 and 7 sequentially outline future research directions and provide a concluding overview of the study.

1.2 Research problem:

Despite the notable advancements in traditional diagnostic techniques for acute lymphoblastic leukemia (ALL), such as manual microscopy and molecular techniques, these methods remain slow, prone to human errors, and impractical in resource-limited environments. In contrast, deep learning techniques have demonstrated their high capability in analyzing and accurately classifying medical images. However, current studies suffer from clear limitations that hinder their practical application; the majority of research lacks distinctive and large-scale medical data for training models, and most focus on binary classification between healthy individuals and those affected without delving into the classification of disease subtypes. Additionally, the limited clinical interpretability of these models still poses a barrier to their adoption in medical practice. These combined challenges highlight the need for a comprehensive review of the latest deep learning techniques in this field, and for proposing more efficient and practical hybrid frameworks for real-world applications, such as models that integrate the capabilities of Convolutional Neural Networks (CNN) with Gated Recurrent Units (GRU) to enhance accuracy and generalization and support deployment in resource-limited computing environments.

1.3 Research Objectives:

Three major goals are the focus of this study. With an emphasis on CNNs, transfer learning, and hybrid models, it first examines and evaluates current deep learning methods used for the identification and categorization of acute lymphoblastic leukemia (ALL). Second, it highlights the main issues and unmet research needs in the existing studies, such as the lack of

interpretability for clinical application, the limited subtype classification, and the scarcity of data. Third, it draws attention to CNN–GRU hybrid architectures, which have been documented in the literature as a promising advancement to improve generalization, accuracy, and viability in environments with limited resources.

1.4 Research Methodology:

- **Search & Selection:** conducted a systematic search throughout PubMed, IEEE Xplore, Scopus, and Web of Science from 2015 to 2023; included deep learning studies on ALL with imaging datasets and performance measures; omitted works written in languages other than English and confined to clinical settings.
- **Classification & Analysis:** According to accuracy, precision, recall, F1-score, ROC-AUC, and efficiency, studies were divided into three categories: CNN-based, hybrid CNN–LSTM/GRU, and ensemble/transfer learning techniques.
- **Justification for CNN–GRU:** While GRU captures temporal dependencies at a lower cost than LSTM, CNN retrieves spatial features. Hybrid models combine the two and, in reviewed works, demonstrated increased accuracy and resilience, frequently through augmentation and supervised training.

2. Image Classification in Deep Learning:

The basic steps in a deep learning-based approach for the detection and classification of leukemia typically involve:

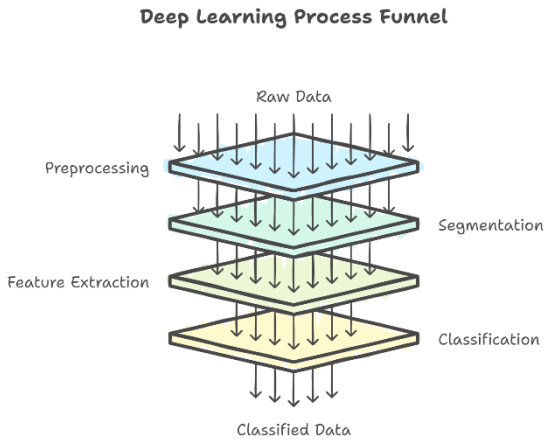

Figure 1: Steps In a Deep Learning-Based

2.1 Step One: Preprocessing

Pre-processing is a technique used to improve the clarity and readability of photographs before they are shown. Medical imaging relies heavily on image processing, which enhances picture quality. Unsatisfactory outcomes are possible if the photos are of low quality; preprocessing leukemia images for deep learning can be a crucial step in improving the performance. The primary goal of preprocessing is to make the photos more suitable for deep learning algorithms by cleaning, normalizing, and transforming the data [8]. Furthermore, the preprocessing technique can include:

a. Image cropping: Removing unnecessary background information from the images can reduce the complexity of the data and make it easier for the model to learn.

b. Image resizing: Resizing images to a consistent size can be important for some deep learning models. It ensures that the model does not have to deal with different image dimensions, which can cause trouble.

c. Data normalization: The deep learning model's performance can be enhanced by scaling pixel values to a predetermined range. In order to

prevent the model from becoming skewed toward particular pixel values, this step is crucial.

d. Data augmentation: Using random modifications like rotation, scaling, and flipping to create new pictures from the existing dataset can assist expand the dataset, lessen overfitting, and broaden the model's applicability.

e. Eliminating noise: Image quality may be enhanced and pictures can be more suited for deep learning by using noise reduction techniques as Gaussian blur, median blur, or morphological processes [4].

CNNs are the best option for a lot of classification tasks. However, the training data has a significant impact on CNN's performance. It may be difficult to collect enough clinical pictures due to patient privacy concerns. We employ a range of data augmentation techniques, including rotation, intensity changes, horizontal and vertical flips, contrast, and brightness adjustment, as recommended by other studies, to address the issue of having insufficient data, which causes the model to overfit [10,11]. As a result, effective system training improves the performance of segmentation or classification.

2.2 Step Two: Segmentation

One of the most important steps in examining blood samples for illness diagnosis is segmenting white blood cells (WBCs) in microscope pictures. Deep learning has recently advanced, making it a preferred approach for image segmentation.

Fully Convolutional Neural Networks and Transfer Learning are the two main categories of deep learning-based segmentation algorithms utilized in ALL detection and classification [49].

2.3 Step Three: Feature Extraction

In order to extract important characteristics that are crucial for ALL detection and classification, we go over a number of deep learning-based feature extraction and feature reduction techniques in this section. In their approach to ALL classification, Vogado et al. used PCA-based

feature reduction approaches after VGG-f-based feature extraction to extract important characteristics [14].

In different research, Vogado et al. extracted features using AlexNet, CaffNet, and VGG-f, then employed a gain ratio to choose more significant features [15].

Shahin et al. provide a brand-new deep convolutional neural network-based white blood cell (WBC) detection method. They employ two transfer learning approaches: fine-tuning of pre-existing deep networks and transfer learning based on deep activation characteristics [12].

Using the best training model among the three modified ResNet models they showed, Das et al. suggested a way to extract key features for ALL classification [9]. The authors of a different publication presented a method for ALL classification using a MobileNetV2-SVM framework, in which SVM is used for effective classification and MobileNetV2 is used for feature extraction [16].

2.4 Step Four: Classification

Deep learning techniques for medical image processing are becoming increasingly popular because of their proven performance efficiency. Convolutional neural networks (CNN), recurrent neural networks (RNN), and autoencoders are the three main categories of deep learning techniques commonly applied in medical image analysis [17].

2.4.1 Convolutional Neural Network (CNN)

The acronym CNN represents "Convolutional Neural Network." This kind of deep learning technique is frequently applied to the analysis of images and videos. The structure and operation of the visual cortex, which is in charge of processing visual input in the brain, served as the model for CNNs.

An input layer, one or more hidden layers (such as pooling, normalization, activation, or convolutional layers), and an output layer make up a CNN. The main component of CNNs is the convolutional layer, which applies a convolution operation to the input picture and assists in feature

extraction by identifying shapes, edges, and textures. To add non-linearity to the network, the activation layer applies a non-linear function to the convolutional layer's output. The purpose of the pooling layer is to improve computational efficiency and decrease the spatial size of the input image. Prior to applying the activation function, the data is normalized by the normalization layer. Using the characteristics that the network has learnt, the output layer is utilized to categorize the picture [18].

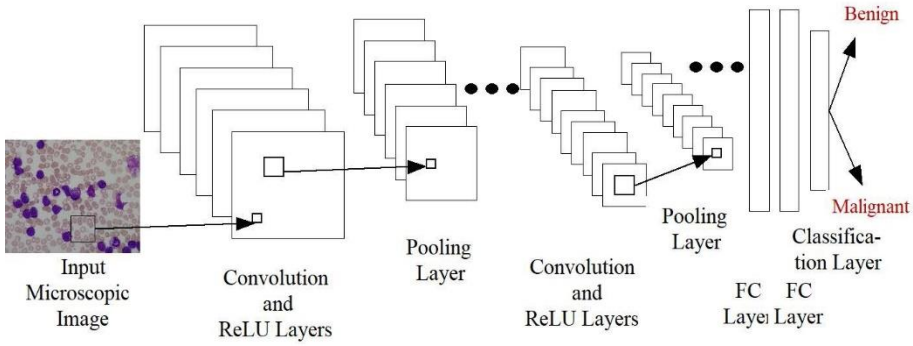

Figure 4: ALL- CNN diagram [49].

a) **Input Layer:** This layer is responsible for receiving the input image and converting it into a format that the network can process. It does not have any parameters and just reshapes the input image to the desired format.
b) **Convolutional Layer:** This layer retrieves information from the input picture by using a convolution operation. It applies a number of filters to the picture, each of which is in charge of identifying a certain aspect of the picture, such as edges, textures, or forms.
c) **To add non-linearity to the network**, the ReLU Layer applies a non-linear function (Rectified Linear Unit) to the convolutional layer's output.
d) **Pooling Layer:** This layer improves computing efficiency by reducing the input image's spatial size. Max pooling is a popular pooling technique that uses the highest value of a set of pixels in the picture.

e) **Fully Connected Layer:** This layer uses the network's learnt characteristics to categorize the picture. It is made up of one or more thick layers that are linked to every neuron in the layers before it.

f) The classification layer is a crucial part of CNN models. It predicts the input's class using the output probabilities of the SoftMax activation function. A number of CNN-based classification techniques for ALL have been put out [18].

g) **Output Layer:** This is where the network makes the final classification. The output is usually a probability distribution over the different classes [13].

Generally, classification layer, it is challenging to achieve optimal performance with conventional CNNs due to the need for large datasets and the limited size of publicly available medical datasets. This makes it hard to properly adjust the network's parameters and extract relevant features, leading to subpar performance.

2.4.1.1 Transfer Learning

It is a technique that expands on the idea of traditional CNN. While transfer learning uses pre-trained networks that have already been trained on a large dataset and refines them on a new, smaller dataset, conventional CNN involves training a neural network from scratch on a new dataset. In transfer learning, a CNN model that has already been trained on a large dataset is used as a starting point and then refined on a smaller dataset to learn the specific features of the new dataset. The pre-trained model can already identify low-level features like edges and textures. It can then be applied to a new dataset to extract features that can be used for classification; this technique is currently becoming more and more popular in the medical imaging industry [22].

Several popular transfer learning approaches are widely used in computer vision and image classification tasks. Some of the most used pre-trained models are:

a) One of the earliest deep convolutional neural networks, AlexNet significantly improved image classification tasks after being trained on a sizable picture dataset.
b) ReLU activations, which introduce nonlinearities, and a max pooling layer, which lessens overfitting, come after each convolution layer [15, 23, 24].
c) **CaffNet**: In 2014, Jia Deng and Wei Dong put out the concept. It is a C++ and CUDA-based neural network library that is open-source. Although the design of CaffNet is similar to that of AlexNet, it is quicker and more effective due to differences in the sequence of the pooling layer and normalization [15].
d) **VGGNet:** This model is known for its good performance in image classification tasks. It has several variations, such as VGG-f, which are popular in medical image analysis.

   Instead of a convolution layer, a block of filters is used. Employing 3 x 3 convolutions in succession increases nonlinearities and optimizes the receptive field [25].
e) **GoogLeNet:** This model, also known as Inception-v1, was the winner of the ImageNet 2014 competition. It employs an inception module with various filter sizes for dimensionality reduction. It has a complex architecture with several branches of convolutional layers, which allows it to extract features from different scales and locations [26].
f) **ResNet:** This model, which has 152 layers and a deeper design, is well-known for successfully training deep networks [27].
g) **MobileNet:** This model was designed to be lightweight and efficient for mobile and embedded vision applications as it demonstrates the effectiveness of depth-wise separable convolution in reducing computational complexity while maintaining accuracy [28].
h) **MobileNetV2:** This is an updated version of MobileNet with a more efficient architecture that uses depth-wise convolutions [29].

i) **Xception:** This model is like Inception models but uses depth-wise separable convolutions to reduce the computational cost [30].

2.4.1.2 Yolo Models Other CNN Advancements

Popular enhanced CNN models for object identification that are intended to be quicker and more effective are called YOLO, YOLOv2, YOLOv3, and YOLOv4 [37, 38, 39, 40]. Although YOLO uses a convolutional layer to anticipate the position of the bounding box, it has a localization error [37]. Anchor boxes, rather than the convolutional layer, are used by YOLOv2 to address this problem [38]. YOLOv3 optimally detects bounding boxes using a prediction method based on logistic regression [39]. Bochkovskiy et al.'s 2020 proposal, YOLOv4, uses a number of performance-enhancing strategies. Classification models have been put forth for effective illness diagnosis, especially cancer detection [40].

While Duggal et al. have described a CNN-based cancer detection system that incorporates a stain deconvolutional (SD) layer to transform microscopic images to Optical Density (OD) space [42], Gehlot et al. have proposed a two-module deep learning ALL classification framework [41].

2.4.2 Recurrent Neural Network (RNN)

Recurrent neural networks, or RNNs, are a deep learning method that uses the output of the current state to predict the output of the future state. This neural network's feedback mechanism aids in processing sequential or streaming input [22]. RNNs are divided into two categories: long short-term memory (LSTM) and bidirectional RNNs [43, 44].

2.4.3 Autoencoder

Reconstructing the input data from a compact representation is the aim of autoencoders, a kind of deep learning neural network used for unsupervised learning. Its two primary components are a decoder that returns the latent code to its original dimension and an encoder that converts the input data to a lower-dimensional representation (also known as a bottleneck or latent code) [22, 23].

3. Datasets of Acute Lymphoblastic Leukemia

The most popular free datasets for studying acute lymphoblastic leukemia (ALL) are ALLIDB1 and ALLIDB2. ALLIDB1 and ALLIDB2 have gained appeal among academics because of their vast amounts of high-quality data, which makes them appropriate for training and evaluating machine learning models across many academic domains. An overview of the most often used datasets is shown in Table 1.

Table 1 presents an overview of the most popular datasets

Dataset	Description
ALLIDB1 [45]	The 108 photos in this collection were gathered in September 2005. It has over 39,000 blood components, and skilled oncologists have designated the cells. A range of microscope magnifications, from 300 to 500, were used to take the pictures.
ALLIDB2 [45]	The purpose of this image collection was to assess how well categorization systems performed. A set of clipped regions of interest from the ALL-IDB1 dataset of normal and blast cells is called ALL-IDB2 version 1.0. With the exception of size, ALL-IDB2 photos have many of the same gray-level characteristics as ALL-IDB1 images.
Atlas [46]	An extensive resource for scholars looking into Acute Lymphoblastic Leukemia (ALL) is the Atlas dataset. It includes single nucleotide polymorphisms (SNPs), copy number variations (CNVs), and structural variants (SVs) for individuals with ALL. Gene expression profiles, which are part of each patient's transcriptome data, give a glimpse of the molecular condition of their leukemia cells at the time of diagnosis.
BCCD [47]	The Blood Cell Characterization and Discovery (BCCD) dataset is helpful for researchers interested in blood cells, such as Acute Lymphoblastic Leukemia (ALL). The BCCD dataset contains high-dimensional imaging and molecular data for many blood cells, including ALL cells. It provides an in-depth look at these cells' molecular and cellular changes. High-resolution images of individual cells are included in the imaging data, providing information on cell size, shape, and texture. Each patient's demographic information, treatment history, and disease outcome are also included in the BCCD dataset.

4. Discussion:

This section compares the performance of different deep learning techniques for detecting and classifying ALL using the results from various studies. Table 2 offers a comparison of the recent advancements in ALL detection and classification using deep learning methods.

Table 2 compares recent advances in deep learning based ALL detection and classification.

Reference	Year	Segmentation	Features		Classifier	Dataset	Accuracy
			Extraction	Reduction			
[7]	2017	marker-based watershed	GLCM	PPCA	RF	ALL-IDB1	96%
[32]	2017	...	...	...	hybrid transfer learning	ImageNet	88.50%
[14]	2017	VGG-F	…	PCA	SVM-MP-RF	ALL-IDB1	100%
[31]	2018	not required	not required	not required	AlexNet	ALL-IDB1 ALL-IDB2	96.06%
[6]	2018	marker-based watershed	GLRL	…	SVM	ALL-IDB1	96.97%
[15]	2018	not required	transfer learning	…	SVM	Many Dataset	99%
[20]	2019	AC-FCM	CH	LDP	SCA-CNN	ALL-IDB2	98.70%
[23]	2019	…	shape-color-texture	…	AlexNet-Autoencoder	Many Dataset	99.80%
[12]	2019	…	transfer learning	chi-squared	SVM-WBCsNet	Many Dataset	96.10%
[19]	2020	k-means	…	…	CNN	BCCD-ALL-IDB2	99.42-98.61%
[33]	2020	…	…	…	DNN	microarray gene data	98.21%
[24]	2020	semantic segmentation	…	…	AlexNet	LISC	98.87%

[21]	2020	…	…	…	Alert Net-RWD	Many Dataset	97.18
[48]	2020	…	…	…	YOLOv2	ALL-IDB1	98.72%
[11]	2021	…	transfer learning	…	…	ALL-IDB2	96.15%
[35]	2021	…	…	…	VGGA-DL	ALL-IDB2	96.80%
[34]	2021	…	transfer learning	…	ShufllNet	ALL-IDB1-ALL-IDB2	96.97%
[16]	2022	…	…	MobileNetV2	…	ALL-IDB1 ALL-IDB2-ASH	99.39%-98.21%

Table 3: Different performance measures are represented mathematically [49].

Performance Measures	**Math Representation**
Actual Negative Rate (TNR)	$\frac{\boldsymbol{TN}}{(\boldsymbol{TN}+\boldsymbol{FP})}$
Actual Positive Rate (TPR)	$\frac{\boldsymbol{TP}}{(\boldsymbol{TN}+\boldsymbol{FP})}$
Precision	$\frac{\boldsymbol{TP}}{(\boldsymbol{TP}+\boldsymbol{FP})}$
Accuracy	$\frac{(\boldsymbol{TP}+\boldsymbol{TN})}{(\boldsymbol{FN}+\boldsymbol{FP}+\boldsymbol{TN}+\boldsymbol{TP})}$
F1 Score	$\frac{\boldsymbol{2(precision*recall}}{(\boldsymbol{precision}+\boldsymbol{recall})}$
False Positive Rate (FPR)	**1-TNR**

Table 4 compares a set of deep learning methods used to detect and classify ALL in the ALL-IDB1 dataset.

Table 4: classification performance in ALL-IDB1.

Method	Classifier	Accuracy%	Sensitivity %	Specify %	Precision %	F1 Score
[34]	Xception	80.61	62	90	93.89	0.7342
	NasNet-Mobile	85.15	76	94.62	96.66	0.8429
	VGG19	90.3	95.33	86.76	86.11	0.9084
	ResNet50	95.15	98	91.97	92.78	0.9489
	ShuffleNet	96.97	98	95.63	96.11	0.968
[36]	AlexNet	93.64	98.64	89.55	89.35	0.9387
	VGG16	85.75	80.35	88.38	89.9	0.8417
	Google-Net	95.15	95.66	94.49	94.67	0.9507
	ResNet18	95.15	96.59	93.79	94.12	0.9517
	Mobile-Net v2	96.06	96.63	95.67	95.81	0.9615
	Hybrid	**99.39**	99.55	**99.33**	**99.74**	**0.9944**
[16]	VGG16	85.94	82.34	85.43	88.36	0.8384
	VGG16+RF	89.62	85.79	89.76	91.36	0.8773
	VGG16+SVM	91.48	87.33	92.03	95	0.8962
	Mobile-Net v2	95.68	96.19	95.72	97	0.9595
	Mobile-Net v2+RF	96.29	96.19	97.33	98	0.9676
	Mobile-Net v2+SVM	**99.39**	**100**	98.57	99	0.9926

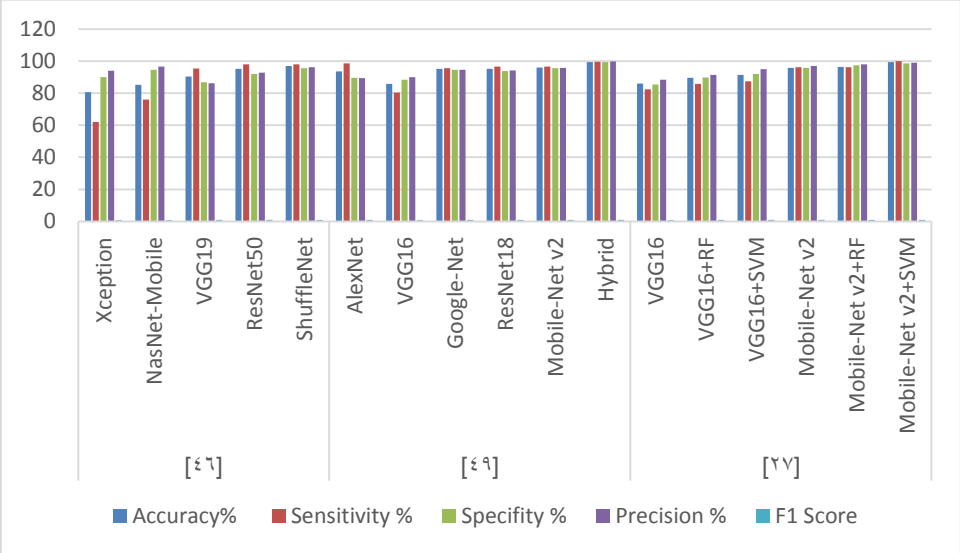

Figure 5: Graphical classification performance in ALL-IDB1.

It's important to acknowledge that most of the research on using deep learning to classify ALL using the ALL-IDB2 dataset has utilized transfer learning. This approach is preferred because it can achieve impressive outcomes even with limited data. Table 5 provides evidence of this trend.

Table 5: classification performance in ALL-IDB2.

Method	Classifier	Accuracy%	Sensitivity %	Specify %	Precision %	F1 Score
[34]	Xception	87.35	83.26	91.11	91.72	0.8701
	NasNet-Mobile	89.48	84.77	94.08	94.32	0.8918
	VGG19	83.59	88.45	80.93	78.76	0.8452
	ResNet50	93.85	92.9	94.82	94.82	0.9385
	ShuffleNet	96.67	96.46	96.95	96.9	0.967
[36]	AlexNet	91.54	94.87	89.77	88.21	0.9225
	VGG16	82.57	86.15	84.6	78.97	0.8537
	Google-Net	91.54	96.41	88.26	86.67	0.9216
	ResNet18	93.34	93.34	93.82	93.34	0.9358
	Mobile-Net v2	92.82	91.8	95.48	93.84	0.936
	Hybrid	97.18	95.9	**98.52**	**98.46**	0.9719
[16]	VGG16	85.22	88.13	83.04	82	0.8551
	VGG16+RF	86.92	87.53	86.42	86.05	0.8697
	VGG16+SVM	90	90.24	89.99	89.68	0.9011
	Mobile-Net v2	92.56	95.95	90.02	89.1	0.9289
	Mobile-Net v2+RF	94.87	96.95	93.58	92.63	0.9524
	Mobile-Net v2+SVM	**98.21**	**98.95**	97.61	97.37	**0.9828**

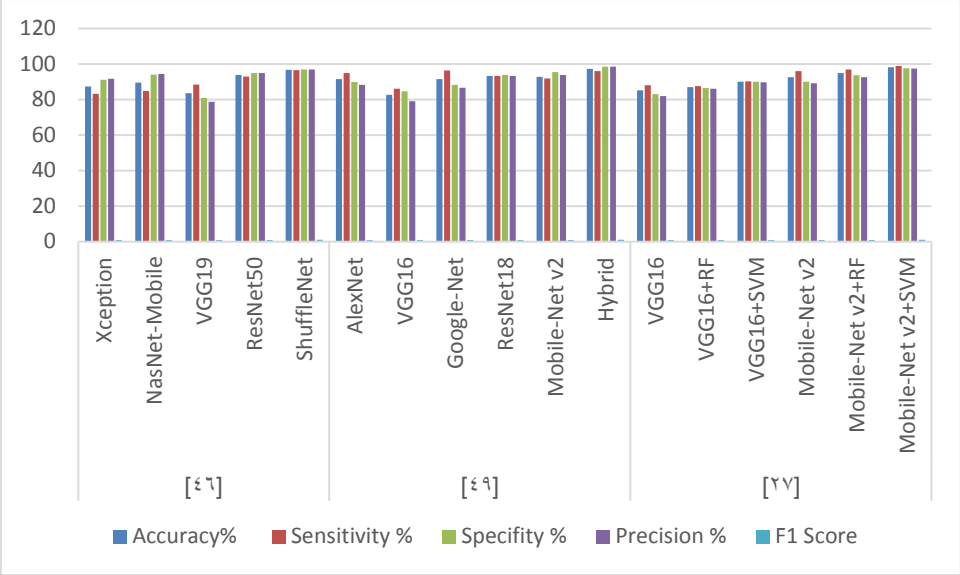

Figure 6: Graphical classification performance in ALL-IDB2.

As suggested, the MobileNetV2-SVM framework-based strategy is the most effective method for categorizing ALL. This is because it achieves the best accuracy and F1 Score performance by effectively fusing the benefits of SVM-based classification with MobileNetV2-based feature extraction. The suggested ALL-detection model, which is based on ShuffleNet, has the second-best specificity, precision, and accuracy performance, while the hybrid model that is being provided attains the second-best overall performance accuracy.

In contrast to conventional machine learning approaches, the deep learning methodology eliminates the need for additional segmentation procedures by completing segmentation, feature extraction, and classification tasks all inside a single neural network system. Therefore, it is challenging to get extraordinary performance using the typical deep learning technique as it requires a large dataset for effective training in order to accomplish efficient categorization.

The problem of needing a big dataset for effective classification has been solved by recent advancements in deep learning, particularly transfer learning techniques. Because of their classification performance and computational economy, AlexNet, MobileNetV2, and ResNet are the most

popular transfer learning techniques. Pretrained networks trained on smaller datasets can provide remarkable results. By combining the benefits of SVM-based classification with MobileNetV2-based feature extraction, the ALL-classification method based on the MobileNetV2-SVM framework has significant promise. Furthermore, Tables 4 and 5 show that the hybrid transfer learning model, which combines the advantages of ResNet18 and MobileNetV2, performs exceptionally well.

The two publicly accessible ALL databases, ALLIDB1 and ALLIDB2, are used in the majority of research to identify and categorize ALL. However, only a small number of studies undertake ALL categorization into all subtypes; the majority of research solely focuses on identifying healthy and ALL-affected patients. One of the most difficult topics for further study in this area is this.

5. CHALLENGING ISSUES:

Acute lymphoblastic leukemia (ALL) detection using deep learning algorithms has advanced significantly, however there are still a number of critical obstacles to overcome. The most significant of them is the challenge of correctly identifying leukemia subtypes, as the majority of existing systems rely on binary categorization (infected vs healthy). The lack of big, annotated datasets, which impedes the creation of highly generalizable models, is another issue. Additionally, a lot of deep learning models are not interpretable, which makes it hard for physicians to use these tools or trust them in practical situations.

Furthermore, in contexts with limited processing power and resources, model performance frequently deteriorates. One major obstacle is still the difficulty to implement big models for urgent clinical applications. In order to improve accuracy, robustness, and efficiency, researchers are increasingly looking into hybrid frameworks that combine the best features of many deep learning models.

A unique hybrid deep learning approach that combines convolutional neural networks (CNNs) and gated recursive units (GRUs) to enhance ALL

classification is proposed in this study in light of these difficulties, particularly the requirement to increase accuracy, generalization ability, and computational economy.

6. PROPOSED FRAMEWORK (CNN-GRU HYBRID MODEL):

The CNN–GRU hybrid approach, as shown in recent works, is highlighted in this review, highlighting its promise as an enhanced framework for ALL classification because of its balance between computing efficiency and accuracy This model improves accuracy, particularly on tiny or unbalanced datasets, by fusing the temporal modeling power of gated recursive units (GRUs) with the spatial feature extraction capabilities of convolutional neural networks (CNNs).

The architecture consists of a CNN block (with convolution, ReLU, and pooling layers), followed by a flattening and resampling phase to feed the GRU layer. This is followed by a dense layer with dropout for normalization, and a final Softmax layer for multi-class prediction. The model is designed to classify acute lymphoblastic leukemia (ALL) and its subtypes and is well-suited for deployment on edge devices.

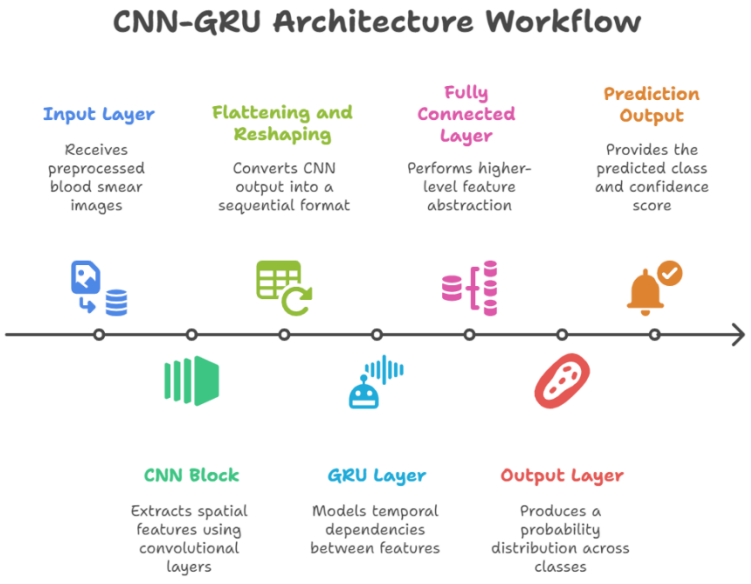

Figure 7: The Proposed Cnn-Gru Architecture

In Figure 7, the proposed CNN-GRU architecture processes blood smear images via a sequential pipeline that integrates spatial and temporal feature learning. The input image is first preprocessed and passed through a series of convolutional and max-pooling layers, which extract low- and high-level spatial features necessary for identifying leukemia patterns. These spatial features are then flattened and reshaped into a sequential form suitable for temporal modeling. This is followed by a gated recurrent unit (GRU) layer, which captures contextual dependencies and relationships across the extracted features. The GRU output is fed into a fully connected layer with dropout regularization to improve generalization. Finally, a softmax output layer performs multi-class classification, enabling the model to distinguish between healthy cells and multiple subtypes of acute lymphoblastic leukemia. This architecture is compact, supports subtype-level classification, and is well-suited for deployment in real-time diagnostic settings, including resource-constrained environments. Improvements such as attention modules, interpretable AI techniques, and multimodal data integration may further improve performance and clinical application.

7. FUTURE SCOPES:

Future research could expand the scope of the proposed CNN-GRU model by enabling subtype-level classification, integrating multimodal data (imaging, genomics), and incorporating explainable AI to facilitate interpretation. Efforts should also focus on immediate deployment to end devices, ensuring generalizability across diverse populations, and addressing ethical considerations such as privacy and fairness.

8. CONCLUSION:

This study examined the most recent deep learning methods for classifying acute lymphoblastic leukemia (ALL), emphasizing the use of autoencoders, convolutional neural networks (CNNs), transfer learning, and regression neural networks (RNNs). It also noted current drawbacks,

including interpretability issues, subtype classification difficulties, and data availability issues.

We suggested a hybrid convolutional neural network-grudging (CNN-GRU) model to fill these gaps by combining temporal and spatial learning to improve diagnostic efficiency and accuracy.. This architecture holds promise for improving traditional binary classification while also enabling subtype-level predictions. Future work should explore extending this architecture with attention modules, multimodal inputs, and edge propagation capabilities, with the goal of developing interpretable, real-time, and clinically applicable diagnostic tools.

References

[1] Della Starza, I., S. Chiaretti, M. S. De Propris, L. Elia, M. Cavalli, L. A. De Novi, … & R. Foà. 2019. "Minimal Residual Disease in Acute Lymphoblastic Leukemia: Technical and Clinical Advances." *Frontiers in Oncology* 9: 726.

[2] Jongen-Lavrencic, M., T. Grob, D. Hanekamp, F. G. Kavelaars, A. Al Hinai, A. Zeilemaker, … & P. J. Valk. 2018. "Molecular Minimal Residual Disease in Acute Myeloid Leukemia." *New England Journal of Medicine* 378 (13): 1189–99.

[3] Weijler, L., F. Kowarsch, M. Wödlinger, M. Reiter, M. Maurer-Granofszky, A. Schumich, & M. N. Dworzak. 2022. "UMAP-Based Anomaly Detection for Minimal Residual Disease Quantification within Acute Myeloid Leukemia." *Cancers* 14 (4): 898.

[4] Hallek, M., B. D. Cheson, D. Catovsky, F. Caligaris-Cappio, G. Dighiero, H. Döhner, … & T. J. Kipps. 2018. "iwCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of CLL." *Blood* 131 (25): 2745–60.

[5] Mishra, S., B. Majhi, & P. K. Sa. 2019. "Texture Feature Based Classification on Microscopic Blood Smear for Acute Lymphoblastic Leukemia Detection." *Biomedical Signal Processing and Control* 47: 303–11.

[6] Mishra, S., B. Majhi, & P. K. Sa. 2018. "GLRLM-Based Feature Extraction for Acute Lymphoblastic Leukemia (ALL) Detection." In *Recent Findings in Intelligent Computing Techniques: Proceedings of the 5th ICACNI 2017, Volume 2*, 399–407. Singapore: Springer.

[7] Rezaei, S., H. Ghasemian, M. H. Behzadi, & S. Jorfi. 2021. "Deep Learning-Based Diagnostic Techniques for Leukemia Detection: A Review." *Artificial Intelligence in Medicine* 113: 102015.

[8] Baig, R., A. Rehman, A. Almuhaimeed, A. Alzahrani, & H. T. Rauf. 2022. "Detecting Malignant Leukemia Cells Using Microscopic Blood Smear Images: A Deep Learning Approach." *Applied Sciences* 12 (13): 6317.

[9] Das, P. K., P. Jadoun, & S. Meher. 2020. "Detection and Classification of Acute Lymphocytic Leukemia." In *Proceedings of the IEEE-HYDCON 2020*, 1–5. IEEE.

[10] Lian, C., M. Liu, J. Zhang, & D. Shen. 2018. "Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis Using Structural MRI." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42 (4): 880–93.

[11] Das, P. K., A. Pradhan, & S. Meher. 2021. "Detection of Acute Lymphoblastic Leukemia Using Machine Learning Techniques." In *Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication*, 425–37. Singapore: Springer.

[12] Shahin, I., Y. Guo, K. M. Amin, & A. A. Sharawi. 2019. "White Blood Cells Identification System Based on Convolutional Deep Neural Learning Networks." *Computer Methods and Programs in Biomedicine* 168: 69–80.

[13] Gehlot, S., A. Gupta, & R. Gupta. 2020. "EDNFC-Net: Convolutional Neural Network with Nested Feature Concatenation for Nuclei-Instance Segmentation." In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1389–93.

[14] Vogado, L. H. S., R. D. M. S. Veras, A. R. Andrade, F. H. D. de Araujo, R. R. V. Silva, & K. R. T. Aires. 2017. "Diagnosing Leukemia in Blood Smear Images Using an Ensemble of Classifiers and Pre-Trained Convolutional Neural Networks." In *Proceedings of the 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 367–73. IEEE.

[15] Vogado, L. H. S., R. M. S. Veras, F. H. D. Araujo, R. R. V. Silva, & K. R. T. Aires. 2018. "Leukemia Diagnosis in Blood Slides Using Transfer Learning in CNNs and SVM for Classification." *Engineering Applications of Artificial Intelligence* 72: 415–22.

[16] Das, P. K., B. Nayak, & S. Meher. 2022. "A Lightweight Deep Learning System for Automatic Detection of Blood Cancer." *Measurement* 191: 110762.

[17] Celard, P., E. L. Iglesias, J. M. Sorribes-Fdez, R. Romero, A. Seara Vieira, L. Borrajo, et al. 2023. "A Survey on Deep Learning Applied to Medical Images: From

Simple Artificial Neural Networks to Generative Models." *Neural Computing and Applications* 35: 2291–2323. https://doi.org/10.1007/s00521-022-07953-4

[18] Houssein, E. H., M. M. Emam, A. A. Ali, & P. N. Suganthan. 2021. "Deep and Machine Learning Techniques for Medical Imaging-Based Breast Cancer: A Comprehensive Review." *Expert Systems with Applications* 167: 114161.

[19] Banik, P. P., R. Saha, & K.-D. Kim. 2020. "An Automatic Nucleus Segmentation and CNN Model-Based Classification Method of White Blood Cell." *Expert Systems with Applications* 149: 113211.

[20] Jha, K. K., & H. S. Dutta. 2019. "Mutual Information-Based Hybrid Model and Deep Learning for Acute Lymphocytic Leukemia Detection in Single Cell Blood Smear Images." *Computer Methods and Programs in Biomedicine* 179: 104987.

[21] Claro, M., L. Vogado, R. Veras, A. Santana, J. Tavares, J. Santos, & V. Machado. 2020. "Convolution Neural Network Models for Acute Leukemia Diagnosis." In *Proceedings of the International Conference on Systems, Signals and Image Processing (IWSSIP)*, 63–68. IEEE.

[22] Mahmud, M., M. S. Kaiser, A. Hussain, & S. Vassanelli. 2018. "Applications of Deep Learning and Reinforcement Learning to Biological Data." *IEEE Transactions on Neural Networks and Learning Systems* 29 (6): 2063–79.

[23] Hegde, R. B., K. Prasad, H. Hebbar, & B. M. K. Singh. 2019. "Comparison of Traditional Image Processing and Deep Learning Approaches for Classification of White Blood Cells in Peripheral Blood Smear Images." *Biocybernetics and Biomedical Engineering* 39 (2): 382–92.

[24] Reena, M. R., & P. M. Ameer. 2020. "Localization and Recognition of Leukocytes in Peripheral Blood: A Deep Learning Approach." *Computers in Biology and Medicine* 126: 104034.

[25] Simonyan, K., & A. Zisserman. 2014. "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv preprint arXiv:1409.1556.

[26] Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, & A. Rabinovich. 2015. "Going Deeper with Convolutions." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9.

[27] He, K., X. Zhang, S. Ren, & J. Sun. 2016. "Deep Residual Learning for Image Recognition." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–78.

[28] Howard, A. G., M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, & H. Adam. 2017. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." arXiv preprint arXiv:1704.04861.

[29] Sandler, M., A. Howard, M. Zhu, A. Zhmoginov, & L.-C. Chen. 2018. "MobileNetV2: Inverted Residuals and Linear Bottlenecks." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4510–20.

[30] Chollet, F. 2017. "Xception: Deep Learning with Depthwise Separable Convolutions." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1251–58.

[31] Shafique, S., & S. Tehsin. 2018. "Acute Lymphoblastic Leukemia Detection and Classification of Its Subtypes Using Pretrained Deep Convolutional Neural Networks." *Technology in Cancer Research & Treatment* 17: 1533033818802789.

[32] Yu, W., J. Chang, C. Yang, L. Zhang, H. Shen, Y. Xia, & J. Sha. 2017. "Automatic Classification of Leukocytes Using Deep Neural Network." In *Proceedings of the IEEE 12th International Conference on ASIC (ASICON)*, 1041–44.

[33] Mallick, P. K., S. K. Mohapatra, G.-S. Chae, & M. N. Mohanty. 2020. "Convergent Learning-Based Model for Leukemia Classification from Gene Expression." *Personal and Ubiquitous Computing* 2: 1–8.

[34] Das, P. K., & S. Meher. 2021. "Transfer Learning-Based Automatic Detection of Acute Lymphocytic Leukemia." In *Proceedings of the National Conference on Communications (NCC)*, 1–6. IEEE.

[35] Genovese, A., M. S. Hosseini, V. Piuri, K. N. Plataniotis, & F. Scotti. 2021. "Acute Lymphoblastic Leukemia Detection Based on Adaptive Unsharpening and Deep Learning." In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1205–09.

[36] Das, P. K., & S. Meher. 2021. "An Efficient Deep Convolutional Neural Network Based Detection and Classification of Acute Lymphoblastic Leukemia." *Expert Systems with Applications* 183: 115311.

[37] Redmon, J., S. Divvala, R. Girshick, & A. Farhadi. 2016. "You Only Look Once: Unified, Real-Time Object Detection." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–88.

[38] Redmon, J., & A. Farhadi. 2017. "YOLO9000: Better, Faster, Stronger." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7263–71.

[39] Redmon, J., & A. Farhadi. 2018. "YOLOv3: An Incremental Improvement." arXiv preprint arXiv:1804.02767.

[40] Bochkovskiy, A., C.-Y. Wang, & H.-Y. M. Liao. 2020. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934.

[41] Gehlot, S., A. Gupta, & R. Gupta. 2020. "SDCT-AuxNet: DCT Augmented Stain Deconvolutional CNN with Auxiliary Classifier for Cancer Diagnosis." *Medical Image Analysis* 61: 101661.

[42] Duggal, R., A. Gupta, R. Gupta, & P. Mallick. 2017. "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 (MICCAI)*, 435–43. Springer.

[43] Schuster, M., & K. K. Paliwal. 1997. "Bidirectional Recurrent Neural Networks." *IEEE Transactions on Signal Processing* 45 (11): 2673–81.

[44] Hochreiter, S., & J. Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9 (8): 1735–80.

[45] Scotti Project. n.d. "ALL-IDB: Acute Lymphoblastic Leukemia Image Database for Image Processing." Accessed October 3, 2024. https://scotti.di.unimi.it/all/.

[46] Hematology Atlas. n.d. "Principal Page." Accessed October 9, 2024. http://hematologyatlas.com/principalpage.htm.

[47] Gan, Sheng. 2018. *BCCD Dataset*. GitHub. Accessed October 12, 2024. https://github.com/shenggan/BCCD_Dataset.

[48] Al-Qudah, R., & C. Y. Suen. 2020. "Synthetic Blood Smears Generation Using Locality Sensitive Hashing and Deep Neural Networks." *IEEE Access* 8: 102530–39.

[49] Das, P. K., Diya V. A., & others. 2022. "A Systematic Review on Recent Advancements in Deep and Machine Learning Based Detection and Classification of Acute Lymphoblastic Leukemia." *International Journal of Intelligent Systems and Applications in Engineering* 10: 1–10.