Development of a Diagnostic Support Model for Acute Leukemias Based on Artificial Intelligence: HematologIA

Authors: A BORNACELLY[1], J SPROCKEL[1,2], MH SOLANO[1,2] and L PARRA[2]
1. Fundación Universitaria Ciencias de la Salud, Bogota, Colombia
2. Hospital de San Jose, Bogota, Colombia

Introduction

Acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) are heterogeneous hematopoietic neoplasms with poor outcomes. The diagnosis of AML and ALL in Colombia has significant challenges due to delays in diagnosis confirmation, which can range from 9 to 24 days, or even longer in areas with limited access to healthcare services. Images analysis throw deep learning using visión transformers techniques can be affordable strategies for developing diagnosis support models for acute leukemias.

Aim

To present results of an Artificial intelligence (AI) based model to differentiate myeloblasts from lymphoblast (hematologIA)

Methods

This was a retrospective, diagnosis study conducted in the Hospital de San José in Bogotá analyzing bone marrow smears of 24 AML, 19 ALL and 5 APL cases attended between 2021 and march 2025 (Table 1). The bone marrow smears images were digitized using the Roche Ventana DP 200 scanner, annotated using the QuPath software and curated by two expert morphologists (Fig. 1). A Phyton-based code was developed for the recognition of blast cells, emphasizing in differentiation between myeloblasts and lymphoblasts using PyTorch library. A second task, the classification of APL cases was also analyzed. Five models were trained, two of them based on vision transformers (ConvNext, Swin) and the other three models based on CNN (EfficientNetB0, VGG16 and ResNet50). Evaluation was performed with the validation set.

Table 1. General characteristics of the population

Variable	AML (n= 24)	ALL (n = 19)
Sex M n (%)	15 (75)	8 (57)
Mean Age ± SD (range)	53 ± 16 (26-80)	40 ± 18 (20-76)
% mean blast count±SD (peripheral blood)	35 ± 25	46 ± 30
% mean blast count ± SD (bone marrow)	48 ± 26	73 ± 15
% mean blast count ± SD (bone biopsy)	57 ± 27	86 ± 12
Diagnostic features (%)		
Anemia	24 (100)	19 (100)
Thrombocytopenia	24 (100)	16 (84)
Neutropenia	20 (83)	16 (84)
Hyperleukocytosis	2 (8.3)	8 (42)
Splenomegaly	1 (4.1)	5 (26)
Comorbidities n (%)		
Cardiovascular	7 (29)	4 (21)
Autoimmune	2 (8.3)	1 (5.2)
Chronic kidney disease	1 (4.1)	3 (15.8)
Molecular or cytogenetic abnormality n (%)		
Normal karyotype	20 (80)	17 (90)
Complex karyotype	--	2 (10)
t(8;21) (q22;q22)	3 (12.5)	--
FLT3 positive	3 (12.5)	--
t11q23 (KMT2A) positive	1 (4.1)	--
NPM1 positive	1 (4.1)	--
t(9;22) positive	--	4 (21)
Extramedullary disease yes, n (%)	1 (4.1)	3 (15.8)
Induction treatment n (%)		
7+3#	18 (75)	--
Azacitidine + venetoclax	3 (12.5)	--
Azacitidine	1 (4.1)	--
Re-induction therapy	6 (25)	--
FRALLE93	--	3 (16)
GRAALL	--	6 (31.5)
HyperCVAD	--	8 (42.1)
Pre-phase only	--	2 (10)
Bone marrow transplantion yes,n (%)	6 (25)	9 (47.3)
Outcome n (%)		
Alive	7 (29)	8 (42.1)
Dead	12 (50)	8 (42.1)
Unknown	5 (21)	3 (15.8)

The median age at diagnosis was 53±16 years in AML, 40±18 years for ALL, and 53±16 years for APL, with a predominance of males (75% in AML cases, 57% in ALL cases and 80% en APL cases).

CNN: convolutional neural networks
APL: acute promyelocytic leukemia

Results

A total of 5454 images from AML and 4689 from ALL were used for training; 606 images from AML and 522 from ALL for validation. The best accuracy for morphological differentiation between myeloblasts and lymphoblast was performed by the models based on vision transformers, the ConvNeXt small model showed 99.67% sensitivity, 98.58% specificity, 99.02% value predictive positive and 99.61% negative predictive value with an AUC-ROC of 0.998 (Table 2 and Fig. 2).

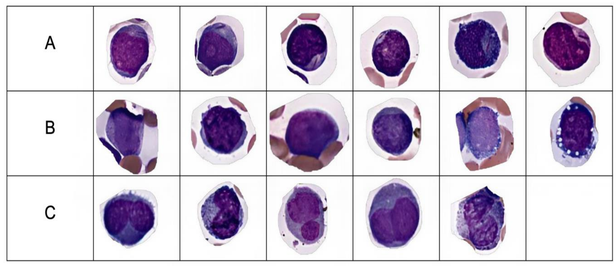


Figure 1. Examples of image patches used for AI model training

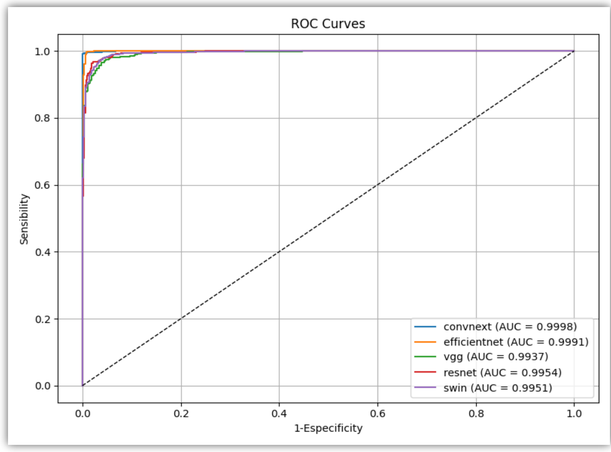


Figure 2. AUC of the different models for the classification of myeloblasts and lymphoblasts

The models EfficientNet B0 and ConvNeXt Small were the least accurate for the differentiation of APL whereas models like ConvNeXt Large, which has a higher capacity architecture achieved a 64.58% accuracy and an AUC-ROC of 0.9407 (Table 2).

Table 2. Results of the trained models for blast classification

Model	Sensit (%)	Spec (%)	Accur (%)	PPV (%)	NPV (%)	LR+	LR-	AUC	AUPresRecall	F1 score (%)	MCC (%)
Models for differentiating lymphoblasts from myeloblasts											
Swin Tiny	89,11	99,23	93,79	99,26	88,7	116,28	0,1098	0,9951	0,9957	93,91	88,15
ConvNext Small	99,67	98,85	99,29	99,02	99,61	86,71	0,0033	0,9998	0,9998	99,34	98,58
EfficientNet B0	99,83	98,47	99,20	98,69	99,81	65,14	0,0017	0,9991	0,9990	99,26	98,40
VGG 16	89,93	98,85	94,06	98,91	89,43	78,24	0,1018	0,9937	0,9947	94,21	88,56
ResNet50	96,70	97,51	97,07	97,83	96,22	38,82	0,0338	0,9954	0,9958	97,26	94,13
Models for differentiating myeloblasts from promyeloblasts											
EfficientNet B0	2,08	99,50	92,35	25,00	92,77	4,20	0,9840	0,9407	0,4283	3,85	5,31
ConvNext Small	0,00	100	92,66	0,00	92,66	---	1,0000	0,8819	0,4060	0,00	0,00
ConvNext Large 1	64,58	92,41	90,37	40,26	97,05	8,50	0,3833	0,9470	0,4290	49,60	46,11
ConvNext Large 2	54,17	93,07	90,21	38,24	96,25	7,81	0,4925	0,9426	0,3802	44,83	40,36

Discussion

The results obtained in this study show that the AI-based bone marrow morphological analysis system can be a tool for the successful differentiation of myeloblasts and lymphoblasts, specially implementing vision transformers-based models. Large-scale validation studies are warranted to further validate the clinical utility of this system.

References

1. Liao H et al. Application of artificial intelligence in laboratory hematology: Advances, challenges, and prospects. Acta Pharmaceutica Sinica B. Chinese Academy of Medical Sciences; 2025
2. Oybek Kizi RF et al. A Review of Deep Learning Techniques for Leukemia Cancer Classification Based on Blood Smear Images. Applied Biosciences. Multidisciplinary Digital Publishing Institute (MDPI); 2025 (4):1–32