Deep-Learning-Based Acute Leukemia Classification Using Imaging Flow Cytometry and Morphology

Jakkrich Laosai*
dept. Telecommunications and Data Communications
Faculty of Industrial Technology
*Ubon Ratchathani Rajabhat University, Thailand
jakkrich.l@ubru.ac.th

Kosin Chamnongthai†
dept. of Electronic and Telecommunication Engineering
Faculty of Engineering
†King Mongkut's University of Technology Thonburi
kosin.cha@kmutt.ac.th

***Abstract*— This paper presents to classify a blood cell into one type out of 11 subtypes of leukemia, which is important for diagnosis. Originally, the classification requires a huge training dataset collected from their races, which depend upon genes. Due to the recent discovery of imaging flow cytometry, it becomes possible to manually classify acute leukemia among different people and races by using a small training dataset. This paper proposes a method of automatic acute leukemia classification using imaging flow cytometry and morphology. The method utilizes intensity and morphology, which are the features from the imaging flow cytometry and blood smear, respectively, to classify a blood cell into one out of three groups; healthy, ALL, and AML ones in the coarse step, and categorizes it into a subtype in the group by deep learning in the fine step. The evaluation has been performed by some samples between Thai and American people, and the results show better accuracy compared with conventional methods. The computer simulations show the proposed system robustly segments and classifies Acute Leukemia based on complete microscopic blood images. We have obtained an accuracy of 99% which is a 4% improvement compared with the conventional method.**

Keywords— imaging flow cytometry (IFC); deep-learning (DL); acute lymphocytic leukemia (ALL); acute myelogenous leukemia (AML).

I. INTRODUCTION

Counting the number of white blood cells or leukocytes plays a crucial role as an indicator in the diagnosis of leukemia, which is one of the most critical diseases. Leukemia is namely, acute myelogenous leukemia (AML), acute lymphoblastic leukemia (ALL) types of leukemia, AML and ALL affect young children, and is an urgent need for diagnosis, especially in the early stages, to treat and cure the patient. For this problem, this paper provides the research of the classification of acute leukemia by using image processing to identify types of acute leukemia. The values pattern of Leukemia cells image after processing will be compared to the values pattern of standard Leukemia cells image automatically and then classified. Our techniques will help medical practitioners diagnose the types of Acute Leukemia faster and with greater efficiency. Our work focuses on the classification of Foil of Bretagne (Lymphoid) and Almeida Lloyd (Myeloid). Classify it into a type among various types of AML and ALL which are M0, M1, M2, M3, M4, M5, M6, and M7, and L1, L2, and L3, respectively.

II. IMAGING FLOW CYTOMETRY AS PREVIOUS WORK

Imaging flow cytometry (IFC) combines the high-throughput, multiparameter capabilities of conventional flow cytometry with morphological and spatial information, all at single-cell resolution. Multichannel digital images of hundreds of thousands of individual cells can be captured within minutes (Fig. 1), and include fluorescence channels, bright field, and dark field. The throughput of IFC means that it is especially well suited to the analysis of rare cell types such as circulating tumor cells and transition states, such as (cell cycle phases) [1].

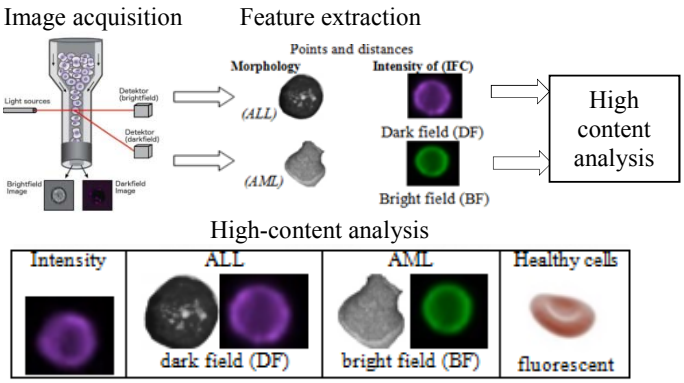

Fig. 1. Imaging flow cytometry (IFC)

IFC Acquires Images of Single Cells in High Throughput [2]. Typical throughput is 5,000 cells/s , but modifications to the instruments can increase this to 100,000 cells/s [3]. The images of each cell captured by charge-coupled device (CCD) detectors as the cell flows past the light sources (left panel) and morphological features are extracted (middle panel). For instance, the patient's blood could be analyzed to distinguish leukemic from normal cells or monitor in vitro or in vivo response to therapeutic intervention (right panel). High-content analysis could be used for personalized diagnosis, prognosis, and therapy [4]. These mentioned manual procedures work well based on the technician skill so that it should be developed in an automatic system. In this paper, Deep Learning [5], which is recently considered as a powerful machine-learning tool is used as a classifier in order to make the manual procedure of IFC become an automatic function.

As verification, this paper employs a network as shown in Fig. 2 containing 4 layers with an input image size of 50x50x3. The convolution layer 2 has the same structure as the convolution layer 1. The Max-Pooling layer 25x25 filter size is 2 and stride is 2. The fully connected layer has 2 neural. The system overview of the proposed method is depicted in Fig. 3, and details of all processes are described in Section 3, 4, and 5.

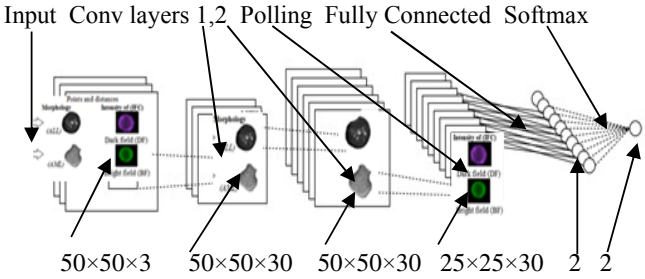

Fig. 2. Illustration of a Convolutional Neural Network.

III. FEATURES BASED ON MEDICAL KNOWLEDGE

Pattern recognition is the selecting of the Intensity image by the numerical values and enabling the automatic system to perform the recognition. Usually many of the features use texture, geometrical and statistical analyses of the image [6]. Feature Extraction Texture, Size, Shape, and Intensity training-testing sets. Those are differentiated based on morphology. Morphology includes the cell size, prominence of nucleoli, the color of a cell and the amount and appearance of cytoplasm. The FAB Features based on medical knowledge classification IFC measures intensity in many different ways, multiparameter capabilities of conventional flow cytometry with morphological and spatial information, all at single-cell resolution, is shown in the Table1. We utilize a CNN to build a regression model between a cell microscopy image and its cell. Location representation: compressed signal IFC. We employ two kinds of CNN architectures [7]. The amplitude values of the Fusion Group b dark field (DF) and bright field (BF). Acquired using a fluorescence microscope.

TABLE I Features based on medical knowledge
Acquired using a fluorescence microscope

Features	ALL	AML	Healthy cells
Texture	No Nucleoli 1_3	No Nucleoli 4_11	No Nucleoli 1
Size	8 µm	12 µm	6 µm
Shape	Oval Shape	Round Shape	No Nucleus
Intensity	 dark field (DF)	 bright field (BF)	 fluorescent

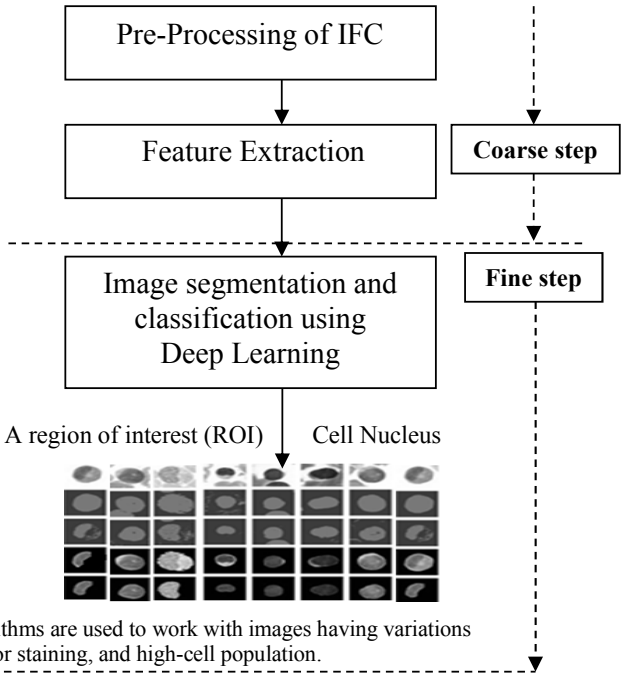

Fig. 3. Systems Overview

IV. PRE-PROCESSING: COLOR-CORRELATION

Coarse step

The preprocessing can be divided into coarse and fine steps as follows**.** The coarse step in the proposed method consists of pre-processing and feature extraction processes with the objective of classification into ALL, AML, and healthy groups as shown in the upper part of Fig.3. Pre-processing Intensity of IFC uses fluorescent microscopy with advances in data-processing algorithms and imaging flow cytometry and Morphology to produce the clearest picture possible. Because the blood-cell image includes noise in general, Feature Extraction determines the Texture, Size, Shape, and Intensity training-testing [8]. Users should select an appropriate ALL and AML based on the morphology and Cytochemical staining of blasts. Those are differentiated based on morphology. The classified acute leukemia groups including ALL and AML are then processed in the fine step.

Fine step

The fine step comprises Image segmentation and classification using Deep Learning and A region of interest (ROI) Cell Nucleus as shown in the lower part of Fig.3. In the nucleus extraction, a blood cell comprising the cytoplasm and nucleus is segmented to obtain its nucleus which is assumed to include important medical information as discussed in Section 3. After pattern recognition comes the selecting of the intensity image determined by the numerical values which will enable the automatic system to perform the recognition shown in Fig.2, it is possible to separate the cell nucleus from the cytoplasm. In the nucleus-segmentation process introduced in subsection algorithms to work with images having variations in color staining, and high-cell population contains. The extracted nucleus of the blood cell, which important

information, undergoes feature extraction (mentioned in Table 1), and the feature-based classification is fed to the classifier in the classification process (explained in Table 2). The output is finally evaluated in with images having variations in color staining and high-cell population. (described in Fig.3).

TABLE II Feature-based Classification (IFC) and Morphology

Signal Prediction by Convolutional Neural Network	
Morphology (ALL)	**Intensity of (IFC)** Extraction Texture
Fusion Group a	Index Fusion Group b (Min)
L1	Signals 1 nm dark field (DF)
L2	Signals 2 nm dark field (DF)
L3	Signals 3 nm dark field (DF)
Morphology (AML)	**Intensity of (IFC)** Extraction Texture
Fusion Group a	Index Fusion Group b (Max)
M0	Signals 4 nm bright field (BF)
M1	Signals 5 nm bright field (BF)
M2	Signals 6 nm bright field (BF)
M3	Signals 7 nm bright field (BF)
M4	Signals 8 nm bright field (BF)
M5	Signals 9 nm bright field (BF)
M6	Signals 10 nm bright field (BF)
M7	Signals 11 nm bright field (BF)

Cellular antigens to be determined (i.e. cell surface, cytoplasm, nuclear), a feature that cannot be achieved by standard flow cytometry. Specific cell populations can be selected for analysis based on combined variables such as fluorescent intensity, cell shape, cell size, and texture. Groups will be created using similar features from channel intensity, such as image segmentation and key point matching. IFC allows the analysis of heterogeneous cell populations.

V. CONVOLUTIONAL NEURAL NETWORKS

In this, nuclei are extracted by the network learns to enhance complex abstraction in images, such as the case of IFC method [9]. In contrast to deep learning, in conventional image analysis, the images are first preprocessed, then cellular objects are identified (segmented) by analysis software, followed by the extraction of the max and min is based on minimizing the objective function is the point of distance.

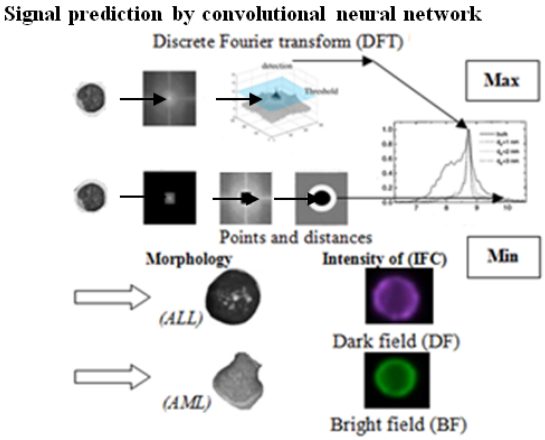

Fig 4. Harmonic field parameterization

In order to parameterize the harmonic field, we used the method proposed [10]. In this algorithm (see Fig. 4), we first solved the sinusoidal using the discrete Fourier transform (DFT) followed by locating harmonic peaks by identifying the largest isolated peaks in the harmonic frequencies [11].

VI. COMPUTER SIMULATION

Deep-Learning (DL) is a powerful tool for data classification based on hyperplane classifier. This classification is achieved by a separating surface (linear or nonlinear) in the input space of the data set. Image Processing and Deep Learning optimize the margin between the classes. The classifier training algorithm is a procedure to find the DL. Using Imaging Flow Cytometry and Morphology-based features are extracted for the image nucleus sample and recorded. Few measurements are tabulated in Table 1. Among all the features, the most relevant features are selected and used to train the conceivable uses of IFC. For instance, a differential diagnosis of acute leukemia diagnostics [12]. Online data hospital management system (HMS) precision data is available at https://github.com/jagonzalez68/leukemiaDB.

VII. PERFORMANCE EVALUATION

Precision, Specificity, Sensitivity, and F-Measure are all defined in relation to the possible outcomes of the classifier system [13]. Healthy cells training - testing using a set of 100 images, 50 normal cell images, and 50 abnormal cell images. The experiment on Matlab and the accuracy rate of the recognition of leukemia by our proposed CNN model. In this experiment, the original ALL-IDB image database which consists of 100 cell image [14], we have obtained accuracy of 99 %.

TABLE III. PERFORMANCE EVALUATION PARAMETERS

No.	Evaluation parameters		
	Parameters	**Formulae**	Value
1	Sensitivity	TP/(TP+FN)	99%
2	Specificity	TN/(TN+FP)	99%
3	Precision.	TP/(TP+FP)	99%
4	F-Measure	$\frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$	99%

TABLE IV. Performance Evaluation Statistics

	Type	Sub-Type	%	SD
Weka [10]	Healthy	Normal cell	95%	5
	ALL,AML	Abnormal cell	90%	10
Deep Learning	Healthy	Healthy	99%	1
	ALL	L1	99%	1
		L2	99%	1
		L3	99%	1
	AML	M0	99%	1
		M1	99%	1
		M2	99%	1
		M3	99%	1
		M4	99%	1
		M5	99%	1
		M6	99%	1
		M7	99%	1

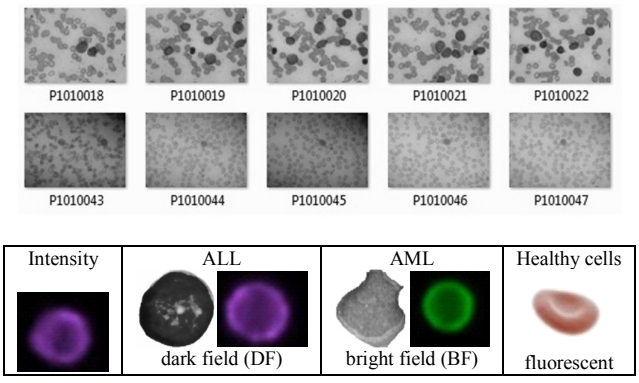

Fig. 5. AML and ALL data set using Intensity of (IFC)

VIII. CONCLUSION

Accurate classification of acute leukemia into corresponding sub-types is very important to properly determine a method of Deep-Learning-Based Acute Leukemia Classification. Using Imaging Flow Cytometry and Morphology Leukemia is divided into 2 categories ALL and AML. The features are extracted from the segmented images and classified using the Deep-Learning (DL). The method has been evaluated using a set of 100 images with 50 abnormal samples and 50 normal samples obtained. The method has been evaluated using Deep-Learning-Based Acute Leukemia Classification Using Imaging Flow Cytometry and Morphology. Features extracted from the IFC are characterized by imaging flow cytometry (IFC) Feature Extraction Texture, Size, Shape, and Intensity training-testing sets on 100 datasets. The classification proposed 3 subtypes of ALL and 8 subtypes of AML We have obtained classification accuracy of 99%, which is an improvement of 4%.

ACKNOWLEDGMENT

The authors would like to extend their appreciation to the medical doctors and technologists at the Division of Hematology and Oncology, Ubonratchathani Cancer Hospital, Thailand for their kind cooperation in sample preparation and in providing high quality image data.

REFERENCES

[1] L.F. Ogle, et al. Imagestream detection and characterisation of circulating tumour cells – a liquid biopsy for hepatocellular carcinoma J. Hepatol., 65 (2016), pp. 305-313

[2] T. Blasi, et al. Label-free cell cycle analysis for high-throughput imaging flow cytometry Nat. Commun., 7 (2016), Article 10256

[3] C.L. Chen, et al. Deep learning in label-free cell classification Sci. Rep., 6 (2016), p. 21471

[4] Blasi T, Hennig H, Summers HD, et al. Label-free cell cycle analysis for high-throughput imaging flow cytometry. Nat Commun . 2016;7:10256

[5] C. Angermueller, et al.Deep learning for computational biology Mol. Syst. Biol., 12 (2016), p. 878

[6] E. Gerdtsson, et al.Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry Converg. Sci. Phys. Oncol., 4 (2018), p. 015002

[7] Jakkrich Laosai and Kosin Chamnongthai "Classification of Acute Leukemia Using CD Markers" 2016 International Conference on Knowledge and Smart Technology (KST2016)

[8] J.C. Caicedo, et al.Data-analysis strategies for image-based cell profiling Nat. Methods, 14 (2017), pp. 849-863

[9] Y. Han, et al.Review: imaging technologies for flow cytometry Lab Chip, 16 (2016), pp. 4639-4647.

[10] Reta C, Altamirano L, Gonzalez JA, DiazHernandez R, Peregrina H, Olmos I, et al. (2015) Segmentation and Classification of Bone Marrow Cells Images Using Contextual Information for Medical Diagnosis of Acute Leukemias. PLoS ONE 10(6): e0130805

[11] M. F. Lohrer et al., "Applying Pattern Recognition to High-Resolution Images to Determine Cellular Signaling Status," in IEEE Transactions on NanoBioscience, vol. 16, no. 6, pp. 438-446, Sept. 2017.

[12] L.F. Grimwade, et al.Applications of imaging flow cytometry in the diagnostic assessment of acute leukaemia Methods, 112 (2017), pp. 39

[13] Hoo-Chang Shin., et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Transactions on Medical Imaging. (2016), Volume: 35, Issue: 5 pp, 1285 – 1298

[14] R.D.Labati, V.Piuri, and F.Scotti, "ALL-IDB: The Acute Lymphoblastic Leukemia Image Database for Image Processing," in 18th IEEE International Conference on Image Processing, 2011.