International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Development of an AI-Based Acute Promyelocytic Leukemia Classification System

Peter Zhu

Abstract: ***Acute promyelocytic leukemia (APL) is an extremely malignant form of acute leukemia and must be treated as soon as possible; however, due to a lack of personnel and equipment, APL is often misdiagnosed in rural regions and developing countries. To solve this problem, I developed an AI-based classification system that determines, from conventional microscopic images of leukocytes, whether or not APL is present.***

Keywords: convolutional neural networks, acute promyelocytic leukemia, machine learning

1. Introduction

Acute promyelocytic leukemia (APL) is a distinct subtype of acute myeloid leukemia with a fatal course of only weeks. Due to coagulopathic complications, it is absolutely necessary to start treatment as soon as APL is suspected in the diagnosis, as even a delay of hours could cause serious impediments and even early hemorrhagic death. There exist many advanced techniques such as flow cytometry, fluorescence in situ hybridization and cytogenetic analysis to physically and chemically determine the presence of APL cells, but the manual microscopic examination of blood smears always remains the first step in leukemia diagnosis. Manual examination of blood films by a technician is effective in a lot of scenarios as a technician can very easily determine irregularities in the patient's blood film; however, with over hundreds of distinct types of irregular leukocytes, a technician, from time to time, would not be able to recognize the specific cell, and in rural regions where there is limited access to sophisticated equipment, this will more than likely lead to transporting the patient's blood slide to the closest tertiary hospital for further diagnosis. This process can take up to several days, and although such a delay would not necessarily present a major problem for most blood conditions, it could definitely compromise the survival rate of a patient affected by APL. In fact, a study done by Silva et al has shown that in Brazil, the 10-year survival rate of APL is lower than 60%, as opposed to the 90-95% seen in Canada and the United States. The purpose of this work is to develop a computer-vision based APL classification system. Such a system will offer two very important features: immediacy and remoteness. Combined with just a smartphone and a microscope, this system can determine, in seconds, whether or not a cell is APL, thus providing crucial assistance to clinics in rural regions.

2. Procedures and Results

Data Acquisition

A total of 17 574 augmented blood smear images, each containing one APL cell and 34 177 augmented blood smear images, each containing one non-APL cell, stained using Wright's stain and labeled by hematopathologists, were acquired from the Dalhousie Morphology Lab, Mount Allison Biology Lab, and Università Degli Studi di Milano's ALL-IDB database.

Stage 1: Image Pre-Processing & Segmentation

The first objective in this work is to develop a method to ensure that all blood smear images, regardless of luminance or color variations, can be processed concordantly. Existing methods, such as the contour tracking method proposed by Lisicki, are mainly geared towards pinpointing leukocytes rather than segmenting the leukocyte for feature extraction. This work proposed new developments such as the use of K-means and Fuzzy C-means clustering to minimize background noise when segmenting a cell, Gray World Assumption algorithm to equalize luminance levels across images of different brightness, and as well as Otsu's thresholding method to separate the cell nucleus from the cytoplasm by locating the peaks on the gray-level intensity histogram. This was done to extract important features from the leukocyte image, such as the nucleus perimeter and area, and the nucleus to cytoplasm ratio (N:C), which will be used in Stage 2.1.

Stage 2.1: Image Classification Using Traditional Machine Learning Methods

The proposed classifier would take certain extracted properties of a pre-processed leukocyte as the input and give a True/False output of whether or not the leukocyte is an APL cell. To achieve this, traditional machine learning classifiers were employed, and morphological features of the leukocyte image, driven by suggestions of hematopathologists, were selected as inputs. A total of 21 morphological features were selected. These features can be generalized under four main categories: Gray-level intensity properties, consisting of granularity and uniformity; shape and size properties, consisting of perimeter, area, convex area, filled area, momentums, solidity, major axis length, minor axis length, orientation, eccentricity, compactness, circular variance, elliptical variance, and elongation; texture properties, consisting of energy and entropy, and color properties, consisting of cell energy and standard deviation. However, in order to avoid overfitting, only the top four most relevant features were used. These features were found by applying forward selection based on k-nearest neighbors classifiers, and evaluating it with the Leave One Out method.

Stage 2.1: Results and Observations

The maximum accuracy achieved with this method was only 86.56% (Fig. 2), and the lack of robustness in this system can be interpreted as an absence of precise features. In most

cases of traditional machine learning, there are no standards for feature extraction and classification. Within each step, approaches could be mixed and matched with any number or type of classier, resulting in a never-ending source of work. Additionally, one of the most defining attributes of an APL cell is the presence of cells with an abundance of Auer rods — large, cytoplasmic inclusion bodies that appear in cells. Unfortunately, traditional feature extraction approaches do not offer a robust technique of precisely locating these lements.

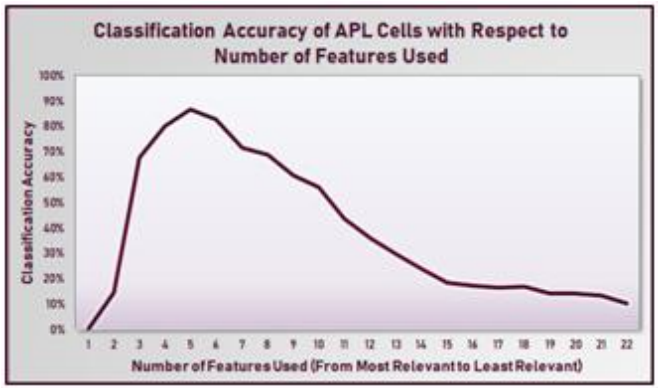

Figure 1: Due to overfitting, the classification accuracy peaks at 4 features

Stage 2.2: Image Classification Using a Convolutional Neural Network

Type	Mean Error	Standard Deviation	Features Used
LDA	0.1344	0.1288	4
LINEAR-B	0.1722	0.1539	4
KNN	0.142	0.1288	4

Figure 2: Classification accuracies of traditional machine learning classifiers

A novel Convolutional Neural Network (CNN) architecture is proposed to replace the process of feature extraction and selection and as well as the use of traditional machine learning classifiers. It was hypothesized that a CNN will increase the accuracy and robustness of the classification system as it is capable of automatic feature extraction and absent of the major flaws of traditional machine learning classifiers shown in Stage 2.1. The architecture of the proposed CNN includes three main types of layers, namely, convolutional layers, pooling layers, and fully-connected (FC) layers. The size of the input image is set to 400x400x3. In the first convolutional layer, 16 different filters with the size 5x5 were employed and the stride was set to 1. The second and third convolutional layers consist of the same structure, albeit a different number of filters — 32 for the second layer, and 64 for the third. Pooling layers were used to spatially decrease the volume of the input, and filters with a size and stride of 2 were employed. Finally, a mini-batch size of 200 was chosen for the training process of the classifier. The CNN was trained with 12 302 images of preprocessed and segmented APL cells and 23 924 images of various types of other leukocytes, consisting mainly of eosinophils, neutrophils, basophils, monocytes and lymphocytes, but comprising of an assortment of different shapes. Additionally, 5 272 unused images of APL cells and 10 253 additional images of non-APL leukocytes were used for testing the classifier.

Stage 2.2: Results and Observations

The CNN significantly improved the system performance, and reached an average accuracy of 99.2%. After training the neural network, an analysis was done on the filters of the first convolution layer and it was shown that one of the filters had very high neuron activations nearing the regions of Auer rods, as represented in Fig. 4. As well, many other features such as texture and color were also recognized by other convolution filters. This confirms the hypothesis of this method, proving that feature extraction is much more effective when done automatically.

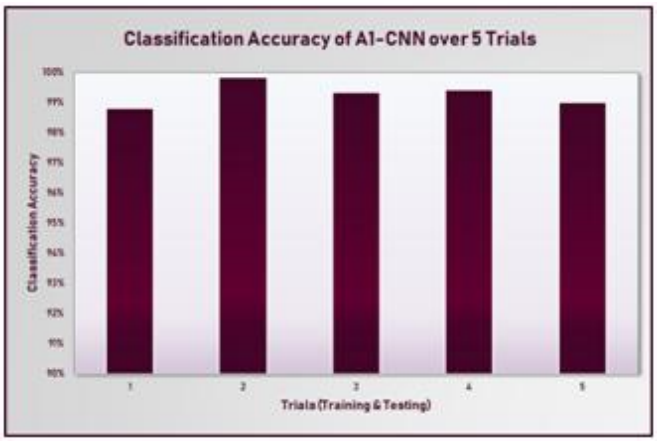

Figure 3: Accuracies of the CNN over 5 trials

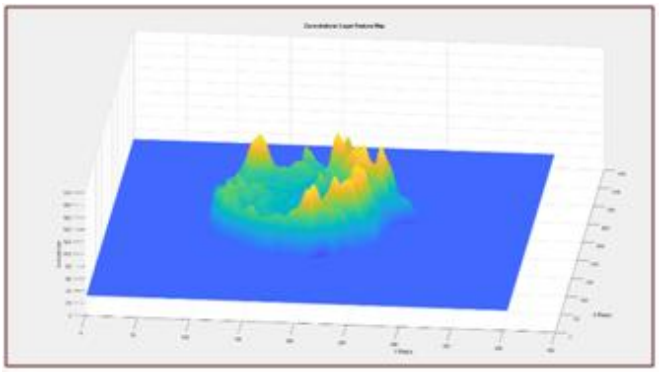

Figure 4: Convolution layer activation map, showing the filter that recognizes Auer rods

3. Discussion

Stage 1 Discussion

In Stage 1, a variety of techniques were used to segment and color correct the leukocyte image for classification. Of the techniques used, it is of interest to note that Otsu's thresholding method, when applied on APL cells containing an abundance of Auer rods (also known as faggot cells), can actually extract certain parts of the Auer rods from both the nucleus and cytoplasm. However, this method does not seem the most intuitive due to some parts of the Auer rods closely resembling the nucleus in intensity, and the fact that most other subtypes of acute myeloid leukemia also carry Auer rods.

Stage 2.1 Discussion

The maximum accuracy achieved with this method was only 86.56%, and the lack of robustness in this system can be interpreted as an absence of precise features. There are quite a bit of overlapping and inseparable data points in 3 and 4-dimensional space, which will decidedly result in a rather high classification error rate. In fact, researchers such as Agaian et al [5] have attempted to address this problem in a similar work, and have even suggested to complicated variables such as

$$\delta = \sum_i \sum_j P^2(i,j) + (\sqrt{-1}) \left(\frac{\sqrt{\sum_{i=1}^{n} (x_l - x')^2}}{n-1} \right)$$

Equation 1: The Cell Energy variable

in the process of feature extraction in order to better detect color variations in leukocytes. However, hand-crafted approaches to feature extraction is time-consuming and does not generalize to broader classification problems. Furthermore, there are no standards for feature extraction and classification. Within each step, approaches could be mixed and matched with any number or type of classier, resulting in a never-ending source of work. Additionally, as mentioned earlier, one of the most defining attributes of an APL is the presence of faggot cells — cells with an abundance of Auer rods. Unfortunately, traditional feature extraction approaches do not offer a robust method of precisely locating these elements.

Future Work

The results of this work could have been more captivating if there was access to more sophisticated equipment. There exists a variety of deep neural networks and training methods in the domain of image classification, however, most of them are extremely demanding in terms of computational power and therefore could not be tested in this work. As well, increasing the training dataset of both APL and non-APL cells would definitely also help the classifier in determining more precise features and reducing classification error. There is also a lot of room for the generalization of this training method on other types of blood cell disease conditions, such as intracellular parasites that are unidentifiable to methods like flow cytometry and cytogenetics, making them only possible of being detected through manual microscopic examination of blood films [14].

4. Applications

This proposed system has been externalized as an Android application (currently only for testing purposes) but it can be easily incorporated onto any other smartphone or similar device. The main goal of the app is to provide ancillary service to technicians in classifying an APL cell (i.e. a technician who's uncertain about a cell could take out his phone, take a picture of the cell through the microscope, and the app would inform the technician of whether or not it is APL.) However, this classification system can be effortlessly applied in many more practices, such as a sophisticated automation software for large hospitals. Furthermore, the methods utilized in this work, especially methods of preprocessing and segmentation, can be applied to a wide range of blood cell analysis developments.

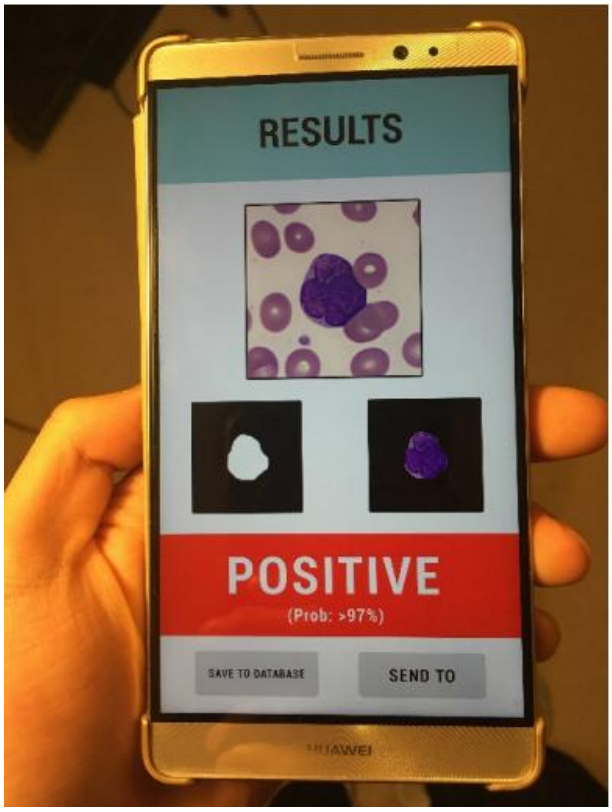

Figure 5: The proposed system as an Android app

5. Conclusions

In this work, a computer vision-based system for the classification of acute promyelocytic leukemia cells was proposed and developed. In Stage 1, K-means clustering, Fuzzy C-Means clustering and Gray World Assumption were used to segment the leukocyte from a larger image and perform color correction. Moreover, in preparation for feature extraction in Stage 2.1, Otsu's method was used to segment the nucleus from the cytoplasm using a gray level intensity histogram. Stage 2.1 consisted of image feature extraction, feature selection and assessing the accuracies of traditional machine learning classifiers. The Linear Discriminant Analysis classifier returned the highest classification accuracy of 86.56%; however, as this accuracy is still far from what was hypothesized, more analyses were done on the classification process and it was determined that the feature extraction system was inefficient. To solve this issue, a convolutional neural network was employed in place of traditional feature extraction, feature selection and machine learning, resulting in a final accuracy of 99.2% averaged across 10 tests (15 525 images each). This novel classification system can be externalized as an app or be used on computers and other portable devices, significantly ameliorating the workflow of medical technicians; for example, a technician could attach their phone to the microscope, use the app to take a picture of an APL resembling leukocyte and get instant feedback on the classification of the cell, which will help them decide whether urgent treatment is required.

References

[1] Wang, Z.-Y. & Chen, Z. Acute promyelocytic leukemia: from highly fatal to highly curable. Blood 111, 2505–2515 (2008).
[2] Giuseppe, A. Newly Diagnosed Acute Promyelocytic Leukemia. NCBI PubMed (2011).
[3] Brescia, M. et al. Early hemorrhagic death before starting therapy in acute promyelocytic leukemia: association with high WBC count, late diagnosis and delayed treatment initiation. Haematologica 95,853–854 (2009).
[4] Acute Leukaemia: Cytology, Cytochemistry and the FAB Classification. Leukaemia Diagnosis 1.56 doi:10.1002/9780470988374.ch1
[5] Mcnerney, R. Diagnostics for Developing Countries. Diagnostics 5, 200–209 (2015).
[6] Silva, W. F. D. et al. Real-life Outcomes on Acute Promyelocytic Leukemia in Brazil – Early Deaths Are Still a Problem. Clinical Lymphoma Myeloma and Leukemia 19, (2019).
[7] Acute Lymphoblastic Leukemia Image Database for Image Processing. Department of Information Technology – Università degliStudi di Milano.
[8] Lisicki, M. A DEEP LEARNING COMPUTER VISION SYSTEM FOR IMAGE BASED CYTOMETRY. (2014).
[9] A. Biondi, G. Cimino, R. Pieters, C. H. Pui, "Biological and therapeutic aspects of infant leukemia", Blood, Vol. 96, No. 1, pp. 24-33, July, 2000
[10] Agaian, S. &Madhukar , M. A new acute leukaemia-automated classification system. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 6, (2016).
[11] A.K. Jain, R.P.W. Duin, and J. Mao, "Statistical pattern recognition: A review," IEEE Transactions on Pattern Analysis and MachineIntelligence, Vol. 22, No. 1, pp. 4-37, 2000