Engineering Applications of Artificial Intelligence 72 (2018) 415–422

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification

Luis H.S. Vogado *, Rodrigo M.S. Veras, Flavio. H.D. Araujo, Romuere R.V. Silva, Kelson R.T. Aires

Federal University of Piauí, Teresina, Brazil

ARTICLE INFO

Keywords:
Medical images
Leukemia diagnosis
Convolutional neural networks
Transfer learning

ABSTRACT

Leukemia is a pathology that affects young people and adults, causing premature death and several other symptoms. Computer-aided systems can be used to reduce the possibility of prescribing inappropriate treatments and assist specialists in the diagnosis of this disease. There is a growing use of Convolutional Neural Networks (CNNs) in the classification and diagnosis of medical image problems. However, the training of CNNs requires a large set of images. To overcome this problem, we use transfer learning to extract images features for further classification. We tested three state-of-the-art CNN architectures and the features were selected according to their gain ratios and used as input to the Support Vector Machine classifier. The proposed methodology aims to correctly classify images with different characteristics derived from different image databases and does not require a segmentation process. We built a new database from the union of three distinct databases presented in the literature to validate the proposed methodology. The proposed methodology achieved hit rates above 99% and outperformed nine methods found in the literature.

1. Introduction

Over the years, multiple medical aid systems have been proposed. Among the diseases aided by computer systems, leukemia is the one that has the highest number of fatalities among adolescents and children, and the risk of developing it is higher in children up to five years of age. Leukemia is a type of cancer that originates in the bone marrow (Fig. 1a) and causes abnormal proliferation of white blood cells (Fig. 1b). To diagnose leukemia, specialists can carry out various tests and exams, including physical examinations, blood tests, blood counts, myelograms, lumbar punctures and bone marrow biopsies. Microscopic analysis is the most economical method of carrying out the initial screening of patients with leukemia. This type of test is done manually, which may generate fatigue in operators. Therefore, a low-cost system that is automatic and robust is necessary to avoid the operator's influence.

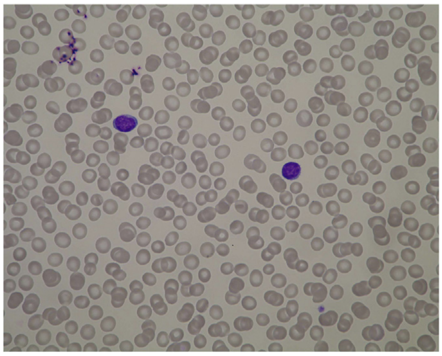
Many computer-aided diagnosis systems were developed with the use of image processing and computational intelligence techniques. These systems usually have some steps such as: preprocessing, segmentation, feature extraction, and classification. Feature extraction and classification are the steps that best define the diagnosis performed by computer-aided diagnosis systems. However, to achieve better results, a proper segmentation can provide an adequate feature extraction and consequently a reasonable classification.

In this work, we propose a leukemia diagnosis system that does not require the segmentation process (commonly used in state-of-the-art techniques). The methodology uses pre-trained Convolutional Neural Network (CNN) models (AlexNet Krizhevsky et al., 2012, Vgg-f Chatfield et al., 2014 and CaffeNet Jia et al., 2014) to extract features directly from the images without any previous preprocessing. Then, the obtained features will be used for the following classification with a Support Vector Machine (SVM) (Cortes and Vapnik, 1995). We used three hybrid datasets to evaluate the performance of the methodology, one with blood smears containing only one leukocyte per image, one with many leukocytes per image, and the last one with both types of images. To demonstrate the robustness of our approach, we compared the results obtained by our methodology with other state-of-the-art methods.

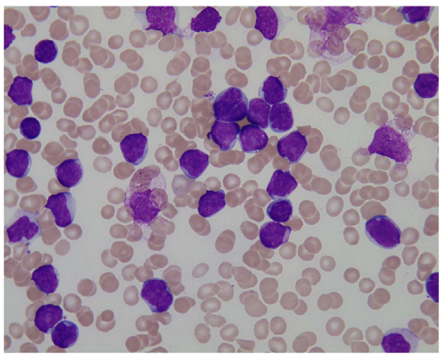
The remainder of the paper is organized as follows: Related works are presented in Section 2, and the proposed method is introduced in Section 3. In Section 4, we present the experiments and also describe the image datasets used in the tests and the evaluation of the method. We discuss the results in Section 5. Finally, the conclusions and perspectives on future works are given in Section 6.

* Correspondence to: Universitária Avenue, Postal code: 64049-550, Brazil.
E-mail addresses: lhvogado@gmail.com (L.H.S. Vogado), rveras@ufpi.edu.br (R.M.S. Veras), flavio86@ufpi.edu.br (F.H.D. Araujo), romuere@ufpi.edu.br (R.R.V. Silva), kelson@ufpi.edu.br (K.R.T. Aires).

https://doi.org/10.1016/j.engappai.2018.04.024
Received 18 October 2017; Received in revised form 14 April 2018; Accepted 20 April 2018



(a)



(b)

Fig. 1. (a) ALL-IDB1 leukemic blood smear sample and (b) ALL-IDB1 non-leukemic blood smear sample.

2. Related works

Several methods for leukemia detection have been proposed over the years and some of these works presented solutions to the classification of the two most common types of leukemia: Acute Myeloid Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL). Some works presenting state-of-the-art technologies only provide diagnosis by using images with one leukocyte per image (Mohaprata et al., 2011; Mohapatra et al., 2014; Neoh et al., 2015; Putzu et al., 2014; Singhal and Singh, 2016) and others that provide diagnosis by using images with multiple leukocytes (Madhukar et al., 2012; Vincent et al., 2015; Patel and Mishra, 2015; Agaian et al., 2018; Singh et al., 2017).

Table 1 present some works on leukemia detection and highlights the techniques of feature extraction and classification used. As can be seen in this table, most of the listed approaches combine texture (Haralick's features Haralick et al., 1973 and Local Binary Pattern (LBP) chen He and Wang, 1990); shape (area, perimeter, compactness, Fractal Dimension Pentland, 1984...) and color features (mean, standard deviation, uniformity, entropy...).

In the classification process, the most frequently classifiers used were: SVM, Multilayer Perceptron (MLP) (Popescu et al., 2009), Random Forest (RF) (Breiman, 2001), K-Nearest Neighbors (KNN) (Aha and Kibler, 1991), Radial Basis Function Network (RBFN) (Schwenker et al., 2001) and Naive Bayesian (NB) (Friedman et al., 1997).

The works presented in Table 1 achieves accuracy rates above 90%. However, they depend on the use of a proper segmentation process and do not show good performance when applied to databases with different characteristics. Recent works reported that the use of deep learning is producing better results than all classical techniques for different types of medical image applications (Castelluccio et al., 2015; Wang et al., 2016; Kumar et al., 2016). A large set of images is required to train a Convolutional Neural Network (CNN), but the available leukemia image databases are small and limited. Thus, to avoid the problem of having a limited number of images, we investigate the use of learning transfer to extract images features and subsequent classification by SVM. It is worth mentioning that this approach does not need the use of the segmentation process and is more robust to the characteristic variation of the images.

Castelluccio et al. (2015) present two methods of transfer learning in their work. The first one consists in fine-tuning the network, where the structure is modified, freezing high-level layers. The other method includes the extraction of the last network fully connected layer, obtained from the input image (Athiwaratkun and Kang, 2015). Then, it uses another classifier in the classification process. In Athiwaratkun and Kang (2015), the authors demonstrated that the layers with activation localized near the output layer could usually extract the best characteristics.

3. Proposed methodology

The method proposed in this work aims to diagnose leukemia using blood smear images. Following the flowchart shown in Fig. 2, it is possible to observe that the system uses an image without any preprocessing or segmentation as input. This is the main difference between our method and the state-of-the-art methods. The CNNs are used to describe the input image and the features are selected and reduced. In the classification step, the SVM is used to classify the images as pathological or not.

3.1. Convolutional neural networks

Currently, researchers working with artificial intelligence use deep learning techniques to create robust computational systems. Among these techniques, CNNs (Fig. 3) have been proposed to aid in diagnosis, and they have outperformed conventional methods of extracting features, obtaining better accuracy rates (Lecun et al., 2015; Litjens et al., 2017). Therefore, we use the power of the CNNs to carry out the construction of our approach.

In recent works, authors have presented two different ways of using the power of CNNs. The first is the usual method in which the training is performed with a large set of data. The second is through the transfer of learning using pre-trained networks (Tajbakhsz et al., 2016). In our work, we used pre-trained CNNs in a large natural image database. Thus, the neural network can assimilate generic features, facilitating its application in small databases. This technique can be used in several types of tasks, for example in the extraction of features from face images, objects, and diseases. The success of the results depends on the similarity of the images from the database used to extract the features and the images from the training set (Shin et al., 2016).

We use CNNs pre-trained on a large natural database of images with 1000 categories, called ImageNet (Russakovsky et al., 2015). Among these categories, the database also contains images of lymphocytes and lymphoblasts. During the development of our approach, three models were selected: AlexNet (Krizhevsky et al., 2012), CaffeNet (Jia et al., 2014), and Vgg-f (Chatfield et al., 2014). These models have similar architectures, and the main difference between them is the size of the filters used in the convolutional layers and the number of neurons in the fully connected layers.

Alexnet: Krizhevsky et al. (2012) developed the architecture for the ILSVRC-2010 competition to carry out training and classification of ImageNet database. It comprises eight layers that need to be trained, five convolutional layers with filters with sizes of 5×5 and 7×7, followed by three fully-connected layers, as well as max-pooling layers.

Caffenet: This architecture was developed by the Berkeley Vision and Learning Center (BVLC) and is considered one of the most popular CNNs in deep learning (Jia et al., 2014). It comprises five convolutional layers, each followed by a pooling layer, and three fully-connected layers. The